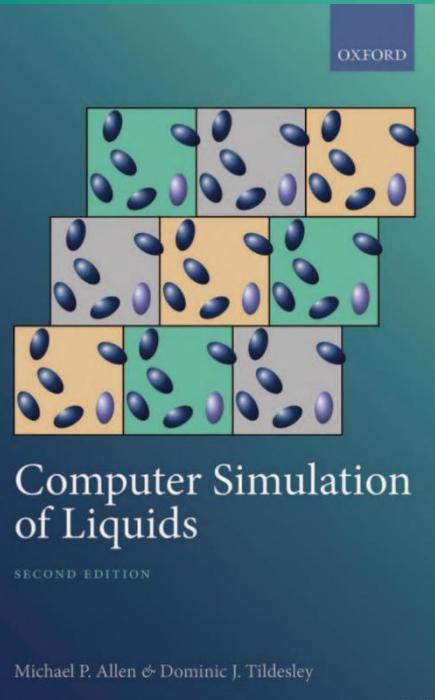


Molecular Dynamics for Neutrons

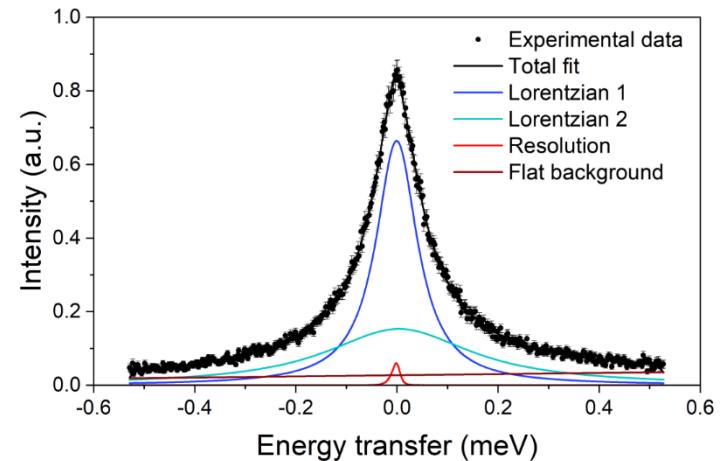
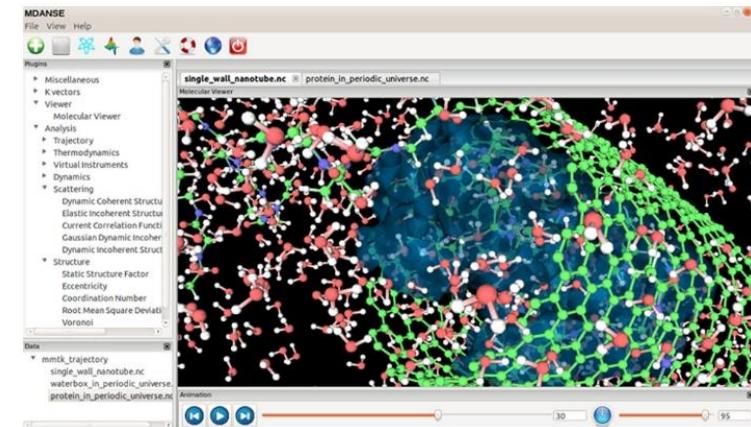
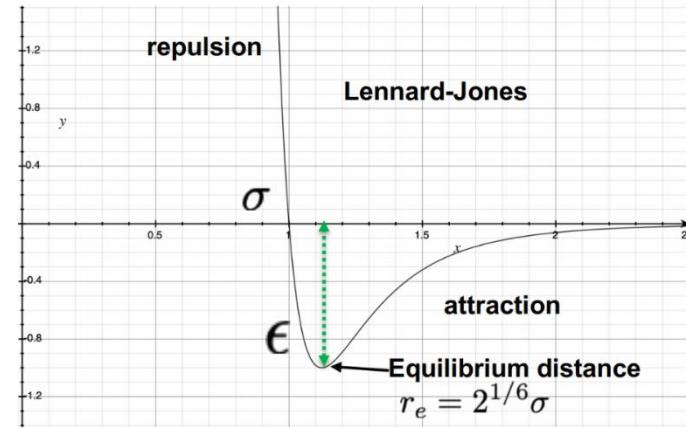
Deeper Reading



Dr. Jeff. Armstrong (ISIS)

What would I like you to take away from today?

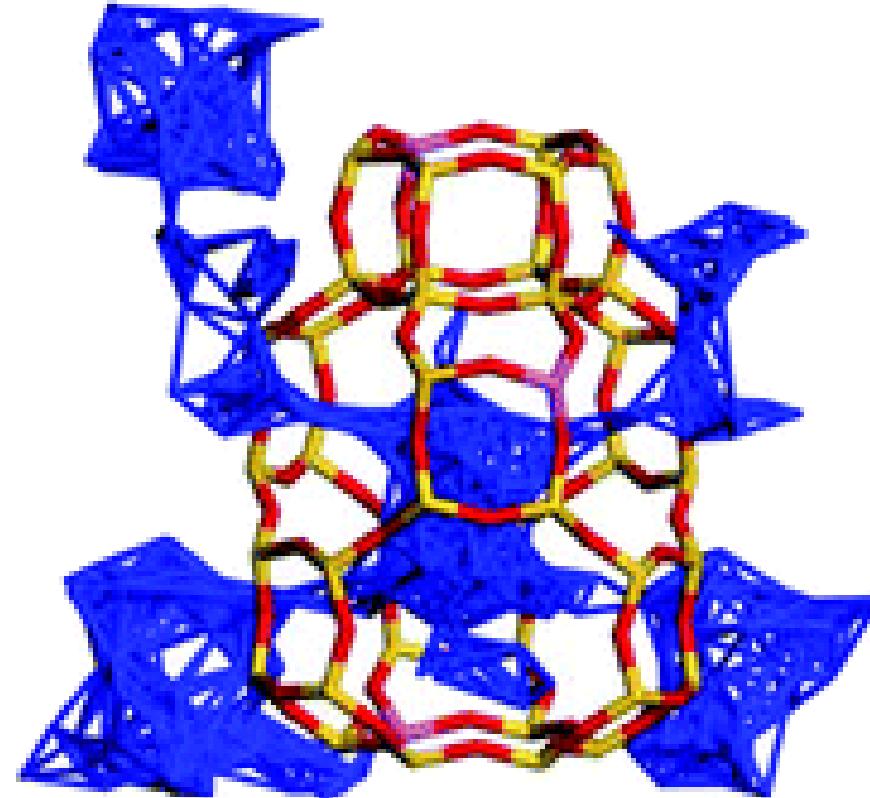
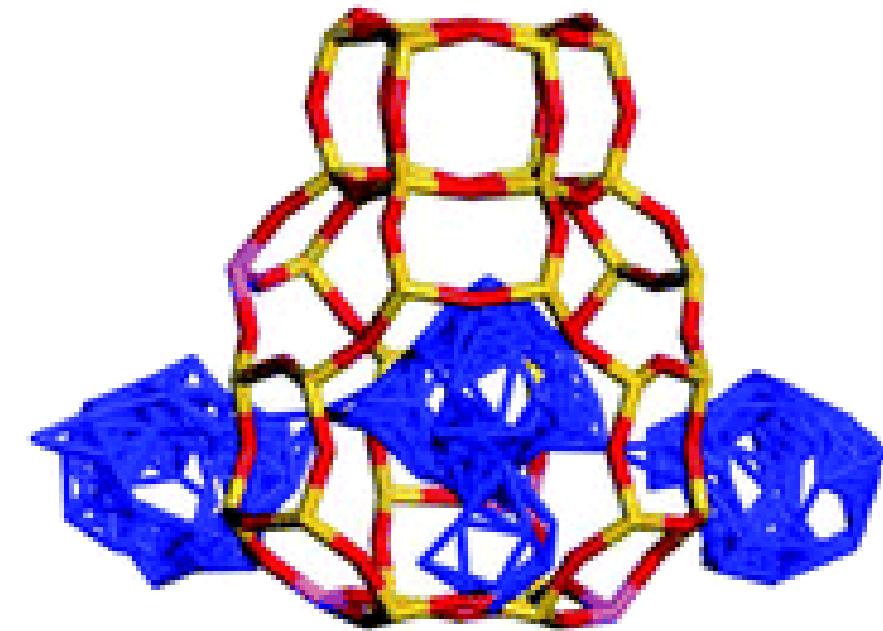
- A basic understanding of the mathematical underpinnings of MD
- Sign posting to the array of software that is available for both setting up and analysing MD
- A convincing argument about how and why MD and Neutron scattering can and should be used to extract more information from one another



What is Molecular Dynamics?

- A method/model that generates a movie at the atomic scale, letting us examine or interrogate the motions/structure of the atoms

Example: Small molecule diffusion in Zeolites

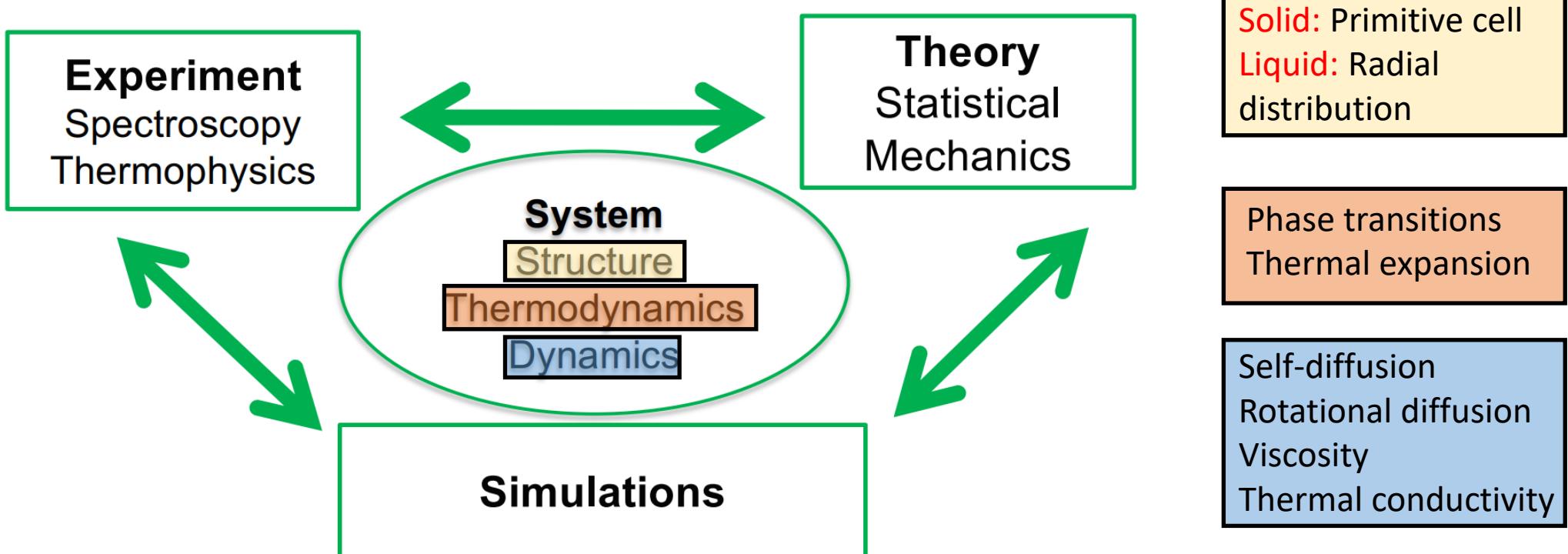


This can help us understand the links between the atomic properties and the macroscopic properties of a system

e.g. Diffusion of products in catalytic zeolite can be a rate limiting effect

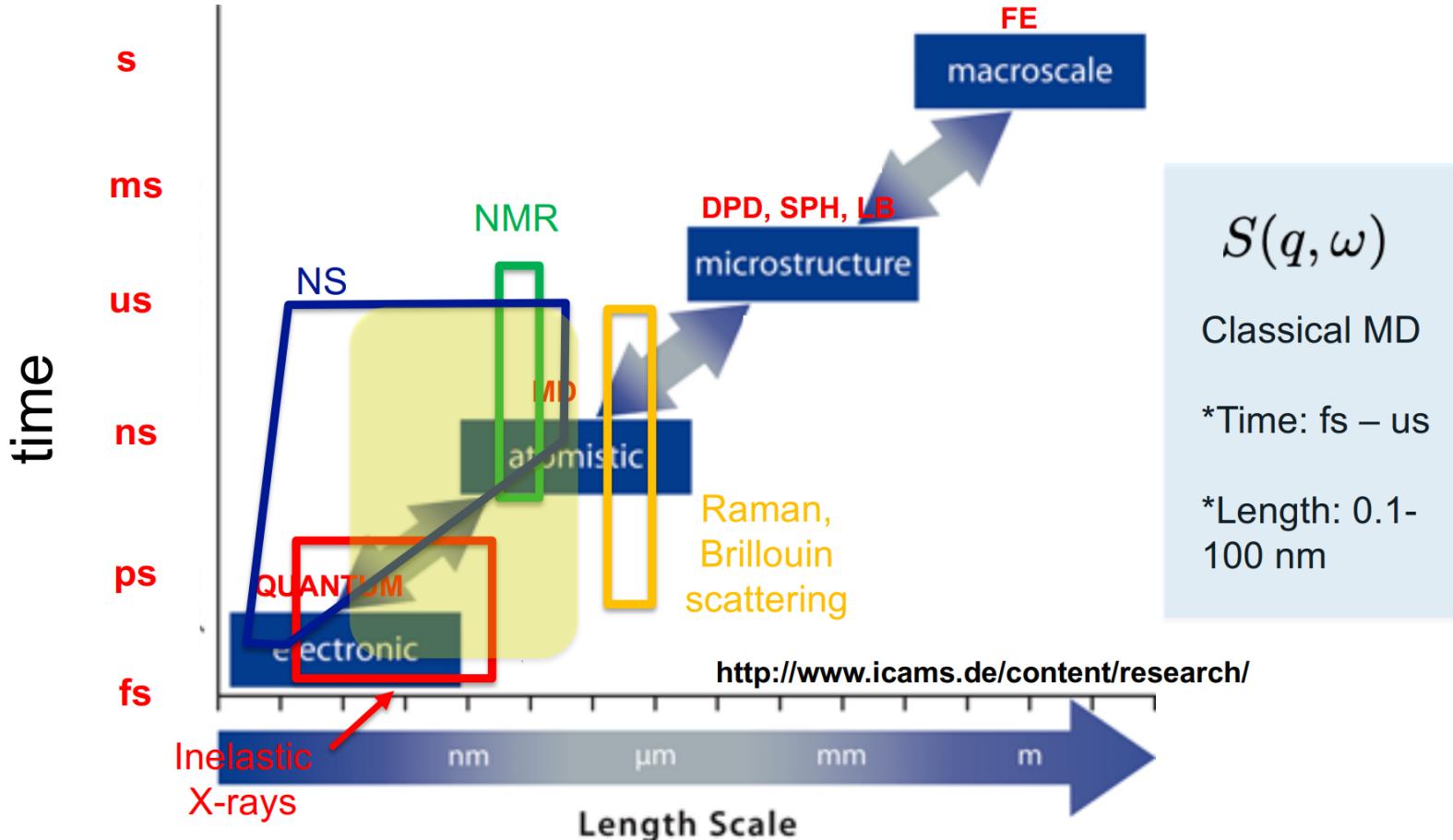
Basic objective of MD (cycle of consistency)

- **Aim:** To predict the equilibrium macroscopic properties of large systems (10^{23} atoms/molecules) using properties of atoms and molecules



- Simulations: supports interpretation of experimental observations, test theories and it is also a **discovery tool => new experiments and theories**

Multiscale ladder (what scale does MD probe?)



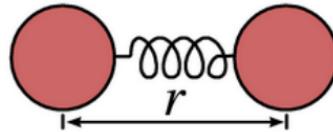
- Dynamics (**QENS**)
- Structure (**SANS/diffraction**)
- Large scale structures/surfaces (**reflectometry**)

➤ Strong overlap in time-lengths scales (q, ω) in neutron scattering and classical Molecular Dynamics

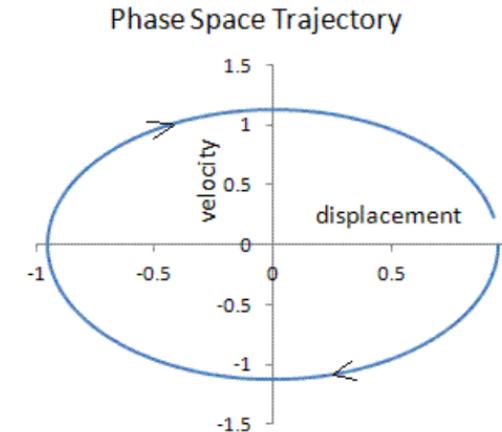
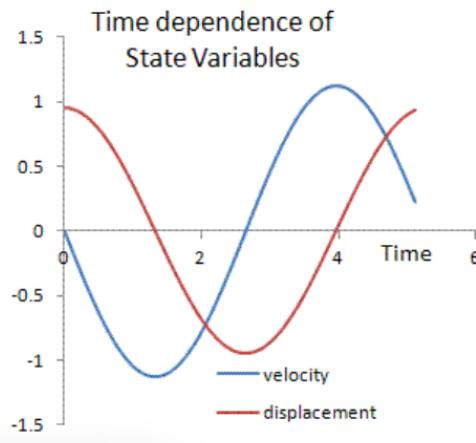
We generate “movies”, but formally we are exploring Phase Space

- Molecular Dynamics provides an approach to solve the dynamic behaviour of atoms and molecules, i.e., to sample the phase space, $\Gamma(q^N, p^N)$, and to compute the properties of a system.

Easily Visualised with the case of a harmonically coupled system



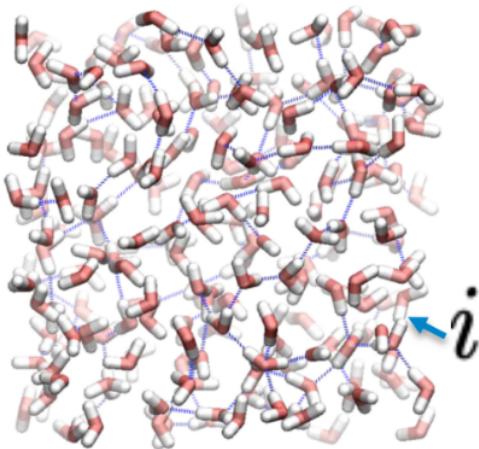
Thermodynamic quantities can be calculated as temporal averages of these state variables



- In MD we assume **ergodicity**: the system can explore “all” the points in the phase space
- We will focus on cartesian coordinates for which Newton’s equations apply

Revisiting Newton (the basic math of MD)

- The force on a given particle i can be represented as the gradient of the potential energy. The time evolution of Newton's equations is given by:



A mathematical term describing the time evolution of the positions

$$\mathbf{F}_i = \frac{d\mathbf{p}_i}{dt} = m_i \frac{d^2\mathbf{r}_i}{dt^2} = -\frac{dU(\mathbf{r}^N)}{d\mathbf{r}_i}$$

Something that is fully determined by the current state of the system (Force-field)

Force-field usually expressed as potential energy

$U \Rightarrow$ Potential energy
 $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$

- For an isolated set of particles Newton's equations conserve the total energy and linear momentum $\Rightarrow (\mathbf{N}, \mathbf{V}, \mathbf{E})$ ensemble

$$E = U + K = U + \sum_i \frac{\mathbf{p}_i^2}{2m_i}$$

$$\frac{dE}{dt} = \sum_i \frac{dU}{d\mathbf{r}_i} \frac{d\mathbf{r}_i}{dt} + \sum_i \frac{d}{dt} \left(\frac{\mathbf{p}_i^2}{2m_i} \right) = 0$$

Energy conservation

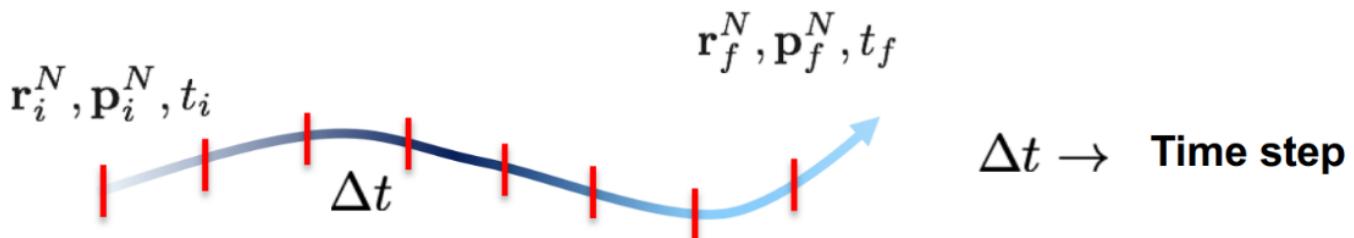
$$\frac{d}{dt} \left(\sum_i \mathbf{p}_i \right) = - \sum_i \frac{dU}{d\mathbf{r}_i} = \sum_i \mathbf{f}_i = 0$$

Linear momentum conservation

How do we solve this in practice?

➤ **Objective:** develop a numerical method to integrate the trajectory of N interacting atoms. Most methods rely on a Taylor expansion of \mathbf{r} about time t . One of the most successful methods was developed by Verlet (1967).

The closer we are to our initial “known” state, the more accurate the truncated Taylor expansion will be (i.e. small time-step).



➤ Perform a Taylor expansion about $\mathbf{r}(t)$:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t \frac{d\mathbf{r}(t)}{dt} + \frac{\Delta t^2}{2} \frac{d^2\mathbf{r}(t)}{dt^2} + \frac{\Delta t^3}{3!} \frac{d^3\mathbf{r}(t)}{dt^3} + \mathcal{O}(\Delta t)^4$$

$$+$$

$$\mathbf{r}(t - \Delta t) = \mathbf{r}(t) - \Delta t \frac{d\mathbf{r}(t)}{dt} + \frac{\Delta t^2}{2} \frac{d^2\mathbf{r}(t)}{dt^2} - \frac{\Delta t^3}{3!} \frac{d^3\mathbf{r}(t)}{dt^3} + \mathcal{O}(\Delta t)^4$$

$$\mathbf{F}_i = m_i \frac{d^2\mathbf{r}_i}{dt^2}$$

$$\mathbf{r}(t + \Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \Delta t^2 \frac{d^2\mathbf{r}(t)}{dt^2} + \mathcal{O}(\Delta t)^4 \quad (1)$$

Time reversible, conserves linear momentum (conservative forces)
“good” energy conservation

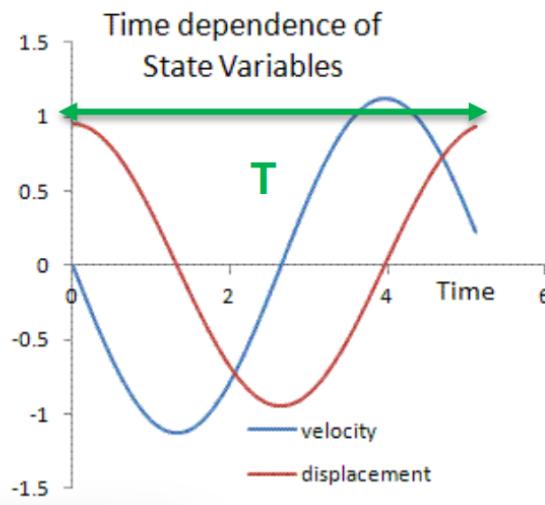
Current step Previous step We have this! F/m

➤ Subtract equations to get:

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} = \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t - \Delta t)}{2\Delta t} + \mathcal{O}(\Delta t)^2 \quad (2)$$

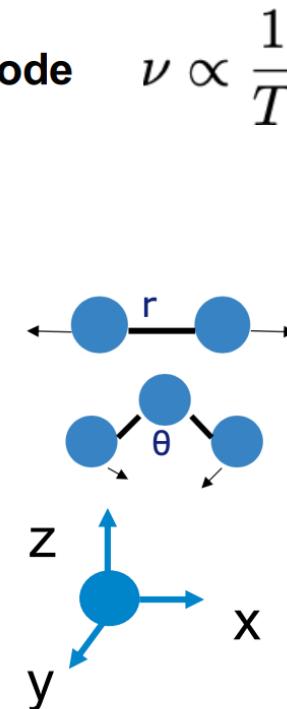
Time steps and energy conservation

- The numerical integration of Newton's equations of motion should conserve energy. **Criterion:** conserve energy in $\sqrt{\langle \delta E^2 \rangle}/E < 1/10^{4-5}$
- In order to conserve energy the time step Δt should be $\leq 1/10^{\text{th}}$ of the fastest time scale:
Useful rule of thumb



$$\Delta T \sim \frac{1}{10\nu} \approx \frac{T}{10} \quad \nu: \text{Frequency of fastest vibrational mode} \quad \nu \propto \frac{1}{T}$$

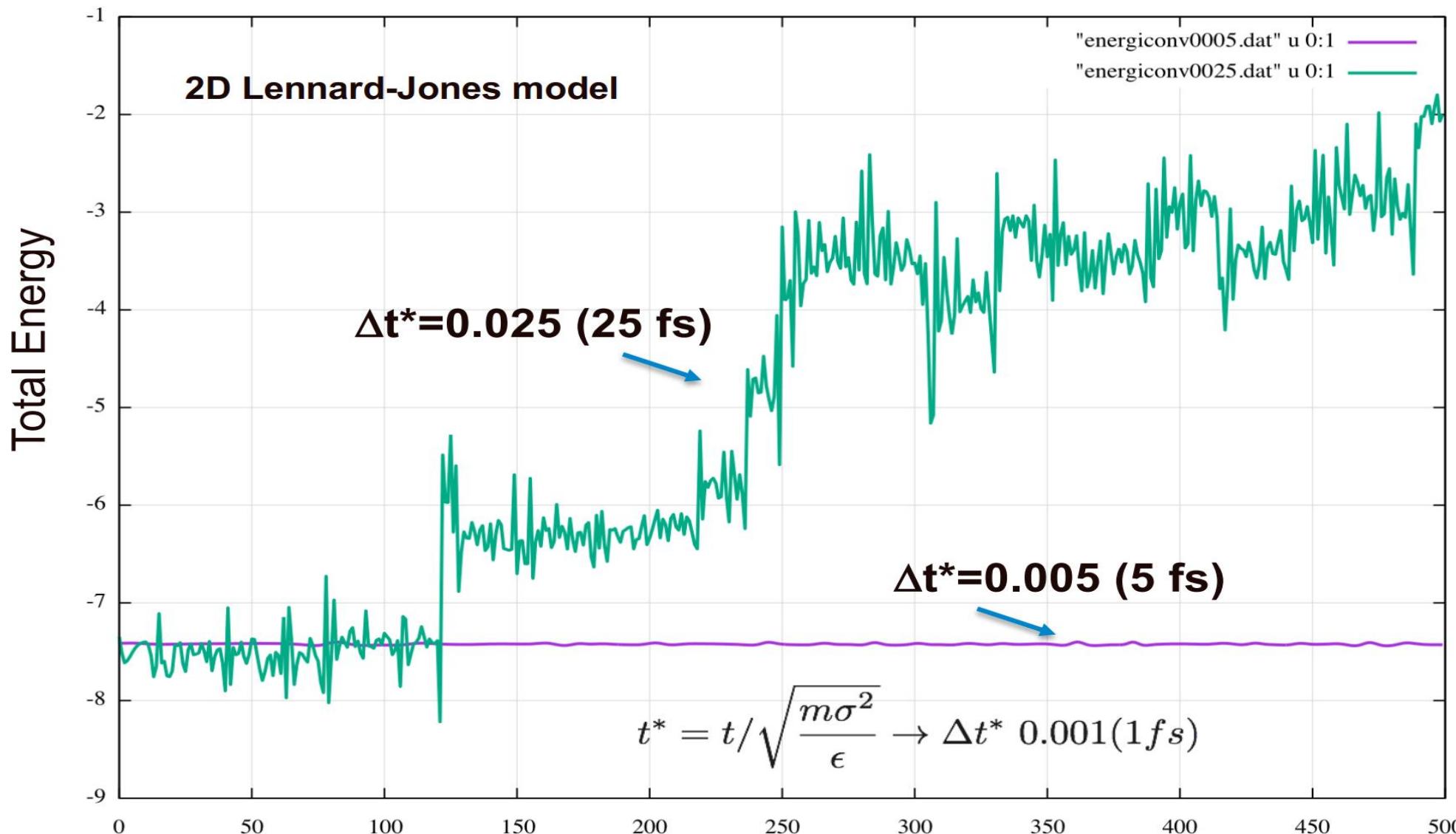
Motion	Time step (fs=10 ⁻¹⁵ s)
Stretching	0.5-1
Bending	2
Translation	5-10



Note: Rigid models are often available, and perform very well in many cases

Example energy conservation

You probably want less than a 1% energy drift over the course of the simulation



Thermostats and Barostats

Guide the system to a given temperature or pressure

➤ **Microscopic state:** defined by the atom positions and momenta $\{q, p\} \Rightarrow$ **phase space** (Γ). A point $\{q, p\}$ defines a **microstate**, and a collection of points satisfying the conditions of a given thermodynamic state define an **ensemble**.

Microcanonical (NVE)	Canonical (NVT)	Isothermal- Isobaric (NPT)
Isolated	q	w, q

➤ Integration of Newton's equations of motion generates microstates in the microcanonical ensemble (NVE). **Ensemble average = Time average (ergodic principle)**

➤ Simulation in other ensembles can be performed by coupling the system to a heat bath (NVT) or to a barostat (NPT)

Nosé-Hoover canonical dynamics (NVT)

$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m_i}$$

$$\frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i - \xi \mathbf{p}_i$$

$$\frac{d\xi}{dt} = \nu_T^2 \left[\frac{\sum_i \mathbf{p}_i^2 / m}{gk_B T} - 1 \right] = \nu_T^2 \left[\frac{T}{T} - 1 \right]$$

Instantaneous
temperature

Relaxation rate $g \approx 3N$ Ensemble temperature

It is often appropriate to run a simulation in NPT at a desired pressure, and from the average box dimensions, switch to a NVE ensemble

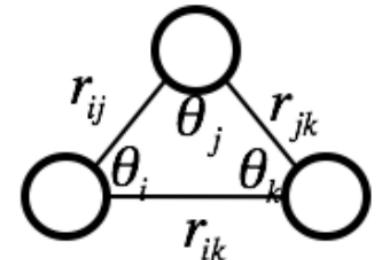
But what is the force-field?

Usually a potential energy that can be expressed as a function of the atomic positons

- The interaction potential can be written as a sum of N -body contributions, including bonding and non-bonding (coulombic, dispersion, polarization, charge transfer) interactions

$$U(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i=1}^N u_1(\mathbf{r}_i) + \sum_{1 \leq i < j \leq N} u_2(\mathbf{r}_{ij}) + \sum_{1 \leq i < j < k \leq N} u_3(\mathbf{r}_{ij}, \mathbf{r}_{jk}, \mathbf{r}_{ki}) + \dots$$

External field **Two body** **Three body**



- Very often we use the pairwise additivity approximation

$$U(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{1 \leq i < j \leq N} u_2(\mathbf{r}_{ij}) = u_2(\mathbf{r}_{12}) + \dots + u_2(\mathbf{r}_{23}) + \dots$$

Higher order than pair potentials are very costly and rarely used

Interatomic potentials

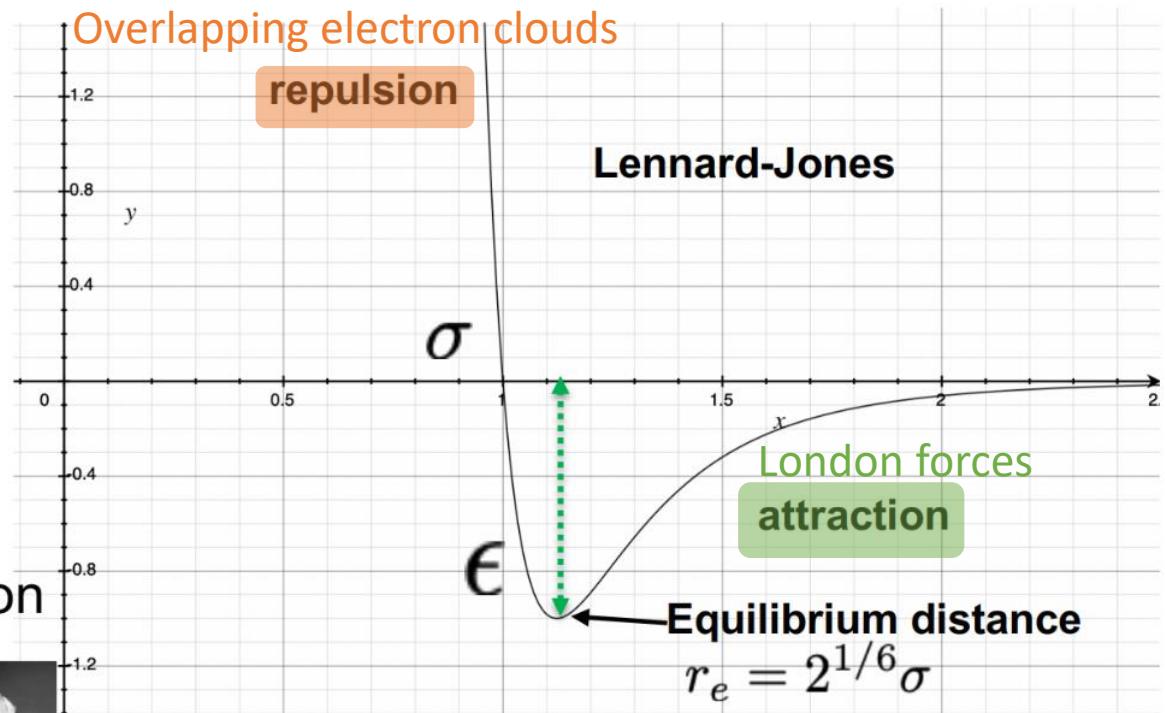
- Hard spheres (no liquid phase)

$$u^{HS}(r_{ij}) = \begin{cases} \infty & \text{if } r_{ij} < \sigma \\ 0 & \text{if } r_{ij} \geq \sigma \end{cases}$$

How would you simulate this!?

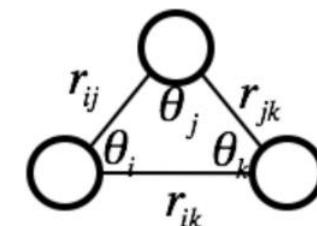
- Lennard-Jones potential => dispersion

$$u(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right]$$



- Three-body: Axilrod-Teller

$$u(\vec{r}_i, \vec{r}_j, \vec{r}_k) = \frac{c(1 + 3\cos\theta_i \cos\theta_j \cos\theta_k)}{r_{ij}^3 r_{jk}^3 r_{ik}^3}$$

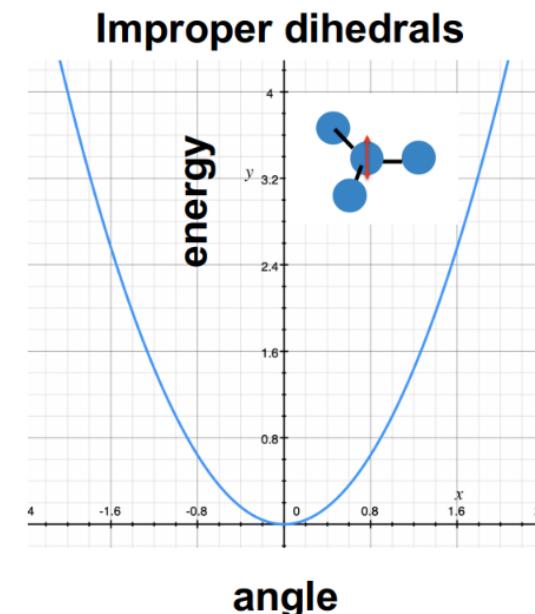
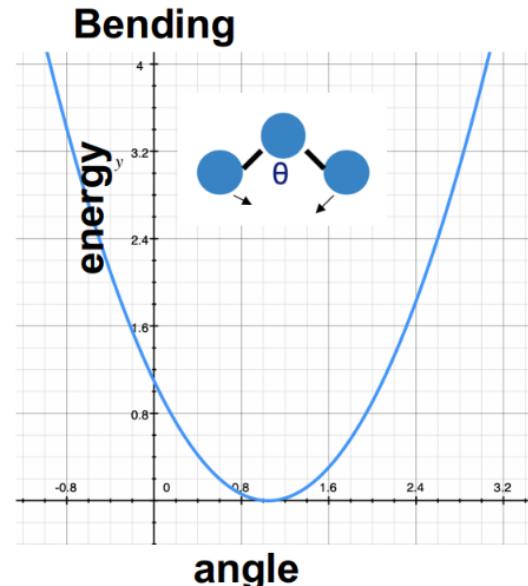
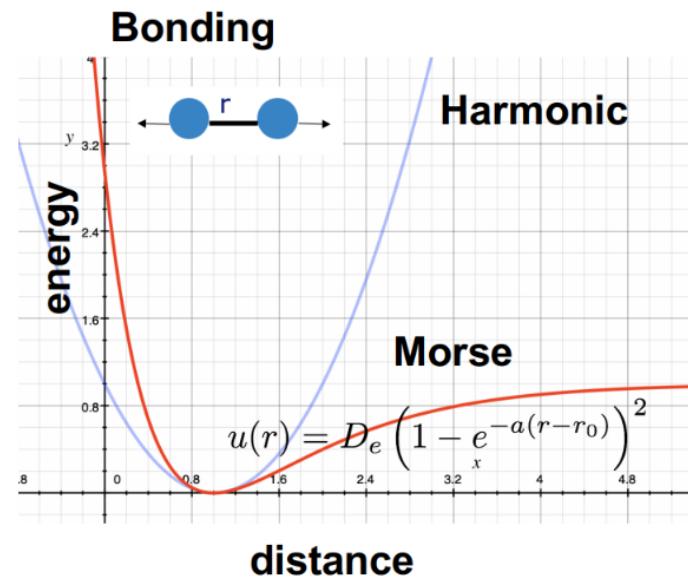


- Coulomb

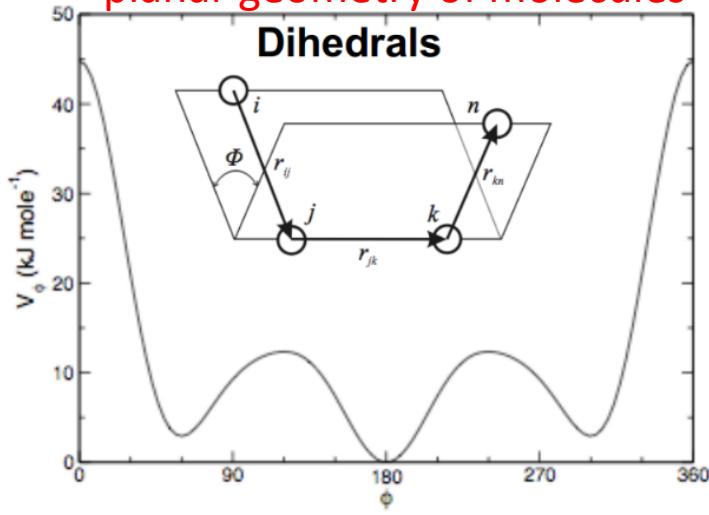
$$u(r) = \frac{q_i q_j}{4\pi\epsilon_0 r}$$

These have to be treated via special methods (e.g. Ewald)
This is due to $1/r^2$ nature (energy drops off at the same rate as volume increases)

Intra-molecular terms



Very important for keeping planar geometry of molecules



$$u_{intra} = \sum_b k_b (r - r_e)^2 + \sum_\theta k_\theta (\theta - \theta_0)^2 + \sum_\xi k_\xi (\xi_{ijkl} - \xi_0)^2$$

$$+ \sum_\phi k_\phi (1 + \cos(n\phi_{ijkl}) - \delta)$$

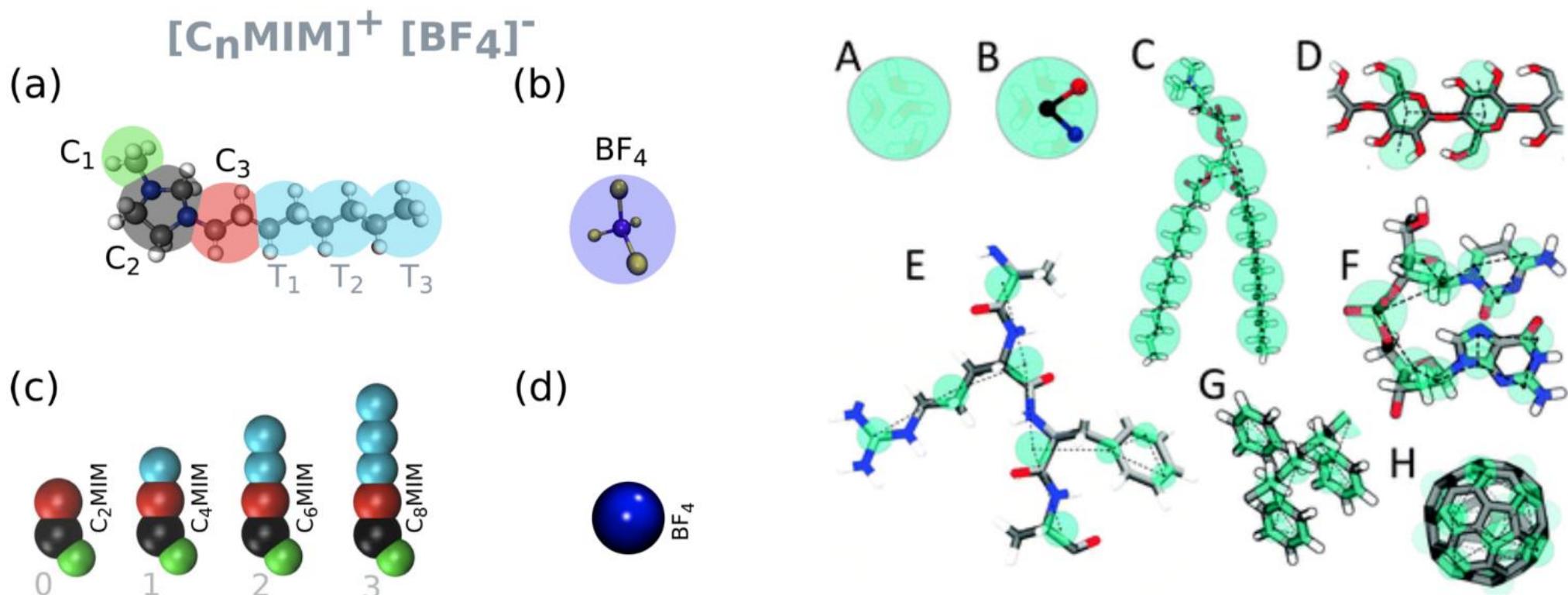
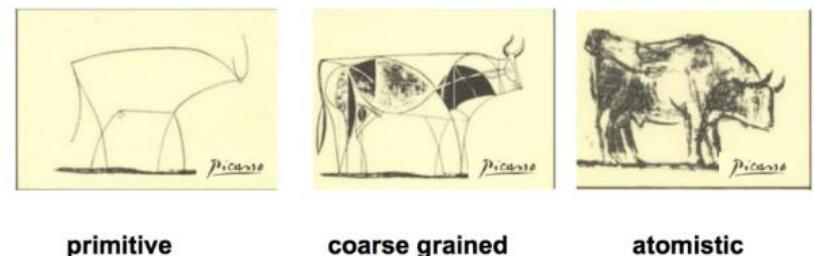
Careful about dihedral convections in different codes/FFs

- At ~300 K vibrations are in the ground state => rigid bonds (constraints)
- AMBER, CHARMM, GROMOS (BIO), OPLS, TraPPE (MOL. LIQUIDS), UFF (GENERIC), FENE (POLYMERS)

Coarse-graining (something to be aware of)

- Particularly important for large systems in biology (lipid bi-layers, proteins etc.)
- Can be important for SANS and reflectometry
- Note that **dynamic quantities** can be **very wrong**

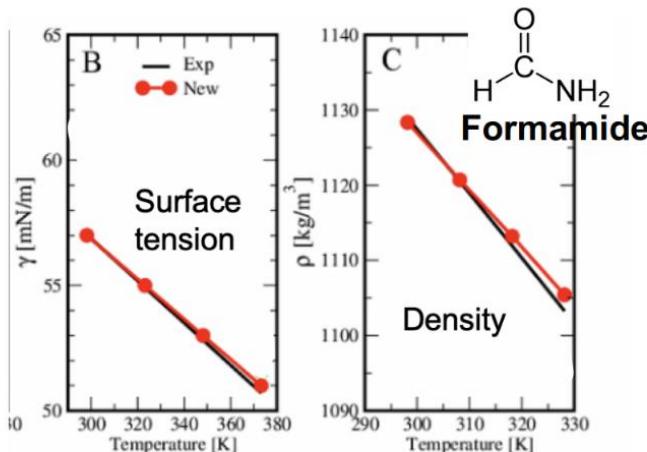
The forcefield jungle



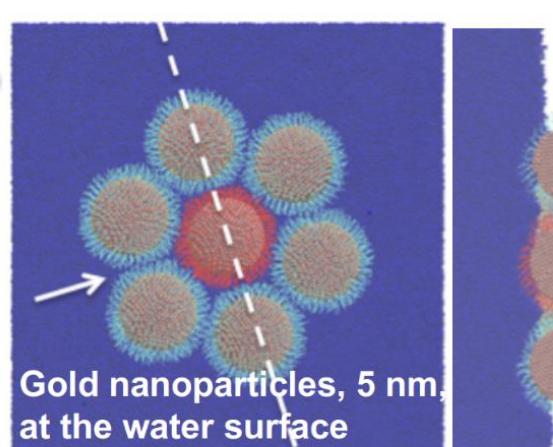
Martini CG vs atomistic

Where do these force-fields come from and what are their weaknesses?

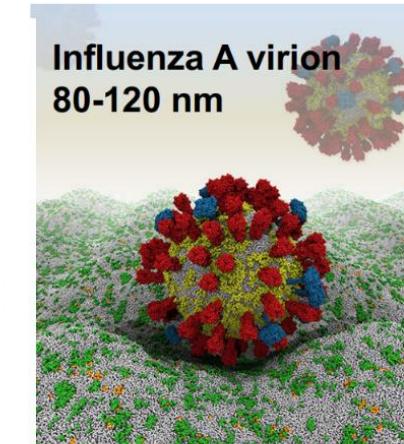
- It is not possible to model chemical reactions (bond breaking formation, electron excitations / charge transfer). **Multi-scale approaches do exist**
- Parametrized to reproduce properties at specific conditions => transferability may be limited. **Phase transitions for example may be shifted**
- Classical FF can be fairly accurate at predicting thermodynamic, structural and dynamic properties + it is possible to investigate large systems



10³ molecules, atomistic
DOI: 10.1021/acs.jctc.5b00080



0.5x10⁶ particles, atomistic
DOI: 10.1039/C5NR00620A



Influenza A virion
80-120 nm
Reddy et al, 10⁶ Particles, CG
DOI: 10.1016/j.str.2014.12.019

Periodic Boundary Conditions

(we can't simulate a whole macroscopic system)

- Simulations are performed using “small” systems $10^2 < N < 10^6$ (CPU limited). Interactions $\sim O(N^2)$
- Large fraction of molecules influenced by **surface effects**: 50% for a system containing 10^3 molecules.
- **Periodic boundary conditions** (Born and von Karman, 1912) => infinite number of copies of the simulation box

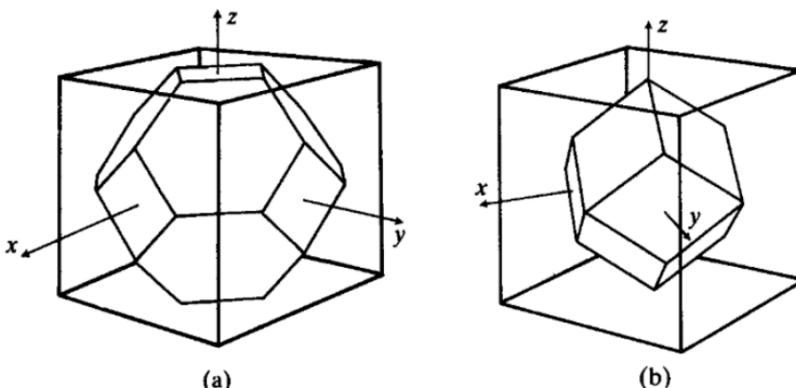
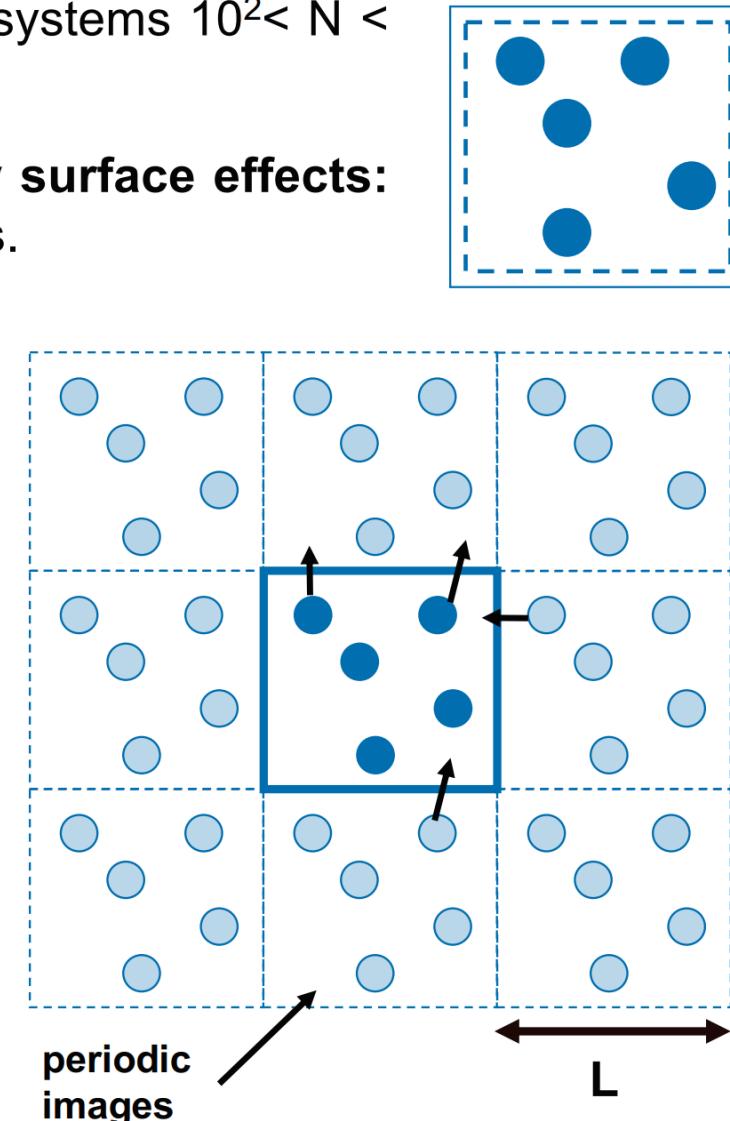
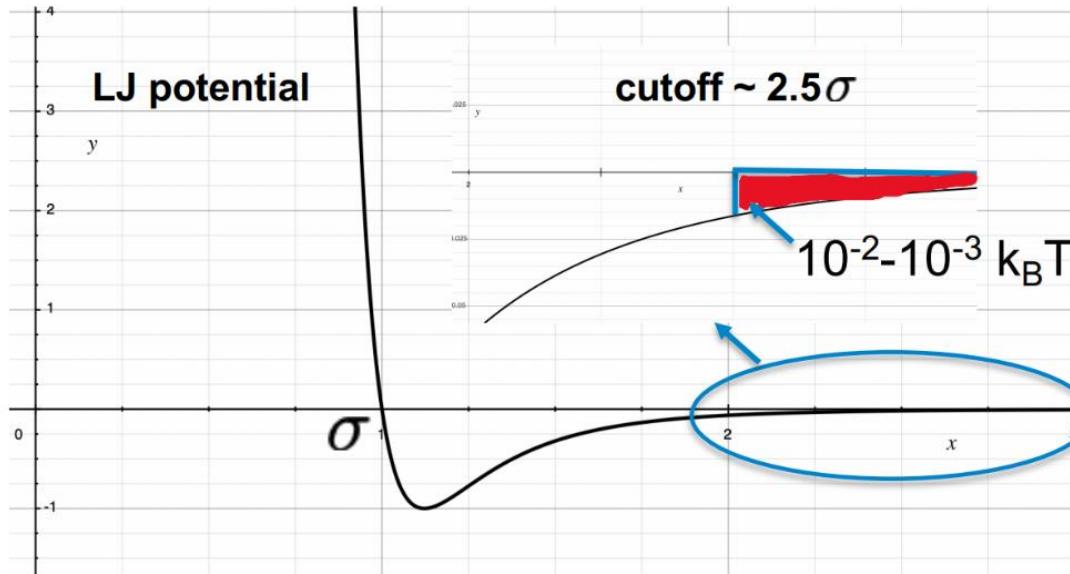
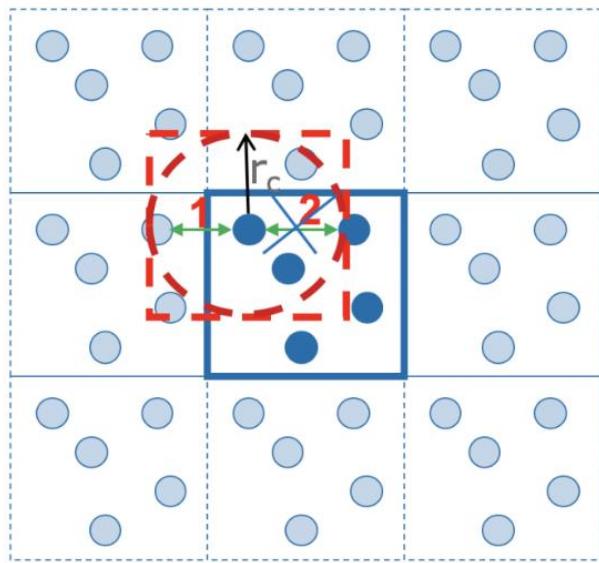


Fig. 1.10 Non-cubic simulation boxes. (a) The truncated octahedron and its containing cube; (b) the rhombic dodecahedron and its containing cube. The axes are those used in microfiche F.1.



Cut-offs and minimum image convention

- Periodic boundary conditions => spherical truncation + minimum image (MI)



- What cutoff, r_c should I use? => as small as possible! and $< (\text{Box length}/2)$ to avoid multiple image interactions: **BEWARE**: short cutoffs modify cohesive energy (e.g. critical point) significantly
- Typical cutoff for dispersion (LJ) interactions 2.5σ => $U(2.5 \sigma) \sim 10^{-2}-10^{-3} k_B T$
- Coulomb: $U(r_c) \sim 10^{-2} k_B T$ at $r_c \sim 100 \sigma$ (water, $0.033 \text{ molecules}/\text{A}^3 \Rightarrow 10^{7-8} \text{ molecules}!!$)=> Special techniques required to compute Coulombic interactions => Ewald summation (Particle Mesh Ewald), Reaction Field, Wolf method (see Allen & Tildesley).

You also need to be aware of neighbour lists, which keep track of which atoms are within a given radial distance of one another, meaning only a sub-set of distances need to be calculated each step

What can we measure?

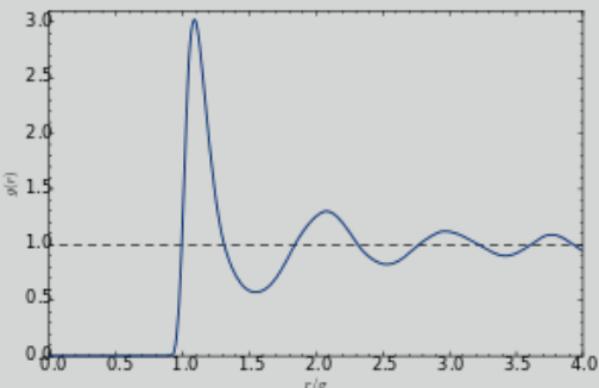
Energy and entropy changes
as a function of T etc.

Thermodynamics

$$E = U + K$$
$$H = U + PV$$

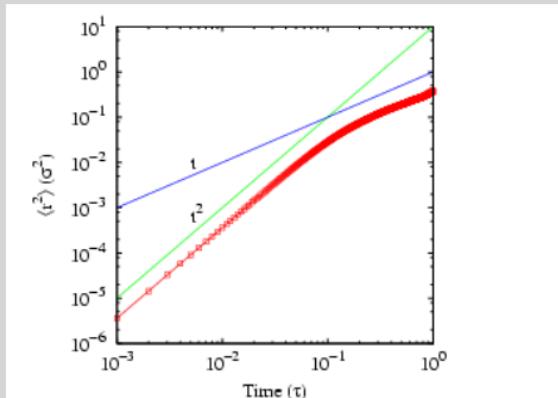
Atomic positions relative to
each other and surfaces

Structure



Individual or collective
motions of atoms

Dynamics

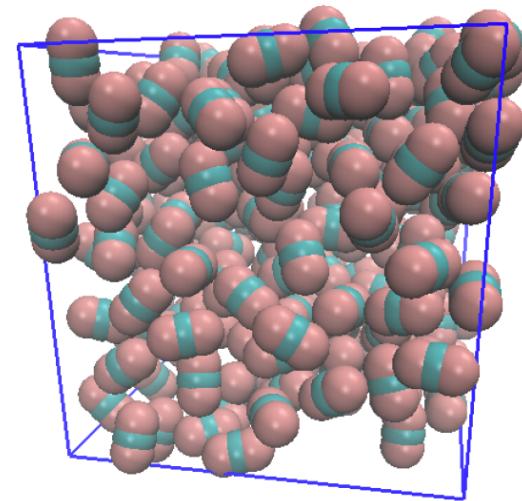
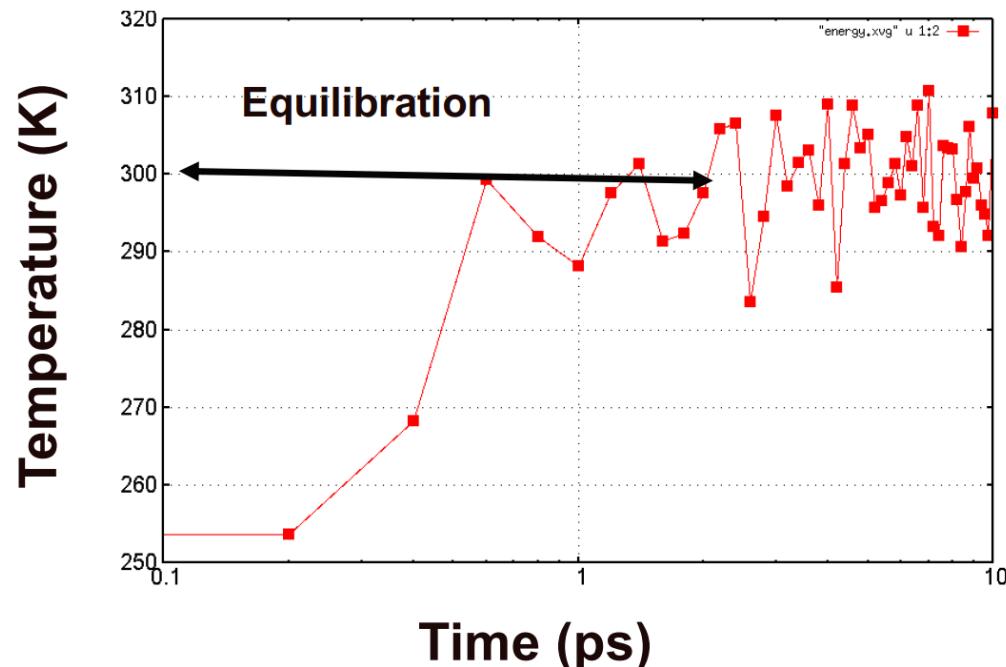


Ergodic principle: $\langle A \rangle_{ensemble} = \langle A \rangle_{time}$

Average over time is the average
over all possible configurations

Equilibration (when are we ready to analyse?)

- When can we calculate averages?: the system needs first to be **equilibrated**. Averages are computed during the **production** process => when the properties fluctuate ($\text{Gaussian} \sim 1/\sqrt{N}$) around a well defined average
Has the running average plateaued?
Is the typical variation consistent with time?



*CO₂ – N= 216 molecules
*NPT: ensemble (1 bar, 300K)

- Estimating the equilibration time: $\tau \sim L^2/D$

e.g.: D~ 10⁻³ nm²/ps; L=0.3 nm => 90 ps; L=3 nm => 9 ns

Thermodynamic quantities

These can be useful quantities to assess equilibration

➤ Potential energy

$$\langle U \rangle = \left\langle \sum_{i=1}^{N-1} \sum_{j=i+1}^N u(r_{ij}) \right\rangle$$

➤ Kinetic Energy

$$\langle K \rangle = \left\langle \frac{1}{2} \sum_{i=1}^N m_i v_i^2 \right\rangle$$

➤ Temperature: $\langle T \rangle = \frac{2 \langle K \rangle}{N_{dof} k_B}$

“ N_{dof} ” is the number of **degrees of freedom**
e.g. N particle moving in 3D (x,y,z)=> 3N

➤ Pressure: $PV = NK_B T + \langle \mathcal{W} \rangle$

ideal

virial

Pairwise interactions

$$\mathcal{W} = \frac{1}{3} \sum_{i=1}^{N-1} \sum_{j>i}^N \mathbf{r}_{ij} \cdot \mathbf{f}_{ij}$$

$\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$
Forces on atom i due to j

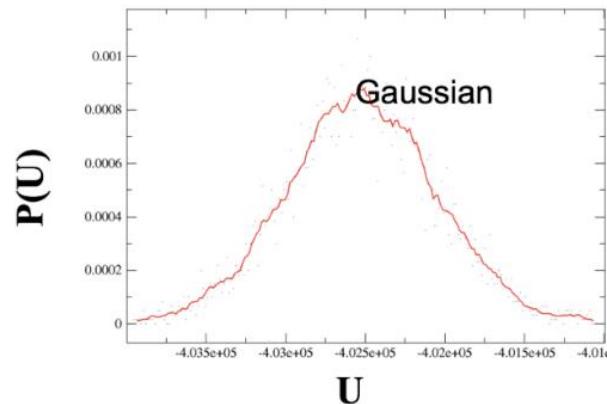
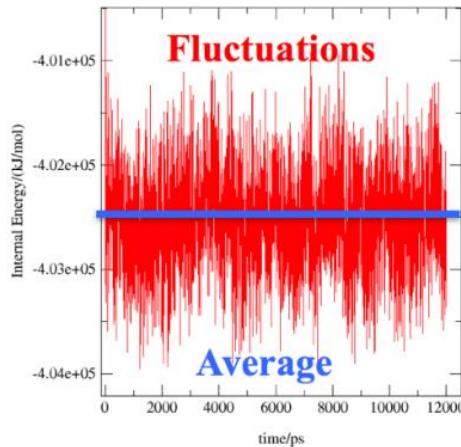
P is a tensor, $\mathbf{P}_{\alpha\beta}$, that can be used to compute surface and interfacial tensions.

➤ Heat Capacity (fluctuations): $C_V = \left(\frac{\partial E}{\partial T} \right)_V = \frac{\langle (E - \langle E \rangle)^2 \rangle}{k_B T^2} = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2}$

Also very sensitive to phase changes of various orders (including dynamical transitions e.g. Rotor phase)

Fluctuations

- Fluctuations are ensemble dependent => Different ensembles require different fluctuation equations.

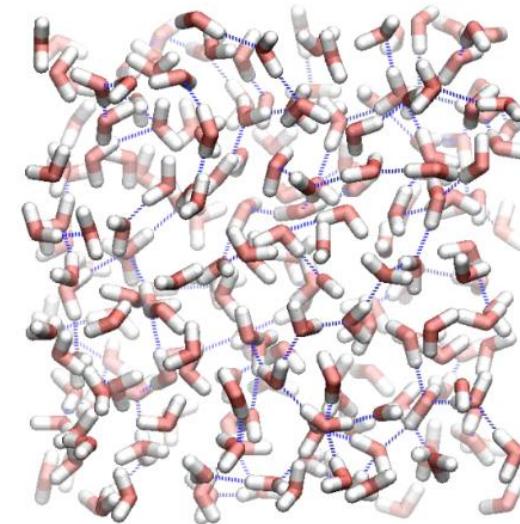


$$C_V = \left(\frac{\partial E}{\partial T} \right)_V = \frac{\langle (E - \langle E \rangle)^2 \rangle}{k_B T^2} = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2}$$

Heat capacity at constant volume, (NVT) ensemble

$$C_p = \frac{\langle H^2 \rangle - \langle H \rangle^2}{k_B T^2}$$

Heat capacity at constant pressure, (NPT) ensemble

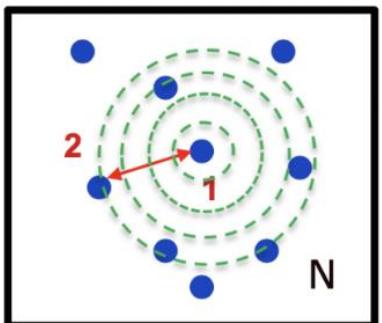


$$\kappa_T = \frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T = \frac{1}{k_B T} \frac{\langle V^2 \rangle - \langle V \rangle^2}{\langle V \rangle}$$

Isothermal compressibility, (NPT) ensemble

Radial Distribution Function

- The radial distribution function or pair correlation function is widely used to investigate the molecular structure of gases, liquids and solids.
- The radial distribution function quantifies the probability of finding two particles 1 and 2 at distance, \mathbf{r}_{12}

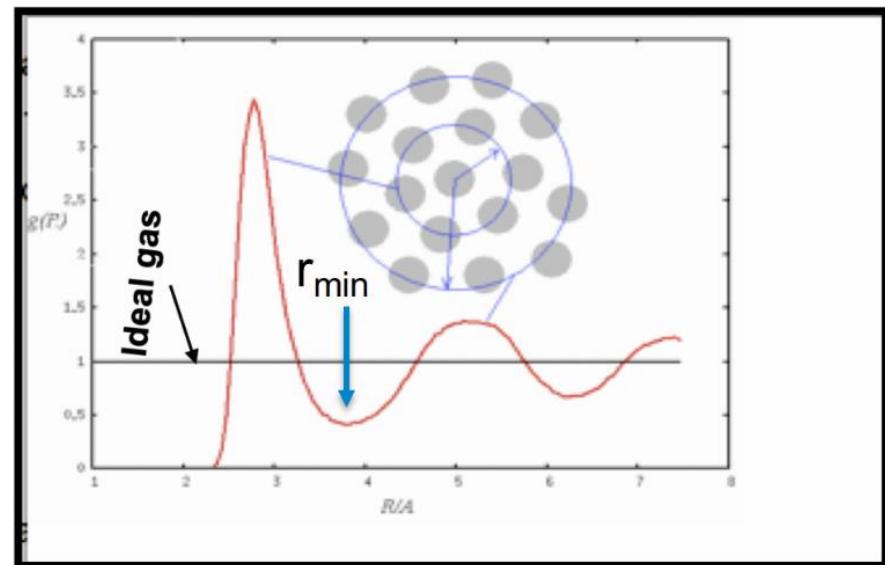


In practice this is counted within given shells

$$g(r) = \frac{1}{\rho} \left\langle \frac{1}{N} \sum_i \sum_{j \neq i} \delta(\mathbf{r} - \mathbf{r}_j + \mathbf{r}_i) \right\rangle$$

For a non interacting set of atoms (ideal gas) gas

$$g(r) = 1 - \frac{1}{N} \approx 1$$



- Coordination number (C.N.) $n(r_{min}) = 4\pi\rho \int_0^{r_{min}} r^2 g(r) dr$

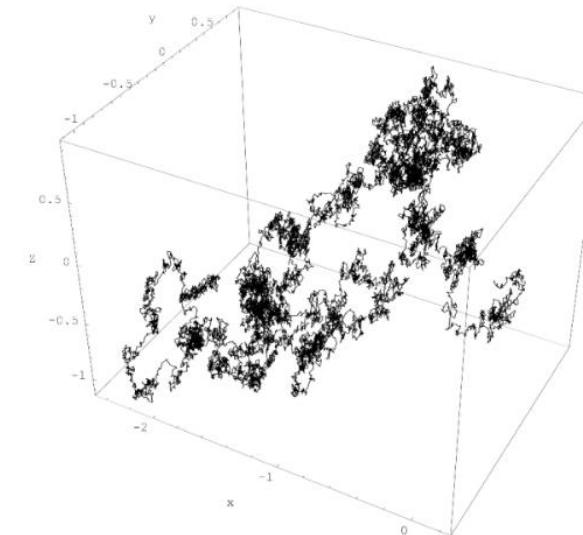
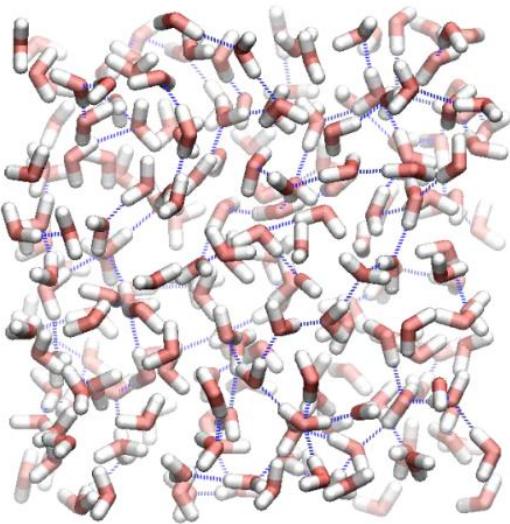
- Structure factor

$$S(q) = 4\pi\rho \int_0^\infty dr r^2 (g(r) - 1) \frac{\sin(qr)}{qr}$$

neutrons

Dynamics (Self-diffusion coefficient)

- Atoms and molecules move continuously changing their positions and velocities
=> Brownian motion



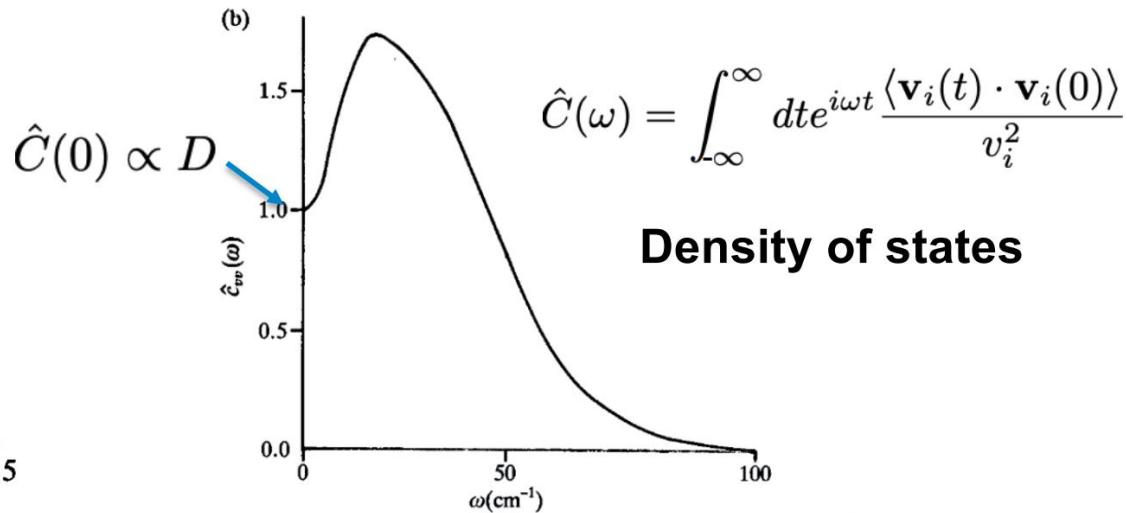
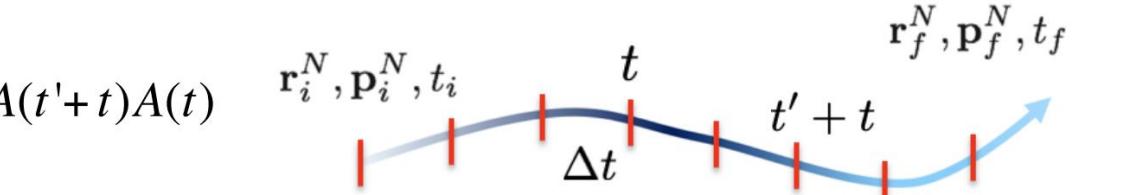
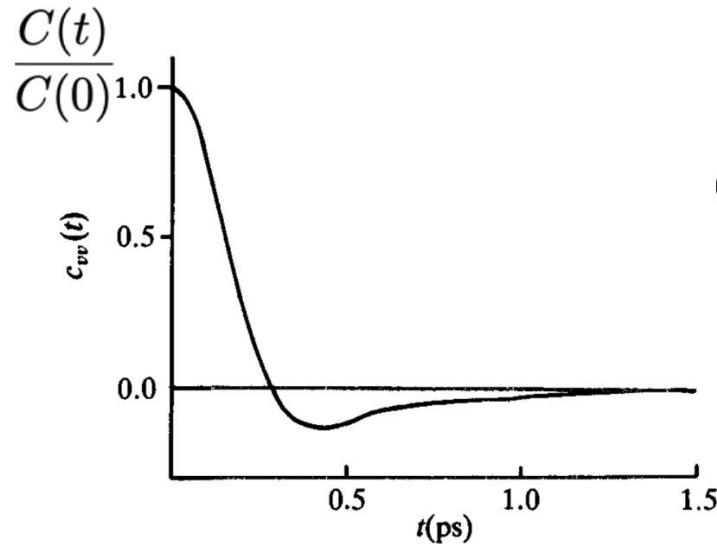
- The diffusion coefficient, D, quantifies how fast molecules/atoms move. There are two approaches to calculate D: mean square displacements and autocorrelation functions

$$MSD \equiv \langle |\mathbf{x}(t) - \mathbf{x}_0|^2 \rangle = \frac{1}{N} \sum_{i=1}^N |\mathbf{x}^{(i)}(t) - \mathbf{x}^{(i)}(0)|^2$$

Correlation functions

- Time correlation functions provide a route to obtain transport coefficients

$$C(t) = \langle A(t')A(0) \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt A(t'+t)A(t)$$



Green-Kubo equation
$$D = \frac{1}{3} \int_0^{\infty} dt \langle \mathbf{v}(0) \cdot \mathbf{v}(t) \rangle$$

Some quantities may also involve cross-correlations

Van Hove (link with Neutrons)

- Van Hove correlation function and dynamic structure factors

$$G(\mathbf{r}, t) = \frac{1}{\rho} \langle \rho(\mathbf{r}, t) \cdot \rho(0, 0) \rangle$$

$$G(\mathbf{r}, t) = G_s(\mathbf{r}, t) + G_d(\mathbf{r}, t)$$

SELF

DISTINCT

$$G_s(\mathbf{r}, t) = \left\langle \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i(t) + \mathbf{r}_i(0)) \right\rangle$$

$$G_s(\mathbf{r}, 0) = \delta(\mathbf{r}) \quad G_d(\mathbf{r}, 0) = \rho g(\mathbf{r})$$

$$G_d(\mathbf{r}, t) = \left\langle \frac{1}{N} \sum_{i=1}^N \sum_{j \neq i}^N \delta(\mathbf{r} - \mathbf{r}_j(t) + \mathbf{r}_i(0)) \right\rangle$$

$$G_s(\mathbf{r}, t \rightarrow \infty) \sim \frac{1}{V} \quad G_d(\mathbf{r}, t \rightarrow \infty) \sim \rho$$

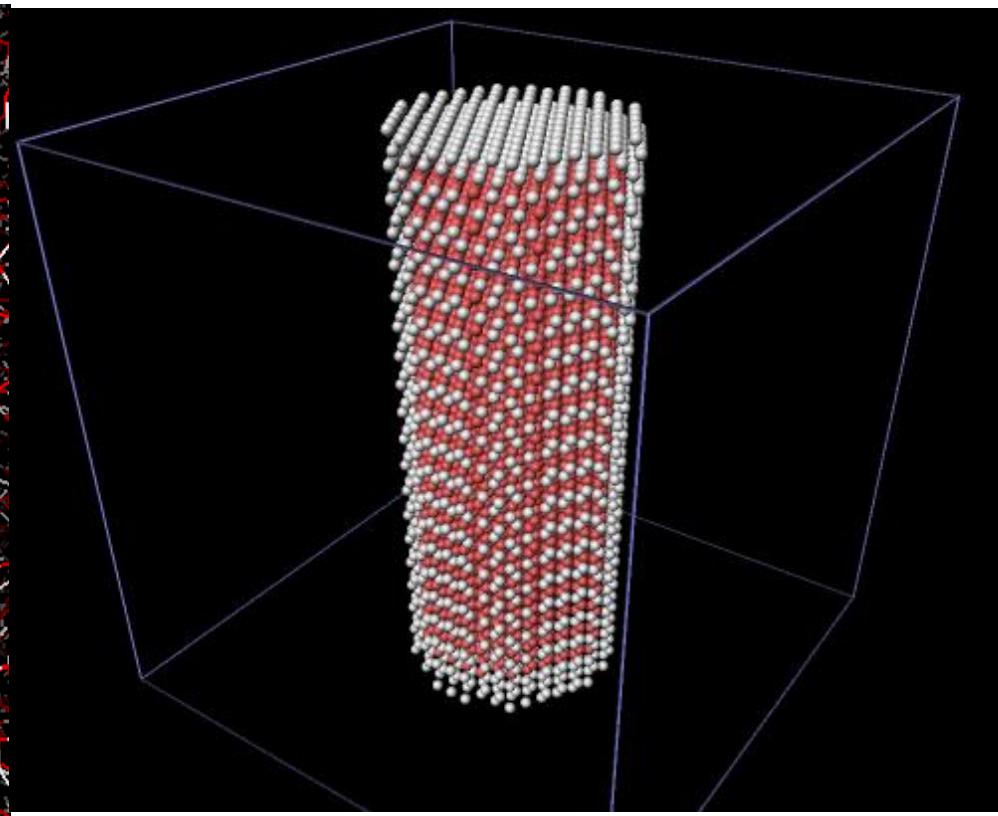
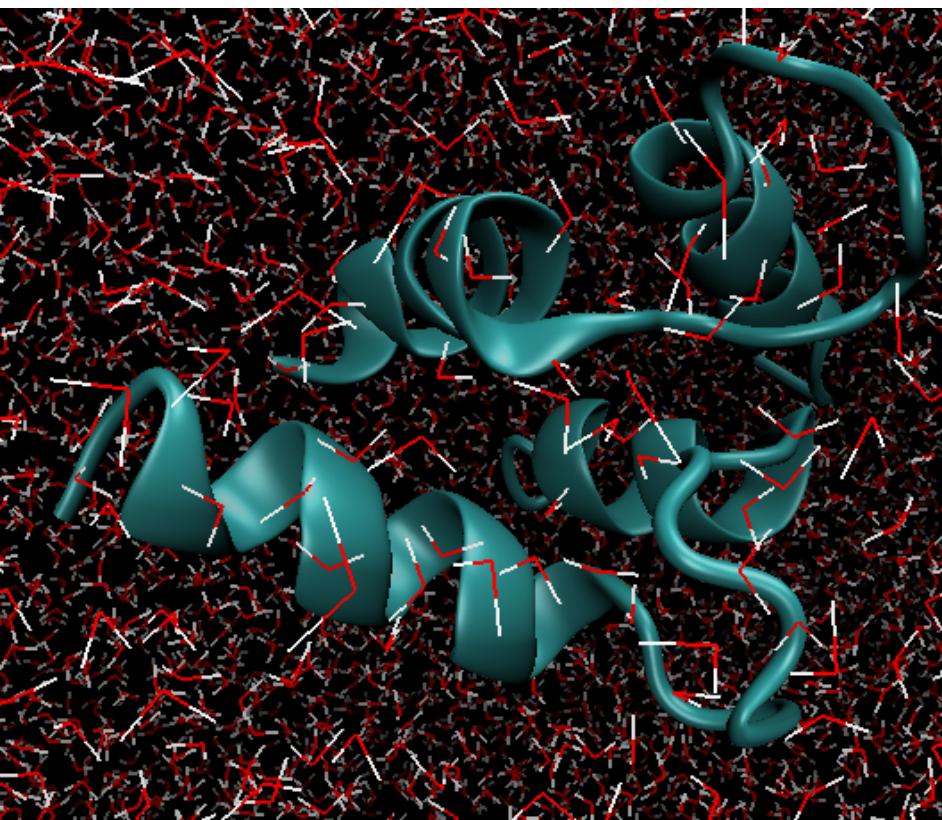
Dynamic structure factor is the spatial and temporal Fourier Transform of the Van Hove Eqn.

$$S_s(q, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt e^{i\omega t} \int G_s(\mathbf{r}, t) e^{-iq \cdot r} d\mathbf{r}$$

Inc.

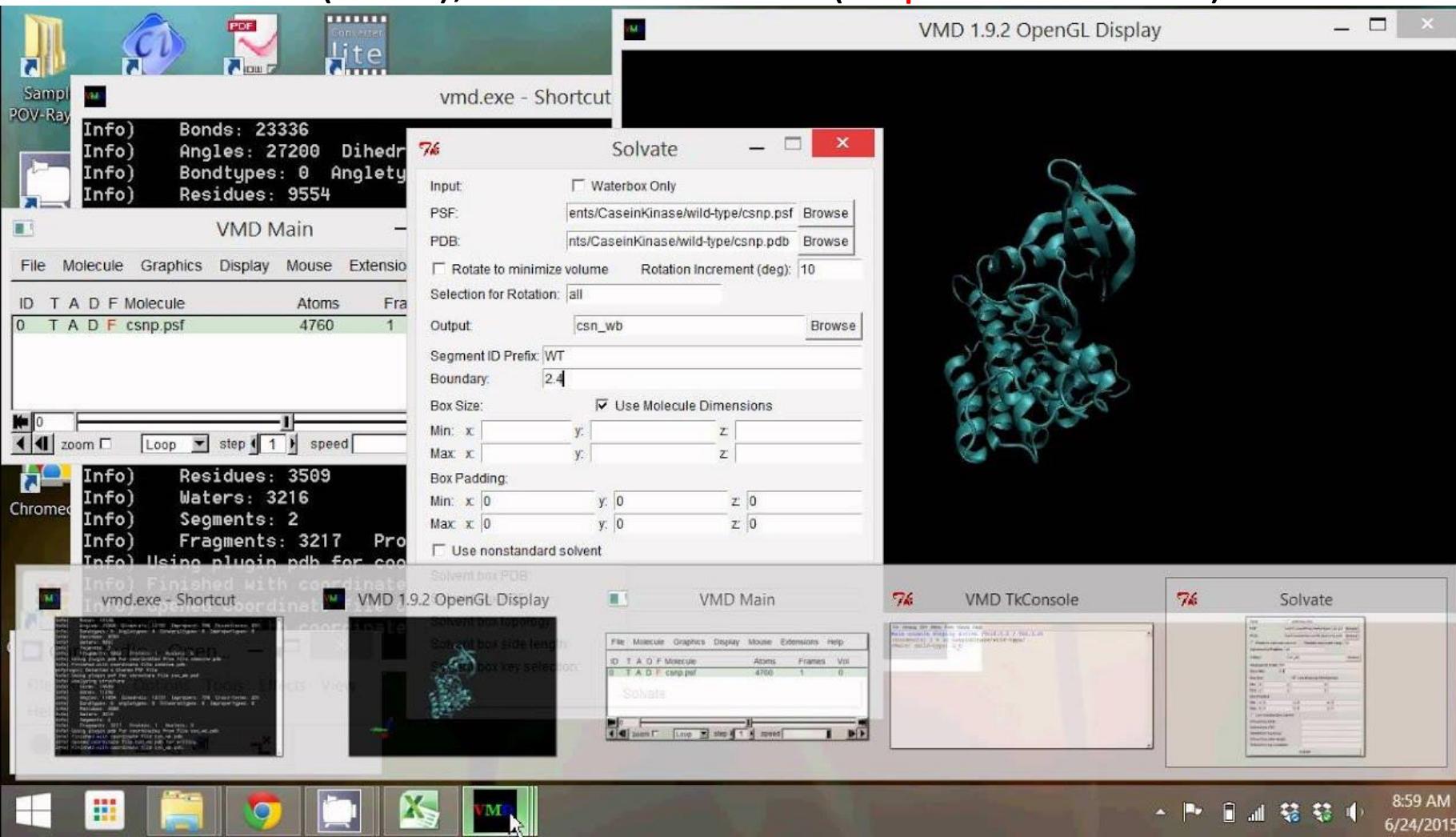
What MD package to use in practice?

- Many flavours of MD (**DL POLY**, **GROMACS**, **LAMMPS**, Not to mention **AB INITIO MD** which also produces trajectories)
- Your choice will depend on your use (**Gromacs**: Fast and therefore good for large systems, but poor flexibility, **LAMMPS** very flexible but slower; good for complex geometries).



Setting up your simulation

- Setting up the configuration for an MD simulation is often the most challenging aspect (especially if a solid/porous material is involved; relaxation of structure)
- TOPOTOOLS (VMD), Materials Studio (requires a licence)

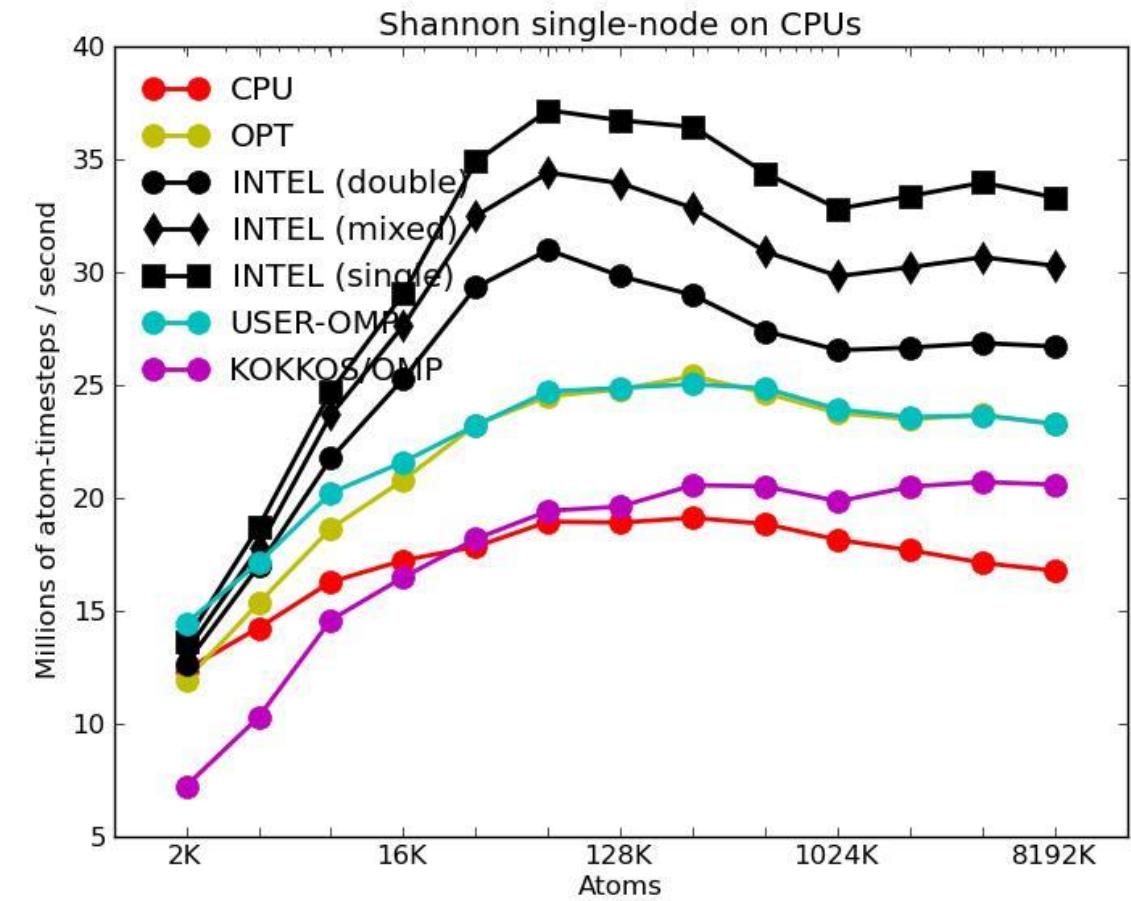


What can these tools do?

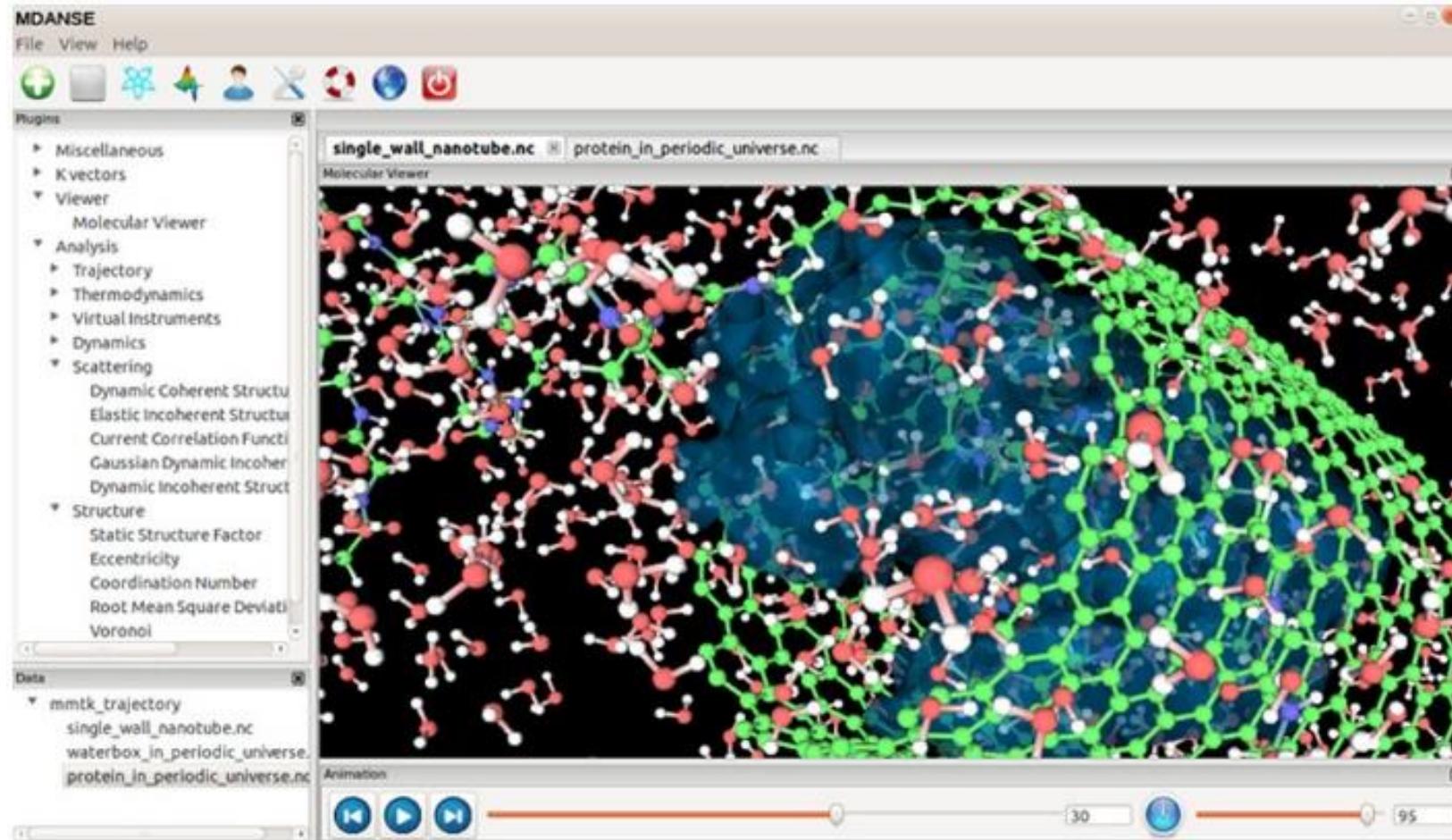
- Set up the geometry
- Set up the force-field
- Analyse results from the trajectory

Running on a cluster and parallelisation

- These calculations usually have to be run on a super computer cluster (**ISIS provides use of the SCARF cluster to users**)
- How many CPUs is efficient? After a certain point more CPUs is bad as the communication between CPUs is more expensive than the gain
- The best method is to run short tests with different numbers of CPUs on your system



We have our trajectory, now how do we compare to Neutron Scattering?



- MDANSE is the go to analysis package for extracting neutron weighted results
- We will now go through some of its features explicitly as there are some important things to consider when choosing different parameters
- Resolution, neutron cross-section weighting, projection onto relevant Q

Structure (Radial distribution function)

- Useful for comparison with SANS (E.g. Nimrod)
- Structure typically favours infrequent printing
(uncorrelated frames)

Pair Distribution Function

trajectory
C:\Users\fpn90659\Dropbox\outputWater.nc

frames
First frame 0 Last frame 19 Frame step 1

r values (nm)
from 0 to 10 by step of 1

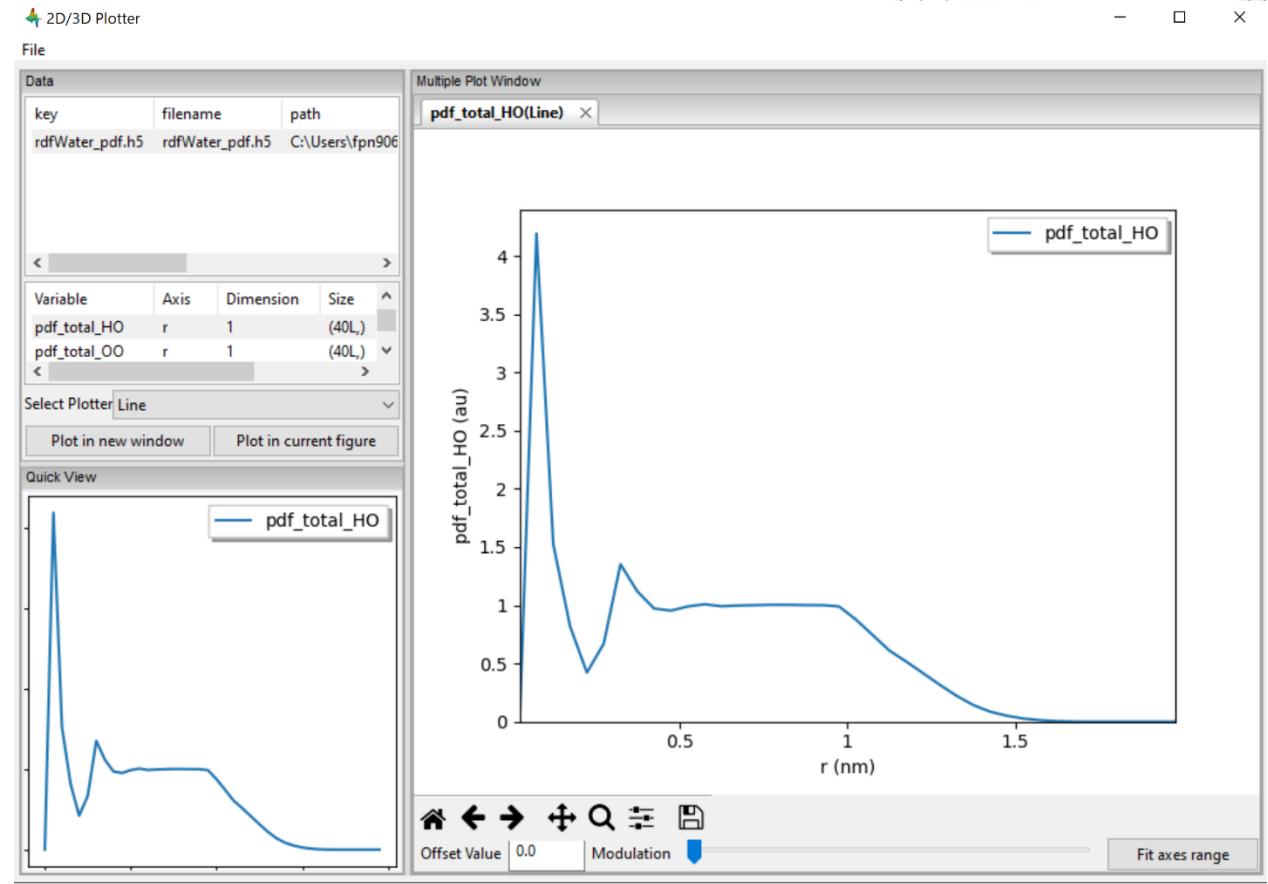
atom selection
Set new selection

atom transmutation
Set new selection

weights
equal

output files
Basename C:\Users\fpn90659\Dropbox\outputWater_pdf output formats

running mode
 monoprocessor multiprocessor 1



Plugins

Structure

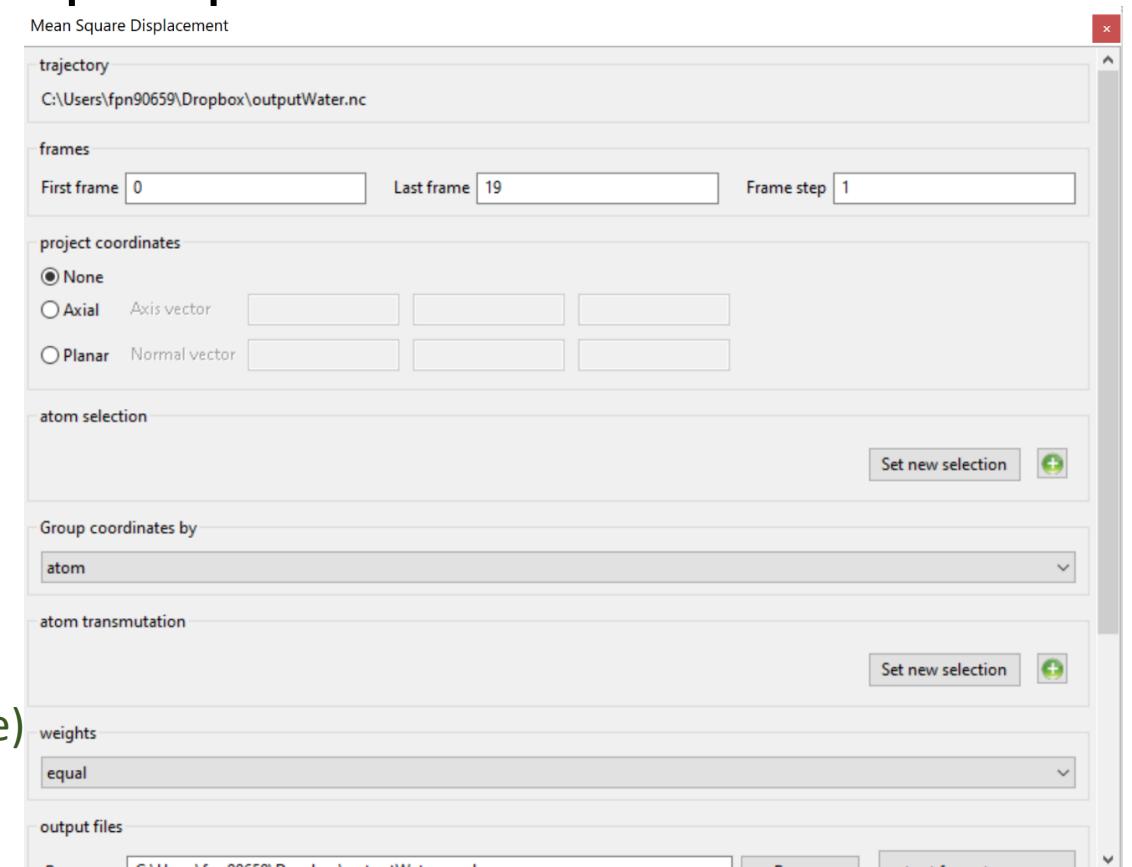
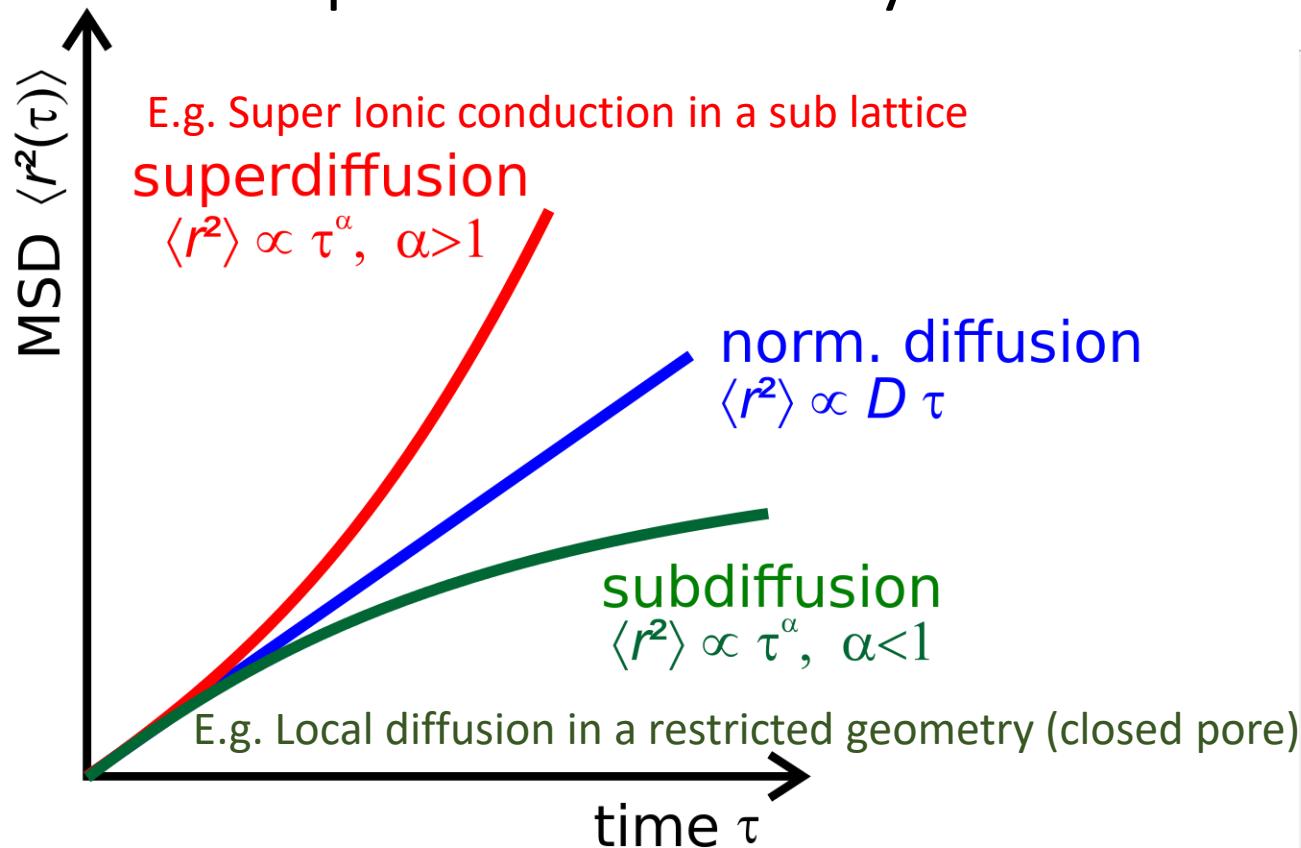
- Area Per Molecule
- Coordination Number
- Density Profile
- Eccentricity
- Molecular Trace
- Pair Distribution Function
- Root Mean Square Deviation
- Root Mean Square Fluctuation
- Radius of Gyration
- Solvent Accessible Surface
- Spatial Density
- Static Structure Factor
- Voronoi
- XRay Static Structure Factor

Thermodynamics

Trajectory

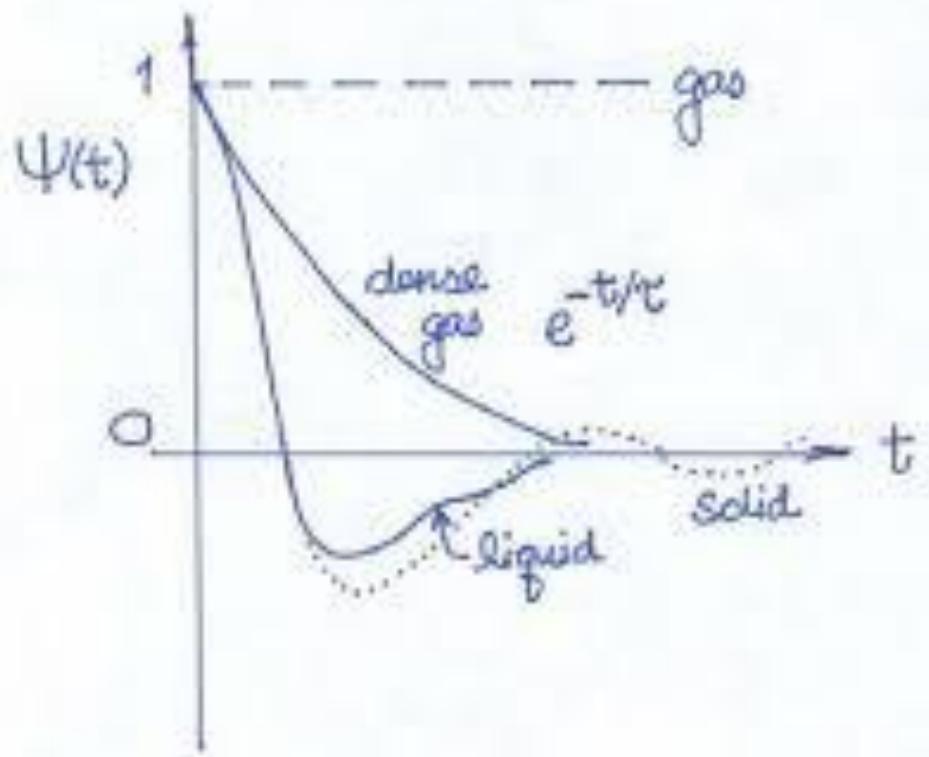
MSD and diffusion coefficients

- A metric to compare with QENS
- Gradient at long time asymptote gives diffusion coefficient
- Typically plotted over a 1 ns period (**depends of speed of diffusion**)
- Shape of MSD can vary for more complex phenomena



Velocity autocorrelation function

- Relates to vibrations, and also diffusion (integral of this is another route to the diffusion coefficient)
- Higher printing frequency required



Velocity AutoCorrelation Function

trajectory
C:\Users\fpn90659\Dropbox\outputWater.nc

frames
First frame 0 Last frame 19 Frame step 1

velocities
interpolation order

project coordinates
 None
 Axial Axis vector
 Planar Normal vector

normalize
 Yes

atom selection
Set new selection

Group coordinates by
atom

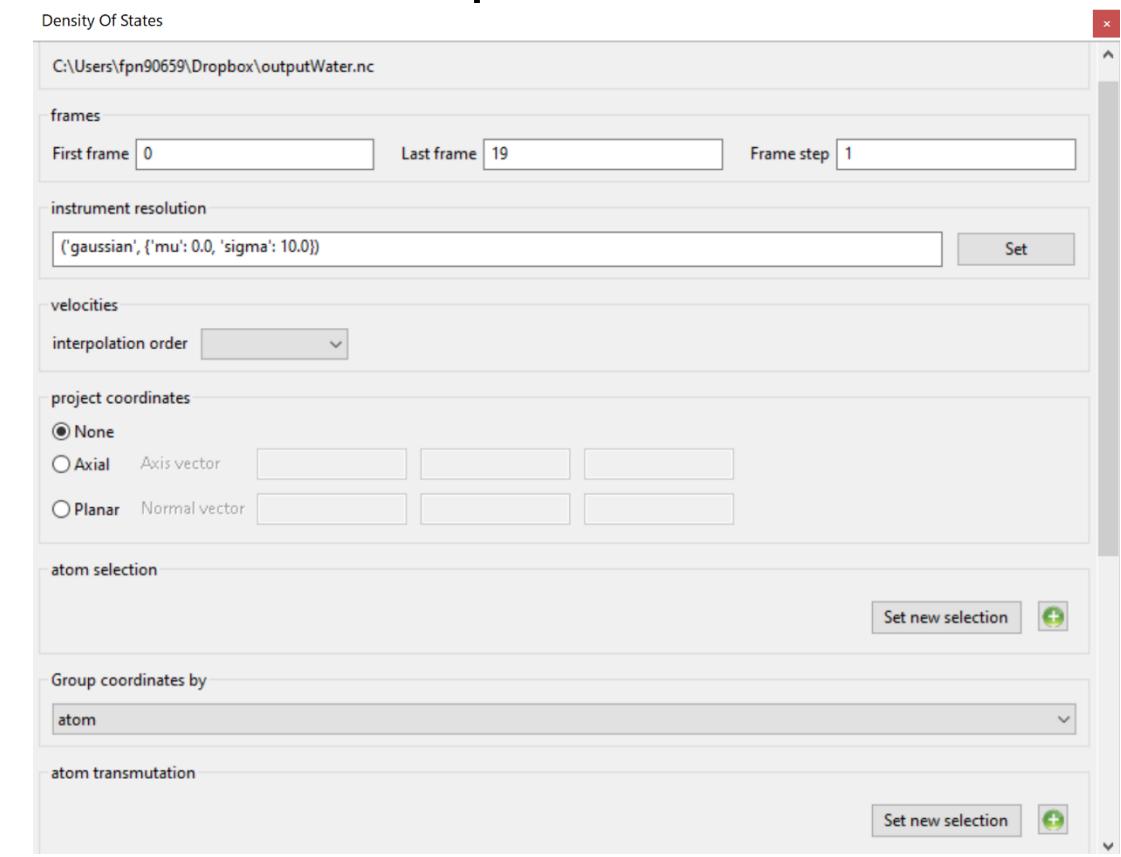
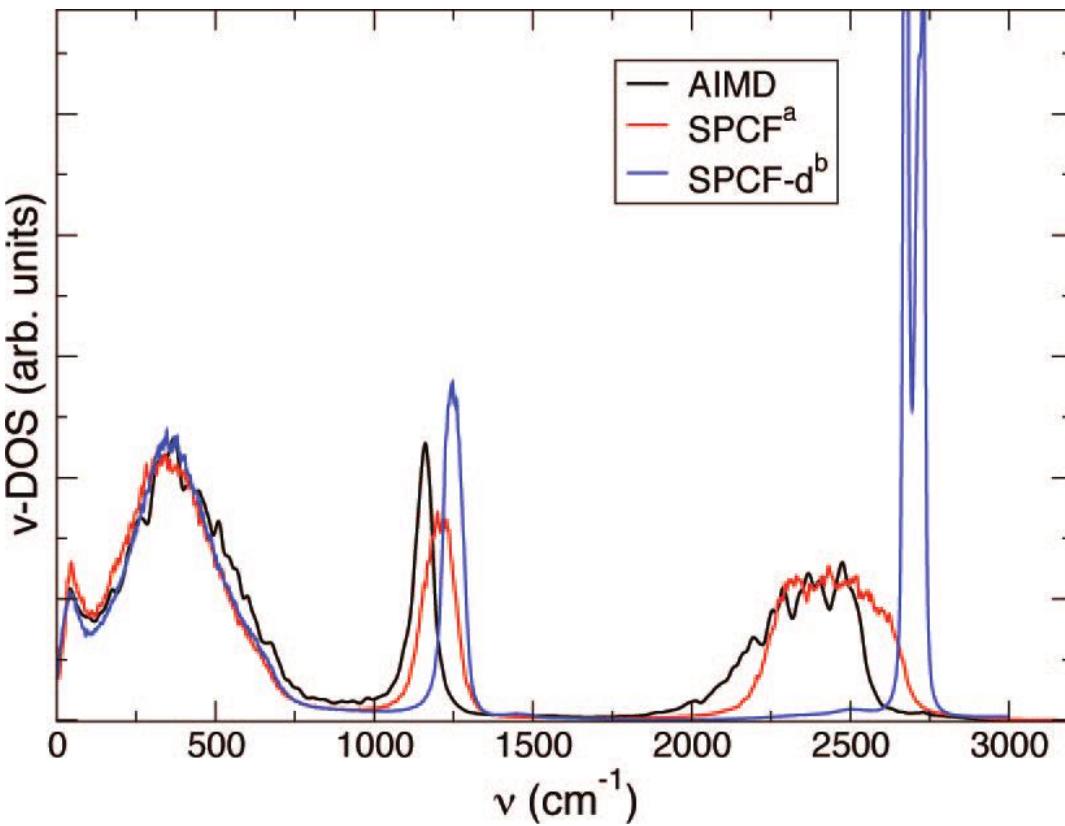
atom transmutation
Set new selection

weights
equal

output files

Vibrational density of states (FT of Vel ACF)

- This can be compared to TOSCA
- Resolution of peaks is included
- Can be weighted by neutron cross-section for correct Expt. relative peak intensity
- High print frequency as we are now interested in vibrations (**fast motions**) as well as quite long if we are dealing with a solid, to include slower phonons



Angular Correlation Function

We must define a frame of reference for our angular vector:

Typically the vector between two atoms in the molecule

Atoms list

Number of atoms 2

Molecules Selected atoms

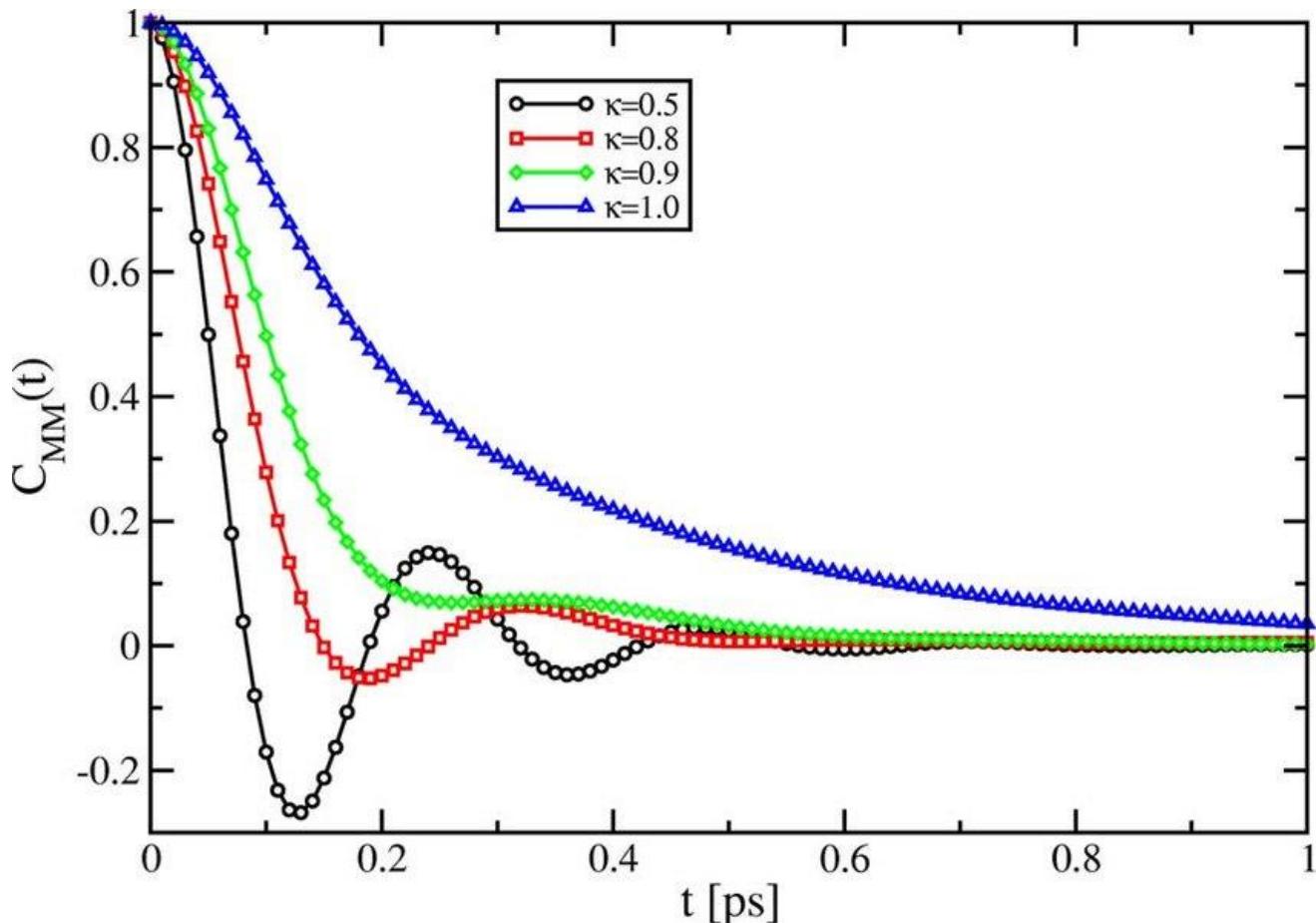
spce water

- HW1
- HW2
- OW0**

Atom 730 : HW1 ; Atom 729 : OW0
Atom 733 : HW1 ; Atom 732 : OW0
Atom 736 : HW1 ; Atom 735 : OW0
Atom 739 : HW1 ; Atom 738 : OW0
Atom 742 : HW1 ; Atom 741 : OW0
Atom 745 : HW1 ; Atom 744 : OW0
Atom 748 : HW1 ; Atom 747 : OW0
Atom 751 : HW1 ; Atom 750 : OW0
Atom 754 : HW1 ; Atom 753 : OW0
Atom 757 : HW1 ; Atom 756 : OW0
Atom 760 : HW1 ; Atom 759 : OW0
Atom 763 : HW1 ; Atom 762 : OW0
Atom 766 : HW1 ; Atom 765 : OW0

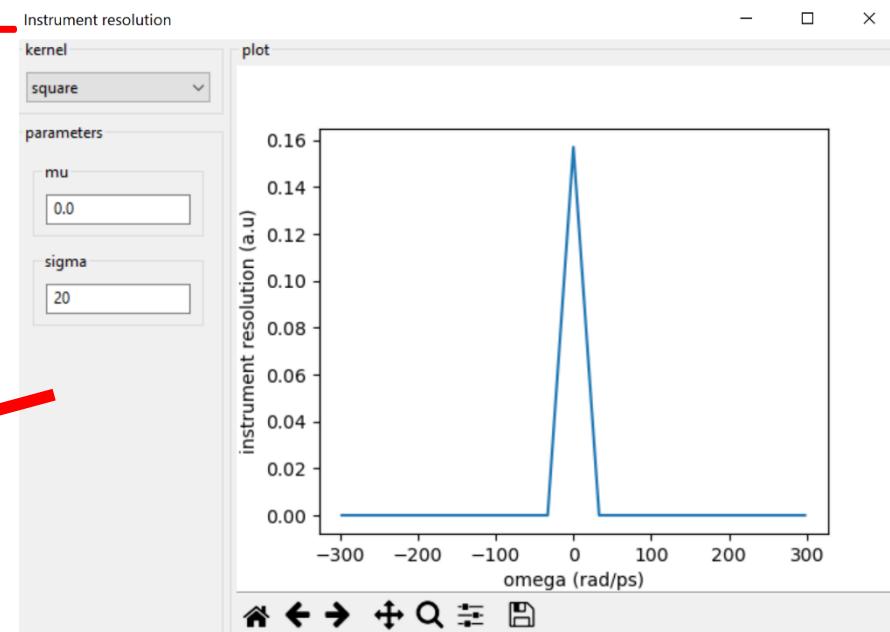
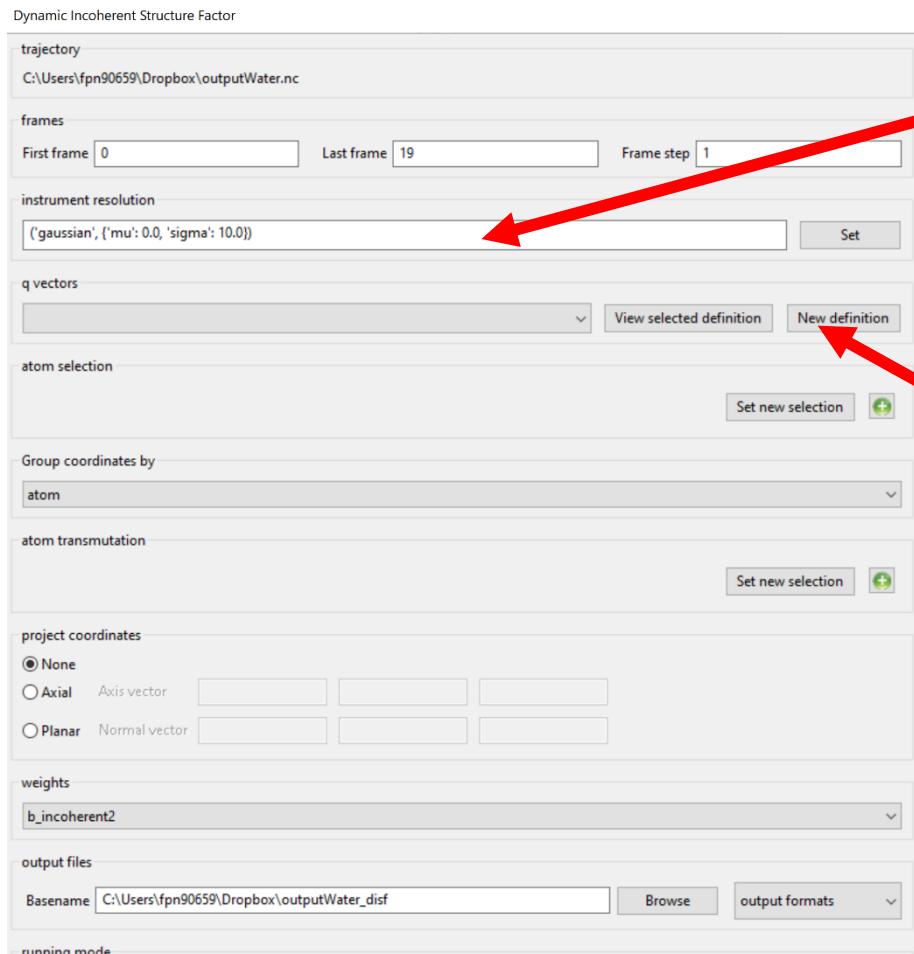
Save

- This can be useful for comparing to QENS
- Extract rotational signal
- Find relaxation times for different molecular units in proteins

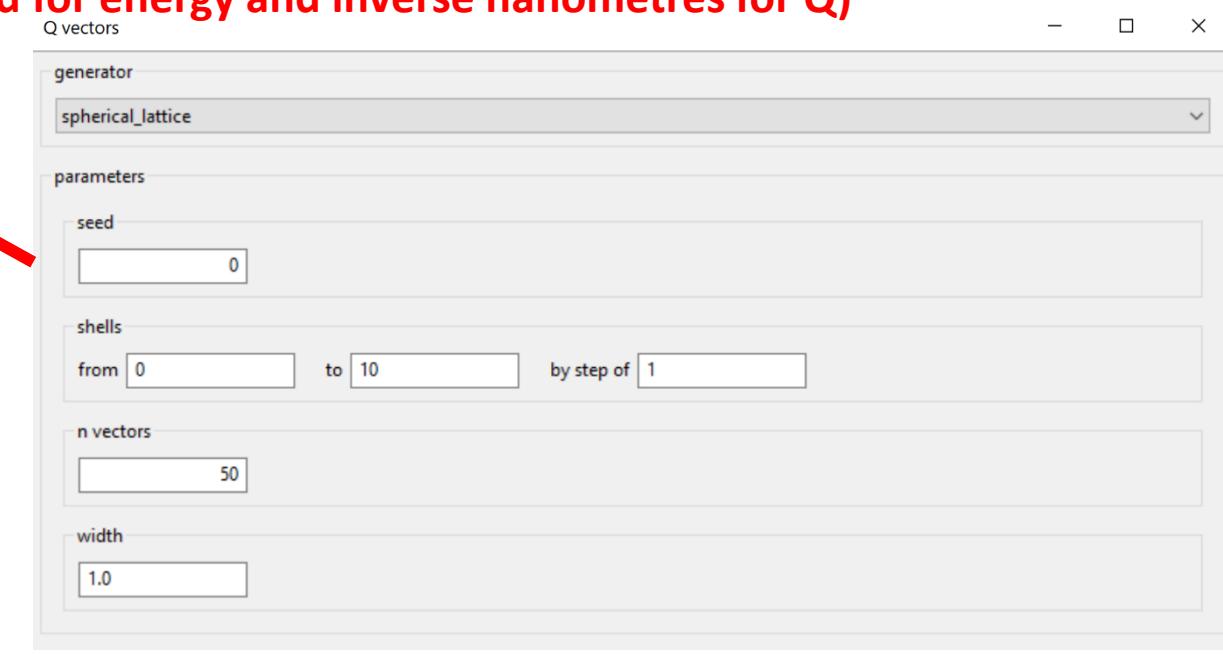


Dynamic Structure factor: IMPORTANT

- Note that **Incoherent** and **Coherent** are calculated separately
- A Q grid must be created (**think carefully about ranges**)
- A resolution function must be specified (**usually a Gaussian**)
- The signals will be weighted by the neutron cross-section

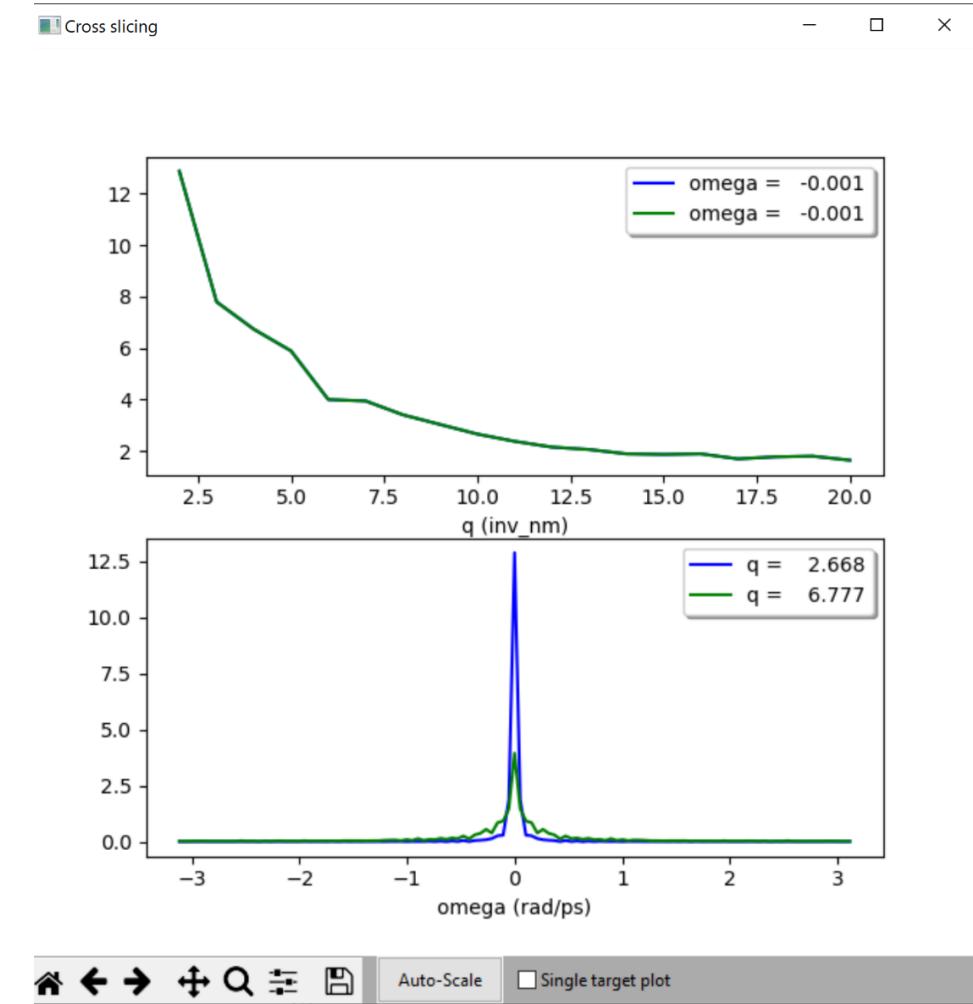
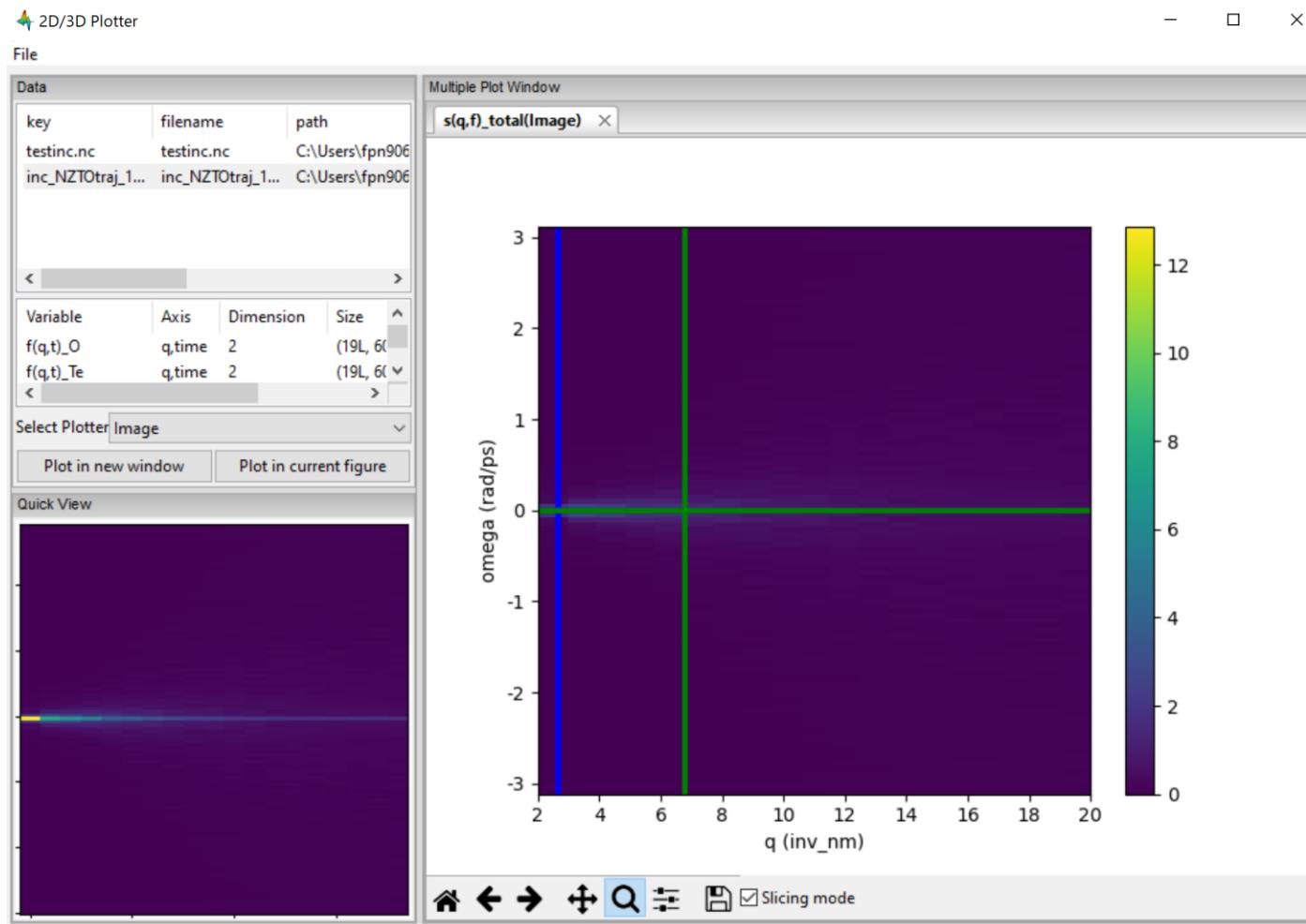


Cautionary warning! MDANSE uses very unusual units (e.g. radians per picosecond for energy and inverse nanometres for Q)



Dynamic Structure factor

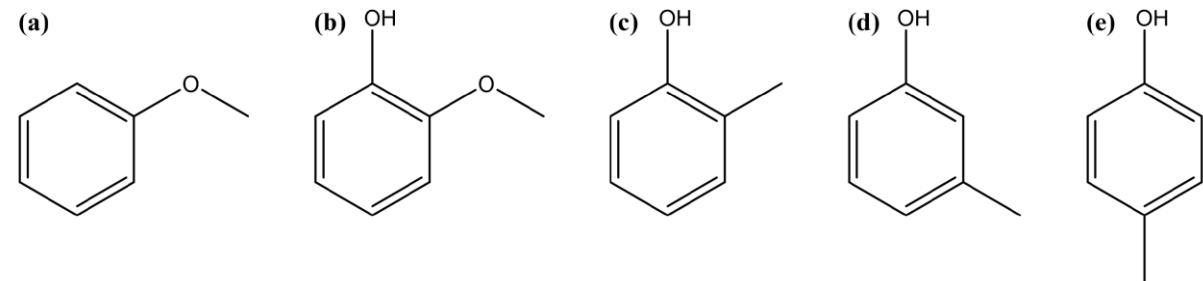
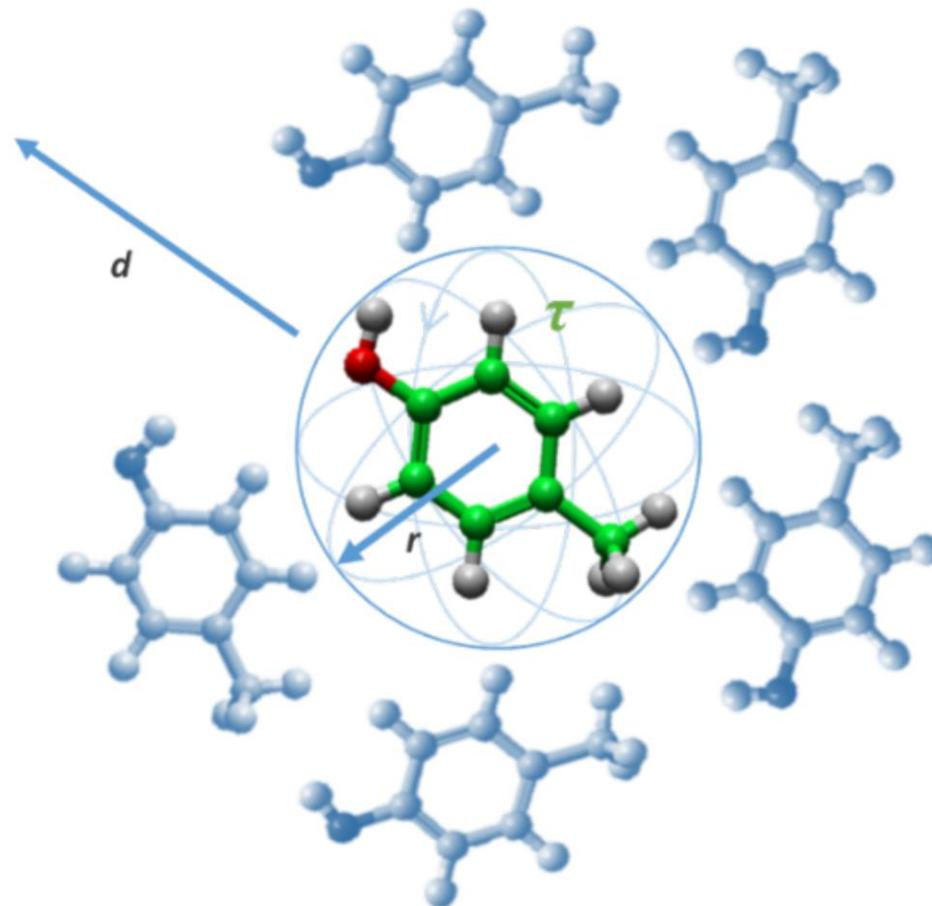
- MDANSE has an internal visualizer, which allows you to take slices in E/Q, which is very useful for deducing trends
- Also useful for working out which regions can be ignored (e.g. Bragg peaks)



Case study: Liquid lignin monomers

What information can we extract?

Potential pitfalls?

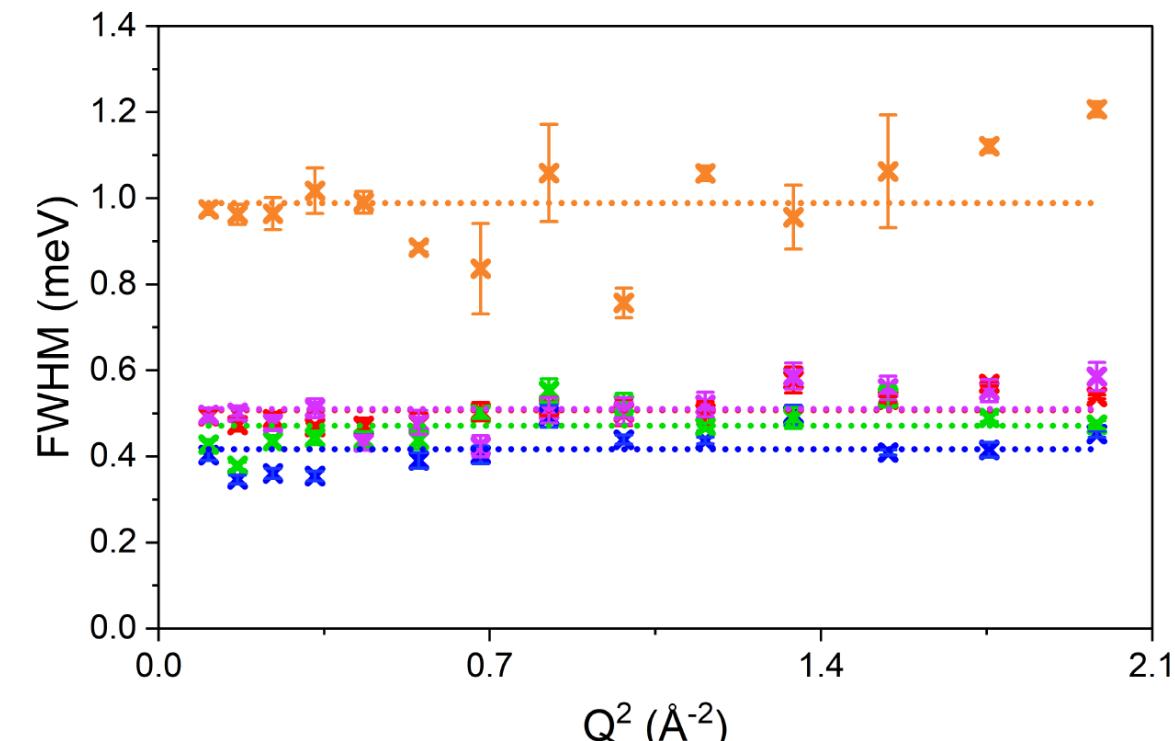
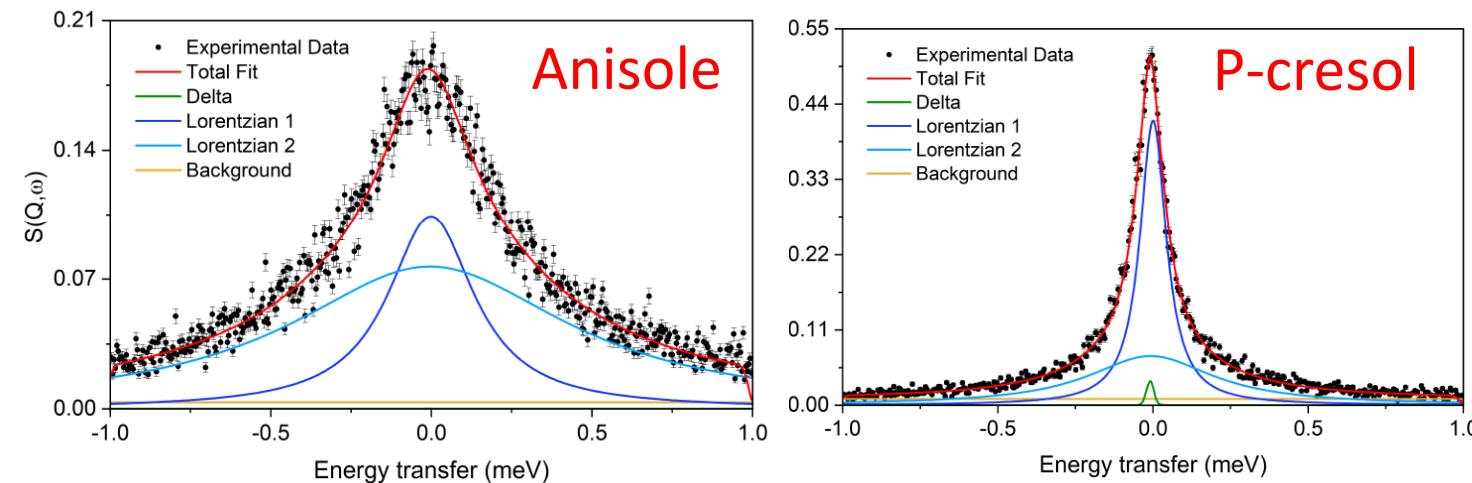
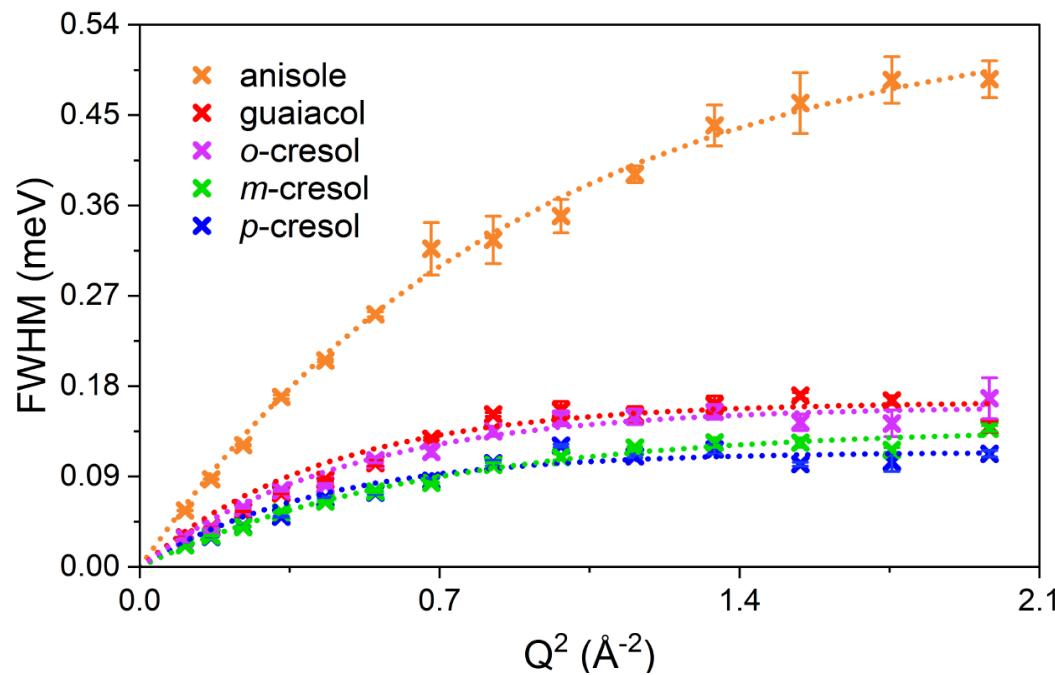


- Simulations performed on ~5000 atoms
- Run times of ~10 ns
- Run in DL POLY
- Trajectories output every 1000 steps for Structural and Slow dynamics, and every 10 for faster dynamics.

QENS Experiment

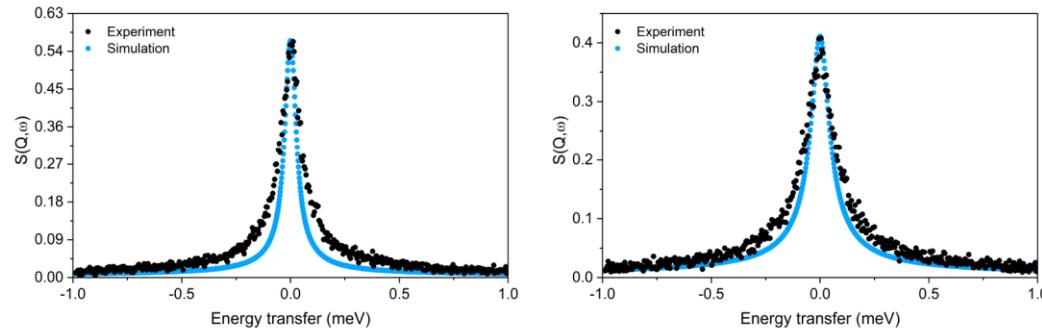
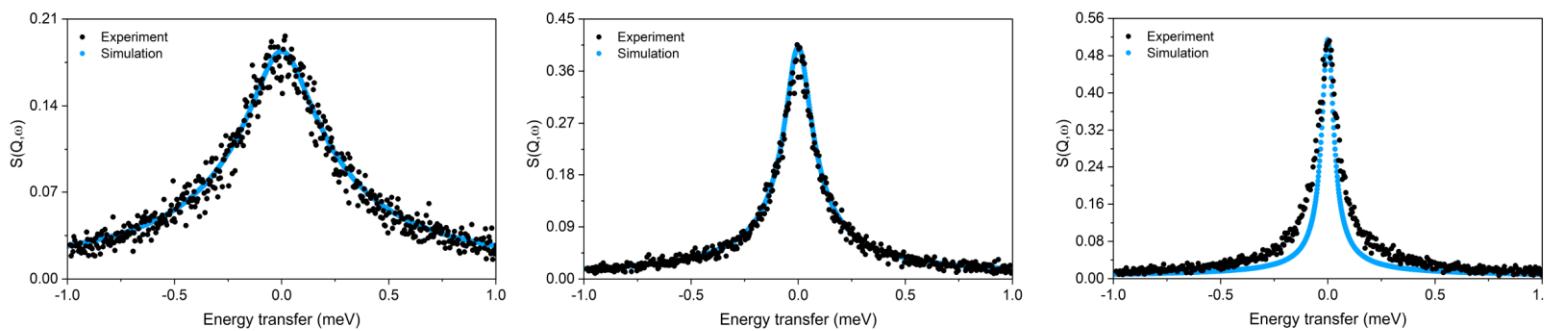
- Fitting of two Lorentzians
- One for **translational hopping**
- One for **local rotation**

This style of fitting can be performed in Mantid (**conv fit**)



We can generate the $S(Q, E)$ from simulation and perform exactly the same analysis

The accuracy of a model can be determined by how well the $S(Q, E)$ matches



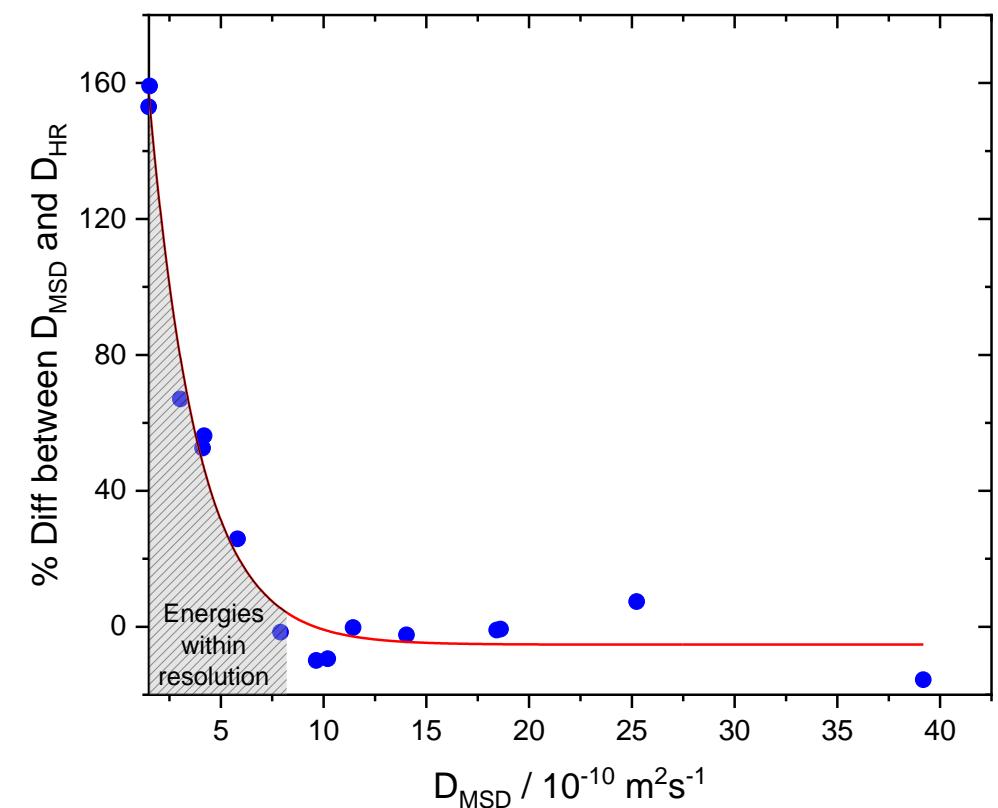
Same fittings applied to Expt and Sim
(Extracted diffusion coefficients)

	Expt HR D	Sim HR D
Anisole	330	19.66
	360	23.345
	390	35.955
E activation	10.61	10.45
	330	6.98
	360	14.375
Guaiacol	390	18.01
	17.11	16.23
	330	9.475
o-cresol	360	12.635
	390	20.71
	E activation	13.705
m-cresol	6.905	17.66
	330	4.46
	360	7.885
<i>p</i> -cresol	390	13.165
	E activation	8.69
	19.29	13.92
	330	6.01
	360	10.33
	390	13.3
E activation	14.28	9.25
	16.08	

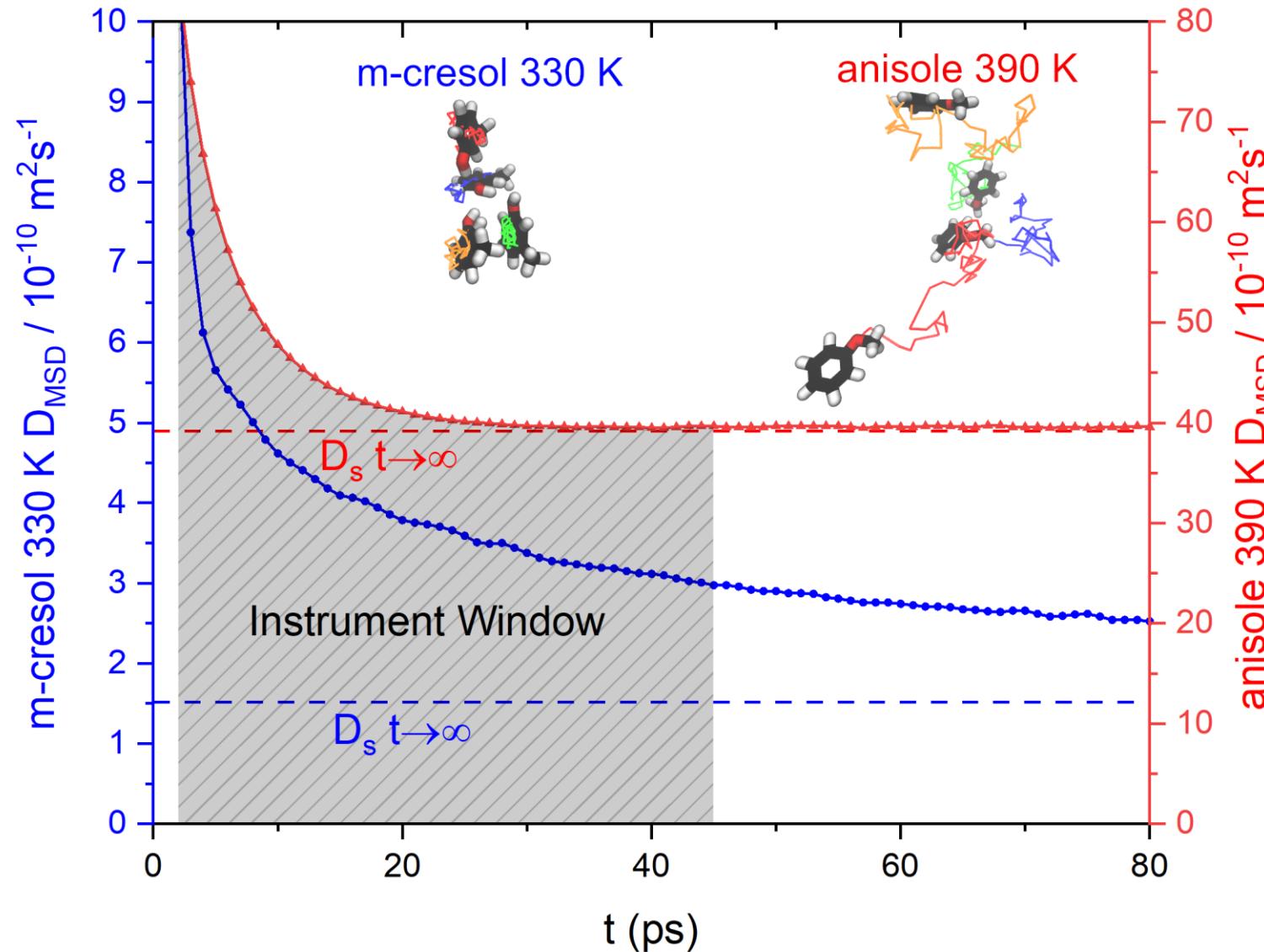
Diffusion from MSD

- MSD provides the “true” self-diffusion coefficient
- Major differences between this value and the fitted models for many simulations
- This same error therefore occurs during experimental fittings

	Temperature (K)	% Difference from MSD
Anisole	330	-0.698549167
	360	7.408874802
	390	-15.58192956
Guaicol	330	25.85910653
	360	-0.262237762
	390	-1.003798155
o-cresol	330	66.99669967
	360	-1.580278129
	390	-2.386039886
m-cresol	330	159.0909091
	360	52.54854369
	390	-9.948186528
p-cresol	330	153.0201342
	360	56.2052506
	390	-9.402546523

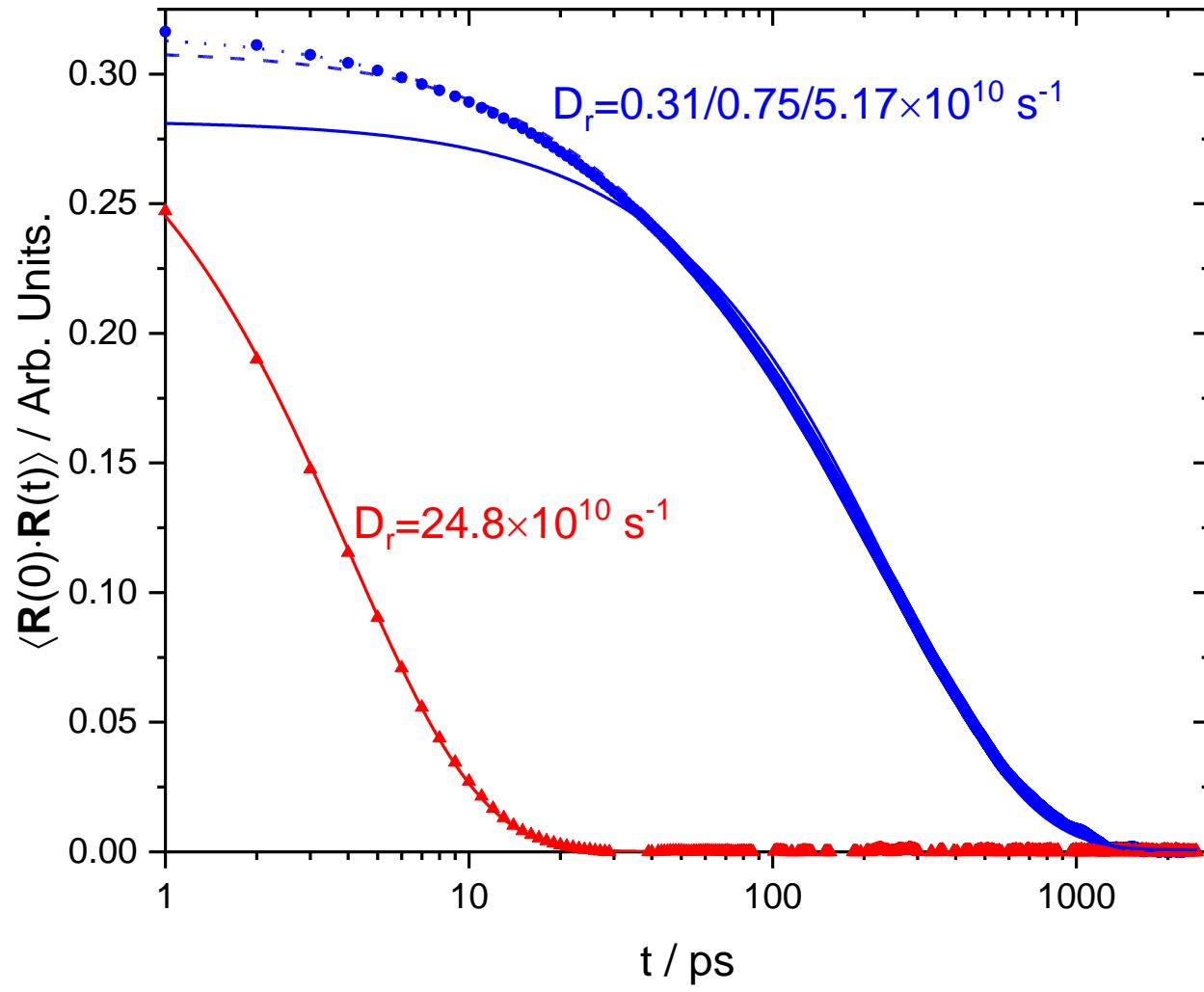


Why do we see this deviation?

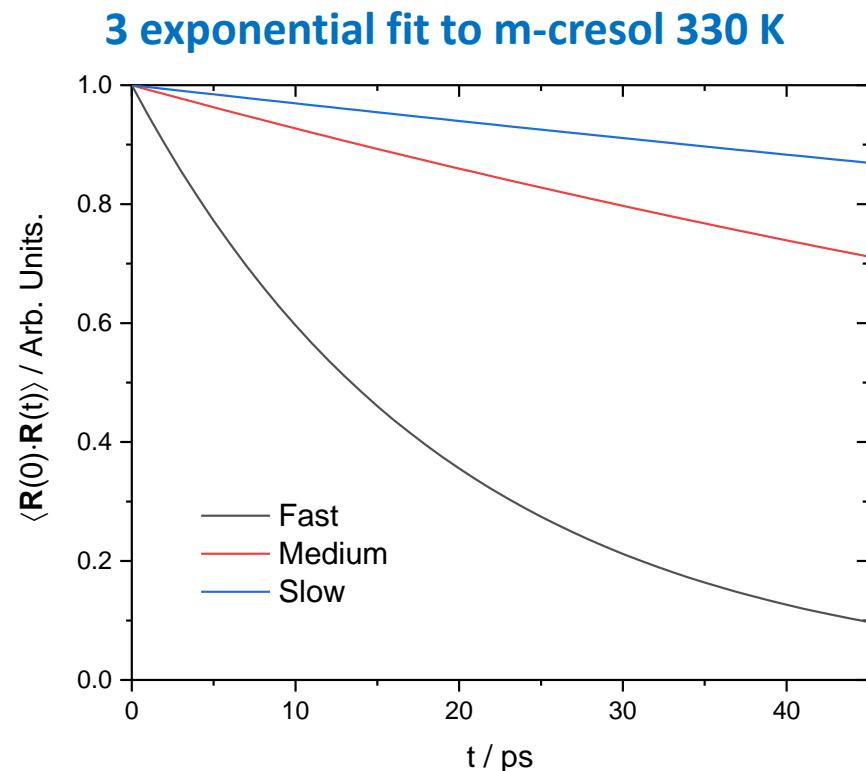


- The first derivative of the MSD serves as a time-dependent diffusion coefficient
- As each instrument has a limited time-scale, it can only sample motions within this range
- The fastest systems in this study arrive at the self-diffusion asymptote within the instrument range, whereas the slower systems have not reached this “Fickian” regime

What about the rotations?

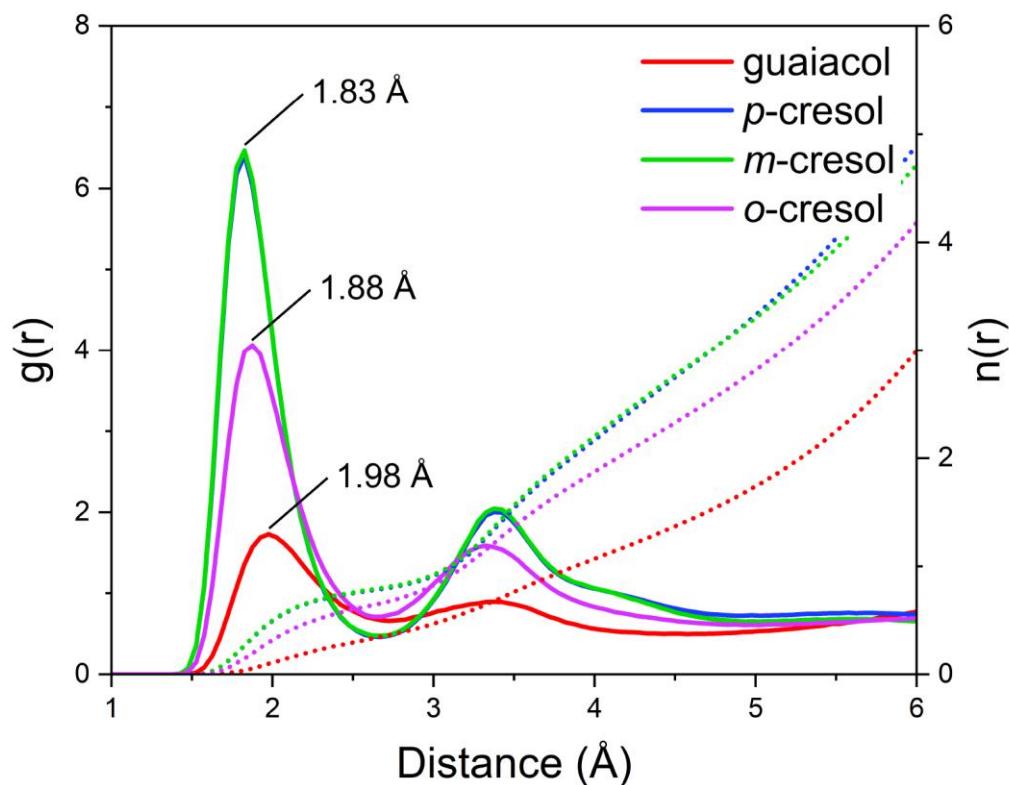


- Accurate for faster systems (E.g. Anisole 390 K)
- Slower systems display multiple exponential decays (~ 3 for m-cresol 330 K) with different time scales
- The largest portion of motions come from the slowest relaxation, which is 30 times slower than that extracted from FWHM fitting
- Fittings in slow regime in general are inappropriate. Likely mostly fitting local motions (delta function probably accounting for some of these very slow motions)



Radial distribution functions can highlight hydrogen bonding trends

- Increased hydrogen bonding indicated by higher O-H peak height
- This can then explain differences in diffusion/rotation between models



p-cresol and *m*-cresol

Strong h-bonding, sharp peaks
at close distances

o-cresol

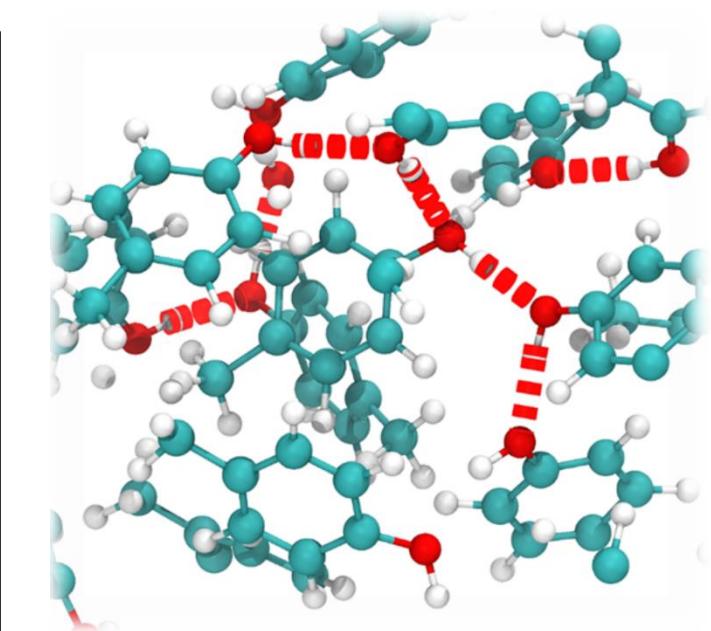
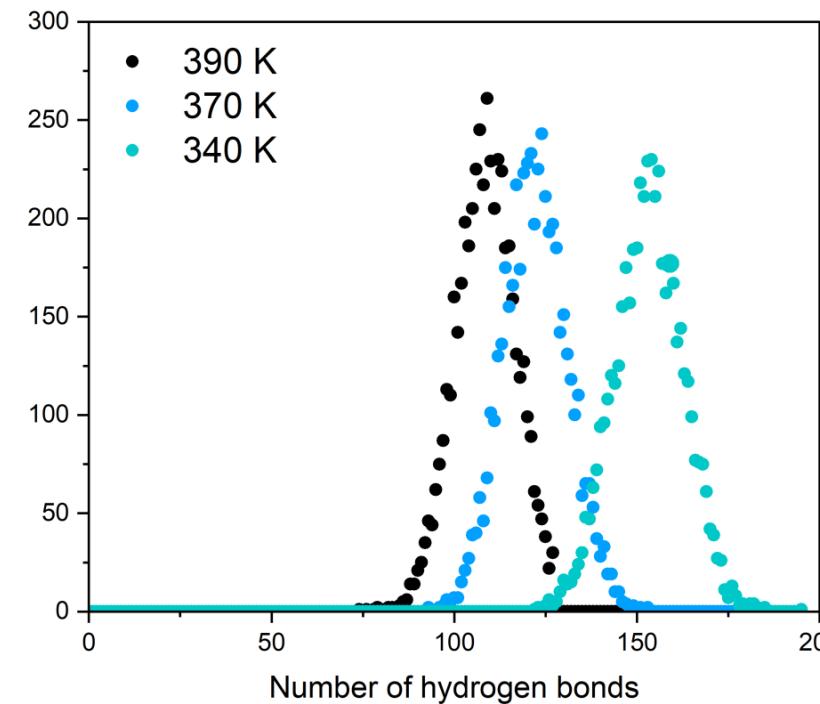
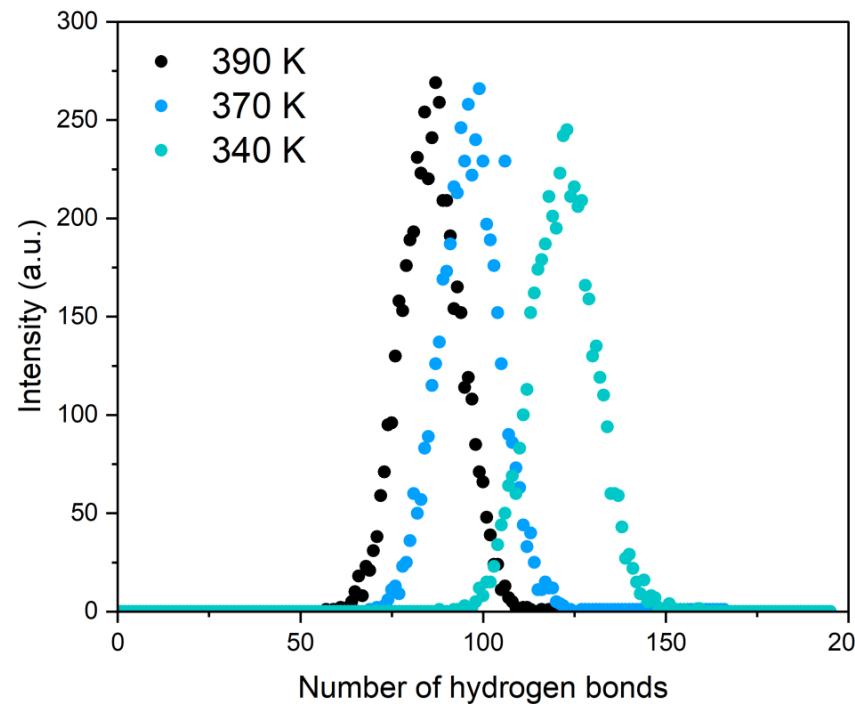
h-bonding restricted by adjacent methyl group,
larger bonding distance

guaiacol

Bulky groups, broader range of bonding
distances due to steric hindrance

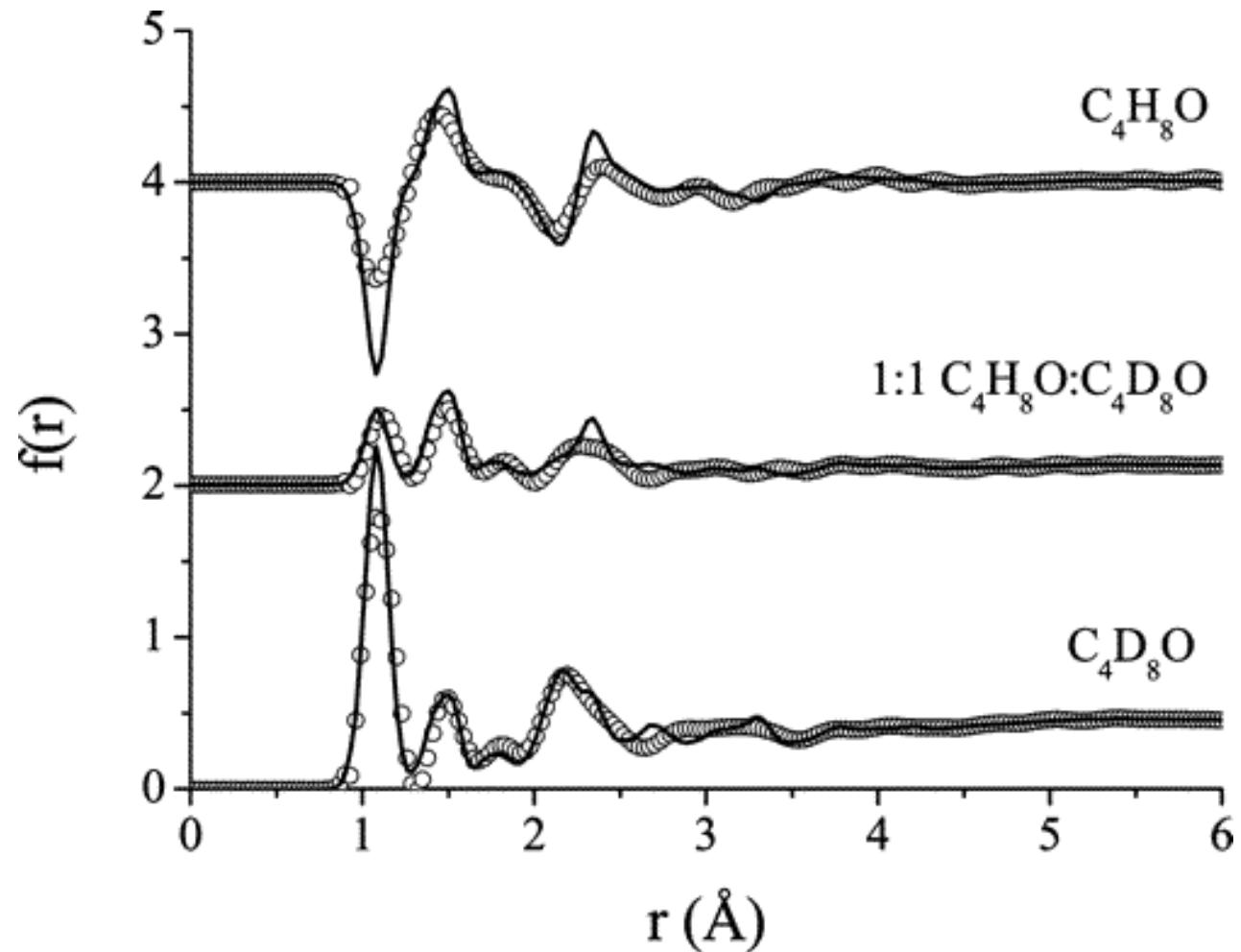
H-bond distributions (from VMD)

- H-bonds can be detected with certain distance and angle criteria
- These distributions can be evaluated over the course of the simulation
- They can also be plotted visually
- More complex method is to look at h-bond lifetimes



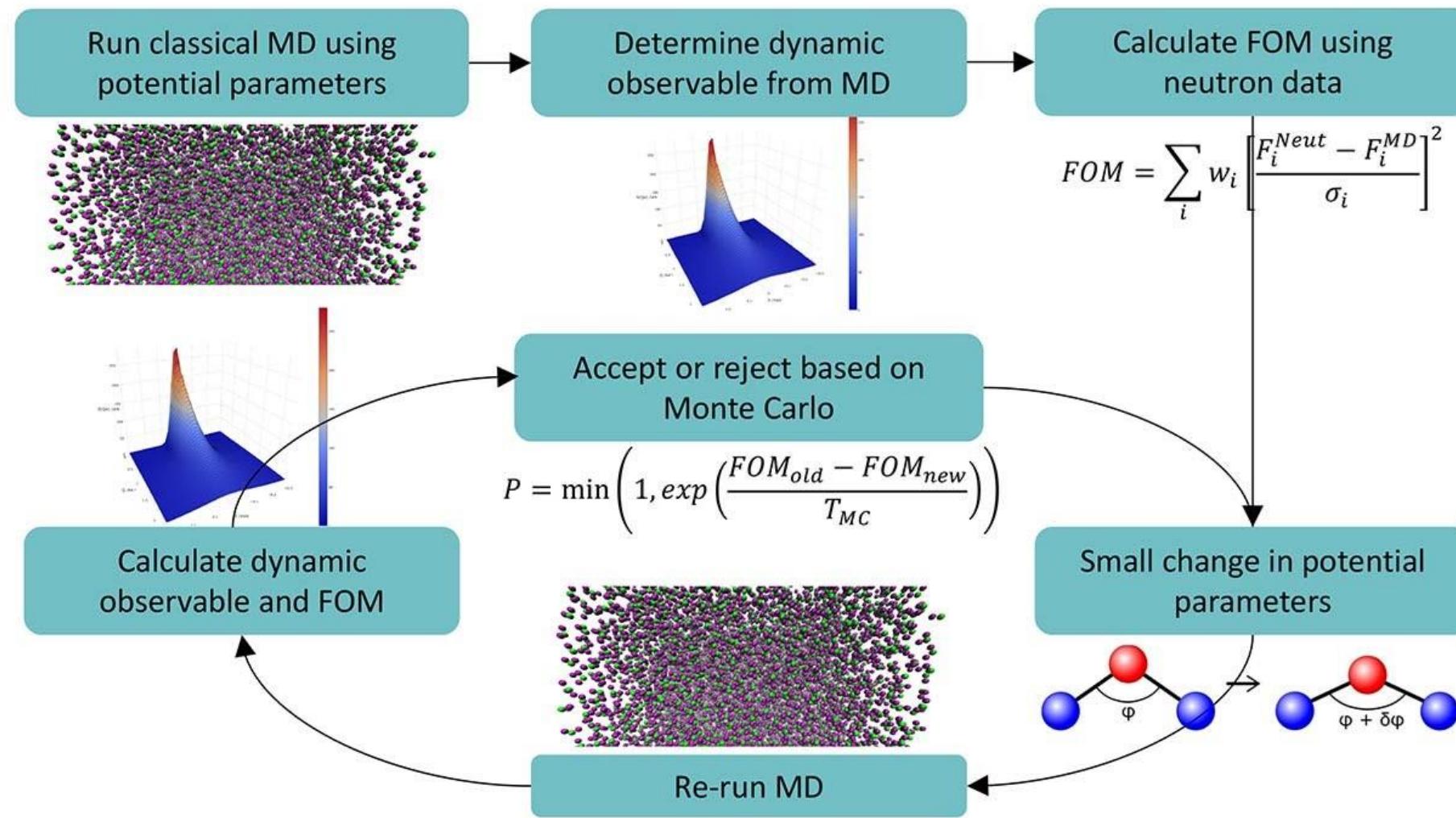
Matching Structure with MD/MC (EPSR)

- Simulations, coupled with a progressive alteration of the force-field parameters via a figure of merit (**deviation from Expt. data**)
- Convergence towards a force-field which reproduces the neutron results
- Force-fields may be meaningless from a dynamics point of view (**configurational space only correct**)



Fitting MD models with Neutron Data (MDMC)

The Algorithm



Questions?

