Day 2: ReaxFF parameterization

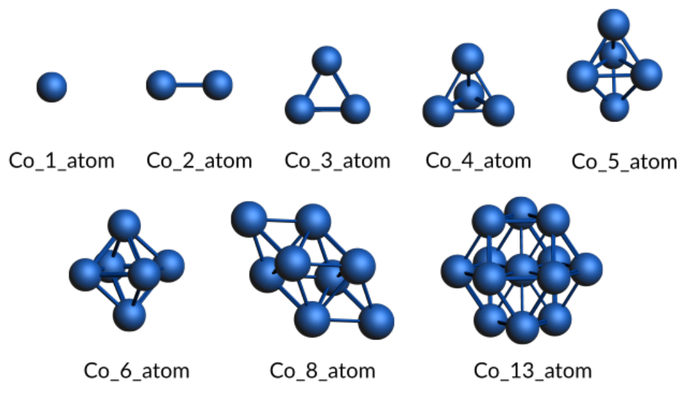
Hands-on workshop Chemistry & Materials with the Amsterdam Modeling Suite

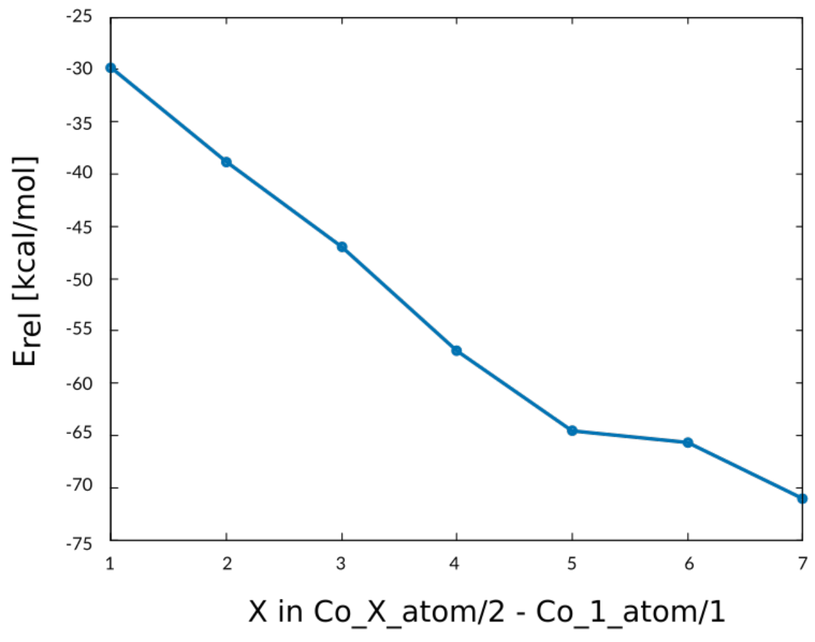
2-day workshops October 2018, Tsinghua & DLUT
Fedor Goumans, goumans@scm.com SCM support: support@scm.com
FermiTech support: wiki, support@fermitech.com.cn

Making Computational Chemistry Work for You

Program

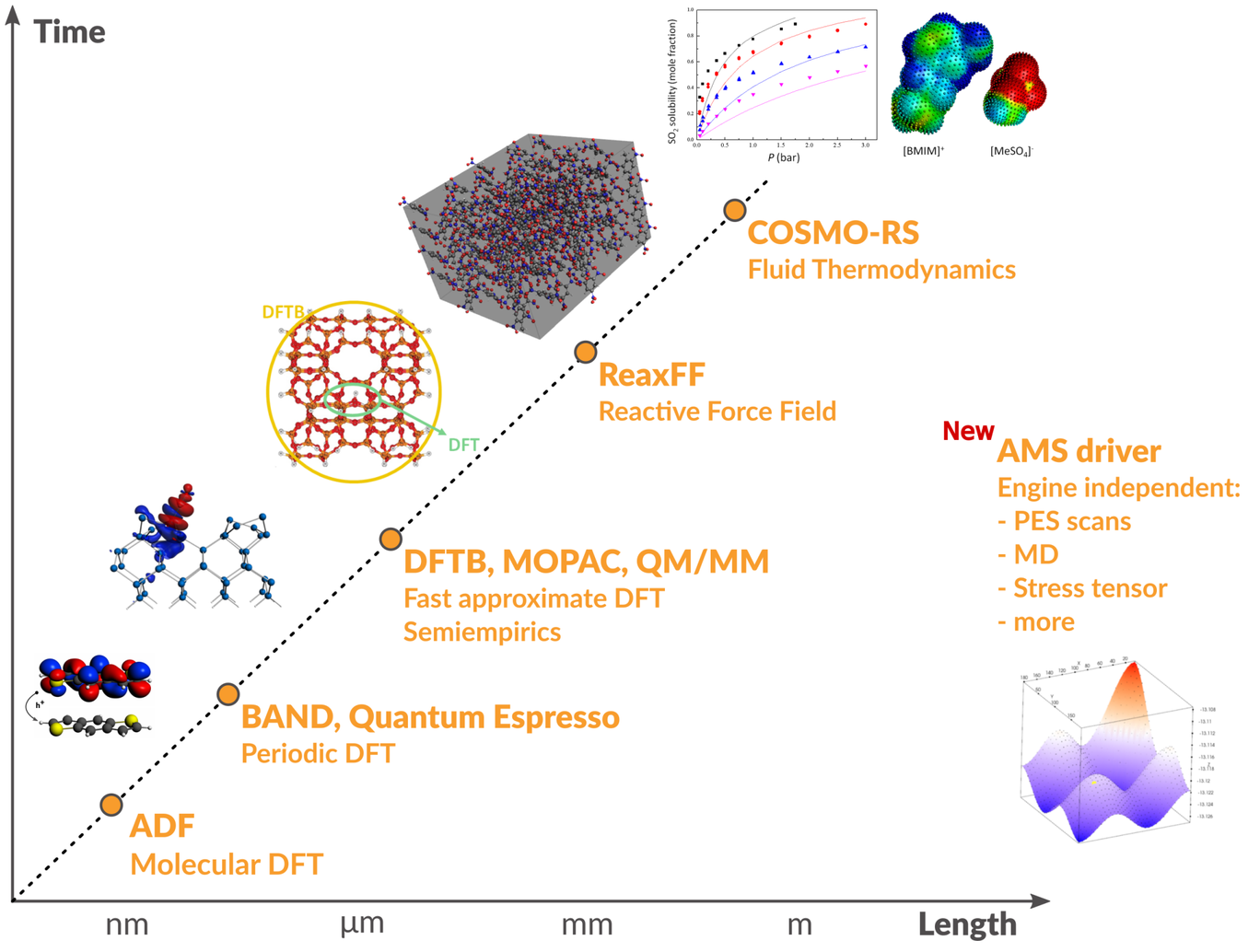
- Introduction: ReaxFF
  - Some advanced exercises, new features
- Acceleration techniques
  - fbMC, GCMC, CVHD, bond boost => polymer structures
- Refitting a force field for parameterization (slides, input files)
  - Scripting on Windows
  - Adding geometries, conformers, reaction pathways, bond scans
  - Optimizing the force field, checking errors & improving





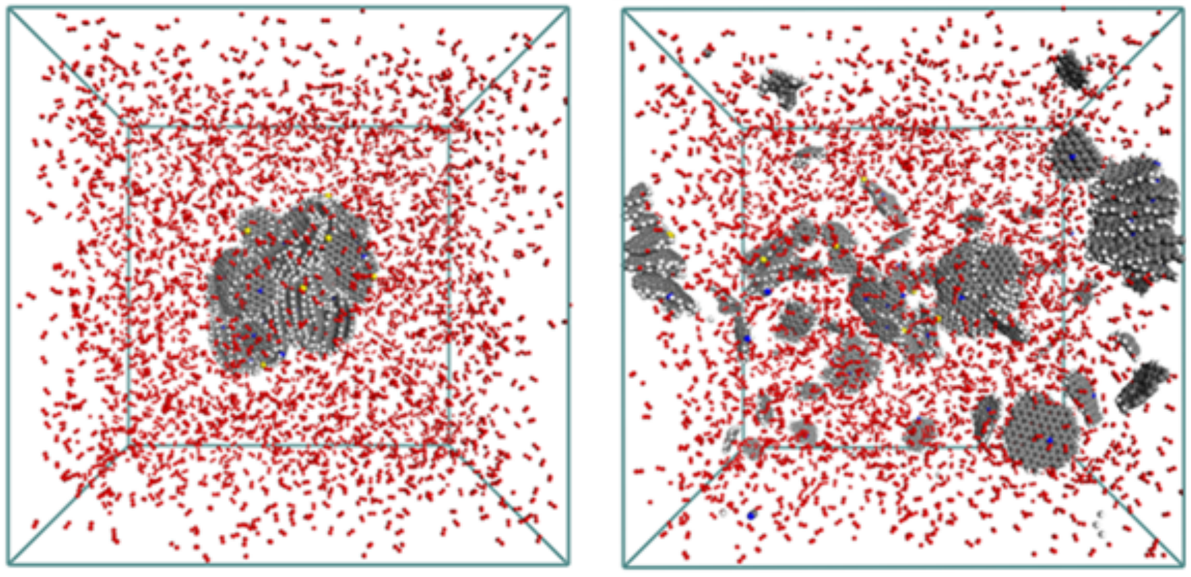
Amsterdam Modeling Suite

- ADF: powerful molecular DFT
  - Spectroscopy: NMR, EPR, VCD, UV, XAS
  - Advanced solvation / environments
- BAND: periodic DFT
  - (2D) Materials
- DFTB: fast approximate DFT
- ReaxFF: Reactive MD
  - Dynamics of large complicated systems
- COSMO-RS: fluid thermodynamics
  - VLE, LLE, logP, solubility
- Integrated GUI – use out of the box
- Scripting: workflows & automation



Computational Chemistry & Materials

$$d^2x/dt^2 = F(x)$$
$$F(x) = - dV(x)/dx$$

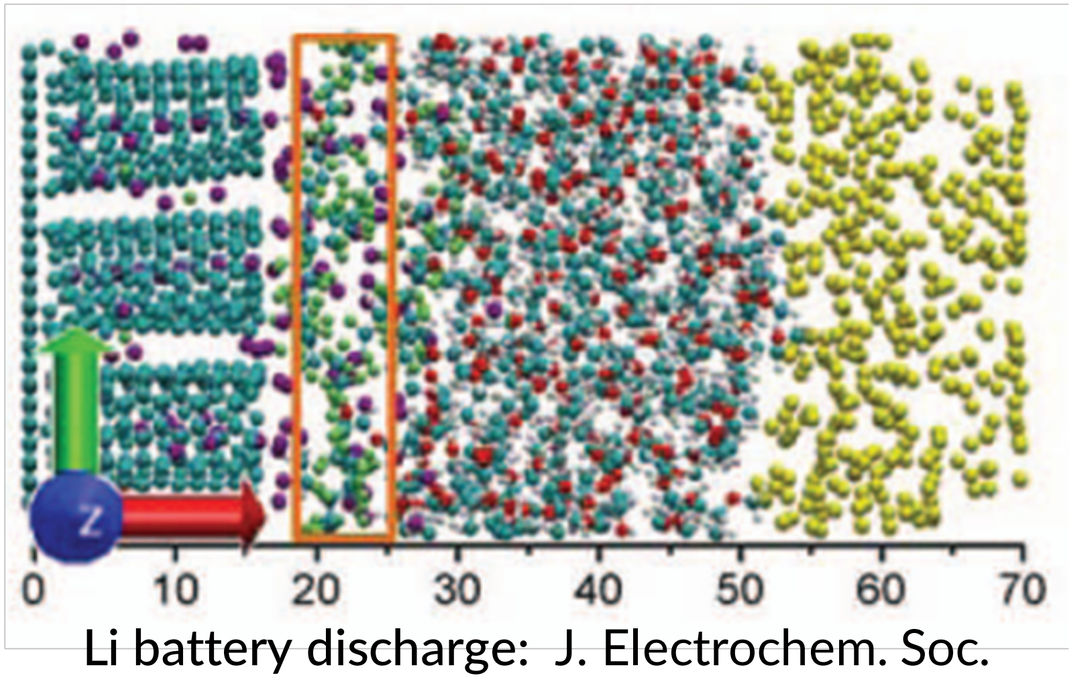


Molecular dynamics: Newton's equations of motion

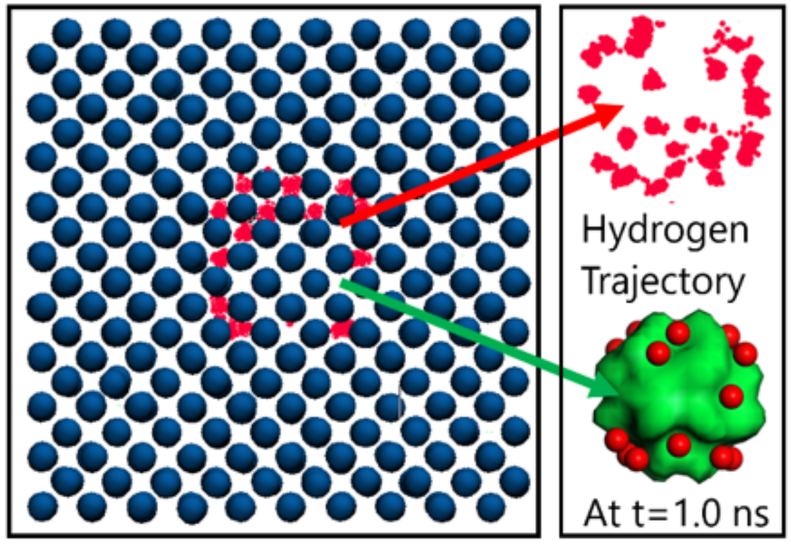
Movement of atoms: solve numerically + propagate,

Properties: reaction rates, diffusion coefficients, stress-strain, ….

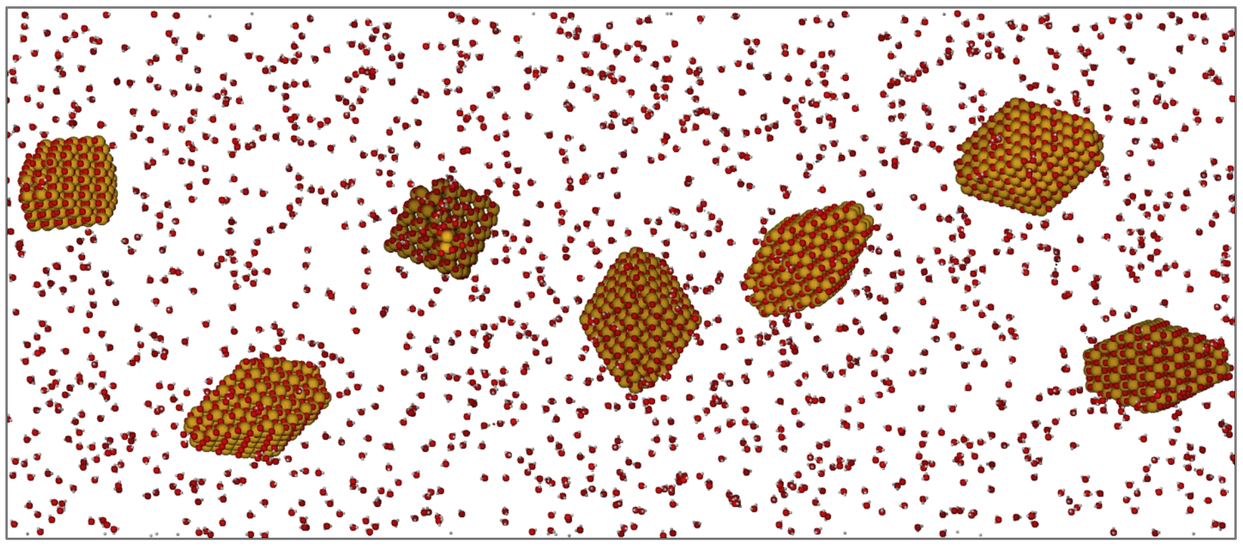
ReaxFF – reactive molecular dynamics



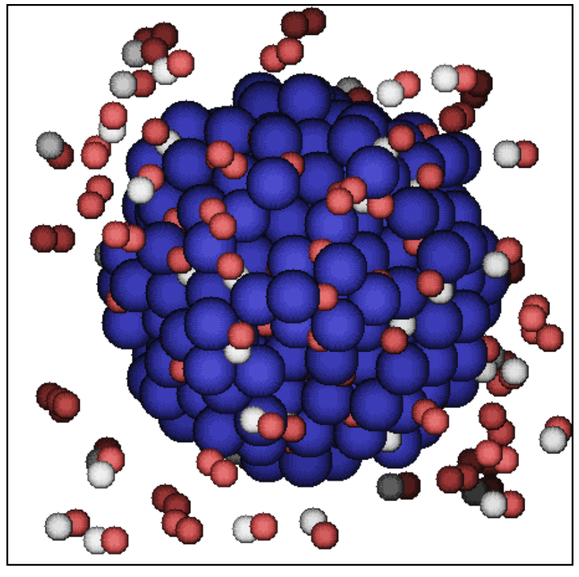
Li battery discharge: J. Electrochem. Soc. **161**, E3009 (2014); PCCP, **17**, 3383 (2015)



Hydrogen embrittlement of steels
Phys. Chem. Chem. Phys. 18 761-771 (2016)



Crystallization TiO_2 nano-particles in water
Nano Lett. **14**, 1836-1842 (2014)



Pd-catalysed CO oxidation GCMC+ReaxFF
J. Chem. Phys., **139** 044109 (2013)

ReaxFF: introduction

- Simulate complex systems at realistic scales
  - Atomistic potentials: bond orders + charge update

A.C.T. van Duin et al ,J. Phys. Chem. A 2001 , 105, 9396-9409.

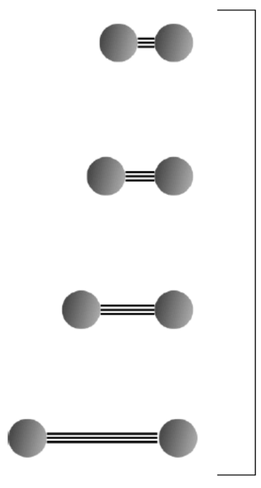
Standard forcefields **vs** *ReaxFF*

Harmonic potentials based on atom distance, bond breaking impossible, e.g.

$E_{bond} \propto (distance)^2$

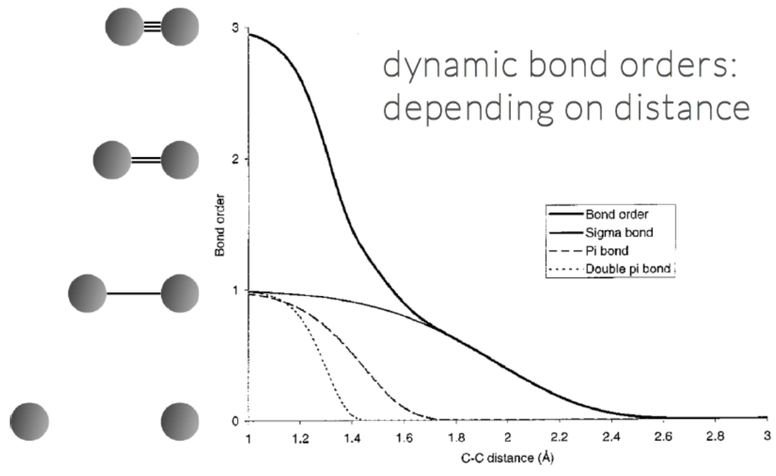
Non-harmonic potentials based on bond orders, bond breaking/forming possible, e.g.

$E_{bond} \propto -(bond\ order) \times \exp[\ (1 - bond\ order)\ ]$



a triple "bond" will always stay a triple "bond"...

($E_{bond} \rightarrow \infty$)



dynamic bond orders: depending on distance

($E_{bond} \rightarrow 0$)

A.C.T. van Duin *et al*,*J. Phys. Chem. A* **2001**, 105, 9396-9409.

See ReaxFF intro slides

General ReaxFF rules

- No discontinuities in energy or forces
- No pre-defined reaction sites or types
- Only 1 atom type per element

$$E_{\text{system}} = E_{\text{bond}} + E_{\text{lp}} + E_{\text{over}} + E_{\text{under}} + E_{\text{val}} + E_{\text{pen}} + E_{\text{coa}} + E_{\text{C}_2} + E_{\text{tors}} + E_{\text{conj}} + E_{\text{H-bond}} + E_{\text{vdWaals}} + E_{\text{Coulomb}}$$

$$BO'_{ij}(r_{ij}) = \exp\left[p_{bo,1} \cdot \left(\frac{r_{ij}}{r_o}\right)^{p_{bo,2}}\right] + \exp\left[p_{bo,3} \cdot \left(\frac{r^{\pi}_{ij}}{r_{o,\pi}}\right)^{p_{bo,4}}\right] + \exp\left[p_{bo,5} \cdot \left(\frac{r^{\pi\pi}_{ij}}{r_{o,\pi\pi}}\right)^{p_{bo,6}}\right]$$

In: distance between atoms, r_{ij}

Out: 1, 2, 1.42, etc...

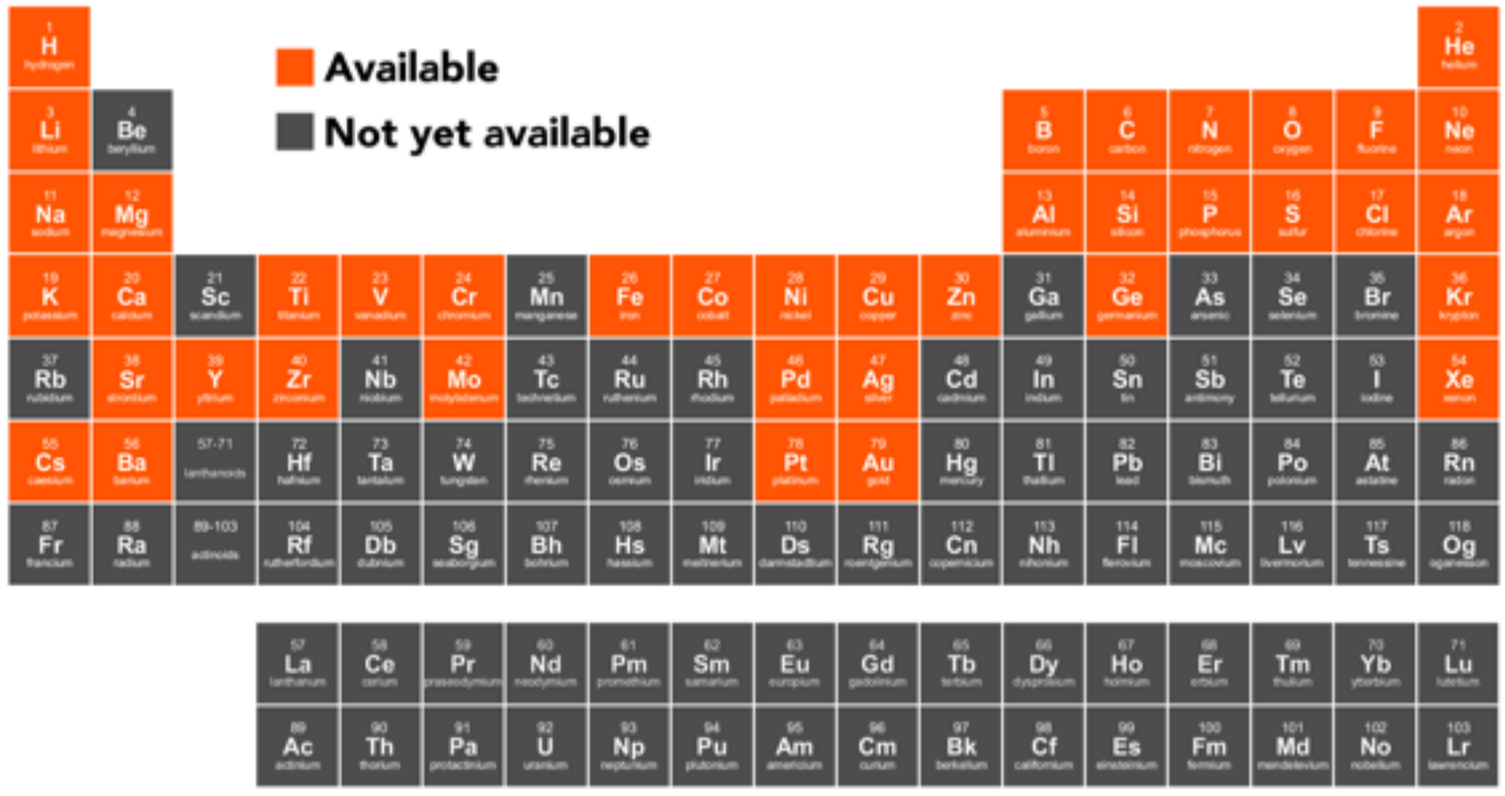
Parameters = 16

$p_{bo,1}, p_{bo,2}, p_{bo,3}, p_{bo,4}, p_{bo,5}, p_{bo,6}, r_o, r_{o,\pi}, r_{o,\pi\pi}$
$val_1, val_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$

Correction terms f_1, f_2, f_3: $\mathbf{BO_{ij}(r_{ij})} = BO'_{ij}(r_{ij}) \cdot f_1(BO'_{ij}) \cdot f_2(BO'_{ij}) \cdot f_3(BO'_{ij})$

ReaxFF parameters, transferability

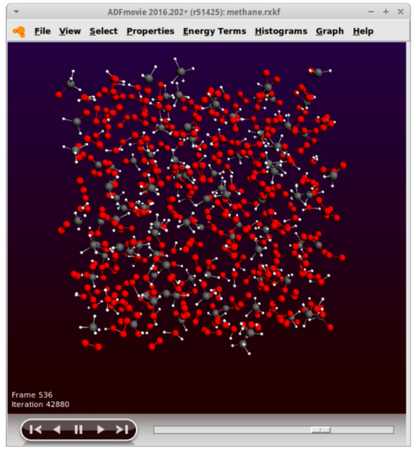
- many elements studied
- each pair needs bonded terms
- validate force field
  - GUI checks
  - training data crucial
  - application specific

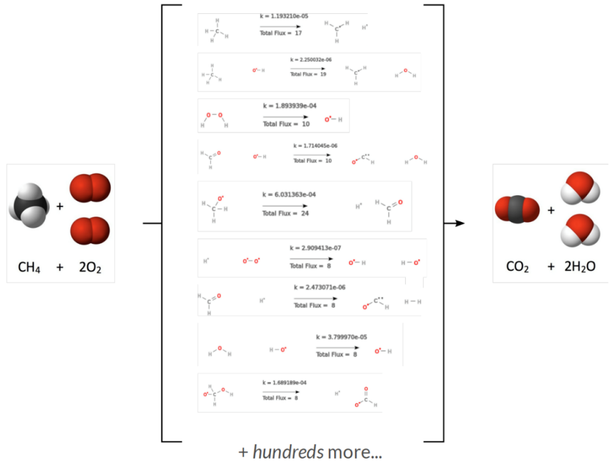


- New parameters
  - ADF 2013: 17 sets, 19 elements
  - ADF 2014: 38 sets, 29 elements
  - ADF 2016: 58 sets, 38 elements
  - ADF 2017: 79 sets, 38 elements
  - AMS2018: 81 sets, 40 elements + Ho/El
  - van Duin, Goddard, others
  - RxFF consulting
  - MCFF & CMA-ES parameterization

ReaxFF tools in Amsterdam Modeling Suite

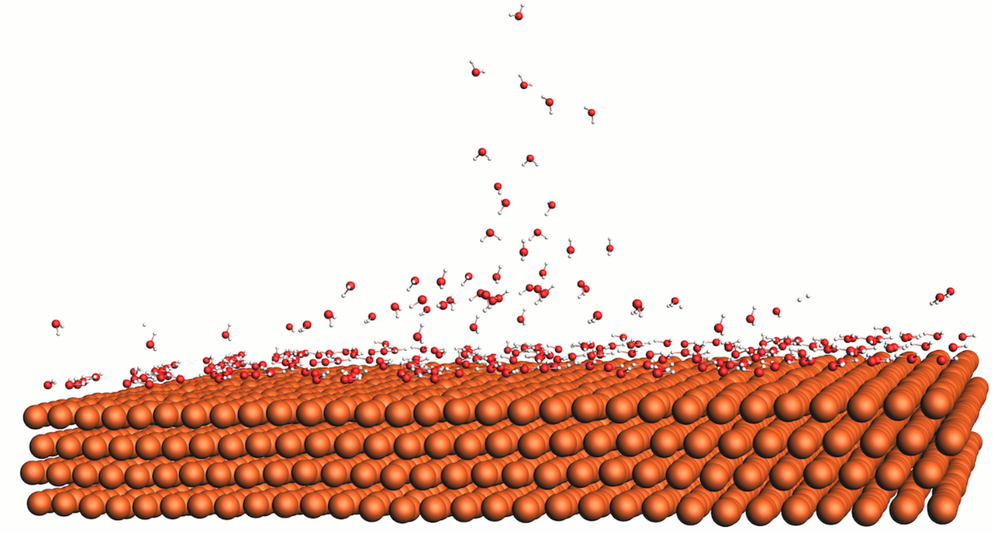
ChemTraYzer: Automated rates & pathways
New in 2018: Analyze surface reactions





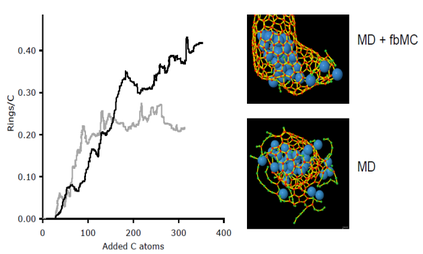
Tools += complete reaction networks
elementary reactions, rate constants, fluxes, timeline

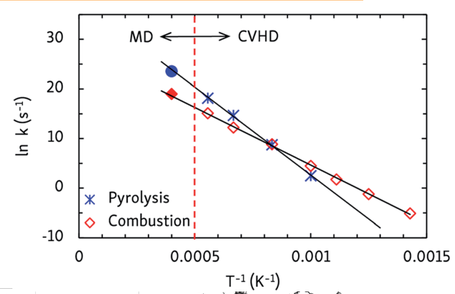
Molecule gun: depositing molecules on surfaces

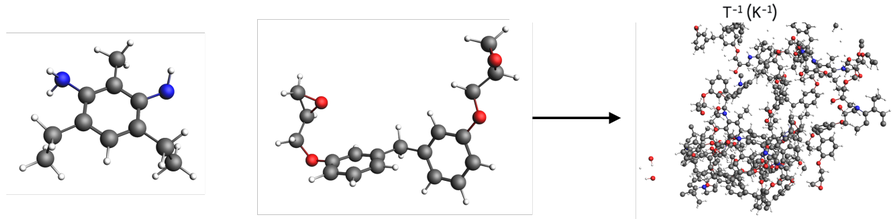


MCFF & CMA-ES parameterize ReaxFF force field

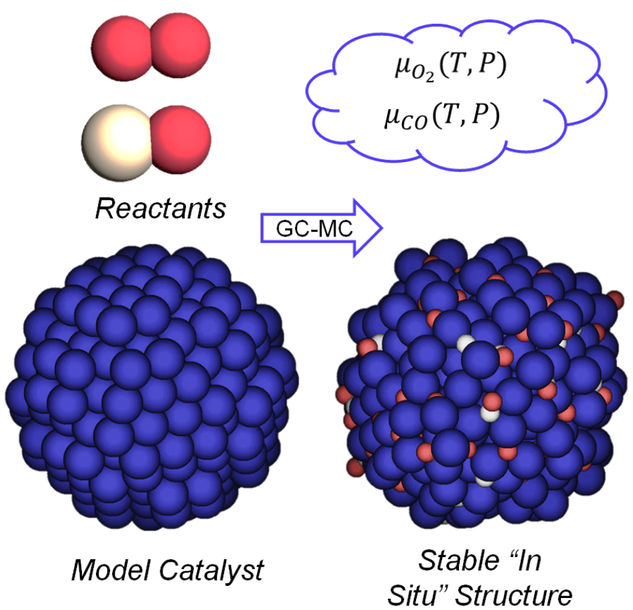
fbMC, CVHD: speed up kinetics



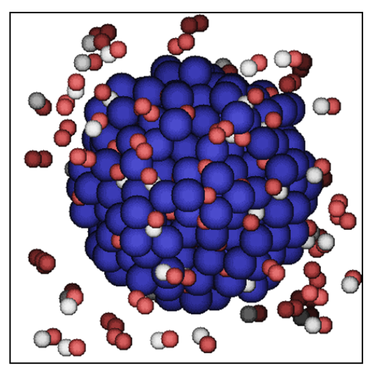




bond boost build polymers



GCMC: speed up thermo



ReaxFF: some tips

- First equilibrate your system before your production run
  - Usually thermostats equilibrate within some dozen ps, barostats take longer
  - For your equilibration, save less frames (Output frequency in Details-> MD)
  - If equilibrated, restart (Details->Restart) or just copy-paste last geometry
  - Check if your system properly equilibrates => damping constants
    - For Berendsen barostat, use a high damping (e.g. 2500fs) , thermostat ~100 fs is OK
    - For NHC thermo/barostat check oscillation and adapt tau (see also manual)

- After importing a structure (cif, database, ..): **relax the system**
  - Geometry optimization with loose criteria
  - OR run a few ps NpT trajectory with a 0.05fs time step at 5K and 0 pressure

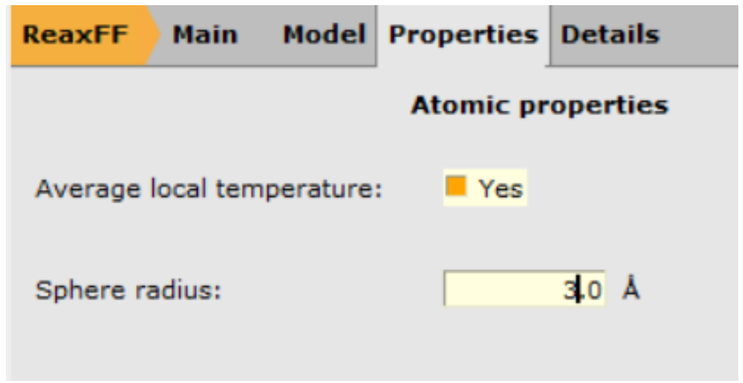
- Avoid having lattice vectors < 10 Å

- See our FAQs, e.g on ReaxFF force field availability / suitability
- Contact support@scm.com

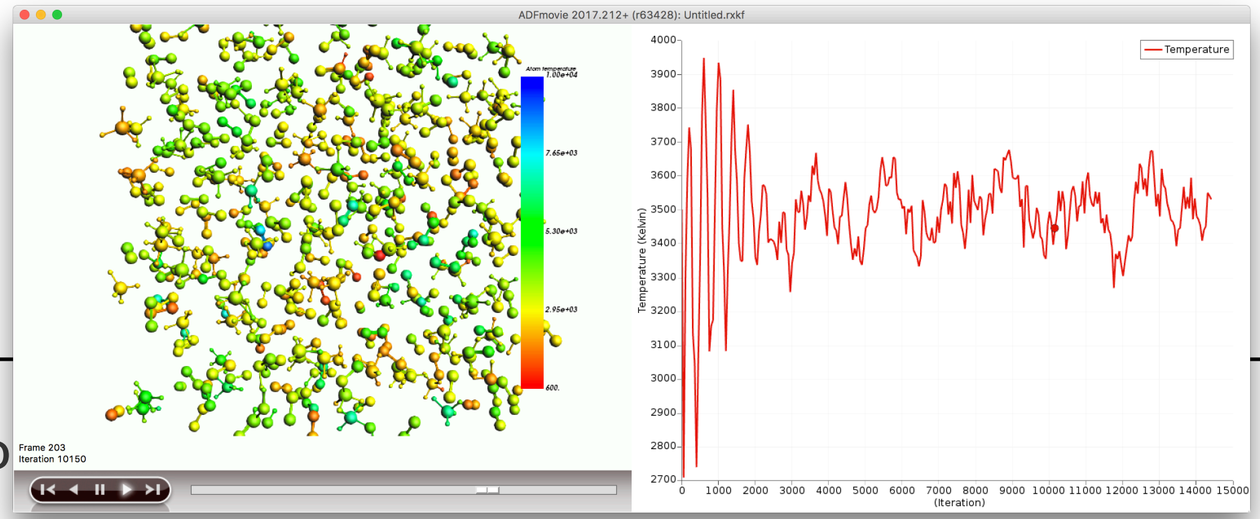
ReaxFF new: local T

Try at home - Exercise 14: seeing local temperature variations

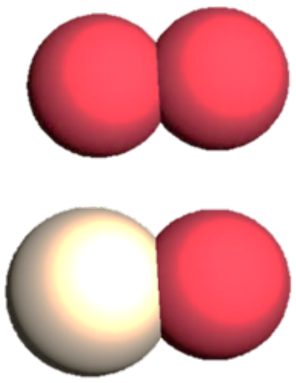
- Start with the methane burning tutorial
  (or with the molecule gun; see also news item with tutorial on visualizing local T)
  - Also local documentation! (Search in GUI)
  - Perhaps reduce the # of steps
- To visualize local temperatures:
  - Properties -> Atomic Properties
  - Tick Average local temperature
  - Set the awareness radius to `3.0 Å`



  - Local T in ADMovie: view -> color atoms by -> Local T
    - Change the axes and the color coding after double-clicking the legend
    - Also do the **ChemTraYzer analysis**



Grand Canonical Monte Carlo + ReaxFF



$\mu_{O_2}(T,P)$

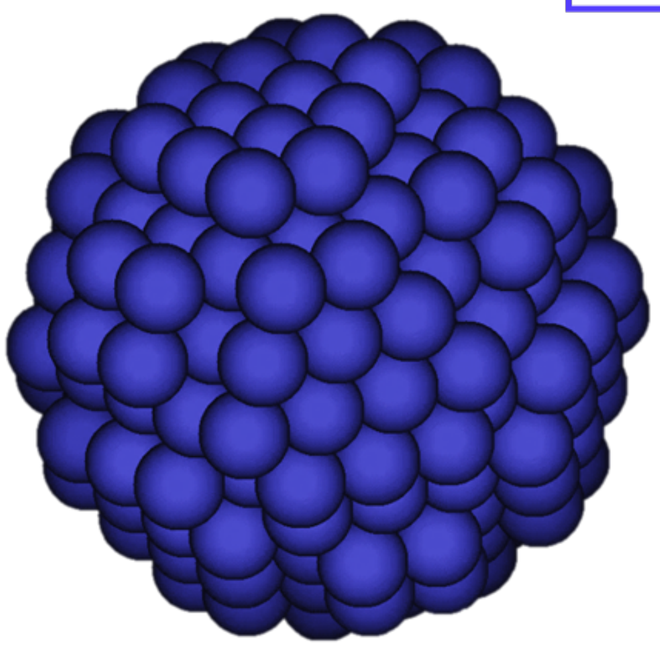
$\mu_{CO}(T,P)$

Reactive MD under actual conditions

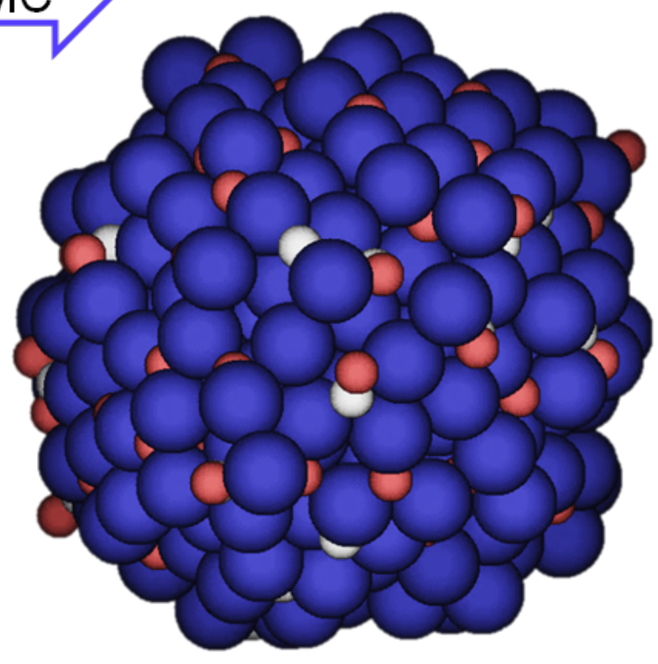
GCMC: What is stable (p,T)?
ReaxFF: What is active?

Reactants

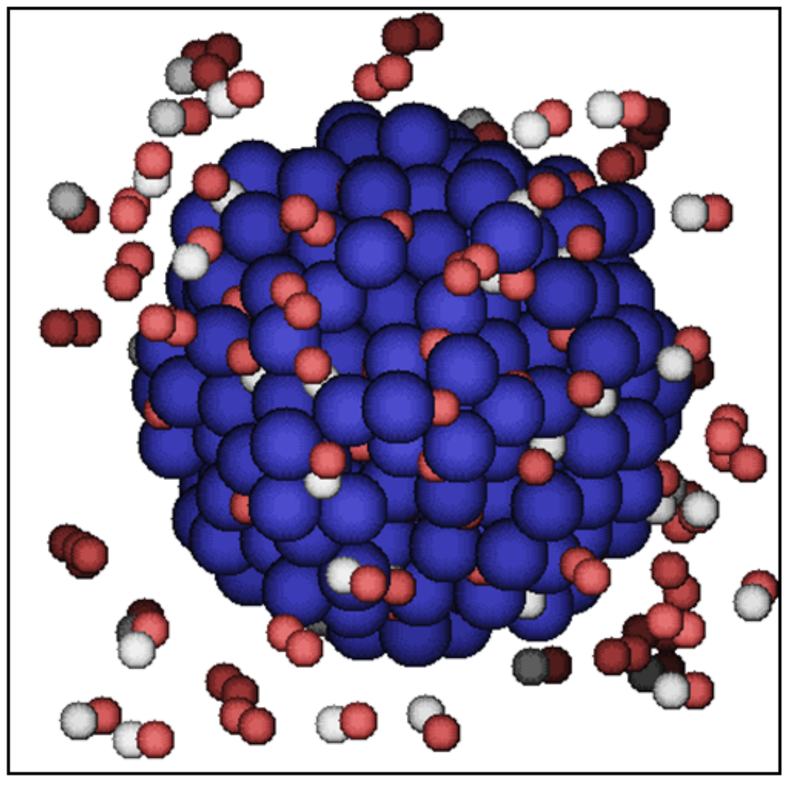
GC-MC



Model Catalyst



Stable "In Situ" Structure



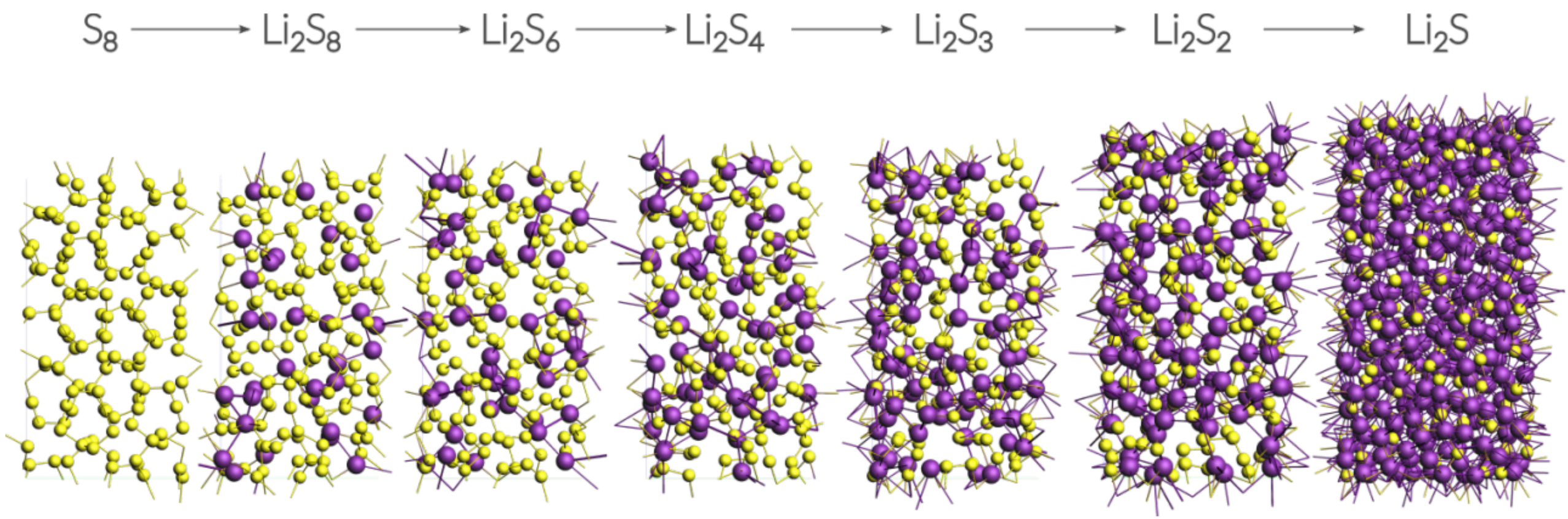
Kinetics

ReaxFF +GCMC **tutorial**

T.P. Senftle, R.J. Meyer, M.J. Janik, A.C.T. van Duin, J. Chem. Phys., **139** (2013) 044109

ReaxFF: GCMC battery voltages

At home: Exercise 15: advanced GCMC battery discharge tutorial inserting Li in S

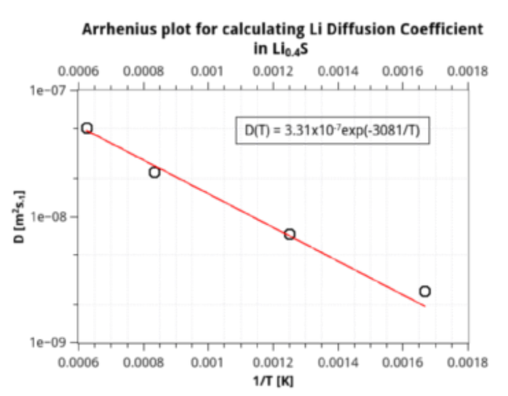


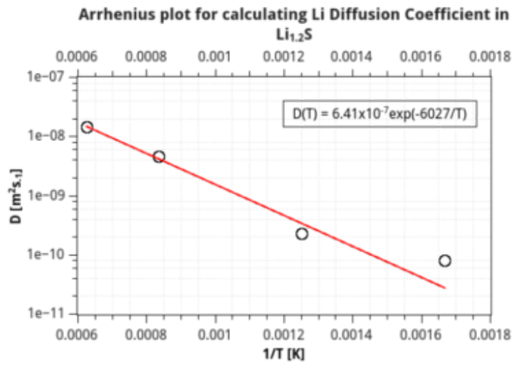
Based on: M. M. Islam, A. Ostadhossein, O. Borodin, A. T. Yeates, W. W. Tipton, R. G. Hennig, N. Kumar, and A. C. T. van Duin, ***ReaxFF molecular dynamics simulations on lithiated sulfur cathode materials***, Phys. Chem. Chem. Phys. **17**, 3383-3393 (2015)

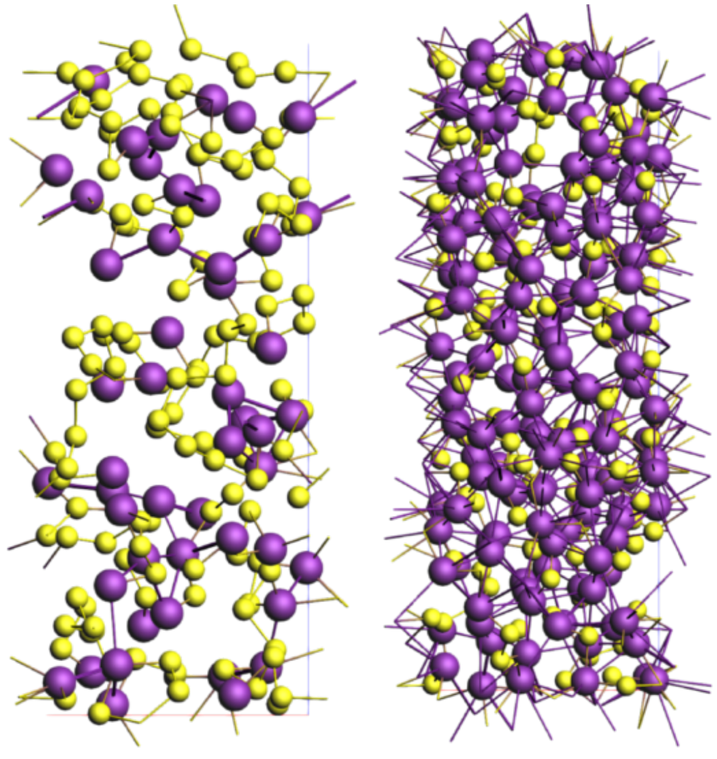
ReaxFF: Li diffusion in batteries

At home: Exercise 16: advanced tutorial Li diffusion (same S_8 coords as exc. 15)

- A more robust way to make the $Li_{1.2}S$ system (step In step 2.2):
  - Take the optimized Li0.4S system, change the lattice c-vector to 38 while ticking 'Adjust atoms'
  - Delete all Li atoms (select one, then select atoms of some type)
  - Now use Packmol ('Builder) to add 154 Li atoms… & do a longer NpT run (!5000 steps? – check Energy)
  - Or even better: take one of the GCMC structures from Exercise 15
  - To properly anneal – a slow cooling rate is recommended (rather than the fast one used here)
  - Check p, T and E during NpT equilibration – stop when it looks 'flat' and 'Update Geometry'



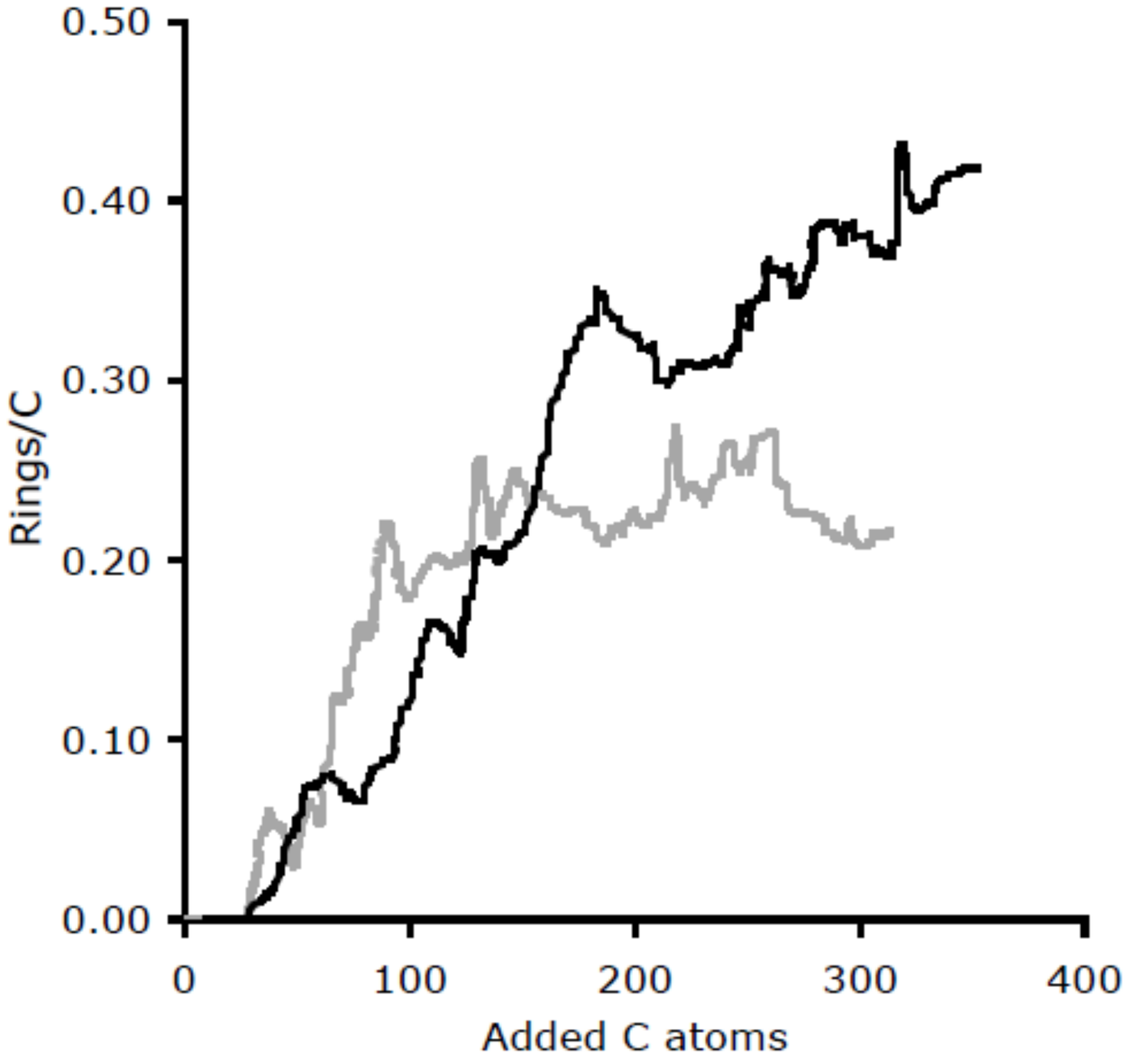


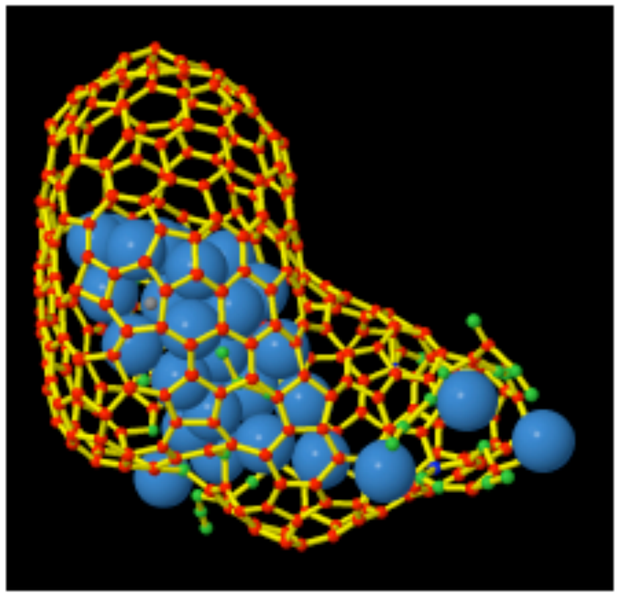


$Li_{0.4}S$ $Li_{1.2}S$

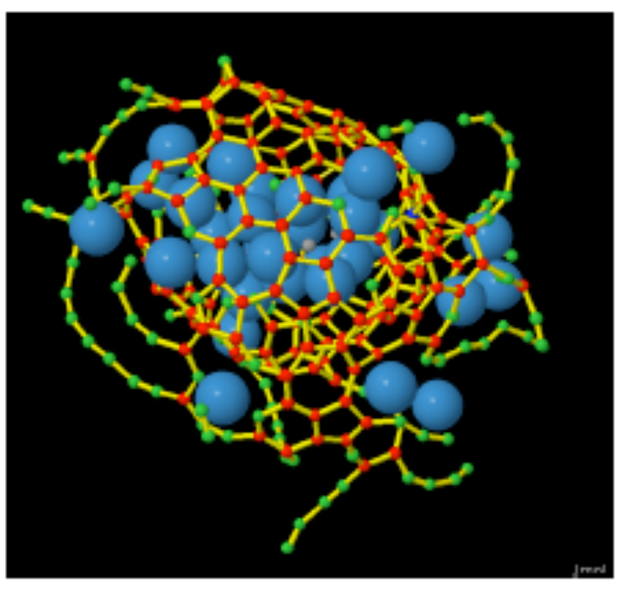
S
Li

Carbon nanotube formation: accelerate MD





MD + fbMC



MD

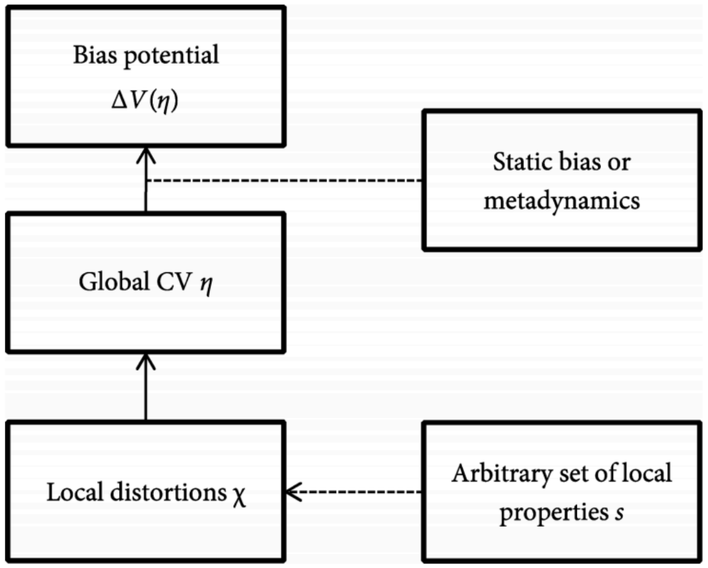
[Molecule gun + fbMC video]

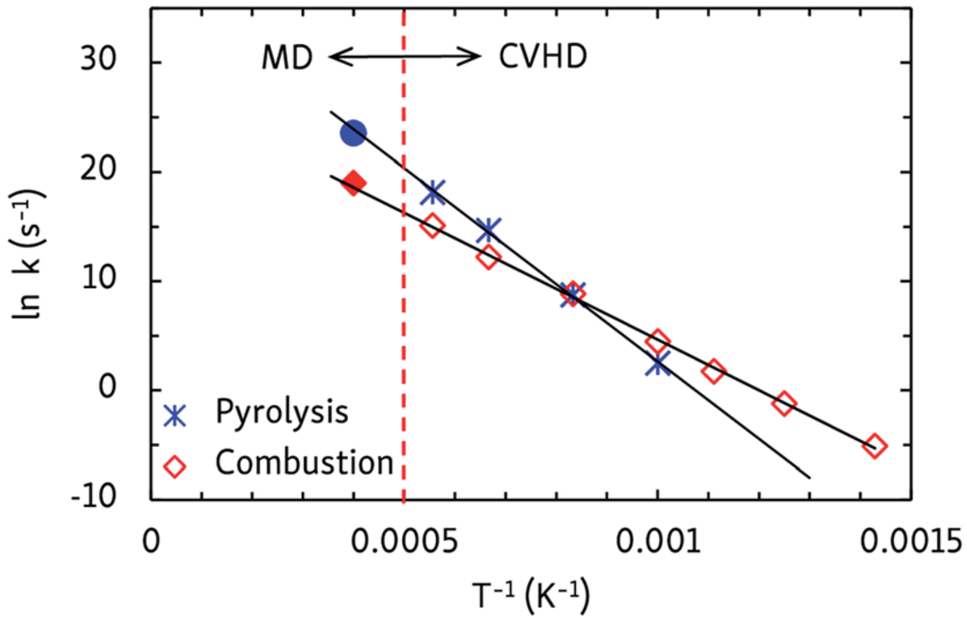
Neyts *et al.*, J. Am. Chem. Soc. **133**, 17225 (2011)

Further accelerating ReaxFF

Collective-Variable driven Hyperdynamics (CVHD)

- Metadynamics meets hyperdynamics; learn CV 'on-the-fly'
- Boosts: 10^6 – 10^9 for pyrolysis / combustion



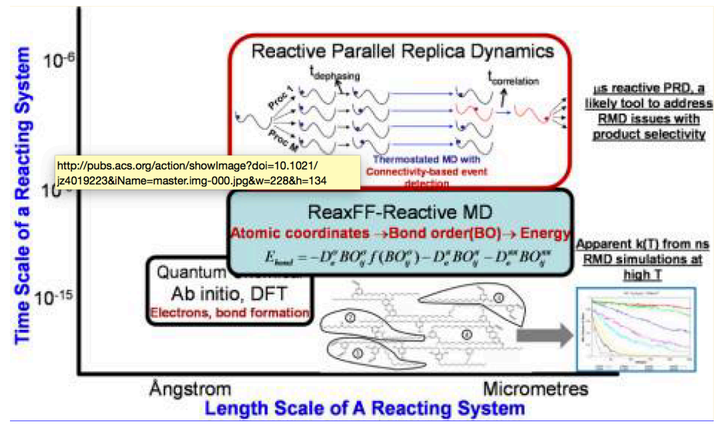


Next week: **tutorial**

Bal & Neyts: J. Chem. Theory Comput. **11**, 4545 (2015); Chem. Sci., **7**, 5280 (2016)

Under development:

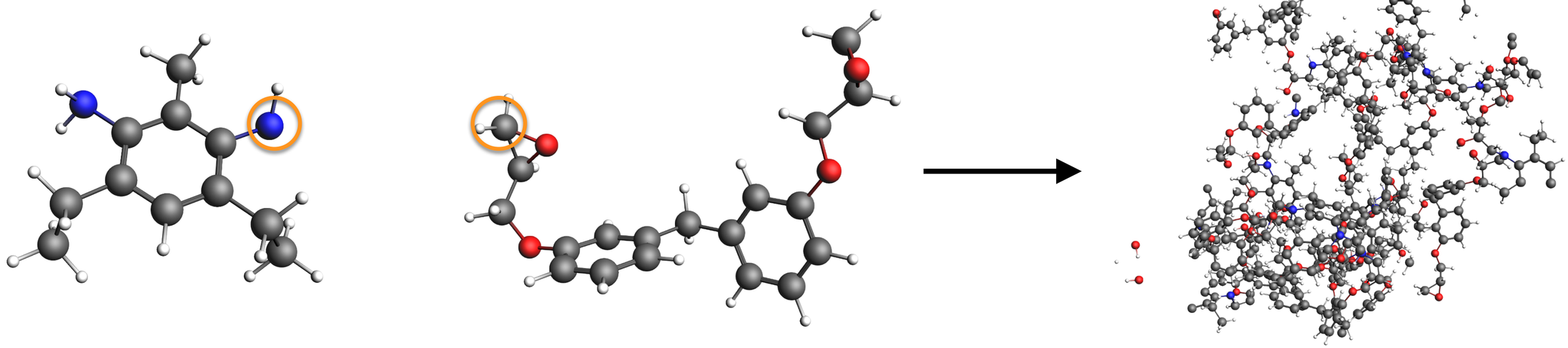
- Multi-replica methods (T-REMD, PRD)
  - Easy set-up, no bias (NB: PRD needs pre-defined event)
- (Path sampling approaches: TPS, TIS, ….)

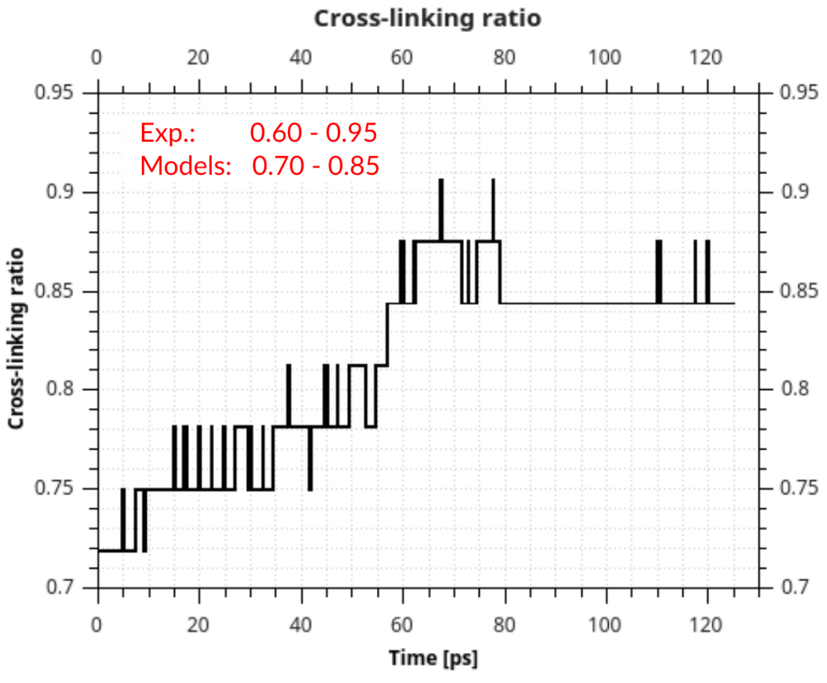


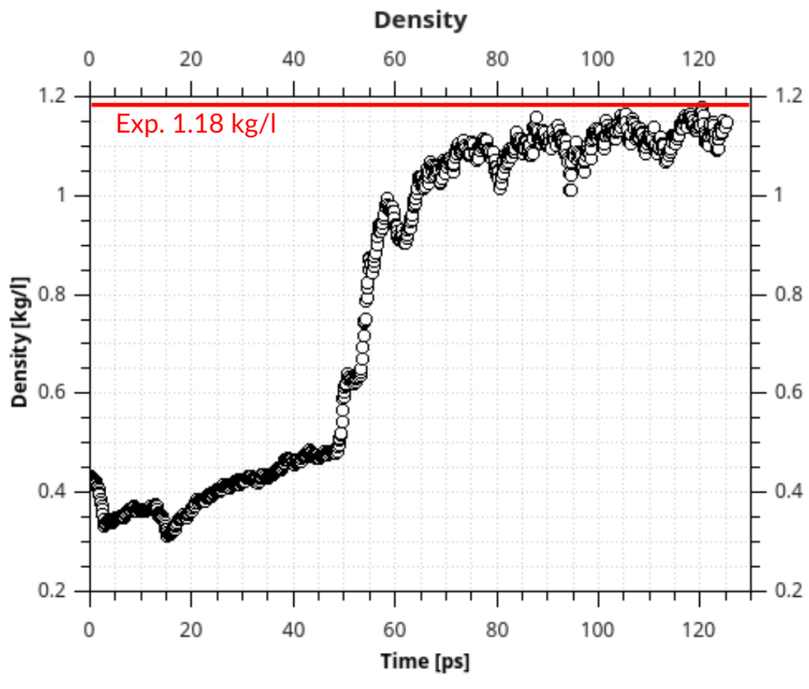
J. Phys. Chem. Lett. **4**, 3792 (2013)

Further accelerating ReaxFF

- ‘Bond boost’ (van Duin)
  - Cross-link polymers (irreversible) – can be extended to other reactions
  - Pre-defined reaction sites, associated rate?



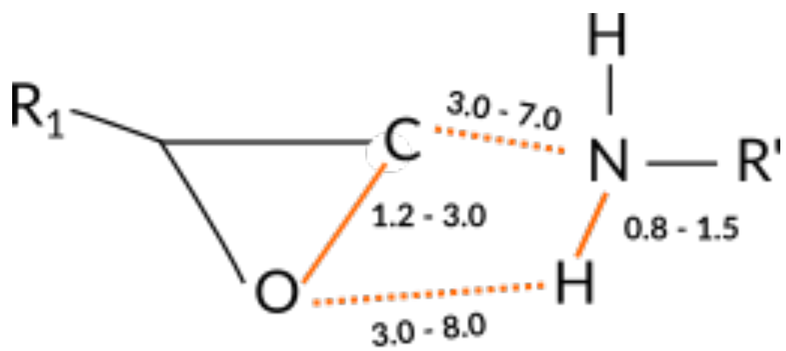




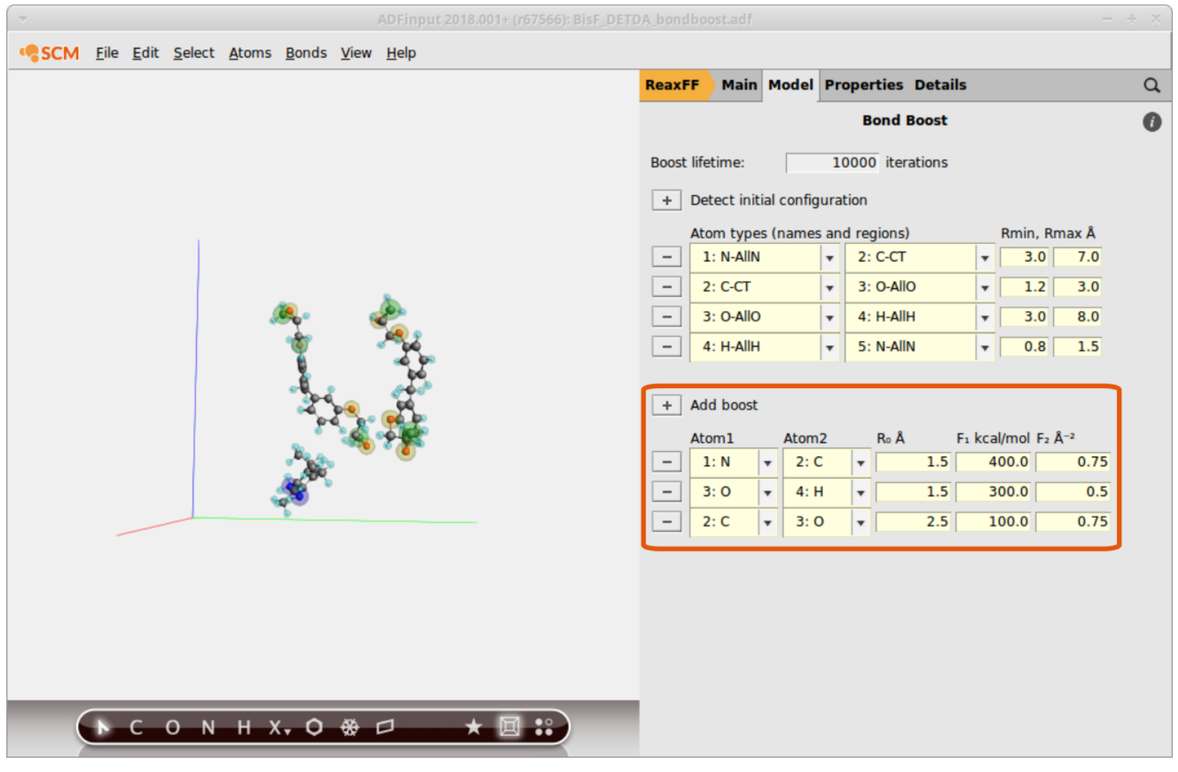
ReaxFF: bond boost

Exercise 17: start with bond boost tutorial

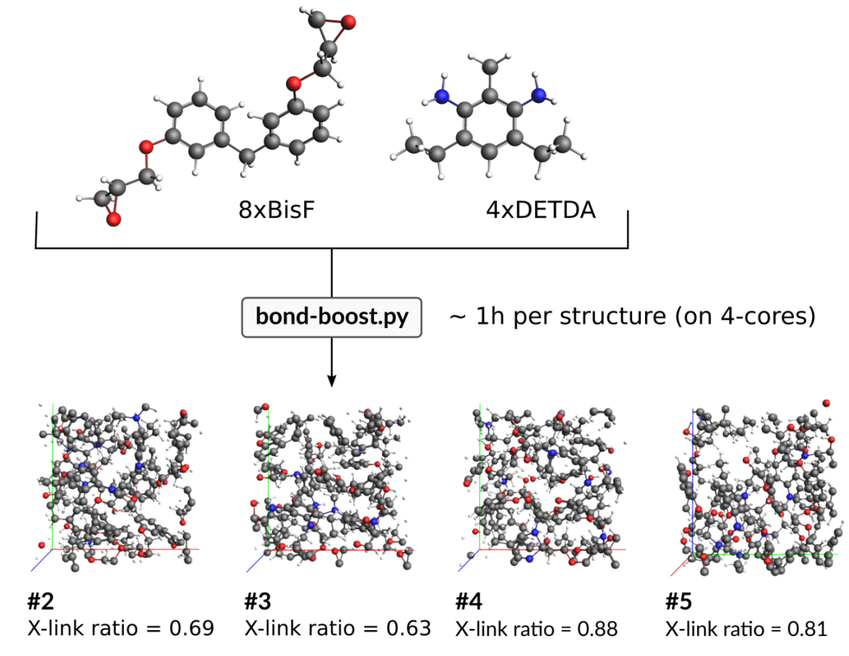
- Add boost potential when reactions are ‘close to reaction’, but not over barrier!
- Track bond distances of certain atoms
- For epoxy polymerization this looks like
- For small systems it’s easy to set up in GUI
- For bigger systems use python scripting
- Also check out mechanical polymer properties



preliminary complex

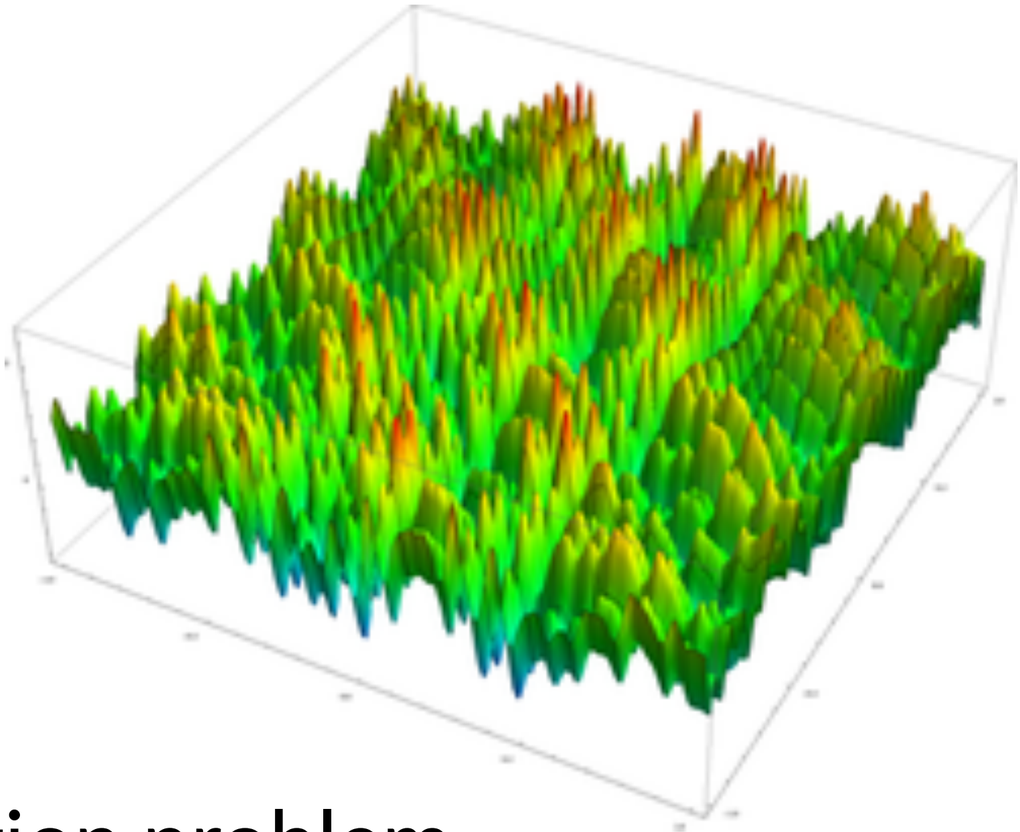


van Duin et al., J. Phys. Chem. A (2018)



Optimizing ReaxFF parameters

- Parameters are
  - interdependent
  - non-linear
  - many
  - not always physically interpretable

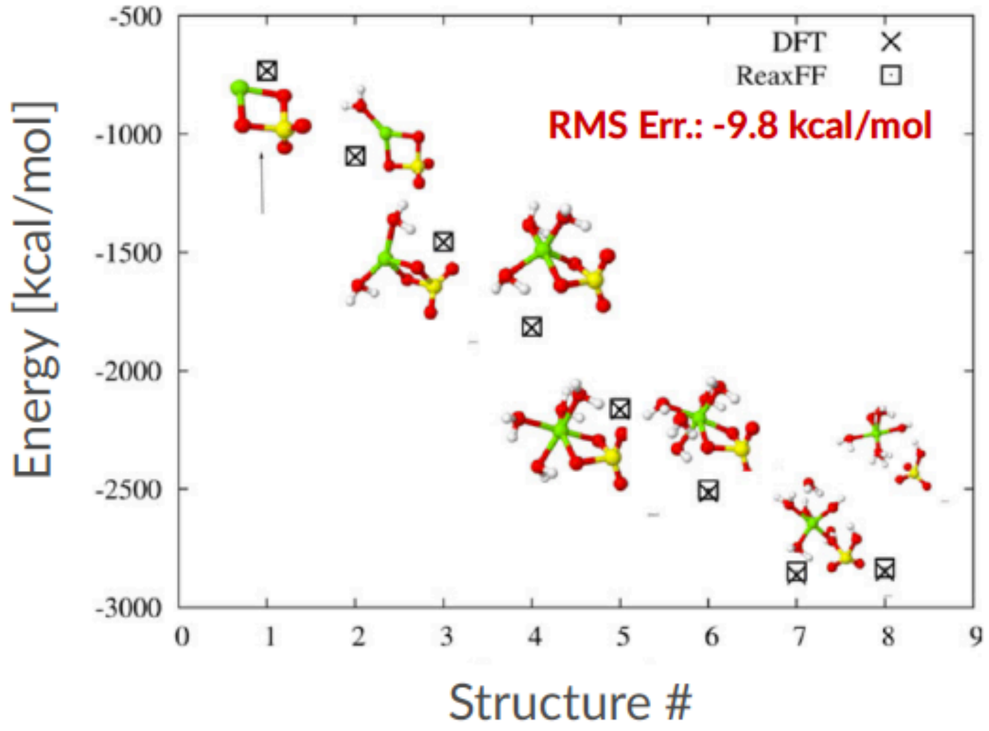


- Highly complex global optimization problem

Monte Carlo with Simulated Annealing

$$Error = \sum_{i=1}^{n} \left[\frac{(x_{i,TS} - x_{i,ReaxFF})}{\sigma_i} \right]^2$$

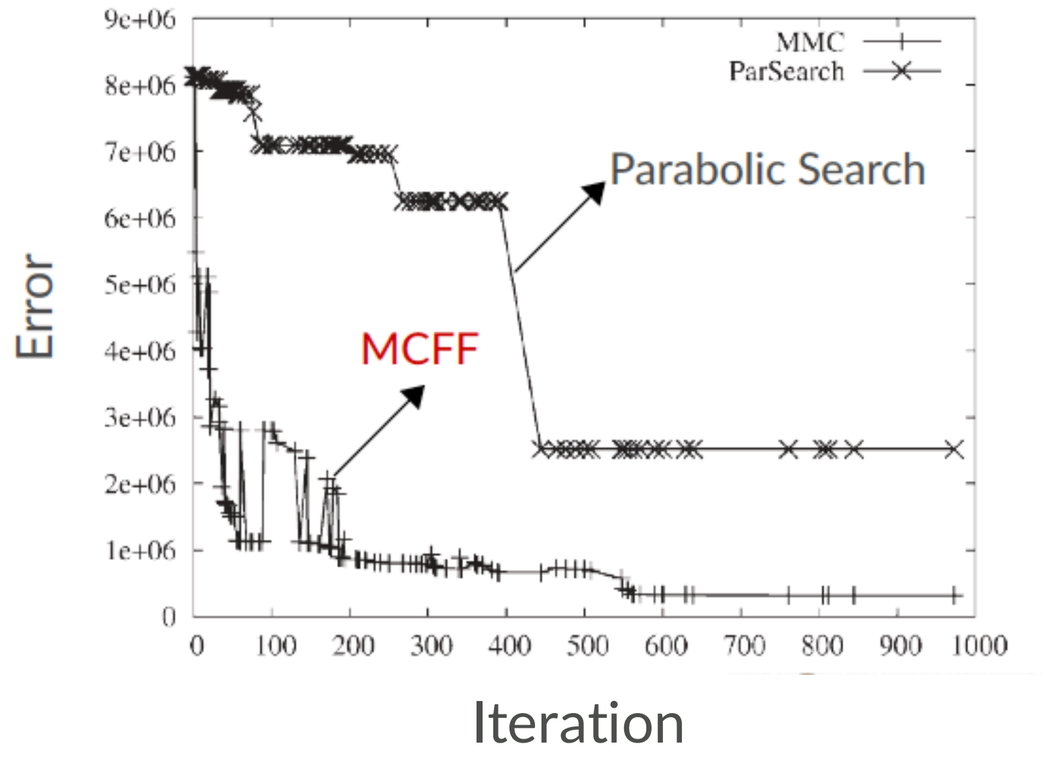
$x_{i,TS}$ = reference
$x_{i,ReaxFF}$ = current ReaxFF
σ_i = weight



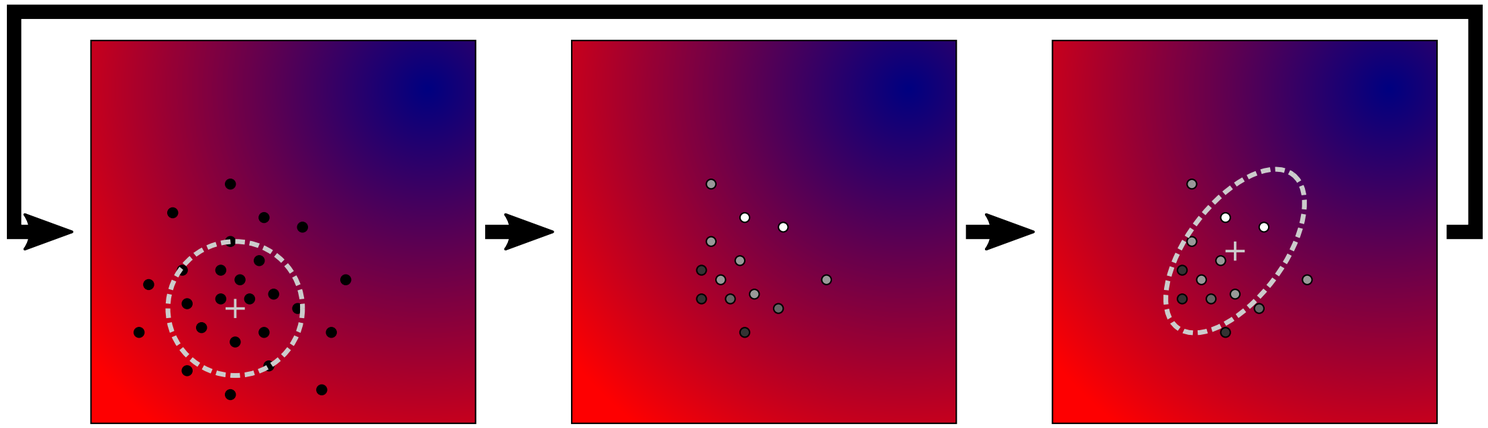
References include:

- Atomic charges
- Heat of formation
- Geometries (Bond/Angle/Dihedrals)
- Cell parameters
- Energies

J. Comp. Chem. **34**, 1143-1154 (2013)



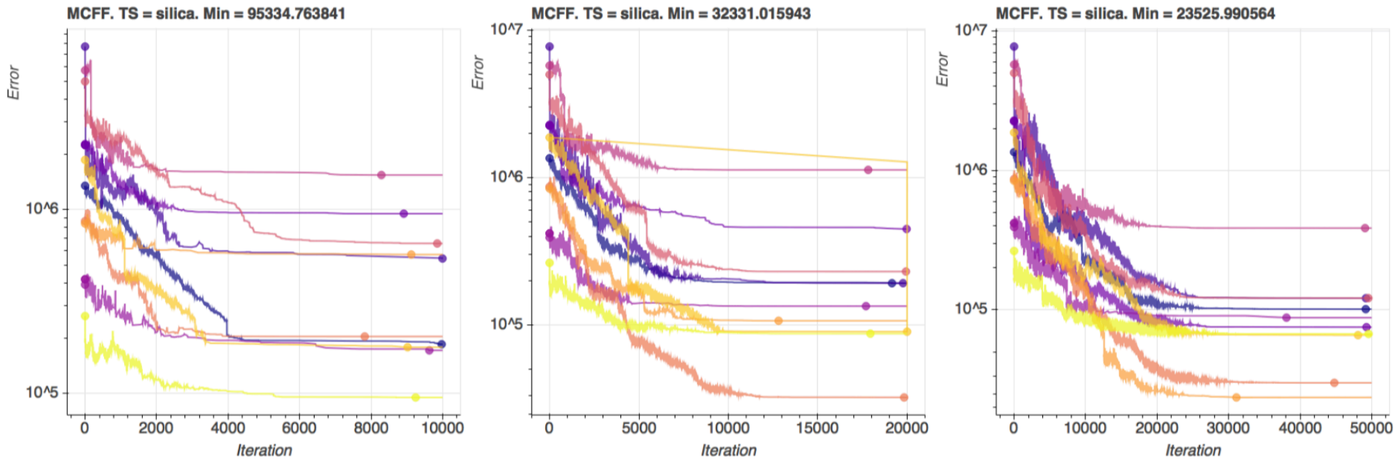
Covariance Matrix Adaptation Evolutionary Strategy



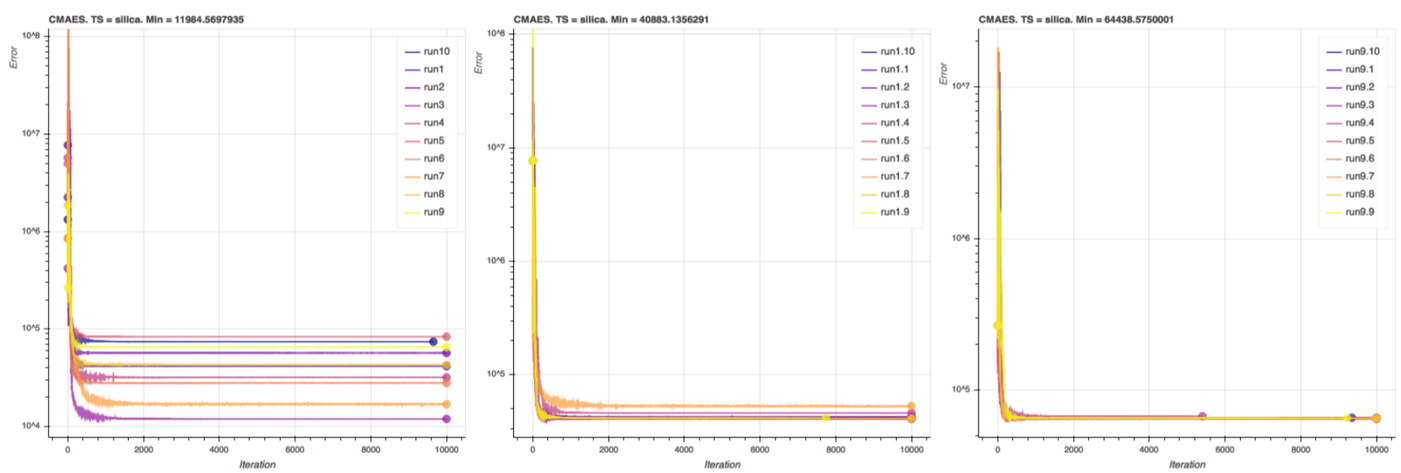
Generate points using multivariate Gaussian distribution with covariance matrix C

Order the points by their objective function value
Assign weights, decreasing from the best point to the median

Set new distribution center to weighted average of successful points
Update covariance matrix



MCFF

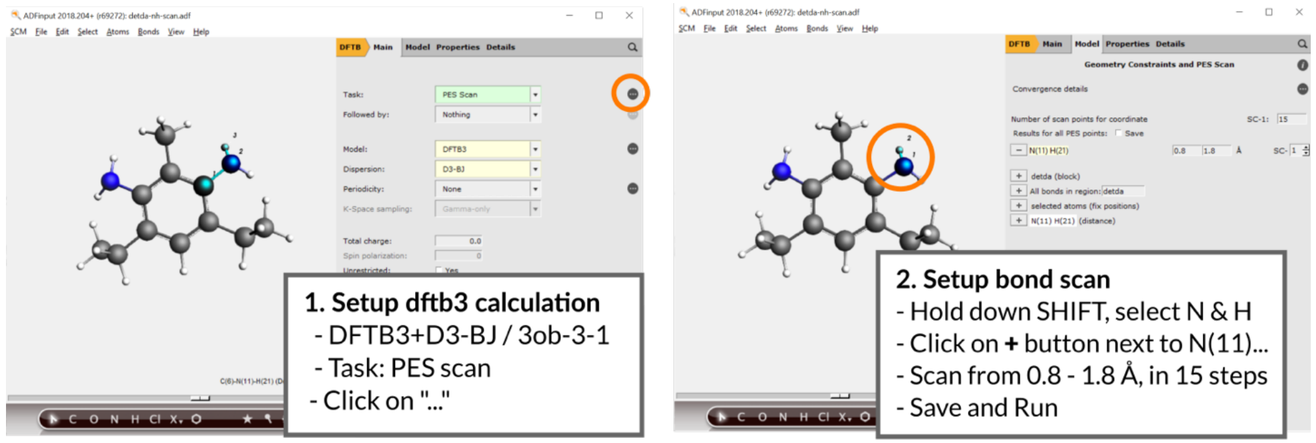


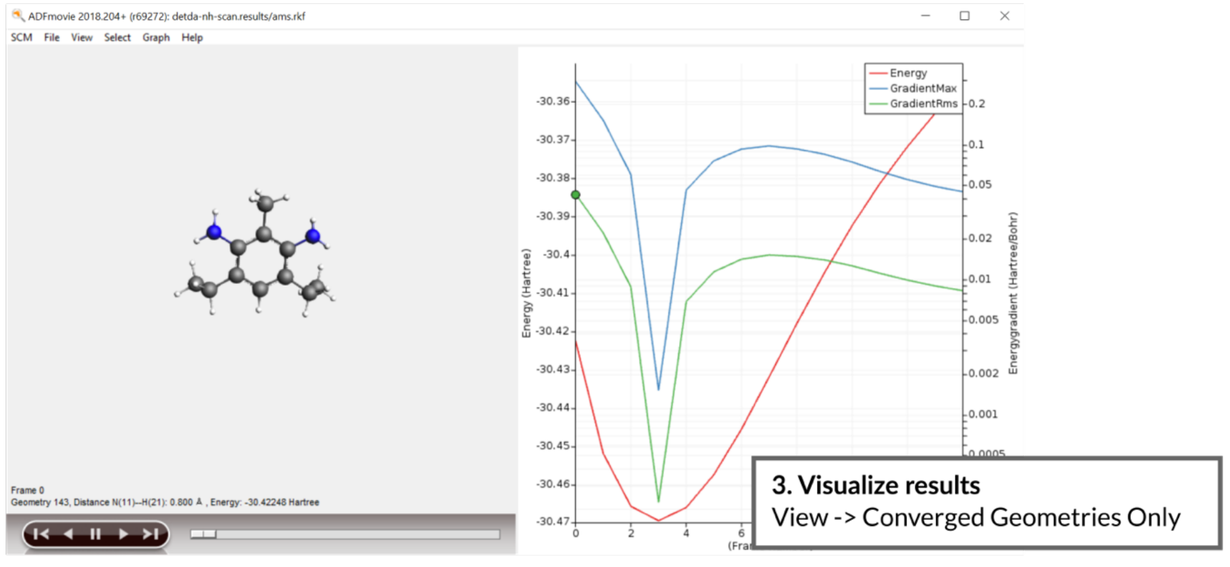
CMA-ES
(Shchygol et al. arxiv)

ReaxFF: reparameterization

Exercise 18: refine ReaxFF parameters

- Follow the tutorial package to build your training set (trainset.in & geo)
  - Take care with adding the geometries into one file and editing the trainset
  - A lot of work, some scripting
- Add geometries
- Add conformers
- Add bond scans
- Run CMA-ES optimization
- Test errors
- Try to further refine
- See also: Co training set





Some notes on (Windows) scripting

Use help -> command-line and type sh to go to a Windows shell with ADF environment variables set.

```
C:\ADF_DATA\CMA-ES-FIT\xyz>set ADFHOME=C:/ADF2018.105/

C:\ADF_DATA\CMA-ES-FIT\xyz>set ADFHOME_=C:\ADF2018.105\

C:\ADF_DATA\CMA-ES-FIT\xyz>set ADFBIN=C:/ADF2018.105//bin

C:\ADF_DATA\CMA-ES-FIT\xyz>set ADFRESOURCES=C:/ADF2018.105//atomicdata

C:\ADF_DATA\CMA-ES-FIT\xyz>set PATH=C:\ADF2018.105\\msys\usr\bin;C:\ADF2018.105\\msys\bin;C:\ADF2018.105\\bin;C:\ADF2018
.105\\bin\tcltk\bin;C:\ADF2018.105\\bin\intelmpi\bin;C:\ADF2018.105\\bin\Putty;C:\ADF2018.105\msys\usr\bin;C:\ADF2018.10
5\bin\intelmpi\bin;C:\ADF2018.105\bin\Putty;C:\ADF2018.105\bin\TclTk\bin;C:\ADF2018.105\bin\openbabel;C:\ADF2018.105\bin
;C:\windows\system32;C:\windows;C:\windows\System32\Wbem;C:\windows\System32\WindowsPowerShell\v1.0\;C:\windows\System32
\OpenSSH\;C:\Program Files (x86)\Intel\Intel(R) Management Engine Components\DAL;C:\Program Files\Intel\Intel(R) Managem
ent Engine Components\DAL;C:\Program Files (x86)\Intel\Intel(R) Management Engine Components\IPT;C:\Program Files\Intel\
Intel(R) Management Engine Components\IPT;C:\Program Files\Intel\WiFi\bin\;C:\Program Files\Common Files\Intel\WirelessC
ommon\;C:\Users\Fedor\AppData\Local\Microsoft\WindowsApps;

C:\ADF_DATA\CMA-ES-FIT\xyz>cmd
Microsoft Windows [Version 10.0.17134.345]
(c) 2018 Microsoft Corporation. Alle rechten voorbehouden.

C:\ADF_DATA\CMA-ES-FIT\xyz>sh
sh-4.3$
```

We now have a basic shell in which can do so some scripting

Functions we will use a lot: cat, ls, pwd, various commands inside $ADFBIN

Some scripting examples

cat dog	output contents of the file named dog to the screen
cat file > file2	output contents of file to a new file, file2
cat file2 >> file3	output file2 and **append** to file3
cat *.bgf > geo	output all files ending in .bgf to a new geo file
cat geo >> ../geo	output geo to the file geo in the directory below
pwd	show in which directory we are
cd dog.results	go one directory up to dog.results
cd ..	go one directory down
cd -	go to the directory you were in before
ls	show which files are in this directory
ls -ltra	show files in directory with more details, order to time

Some other useful shell tips

arrow up / down	scroll through previous commands the directory below
for; do; done	Loop. Example: for i in 1 2 3; do cat $i/geo >> geo; done (append 1/geo 2/geo and 3/geo to geo)
!?	reuse last argument. Example: ls geo cp !$ geo.1 (== cp geo geo.1)
<TAB>	Autocomplete. Example: you want to copy trainset.in cp tr<TAB> will search for all files here starting with ‘tr’ if only 1: it completes to that, otherwise prints a list

Acknowledgements

The workshop materials were prepared with the aid of EU funding:

The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 764810 (S4CE).