Hands-on workshop ReaxFF Acceleration & Parameterization with the Amsterdam Modeling Suite

Tokyo, 1+2 November
Fedor Goumans, goumans@scm.com SCM support: support@scm.com
Molsis: ms-support@molsis.co.jp

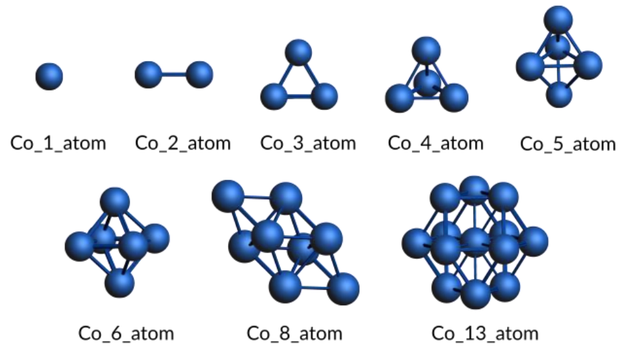
Making Computational Chemistry Work for You

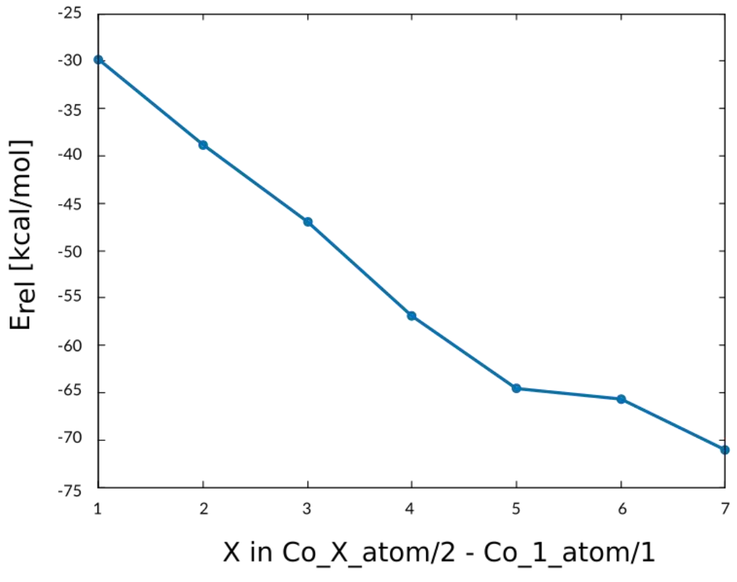
Program

- Introduction: ReaxFF
  - Some advanced exercises, new features

- Acceleration techniques
  - Idea, ongoing developments
  - **Day 1 hands-on**: CVHD, bond boost => polymer structures, fbMC

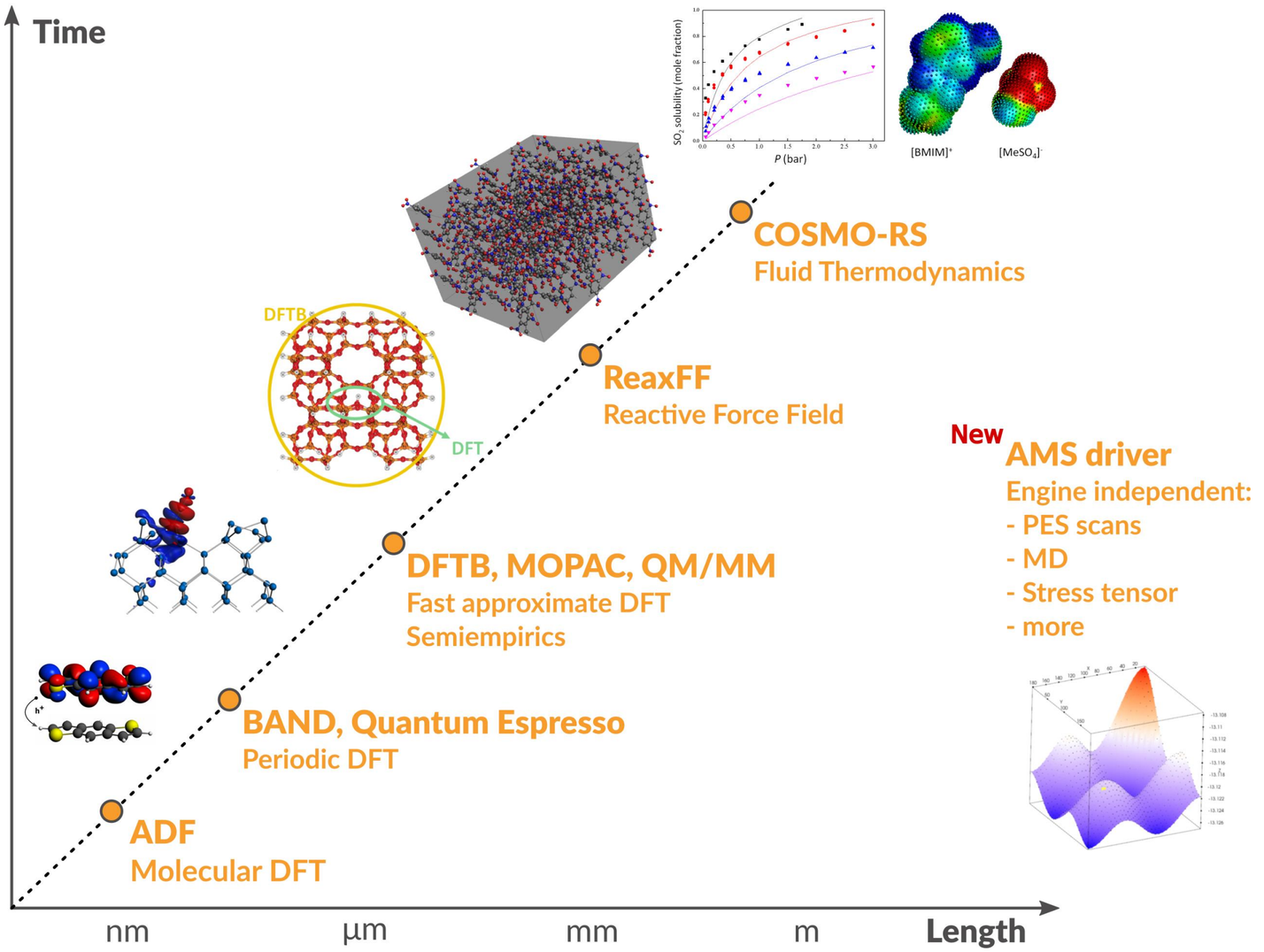
- **Day 2:** Refitting a force field for parameterization (slides, input files)
  - Scripting on Windows
  - Adding geometries, conformers, reaction pathways, bond scans
  - Optimizing the force field, checking errors & improving
  - Extending to materials, alloys





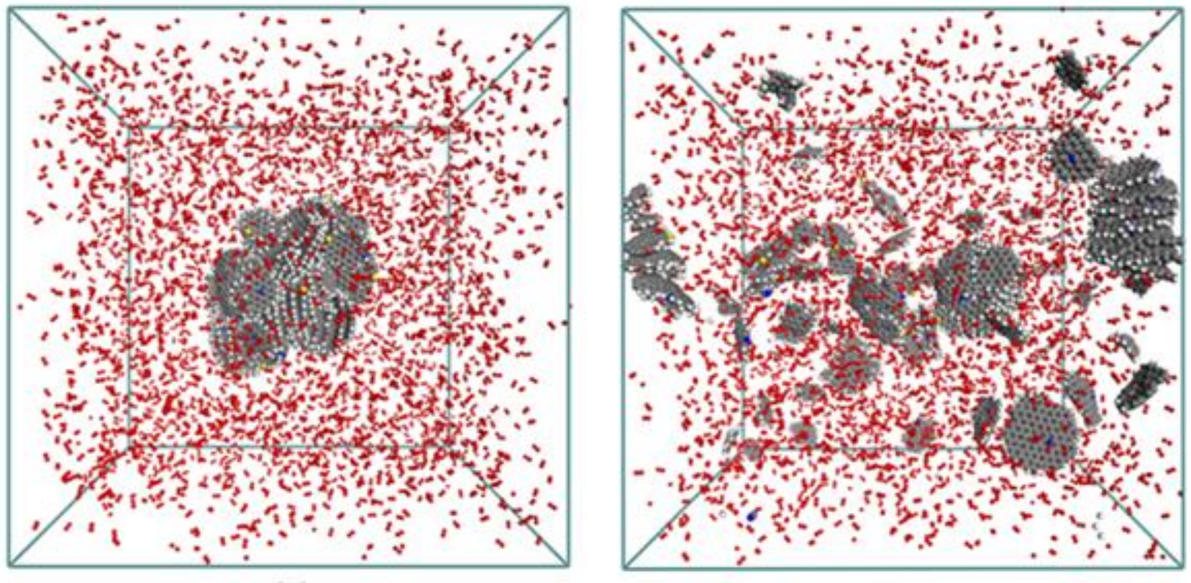
Amsterdam Modeling Suite

- ADF: powerful molecular DFT
  - Spectroscopy: NMR, EPR, VCD, UV, XAS
  - Advanced solvation / environments
- BAND: periodic DFT
  - (2D) Materials
  - (Orbital) analysis, spectroscopy
- Interface to Quantum ESPRESSO
- 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

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

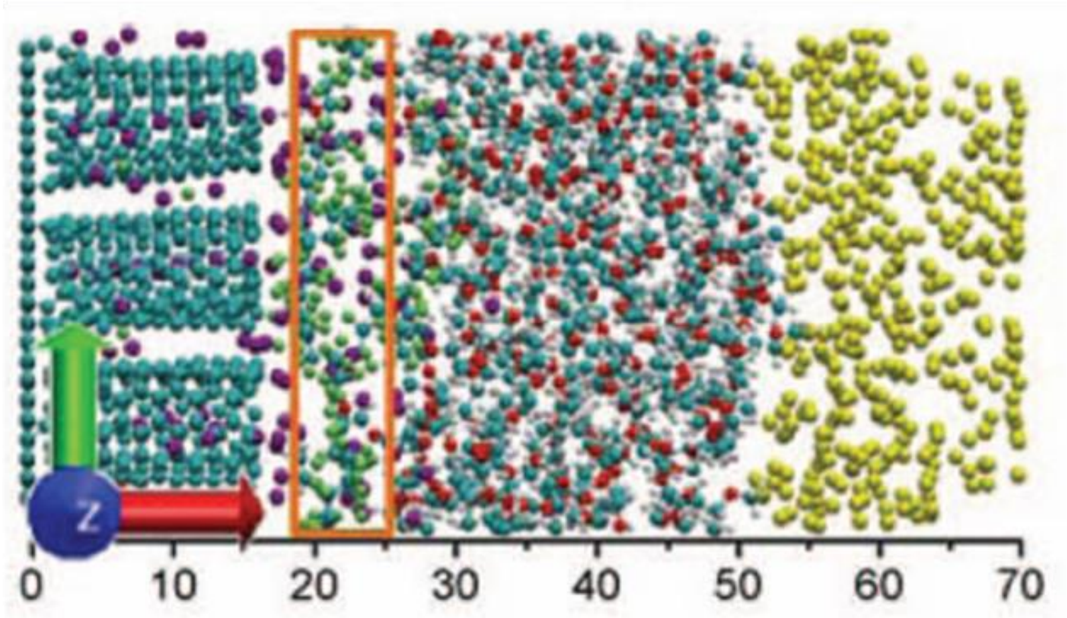


Molecular dynamics: Newton's equations of motion

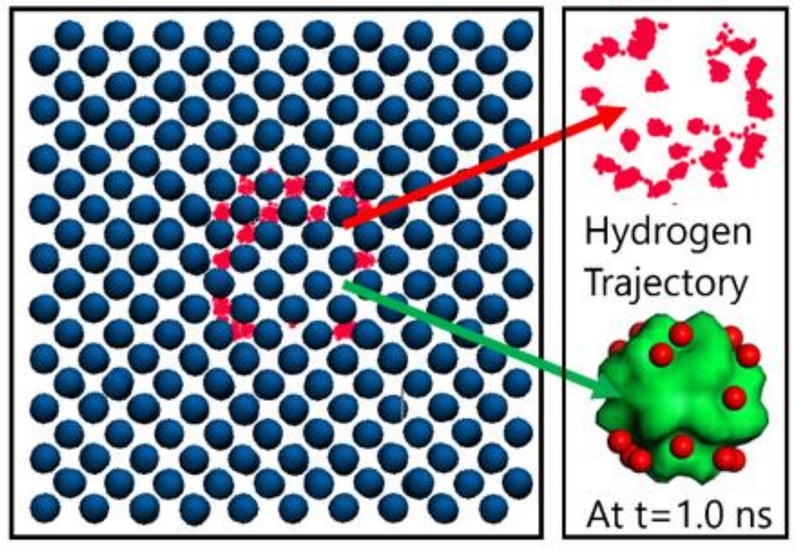
Movement of atoms: solve numerically + propagate
Forces can come from 'anywhere': potentials, QM code, …

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

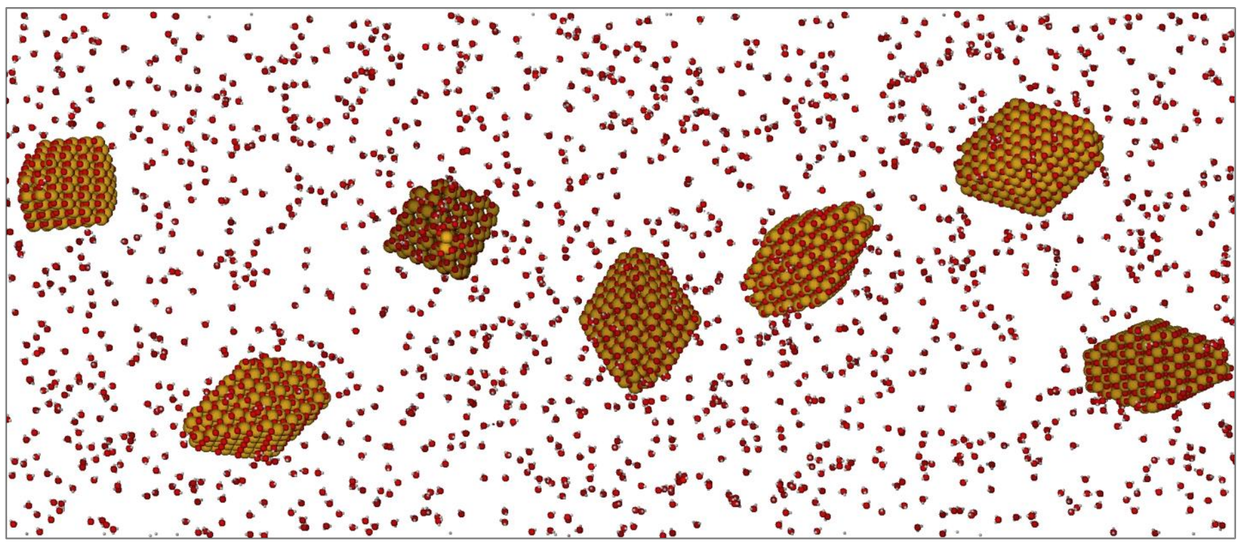
ReaxFF – reactive molecular dynamics



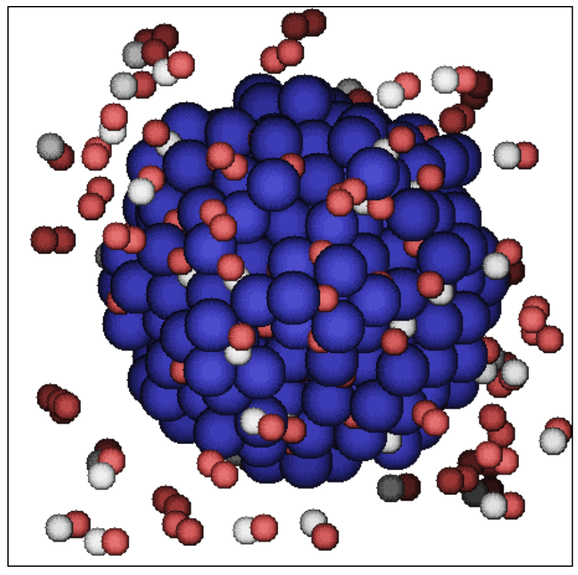
Li battery discharge: J. Electrochem. Soc. **161**, E3009 (2014); PCCP, **17**, 3383 (2015)



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.

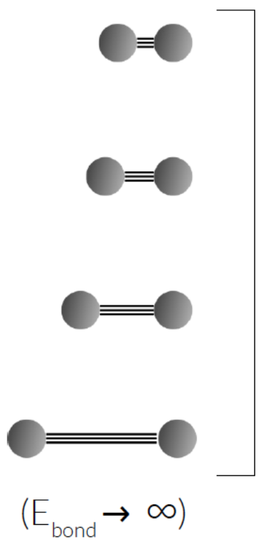
Standard forcefields **vs** *ReaxFF*

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

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

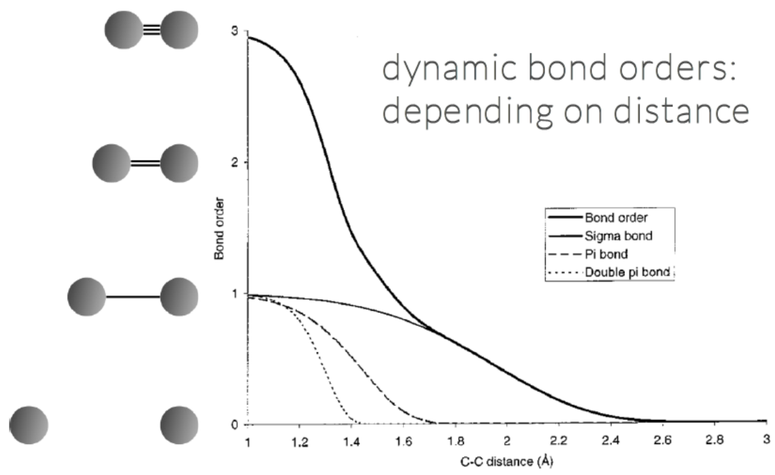
Non-harmonic potentials based on bond orders, bond breaking/forming possible, e.g.

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



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...

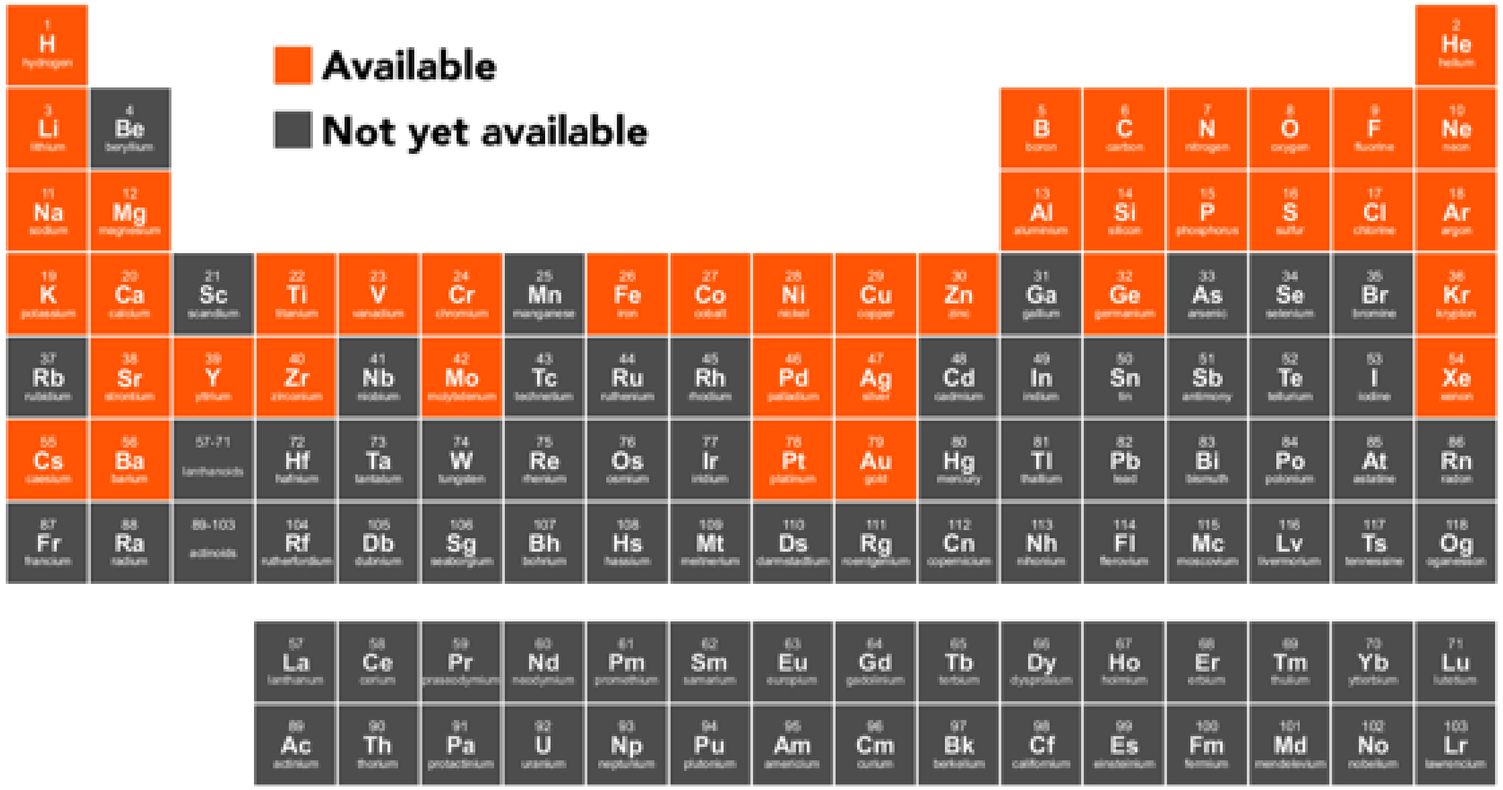
Parameters = 16

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

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

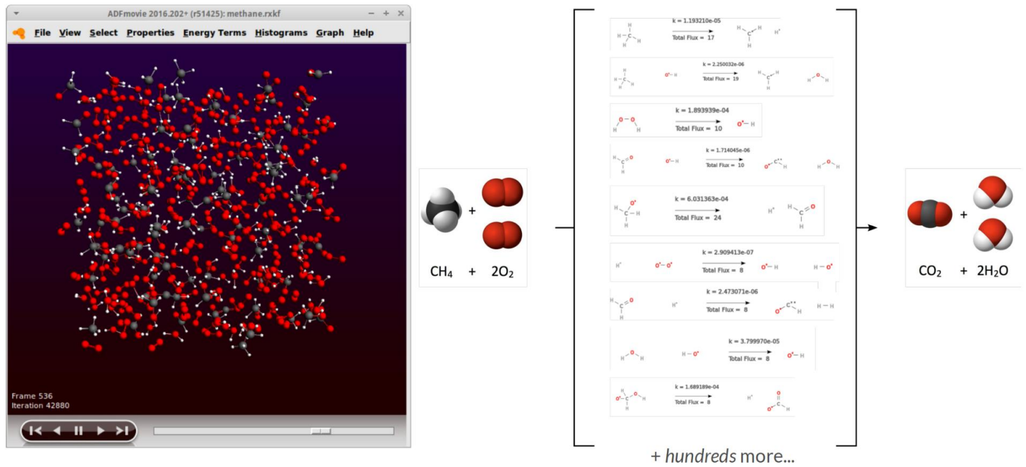
- many elements studied
- each pair needs bonded terms
- validate force field
  - GUI checks
  - training data crucial
  - application specific



- 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

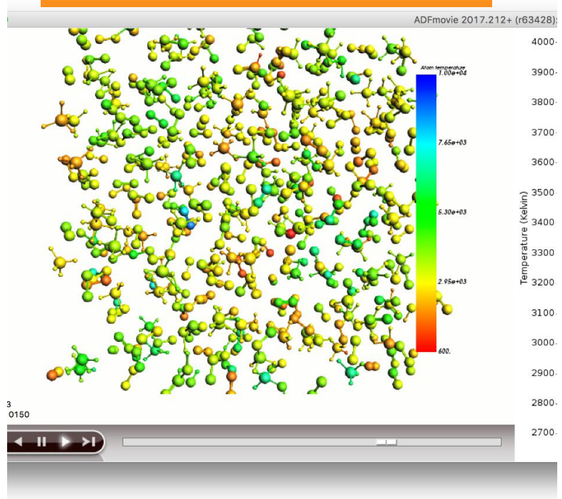
ReaxFF tools in Amsterdam Modeling Suite

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

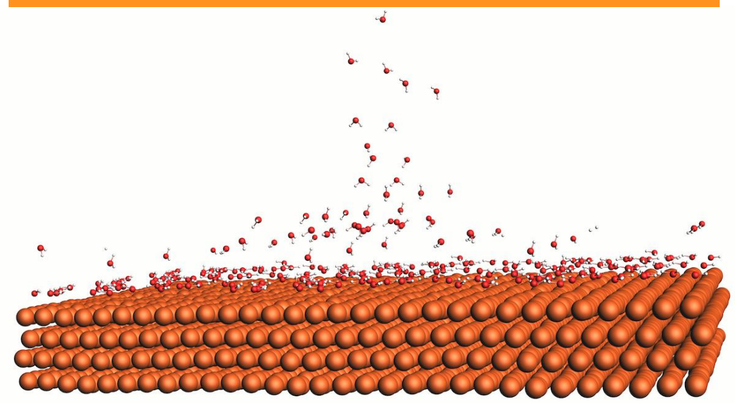


Tools += complete reaction networks
elementary reactions, rate constants, fluxes, timeline

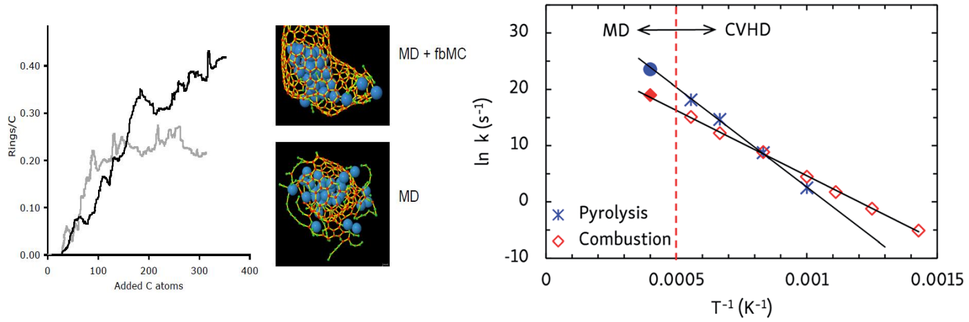
T-NEMD, local T:
heat transport

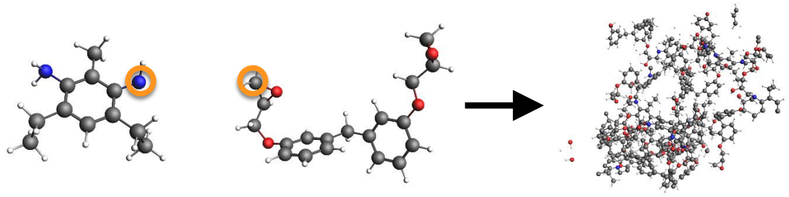


Molecule gun: depositing molecules on surfaces



fbMC, CVHD: speed up kinetics

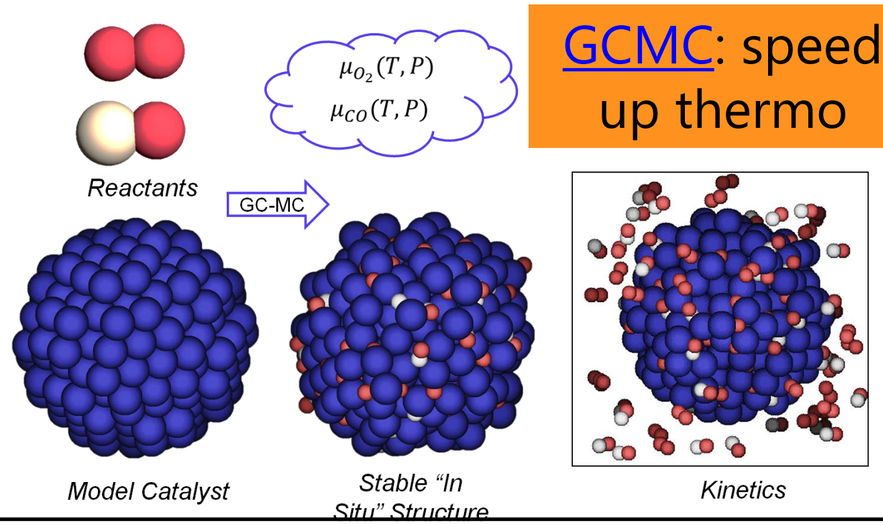




bond boost
build polymers

MCFF & CMA-ES parameterize ReaxFF force field

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

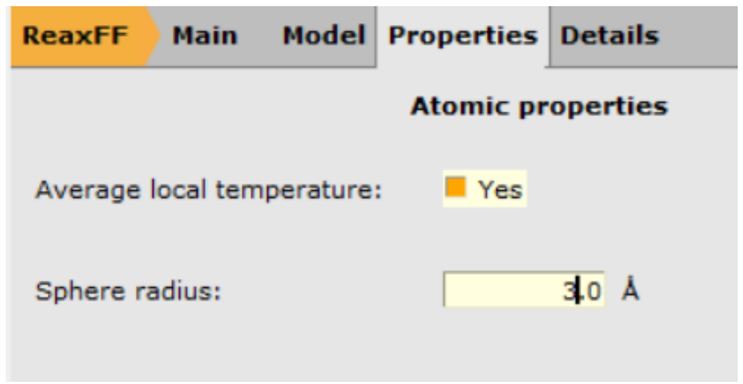
- Avoid having lattice vectors < 10 Å

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

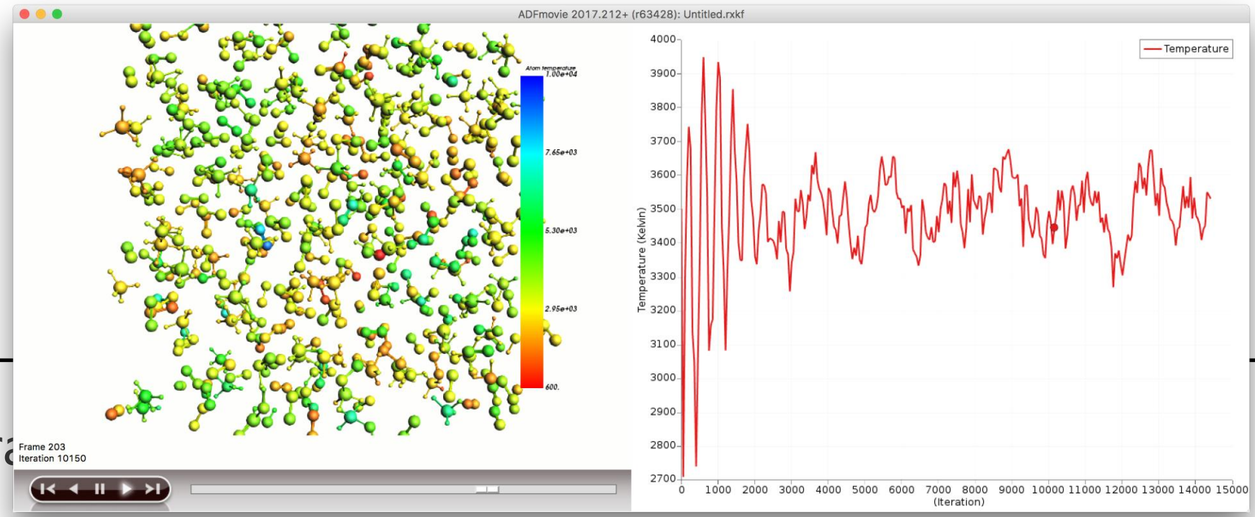
ReaxFF new: local T

Try yourself: seeing local temperature variations

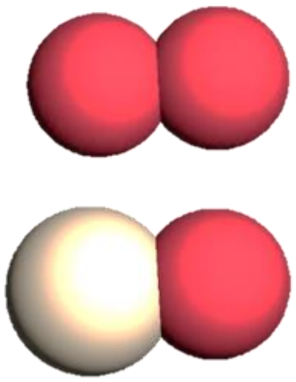
- Start with the methane burning tutorial
  (or with the molecule gun; see also news item with tutorial on visualizing local T)
  - Also local documentation! (Search in GUI)
  - Perhaps reduce the # of steps
- To visualize local temperatures:
  - Properties -> Atomic Properties
  - Tick Average local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



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

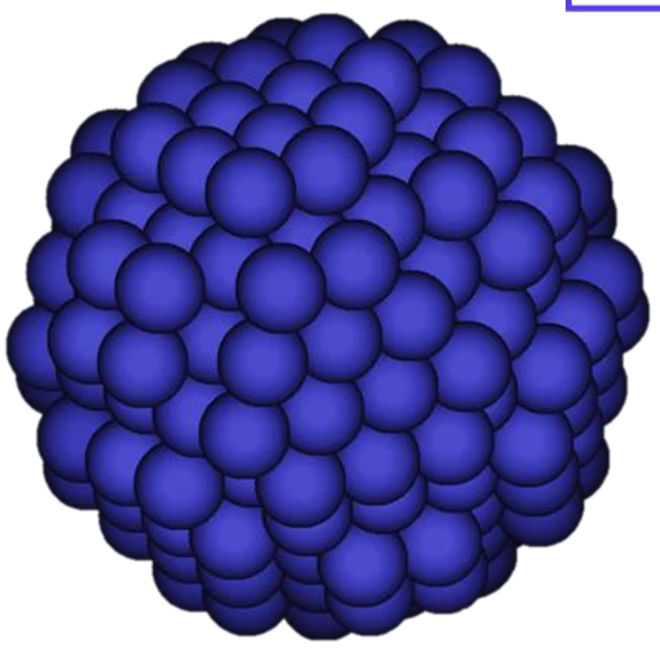
$\mu_{CO}(T,P)$

Reactive MD under actual conditions

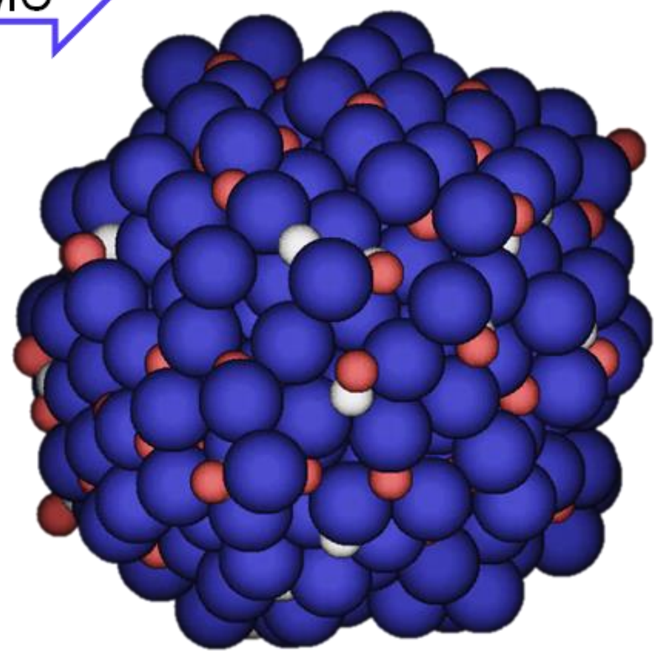
GCMC: What is stable (p,T)?
ReaxFF: What is active?

Reactants

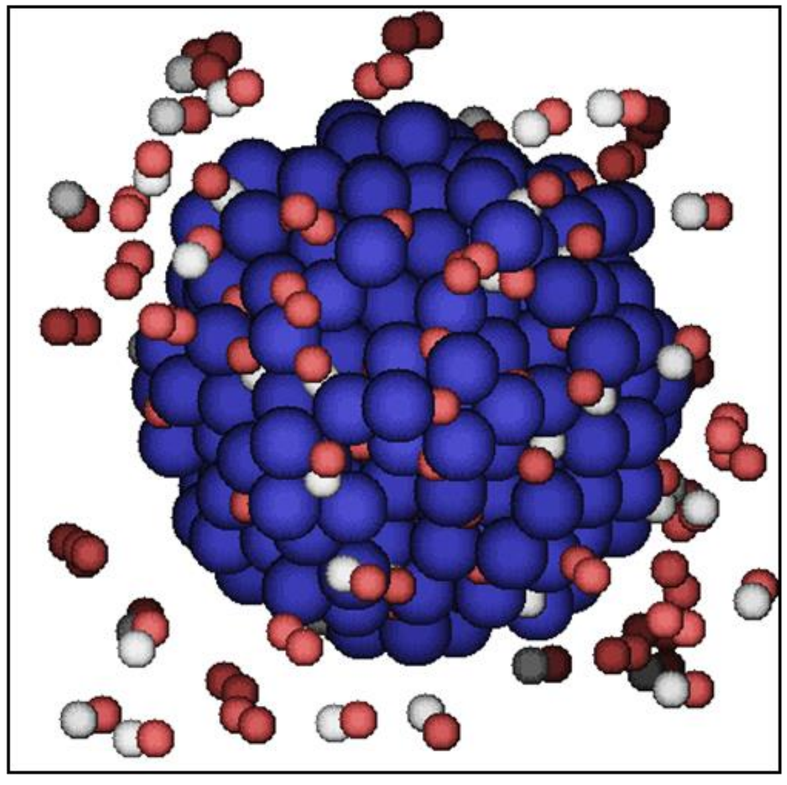
GC-MC



Model Catalyst



Stable "In Situ" Structure



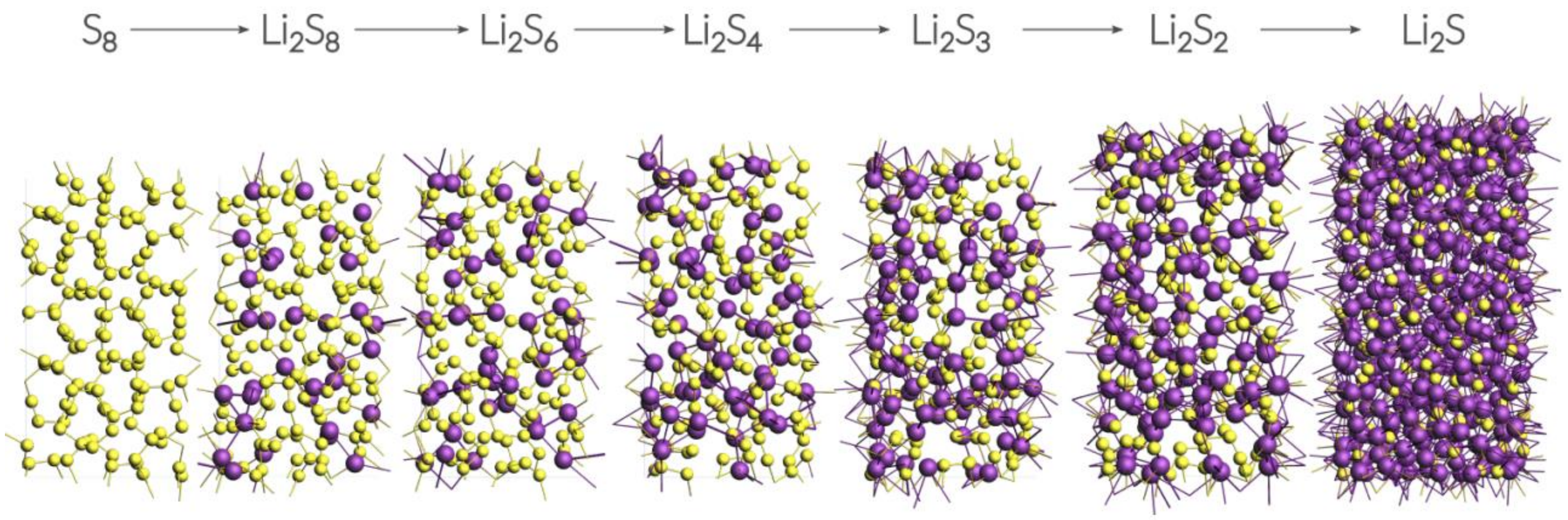
Kinetics

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

Try yourself: advanced GCMC battery discharge tutorial inserting Li in S

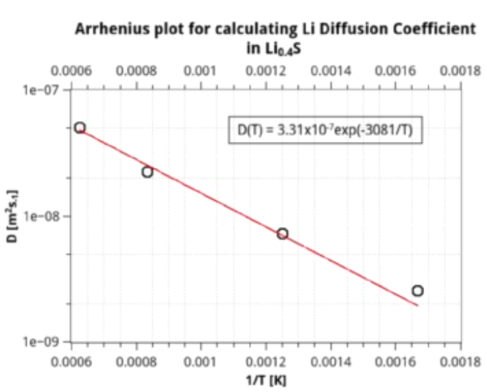


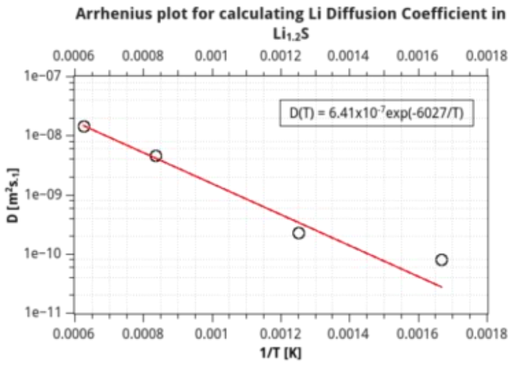
Based on: M. M. Islam, A. Ostadhossein, O. Borodin, A. T. Yeates, W. W. Tipton, R. G. Hennig, N. Kumar, and A. C. T. van Duin, ***ReaxFF molecular dynamics simulations on lithiated sulfur cathode materials***, Phys. Chem. Chem. Phys. **17**, 3383-3393 (2015)

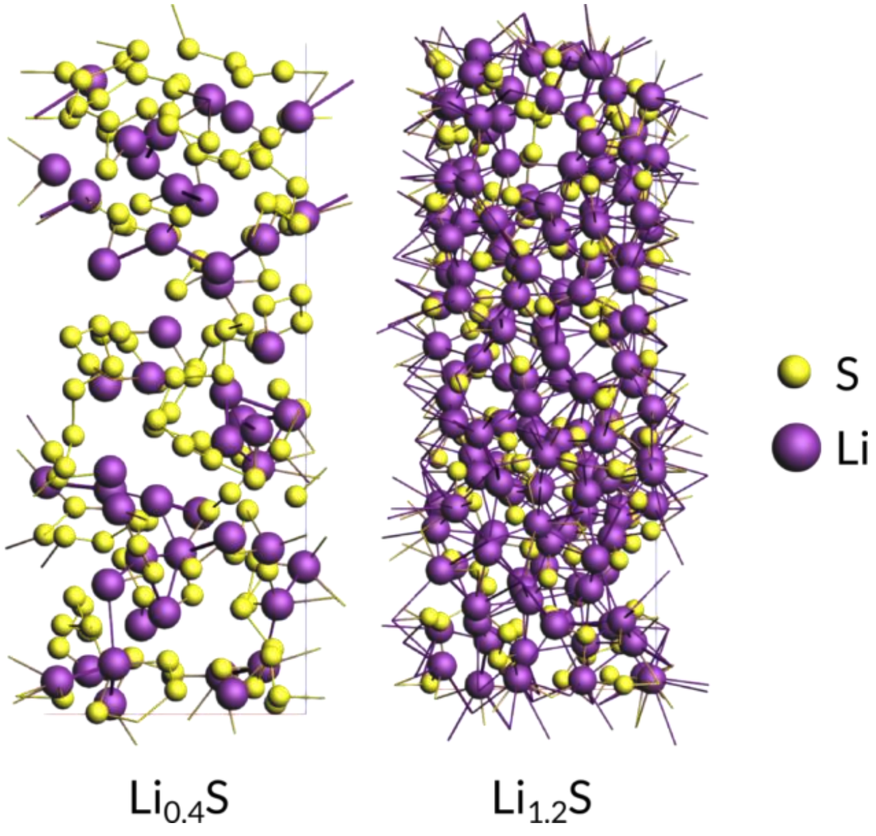
ReaxFF: Li diffusion in batteries

Try yourself: advanced tutorial Li diffusion (same S_8 coords as GCMC)

- 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'







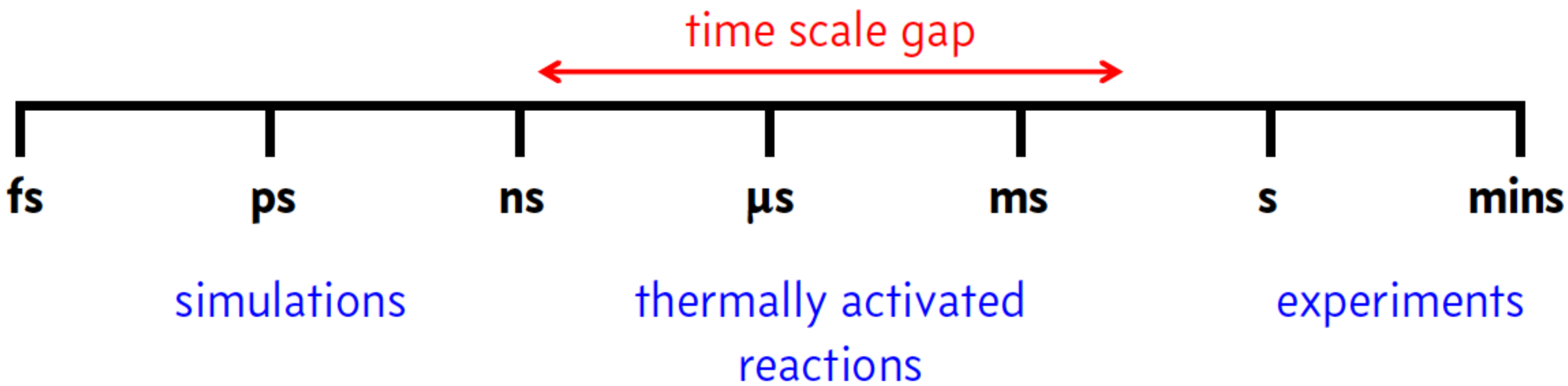
Travis is a bit difficult to use:

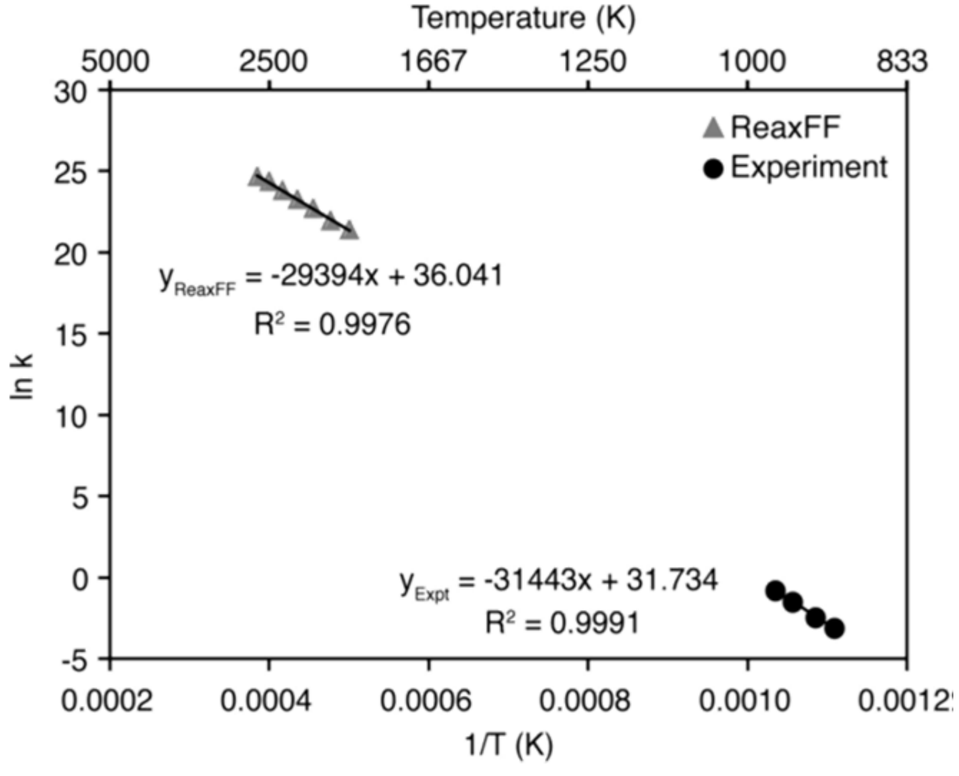
We will develop native analysis tools

You can also write script in PLAMS

The Time Scale Problem

- Even with fast (reactive) FF methods, there are still time limitations!



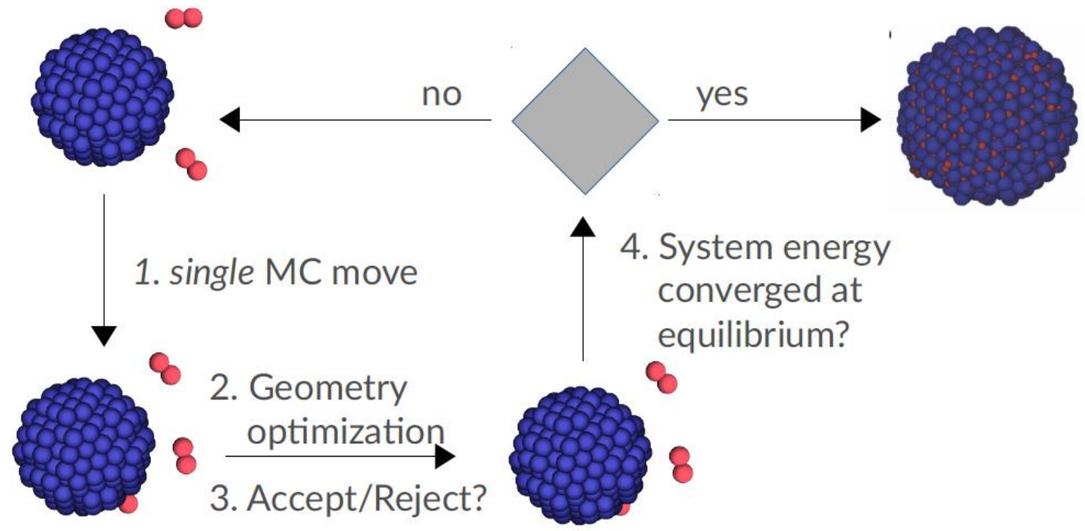


Usually/practically: increase T

But: do dynamics change?

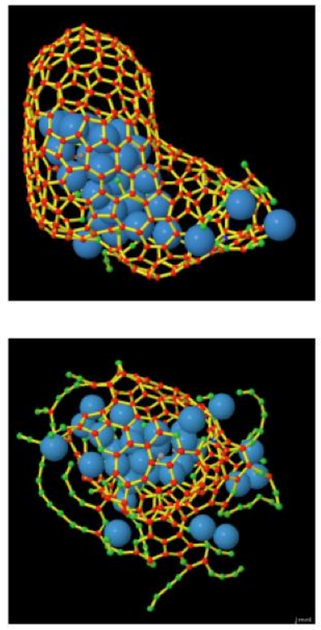
Accelerated dynamics

- Temperature - dynamics OK?
- Parallel replica dynamics (linear scale)
- <u>Monte Carlo</u>
  - Grand-canonical: thermodynamics
  - force bias – associated time scale?



J. Chem. Phys., 139 (2013) 044109

- Bond boost *J. Phys. Chem. A*, **2018**, *122*, 6633
- Metadynamics: collective variables
- Hyperdynamics: bias potential
- CVHD = MD + HD

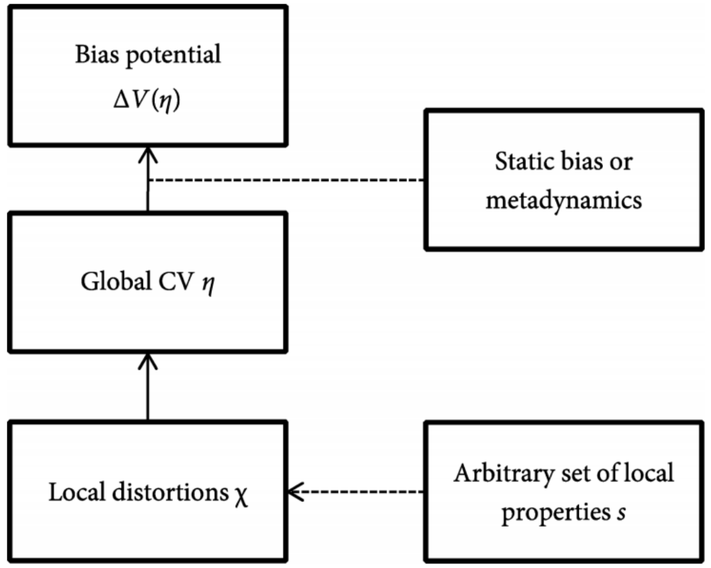


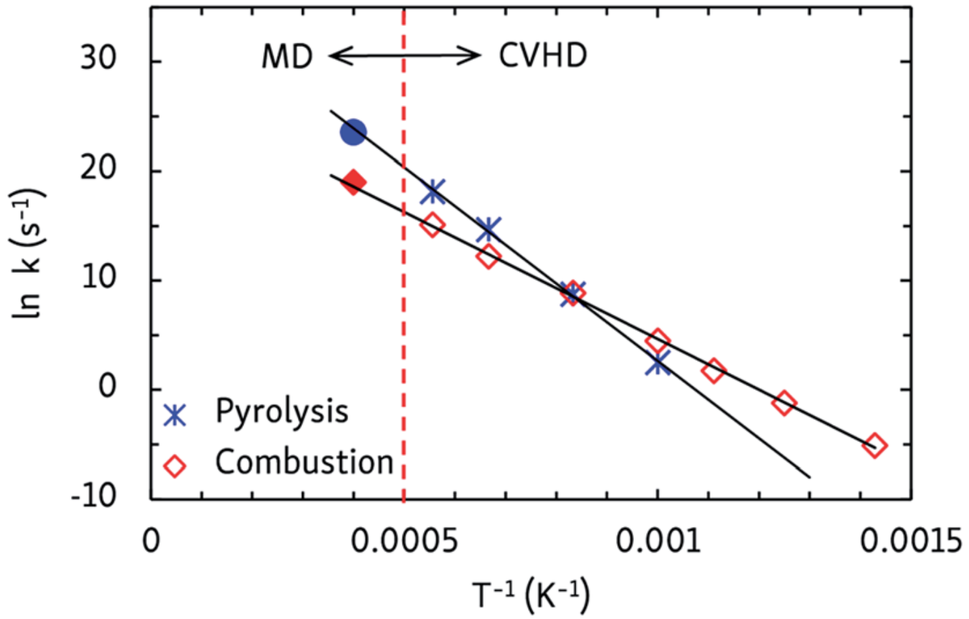
J.Am. Chem. Soc. **134**, 1256–1260 (2012)

Further accelerating ReaxFF

Collective-Variable driven Hyperdynamics (CVHD) => More CV's

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

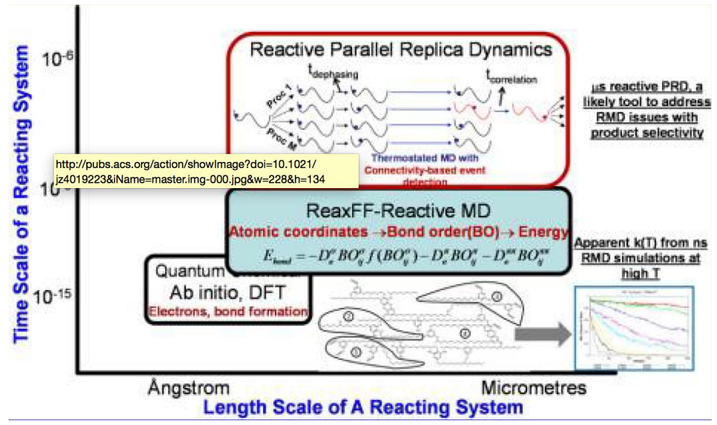




Bal & Neyts: J. Chem. Theory Comput. **11**, 4545 (2015); Chem. Sci., **7**, 5280 (2016)

Under development:

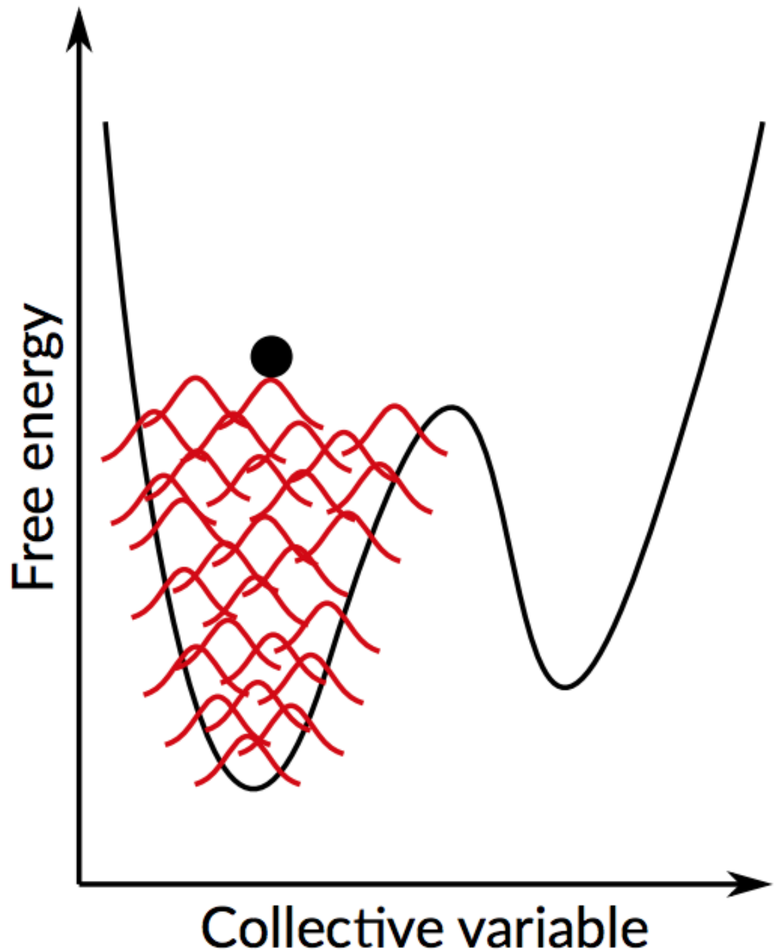
- Multi-replica methods (T-REMD, PRD)
  - Easy set-up, no bias (NB: PRD needs pre-defined event)
- (Path sampling approaches: TPS, TIS, ….)



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

Acceleration: biased MD

- **Metadynamics**: define a reaction coordinate (collective variable)
- Slowly add a Gaussian bias everywhere => Correct $\Delta G^{\#}$, but **no time / rate**

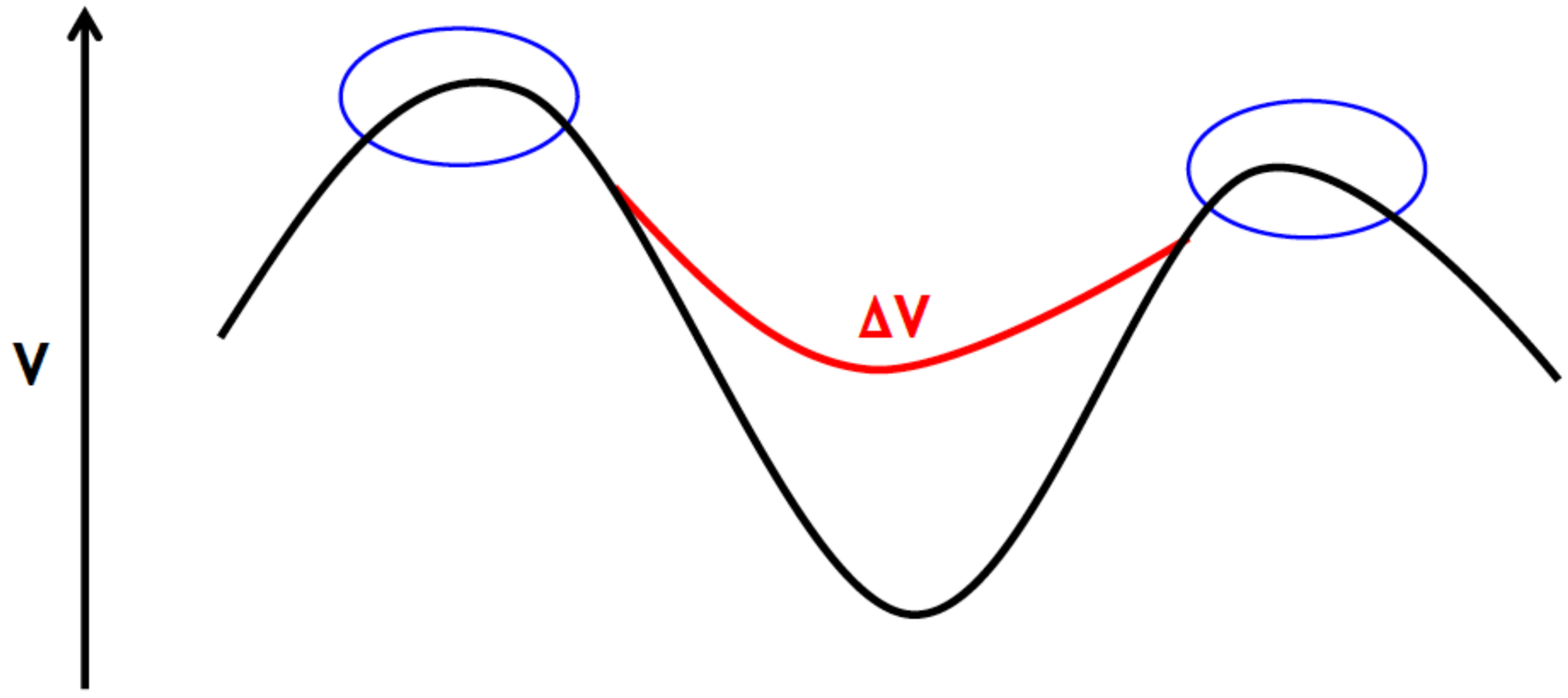


$$V_G\left(S\left(x\right),t\right)=\sum_{t'=t_G,2t_G,3t_G,\cdots} w\exp\left(-\frac{\left(S\left(x\right)-s_{t'}\right)^2}{2\delta s^2}\right)$$

A. Laio, http://people.sissa.it/~laio/Research/Res_metadynamics.php

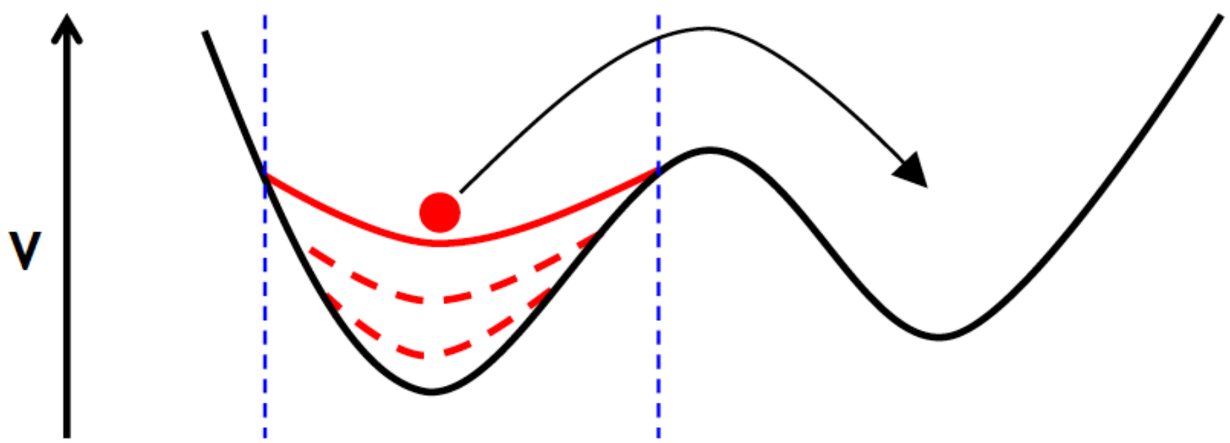
Acceleration: biased MD

- **Hyperdynamics**: Add bias potential ΔV to PES (not touching TS)
- MD on modified PES $V^*(\mathbf{R}) = V(\mathbf{R}) + \Delta V(\mathbf{R})$ => state-to-state rel. dynamics
- Boost MD time: **hypertime** is MD time x **boost factor** $e^{\beta\Delta V}$
- **How to define & parameterize V(R)?**
  - Hessian low modes, V itself, bond distortion, ….



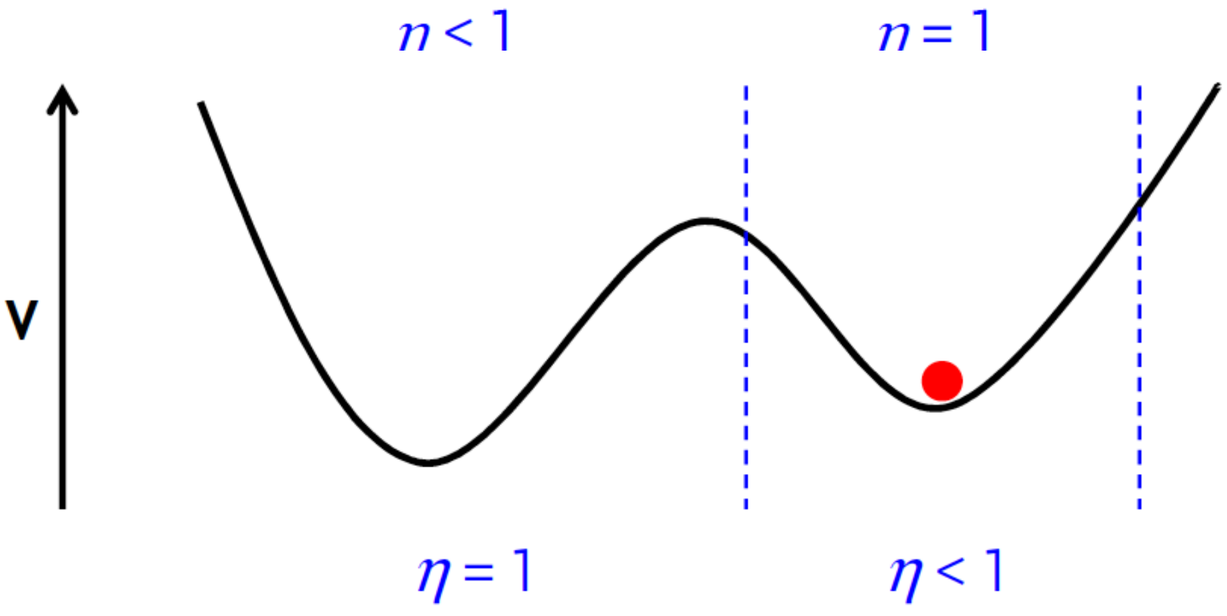
Collective-Variable driven HyperDynamics

- Hyperdynamics on a self-learning bias as function of CV: V(η)
- Get real dynamics without having to construct V(R) a priori



$$\Delta V = \sum_{i}^{hills} e^{-\frac{(\eta-\eta(t_i))^2}{2\delta^2}}$$

$$\text{acceleration} = \left\langle e^{\Delta V/kT} \right\rangle$$



Reset bias after transition

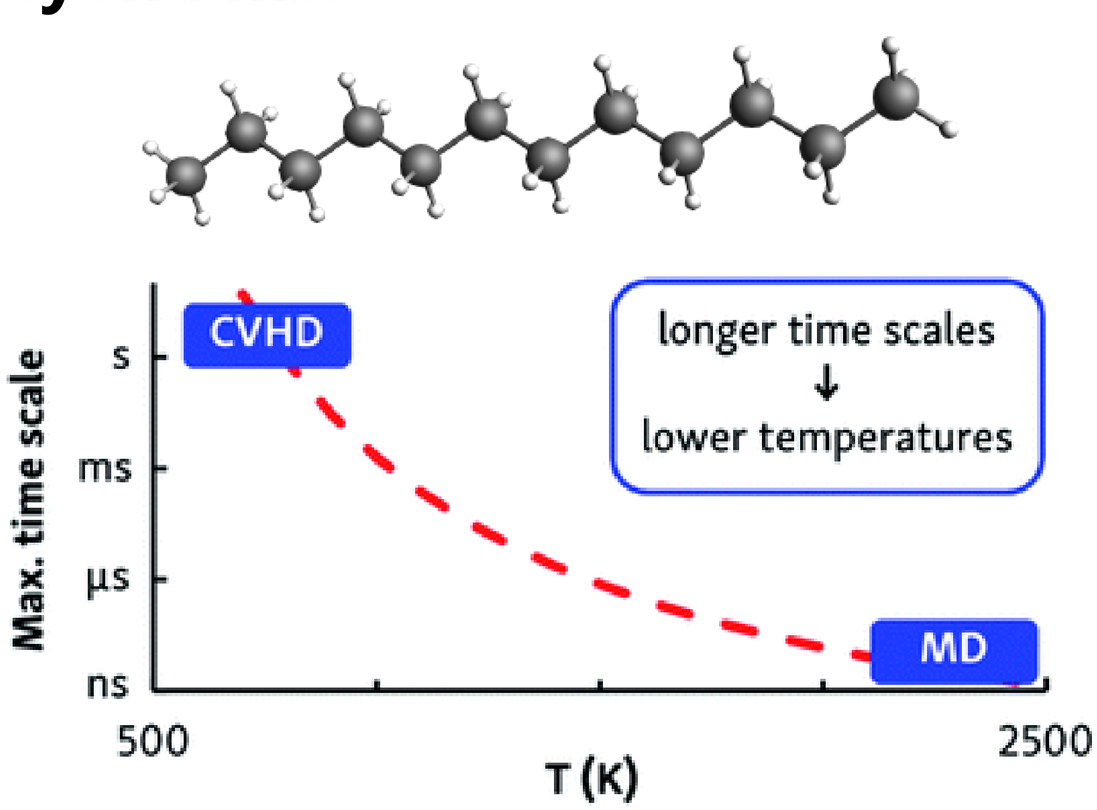
Bal & Neyts, J. Chem. Theory Comput. 11, 4545 (2015)

CVHD for pyrolysis & combustion

- Combustion studied often with ReaxFF, but at high T (~2000K)!
- Complex temperature-dependent pathways and products
- Barriers from ~30 kcal/mol (propagation) to ~80 kcal/mol (initiation)
- Parallel replica only got to 1 μs at 1350 K (Joshi et al.)

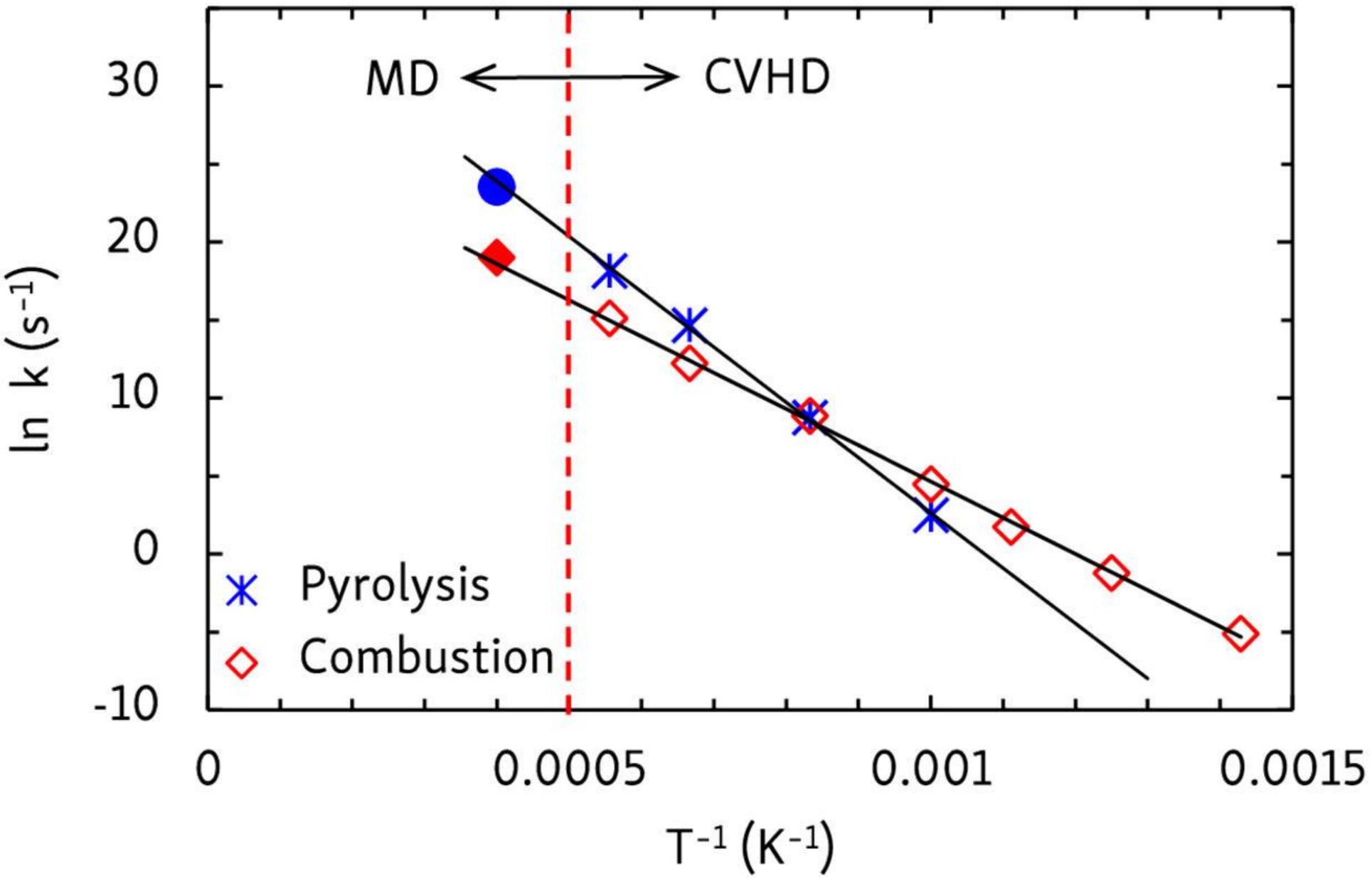
***n*-dodecane test system: CVHD reaches beyond ms!**

	Pyrolysis	Combustion
Lowest T	1000 K	700 K
Longest t	57 ms	39 s
Largest boost	6.3×10^6	1.3×10^9



Bal & Neyts, Chem. Sci. 7, 5280 (2016).

CVHD for pyrolysis & combustion



Day 1: CVHD **tutorial**

Fixing the CVHD scripts

If with one of the fixes you get an error:

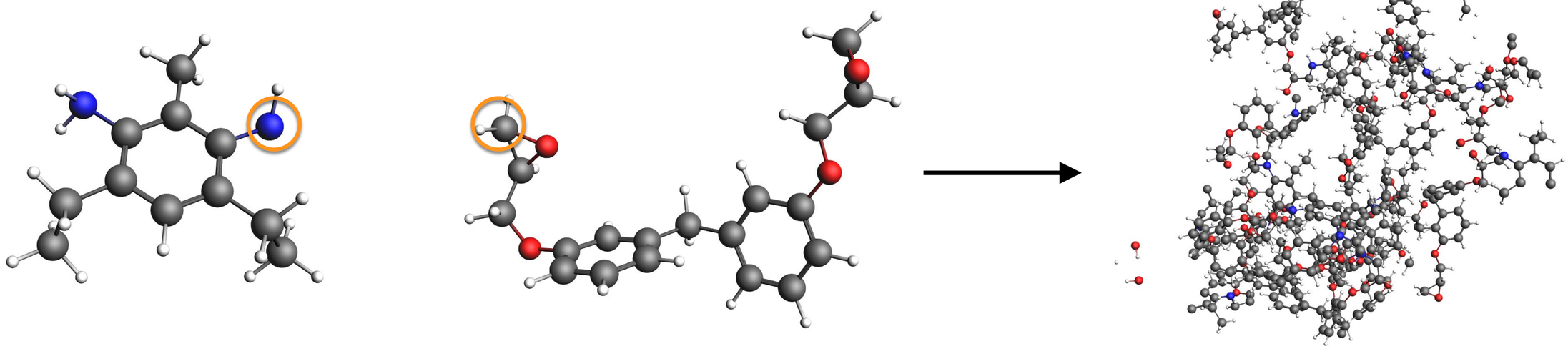
Run the cvhd-hills c.q. cvhd-hypertime script with an extra argument, e.g.
cvhd-hills myhills.csv
cvhd-hypertime my.logfile hypertime.csv

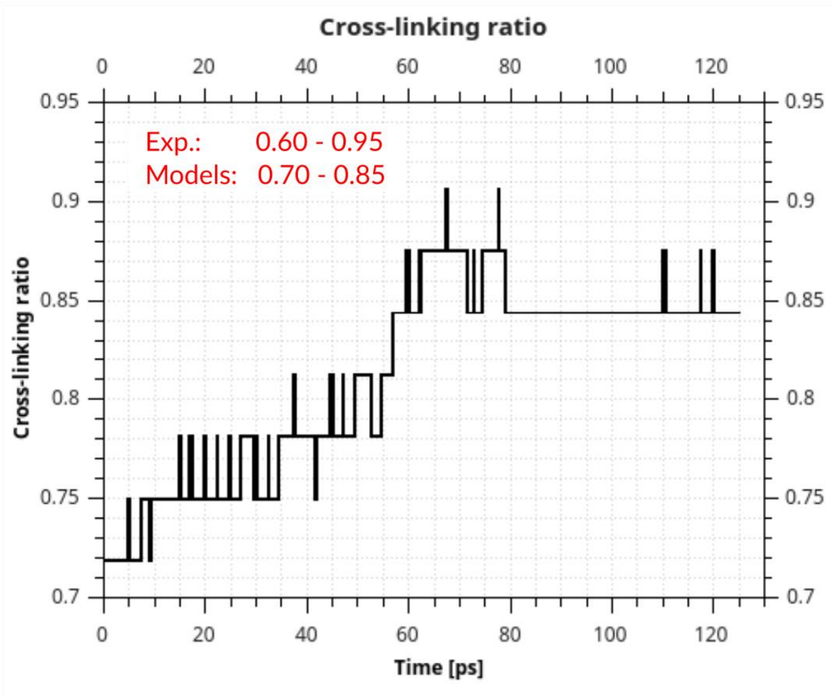
follow by one of these cmd

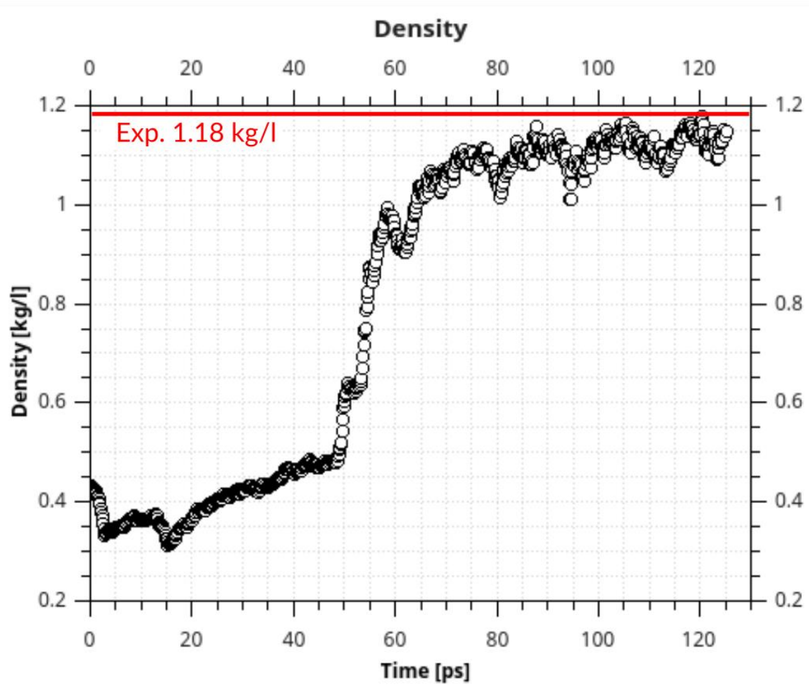
```
$ADFBIN/adfgraphs.exe myhills.csv
$ADFBIN/adfgraphs.exe hypertime.csv
```

Further accelerating ReaxFF

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



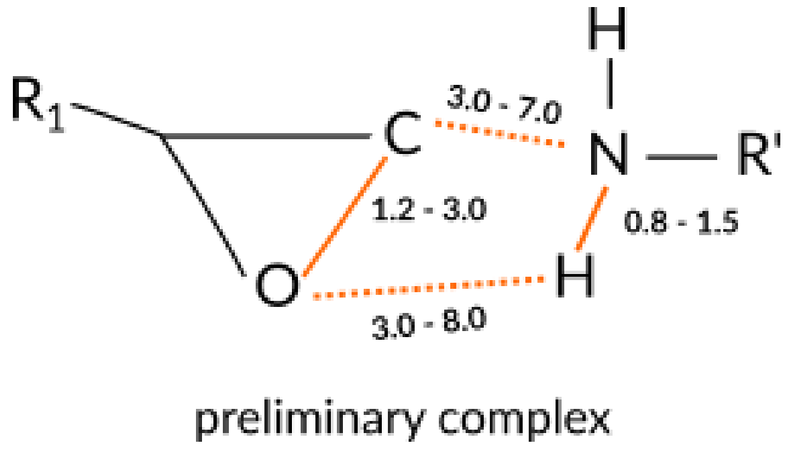




ReaxFF: bond boost

Day 1 hands on: start with bond boost tutorial (part is used in ReaxFF training set)

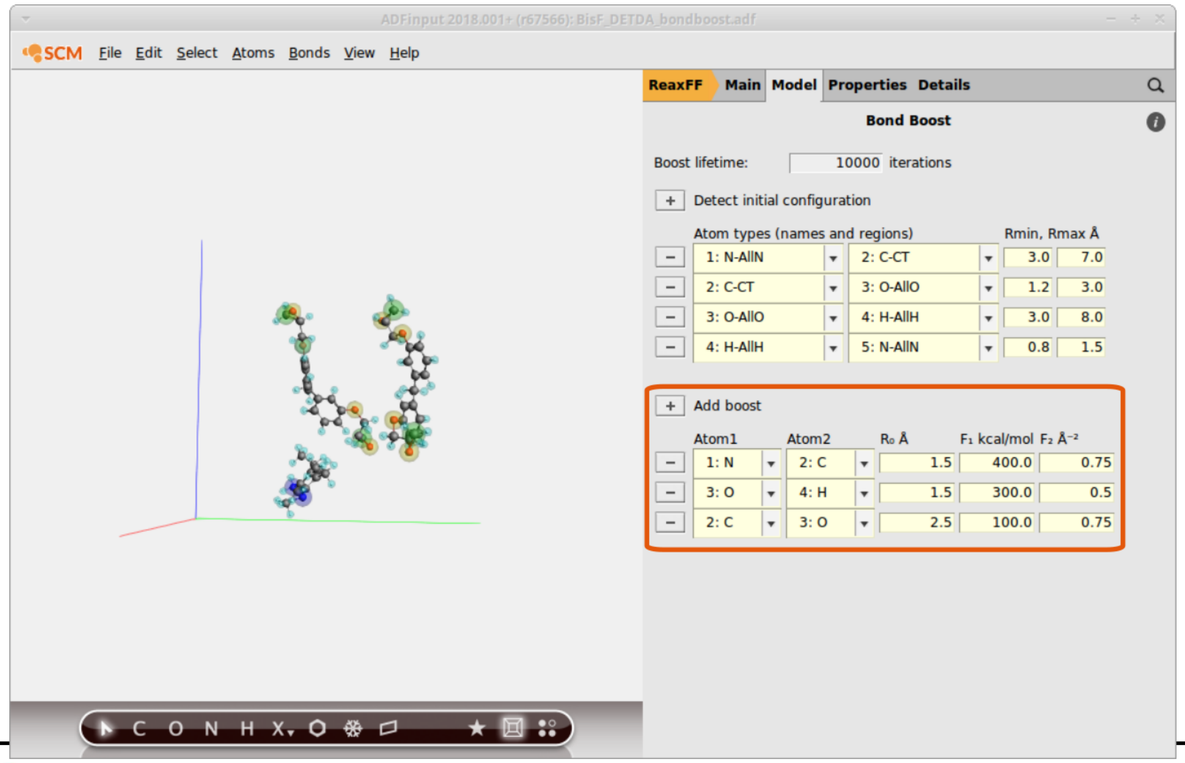
- Add boost potential when reactions are ‘close to reaction’, but not over barrier!
- Track bond distances of certain atoms =>
  - Needs tweaking for new ff or rxn => new param
- For small systems: easy to set up in GUI
- For bigger systems use **python scripting**

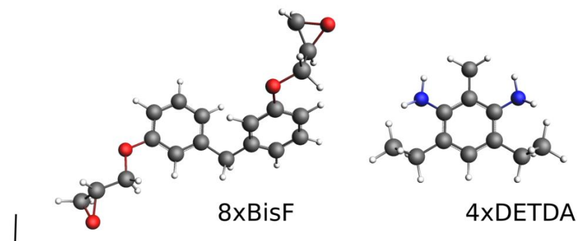


preliminary complex

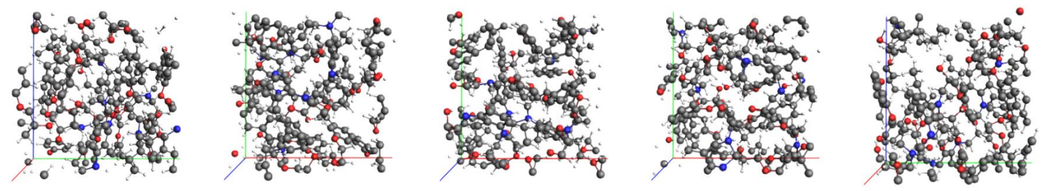
- Also check out mechanical polymer properties

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



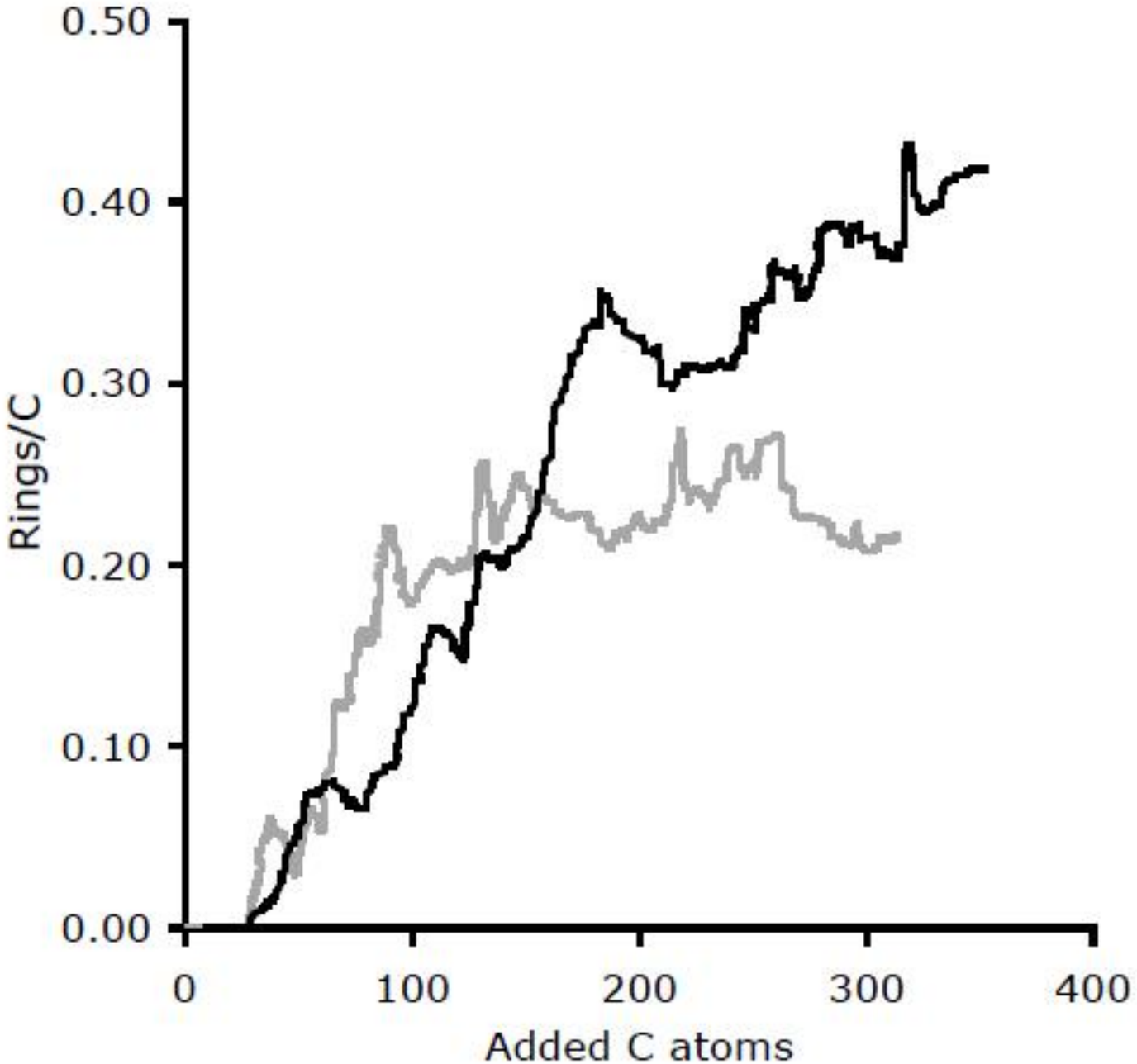


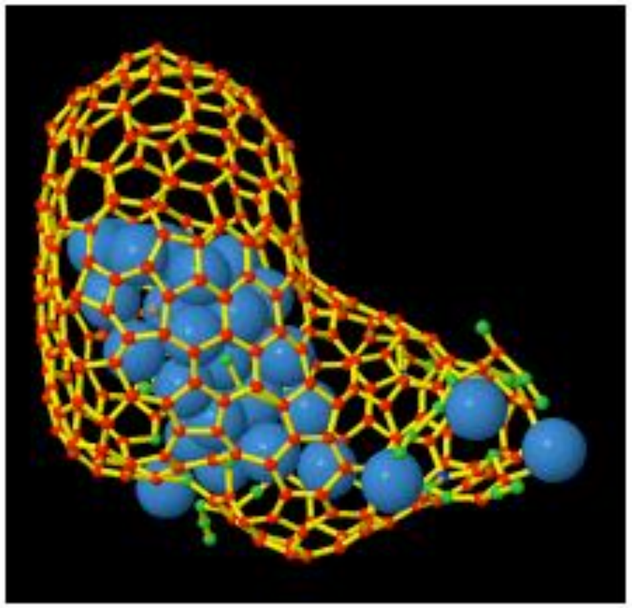
bond-boost.py ~ 1h per structure (on 4-cores)



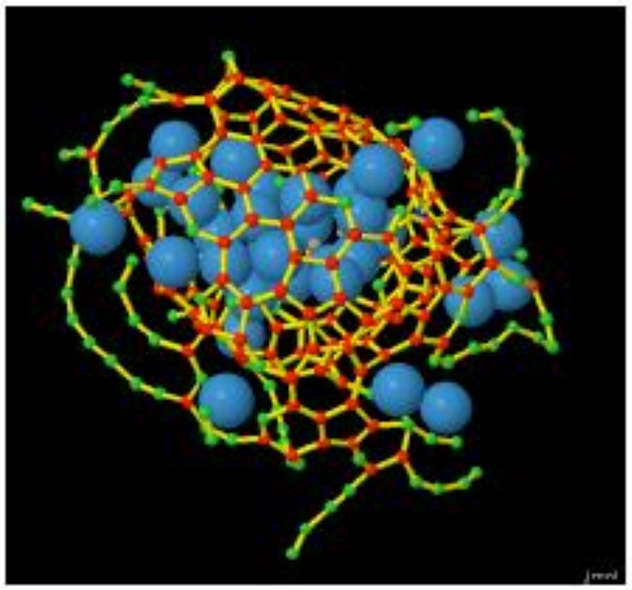
#1 X-link ratio = 0.81
#2 X-link ratio = 0.69
#3 X-link ratio = 0.63
#4 X-link ratio = 0.88
#5 X-link ratio = 0.81

Carbon nanotube formation: fbMC





MD + fbMC

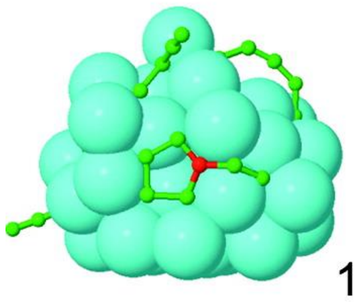


MD

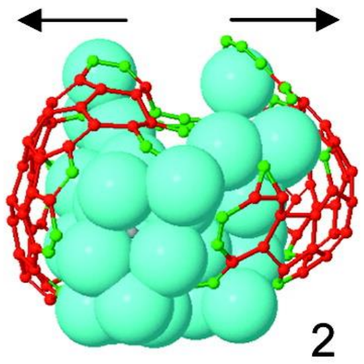
[Molecule gun + fbMC video]

Carbon nanotube formation: accelerate MD

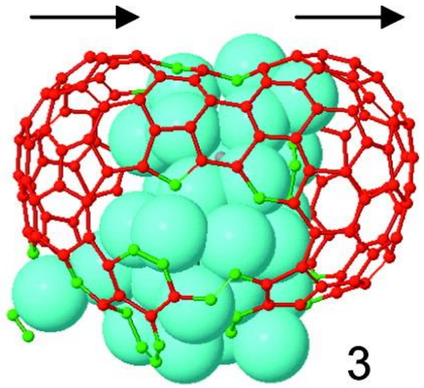
pentagon formation



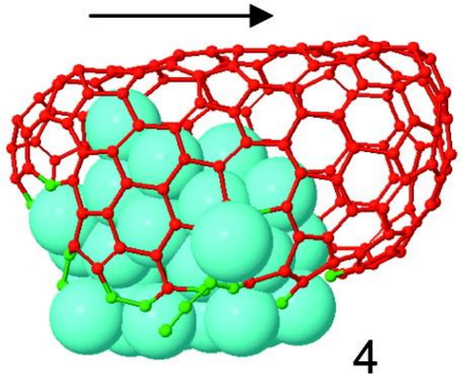
graphitic patches



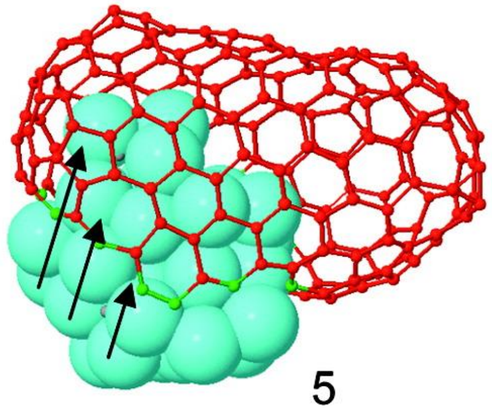
cap development



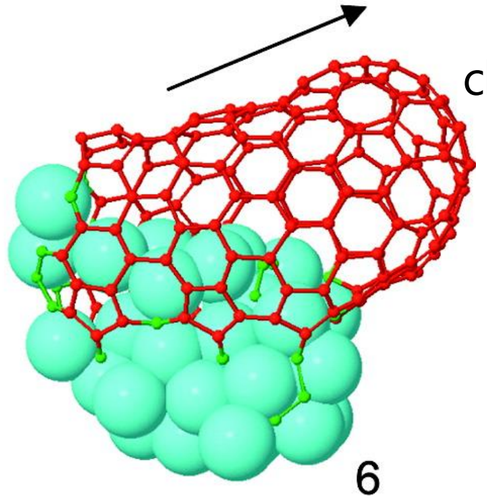
carbon network slides over catalyst



part of metal gets freed



end of sliding, chirality obtained

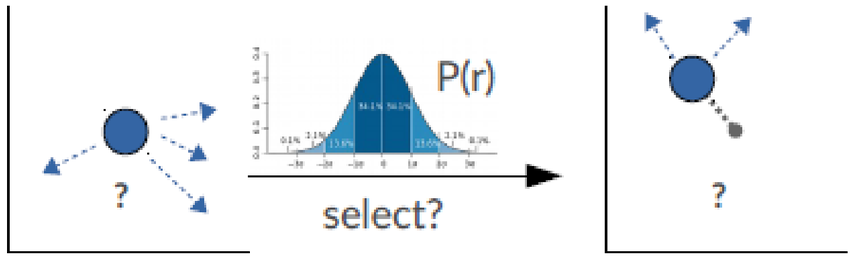


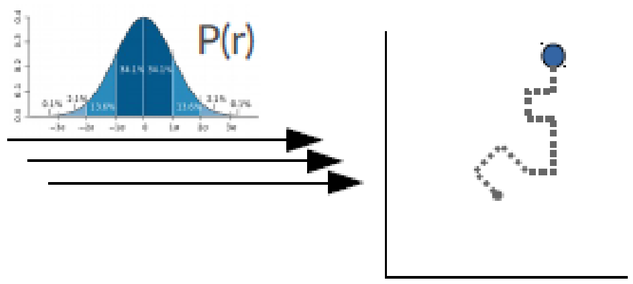
ReaxFF-MD + fbMC: healing of defects during sliding of network (steps 4-5).
First simulation showing growth of an armchair SWNT with **definite chirality** on catalyst.

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

Graphene healing with fbMC

Metropolis MC

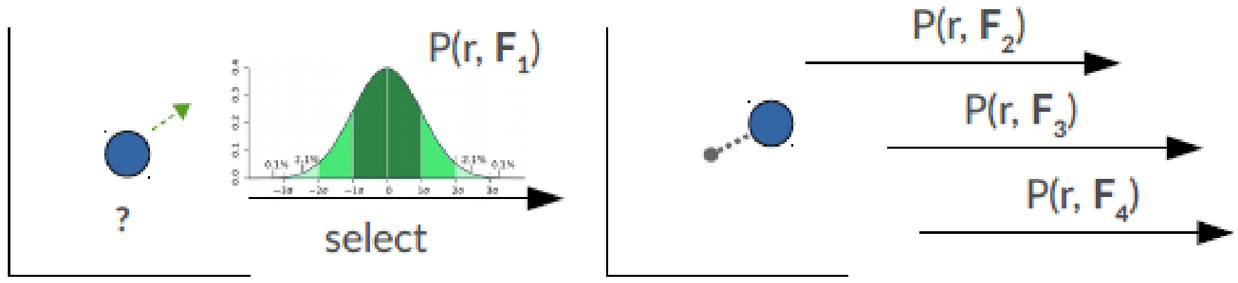




Sampling structural *Phase Space*

according to *one* well known
ensemble distribution function (e.g. NVT)

fbMC



Sampling the *Dynamics*

- each change driven by "instantaneous" and "local" Probability Distributions
- irrespective from distance to equilibrium
- Limits of P(r,F)
  T >> F → completely random movement
  T << F → Particle moves exactly in direction of force

Derivation……Timonova *et al.*, Phys. Rev. B 81, 144107 (2010)
Applicability…Bal and Neyts, J. Chem. Phys. 141, 204104 (2014)
Application…..Mees *et al.*, Phys. Rev. B 85, 134301

How fbMC works & input settings

During a *single* fbMC step *each* atom is displaced by $\xi_{i,j} \cdot \Delta_i$ in *every* cartesian direction, with $\xi_{i,j}$ stochastically distributed as follows:

$$p(\xi_{i,j}) = \begin{cases} \dfrac{e^{\gamma_{i,j}(2\xi_{i,j}+1)} - e^{-\gamma_{i,j}}}{e^{\gamma_{i,j}} - e^{-\gamma_{i,j}}} & \text{if} \quad \xi_{i,j} \in [-1, 0[\\ \dfrac{e^{\gamma_{i,j}} - e^{\gamma_{i,j}(2\xi_{i,j}-1)}}{e^{\gamma_{i,j}} - e^{-\gamma_{i,j}}} & \text{if} \quad \xi_{i,j} \in]0, 1] \end{cases}$$

$$\gamma_{i,j} = \frac{F_{i,j}\Delta_i}{2k_B T}, \quad \Delta_i = \Delta\sqrt{\frac{m_{\min}}{m_i}},$$

$F_{i,j}$: Force along component j acting on atom i

m_i : mass of atom i

m_{min}: smallest mass in the system

Δ : system wide parameter

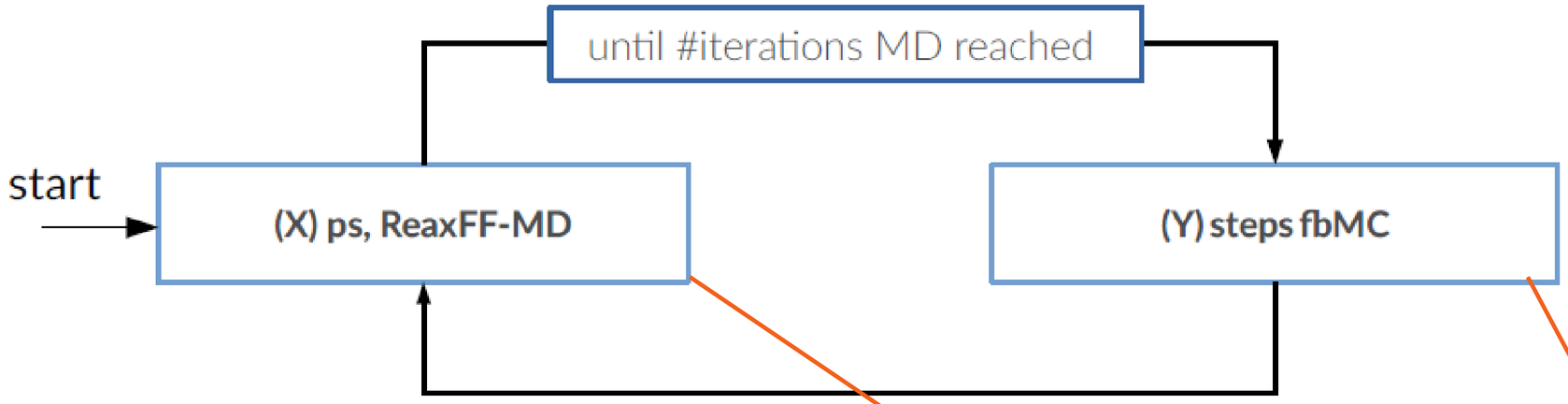
*** required input: T, Δ ***

How to choose Δ? Balance speed & 'physicality'

large Δ → faster, but larger violation detailed balance

formally correct only for *infinitesimal small* Δ

How fbMC works & input settings

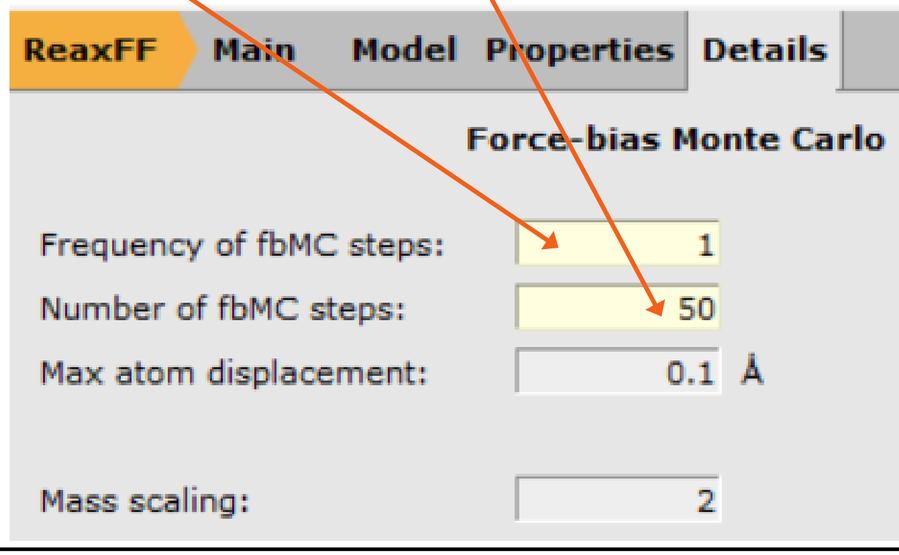


(X) : "Frequency of fbMC steps" = start fbMC from last MD structure after (X) steps
(Y) : "Number of fbMC steps" = make (Y) fbMC steps before restarting the MD

Example:
(X) = 1000
(Y) = 500
Number of Iterations = 10 000

1. 1000 steps MD
2. 500 steps fbMC
3. 1000 steps MD
4. 500 steps fbMC
5. .....

Σ_{MD} = 10 000 steps
Σ_{fbMC} = 5000 steps



Kinetics? Flexible mass scaling

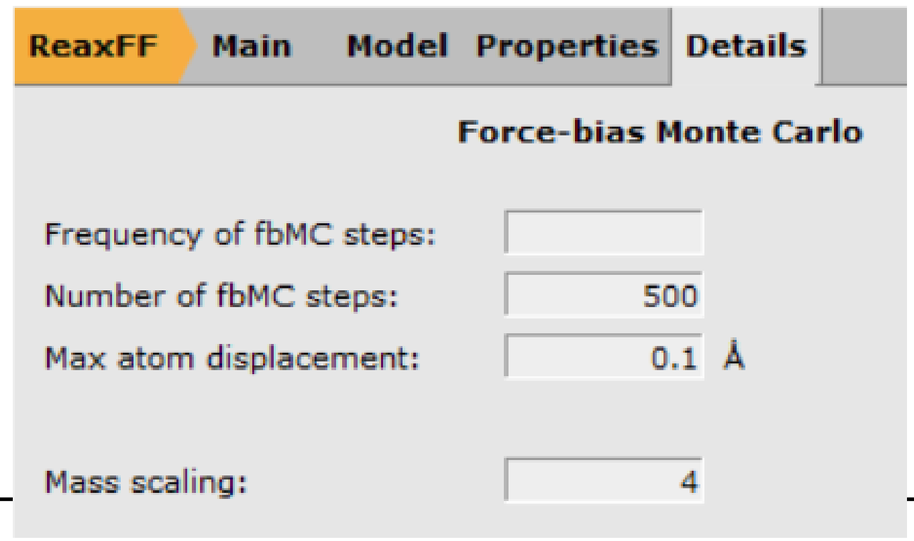
fbMC does not scale time uniformly but 'pseudodynamics' very close to physical dynamics (J. Chem. Phys. 141, 204104 (2014)).

time-stamped force-bias Monte Carlo? **(tfMC**)

Test events 'real dynamics' vs tfMC $\langle \Delta t \rangle_{\mathrm{tfMC}} = \Delta t_{\mathrm{MD}} \frac{\mathrm{MSD}_{\mathrm{tfMC}}}{\mathrm{MSD}_{\mathrm{MD}}}$

diffusion coefficients, reaction rates scale ~$m^{1/2}$
imcroo=4 in control file or set in fbMC panel in GUI

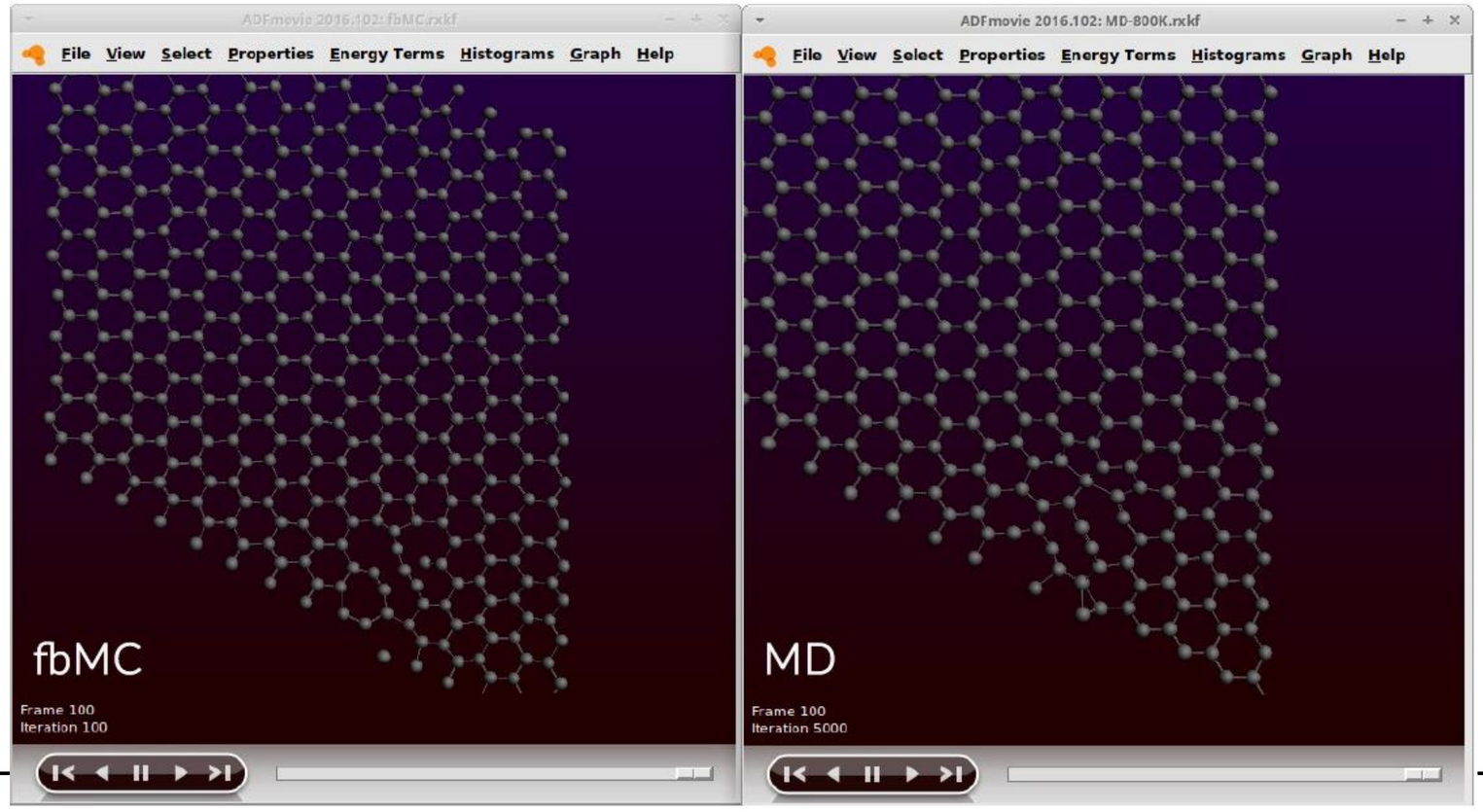
$$\Delta_i = \Delta \left(\frac{m_{\min}}{m_i} \right)^{1/4}$$



Healing graphene

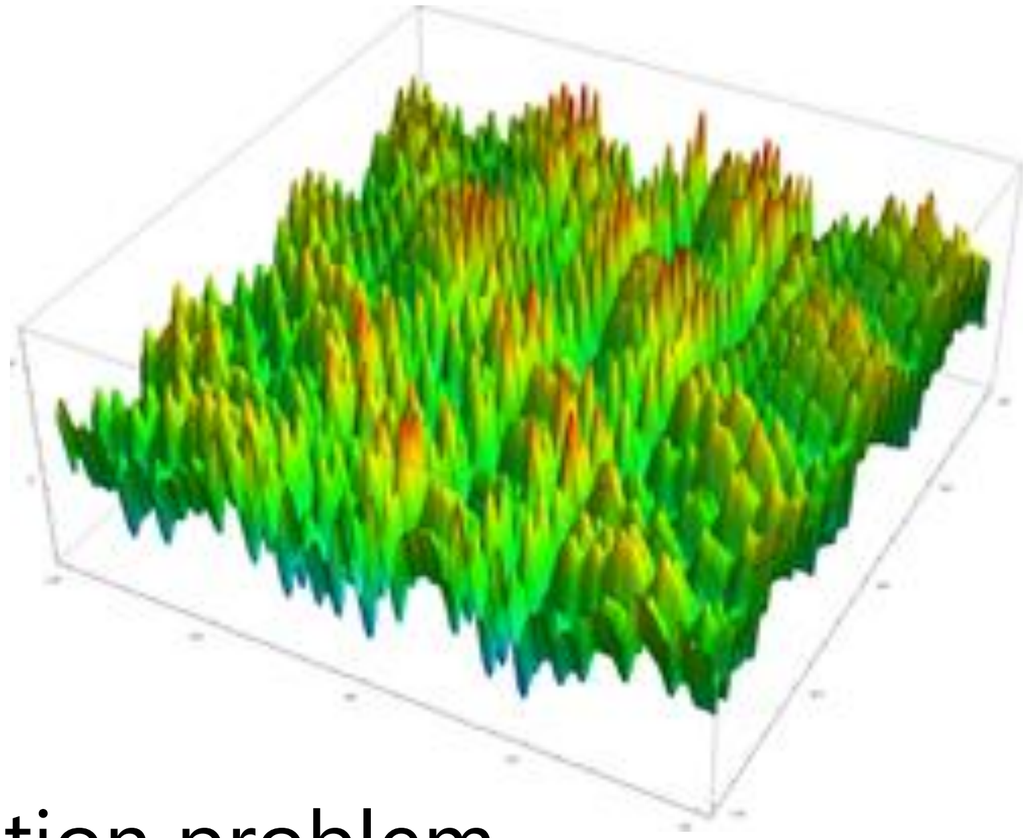
Hands-on day 1: **fbMC tutorial**

- build a defected graphene
- Run healing at 800 K with fbMC, tfMC & regular ReaxFF
- Try a few different Δ (probably not much larger!)
- Compare healing visually (+ PLAMS script?)
- Any better (newer) force field?



Optimizing ReaxFF parameters

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

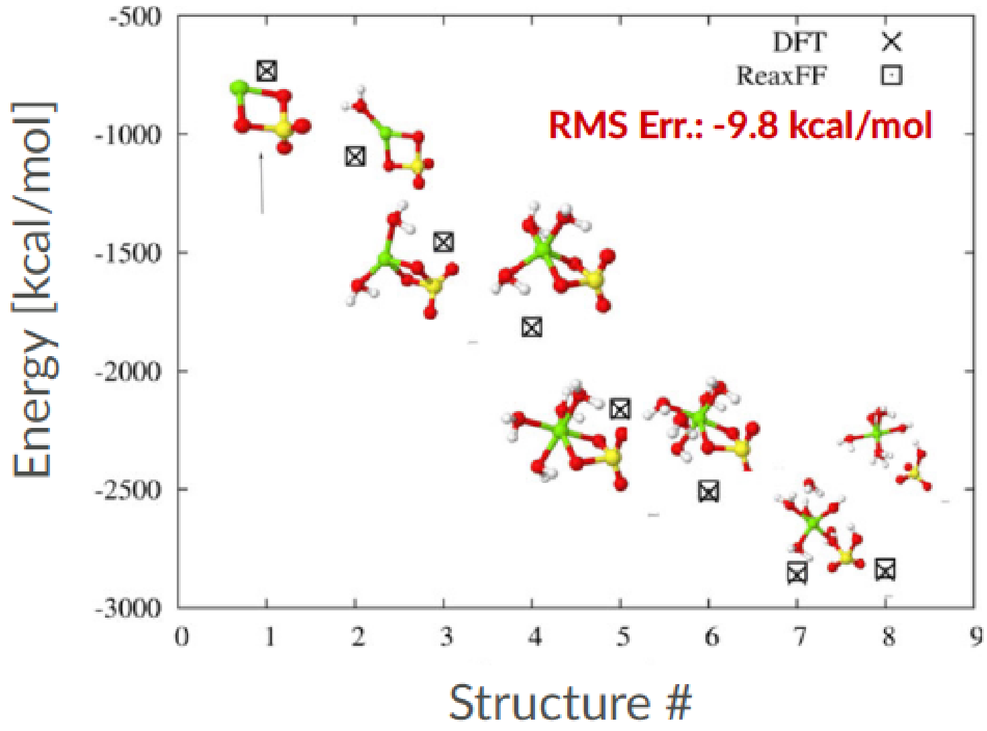


- Highly complex global optimization problem

Monte Carlo with Simulated Annealing

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

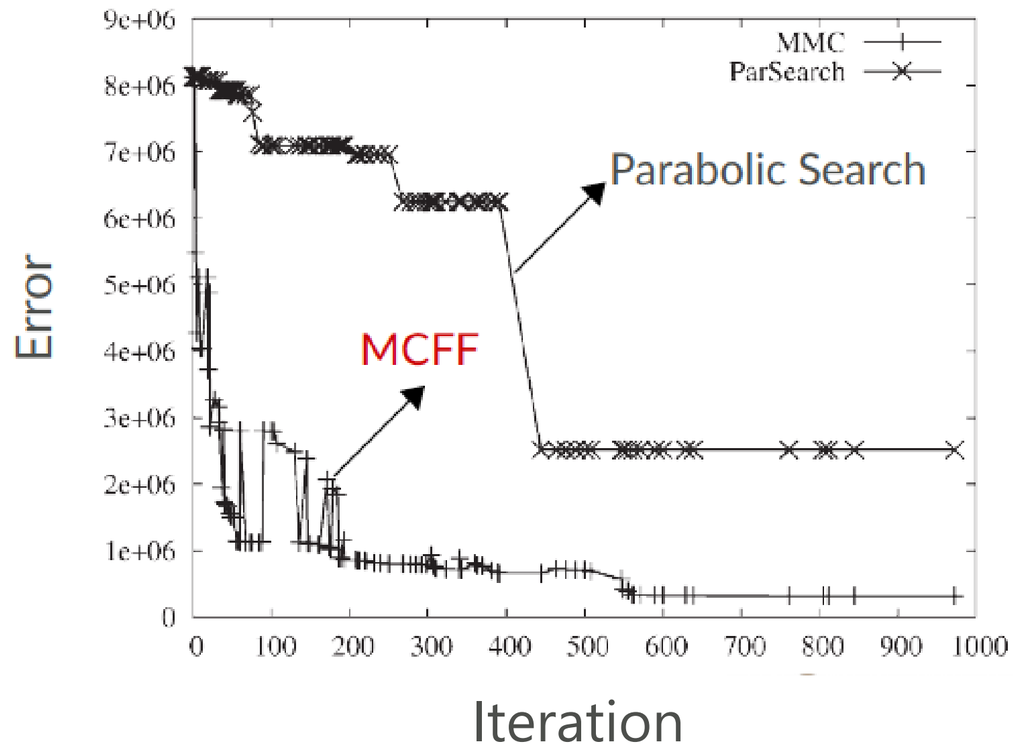
$x_{i,TS}$ = reference
$x_{i,ReaxFF}$ = current ReaxFF
σ_i = weight



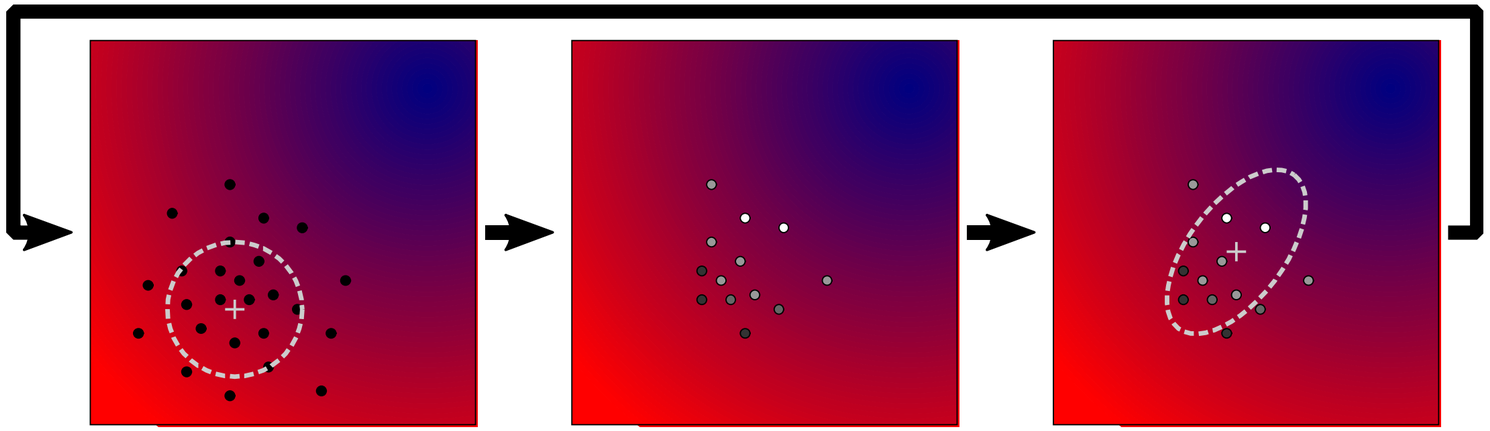
References include:

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

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



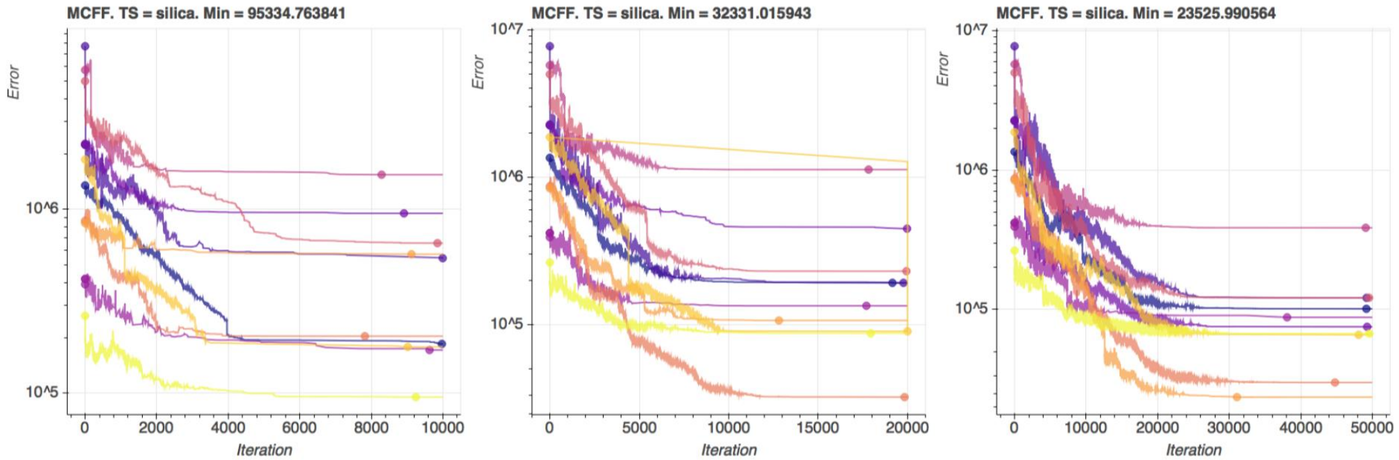
Covariance Matrix Adaptation Evolutionary Strategy



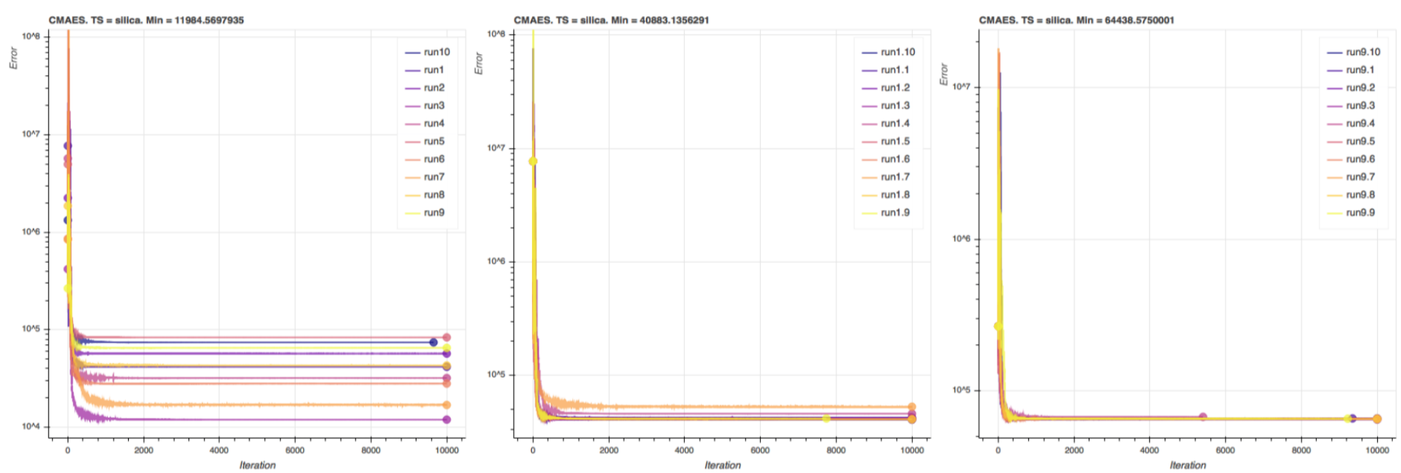
Generate points using multivariate Gaussian distribution with covariance matrix C

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

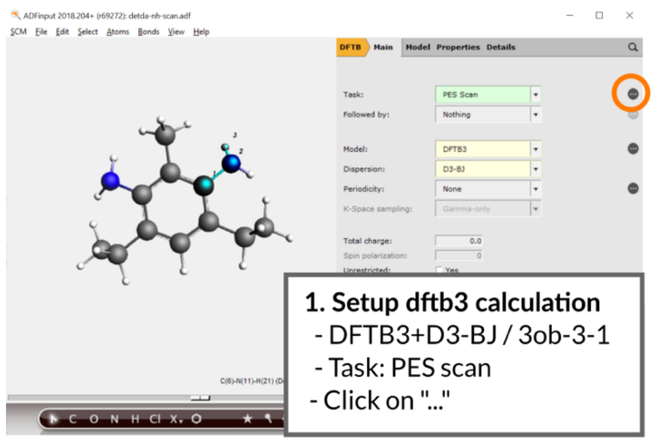


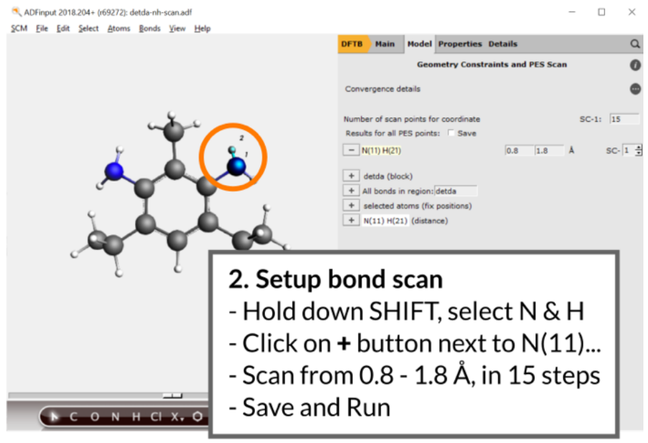
CMA-ES
(Shchygol et al. arxiv)

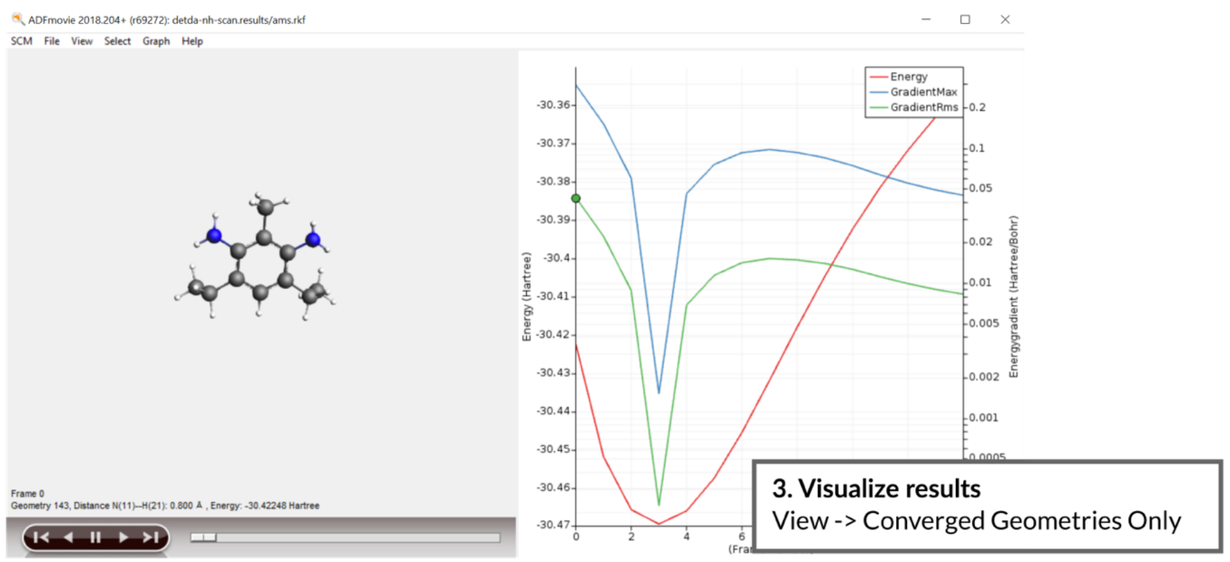
ReaxFF: reparameterization

[Refine ReaxFF parameters for cross-linking polymers]

- 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, trajectories
- Add bond scans
- Run CMA-ES optimization
- Test errors, cross-validate
- Try to further refine
- See also: Co training set
  - Relative crystal energies
  - Equation of State, elastic tensor
  - Cohesive energy
  - Defect, adsorption energy
  - Surface energies

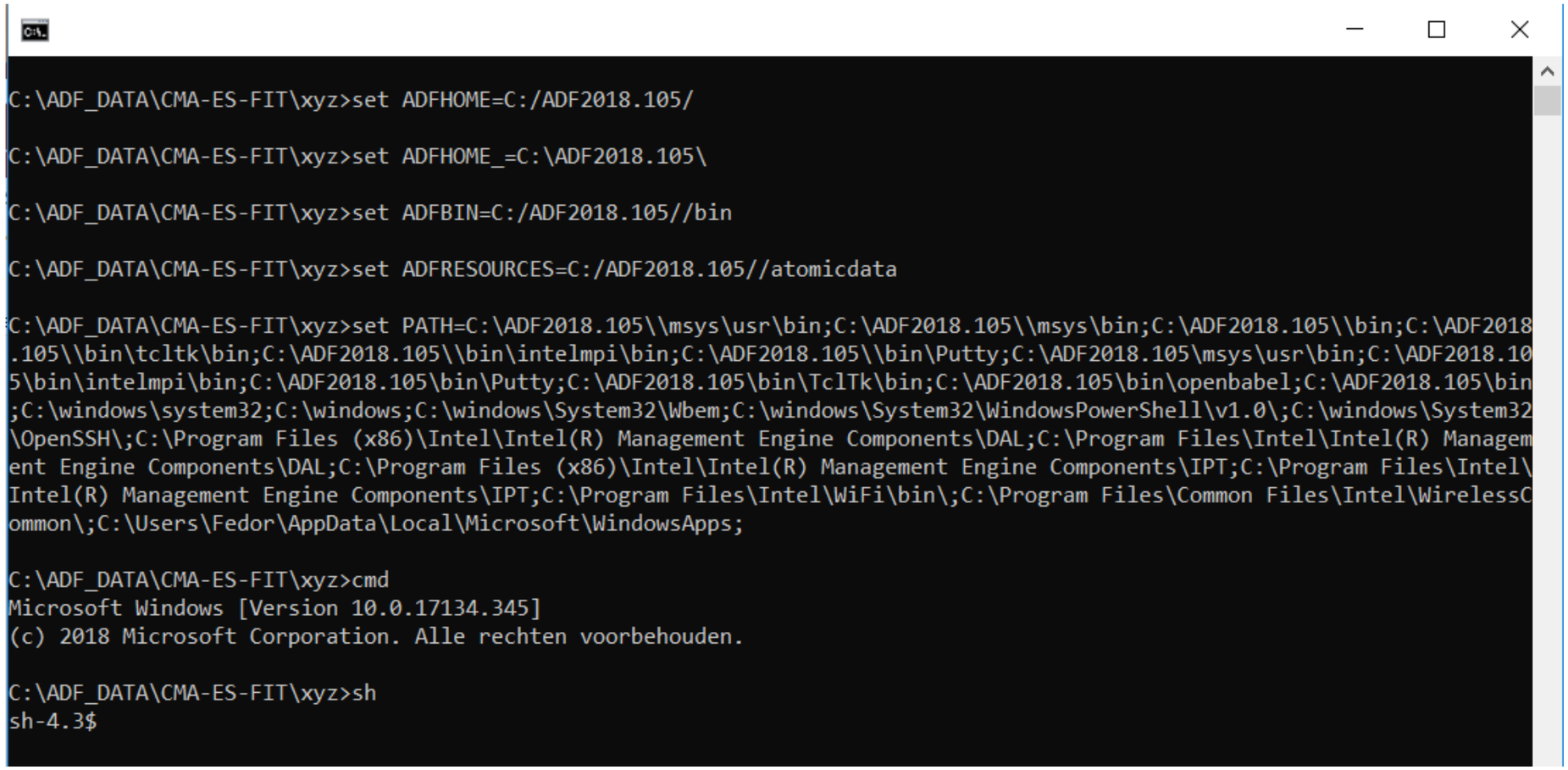






Some notes on (Windows) scripting

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



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