Computer simulation of glasses
(a brief introduction)

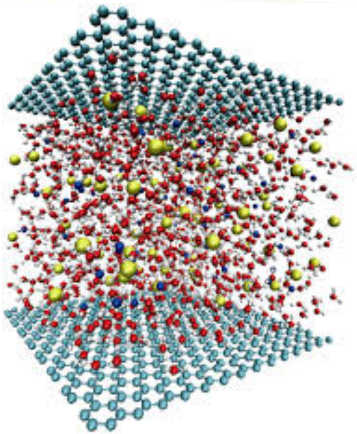
Miguel A. Gonzalez

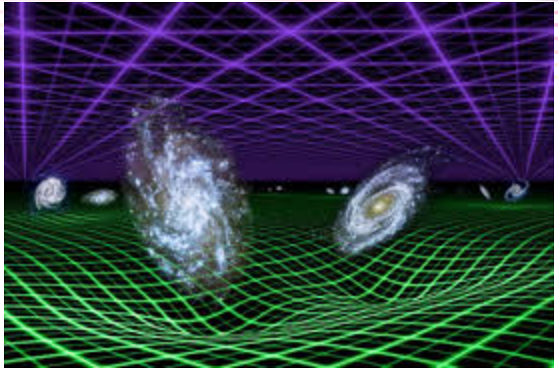
Institut Laue-Langevin, Grenoble (France)

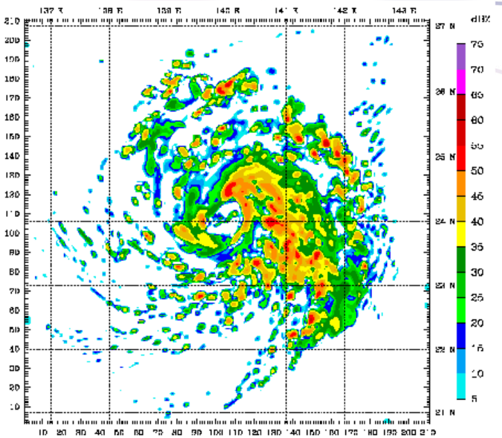
Plan

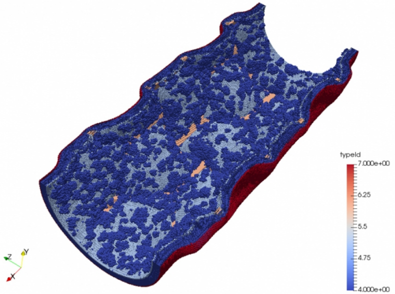
- Introduction
  - ✓ Do we need computer modelling? What kind of simulations?
- Representing intermolecular forces
  - ✓ Electronic structure methods → Density functional theory **(DFT)**
  - ✓ **Force fields**
- Sampling phase space
  - ✓ **Monte Carlo**
  - ✓ **Molecular Dynamics**
- Analysis
- Some examples

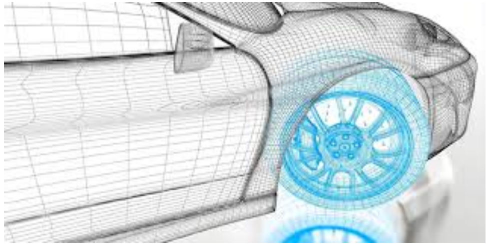
Google search: Computer simulation

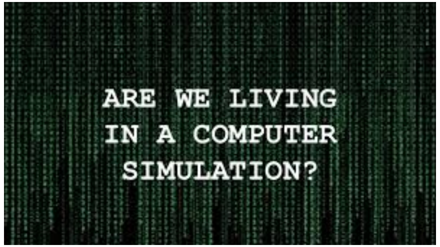












Introduction

Do we need computer simulations?

EXCELLENT COMPUTER SIMULATIONS ARE DONE FOR A PURPOSE. THE MOST VALID PURPOSES ARE TO EXPLORE UNCHARTED TERRITORY, TO RESOLVE A WELL-POSED SCIENTIFIC OR TECHNICAL QUESTION, OR TO MAKE A GOOD DESIGN CHOICE.

Leo P. Kadanoff: “Excellence in computer simulation”, *Comp. Sci. Eng.* **6**, 57-67 (2004).

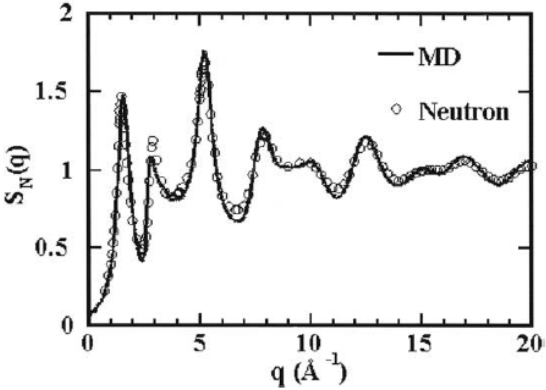
- Only in rare cases simulations have discovered new phenomena (e.g. long time tails)
- Typical role is to work with theory and experiment and serve to test models and theories in order to understand a given phenomenon
- Caution: Quite often simulations can be wrongly used to support our previous beliefs

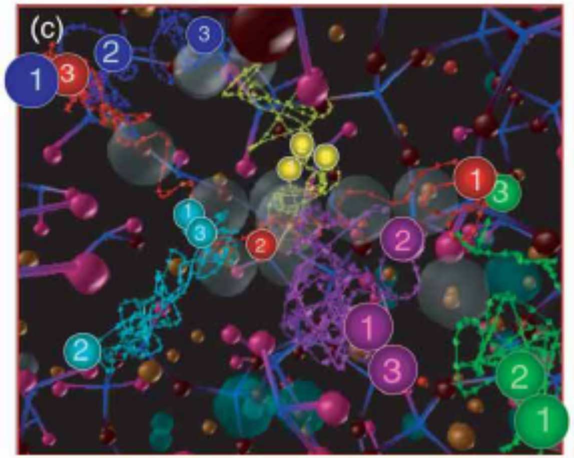
“Experiment is very good at finding unexpected behavior and describing its overall characteristics; theory often can explain what’s going on. After an appropriate pause for algorithm development, simulations then can test the ideas and fill in the details.”

Introduction

Do we need computer simulations?

- Experimental information is insufficient to determine fully the structure of a glass (spherical average).
- Simulations can supply very detailed information at the atomic level that serves to:
  - Interpret and understand the experiment.
  - Design new experiments and improve experimental conditions.
  - Test theories and predictions.
- The main purpose is not to reproduce exactly the experimental result, but to provide a useful model of our system (but naturally the computer model has to be able to reproduce experimental results reasonably well in order to be considered a realistic model).





Introduction

What kind of simulations? Compromise: Accuracy vs efficiency

From Kermode et al., in *Multiscale Simulation Methods in Molecular Sciences,*
J. Grotendorst, N. Attig, S. Blügel, D. Marx (Eds.), NIC Series, Vol. 42, pp. 215-228 (2009).

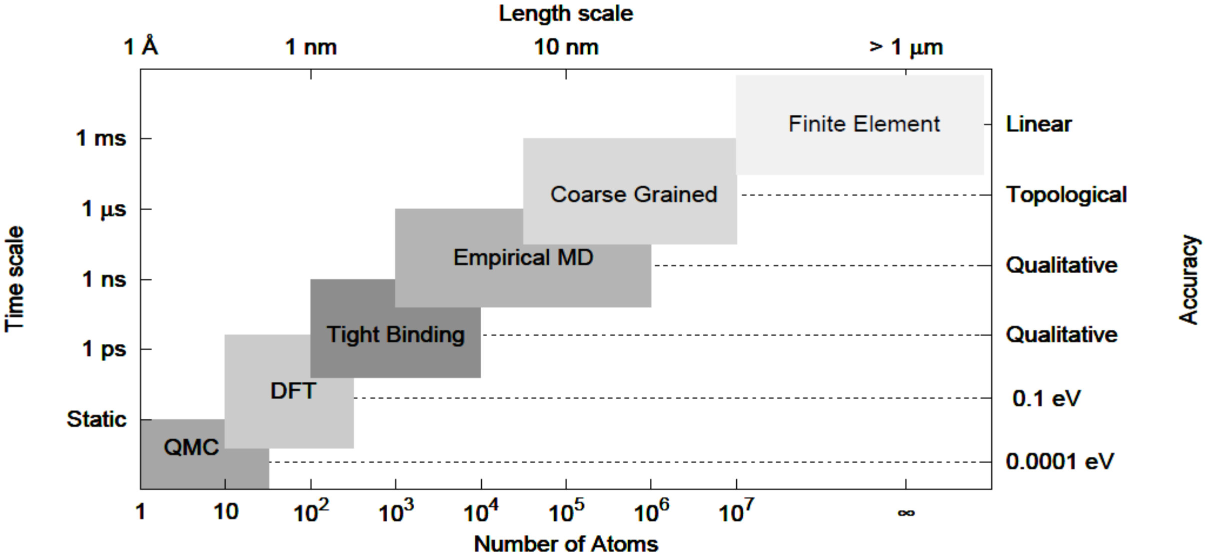


Figure 1. Schematic representation of the range of length- and time-scales accessible to a variety of modelling methods, from quantum Monte Carlo (QMC) for very accurate, very expensive static calculations through to approximate methods such as finite-element modelling.

Introduction

What kind of simulations?

Choose way of computing interatomic forces

Hartree-Fock methods

Semiempirical methods

Density Functional Theory

Tight-binding

Empirical force fields

Coarse graining

Brownian Dynamics

Langevin Dynamics

Dissipative Particle Dynamics

Choose way of sampling the phase space

Lattice Dynamics

$$D_{ij}(\mathbf{k}) = \frac{1}{\sqrt{m_i m_j}} \sum_{\ell} \Phi_{ij}(0\ell) \exp\left(i\mathbf{k}\cdot[\mathbf{r}_i(0) - \mathbf{r}_j(\ell)]\right)$$

Thermodynamics
Phonons

Monte Carlo

$$\pi_a(\Gamma \to \Gamma') = \min(1, \exp(-\beta \Delta H))$$

Thermodynamics
Structure

Molecular Dynamics

Computing energies and forces

1. Electronic structure methods

Interatomic forces

Electronic structure

Solve Schrödinger equation:
Generally using Born-Oppenheimer approximation (and non relativistic)

$$\boxed{\hat{H}\Psi = [\hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{int}} + E_{\text{nucl}}]\Psi = E\Psi}$$

⬇

Obtain potential energy surface

$$\hat{T} = \text{kinetic energy of electrons} = \sum_i -\frac{1}{2}\nabla_i^2$$

$$E_{\text{nucl}} = \text{classical interaction of nuclei} = \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J}{|R_I - R_J|} = \text{constant}$$

$$\hat{V}_{\text{ext}} = \text{potential due to nuclei} = -\sum_{i,I}\frac{Z_I}{|r_i - R_I|}$$

$$\boxed{\hat{V}_{\text{int}} = \text{electron−electron interaction} = \frac{1}{2}\sum_{i\neq j}\frac{1}{|r_i - r_j|}} \longrightarrow$$

Many-body problem
Difficult!

Electronic structure

Hartree-Fock methods

$$\Psi \approx \phi_1 \phi_2 \dots \phi_N \quad \longrightarrow \quad \hat{H}_{\text{eff}}\,\phi_i = \left[-\frac{1}{2}\nabla_i^2 + V_{\text{eff}}\right]\phi_i = \varepsilon_i\,\phi_i$$

$$\text{with } V_{\text{eff}} = -\sum_I \frac{Z_I}{|r_i - R_I|} - \int \frac{n(r')}{|r' - r_i|}\,dr' = \boldsymbol{\hat{V}_{\text{ext}}(\{R_I\}, r_i) + \hat{V}_{\text{H}}(r_i)}$$

A first improvement needed is to force Ψ to be antisymmetric (Pauli principle):

$$\Psi \propto \begin{vmatrix} \emptyset_1(r_1) & \emptyset_1(r_2) & \cdots & \emptyset_1(r_N) \\ \emptyset_2(r_1) & \emptyset_2(r_1) & \cdots & \emptyset_2(r_1) \\ \vdots & \vdots & \vdots & \vdots \\ \emptyset_N(r_1) & \emptyset_N(r_1) & \cdots & \emptyset_N(r_1) \end{vmatrix} \quad \longrightarrow \quad \begin{aligned} &\left[-\frac{1}{2}\nabla_i^2 + \hat{V}_{\text{ext}}(\{R_I\}, r_i) + \hat{V}_{\text{H}}(r_i)\right]\phi_i \\ &- \sum_j \int \frac{\emptyset_j^*(r')\emptyset_i(r')}{|r' - r_i|}\,\emptyset_j(r)\,dr' = \varepsilon_i\,\phi_i \end{aligned}$$

Slater determinant

Exchange potential

Electronic structure

Hartree-Fock methods

- Solve self-consistently single-electron effective Hamiltonian (SCF method)
- Misses correlation energy (except for the exchange energy)
- HF predicts reasonable molecular geometries and vibrational frequencies for many molecules, but total energies are less satisfactory
- Not suitable for metals
- Accuracy can be improved→ Post-HF methods: Perturbation theory (MP2, MP3, MP4), Configuration interaction, Coupled cluster, … and Quantum Monte Carlo
- **But computational cost makes those methods impractical for the typical problems in solid state physics, soft-matter or biophysics studied with neutron scattering**

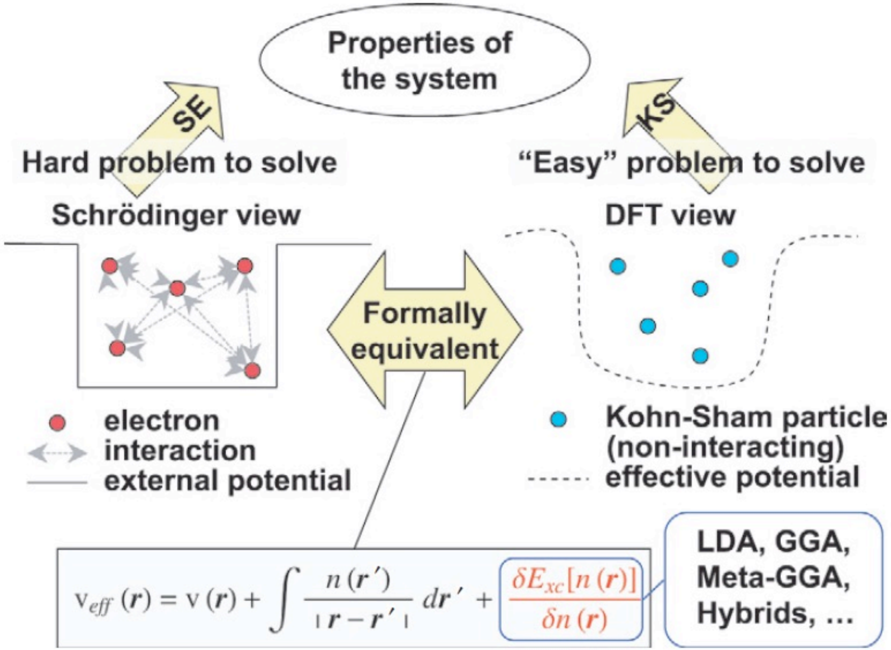
Electronic structure

DFT

The Kohn-Sham equations are formally **equivalent** to the Schrödinger equation.

They recast the many-body problem into a set of coupled, **single-particle** equations, and isolate all our ignorance about many-body interactions into the **exchange-correlation functional**.

We solve a much simpler, but equivalent, fictitious system of non-interacting "Kohn-Sham" particles in an effective potential.



$$v_{eff}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}' + \frac{\delta E_{xc}[n(\mathbf{r})]}{\delta n(\mathbf{r})}$$

A. E. Mattsson et al.: "Designing meaningful density functional theory calculations in materials science – a primer", *Modelling Simul. Mater. Sci. Eng.* **13**, R1-R31 (2005)

DFT The Hohenberg-Kohn theorems

P. Hohenberg and W. Kohn: "Inhomogeneous electron gas", *Phys. Rev.* **136**, B864 (1964)

I. For any system of interacting particles in an external potential $V_{ext}(\mathbf{r})$, the potential is determined uniquely by the ground state particle density, $n_0(\mathbf{r})$ → ***All the properties of the system are completely determined given only the ground state density***

II. A universal functional for the energy E[n] in terms of the density $n(\mathbf{r})$ can be defined, valid for any $V_{ext}(\mathbf{r})$. The density that minimizes the functional is the exact ground state density → ***The functional E[n] alone is sufficient to determine the exact ground state and energy***

But now we need a way to find the functional E[n]!

The Kohn-Sham ansatz

W. Kohn and L. J. Sham: “Self-consistent equations including exchange and correlation effects”, *Phys. Rev.* **140**, A1133 (1965)

$$\hat{H}_{\mathrm{aux}} = -\frac{1}{2}\nabla^2 + V_{\mathrm{eff}}(\boldsymbol{r}) \qquad \text{and} \qquad n(\boldsymbol{r}) = \sum_{i=1}^{N} |\Psi_i(\boldsymbol{r})|^2$$

We write the ground state energy functional as:

$$E_{\mathrm{KS}} = T_s[n] + \int d\boldsymbol{r} V_{ext}(\boldsymbol{r}) n(\boldsymbol{r}) + E_{\mathrm{Hartree}}[n] + E_{\mathrm{II}} + \boxed{E_{\mathrm{xc}}[n]}$$

T_s = independent-particle energy = $\frac{1}{2}\sum_{i=1}^{N}|\nabla\Psi_i|^2$

V_{ext} = external potential due to nuclei + other external fields

E_{II} = interaction between nuclei

$E_{Hartree}$ = classical Coulomb interaction energy of the electron density interacting with itself = $\frac{1}{2}\int d^3r d^3r' \frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}$

Contains all many-body effects of exchange and correlation. Unknown!

But it can be approx. as a local or nearly local functional of the density

DFT

The main problem: Finding good functionals

We can define a hierarchy of functionals. Each extra level adds increased sophistication and (hopefully) accuracy:

- **LDA** (Local Density Approximation):
  - Uses only n(**r**) at a point **r**
  - Assumes the functional is the same as in the homogeneous electron gas
  - Even if this seems a huge approx., LDA reproduces reasonably well chemical bonding in solids, molecules, surfaces and defects
- **GGA** (Generalized Gradient Approximation)
  - Uses n(**r**) and $|\nabla n(\mathbf{r})|$, introducing some non-locality
  - Generally more accurate, corrects overbinding of LDA
  - Can be parameter free (PBE, PW91) or contain some parameters determined from fits to expt data (BLYP)
- **Hybrid**
  - Combine some fraction of orbital-dependent HF
  - More accurate functional to compute energies and more popular in chemistry community (B3LYP, PBE0)
- Also Meta-GGA, hyper-GGA, +U, dispersive, …

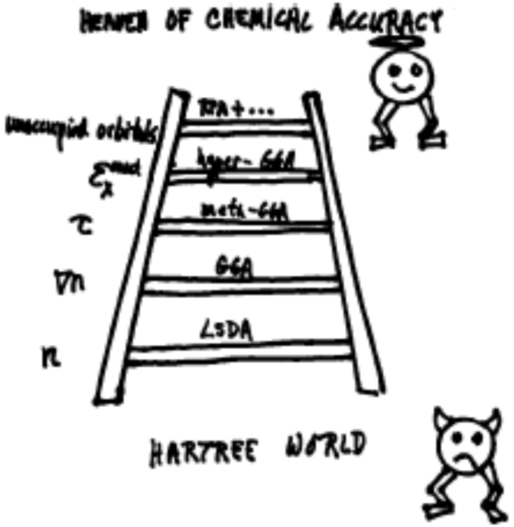
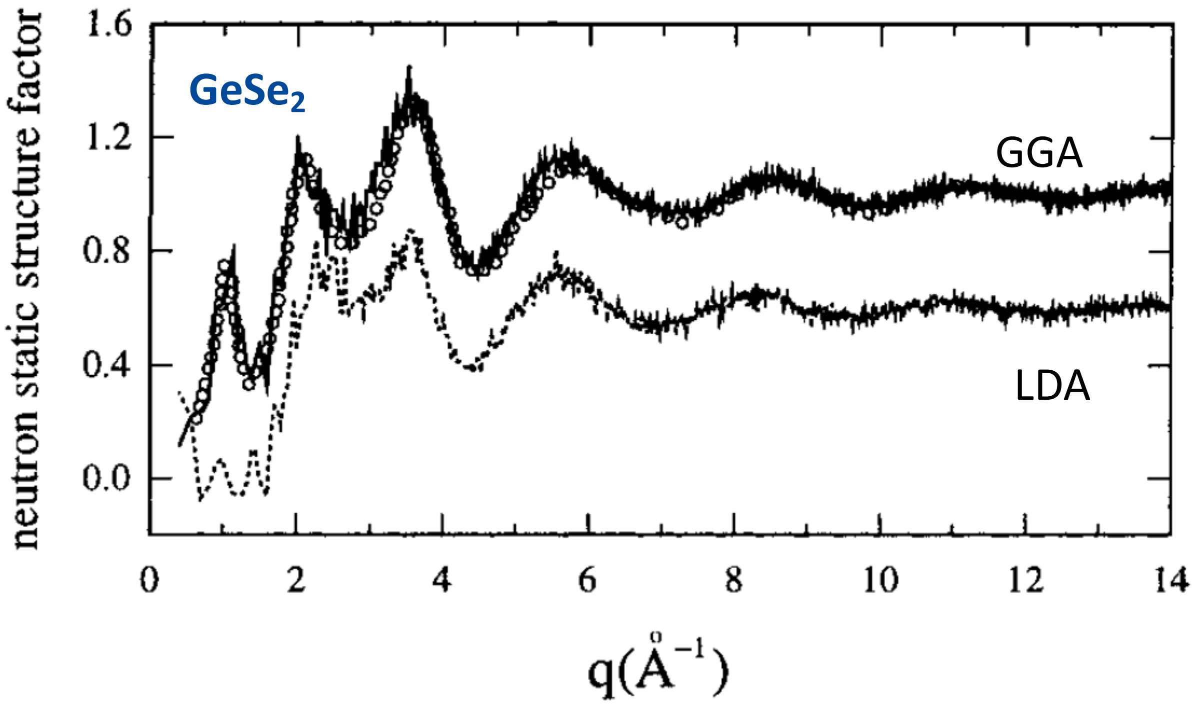


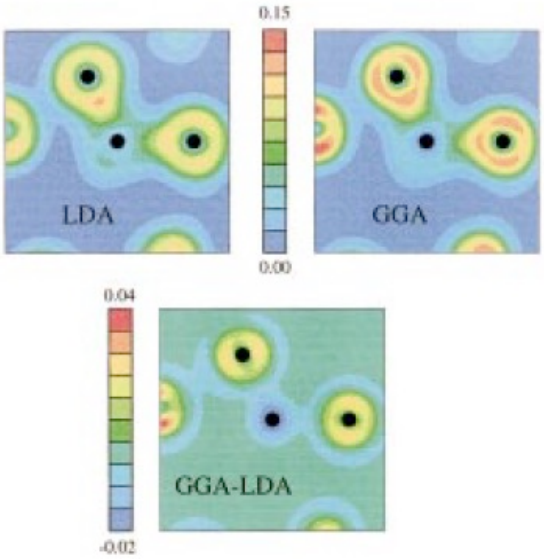
Figure 2. The Jacob's ladder of density functional approximations to the exchange-correlation energy adds local ingredients successively, leading up in five steps from the Hartree world (E_{xc} = 0) of weak or no chemical bonding to the heaven of chemical accuracy (with errors in energy differences of order 1 kcal/mol = 0.0434 eV).

From Perdew, *J. Chem. Theory Comput.* **5**, 902 (2009)

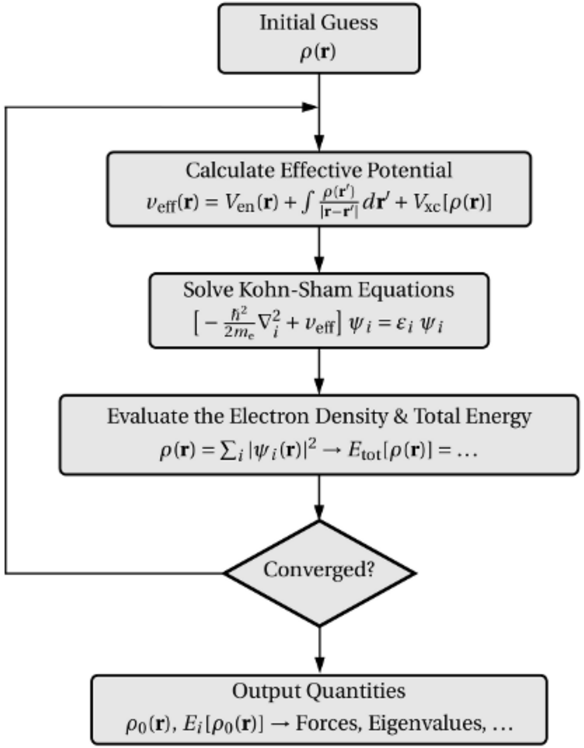
DFT The main problem: Finding good functionals

Massobrio et al., *J. Am. Chem. Soc.* **121**, 2943 (1999)





DFT How?



Initial Guess
$\rho(\mathbf{r})$

Calculate Effective Potential
$v_{\text{eff}}(\mathbf{r}) = V_{\text{en}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + V_{\text{xc}}[\rho(\mathbf{r})]$

Solve Kohn-Sham Equations
$\left[-\frac{\hbar^2}{2m_e}\nabla_i^2 + v_{\text{eff}}\right] \psi_i = \varepsilon_i \psi_i$

Evaluate the Electron Density & Total Energy
$\rho(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2 \rightarrow E_{\text{tot}}[\rho(\mathbf{r})] = \ldots$

Converged?

Output Quantities
$\rho_0(\mathbf{r}), E_i[\rho_0(\mathbf{r})] \rightarrow$ Forces, Eigenvalues, …

Planewaves codes

VASP, CASTEP, Abinit, CPMD, QE

Gaussian basis

Crystal, CP2K

Numerical basis

Dmol3

LAPW

Wien2k

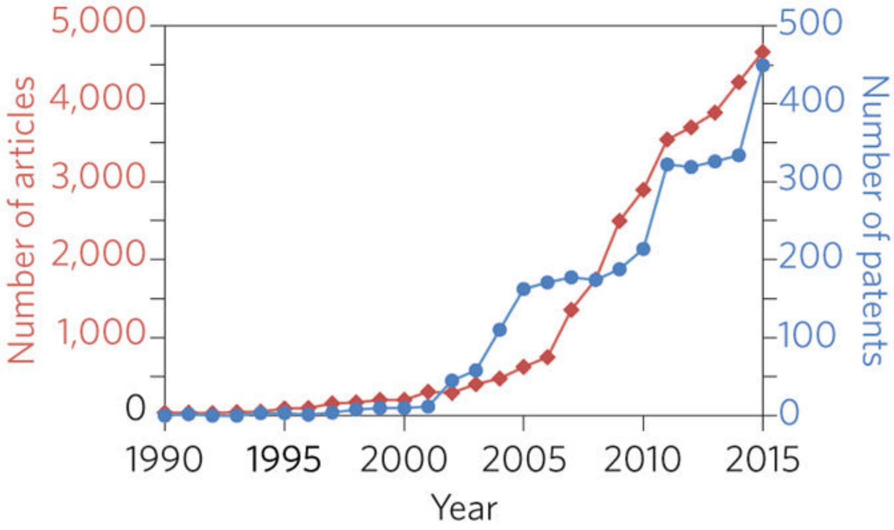
Linear scaling

Siesta, Onetep, BigDFT

And many more. See *http://www.psi-k.org/codes.shtml and https://en.wikipedia.org/wiki/List_of_quantum_chemistry_and_solid-state_physics_software*

DFT Becoming a vital tool in materials modelling

Editorial: "Boosting materials modelling", *Nature Mater.* **15**, 365 (2016)



- Today DFT is the reference method in solid state physics and materials simulation
- Typically, DFT can be used to model systems containing ~100-1000 atoms, times ~100 ps

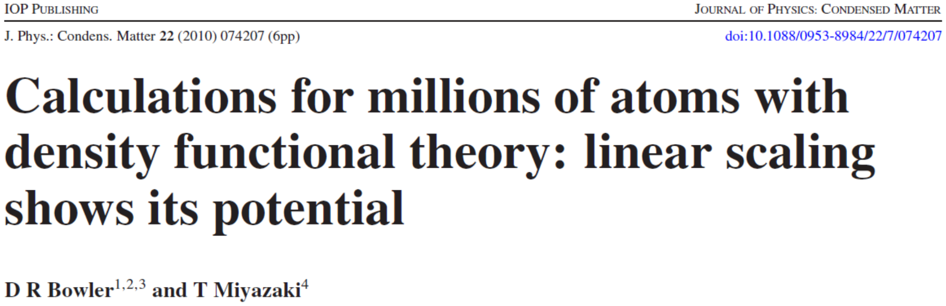
See 'big data' projects: https://materialsproject.org/ (USA),
https://nomad-coe.eu/ (EU)

DFT Problems, limitations and ongoing research

A. J. Cohen et al.: "Insights into current limitations of DFT", *Science* **321**, 792 (2008)

K. Burke: "Perspective on density functional theory", *J. Chem. Phys.* **136**, 150901 (2012)

- Incomplete treatment of dispersion forces → Non-local (**Hybrid**) functionals
- Fails for strongly correlated systems → **LDA+U, DMFT**
- Band gaps in semiconductors underestimated → Hybrid
- Computationally expensive, i.e. limited system size → **linear scaling DFT**

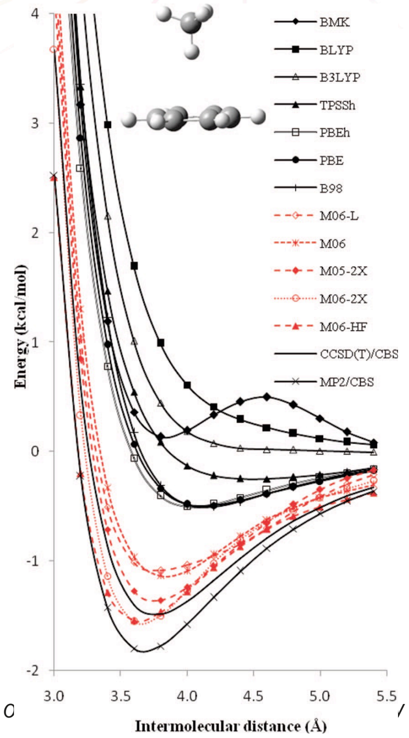


IOP PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **22** (2010) 074207 (6pp) doi:10.1088/0953-8984/22/7/074207

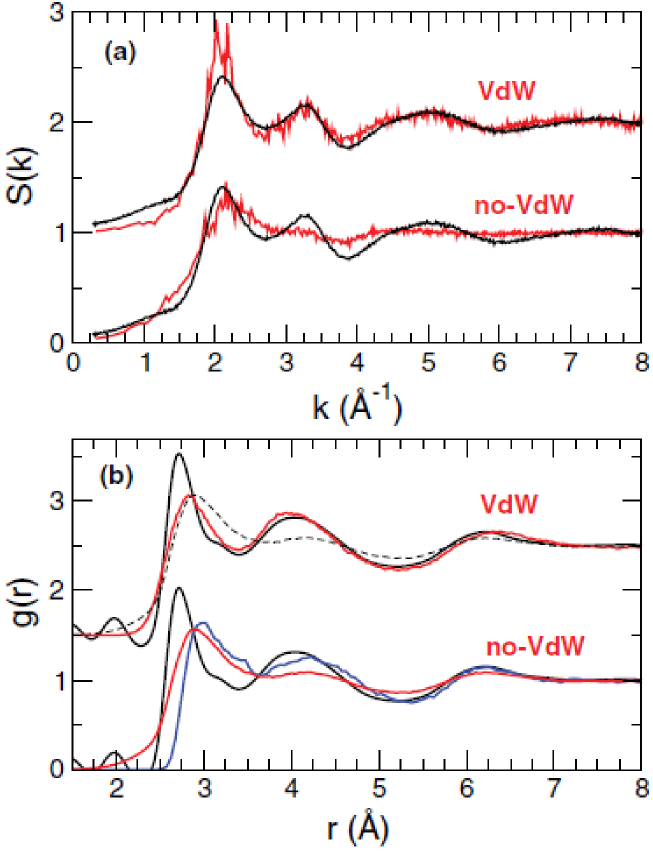
Calculations for millions of atoms with density functional theory: linear scaling shows its potential

D R Bowler[1,2,3] and T Miyazaki[4]



DFT Effect of adding dispersion

Micoulaut, *J. Chem. Phys.* **138**, 061103 (2013)



- **$Ge_{15}Te_{85}$ – 200 atoms – 733 K – 25 ps trajectories**
- GGA overestimates bond distances
- Adding an empirical dispersion (Grimme) correction:

$$E_{disp} = -s_6 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{C_{ij}}{R_{ij}^6} f_{damp}(R_{ij})$$

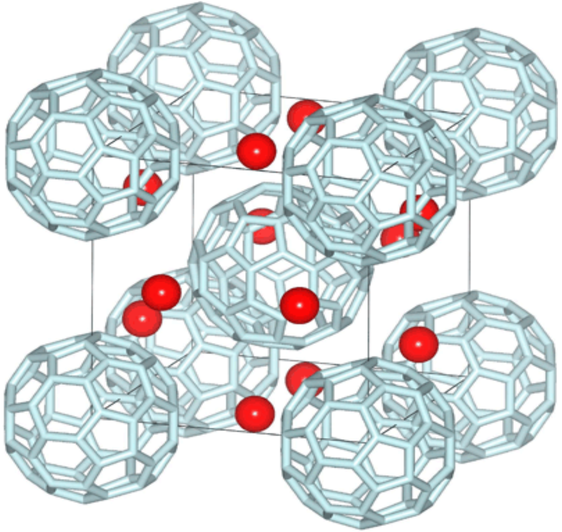
gives a better overall description of ND data

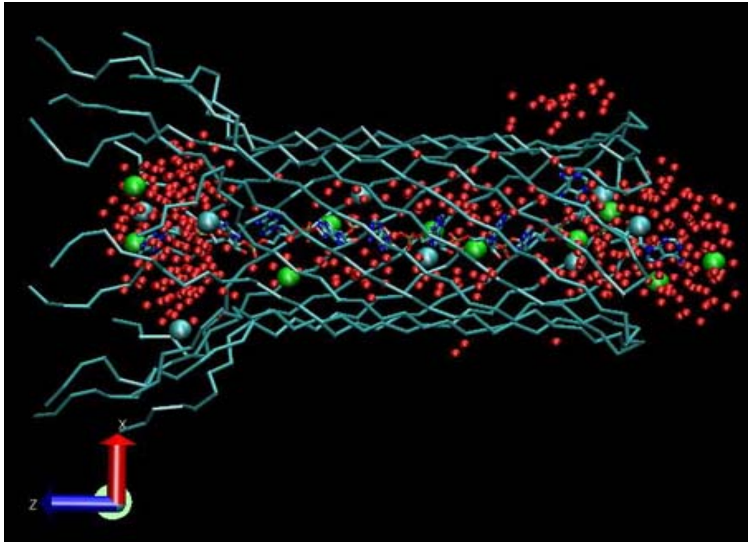
- No changes in angle distributions

Computing energies and forces
2. Empirical potentials

Molecular Mechanics approach

- No electrons
- Molecules represented as ‘balls’ + ‘springs’
- Empirical potential to account for intra- and intermolecular interactions
  →**Force Field (FF)**
- ~5-6 orders of magnitude faster than DFT → ~10^4-10^5 atoms, ~1-100 ns

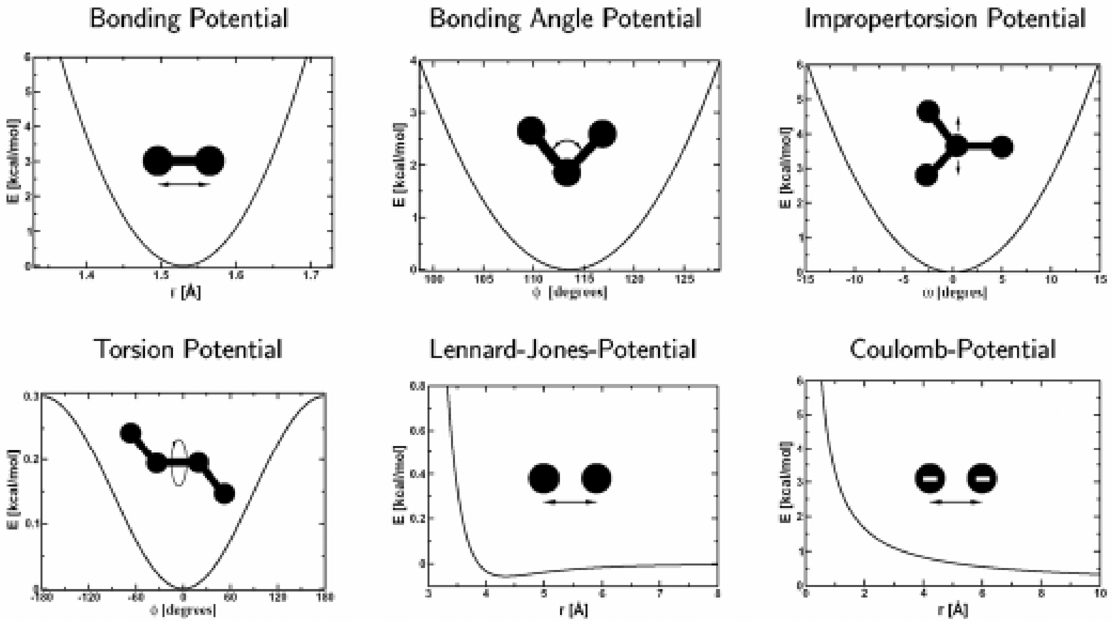




What is a Force Field?

- A mathematical expression describing the dependence of the energy of a system on the coordinates of the atoms
- Consists of:
  1. An analytical form of the interatomic potential energy, $U(r_1 \dots r_N)$
  2. A set of parameters entering $U(r_1 \dots r_N)$
- The parameters are obtained from *ab initio* calculations and/or fitting to experimental data
- Molecules are defined as collection of atoms held together by simple elastic (harmonic) forces

A standard FF



$$
\begin{aligned}
E &= \sum_{b} k_b (r - r_{\mathrm{b}})^2 + \sum_{\theta} k_\theta (\theta - \theta_{\mathrm{O}})^2 + \sum_{\omega} k_\omega (\omega - \omega_{\mathrm{O}})^2 \\
&+ \sum_{\phi} k_\phi (1 - \cos(n\phi - \delta)) + \sum_{i<j} 4\, \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^{6} \right] + \sum_{i<j} \frac{q_i q_j}{r}
\end{aligned}
$$

Some examples of FFs for inorganic or metallic glasses

$$U(r_{ij}) = \frac{q_i q_j}{r_{ij}} + A_{ij} e^{-b_{ij} r_{ij}} + C_{ij} r_{ij}^{-6}$$ (Coulomb + Buckingham)

e.g. Van Beest et al., *Phys. Rev. Lett.* **64**, 1955 (1990) for SiO_2 and $AlPO_4$

$$U(r_{ij}) = \frac{q_i q_j}{r_{ij}} + A_{ij} e^{\left(\frac{\sigma_{ij} - r_{ij}}{\rho_{ij}}\right)} - C_{ij} r_{ij}^{-6} + D_{ij} r_{ij}^{-8}$$ (Born-Mayer-Huggins)

$$U_{ijk}(r_{ij}, r_{ik}, \theta_{ijk}) = \lambda_{ijk} e^{\left(\frac{\gamma_{ij}}{r_{ij} - r_{ij}^0} + \frac{\gamma_{ik}}{r_{ik} - r_{ik}^0}\right)} \left(\cos\theta_{ijk} - \cos\theta_{ijk}^0\right)^2$$ (3-body)

e.g. Bauchy, *J. Chem. Phys* **141**, 024507 (2014) for $(SiO_2)_{0.6}(Al_2O_3)_{0.1}(CaO)_{0.3}$

$$E_{\text{tot}} = \sum_i F_i \left(\sum_{j \neq i} \rho_j^a(r_{ij})\right) + \frac{1}{2} \sum_{i \neq j} U_{ij}(r_{ij})$$ (EAM: embedded-atom method)

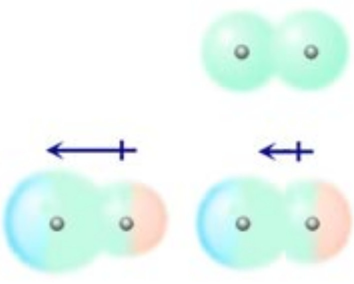
e.g. Trady et al., *J. Non-Cryst. Solids* **443**, 136 (2016) for Ni **metallic** glass

Adding the effect of polarizability

- Surrounding molecules will induce a charge redistribution
- This can be modelled using fluctuating charges, shell models (Drude particle) or induced point dipoles

$$U_{pol} = -\frac{1}{2}\sum_i \mu_i \boldsymbol{E_i^0}$$

$$\mu_i = \alpha_i \boldsymbol{E_i} = \alpha_i \left[\boldsymbol{E_i^0} + \sum_{j \neq i}^{N} \boldsymbol{T_{ij}} \mu_j \right]$$



- Iterative calculation:
  - Evaluate induced dipoles due to permanent charges
  - Adjust charges or reevaluate induced dipoles due to q+μ^{ind}
  - Repeat until convergence
  - Reiterate every time one molecule is moved

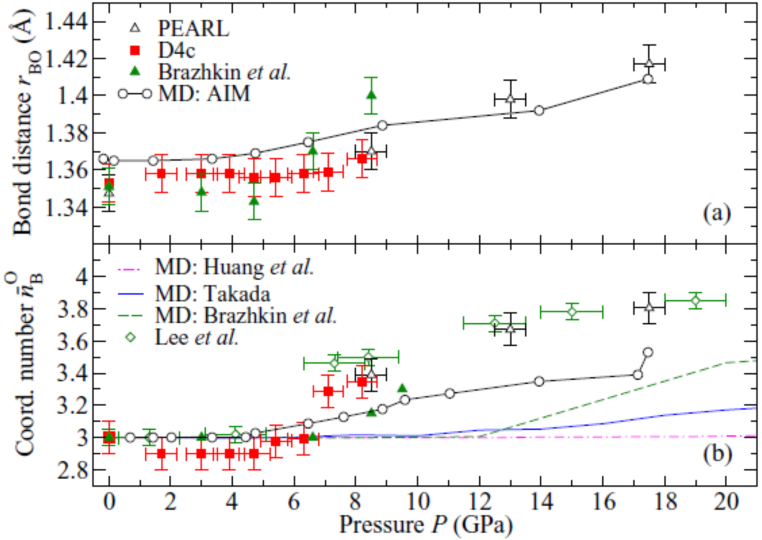
Polarizable FFs: Aspherical Ion Model

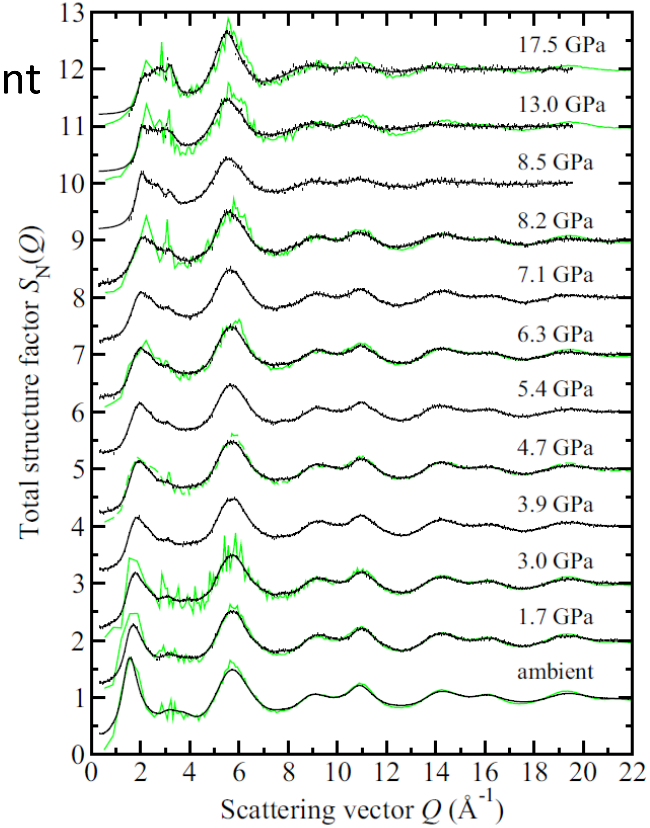
B_2O_3: Structural changes with pressure

Zeidler et al., *Phys. Rev. B* **90**, 024206 (2014)

Shape of ions allowed to change with coordination environment

$$V^{tot} = V^{rep} + V^{disp} + V^{coul} + V^{pol}$$





FF parameterization

An elaborate job

- Goal: Describe in classical terms the quantum mechanical facts by partitioning the total electronic energy into well separated atom-atom contributions
- As it is impossible to fully separate the intricate electronic effects, this means applying significant approximations → empirical potentials
- We will obtain different FFs depending on the set of data (experimental or QM calculations) and the procedure employed to optimize the parameters
- "**FF development is still as much a matter of art as of science**" (T. Halgren, *Curr. Opin. Struct. Biol.*, 1995)

FF parameterization

Components

- Functional forms of the components of the energy expression
- A list of atom types
- A list of atomic charges
- Rules for atom types
- A set of parameters for the function terms
- (Rules to generate parameters that have not been defined explicitly)

Generally developed by specialized groups and available in the literature as data tables

FF example

Amber (other general FFs: CHARMM, OPLS-AA, COMPASS, …)

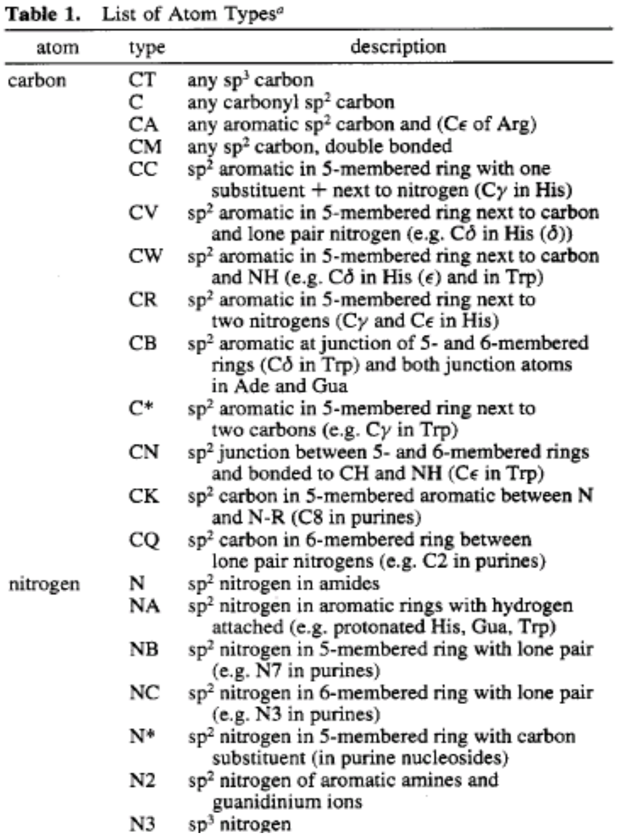
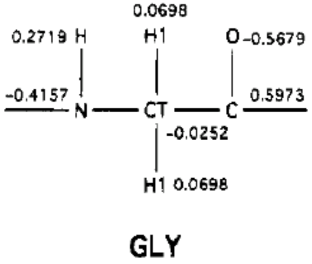
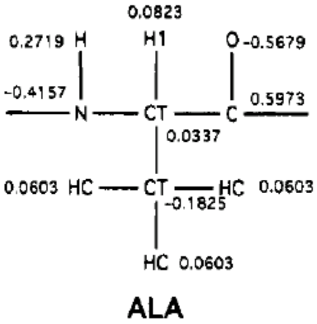


Table 1. List of Atom Types[a]

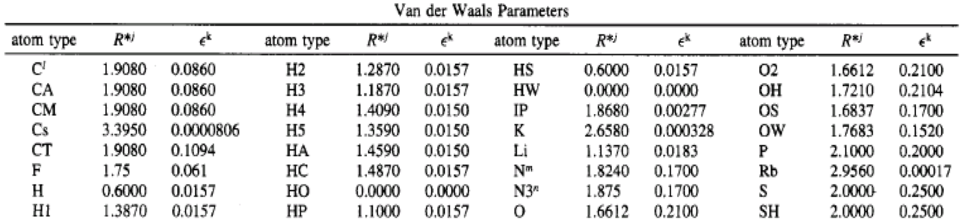
atom	type	description
carbon	CT	any sp^3 carbon
	C	any carbonyl sp^2 carbon
	CA	any aromatic sp^2 carbon and (Cε of Arg)
	CM	any sp^2 carbon, double bonded
	CC	sp^2 aromatic in 5-membered ring with one substituent + next to nitrogen (Cγ in His)
	CV	sp^2 aromatic in 5-membered ring next to carbon and lone pair nitrogen (e.g. Cδ in His (δ))
	CW	sp^2 aromatic in 5-membered ring next to carbon and NH (e.g. Cδ in His (ε) and in Trp)
	CR	sp^2 aromatic in 5-membered ring next to two nitrogens (Cγ and Cε in His)
	CB	sp^2 aromatic at junction of 5- and 6-membered rings (Cδ in Trp) and both junction atoms in Ade and Gua
	C*	sp^2 aromatic in 5-membered ring next to two carbons (e.g. Cγ in Trp)
	CN	sp^2 junction between 5- and 6-membered rings and bonded to CH and NH (Cε in Trp)
	CK	sp^2 carbon in 5-membered aromatic between N and N-R (C8 in purines)
	CQ	sp^2 carbon in 6-membered ring between lone pair nitrogens (e.g. C2 in purines)
nitrogen	N	sp^2 nitrogen in amides
	NA	sp^2 nitrogen in aromatic rings with hydrogen attached (e.g. protonated His, Gua, Trp)
	NB	sp^2 nitrogen in 5-membered ring with lone pair (e.g. N7 in purines)
	NC	sp^2 nitrogen in 6-membered ring with lone pair (e.g. N3 in purines)
	N*	sp^2 nitrogen in 5-membered ring with carbon substituent (in purine nucleosides)
	N2	sp^2 nitrogen of aromatic amines and guanidinium ions
	N3	sp^3 nitrogen



GLY



ALA



Van der Waals Parameters

atom type	R^{*j}	ϵ^k	atom type	R^{*j}	ϵ^k	atom type	R^{*j}	ϵ^k	atom type	R^{*j}	ϵ^k
C^l	1.9080	0.0860	H2	1.2870	0.0157	HS	0.6000	0.0157	O2	1.6612	0.2100
CA	1.9080	0.0860	H3	1.1870	0.0157	HW	0.0000	0.0000	OH	1.7210	0.2104
CM	1.9080	0.0860	H4	1.4090	0.0150	IP	1.8680	0.00277	OS	1.6837	0.1700
Cs	3.3950	0.0000806	H5	1.3590	0.0150	K	2.6580	0.000328	OW	1.7683	0.1520
CT	1.9080	0.1094	HA	1.4590	0.0150	Li	1.1370	0.0183	P	2.1000	0.2000
F	1.75	0.061	HC	1.4870	0.0157	N^m	1.8240	0.1700	Rb	2.9560	0.00017
H	0.6000	0.0157	HO	0.0000	0.0000	$N3^n$	1.875	0.1700	S	2.0000	0.2500
H1	1.3870	0.0157	HP	1.1000	0.0157	O	1.6612	0.2100	SH	2.0000	0.2500

FF examples

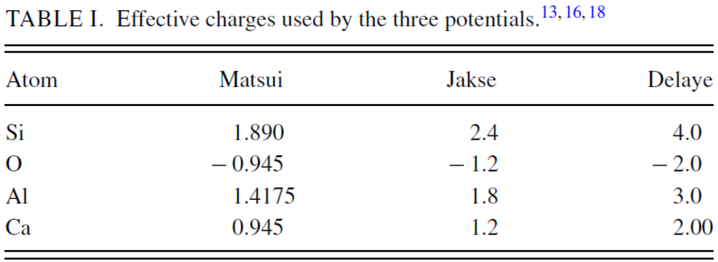


TABLE I. Effective charges used by the three potentials.[13, 16, 18]

Atom	Matsui	Jakse	Delaye
Si	1.890	2.4	4.0
O	− 0.945	− 1.2	− 2.0
Al	1.4175	1.8	3.0
Ca	0.945	1.2	2.00

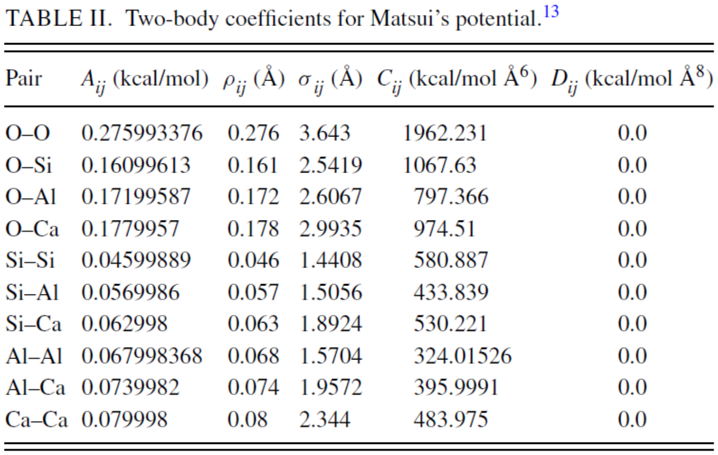


TABLE II. Two-body coefficients for Matsui's potential.[13]

Pair	A_{ij} (kcal/mol)	ρ_{ij} (Å)	σ_{ij} (Å)	C_{ij} (kcal/mol Å^6)	D_{ij} (kcal/mol Å^8)
O–O	0.275993376	0.276	3.643	1962.231	0.0
O–Si	0.16099613	0.161	2.5419	1067.63	0.0
O–Al	0.17199587	0.172	2.6067	797.366	0.0
O–Ca	0.1779957	0.178	2.9935	974.51	0.0
Si–Si	0.04599889	0.046	1.4408	580.887	0.0
Si–Al	0.0569986	0.057	1.5056	433.839	0.0
Si–Ca	0.062998	0.063	1.8924	530.221	0.0
Al–Al	0.067998368	0.068	1.5704	324.01526	0.0
Al–Ca	0.0739982	0.074	1.9572	395.9991	0.0
Ca–Ca	0.079998	0.08	2.344	483.975	0.0

Bauchy, *J. Chem. Phys* **141**, 024507 (2014)

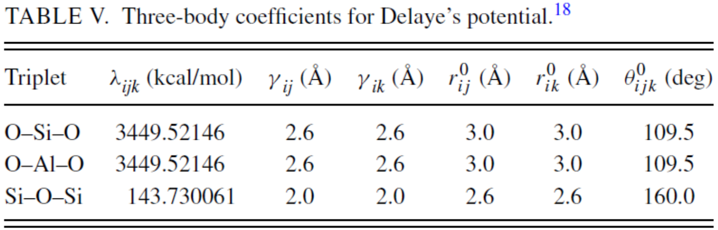


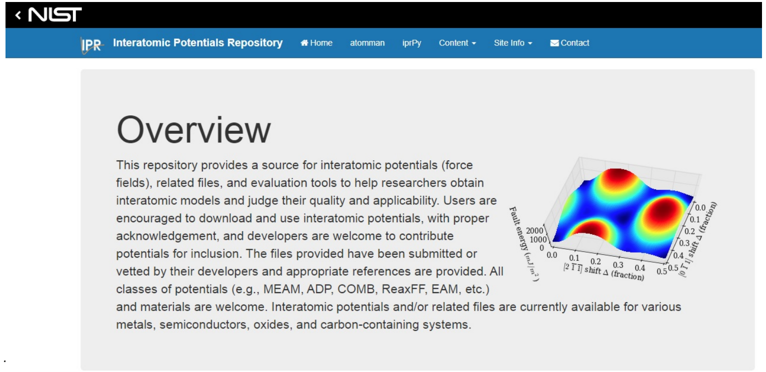
TABLE V. Three-body coefficients for Delaye's potential.[18]

Triplet	λ_{ijk} (kcal/mol)	γ_{ij} (Å)	γ_{ik} (Å)	r^0_{ij} (Å)	r^0_{ik} (Å)	θ^0_{ijk} (deg)
O–Si–O	3449.52146	2.6	2.6	3.0	3.0	109.5
O–Al–O	3449.52146	2.6	2.6	3.0	3.0	109.5
Si–O–Si	143.730061	2.0	2.0	2.6	2.6	160.0

Welcome to the potfit wiki

potfit is a free implementation of the force-matching algorithm to generate effective potentials from ab-initio reference data.

S T I



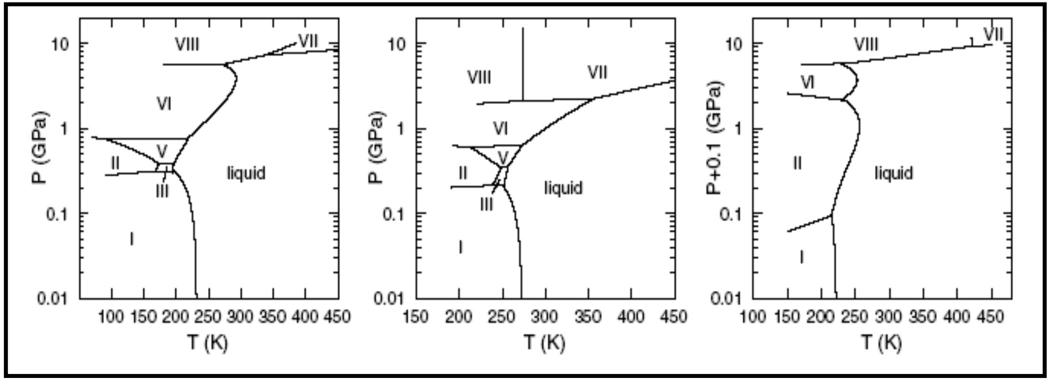
Overview

This repository provides a source for interatomic potentials (force fields), related files, and evaluation tools to help researchers obtain interatomic models and judge their quality and applicability. Users are encouraged to download and use interatomic potentials, with proper acknowledgement, and developers are welcome to contribute potentials for inclusion. The files provided have been submitted or vetted by their developers and appropriate references are provided. All classes of potentials (e.g., MEAM, ADP, COMB, ReaxFF, EAM, etc.) and materials are welcome. Interatomic potentials and/or related files are currently available for various metals, semiconductors, oxides, and carbon-containing systems.

Choosing a FF

Highly subjective, e.g. water case

Sanz, PRL (2004)



TIP4P **Exp** **SPC/E**

But ...

$D_{SPC/E} = 2.49$, $D_{TIP4P} = 3.29$, $D_{exp} = 2.4$ (10^{-5} cm^2/s)

$\varepsilon_{SPC/E} = 71$, $\varepsilon_{TIP4P} = 53$, $\varepsilon_{exp} = 78$

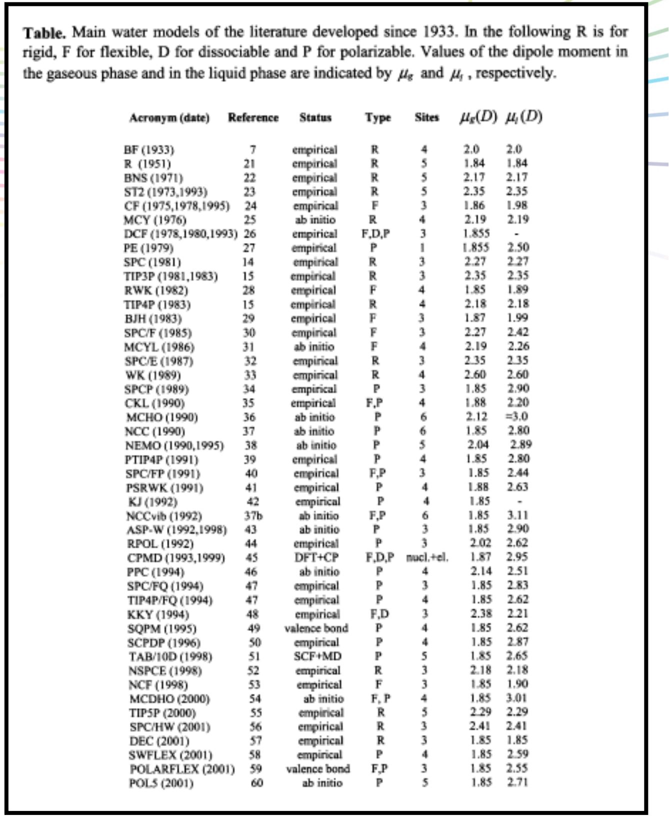


Table. Main water models of the literature developed since 1933. In the following R is for rigid, F for flexible, D for dissociable and P for polarizable. Values of the dipole moment in the gaseous phase and in the liquid phase are indicated by μ_g and μ_l, respectively.

Acronym (date)	Reference	Status	Type	Sites	$\mu_g(D)$	$\mu_l(D)$
BF (1933)	7	empirical	R	4	2.0	2.0
R (1951)	21	empirical	R	5	1.84	1.84
BNS (1971)	22	empirical	R	5	2.17	2.17
ST2 (1973,1993)	23	empirical	R	5	2.35	2.35
CF (1975,1978,1995)	24	empirical	F	3	1.86	1.98
MCY (1976)	25	ab initio	R	4	2.19	2.19
DCF (1978,1980,1993)	26	empirical	F,D,P	3	1.855	-
PE (1979)	27	empirical	P	1	1.855	2.50
SPC (1981)	14	empirical	R	3	2.27	2.27
TIP3P (1981,1983)	15	empirical	R	3	2.35	2.35
RWK (1982)	28	empirical	F	4	1.85	1.89
TIP4P (1983)	15	empirical	R	4	2.18	2.18
BJH (1983)	29	empirical	F	3	1.87	1.99
SPC/F (1985)	30	empirical	F	3	2.27	2.42
MCYL (1986)	31	ab initio	F	4	2.19	2.26
SPC/E (1987)	32	empirical	R	3	2.35	2.35
WK (1989)	33	empirical	R	4	2.60	2.60
SPCP (1989)	34	empirical	P	3	1.85	2.90
CKL (1990)	35	empirical	F,P	4	1.88	2.20
MCHO (1990)	36	ab initio	P	6	2.12	≈3.0
NCC (1990)	37	ab initio	P	6	1.85	2.80
NEMO (1990,1995)	38	ab initio	P	5	2.04	2.89
PTIP4P (1991)	39	empirical	P	4	1.85	2.80
SPC/FP (1991)	40	empirical	F,P	3	1.85	2.44
PSRWK (1991)	41	empirical	P	4	1.88	2.63
KJ (1992)	42	empirical	P	4	1.85	-
NCCvib (1992)	37b	ab initio	F,P	6	1.85	3.11
ASP-W (1992,1998)	43	ab initio	P	3	1.85	2.90
RPOL (1992)	44	empirical	P	3	2.02	2.62
CPMD (1993,1999)	45	DFT+CP	F,D,P	nucl.+el.	1.87	2.95
PPC (1994)	46	ab initio	P	4	2.14	2.51
SPC/FQ (1994)	47	empirical	P	3	1.85	2.83
TIP4P/FQ (1994)	47	empirical	P	4	1.85	2.62
KKY (1994)	48	empirical	F,D	3	2.38	2.21
SQPM (1995)	49	valence bond	P	4	1.85	2.62
SCPDP (1996)	50	empirical	P	4	1.85	2.87
TAB/10D (1998)	51	SCF+MD	P	5	1.85	2.65
NSPCE (1998)	52	empirical	R	3	2.18	2.18
NCF (1998)	53	empirical	F	3	1.85	1.90
MCDHO (2000)	54	ab initio	F, P	4	1.85	3.01
TIP5P (2000)	55	empirical	R	5	2.29	2.29
SPC/HW (2001)	56	empirical	R	3	2.41	2.41
DEC (2001)	57	empirical	R	3	1.85	1.85
SWFLEX (2001)	58	empirical	P	4	1.85	2.59
POLARFLEX (2001)	59	valence bond	F,P	3	1.85	2.55
POL5 (2001)	60	ab initio	P	5	1.85	2.71

Guillot, J. Mol. Liq. (2002)

Choosing a FF

Highly subjective, e.g. water case

- Compute 17 different properties and note each of them
- TIP4P/2005 is probably close to the best that can be achieved with a rigid non-polarizable model

Vega & Abascal, *Phys. Chem. Chem. Phys.* **13**, 19663-19688 (2011)

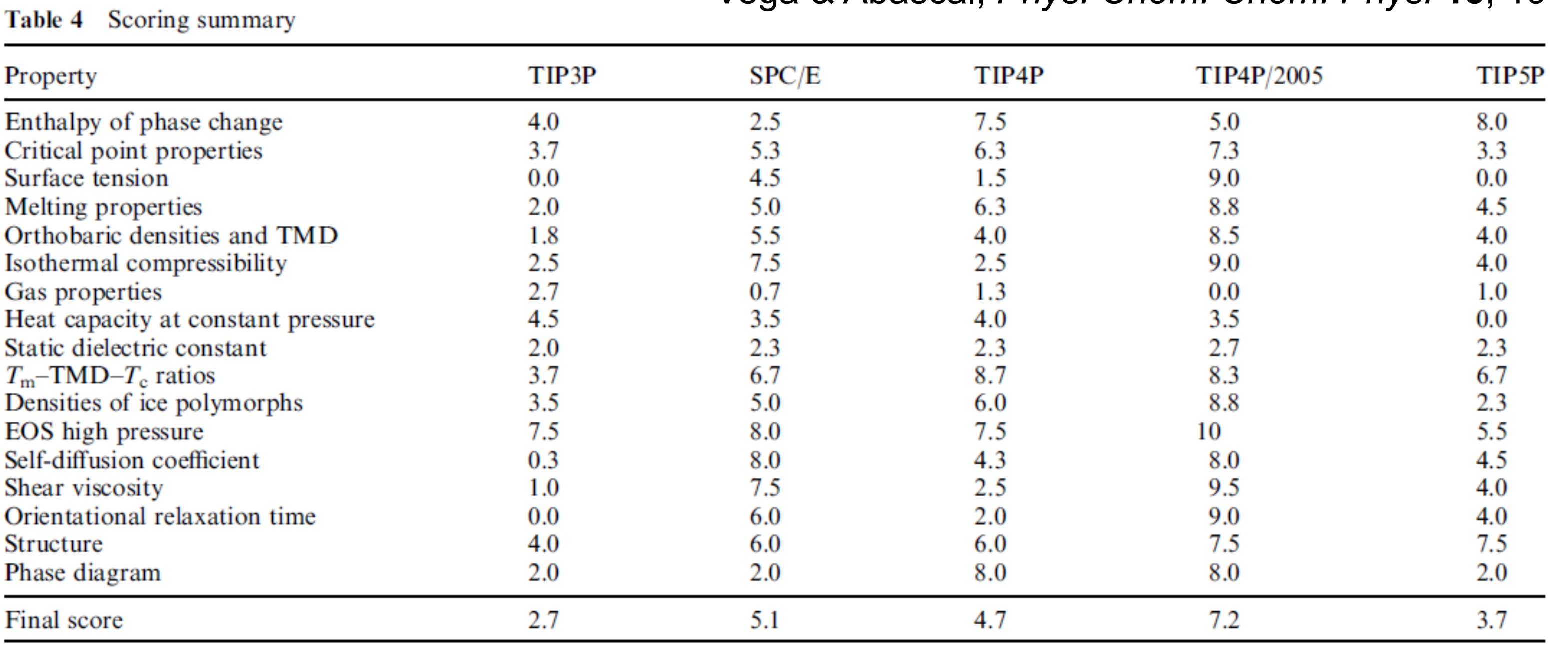


Table 4 Scoring summary

Property	TIP3P	SPC/E	TIP4P	TIP4P/2005	TIP5P
Enthalpy of phase change	4.0	2.5	7.5	5.0	8.0
Critical point properties	3.7	5.3	6.3	7.3	3.3
Surface tension	0.0	4.5	1.5	9.0	0.0
Melting properties	2.0	5.0	6.3	8.8	4.5
Orthobaric densities and TMD	1.8	5.5	4.0	8.5	4.0
Isothermal compressibility	2.5	7.5	2.5	9.0	4.0
Gas properties	2.7	0.7	1.3	0.0	1.0
Heat capacity at constant pressure	4.5	3.5	4.0	3.5	0.0
Static dielectric constant	2.0	2.3	2.3	2.7	2.3
T_m–TMD–T_c ratios	3.7	6.7	8.7	8.3	6.7
Densities of ice polymorphs	3.5	5.0	6.0	8.8	2.3
EOS high pressure	7.5	8.0	7.5	10	5.5
Self-diffusion coefficient	0.3	8.0	4.3	8.0	4.5
Shear viscosity	1.0	7.5	2.5	9.5	4.0
Orientational relaxation time	0.0	6.0	2.0	9.0	4.0
Structure	4.0	6.0	6.0	7.5	7.5
Phase diagram	2.0	2.0	8.0	8.0	2.0
Final score	2.7	5.1	4.7	7.2	3.7

FF vs DFT

Disadvantages

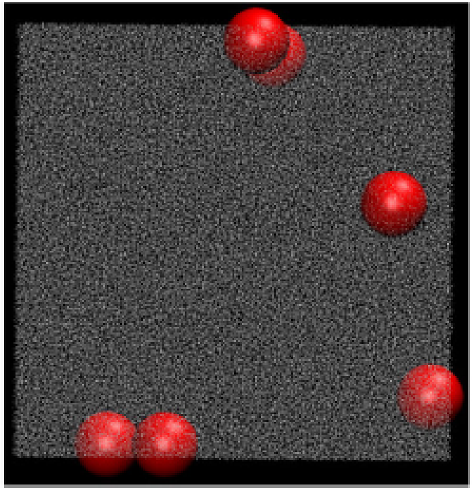
- **Intrinsic limitations:**
  - No electronic (or magnetic) structure information is
  - Unable to handle reactions (bond breaking/fo transfer, ...)
  - Bad in reproducing vibrational spectra (I

- **Limited prediction power:**
  - Accuracy depends on the tion
  - Limited to systems nal groups included in parameterization
  - Limited transf e FF (care needed when applying a FF under conditions different from the conditions used in its para e.g. P, T)

FFs need experimental validation!

Empirical force fields Advantages

- Analysis of energy contributions can be done at the level of individual interactions or classes of interactions
- Possible to modify the energy expression to bias the calculation
- Allow to handle large systems and simulate relatively long times (several orders of magnitude faster than *ab initio*):
  - Small cluster (16 processors) ~ >10^4 atoms @ 1 ns/day
  - BlueGene/L (131072 processors) ~ 320 billion atoms (a cubic piece of metal of side ~1 μm) @ 10 ps/day (Kadau, *Int. J. Modern Physics C* (2006))
  - Anton (specialized machine for MD simulations) ~ 10^4 atoms @ 10 μs/day (Klepeis, *Curr. Opin. Str. Biol.* (2009))
  - NCS (256 nodes) ~ 10^6 atoms (all-atom satellite tobacco mosaic virus, NAMD) @ 1ns/day (Freddolino, *Structure* (2006))
  - Perspectives: multimillion-atom (~100 nm scale) @ 30 ns/day on a Cray XT5 (10^5 cores) (Schulz, *J. Chem. Theory Comput.* (2009))



Adkins & Cormack, JNCS **357**, 2538 (2011)

$(Na_2O)_x(SiO_2)_{1-x}$ simulation: 10^6 atoms (116053 CPU-hours)

Sampling phase space

1. Monte Carlo

Some statistical mechanics

- **Ensemble average:** Thermodynamic properties can be calculated as averages over all the possible microstates of the system (the ensemble of possible states) as:

$$\bar{A} = \sum_i p_i A_i \quad \left(\text{e.g. } p_i = \frac{e^{-E_i/kT}}{\sum_i e^{-E_i/kT}} \right) \quad \text{or} \quad \bar{A} = \int \dots \int p(\boldsymbol{r}_N, \boldsymbol{p}_N) A(\boldsymbol{r}_N, \boldsymbol{p}_N) d\boldsymbol{r}_N d\boldsymbol{p}_N$$

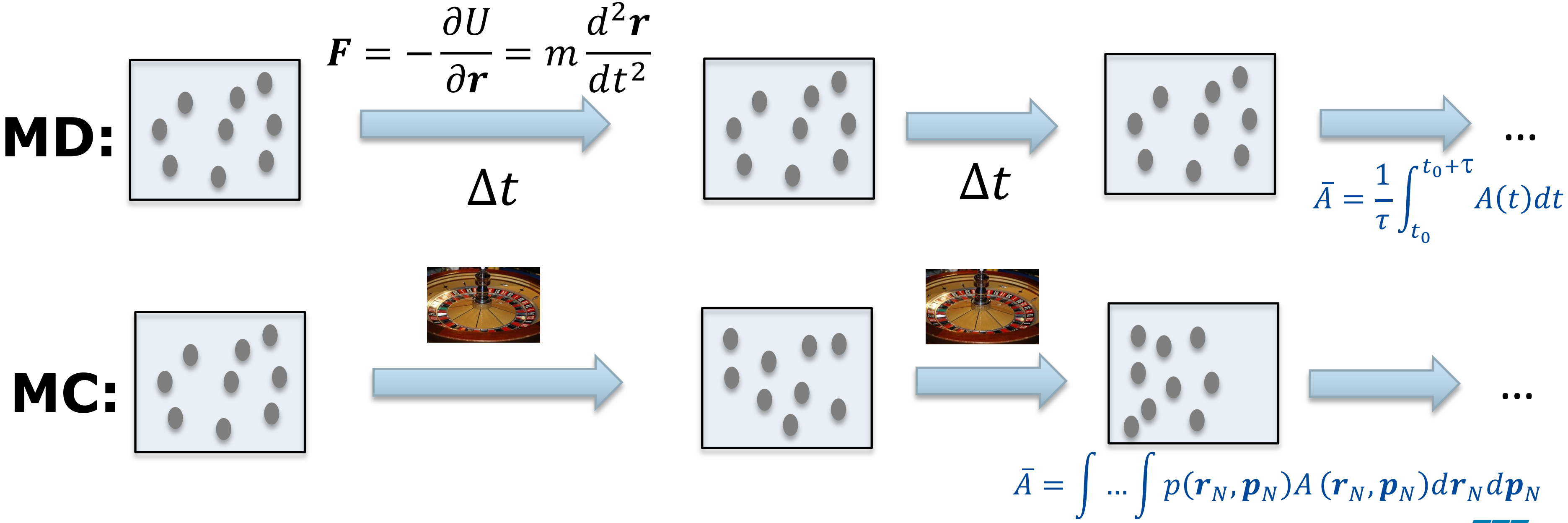
- **Ensembles:** Microcanonical (isolated system: constant NVE), canonical (NVT), isothermal-isobaric (NPT), grand canonical (μVT)

- **Ergodic principle:** The time average of a quantity is equal to its statistical average:

$$\bar{A} = \int \dots \int p(\boldsymbol{r}_N, \boldsymbol{p}_N) A(\boldsymbol{r}_N, \boldsymbol{p}_N) d\boldsymbol{r}_N d\boldsymbol{p}_N = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} A(t) dt$$

Some statistical mechanics

- Impossible to explore all possible states of the system
- But it is enough to generate a set of representative states



MD:

$$\boldsymbol{F} = -\frac{\partial U}{\partial \boldsymbol{r}} = m\frac{d^2\boldsymbol{r}}{dt^2}$$

Δt

Δt

…

$$\bar{A} = \frac{1}{\tau}\int_{t_0}^{t_0+\tau} A(t)dt$$

MC:

…

$$\bar{A} = \int \dots \int p(\boldsymbol{r}_N, \boldsymbol{p}_N) A\,(\boldsymbol{r}_N, \boldsymbol{p}_N) d\boldsymbol{r}_N d\boldsymbol{p}_N$$

Monte Carlo The basics

- **Simplest approach:**
  - Generate randomly N configurations
  - Calculate their probabilities: $p_i = \frac{e^{-E_i/kT}}{\sum_i^N e^{-E_i/kT}}$
  - Calculate the ensemble average: $\bar{A} = \sum_i^N p_i A_i$
  - Problem: For many states, $p_i \approx 0$ → very large N to get a good estimate for Ā
- **Importance sampling:**
  - Generate N configurations that already follow the Boltzmann distribution
  - The ensemble average is then calculated directly: $\bar{A} = \frac{1}{N} \sum_i^N A_i$

Monte Carlo Markov chain

N. Metropolis et al.: “Equations of state calculations by fast computing machines”, *J. Chem. Phys.* **21**, 1087 (1953)

W. K. Hastings: “Monte Carlo sampling methods using Markov chains and their applications”, *Biometrika* **57**, 97 (1970)

Method to generate configurations that follow a given probability

- Given a configuration *i*
- Create randomly a new configuration *j*
- Compute the probability W(*i*→*j*) of going from state *i* to state *j*, according to the desired probability distribution
- W(*i*→*j*) depends only on *i* and *j*: **No memory!**
- Random choice according to W(*i*→*j*) ⇒ Accept new configuration or remain on state *i*.

Monte Carlo Implementing Metropolis algorithm

- Given a configuration *i*, create a new configuration *j*
- New configuration can be created in any way, e.g. fully randomly. However in most cases this will result in unphysical configurations with ${}^{P_j}/_{P_i} \ll 1$
- Normally, configuration *j* is built from *i* by just adding small modifications (trial moves)
- Typical moves:
  - Translation of atoms or molecules
  - Rotation of atoms, molecules, or molecular groups
  - But also possible to try unphysical moves: swap atoms, cut and recombine parts of molecular chains (e.g. in polymers), etc.
- Calculate energy of new configuration, E_j and ${}^{P_j}/_{P_i} = \exp\left(-\frac{E_j - E_i}{k_B T}\right)$
- Generate a random number R from uniform distribution (0,1):

$$\text{If } R < \frac{P_j}{P_i} \text{ accept new configuration, otherwise keep state } i$$

Monte Carlo Working on different ensembles

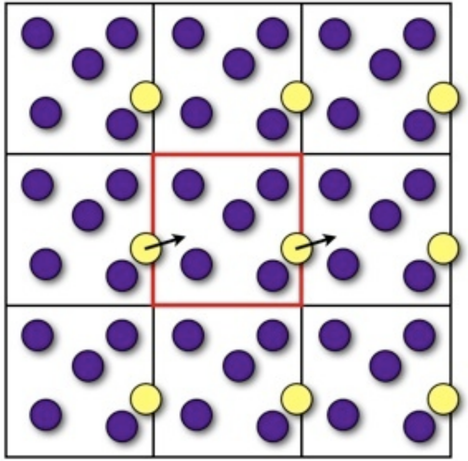
- Standard MC samples canonical ensemble (NVT): $P_{acc}(i \rightarrow j) = min\left(1, \exp\left(-\frac{E_j - E_i}{k_B T}\right)\right)$
- But using new moves and probability dist., we can adapt MC to other ensembles
- **Isothermal-isobaric (NPT) ensemble:**
  - New move: volume change $V_i \rightarrow V_i + (2R - 1) \cdot \delta V$
  - New acceptance probability: $P_{acc}(i \rightarrow j) = \min\left(1, \left(\frac{V_j}{V_i}\right)^N \exp\left(-\frac{E_j - E_i}{k_B T} + \frac{P(V_j - V_i)}{k_B T}\right)\right)$
- **Grand-canonical (μVT) ensemble:**
  - New moves: Insertion and deletion of particles
  - New acceptance probabilities for insertion and deletion:

$$P_{acc}(i \rightarrow j) = \min\left(1, \frac{V}{\Lambda^3 (N_i + 1)} \exp\left(-\frac{E_j - E_i}{k_B T} + \frac{\mu}{k_B T}\right)\right) \quad \text{with } \Lambda = \sqrt{\frac{h^2}{2\pi m k_B T}}$$

$$P_{acc}(i \rightarrow j) = \min\left(1, \frac{N_i \Lambda^3}{V} \exp\left(-\frac{E_j - E_i}{k_B T} - \frac{\mu}{k_B T}\right)\right)$$

Monte Carlo Practical aspects

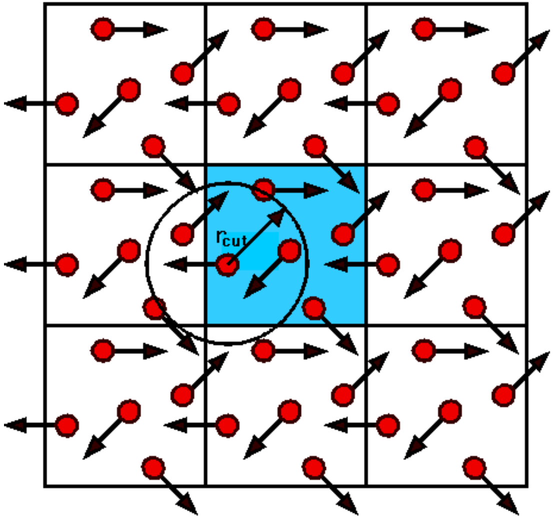
- Same as for Molecular Dynamics simulations
- Use **periodic boundary conditions (PBC)** to avoid surface effects
  - Atoms in surface ~ $N^{-1/3}$ (49% for a box containing 1000 atoms, 6% for 10^6)
  - Need PBC to get bulk properties using a limited number of atoms
  - When an atom leaves the simulation cell, it is replaced by another with the same velocity, entering from the opposite cell face (N conserved)



- Normally OK, but beware of artificial periodicity
  - Only allowed fluctuations: λ compatible with box length
  - Longest wavelength λ = L → problem if long wavelength fluctuations are important (e.g. phase transitions)
  - When simulating solids, the strain field generated by inhomogeneities will be artificially truncated and modified by the boundary
  - A macromolecule may interact with its own image if the box is too small → need to add layer of water sufficiently large
  - In some cases, spurious correlations may appear when dealing with charged or strongly polar solutes

Monte Carlo Practical aspects II

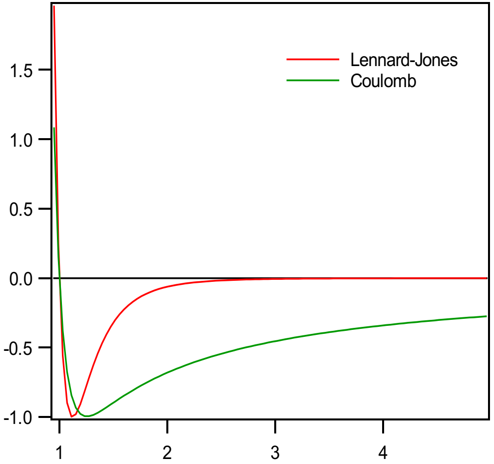
- Potential truncation of non-bonding interactions
- This is the most time-consuming part of the simulation
  - For bonding interactions → O(N)
  - For non-bonding interactions → $O(N^2)$
- Unfeasible to compute interactions with a very large number of images



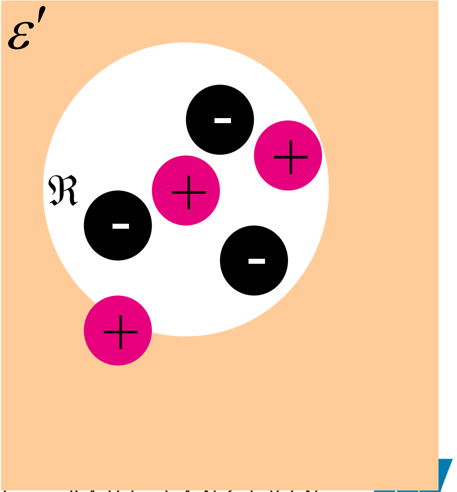
- Minimum image convention
  - Each atom only sees the closest image of the remaining N-1 atoms
  - “Truncation” surface not spherical
- Or apply an spherical cut-off
  - Only interactions with atoms at $r < R_c$ are considered
  - $R_c < L/2$. Typically
  - Introduces a discontinuity in energies and forces (for MD) at R_c. Often the potential is modified in order to bring it smoothly to 0 at R_c (switching function)

Monte Carlo Practical aspects III

- Treatment of long-range (electrostatic) interactions
- Truncation method is not appropriate, as they extend beyond the primitive cell
- **Reaction Field:**
  - Use spherical truncation and compute interactions for $r < R_c$
  - Beyond R_c, consider the medium as a dielectric continuum that will react to the dipole moment of the sphere given by $r < R_c$

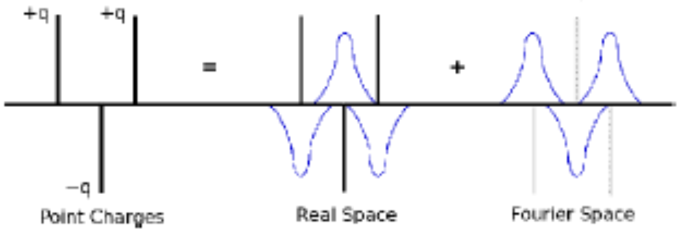


$$U_{rf} = \frac{1}{2}\sum_{i=1}^{N}\mu_i \cdot \left[\frac{2(\varepsilon'-1)}{2\varepsilon'+1}\frac{1}{r_c^3}\sum_{j\in\Re_i}\mu_j \right]$$



Monte Carlo Practical aspects IV

- Treatment of long-range (electrostatic) interactions
- Compute the full lattice sum over the infinite set of images
- **Ewald sum:**
  - Reference method, often in one of its particle-mesh variants (PME, SPME, PPPE)
  - At each point charge, add a Gaussian distribution of opposite sign to screen it
  - Screened charges can be treated in real space (short-range interactions)
  - Compensating Gaussian distribution treated in reciprocal space



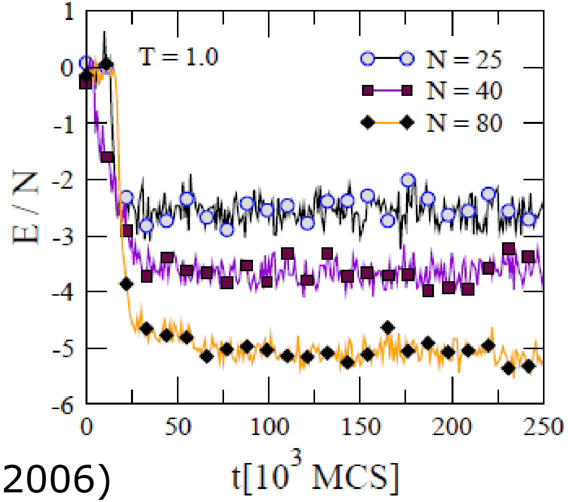
$$\rho = -q_i \left(\frac{\alpha}{\pi}\right)^{3/2} \exp\left(-\alpha r^2\right) \qquad U_{sr} = \frac{1}{2}\sum_{i \neq j}^{N} q_i q_j \operatorname{erfc}\left(\sqrt{\alpha}\, r_{ij}\right) / r_{ij}$$

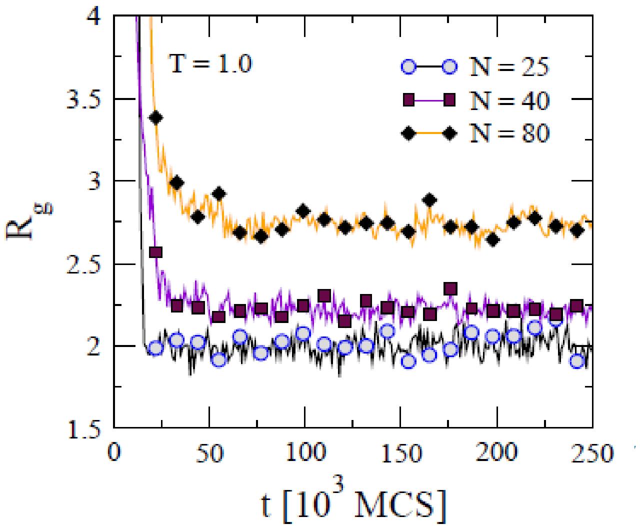
$$U_{lr} = \frac{1}{2V}\sum_{\boldsymbol{k} \neq 0} \frac{4\pi}{k^2} |\rho(\boldsymbol{k})|^2 \exp\left(-k^2/4\alpha\right) \qquad U_{\text{self}} = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{2}} \sum_{i=1}^{N} q_i^2$$

$$\text{with } \rho(\boldsymbol{k}) = \sum_{i=1}^{N} q_i \exp\left(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}_i\right)$$

Monte Carlo Practical aspects V

- Usually the initial configuration is not at equilibrium (e.g. when starting from an ordered configuration to simulate a liquid)
- Values of the properties computed during the first steps of the simulation will not be representative of the true values of the state we want to study
- Need to ensure that the system is equilibrated before starting the production phase, used to compute the desired properties

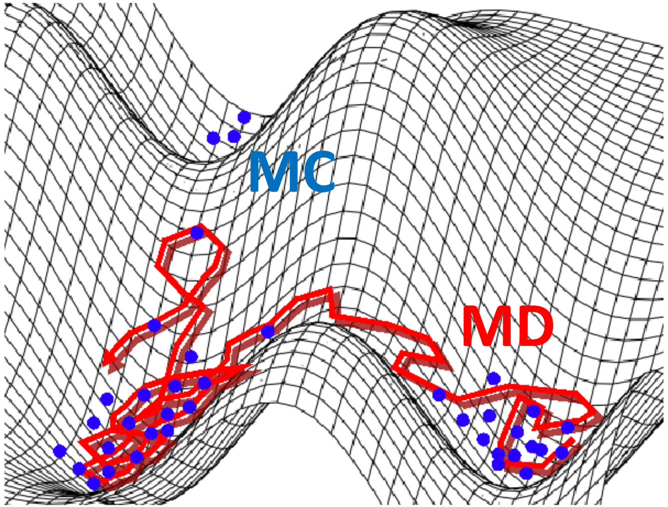




Seaton et al., Braz. J. Phys. (2006)

Monte Carlo Pros and cons

- Extremely useful to compute thermodynamic and structural properties
- Able to treat different thermodynamic ensembles in a simple way
- Allows exploring regions not accessible when using standard MD
- **But dynamic properties not available!**

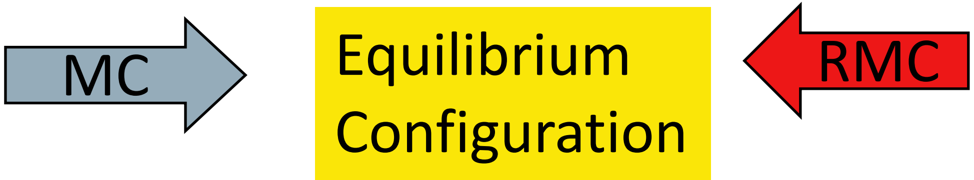


Monte Carlo Reverse Monte Carlo

R. L. McGreevy: “Reverse Monte Carlo modelling”, *J. Phys.: Condens. Matter* **13**, R877 (2001)

- We can use any “measure” that we want to accept or discard new states
- E.g. the “distance” between the simulated and measured structure factors:

$$\chi^2_{S(Q)} = \sum_i \frac{\left[S_{\exp}(Q_i) - S_{\text{sim}}(Q_i)\right]^2}{\sigma^2(Q_i)}$$



Interaction Potential

MC

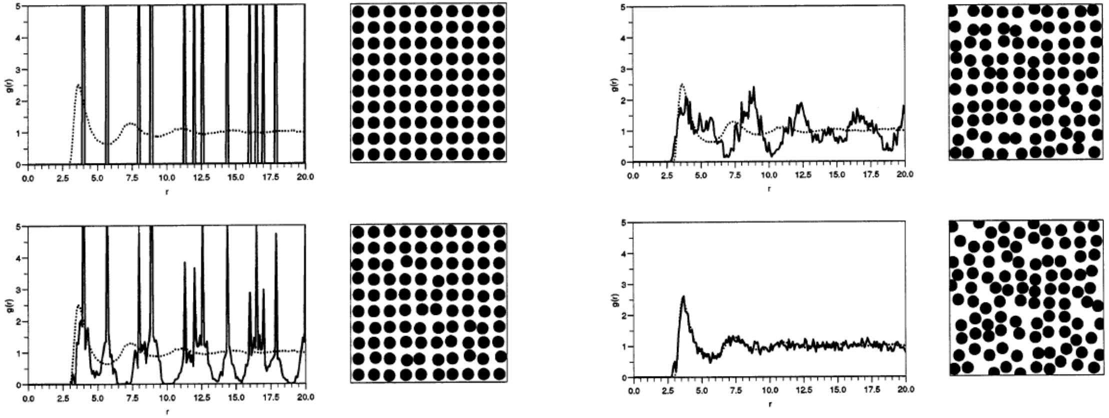
Equilibrium Configuration

RMC

Structural Data (e.g., S(Q))

compare

Experiment

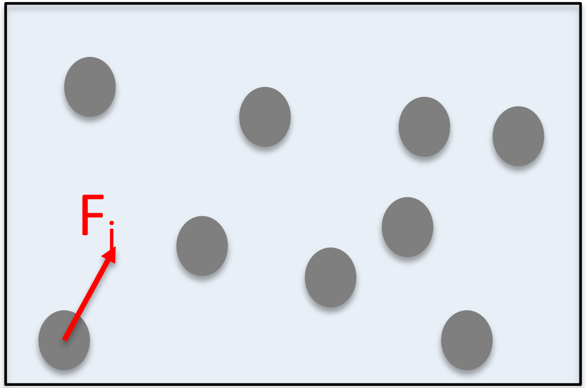


Sampling phase space

2. Molecular Dynamics

Molecular Dynamics The basics

- Generate a dynamical trajectory by integrating Newton's equations of motion, with suitable initial and boundary conditions
- We need a good way to determine the forces acting on each atom (e.g. FF or DFT)
- Also need an accurate numerical method to integrate the equations of motion



$$\boldsymbol{F}_i = -\frac{\partial U(\boldsymbol{r}^N)}{\partial \boldsymbol{r}_i} = m_i \frac{d^2 \boldsymbol{r}_i}{dt^2}$$

- System of N coupled ODEs
- No exact solution
- Step-by-step numerical integration

Molecular Dynamics The basics II

$$\boxed{\begin{aligned} \frac{d\mathbf{r}_j}{dt} &= \frac{\mathbf{p}_j}{m} \\ \frac{d\mathbf{p}_j}{dt} &= \mathbf{F}_j \end{aligned}}$$

- **Discretize solution, using a time step Δt**

$$\boldsymbol{r}_i(t_0) \rightarrow \boldsymbol{r}_i(t_0 + \Delta t) \rightarrow \boldsymbol{r}_i(t_0 + 2\Delta t) \rightarrow \cdots \rightarrow \boldsymbol{r}_i(t_0 + n\Delta t)$$

- **Important features of our numerical integrator:**
  - minimal need to compute forces (a very expensive calculation)
  - good stability for large time steps
  - good accuracy
  - conserves energy and momentum
  - time-reversible
  - symplectic: conserves volume in phase space

- **Taylor expansion?**

$$\boldsymbol{r}_i(t_0 + \Delta t) = \boldsymbol{r}_i(t_0) + \boldsymbol{v}_i(t_0)\Delta t + \frac{1}{2}\boldsymbol{a}_i(t_0)\Delta t^2 + O(\Delta t^3)$$

$$\boldsymbol{v}_i(t_0 + \Delta t) = \boldsymbol{v}_i(t_0) + \boldsymbol{a}_i(t_0)\Delta t + O(\Delta t^2)$$

Unstable and inaccurate!

Molecular Dynamics The basics III

Verlet algorithm:

$$\boldsymbol{r}_i(t_0+\Delta t) = \boldsymbol{r}_i(t_0) + \boldsymbol{v}_i(t_0)\Delta t + \frac{1}{2}\boldsymbol{a}_i(t_0)\Delta t^2 + \frac{1}{3!}\dddot{\boldsymbol{r}}_i(t_0)\Delta t^3 + O(\Delta t^4)$$

$$\boldsymbol{r}_i(t_0-\Delta t) = \boldsymbol{r}_i(t_0) - \boldsymbol{v}_i(t_0)\Delta t + \frac{1}{2}\boldsymbol{a}_i(t_0)\Delta t^2 - \frac{1}{3!}\dddot{\boldsymbol{r}}_i(t_0)\Delta t^3 + O(\Delta t^4)$$

$$\boldsymbol{r}_i(t_0+\Delta t) + \boldsymbol{r}_i(t_0-\Delta t) = 2\boldsymbol{r}_i(t_0) + \boldsymbol{a}_i(t_0)\Delta t^2 + O(\Delta t^4)$$

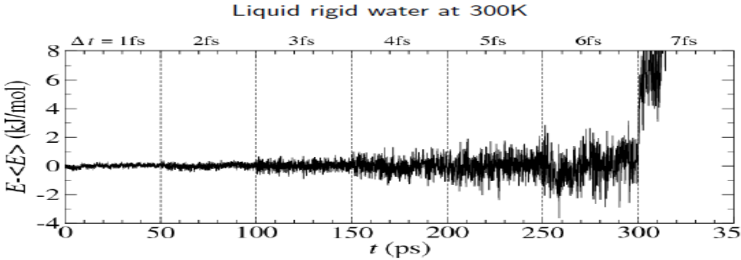
- Positions at $t_0+\Delta t$ computed from actual positions and forces, and previous positions
- Error is $O(\Delta t^4)$
- Velocities not available, but easy to derive as $\boldsymbol{r}_i(t_0+\Delta t) - \boldsymbol{r}_i(t_0-\Delta t) = 2\boldsymbol{v}_i(t_0) + O(\Delta t^3)$
- Verlet and equivalent algorithms (e.g. leap-frog) are simple, efficient, stable and reasonably accurate. Also time-reversible and symplectic, giving low drifts in energy
- Used in most MD software
- Predictor-corrector algorithms are also used occasionally. They are more accurate for small Δt, but not time reversible, introducing larger drifts in E_{total}

Molecular Dynamics The basics IV

Time step:

- Compromise between efficiency and accuracy of the integration
- Small Δt → Inefficient phase space sampling
- Large Δt → Numerical instabilities
- Hints:
  - Δt < mean time between collisions or shorter relaxation time
  - LJ atomic system ~ 10 fs
  - Flexible molecules with rigid bonds ~ 2 fs
  - Flexible bonds ~ 0.5-1 fs
  - Good test: Total energy should be conserved in NVE simulations. Acceptable $\sqrt{\langle \delta E^2 \rangle}/_E < 10^{-4}$

Frenkel & Smit: "Understanding molecular simulation", Academic Press (2002)



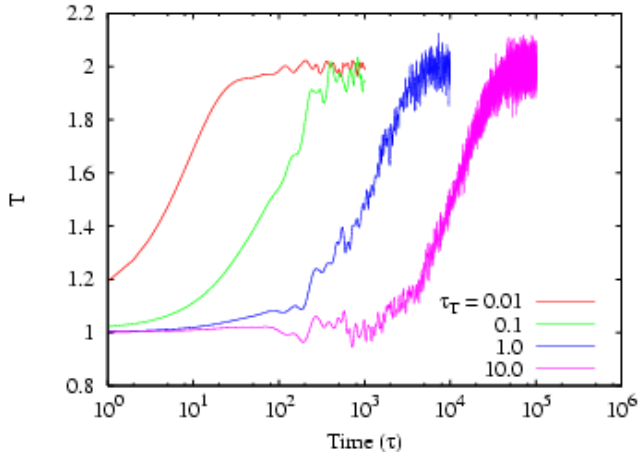
Molecular Dynamics The basics V

- **Periodic Boundary Conditions**
- **Potential truncation**
  - May need to ensure that forces (and not only energy) are 0 at R_c
- **Electrostatic forces**
- **Thermodynamic ensembles**
  - The integration of the equations of motion keeps constant N, V, and E (microcanonical ensemble)
  - Integration errors, force fluctuations and inconsistencies in the forces (e.g. generated by the cutoff) may cause slow drifts in the total energy: check E_{total} conservation!
  - E is constant in the NVE ensemble, but not K and U, so systems not in equilibrium will go to equilibrium while the temperature changes
  - Need to modify the Lagrangian or couple the system to a heat or pressure bath to keep T and/or P constant

Molecular Dynamics The basics VI

Thermodynamic ensembles

- Several thermostats allow to do **NVT** simulations: velocity scaling, Berendsen, Andersen, Nosé-Hoover, but not all of them sample strictly the correct thermodynamic ensemble.
- There are also different barostats to do **NPT** simulations.



Berendsen: Simple, fast response and flexible, but no canonical ensemble

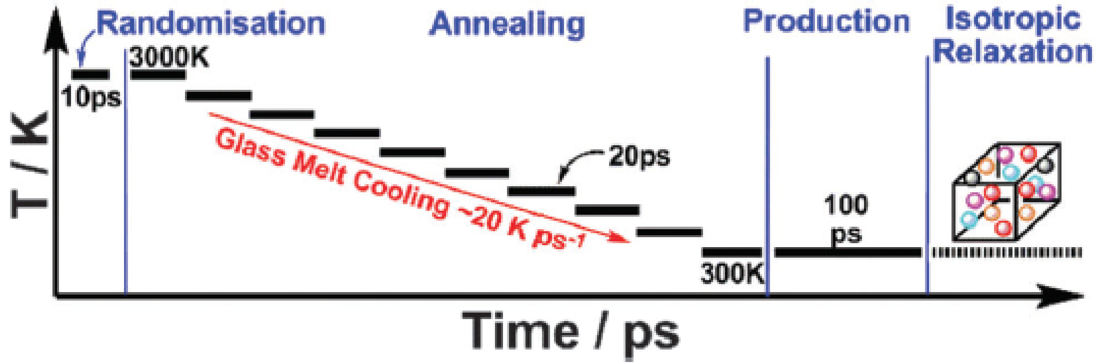
$$\boldsymbol{a} = \boldsymbol{F}/\boldsymbol{m} + \frac{1}{2\tau_T}\left(\frac{T_B}{T(t)} - 1\right)\boldsymbol{v}$$

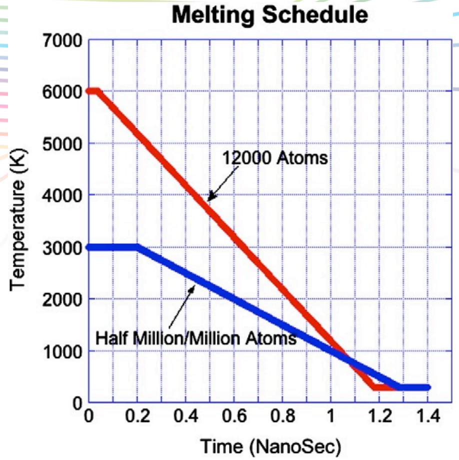
Nosé-Hoover: Extended lagrangian. Sample correct canonical ensemble, but goes to T_{ref} in an oscillatory way (poorer T control)

Molecular Dynamics Protocol

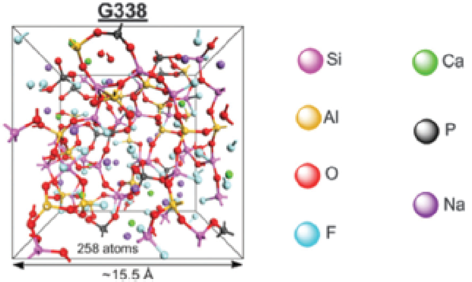
- Create initial system:
    - N atoms in random positions or crystal structure (either true or not) or from previous simulation
    - Assign random velocities corresponding to desired T (ensure that $v_{\mathrm{Box,CM}}=0$)
- Compute forces on each atom:
    - From selected potential, $\vec{F}_i = \sum_{j\neq i} -\nabla \mathrm{U}\left(\vec{r}_{ij}\right) = \sum_{j\neq i} -\frac{\partial U(r)}{\partial r_{ij}} \frac{\vec{r}_{ij}}{r_{ij}}$
    - Or from electronic structure determined by DFT
- Update positions and velocities of each atom with selected time step:
    - E.g. leapfrog: $\vec{v}_i(t + {}^{\Delta t}/_2) = \vec{v}_i(t - {}^{\Delta t}/_2) + \frac{\vec{F}_i(t)}{m_i}\Delta t,\ \vec{r}_i(t+\Delta t) = \vec{r}_i(t) + \vec{v}_i(t + {}^{\Delta t}/_2)\cdot \Delta t$
    - Note: Equivalent for BOMD using DFT, but not for Car-Parrinello MD
    - More complex expressions when using a thermostat or a barostat
- Recompute forces for new positions and propagate trajectory
- Equilibration: Initial configurations are not representative of the real system → Equilibrate your system (and check equilibration has been reached) before starting a production run and generate a trajectory ready to be analyzed.

Molecular Dynamics Protocol





NVT cooling + NPT equilibration
Adkins & Cormack, JNCS (2011)



DFT NVT simulations
Tian et al., PCCP **18**, 837 (2016)

Typical cooling rates ~ $10^{12} – 10^{10}$ K/s → Fictive T of simulated glass typically much higher than that of real glass!

E.g. $T_g^{\mathrm{sim}}\,(\mathrm{SiO_2}) \sim 2500\ \mathrm{K}\quad vs\quad T_g^{\mathrm{exp}} \sim 1450\ \mathrm{K}$

Software

- **DFT/QM**

  https://en.wikipedia.org/wiki/List_of_quantum_chemistry_and_solid-state_physics_software

- **Monte Carlo**

  https://en.wikipedia.org/wiki/List_of_software_for_Monte_Carlo_molecular_modeling

- **Molecular mechanics (MD)**

  https://en.wikipedia.org/wiki/Comparison_of_software_for_molecular_mechanics_modeling

- **Materials modelling (DFT, AIMD, MD, MC, BD, DPD, etc.)**

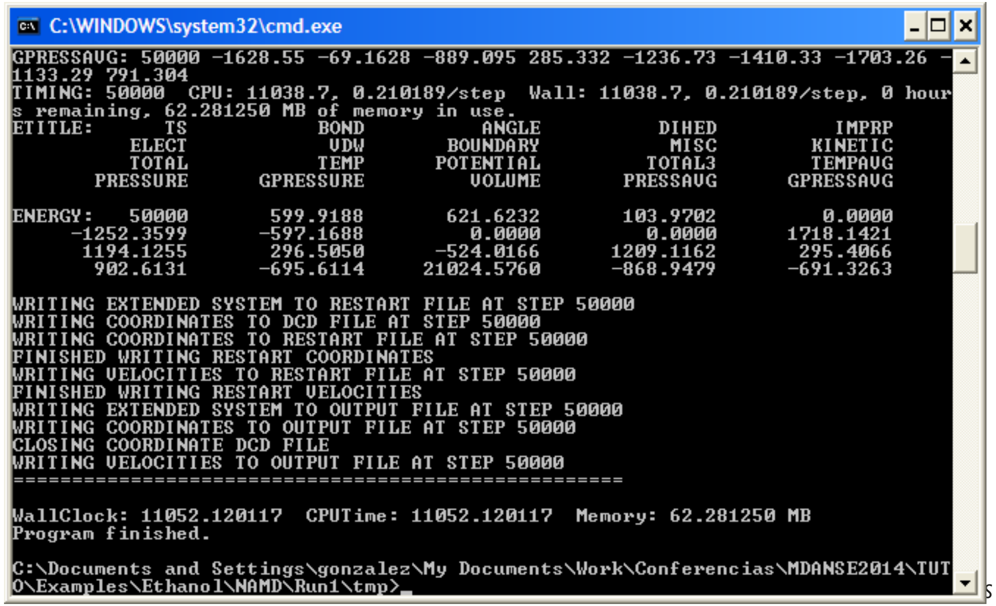
  www.sklogwiki.org/SklogWiki/index.php/Materials_modelling_and_computer_simulation_codes

At ILL: VASP, Castep, Quantum Espresso, Crystal (DFT)
DL_POLY, LAMMPS, NAMD, Gromacs (MD)
Materials Studio (commercial)

Analysis

Once our MD run is over …

- Snapshot of the final state of the system
- Trajectory file containing positions (and velocities, if requested) at selected intervals
- File containing thermodynamic information (T, P, energies, etc.)



```
GPRESSAVG: 50000 -1628.55 -69.1628 -889.095 285.332 -1236.73 -1410.33 -1703.26 -
1133.29 791.304
TIMING: 50000  CPU: 11038.7, 0.210189/step  Wall: 11038.7, 0.210189/step, 0 hour
s remaining, 62.281250 MB of memory in use.
ETITLE:      TS           BOND          ANGLE          DIHED          IMPRP
          ELECT            VDW       BOUNDARY           MISC        KINETIC
          TOTAL           TEMP      POTENTIAL         TOTAL3        TEMPAVG
       PRESSURE      GPRESSURE         VOLUME       PRESSAVG      GPRESSAVG

ENERGY:   50000       599.9188       621.6232       103.9702         0.0000
     -1252.3599      -597.1688         0.0000         0.0000      1718.1421
      1194.1255       296.5050      -524.0166      1209.1162       295.4066
       902.6131      -695.6114     21024.5760      -868.9479      -691.3263

WRITING EXTENDED SYSTEM TO RESTART FILE AT STEP 50000
WRITING COORDINATES TO DCD FILE AT STEP 50000
WRITING COORDINATES TO RESTART FILE AT STEP 50000
FINISHED WRITING RESTART COORDINATES
WRITING VELOCITIES TO RESTART FILE AT STEP 50000
FINISHED WRITING RESTART VELOCITIES
WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 50000
WRITING COORDINATES TO OUTPUT FILE AT STEP 50000
CLOSING COORDINATE DCD FILE
WRITING VELOCITIES TO OUTPUT FILE AT STEP 50000
====================================================

WallClock: 11052.120117  CPUTime: 11052.120117  Memory: 62.281250 MB
Program finished.

C:\Documents and Settings\gonzalez\My Documents\Work\Conferencias\MDANSE2014\TUT
O\Examples\Ethanol\NAMD\Run1\tmp>_
```

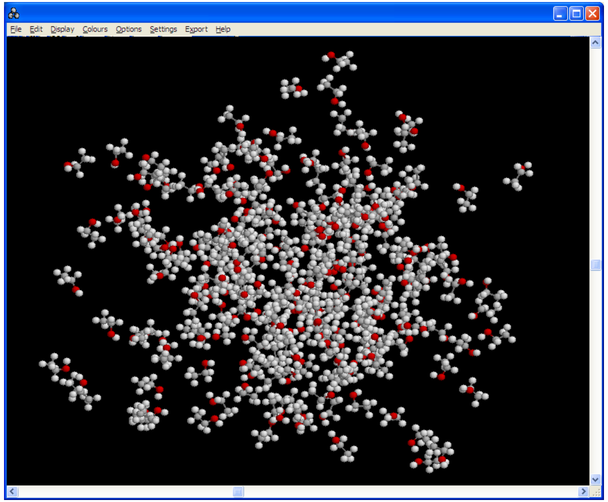
Did everything go well?

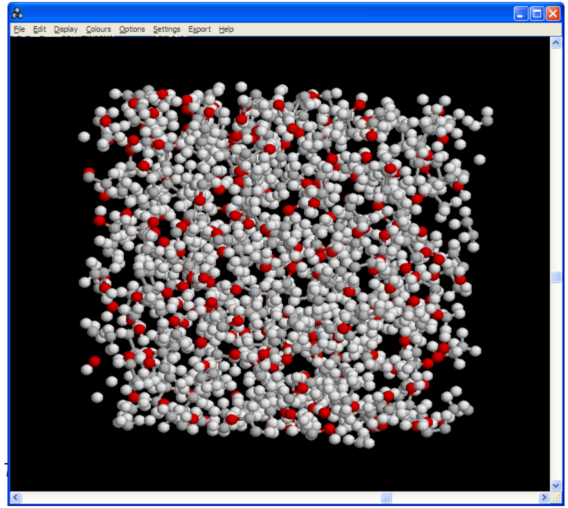
Can I get the information I'm looking for?

Once our MD run is over …

Visual checking of last configuration:

- Many visualization tools available (VMD, Rasmol, MDANSE, …)
- Don’t be puzzled by PBC’s (remember that the simulated system is really an infinite replica of images of the primary box)
- Are molecular geometries reasonable?
- Are there pairs of atoms that come ‘too close’ to each other?

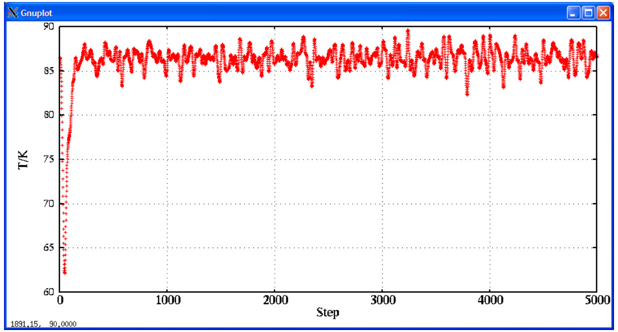


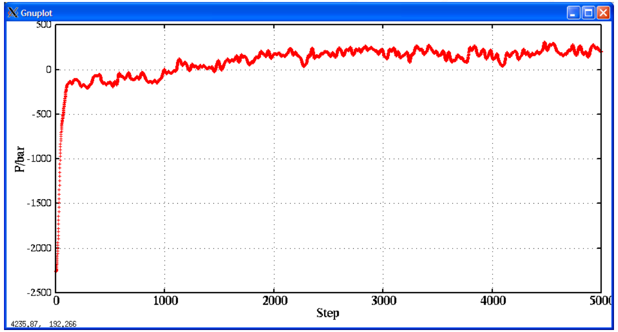


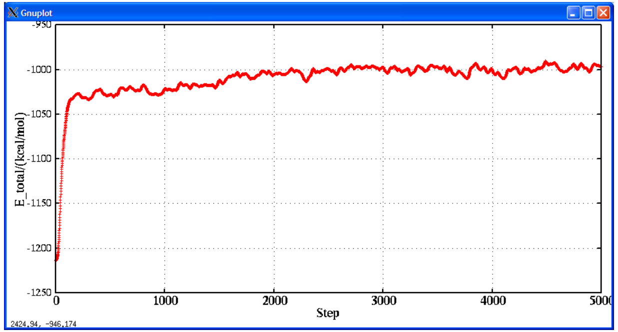
Once our MD run is over …

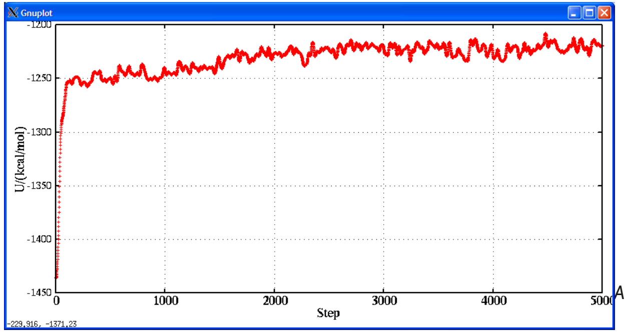
Is the system well equilibrated?

- Not all the properties converge at same speed





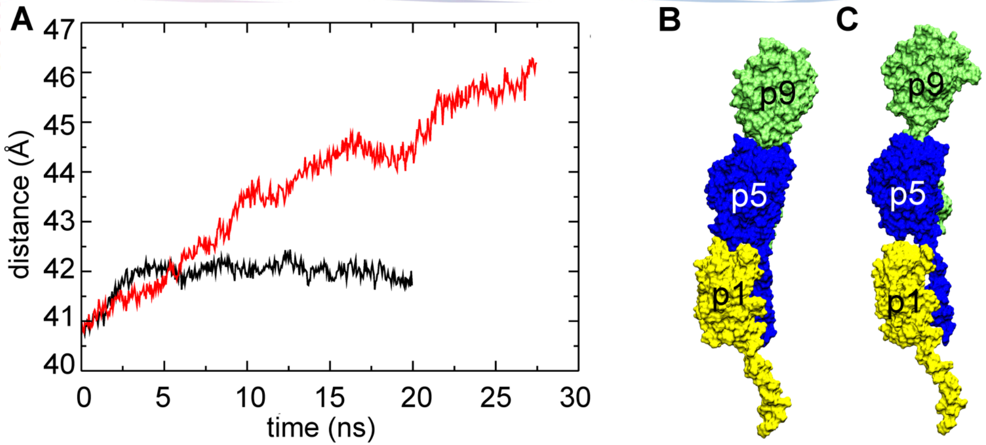




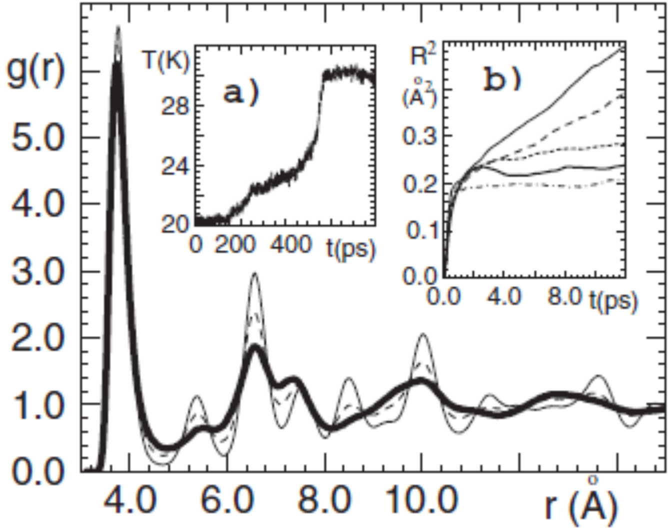
Once our MD run is over …

Does the system change during the simulation?

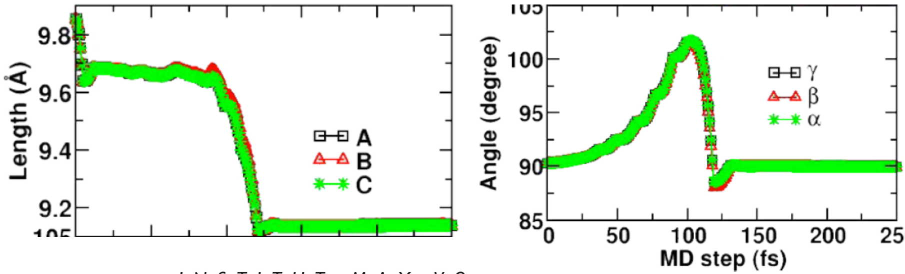
Baker, PLOS Comp. Biol. (2013)



De Santis, JPCM **14** (2002)



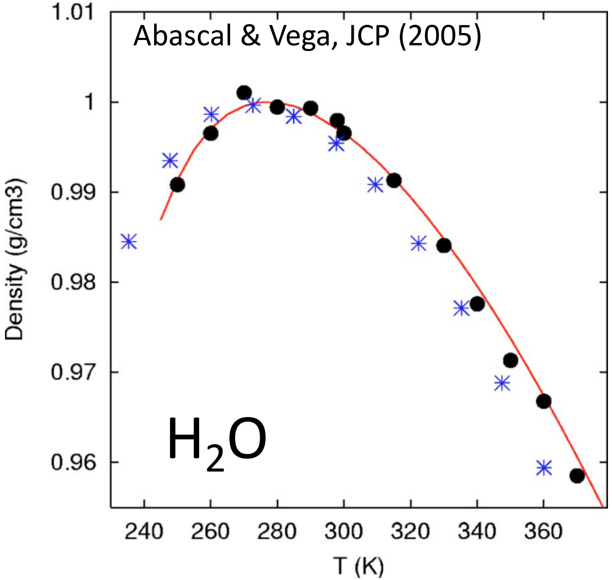
Durandurdu, JPCM **21** (2009)



Thermodynamic averages

Does our model give reasonable average values for T, P, density, …?

- T, P, V, energy contributions are usually directly given by the MD code
- Other thermodynamic properties may need to be calculated later
- If possible, compare with as much experimental data as possible
- Note that formulae based on fluctuations depend on the ensemble



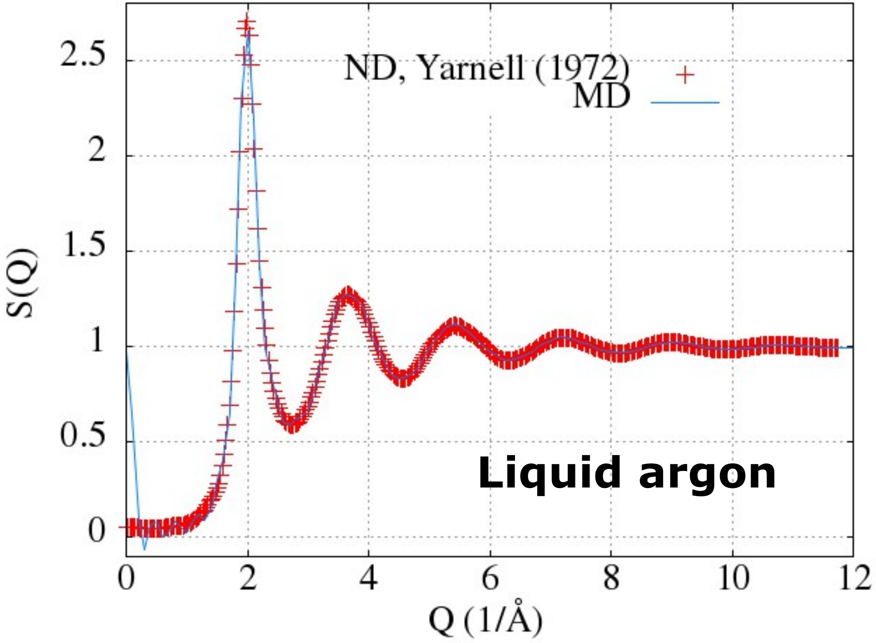
$$\Delta H_{\mathrm{vap}} = H_{\mathrm{gas}} - H_{\mathrm{liq}} = (U+PV)_{\mathrm{gas}} - (U+PV)_{\mathrm{liq}}$$

$$C_V = \frac{\left\langle \left(H - \langle H \rangle\right)^2 \right\rangle}{k_B T^2} = \frac{\left\langle \left(U - \langle U \rangle\right)^2 \right\rangle}{k_B T^2} + \frac{3}{2} N k_B T \qquad \text{(NVT)}$$

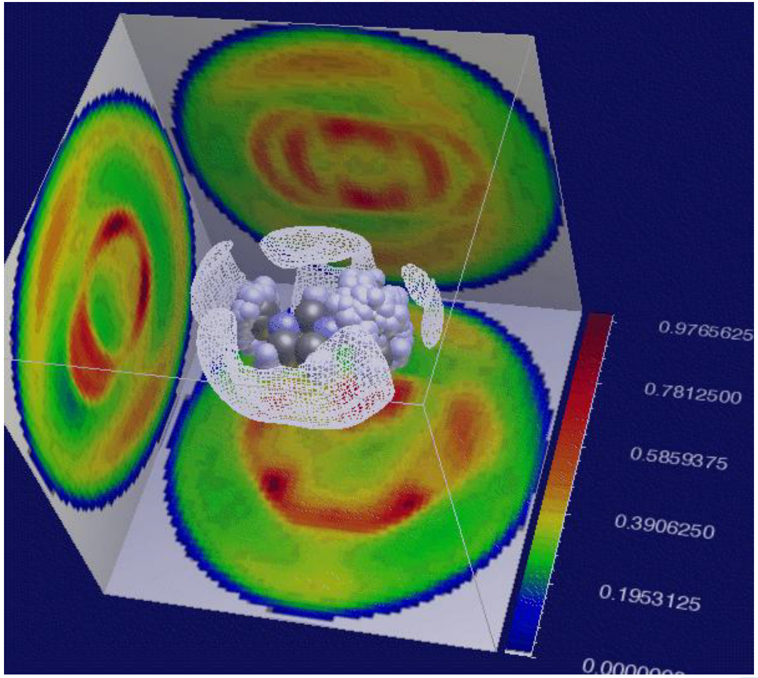
$$C_V = \frac{9 N k_B^3 \langle T \rangle^2}{6 k_B^2 \langle T \rangle^2 - 4 \left\langle \left(U - \langle U \rangle\right)^2 \right\rangle} \qquad \text{(NVE)}$$

Structural properties

The static structure factor can be easily calculated and compared to neutron or x-ray diffraction data



Liquid argon



Structural properties

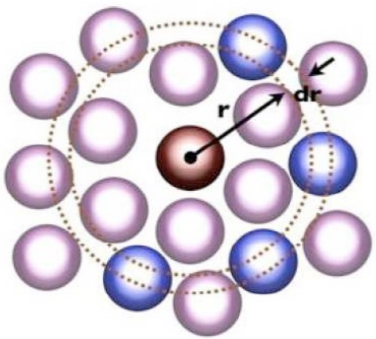
$$g_{\alpha\alpha}(r)=\frac{N}{\rho {N_\alpha}^2}\sum_{i=1}^{N_\alpha}\sum_{k\neq i}^{N_\alpha}\langle\delta(r-|r_k-r_i|)\rangle$$

$$g_{\alpha\beta}(r)=\frac{N}{\rho N_\alpha N_\beta}\sum_{i=1}^{N_\alpha}\sum_{k}^{N_\beta}\langle\delta(r-|r_k-r_i|)\rangle$$

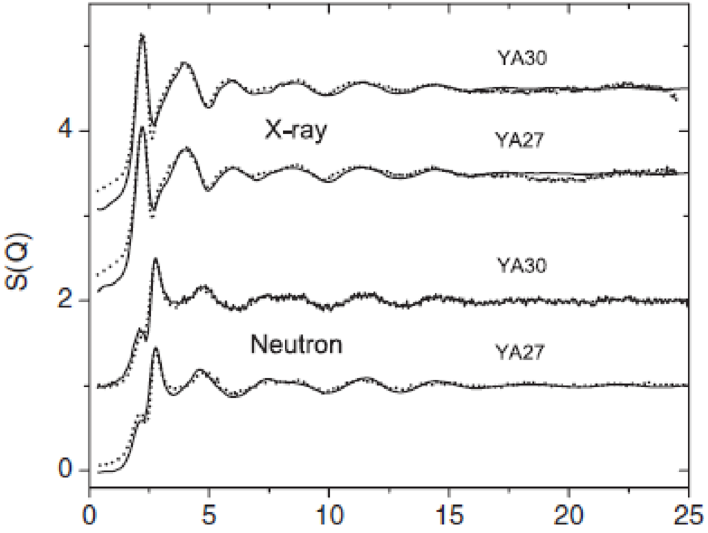
$$S_{\alpha\beta}(q)=S_{\beta\alpha}(q)=1+\frac{1}{c_\alpha c_\beta N}\left\langle\sum_{i,j\neq i}^{N_\alpha,N_\beta}\frac{\sin(qr_{ij})}{(qr_{ij})}\right\rangle,$$

$$F(q)\overset{\text{def}}{=}\sum^{n}c_\alpha c_\beta\overline{b}_\alpha\overline{b}^*_\beta[S_{\alpha\beta}(q)-1],$$

$$\tilde{F}_i(q)=\sum_{\alpha,\beta}^{n}c_\alpha c_\beta b_{\alpha i}(q)b^*_{\beta i}(q)[S_{\alpha\beta}(q)-1],$$

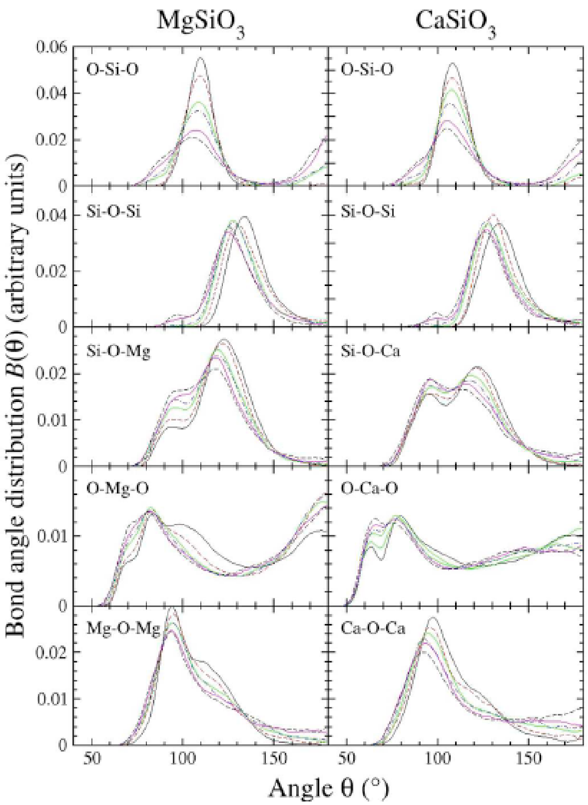


Set (100s-1000s) of configurations from MD or MC

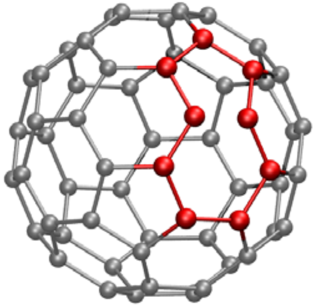


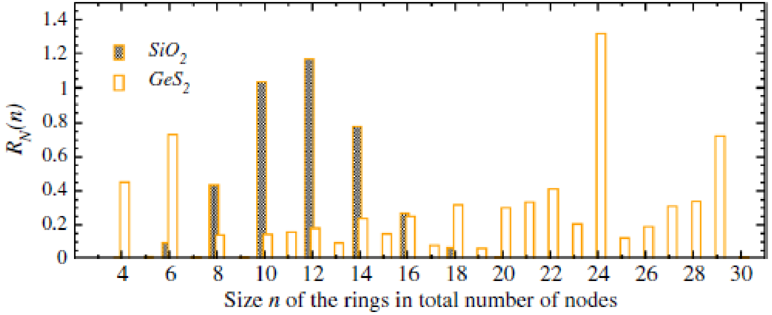
Du et al., J. Phys.: Condens. Matter 21 (2009) 205102

Structural properties



Bond angle distribution (changes with pressure)
Salmon et al., JNCS (2019)





Ring statistics
Le Roux & Jund, Comp. Mater. Sci. (2010)

R.I.N.G.S.

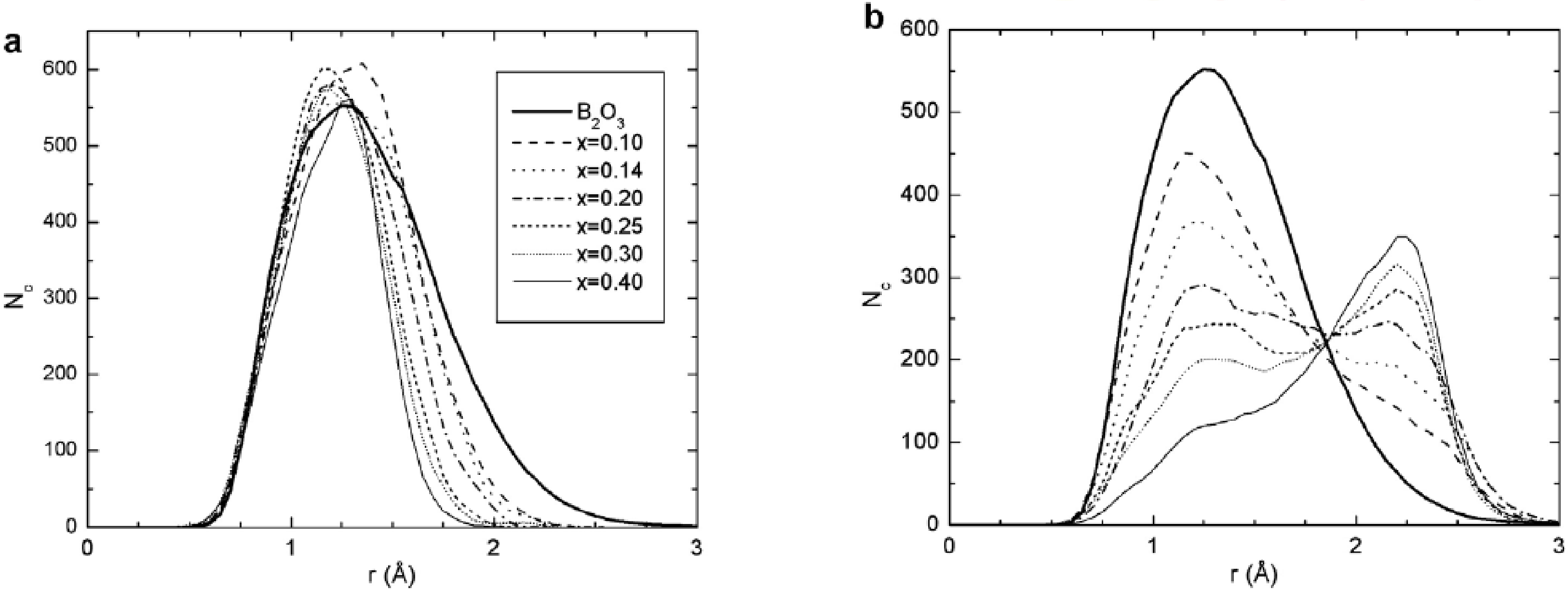
Rigorous Investigation of Networks Generated using Simulations

Structural properties

Voronoi-Delaunay tessellation to determine size of voids in the network

$(Li_2O)_x(B_2O_3)_{1-x}$

$(Cs_2O)_x(B_2O_3)_{1-x}$



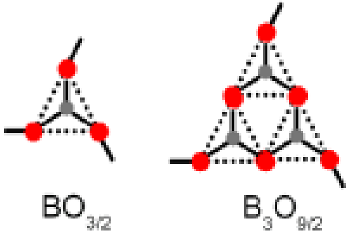
Distribution of voids in alkali borate glasses

Gonzalez et al., JNCS (2008)

Examples

Boroxol rings in B_2O_3

G. Ferlat et al., *Phys. Rev. Lett.* **101**, 065504 (2008)



Which is the fraction of B atoms inside boroxol rings?

Raman: Very intense peak at 808 cm^{-1} → f = 0.65-0.85

Diffraction: f = 0 – 0.82?

Diffraction (RMC): f < 0.30 if experimental density imposed

MD: f < 0.30 (problems: too high quench rate, system size, model)

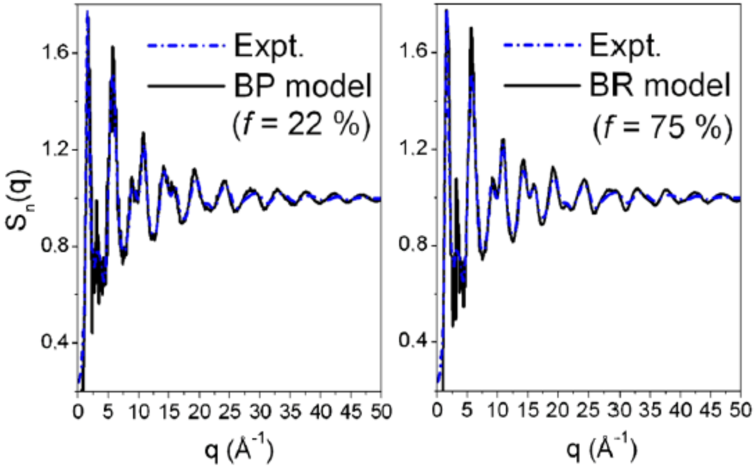
New DFT simulations:

100 atoms – 20 ps at 2000 K - ρ = 1.84 g/cm^3 (glass density at RT)

Quench to 300 K + 7 ps simulation at 300 K → f = 0.22

Create and simulate also artificial model with f = 0.75

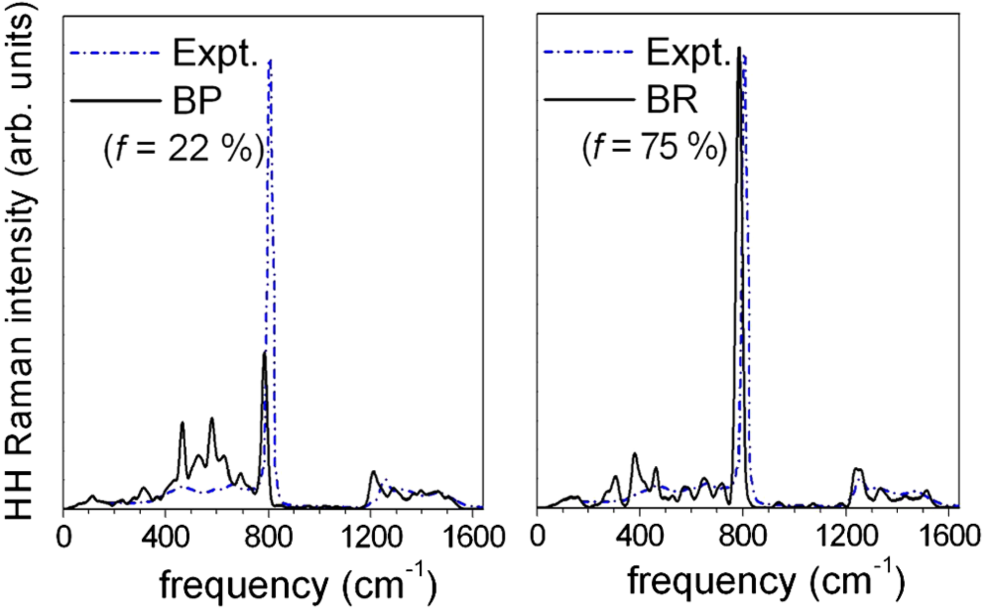
Both models are consisted with diffraction density and experimental density

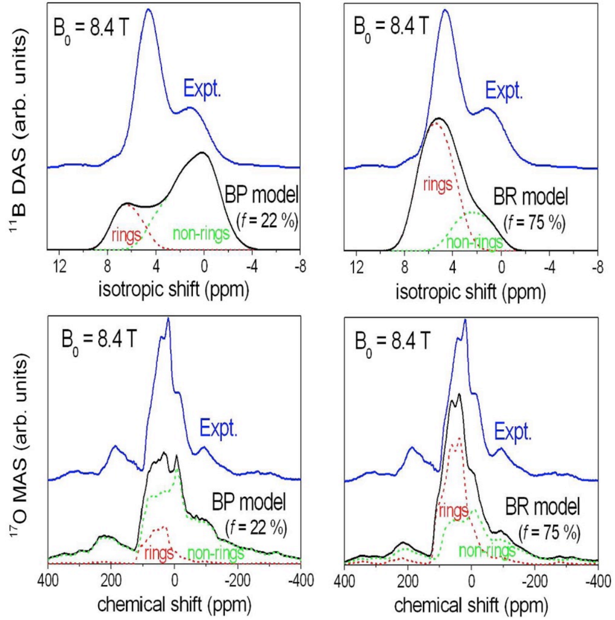


Boroxol rings in B_2O_3

G. Ferlat et al., *Phys. Rev. Lett.* **101**, 065504 (2008)

Comparison with diffraction or IR data does not discriminate between both models.
But comparison with NMR and specially Raman, indicates clearly that the boron rich model is the correct one!





Combining diffraction data and AIMD

A. Pandey et al., *Sci. Rep.* **6**, 33731 (2016)

“Simulation paradigm” = Use MC/MD simulation to produce a structural model

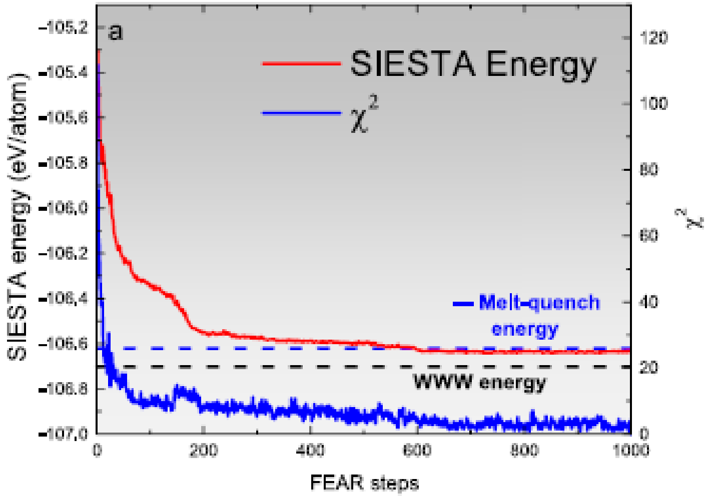
“Information paradigm” = Use XRD or ND to create the model (through RMC)

Each approach has their own limitations, so **can we combine both?**
AIFEAR: Ab initio Force Enhanced atomic refinement

Method:

1. Prepare a random model
2. Do RMC (M moves) → minimize χ^2
3. Do AIMD (N steps) →minimize energy
4. Go to 2 and repeat until reaching desired accuracy

$$\chi^2 = \sum_i \left[\frac{F_E(k_i) - F_M(k_i)}{\sigma(k_i)}\right]^2,$$



Combining diffraction data and AIMD

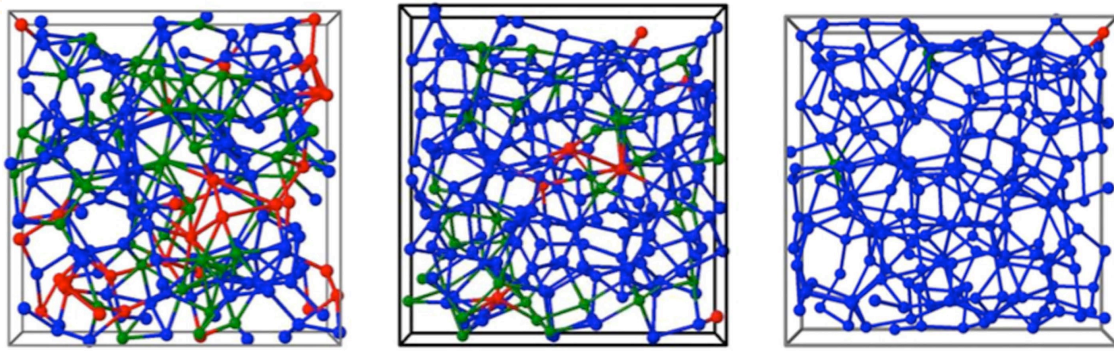
A. Pandey et al., *Sci. Rep.* **6**, 33731 (2016)

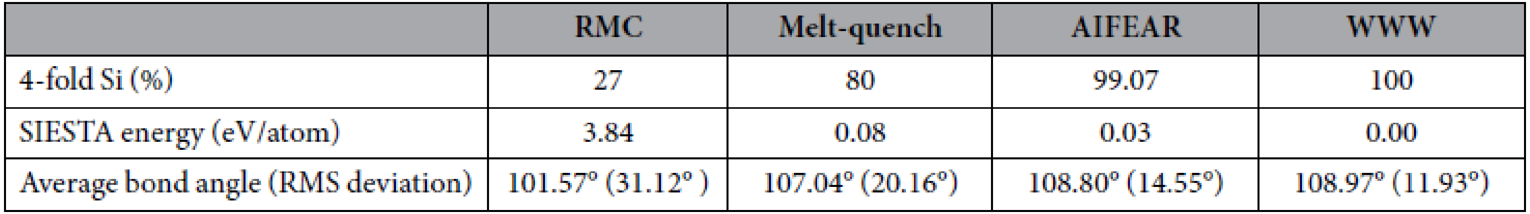
216-atom model of amorphous Si

Coordination: ● 3 ● 4 ● 5

RMC = model w too many defects

DFT alone (melt-quench) = less defects, but worse agreement for g(r)


a) b) c)

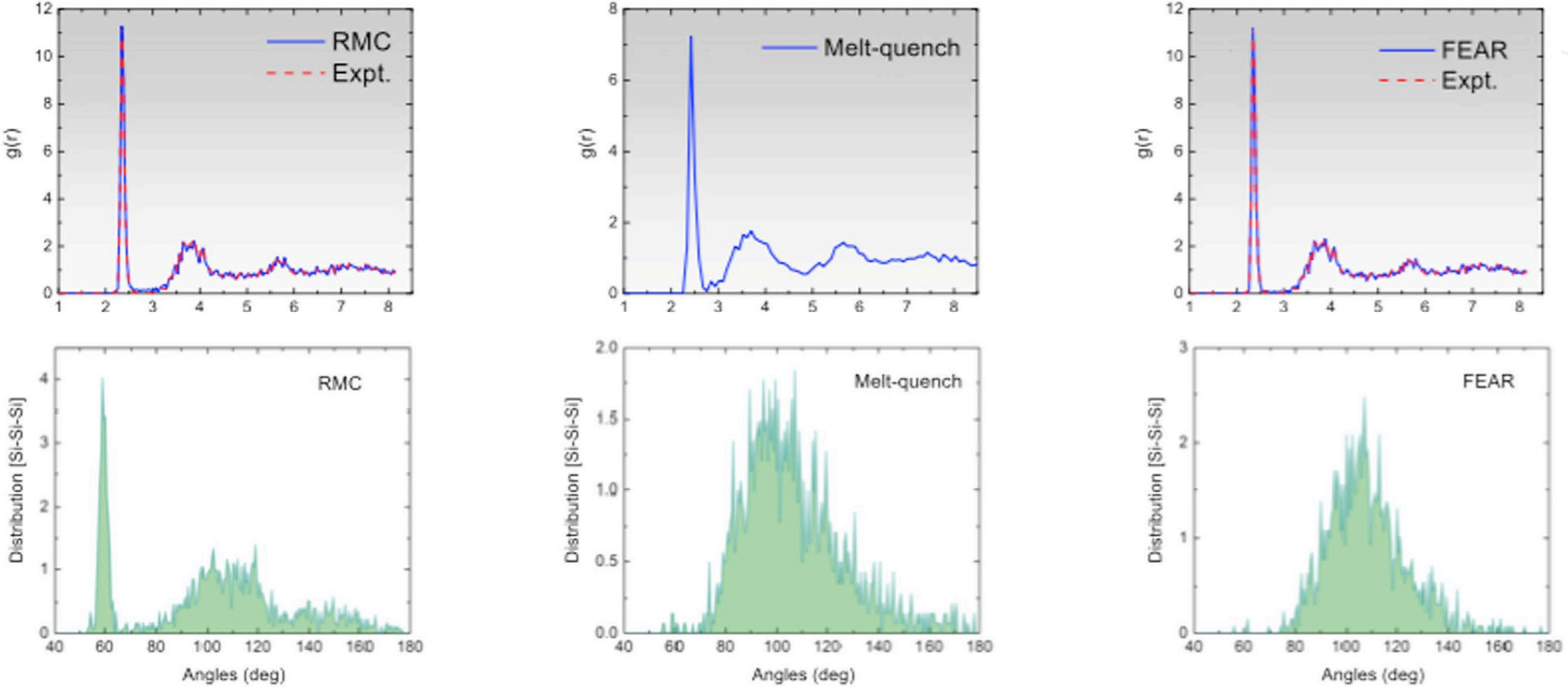


	RMC	Melt-quench	AIFEAR	WWW
4-fold Si (%)	27	80	99.07	100
SIESTA energy (eV/atom)	3.84	0.08	0.03	0.00
Average bond angle (RMS deviation)	101.57° (31.12°)	107.04° (20.16°)	108.80° (14.55°)	108.97° (11.93°)

Table 1. Total energy and key structural properties of *a*-Si models. The energy per atom is expressed with reference to the energy of the WWW model.

Combining diffraction data and AIMD

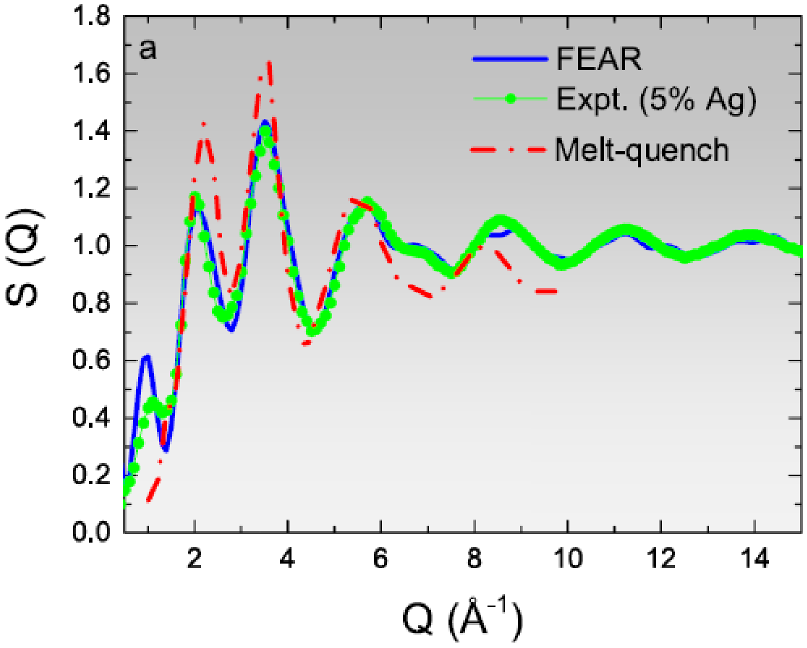
A. Pandey et al., *Sci. Rep.* **6**, 33731 (2016)

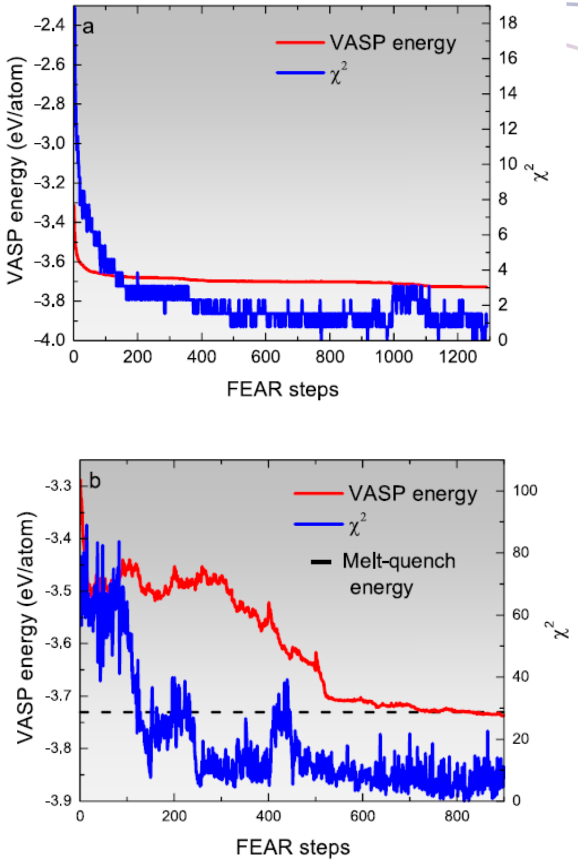


Combining diffraction data and AIMD

A. Pandey et al., *Sci. Rep.* **6**, 33731 (2016)

$Ag_x(GeSe_3)_{1-x}$ with x = 0.05 and 0.07





Cooling rate effects in sodium silicates

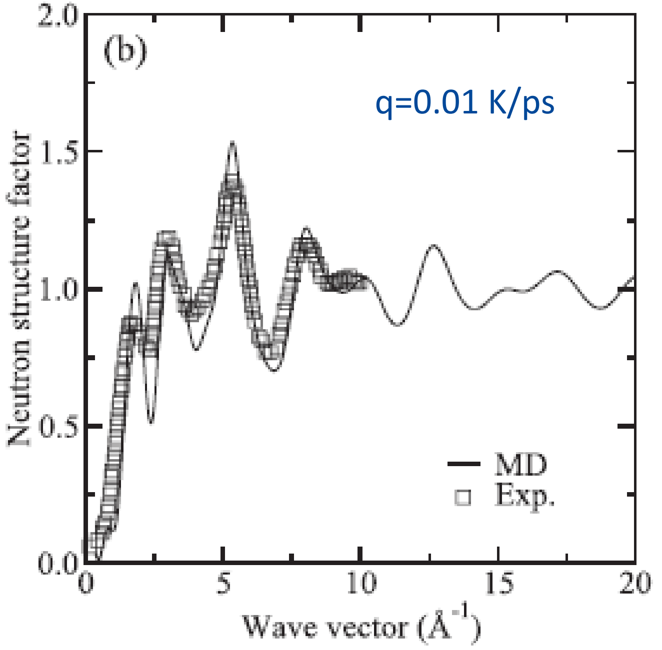
X. Li et al., *J. Chem. Phys.* **147**, 074501 (2017)

Glass properties depend on their thermal history. So can we trust MD simulations where the cooling rate is ~ 10^{12} K/s vs typical experimental cooling rates of 1-100 K/s?

Study $(Na_2O)_{30}(SiO_2)_{70}$ at different cooling rates!

Protocol:

- 3000 atoms melted at 4000 K (10 ps NVT + 100 ps NPT)
- NPT cooling to 300 K at q = 100 (37 ps), 10, 1, 0.1 and 0.01 (**370 ns**) K/ps
- Relax at 300 K (10 ps NPT) + production run (100 ps NVT)
- Analysis of 100 configurations saved during production run

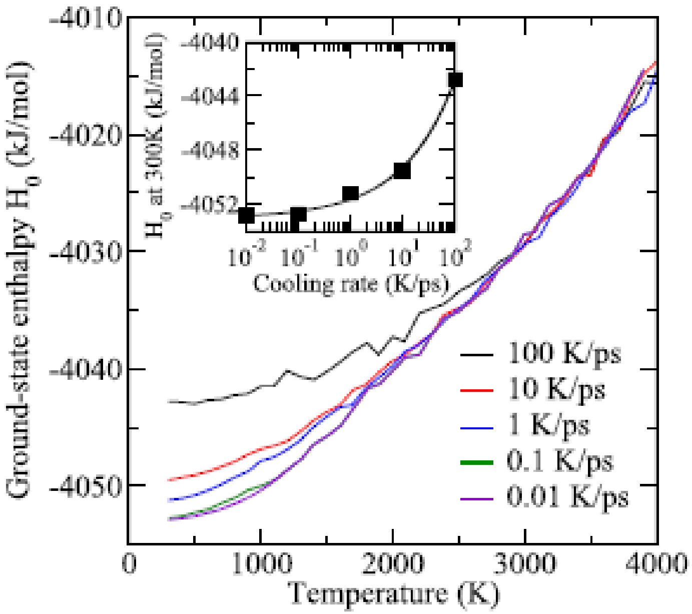


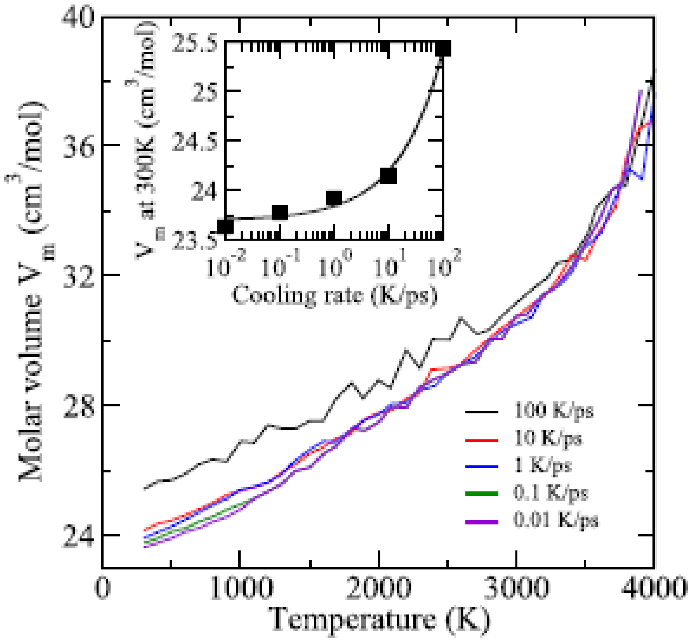
Reasonable agreement between MD and neutron diffraction

Cooling rate effects in sodium silicates

X. Li et al., *J. Chem. Phys.* **147**, 074501 (2017)

- Enthalpy and volume show the expected dependence on q, with the fictive T decreasing for slower quench rates

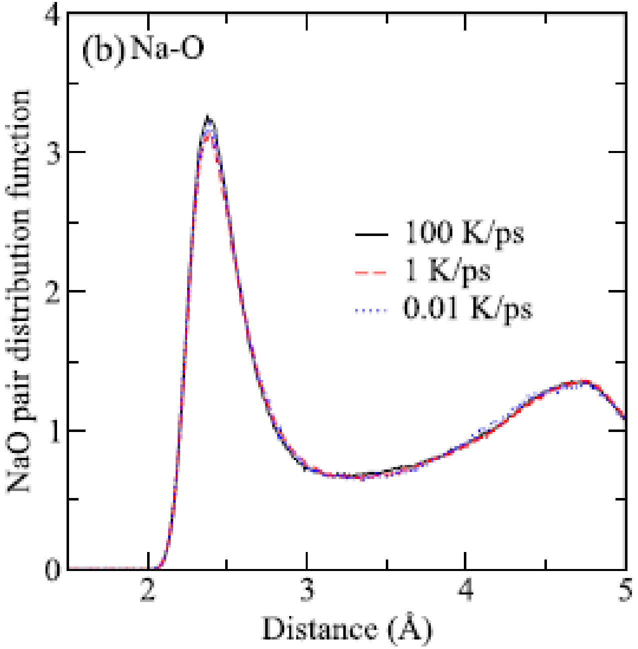


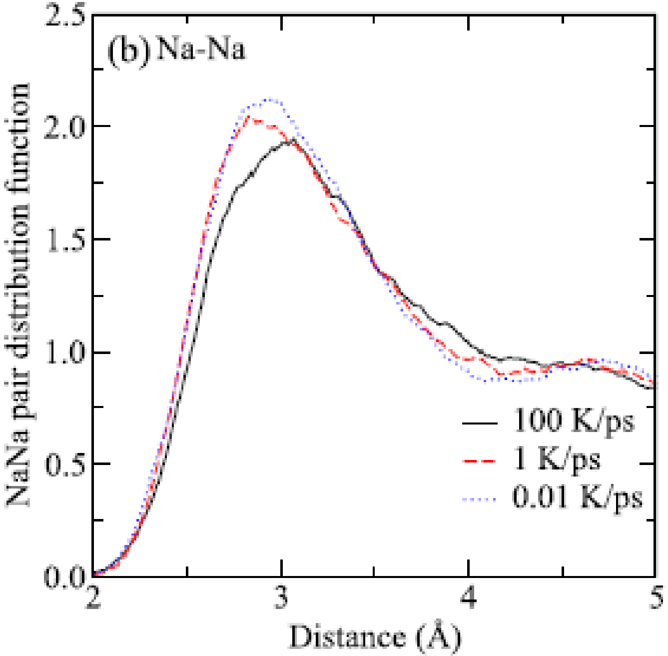


Cooling rate effects in sodium silicates

X. Li et al., *J. Chem. Phys.* **147**, 074501 (2017)

- No appreciable changes in local environment of Si and Na (Si-O and Na-O pdfs)
- Trends observed in Si-Si and Na-Na pdfs indicate a tendency to form a more ordered network with decreasing cooling rate



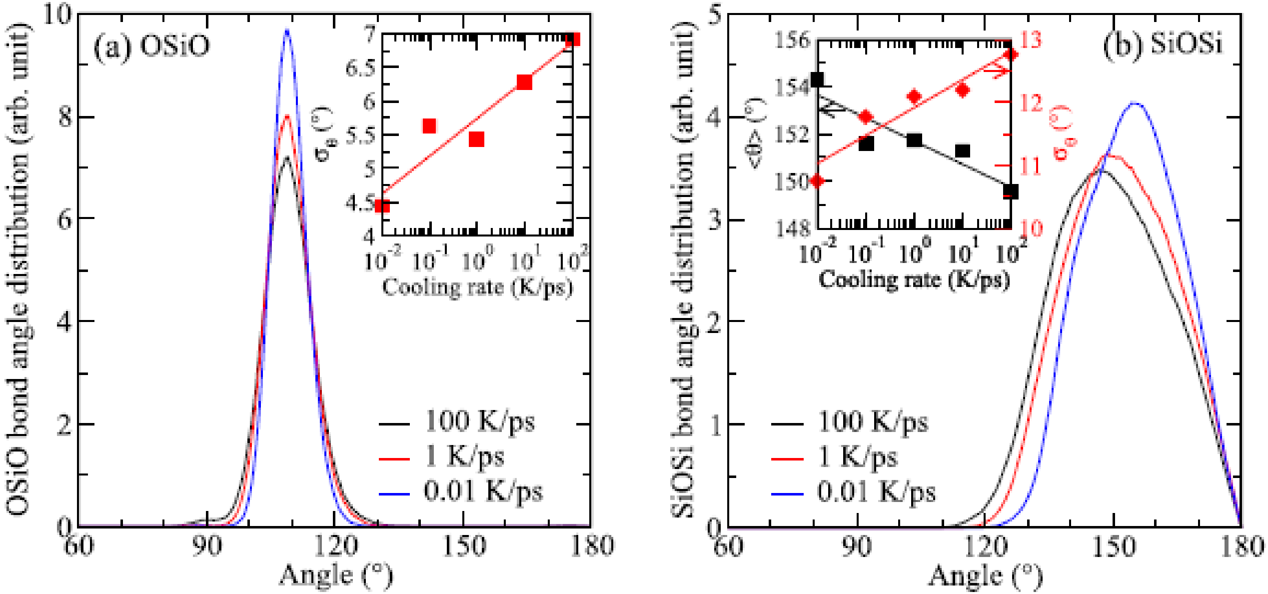


NEUTRONS FOR SOCIETY®

Cooling rate effects in sodium silicates

X. Li et al., *J. Chem. Phys.* **147**, 074501 (2017)

- Effect of cooling rate is more visible in the bond angle distributions
- Lower cooling rates result in higher angular order and Si-O-Si shows a clear shift



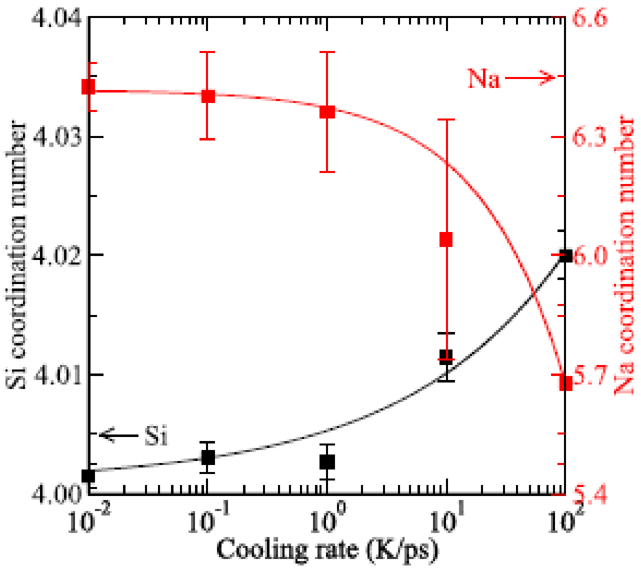
ANGEVIN

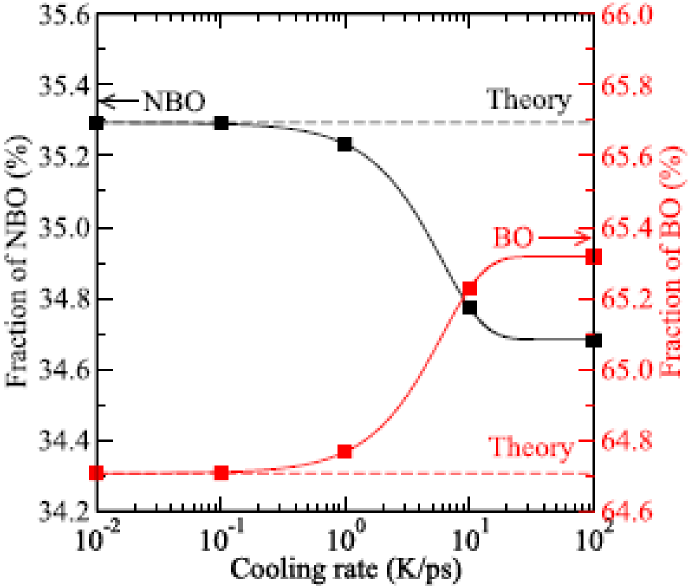
NEUTRONS FOR SOCIETY®

Cooling rate effects in sodium silicates

X. Li et al., *J. Chem. Phys.* **147**, 074501 (2017)

- Coordination defects disappear at lower cooling rates
- Na tends to form local clusters with lower cooling rate
- Fraction of non-bridging O increases as cooling rate decreases, approaching the theoretical value obtained by assuming that each Na creates 1 NBO

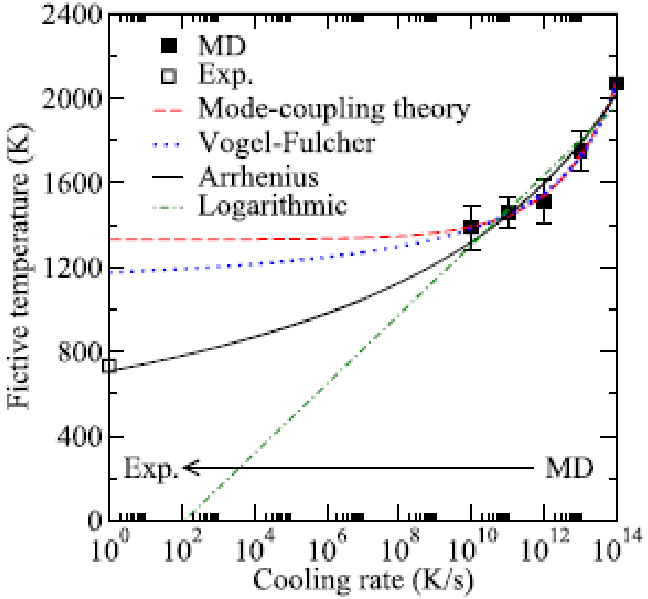


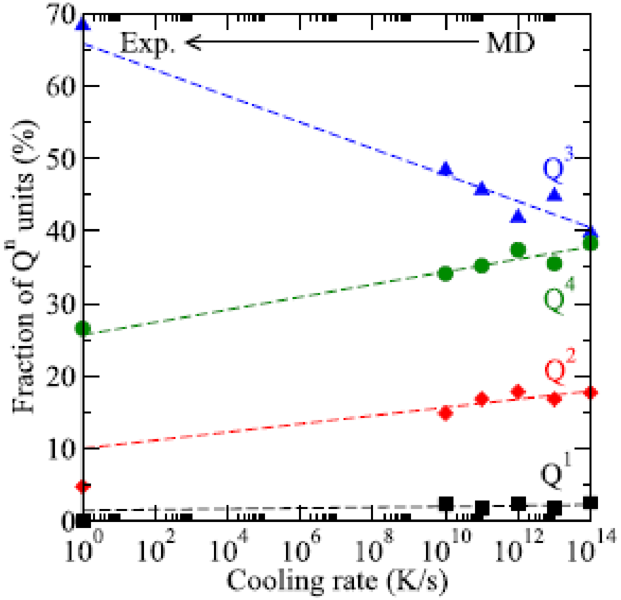


Cooling rate effects in sodium silicates

X. Li et al., *J. Chem. Phys.* **147**, 074501 (2017)

- Q^n distribution (obtained from NMR) is one of most challenging properties to predict for MD
- Extrapolating with a log function gives results in good agreement with experiment
- MD are intrinsically limited to high cooling rates, but their output can be compared to real glasses by extrapolation





Medium range structure of sodium silicates

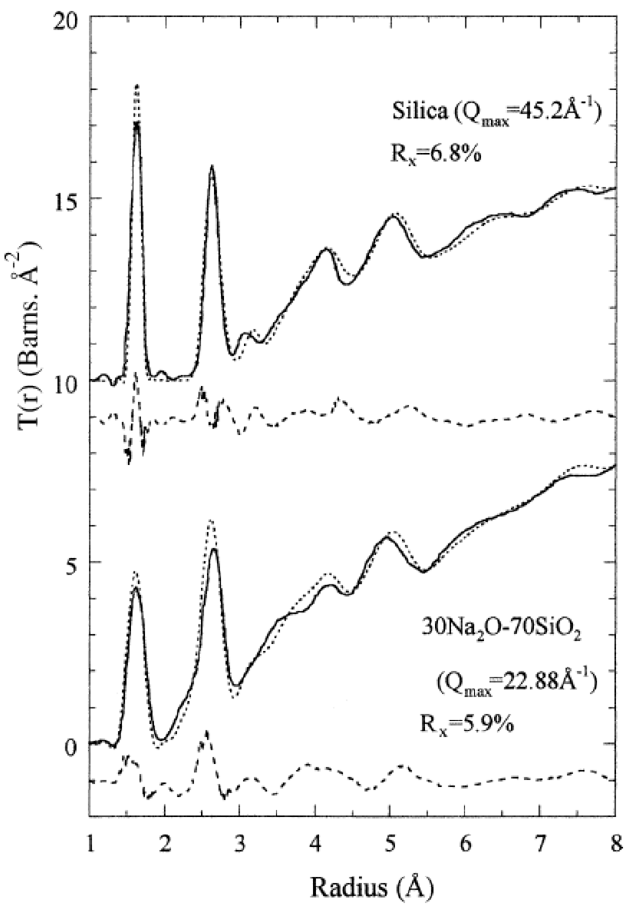
Du & Cormack, *J. Non-Cryst. Solids* **349**, 66 (2004)

Study of series $(Na_2O)_x(SiO_2)_{1-x}$ glasses with x = 0, 0.1, 0.2, 0.3, 0.4 and 0.5

3000 atoms → L = 34.08 – 35.23 Å

Protocol:

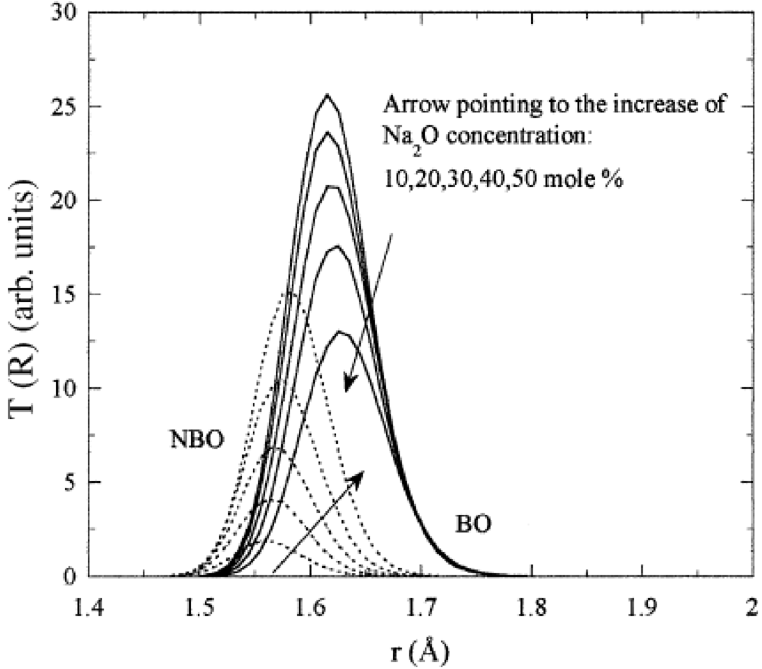
- Melt at 6000 K (80 ps)
- Quench to 300 K (570 ps → q = 10 K/ps)
- Equilibration (20 ps)
- Production (20 ps)

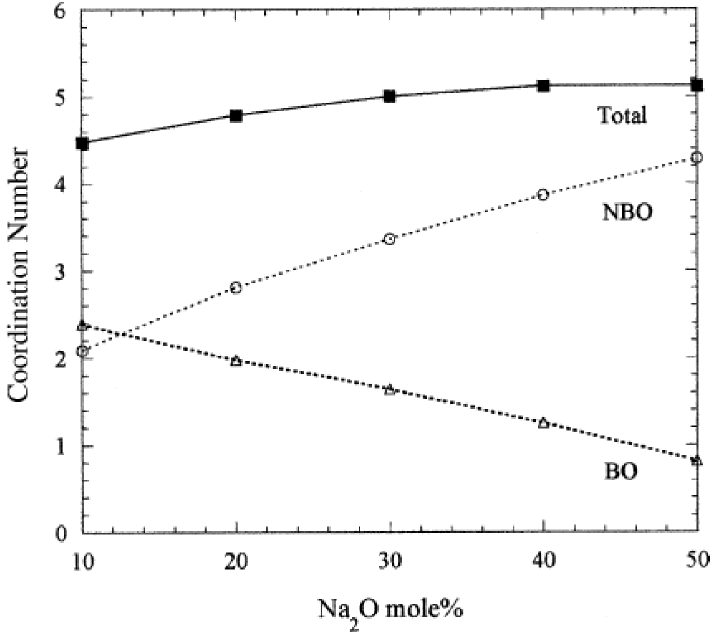


Medium range structure of sodium silicates

Du & Cormack, *J. Non-Cryst. Solids* **349**, 66 (2004)

- Si-O bond length decreases only slightly, but changes in Si-NBO and Si-BO are stronger (and both increase length)
- Coordination numbers consistent with EXAFS



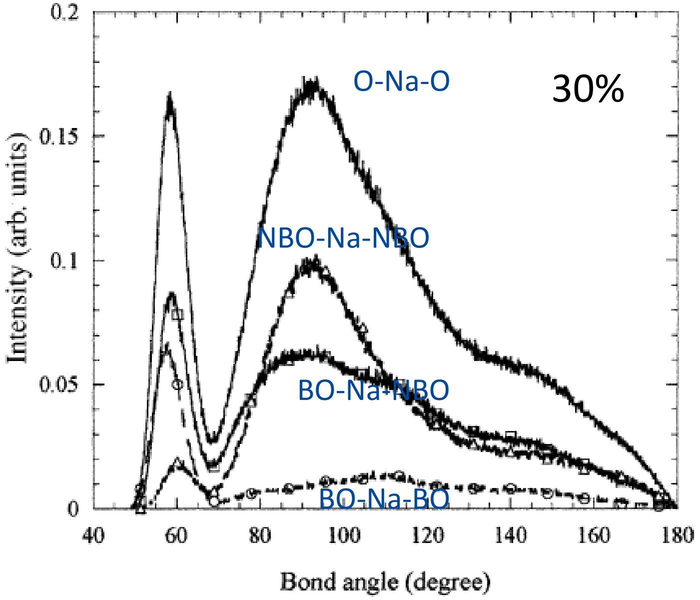


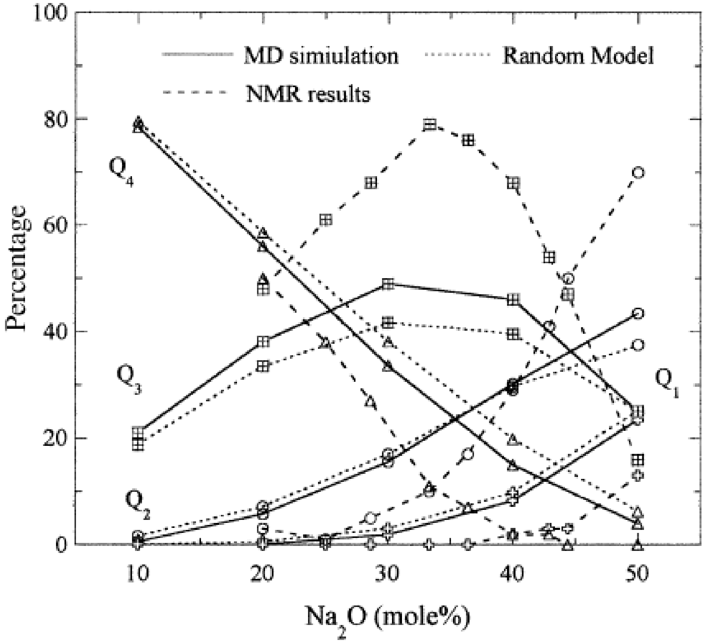
I N S T I

Medium range structure of sodium silicates

Du & Cormack, *J. Non-Cryst. Solids* **349**, 66 (2004)

- Trends in Q^n agree well with experiment (e.g. max in Q^3), but no quantitative agreement (possibly due to very fast quenching, as we saw before)

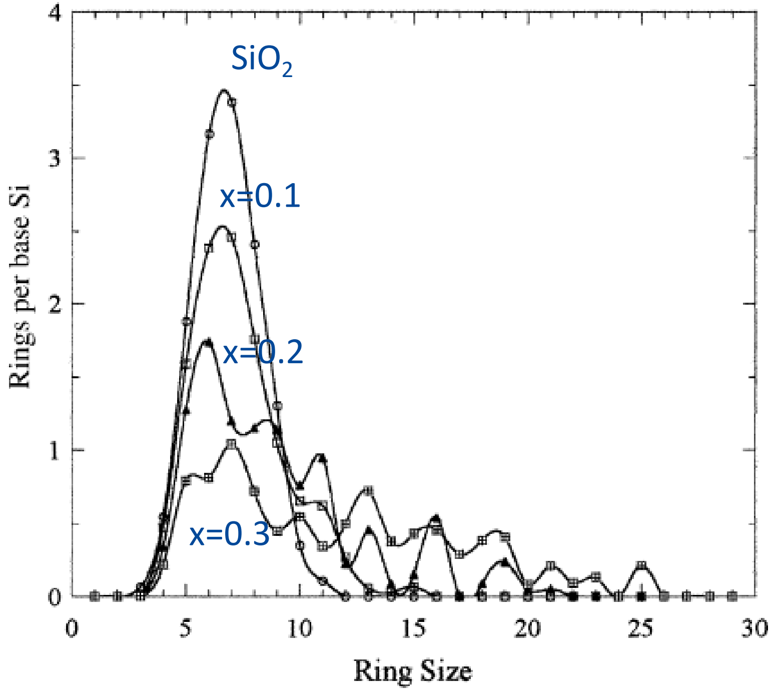


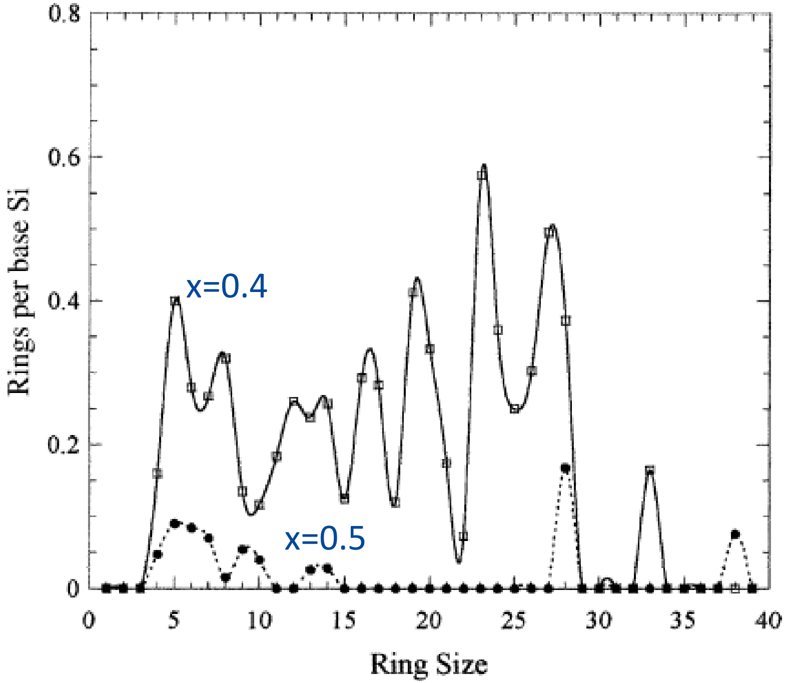


Medium range structure of sodium silicates

Du & Cormack, *J. Non-Cryst. Solids* **349**, 66 (2004)

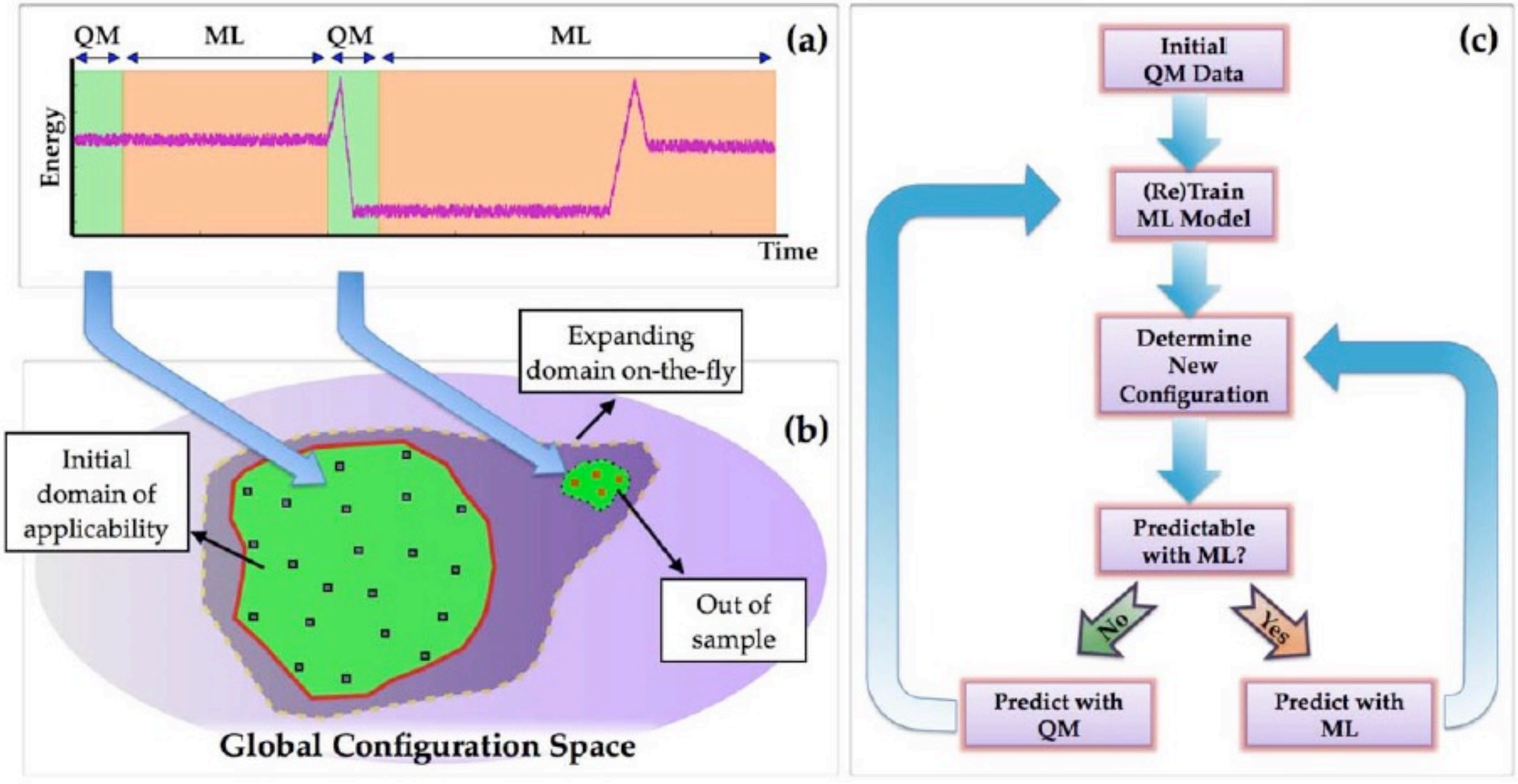
- Decrease of 6-rings and appearance of larger rings
- Q^n and ring distributions interpreted as sign of inhomogeneities → silica rich and sodium rich areas





Perspectives: Machine Learning to accelerate AIMD

Botu & Ramprasad, *Int. J. Quantum Chem.* **115**, 1074 (2015)

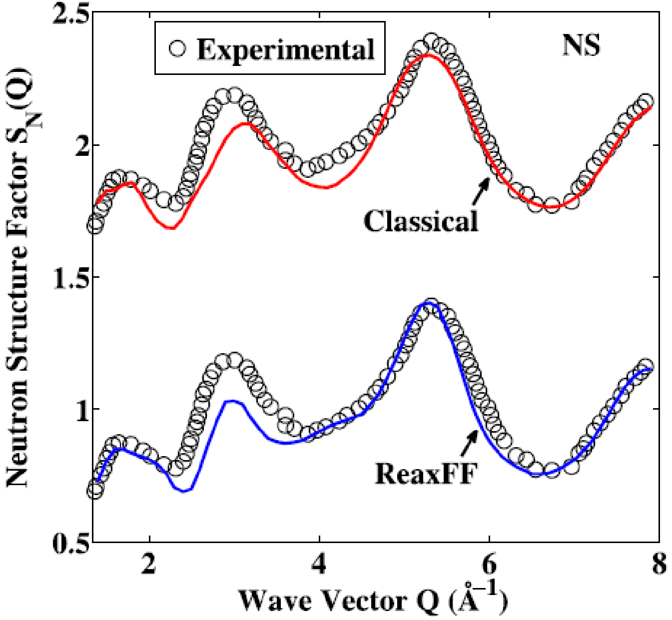


Perspectives: Reactive potentials

Yu et al., *Int. J. Appl. Glass Sci.* **8**, 276 (2017)

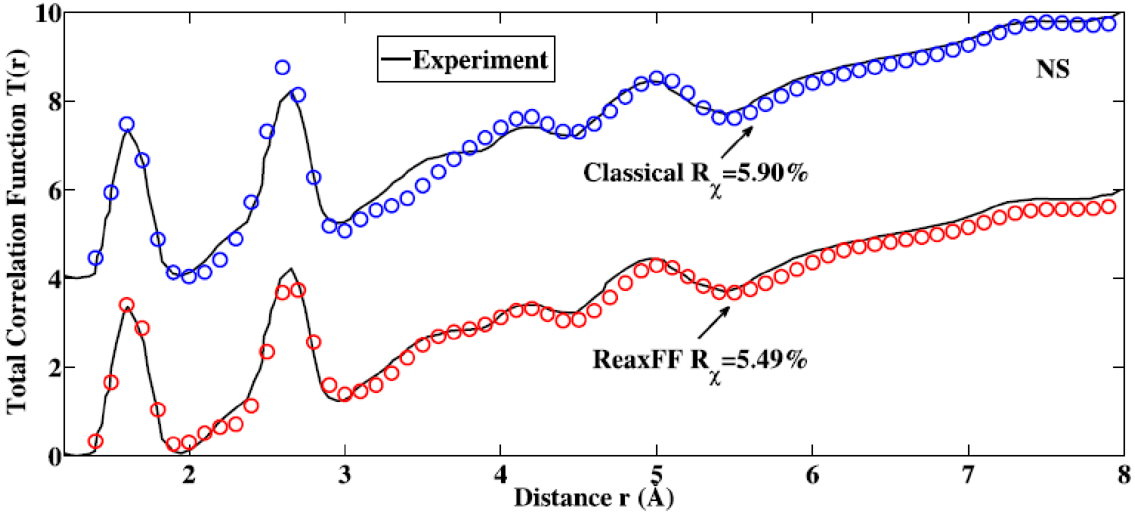
$$E_{\text{system}} = E_{\text{bond}} + E_{\text{vdWaals}} + E_{\text{Coulomb}} + E_{\text{under}} + E_{\text{over}} + E_{\text{lp}} + E_{\text{val}} + E_{\text{pen}} + E_{\text{torsion}} + E_{\text{conj}}$$

- ReaxFF allows to reproduce bond breaking and forming
- Charges are assigned dynamically
- Potential energies adapted to local order
- 10 times slower than standard MD
- Application to $(Na_2O)_{0.3}(SiO_2)_{0.7}$:
  - 3000 atoms
  - Melt at 4000 K (NPT 100 ps)
  - NPT cool to 300 K at 1 K/ps using standard Teter FF
  - Relax using ReaxFF for 1 ns



Perspectives: Reactive potentials

Yu et al., *Int. J. Appl. Glass Sci.* **8**, 276 (2017)



- ReaxFF overcomes the limited transferability of empirical potential by accounting for the changes in the environment of each atom
- Demonstrated by ability to model sodium silicate and glassy silica with same parameters
- ReaxFF captures the destabilization of Si-O bonds caused by addition of alkali modifiers

Thank you!

Questions?