From electrons to the simulation of materials

Ralf Drautz
ICAMS, Ruhr-Universität Bochum, Germany

Workshop: Complex Scientific Workflows at Extreme Computational Scales
Part of Long Program: New Mathematics for the Exascale: Applications to Materials Science
IPAM, UCLA, Los Angeles, 3 May 2023

ICAMS Team

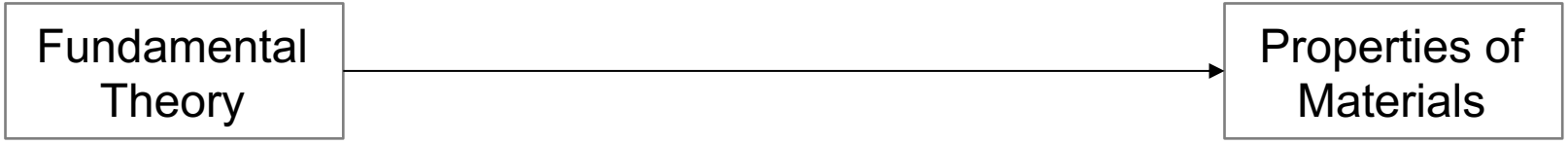
- Yury Lysogorskiy
- Matous Mrovec
- Anton Bochkarev
- Minaam Qamar
- Matteo Rinaldi
- Yanyan Liang
- Eslam Ibrahim

Related presentations at program

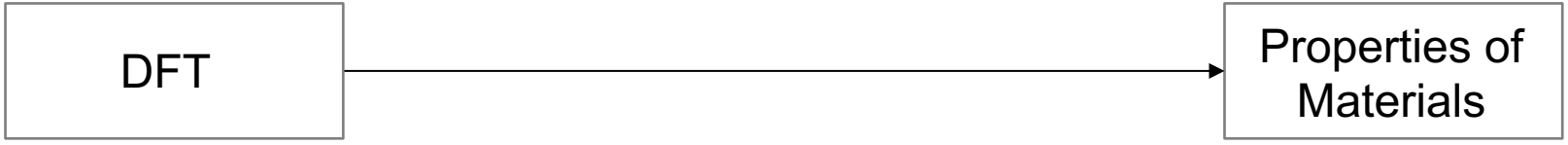
- Michele Ceriotti
- Jan Janssen
- Jörg Neugebauer

- Gábor Csányi
- James Kermode
- Boris Kozinsky
- Ivan Oleynik
- Christoph Ortner
- Aidan Thompson

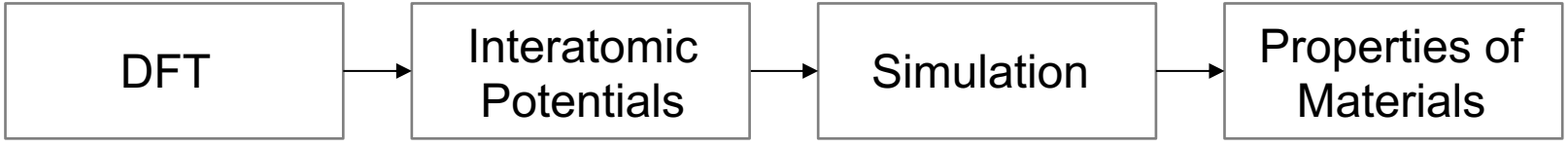
Outline



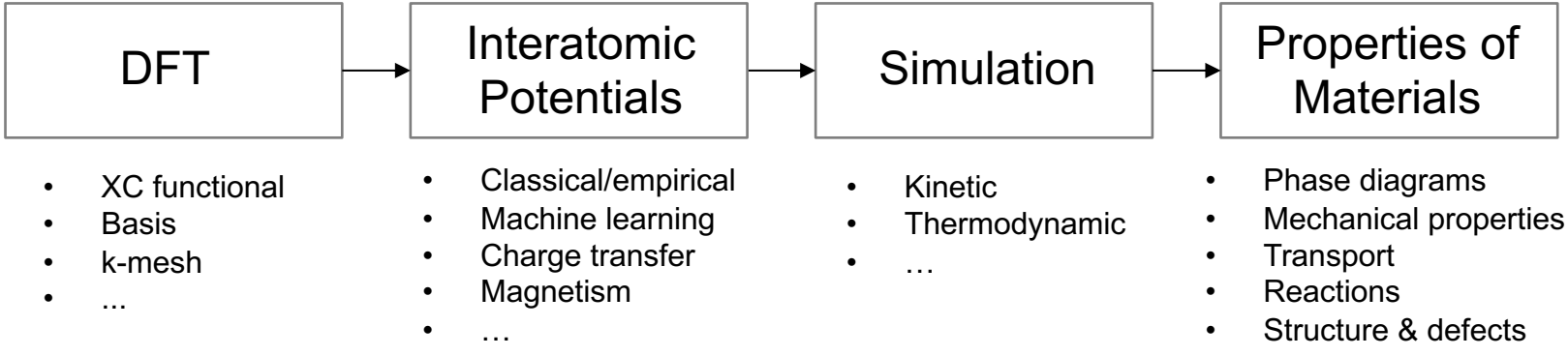
Outline



Outline


DFT
Interatomic Potentials
Simulation
Properties of Materials

Outline



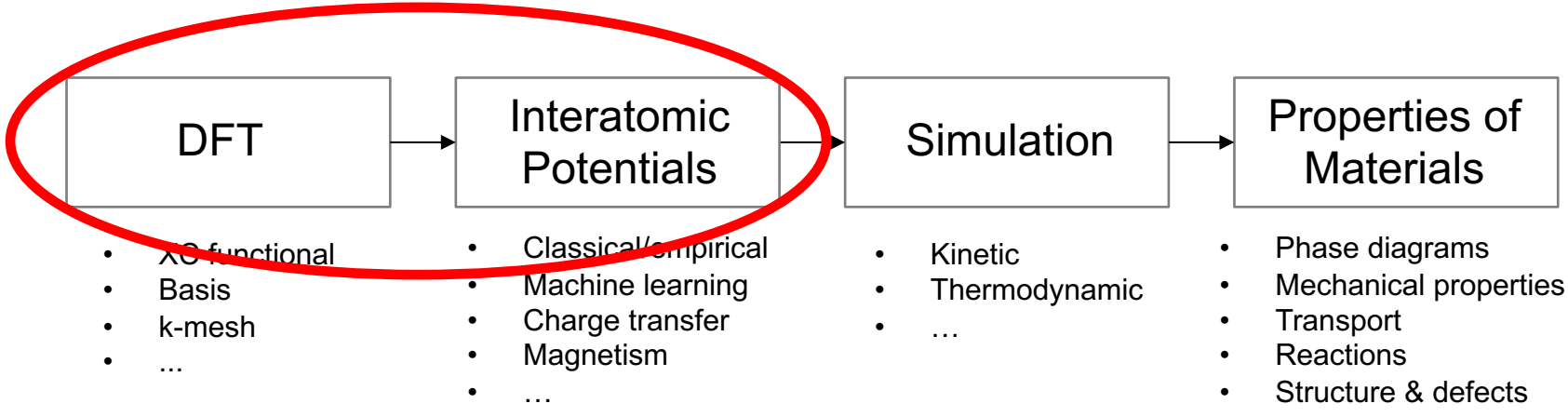
- Would like to go from DFT to properties within few days.

➜ Efficient workflow management is critical.

c.f. talks Jörg Neugebauer and Jan Janssen

pyiron
www.pyiron.org

Outline



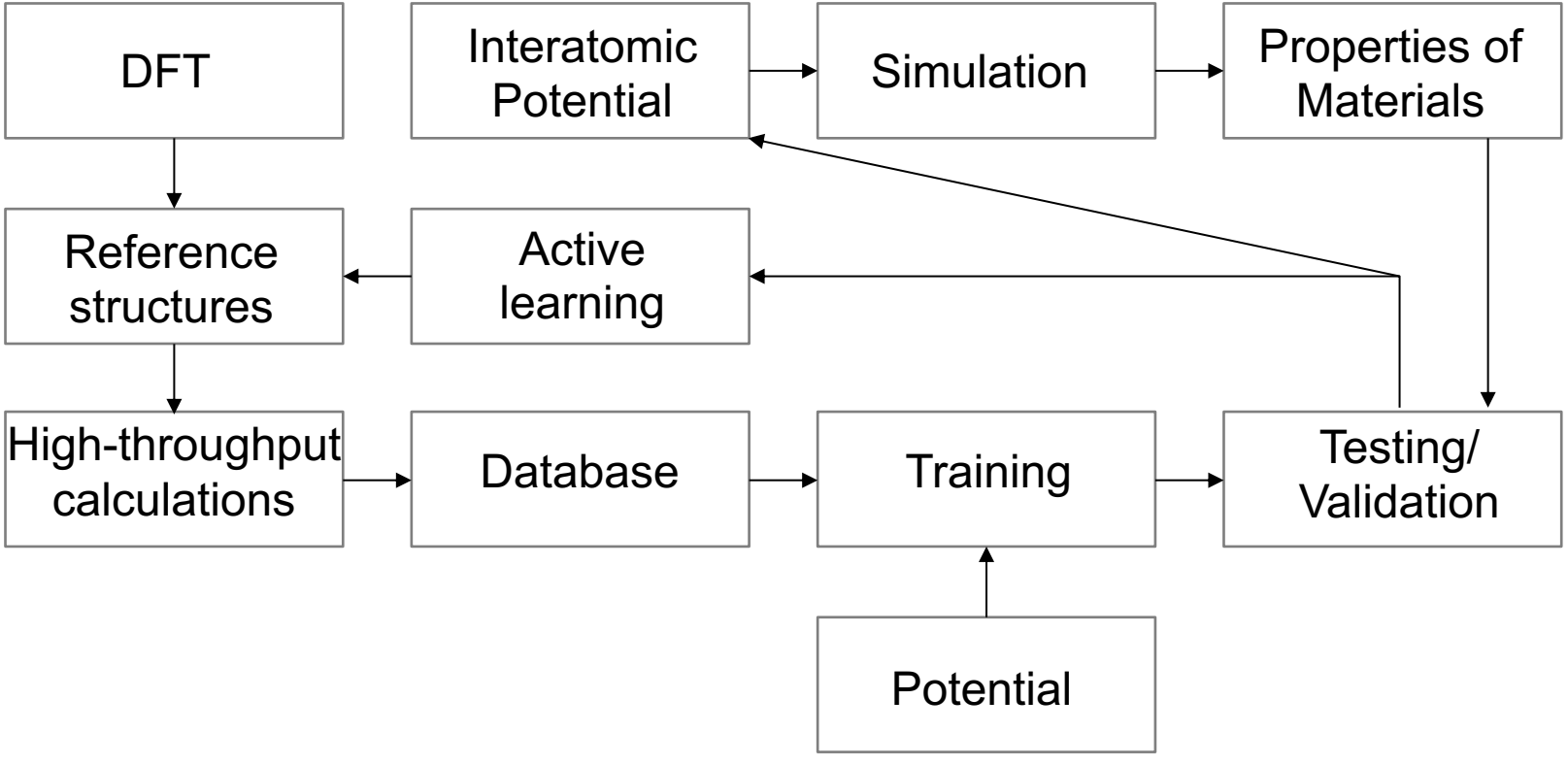
- Would like to go from DFT to properties within few days.

➜ Efficient workflow management is critical.

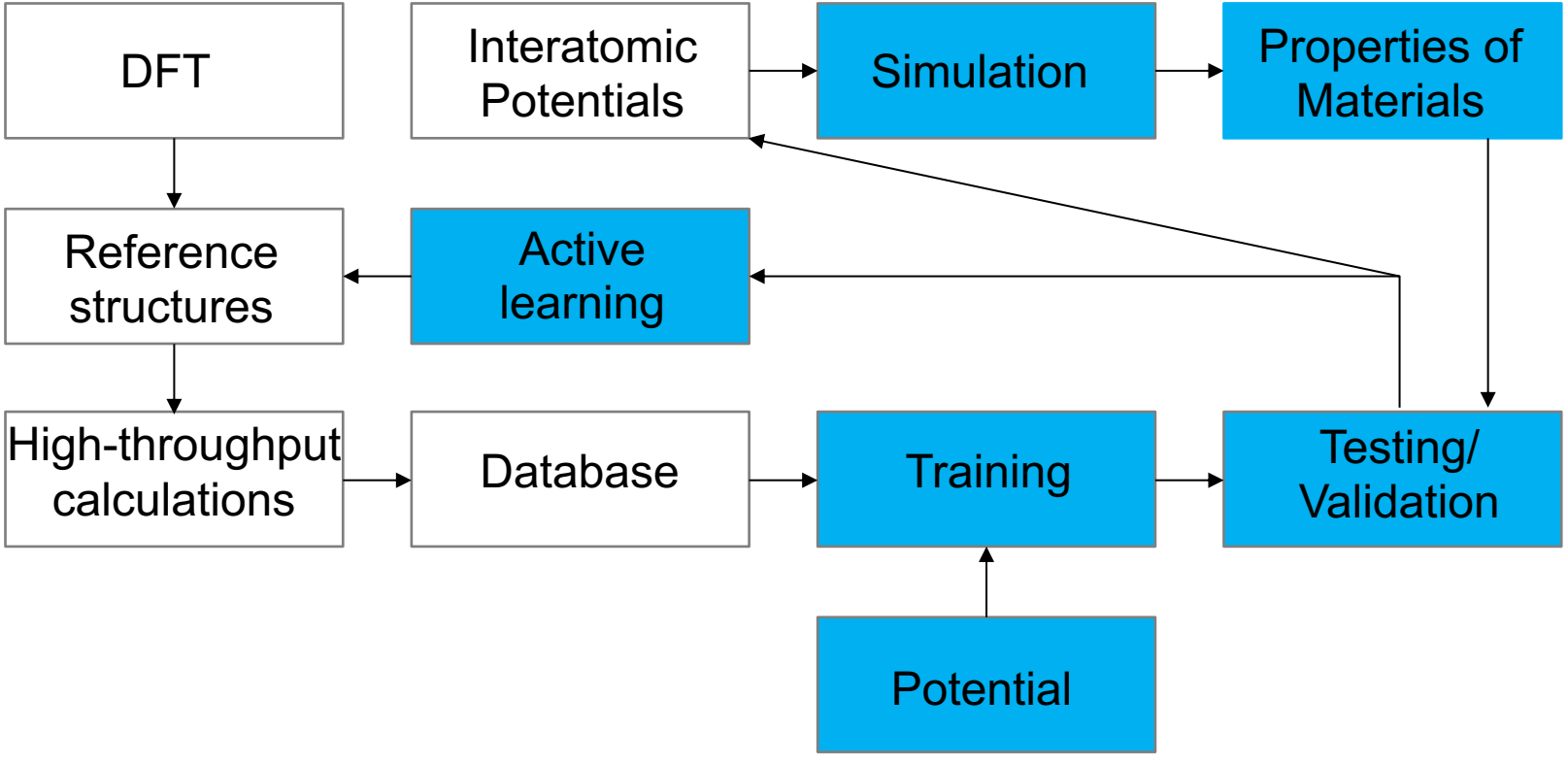
c.f. talks Jörg Neugebauer and Jan Janssen

pyiron
www.pyiron.org

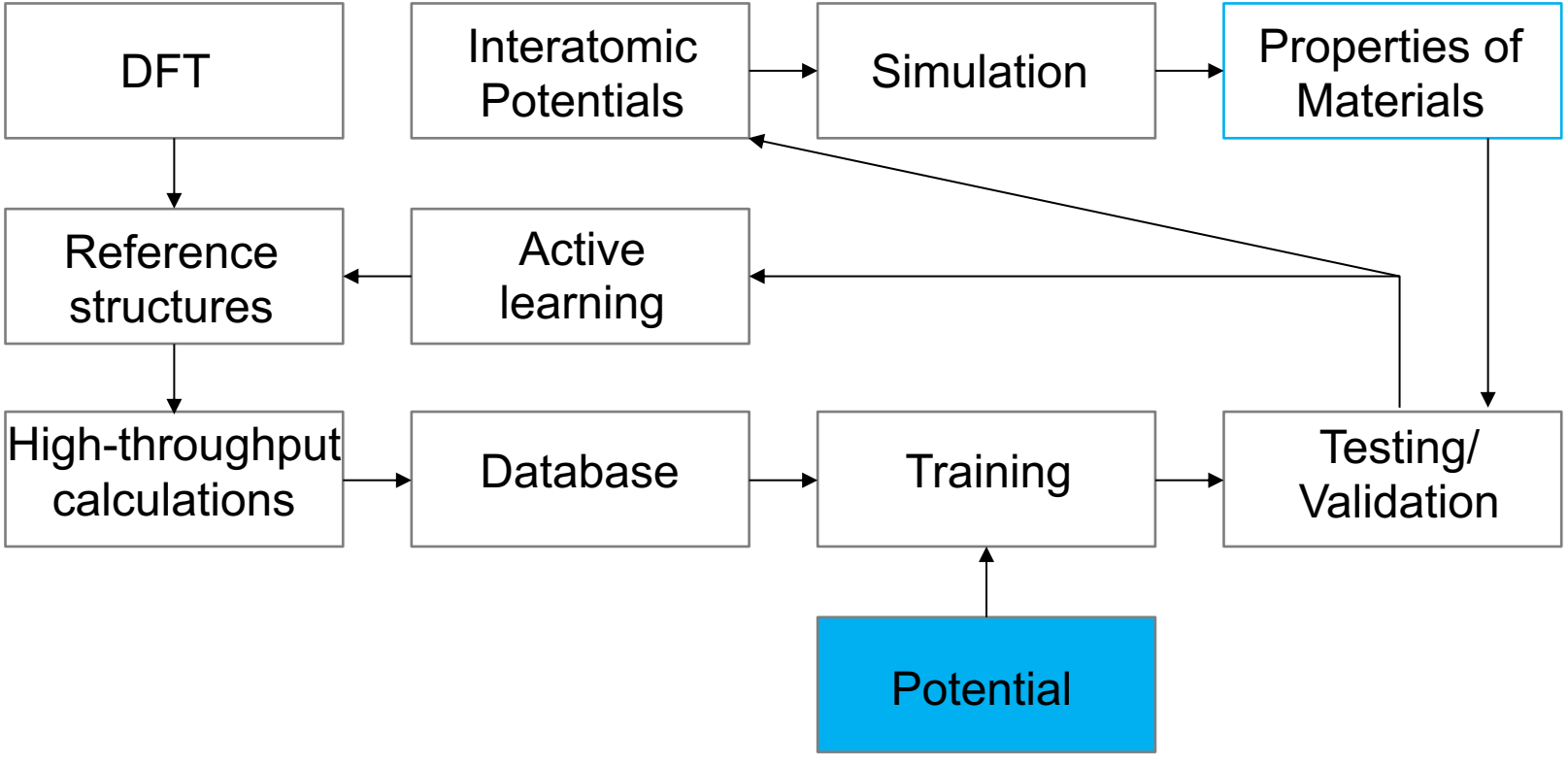
From DFT to Potentials



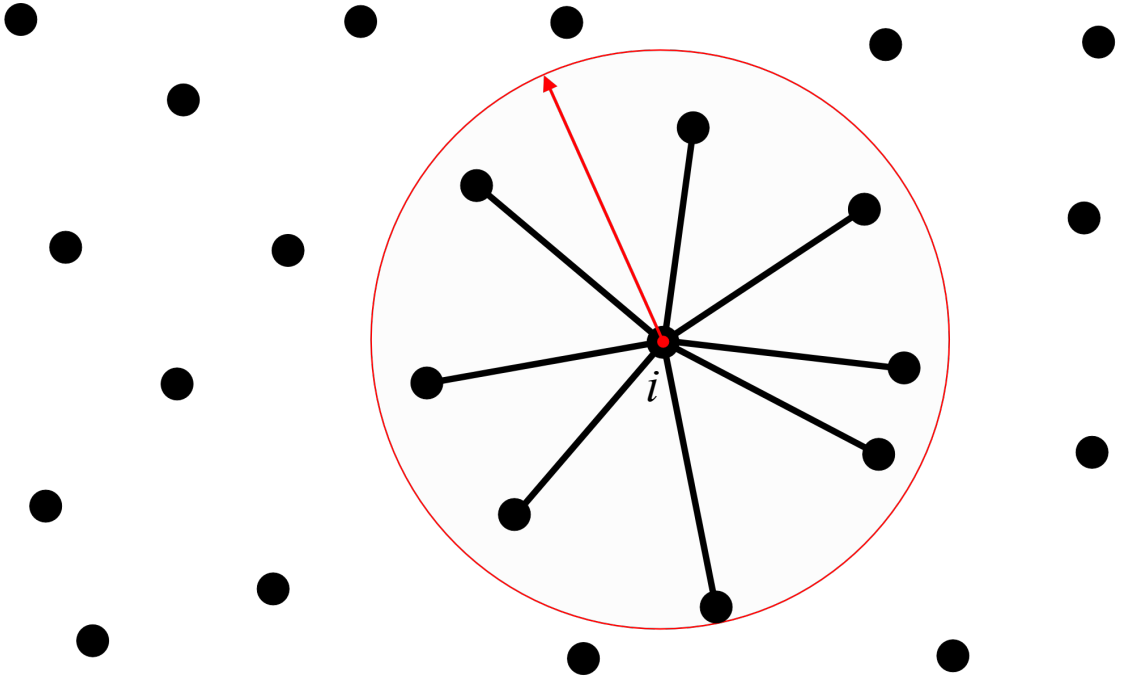
From DFT to Potentials



1. Interatomic potentials: Atomic Cluster Expansion



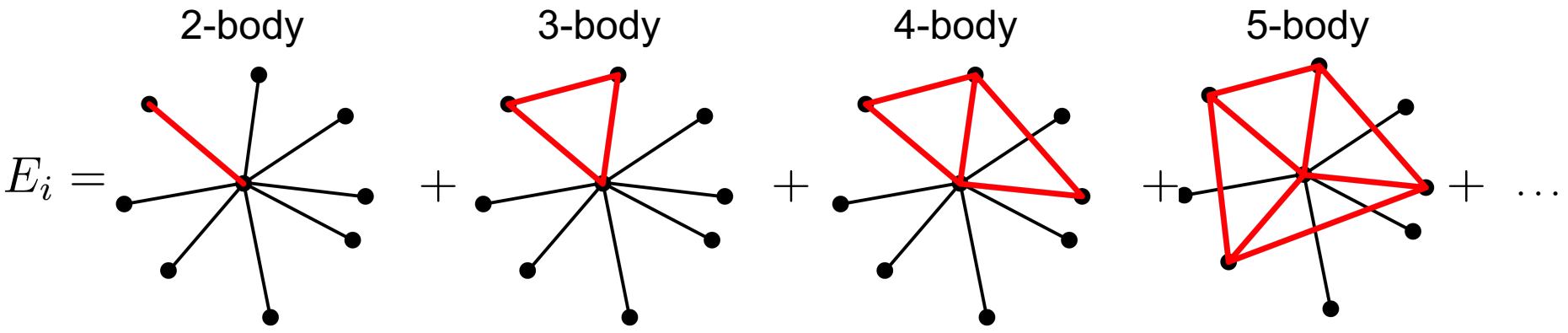
Locality



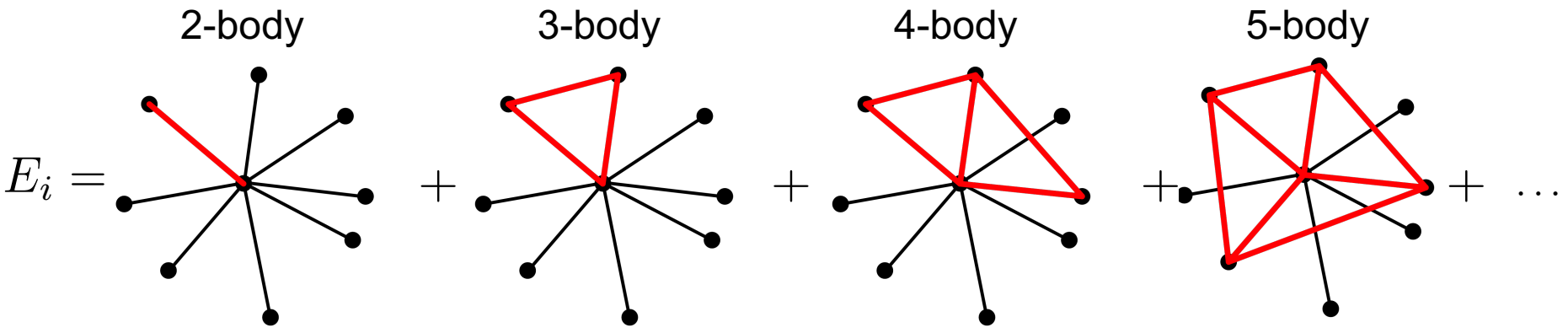
$$E = \sum_i E_i$$

➜ Focus on atomic energies

Naïve body-ordered expansion



Naïve body-ordered expansion



Example: 100 neighboring atoms

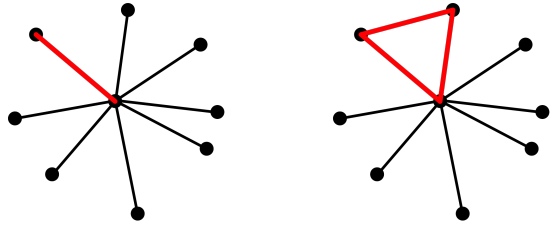
- Number of terms/operations:

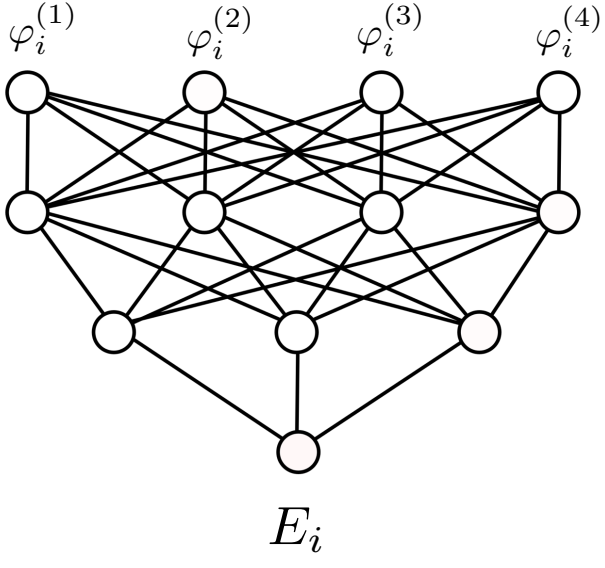
100	10000	1000000	100000000

➜ Computationally not efficient/feasible

Neural network potentials solution

- Limit to 2-body and 3-body contributions
- Determine higher order terms from HDNN



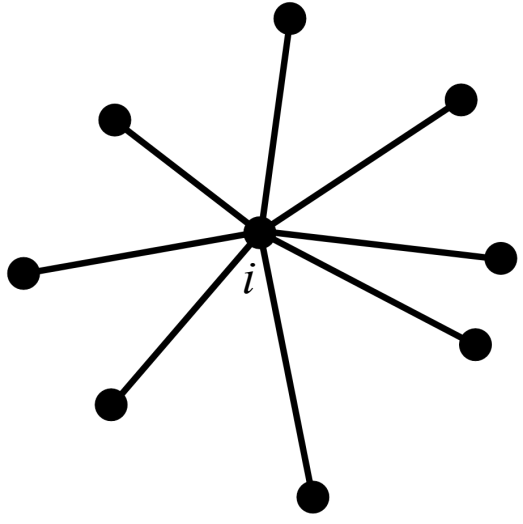


Behler and Parrinello, PRL 98 (2007) 146401

Atomic Cluster Expansion

- Atomic energy is fully characterized by vectors to all other atoms

$$\boldsymbol{\sigma} = \{\boldsymbol{r}_{1i}, \boldsymbol{r}_{2i}, \ldots, \boldsymbol{r}_{Ni}\}$$



➜ Atomic energy

$$E_i(\boldsymbol{\sigma}) = E_i(\boldsymbol{r}_{1i}, \boldsymbol{r}_{2i}, \ldots, \boldsymbol{r}_{Ni})$$

with $\boldsymbol{r}_{ji} = \boldsymbol{r}_j - \boldsymbol{r}_i$

Basis

- Inner product

$$\langle f|g\rangle = \int f^*(\boldsymbol{\sigma}) g(\boldsymbol{\sigma})\, d\boldsymbol{\sigma}$$

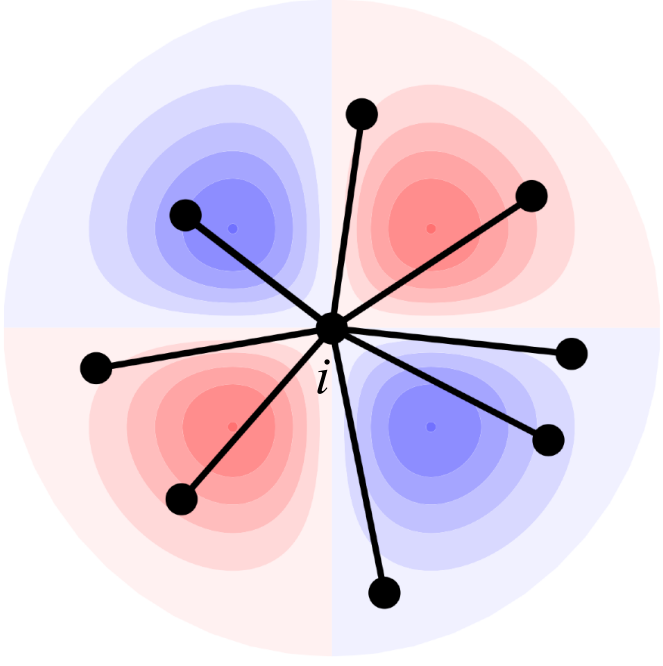
- Choose single-particle basis functions

- Orthonormal and complete

$$\langle \phi_v | \phi_u \rangle = \delta_{vu}$$

$$\sum_v |\phi_v\rangle\langle\phi_v| = \delta(\boldsymbol{r} - \boldsymbol{r'})$$

- Each basis function $\phi_v(\boldsymbol{r}_{ji})$ depends only on single bond vector



Cluster Expansion

- Cluster basis function

$$\Phi_{\boldsymbol{v}} = \phi_{v_1}(\boldsymbol{r}_{j_1 i})\phi_{v_2}(\boldsymbol{r}_{j_2 i})\phi_{v_3}(\boldsymbol{r}_{j_3 i})\ldots$$

- Orthogonality and completeness follow

$$\langle \Phi_{\boldsymbol{v}} | \Phi_{\boldsymbol{u}} \rangle = \delta_{\boldsymbol{vu}}$$

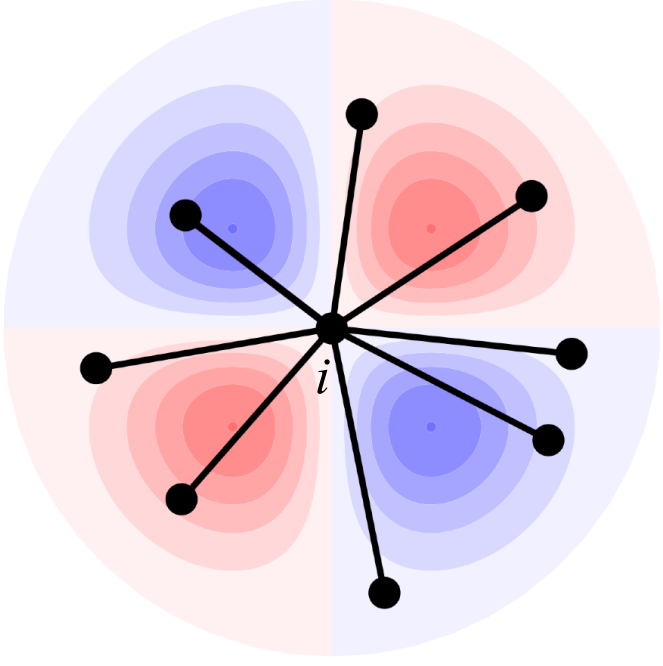
$$\sum_{\boldsymbol{v}} |\Phi_{\boldsymbol{v}}\rangle\langle\Phi_{\boldsymbol{v}}| = \delta(\boldsymbol{\sigma} - \boldsymbol{\sigma'})$$

- Cluster expansion

$$E_i(\boldsymbol{\sigma}) = \sum_{\boldsymbol{v}} J_{\boldsymbol{v}} \Phi_{\boldsymbol{v}}$$

- Expansion coefficients by projection

$$J_{\boldsymbol{v}} = \langle \Phi_{\boldsymbol{v}} | E_i \rangle$$



Drautz, Phys. Rev. B 99 (2019) 014104

Complexity

- Cluster basis function can be computed efficiently

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N}\phi_i\phi_j\phi_k\phi_l \quad = \quad \left(\sum_{i=1}^{N}\phi_i\right)^4$$

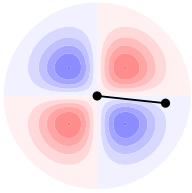
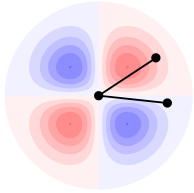
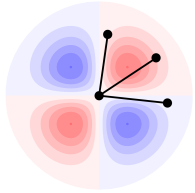
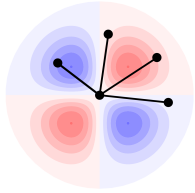
100000000 operations

100 operations

- High body order is no longer a problem
- Recursive evaluation: one operation ➜ one basis function

➜ fast and accurate representations possible

Density trick, recursive evaluation

$E =$  $+ \ldots +$  $+ \ldots +$  $+ \ldots +$  $+ \ldots$

- Atomic base

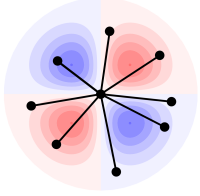
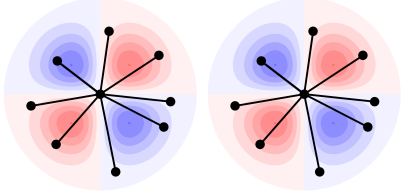
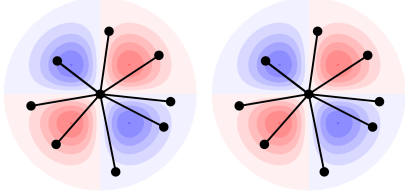
$$A_v = \sum_j \phi_v(\boldsymbol{r}_{ji})$$

- Cluster expansion ➔ polynomial

$$E_i(\boldsymbol{\sigma}) = \sum_{\mathbf{v}} \tilde{c}_{\mathbf{v}} \boldsymbol{A}_{\mathbf{v}} \quad \text{with} \quad \boldsymbol{A}_{\mathbf{v}} = A_{v_1} \ldots A_{v_N}$$

$\text{CPU time} \propto N^{|\boldsymbol{v}|}$ — leading order; number of neighbors

$\text{CPU time} \propto N$ — fast ✓; high body-order ✓

$E =$  $+$  $+$  $+ \ldots$

TRIP covariance

- Employ LCAO basis (as irreducible basis of rotation group)

$$\phi_v(\boldsymbol{r}) = R_{nl}(r) Y_l^m(\hat{\boldsymbol{r}})$$

- Atomic base as before

$$A_v = \langle \rho_i | \phi_v \rangle = \sum_j \phi_v(\boldsymbol{r}_{ji})$$

- Rotationally covariant basis functions

$$\boldsymbol{B} = \mathcal{C}\boldsymbol{A}$$

Generalized Clebsch-Gordan coefficients

- **Atomic cluster expansion**

$$E_i(\boldsymbol{\sigma}) = \sum_{\boldsymbol{v}} c_{\boldsymbol{v}} \boldsymbol{B}_{\boldsymbol{v}}$$

➔ Complete expansion for scalar, vectorial or tensorial properties

TRIP:
- Translation
- Rotation
- Inversion
- Permutation

Drautz, PRB 99 (2019) 014104; Dusson, et al, arXiv 1911.03550 (2020)

Other descriptors and potentials

- Many other potentials and descriptors can be cast in the form of ACE
  - Steinhardt parameters
    Steinhardt, Nelson, and Ronchetti, PRB 28, 784 (1983)
  - Symmetry functions
    Behler, J. Chem. Phys. 134, 074106 (2011)
  - Smooth Overlap of Atomic Positions (SOAP)
    Bartok, Kondor, and Csanyi, PRB 87, 184115 (2013)
  - Spectral Neighbor Analysis Potential (SNAP)
    Thompson, et al., J. Comp. Phys. 285, 316 (2015)
  - Moment Tensor Potential
    Shapeev, Multiscale Model. Simul. 14, 1153 (2016)
  - …

Representation of energy

- Use several atomic properties

$$\varphi_i^{(p)} = \sum_{\boldsymbol{v}} c_{\boldsymbol{v}}^{(p)} \boldsymbol{B}_{i\boldsymbol{v}}$$

- Energy from non-linear function

$$E_i = \mathcal{F}(\varphi_i^{(1)}, \ldots, \varphi_i^{(P)})$$

- Choice of non-linear function $\mathcal{F}$: double convergence

SUB:
- Scale
- Universal
- Basis

TRIP:
- Translation
- Rotation
- Inversion
- Permutation

➔ SUB-TRIP covariant representation of energy

Representation of energy

$$E_i = \varphi_i^{(1)} + \sqrt{\varphi_i^{(2)}}$$

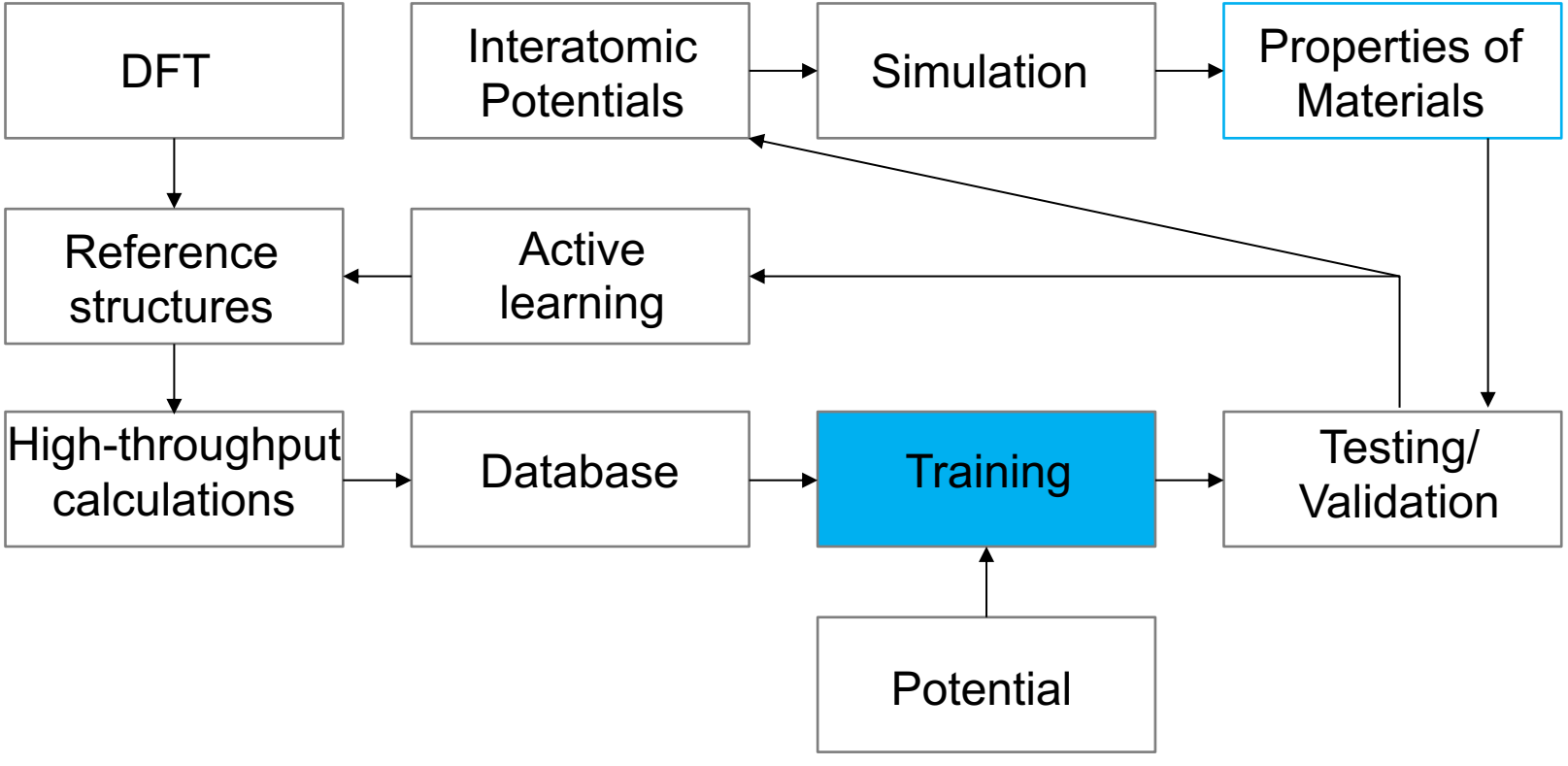
Atomic Cluster Expansion

Extension of Finnis-Sinclair/
Embedded Atom Method Potential

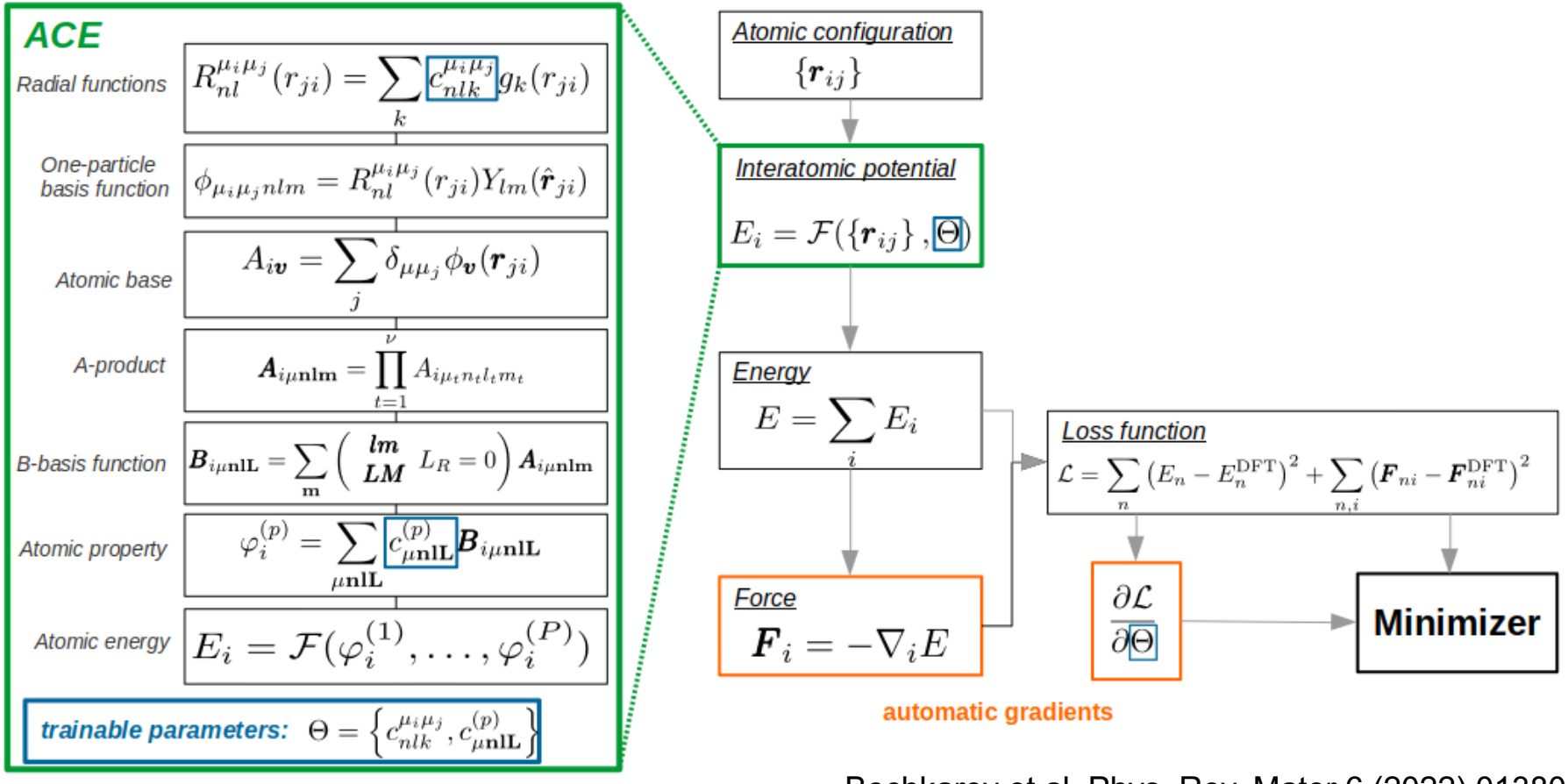
- Two complete ACE descriptors
- Physics-motivated mild non-linearity
- Universal and scale-invariant

Drautz, PRB 99 (2019) 014104

2. Training: PACEmaker

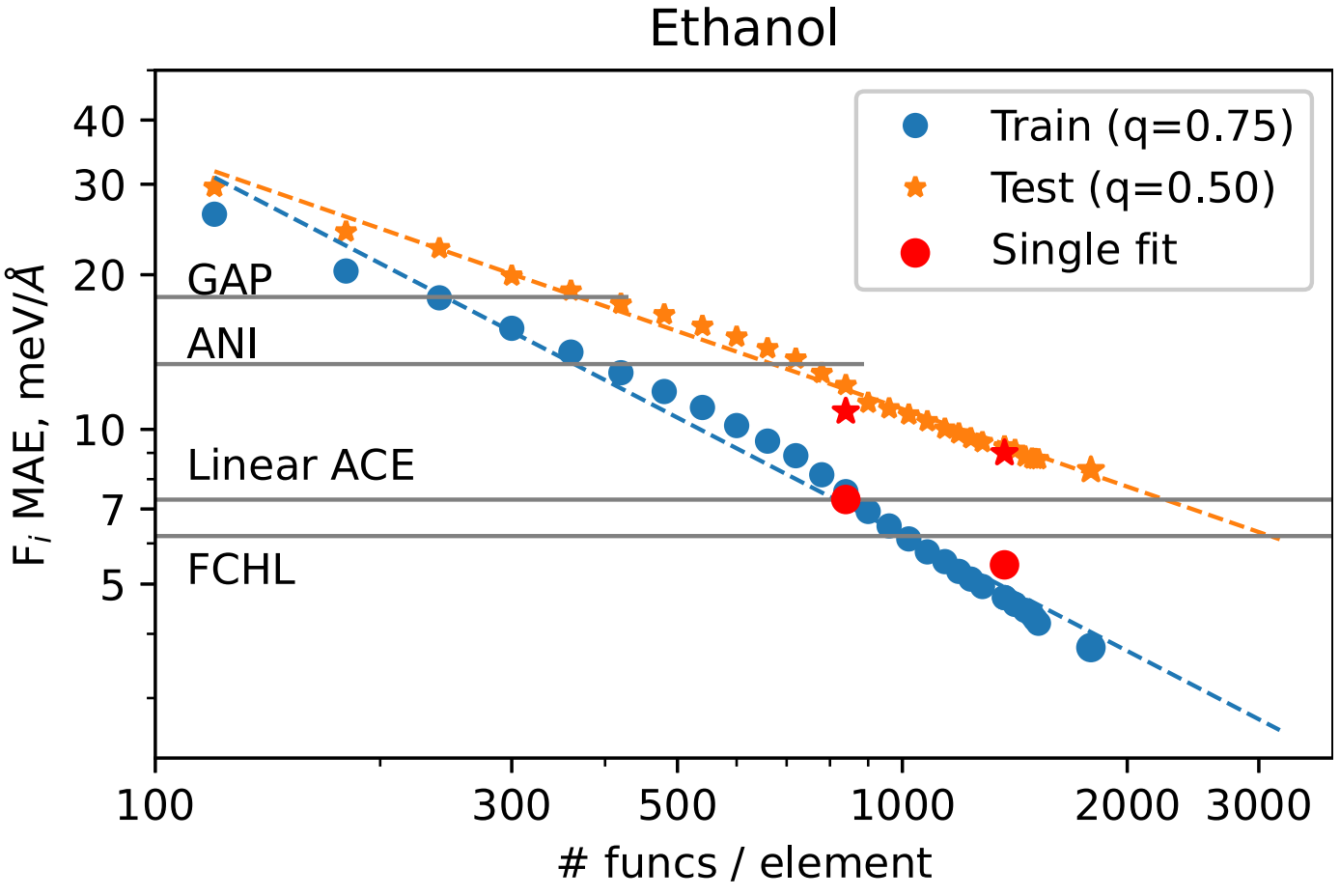


Training: PACEmaker workflow

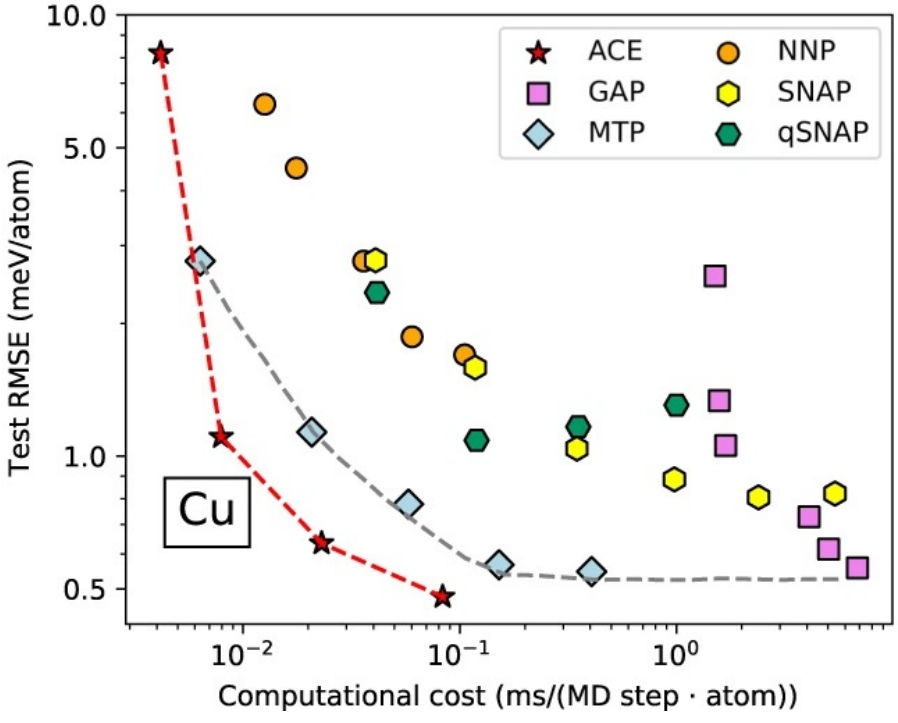


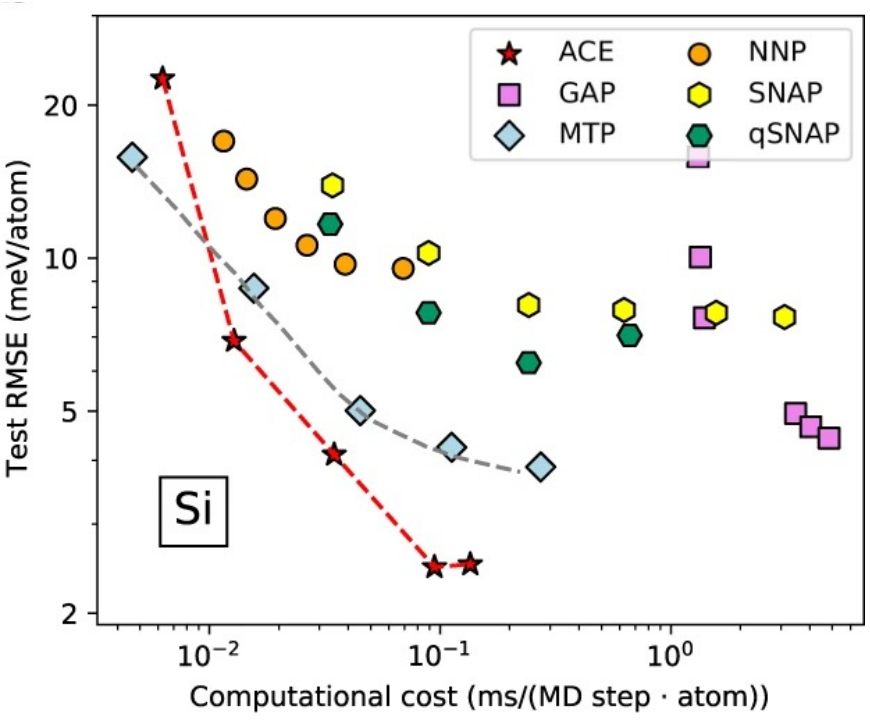
Bochkarev et al, Phys. Rev. Mater 6 (2022) 013804

Training: feature curve



Accuracy – Efficiency Pareto Front



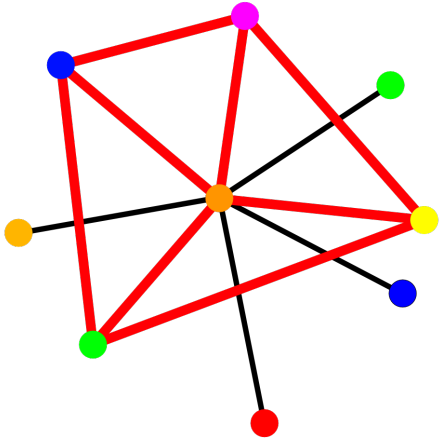


Zuo, et al., J. Phys. Chem. A 124 (2020) 731

Lysogorskiy et al., npj Comput. Mater. 7 (2021) 97

Chemical Complexity

- Assume different chemical species in multi-body interactions



For five elements
$5^5 = 3125$
times more effort than
for a single element

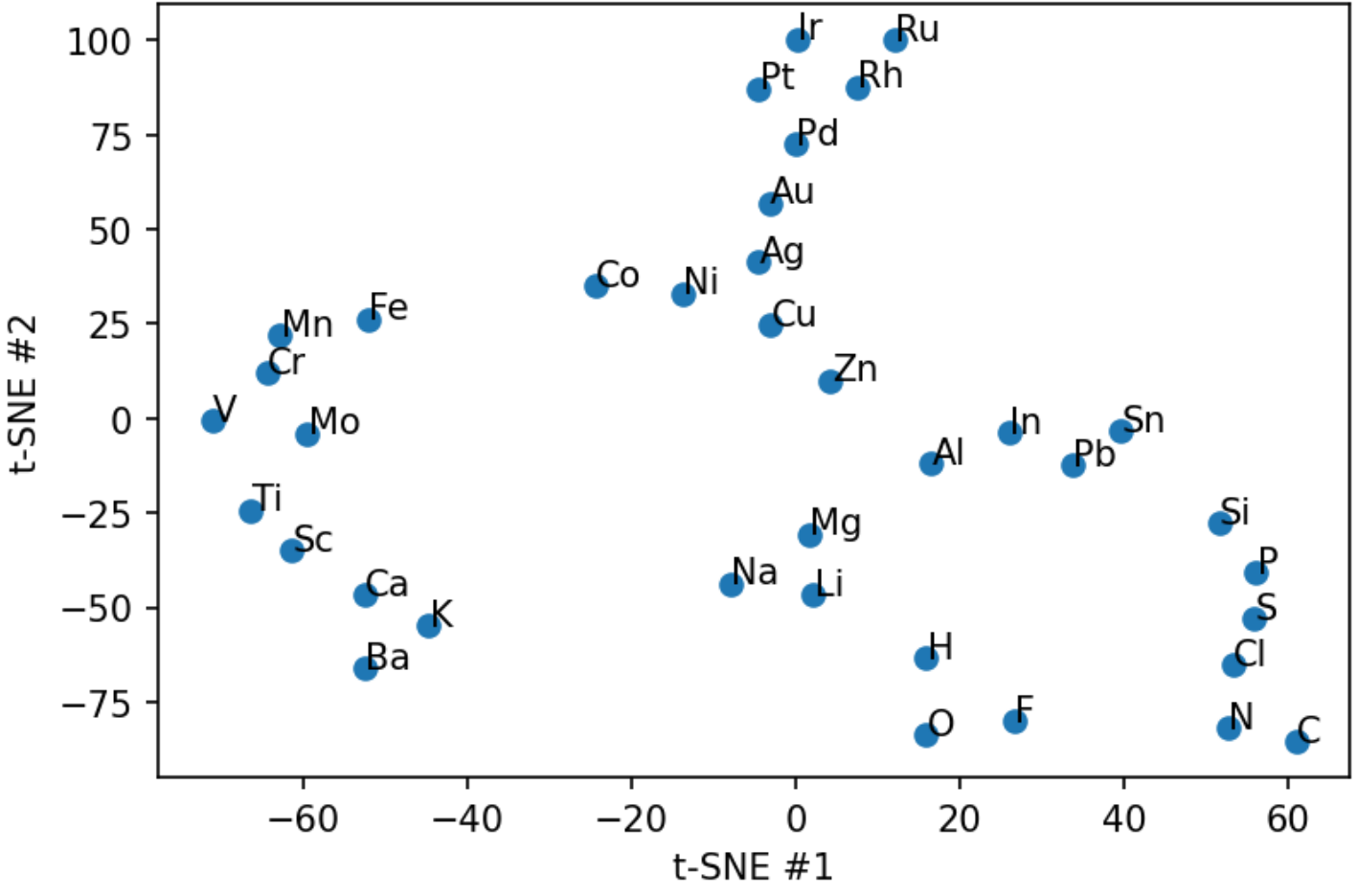
- Decompose into low-rank tensors (exact decomposition)

$$c_{ijnm} = \sum_k \lambda_k v_j^{(k)} v_j^{(k)} v_n^{(k)} v_m^{(k)}$$

Darby et al.,arXiv:2210.01705

➜ multiple chemical elements possible

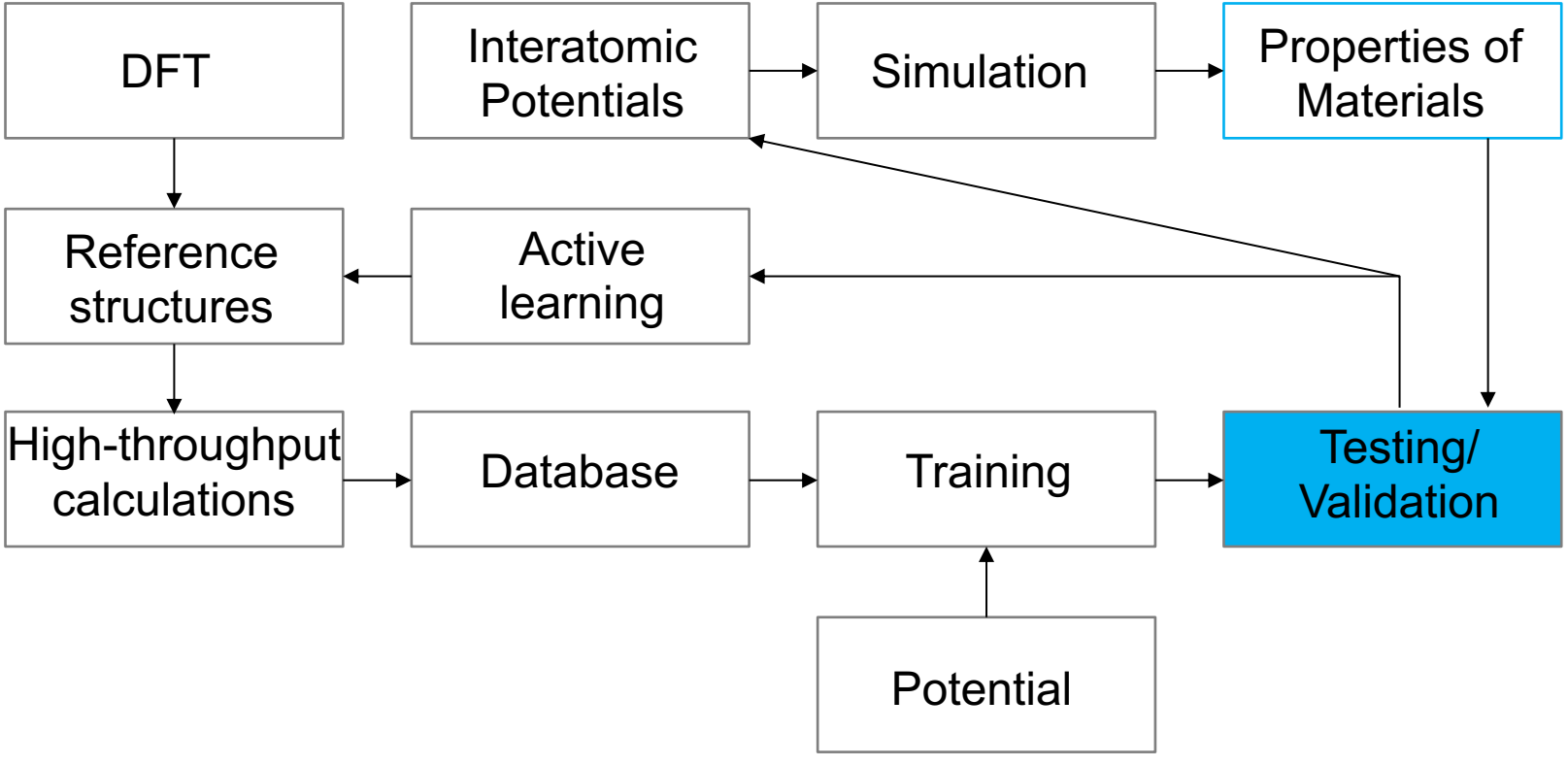
Multi-element parameterization



tSNE of ACE embeddings with 38 elements

Dataset from Takamoto et al., HME21 (2022)

3. Testing and validation

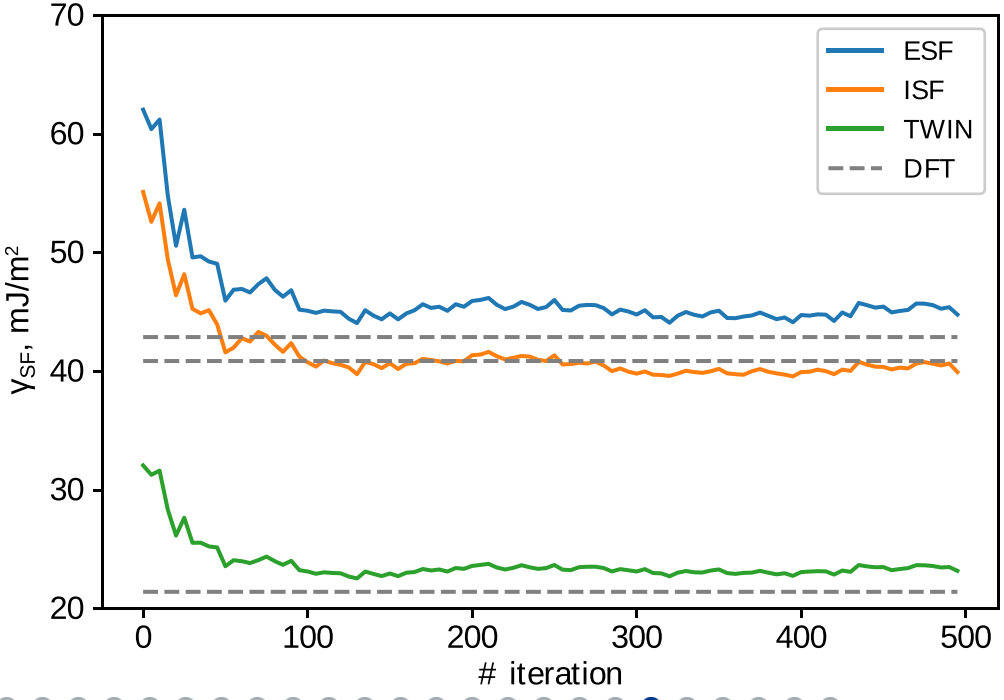


Testing and validation

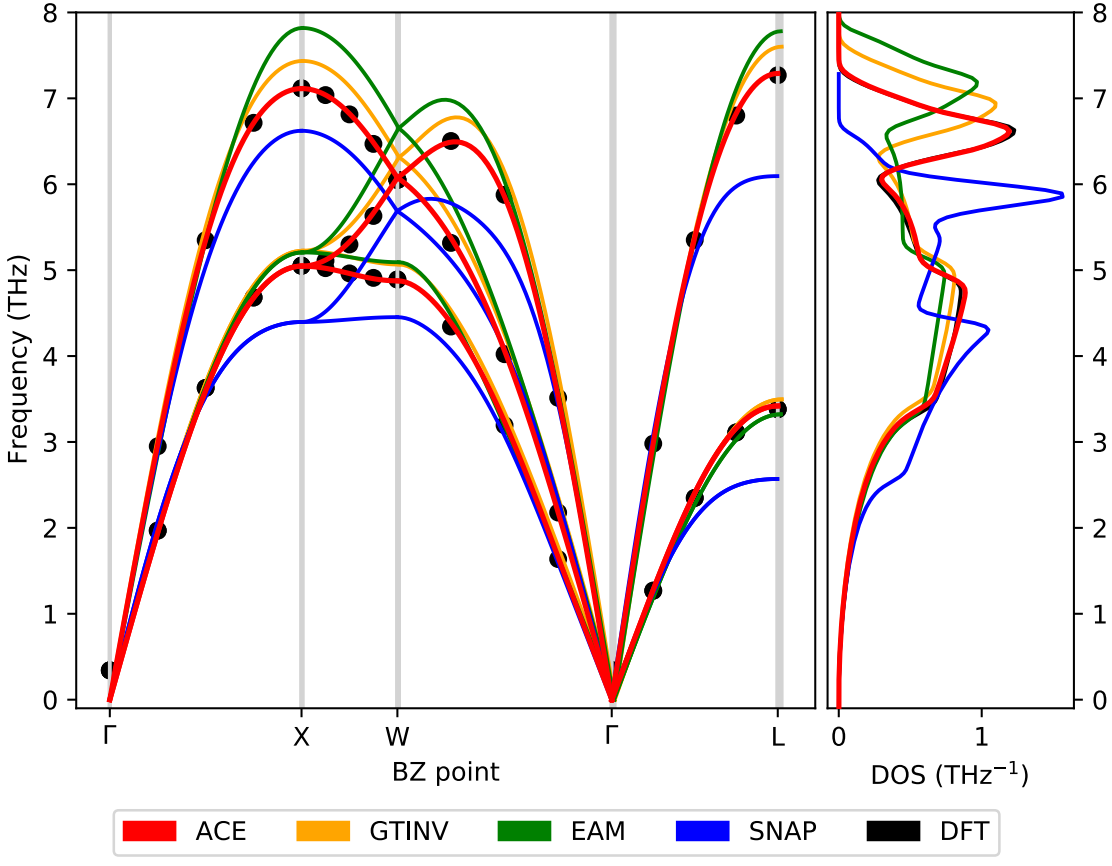
- Level I: comparison to DFT energies and forces
- Level II: comparison to DFT predicted properties
- Level III: validation by application

POTENTIALS collaboration

- On-the-fly validation of stacking fault energies in Cu during training

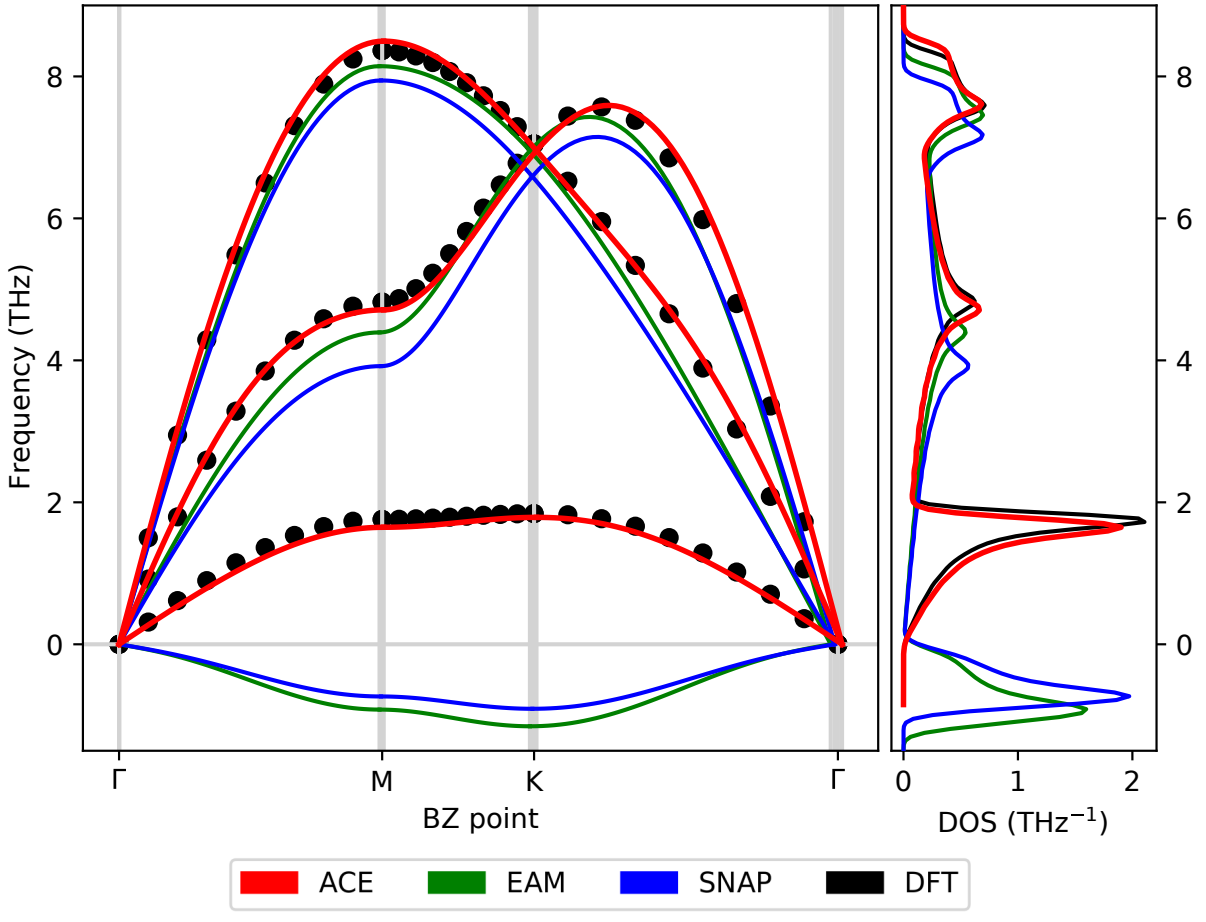


Testing and validation: phonons in copper

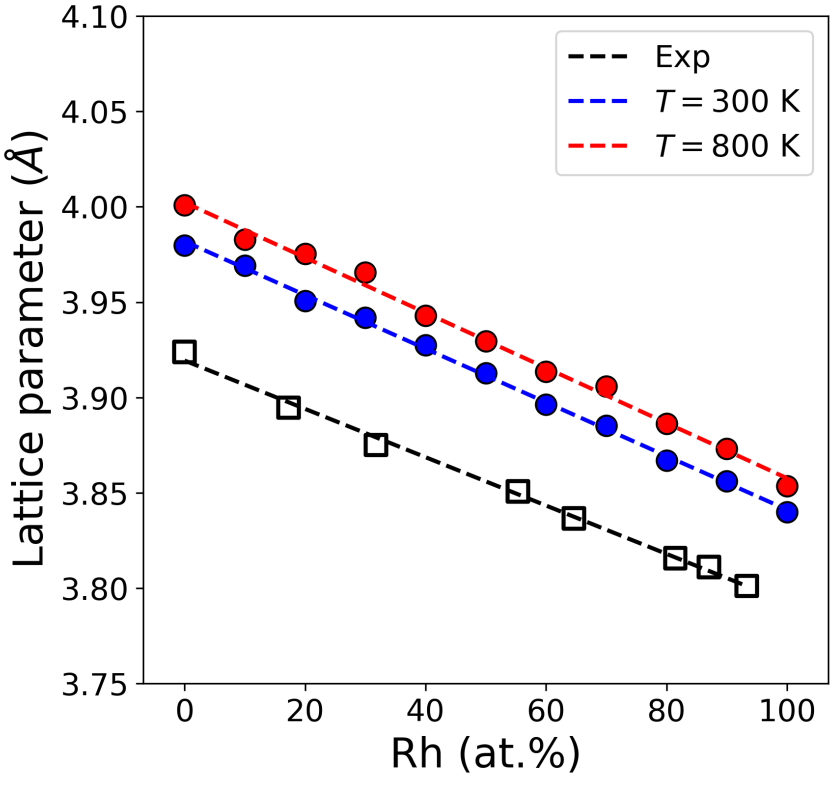


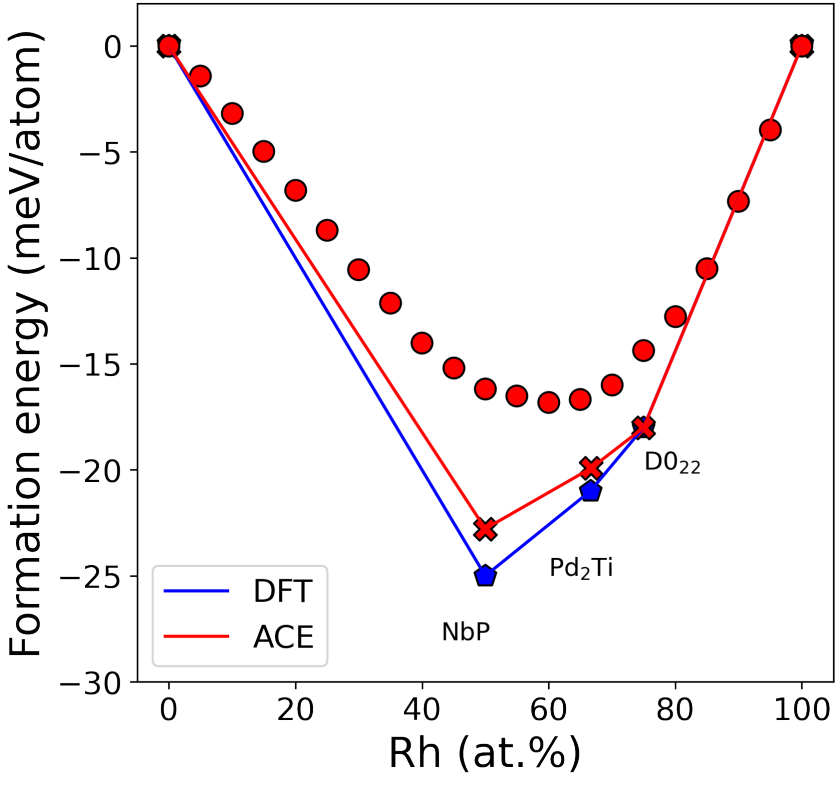
Lysogorskiy et al., npj Comput. Mater. 7 (2021) 97

Testing and validation: phonons in 2d structure

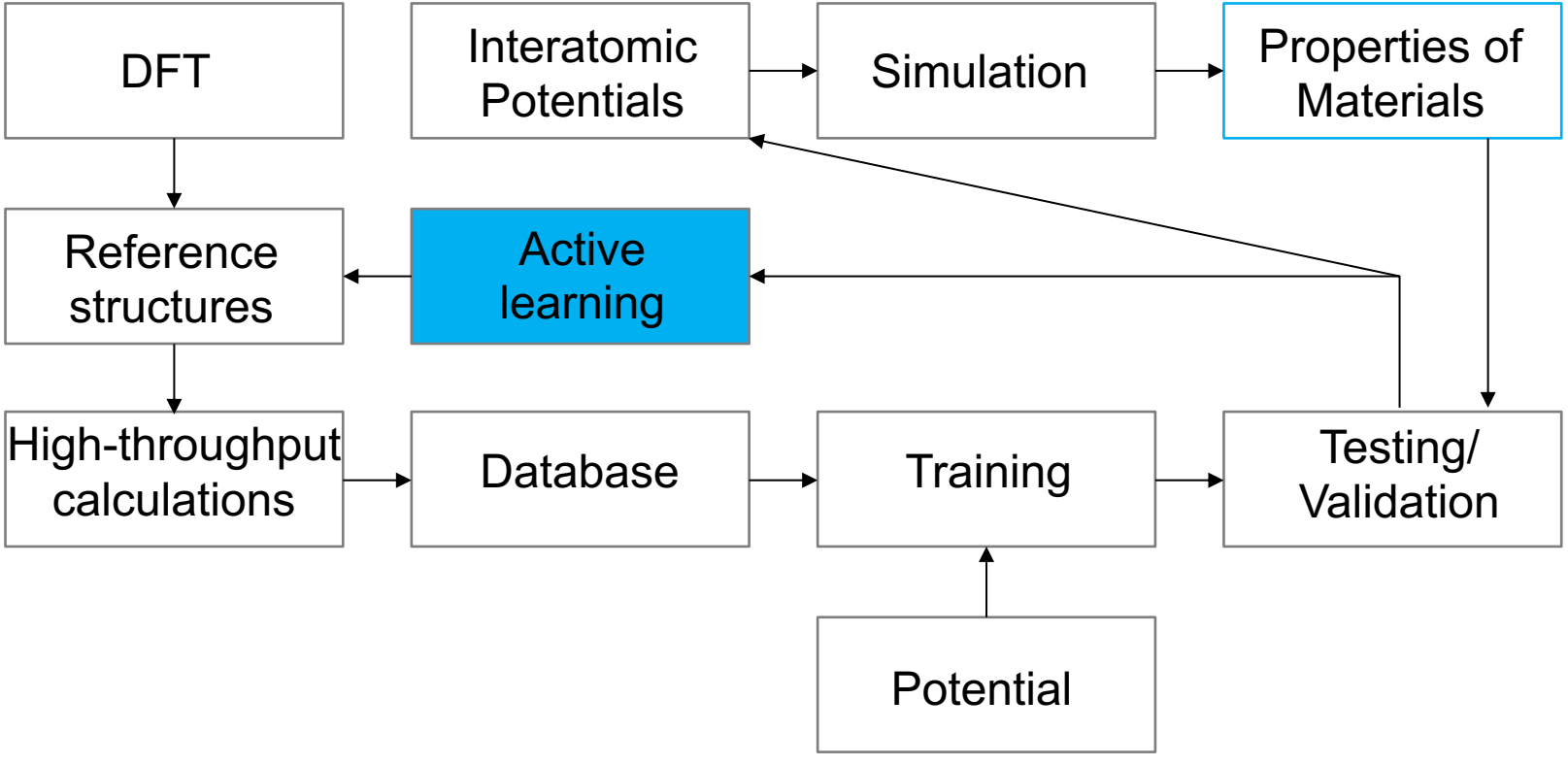


Testing and validation: Pt-Rh alloy



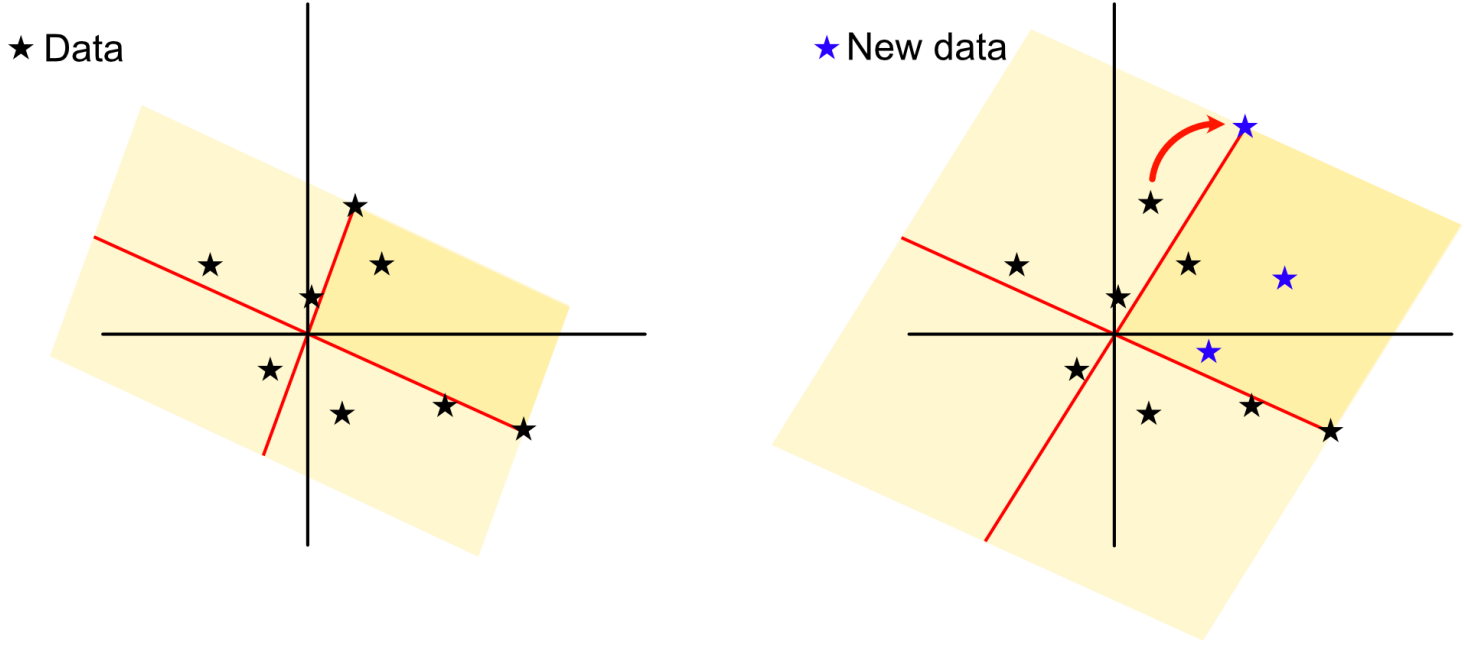


4. Active learning



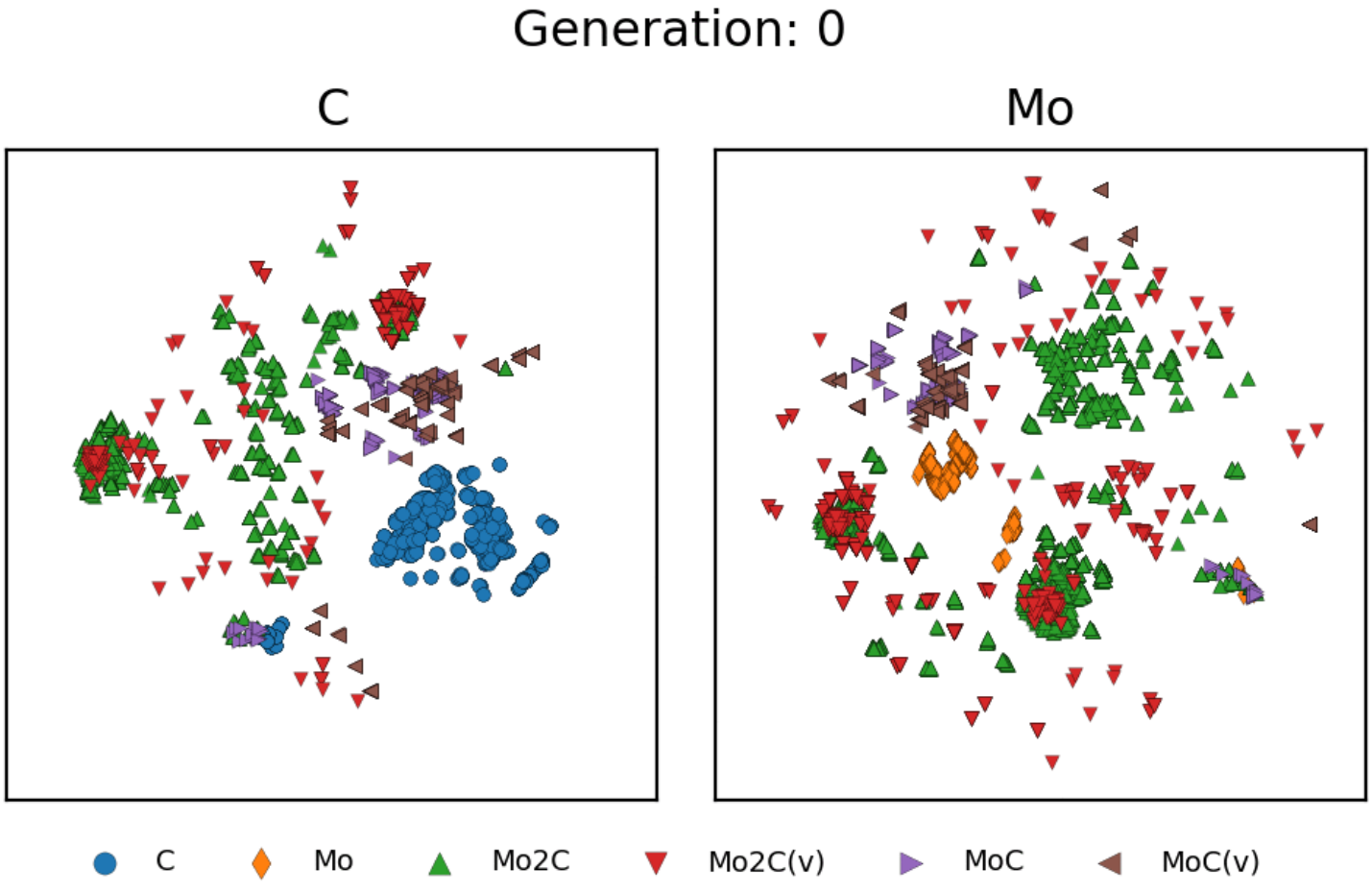
Active learning

- Uncertainty prediction based on D-optimality



Lysogorskiy et al., Phys. Rev. Mater. 7 (2023) 043801

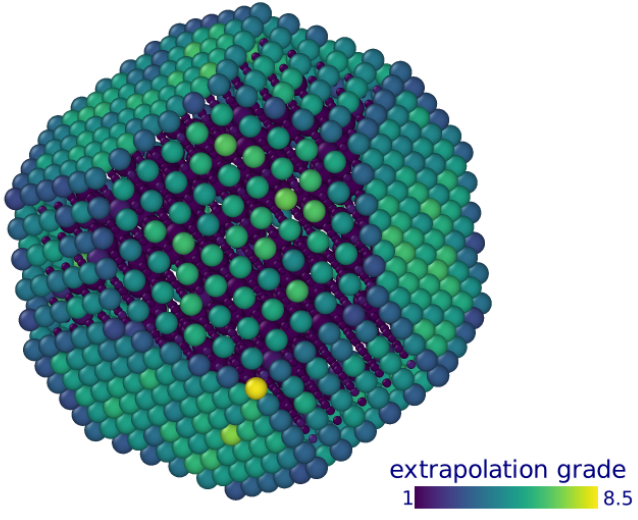
Active learning



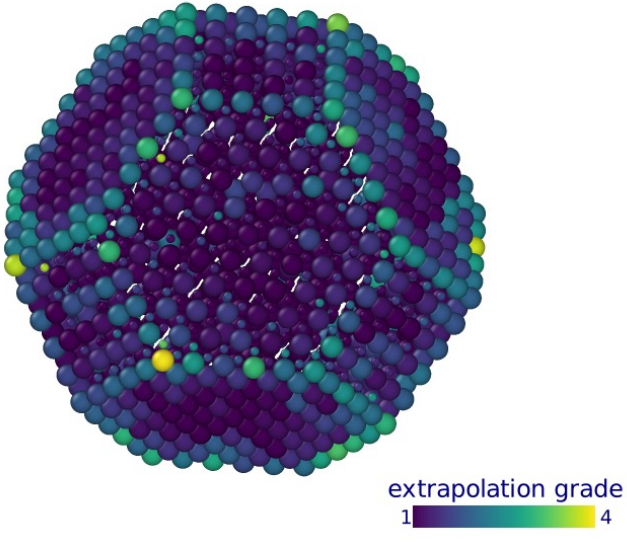
Active learning: Pt-Rh clusters

- Extrapolation grade

Before active learning



Updated reference data using active learning



➜ Extrapolation grade for robust uncertainties

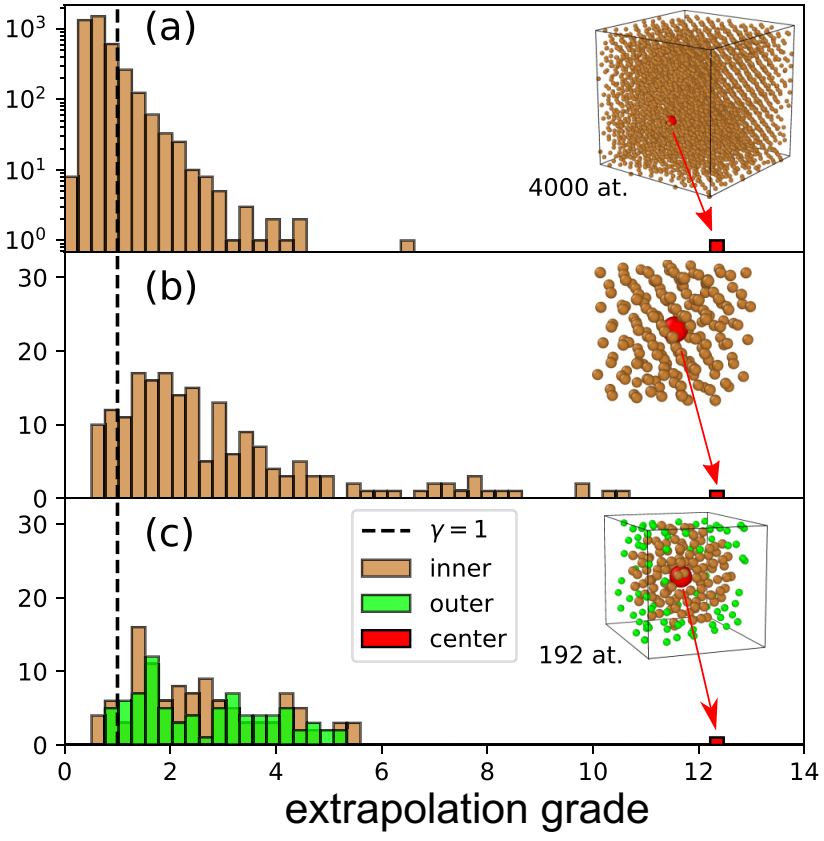
Active learning: training from large simulations

Workflow for training from large simulations

(a) Detection of atom with large extrapolation grade

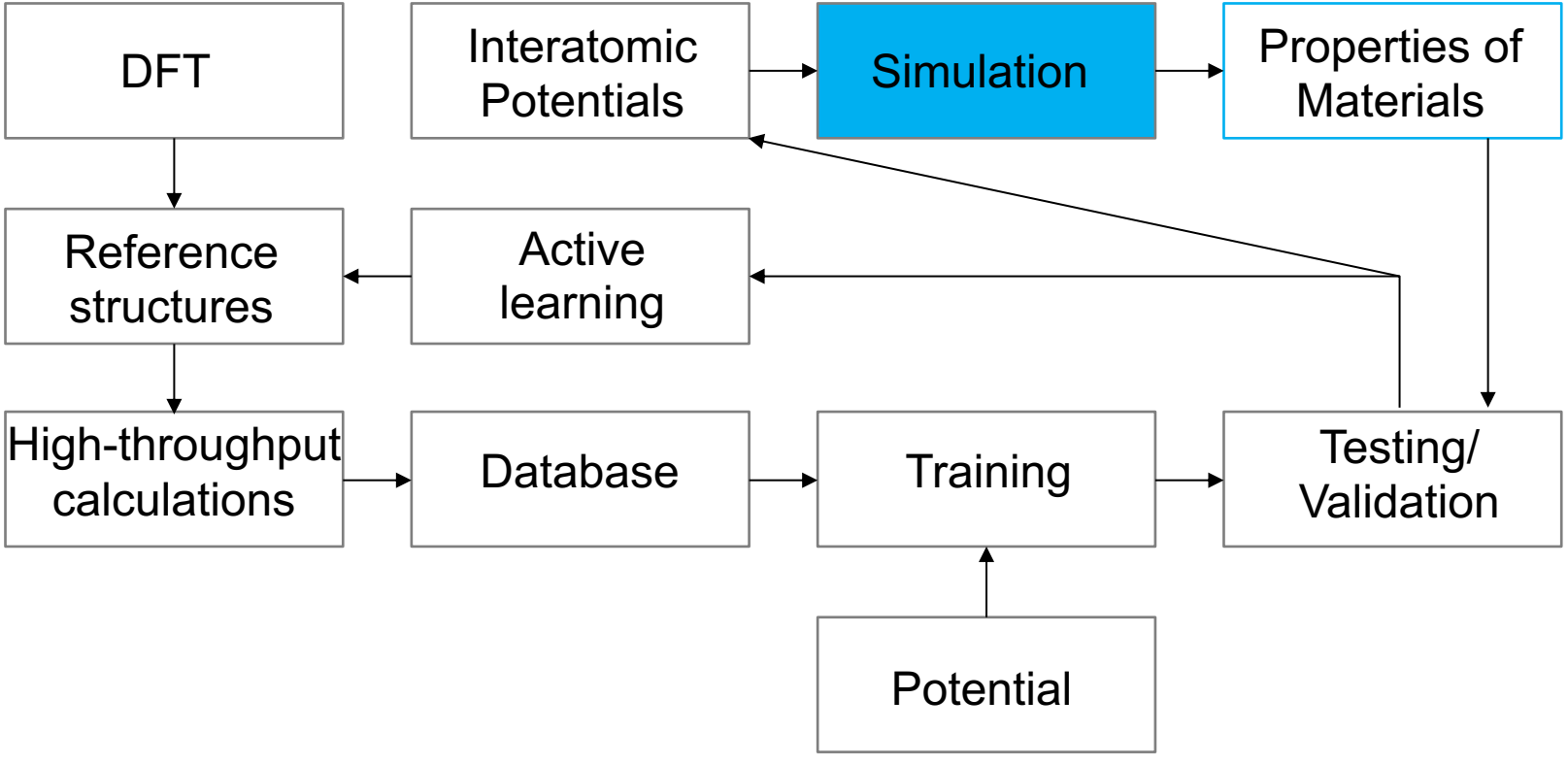
(b) Cutting out relevant atom

(c) Padding for DFT calculations



➜ Simulations with on-the-fly uncertainty prediction (available with LAMMPS)

5. Simulation



Simulation

- Molecular dynamics
- Monte Carlo and combined MD/MC
- Free energy computation

- Heavily rely on LAMMPS
- CALPHY for free energy and phase diagrams

CALPHY workflow for non-equilibrium free energy computation

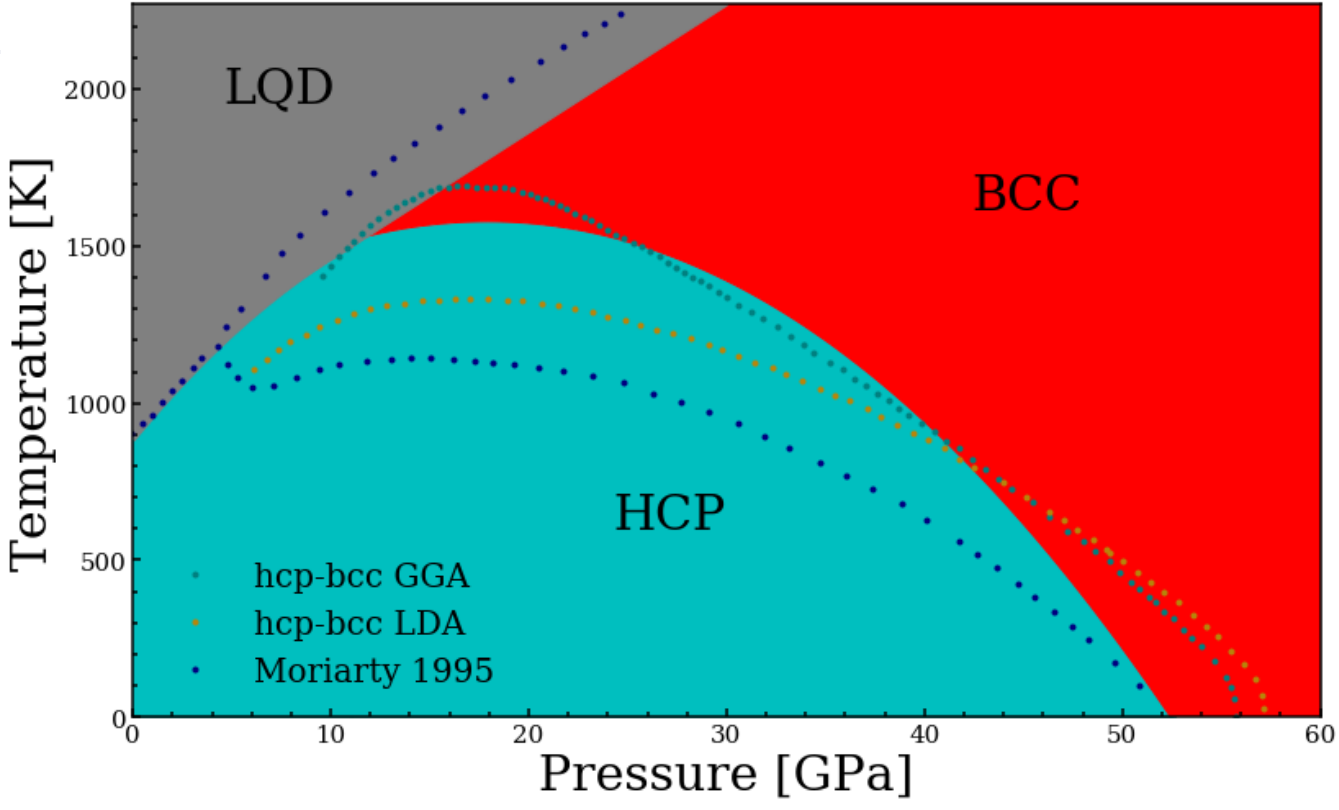
if constant V **then**
  $F(N, V, T_i)$ from Algorithm 1
else if constant P
  $F(N, V, T_i)$ from Algorithm 1
  calculate $G(N, P, T_i) = F(N, V, T_i) + PV_i$
if n independent runs **then**
  constant V
    equilibrate for time t_{eq} in NVT ensemble
    switch $\lambda : 1 \rightarrow T_i/T_f$ over time t_{sw}
    calculate work $W^s_{i\rightarrow f}$ [Eq. (C3)]
    equilibrate for time t_{eq} in NVT ensemble
    switch $\lambda : T_i/T_f \rightarrow 1$ over time t_{sw}
    calculate work $W^s_{f\rightarrow i}$ [Eq. (C3)]
else if constant P **then**
    equilibrate for time t_{eq} in NPT ensemble
    switch $\lambda : 1 \rightarrow T_i/T_f$ over time t_{sw}
    calculate work $W^s_{i\rightarrow f}$ [Eq. (C6)]
    equilibrate for time t_{eq} in NPT ensemble
    switch $\lambda : T_i/T_f \rightarrow 1$ over time t_{sw}
    calculate work $W^s_{f\rightarrow i}$ [Eq. (C6)]
if constant V
  average over n independent runs $\Delta F = \frac{1}{2}(W^s_{i\rightarrow f} - W^s_{f\rightarrow i})$
  calculate $F(N, V, T_f) = F(N, V, T_i) - \frac{3}{2}k_B T_f N \ln \frac{T_f}{T_i} + \frac{T_f}{T_i}\Delta F$
else if constant P **then**
  average over n independent runs $\Delta G = \frac{1}{2}(W^s_{i\rightarrow f} - W^s_{f\rightarrow i})$
  calculate $G(N, P, T_f) = G(N, P, T_i) - \frac{3}{2}k_B T_f N \ln \frac{T_f}{T_i} + \frac{T_f}{T_i}\Delta G$
  calculate S and C_P using Eqs. (13) and (14)

Menon et al, Phys. Rev. Mater 5 (2021) 103801

Simulation: ACE Mg phase diagram

- General purpose ACE
- Fitted to PBE data from FHIaims

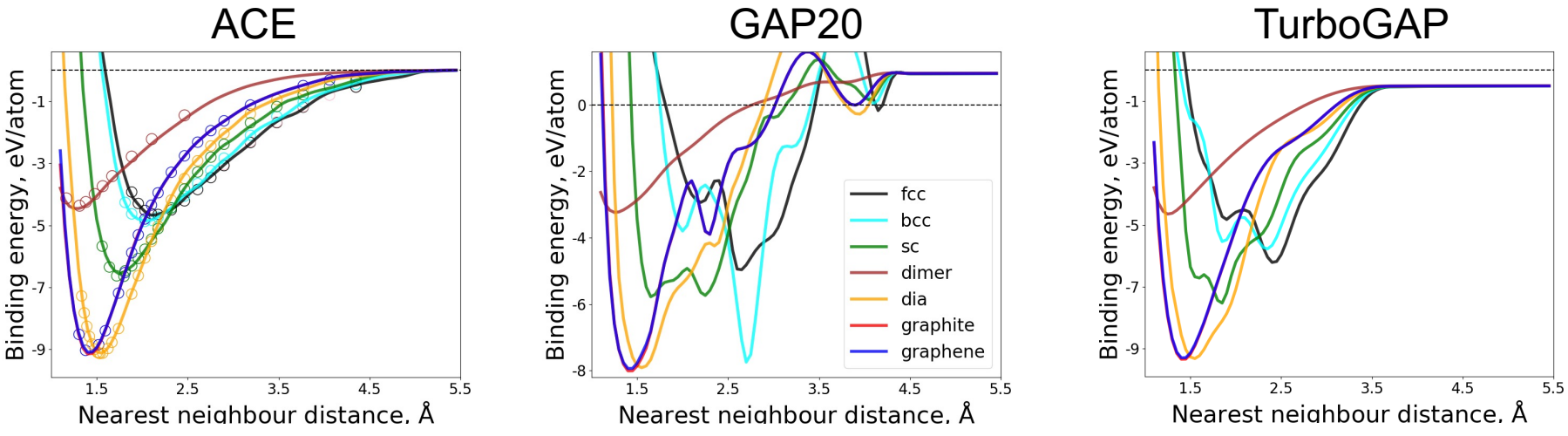
- Melting temperature
  862 K (ACE)
  923 K (EXP)



Moriarty and Althoff, PRB 51 (1995) 5609
GGA/LDA: Mehta, Price, Alfè, J. Chem. Phys. 125 (2006) 194507

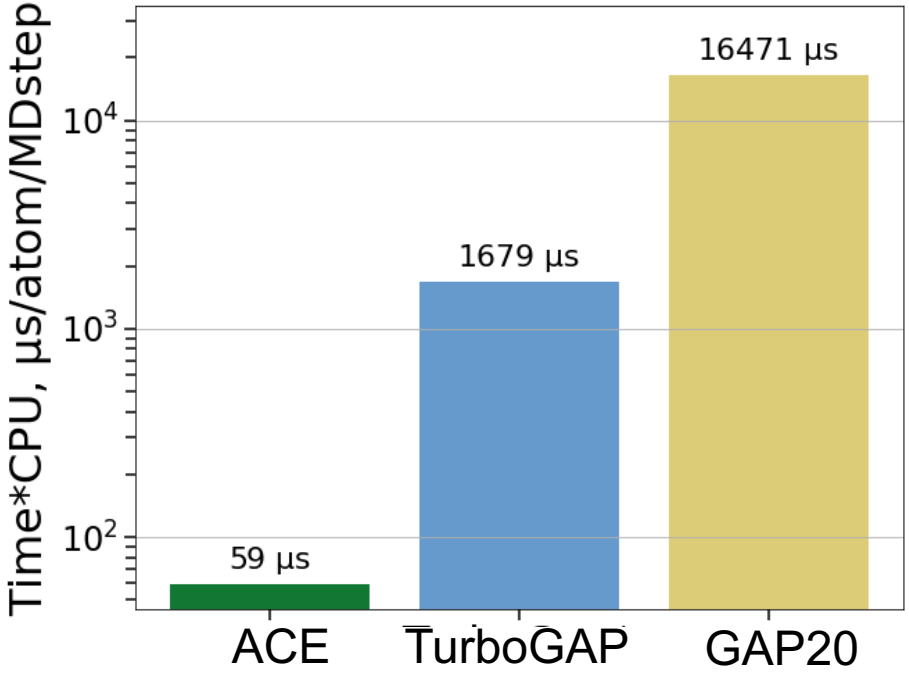
Carbon

Carbon – structural stability



ACE: Qamar, et al., arXiv:2210.09161v2
GAP20: Rowe, et al., J. Chem. Phys. 153 (2020) 034702
TurboGAP: Wang, et al., Chem. Mater. 34 (2022) 617

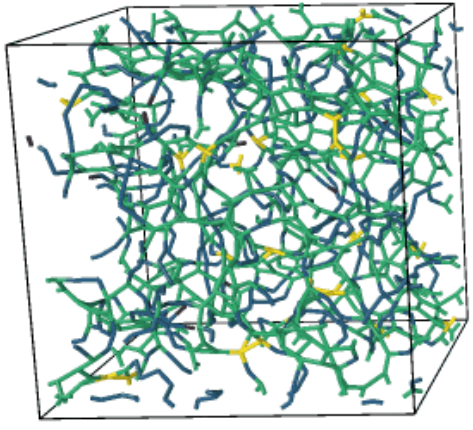
Carbon – timing



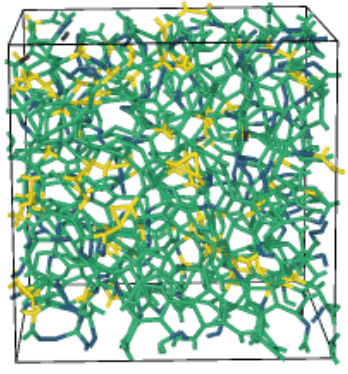
GAP20: Rowe, et al., J. Chem. Phys. 153 (2020) 034702
TurboGAP: Wang, et al., Chem. Mater. 34 (2022) 617

Carbon – quench at different densities

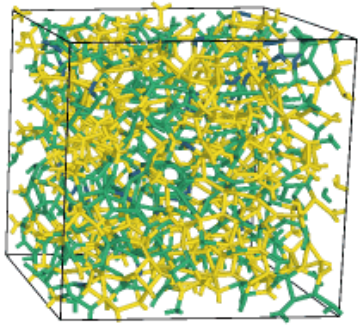
ρ = 1.4 g/cc



ρ = 2.2 g/cc



ρ = 3.2 g/cc

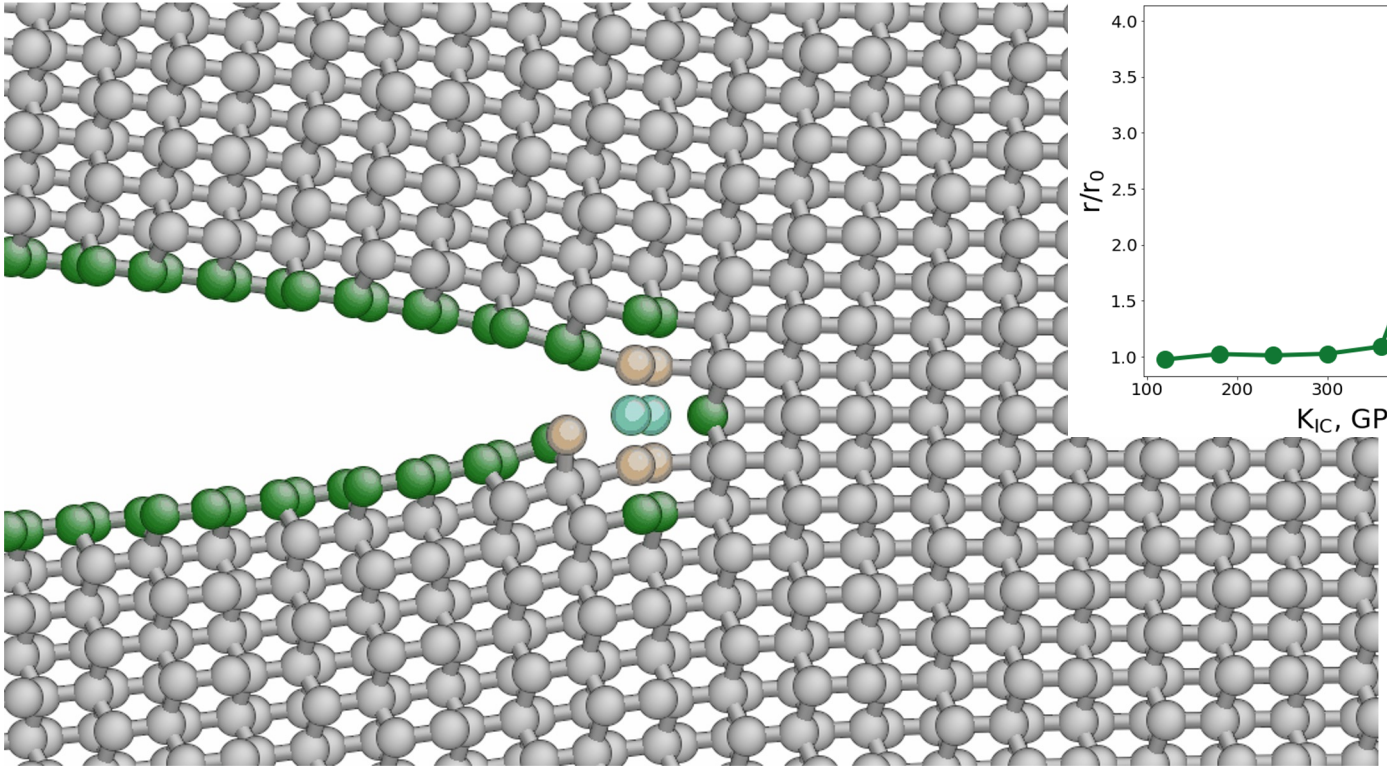


111.2 ps

4140 K

Carbon – diamond fracture

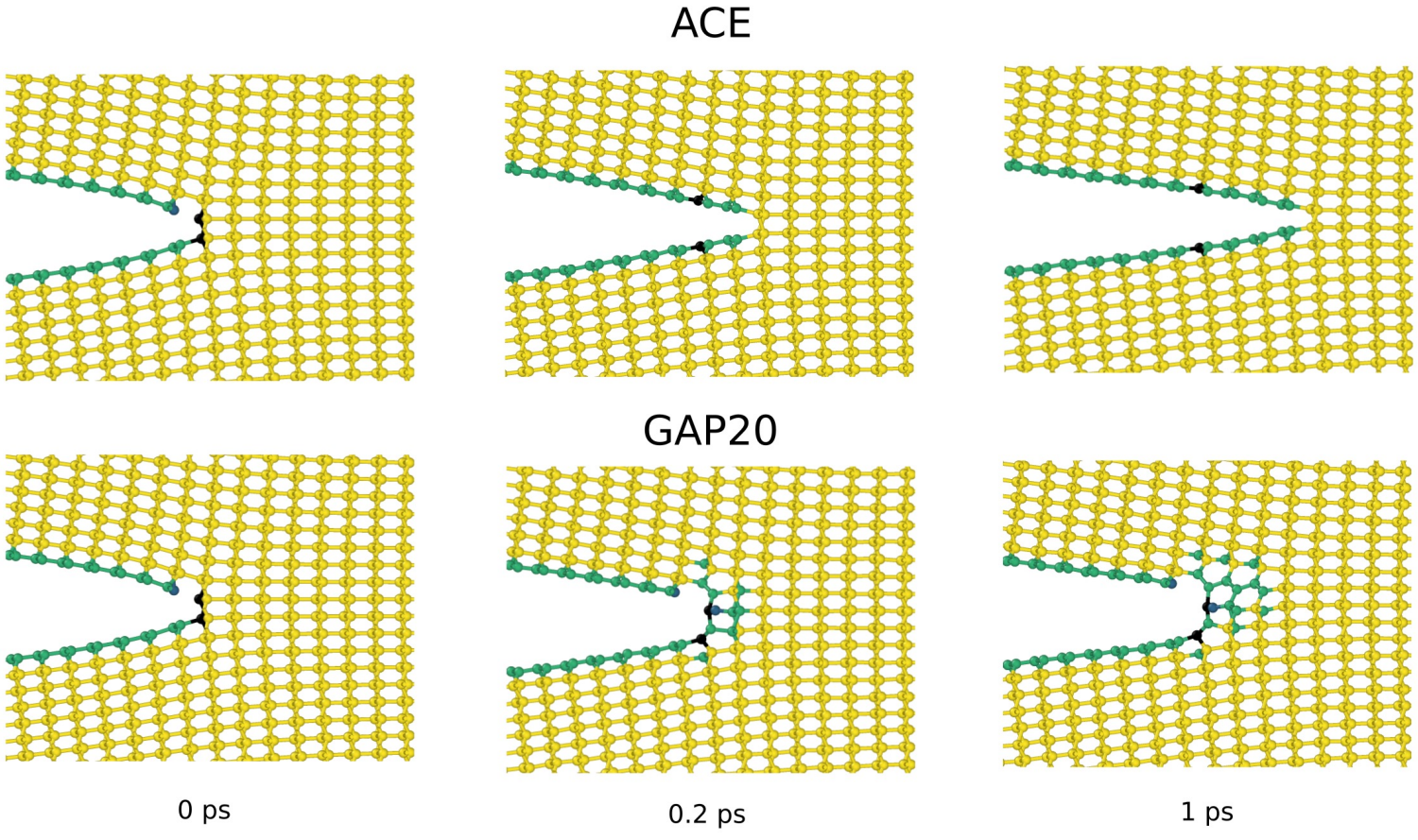
{110}



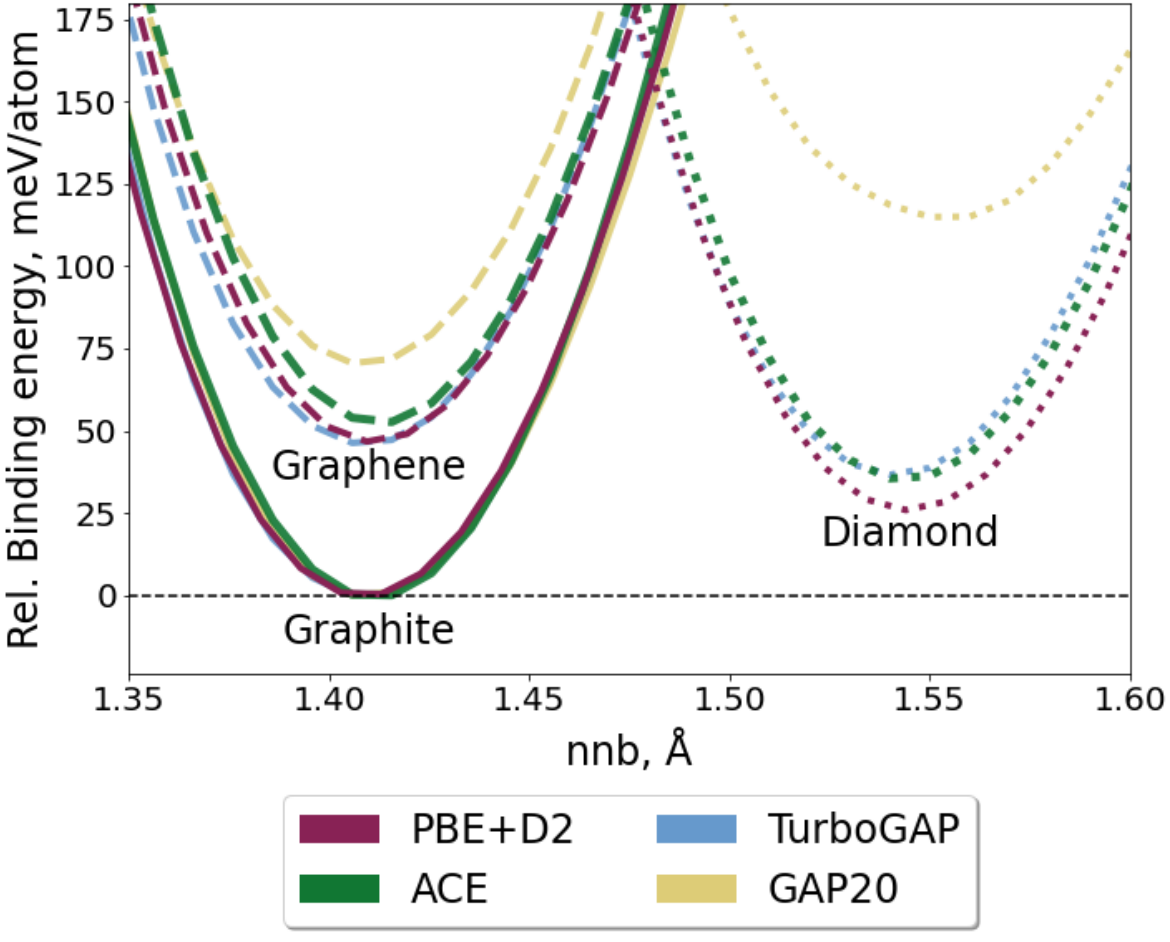
{111}

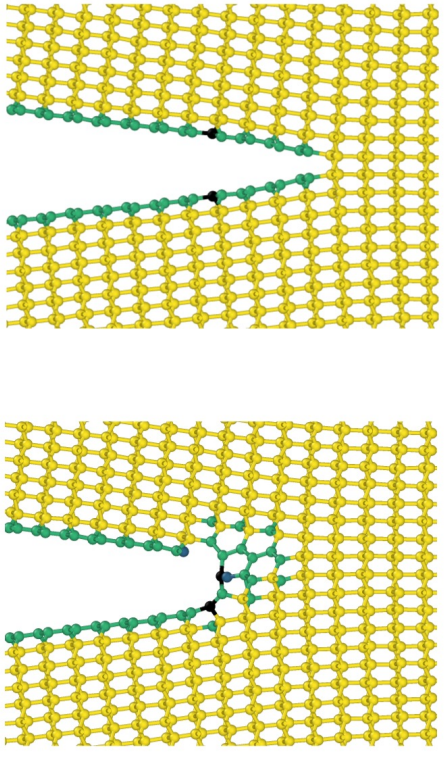
K_{IC} = 540 GPa. $A^{1/2}$

Carbon – diamond fracture

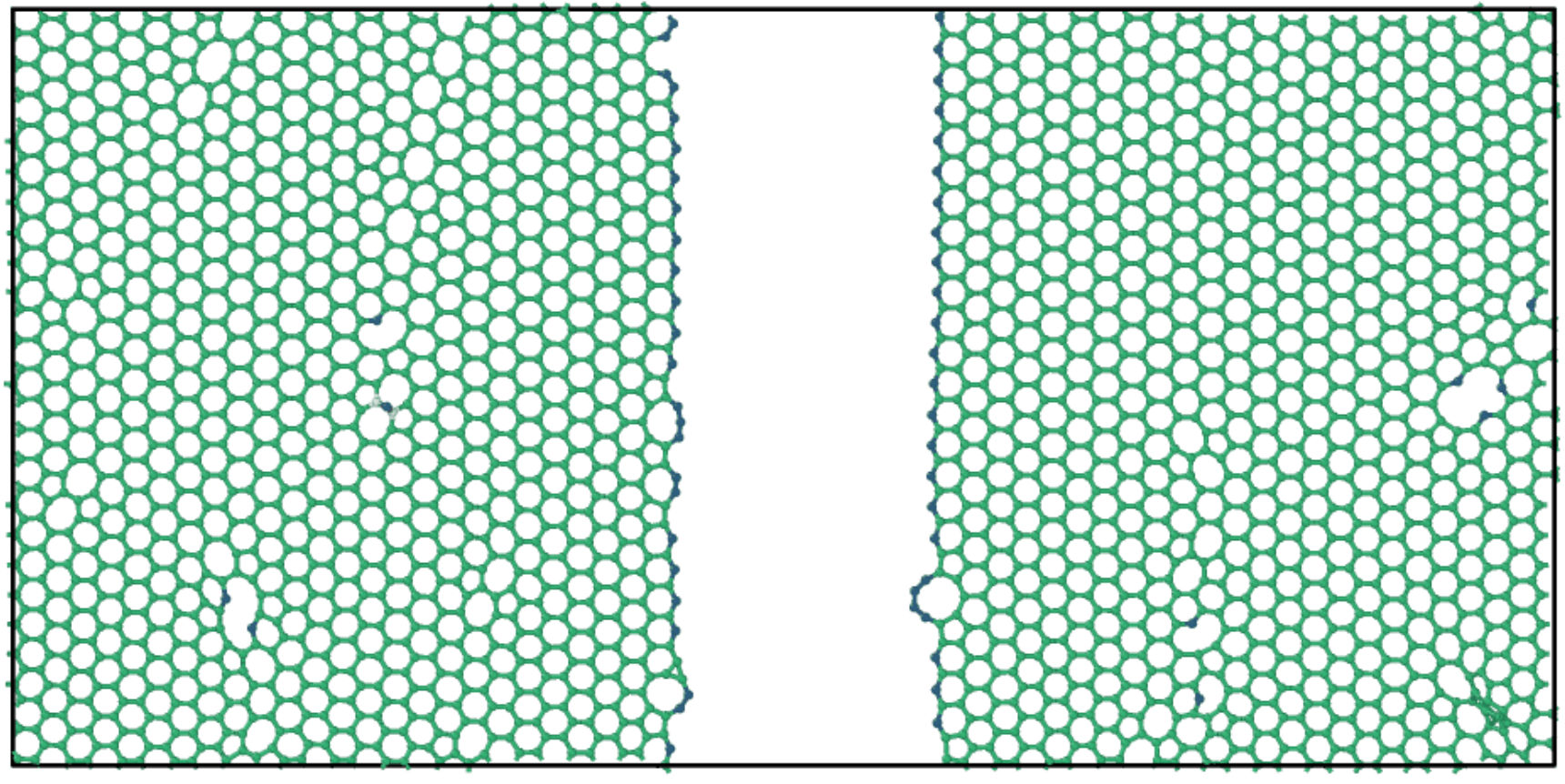


Carbon – diamond fracture

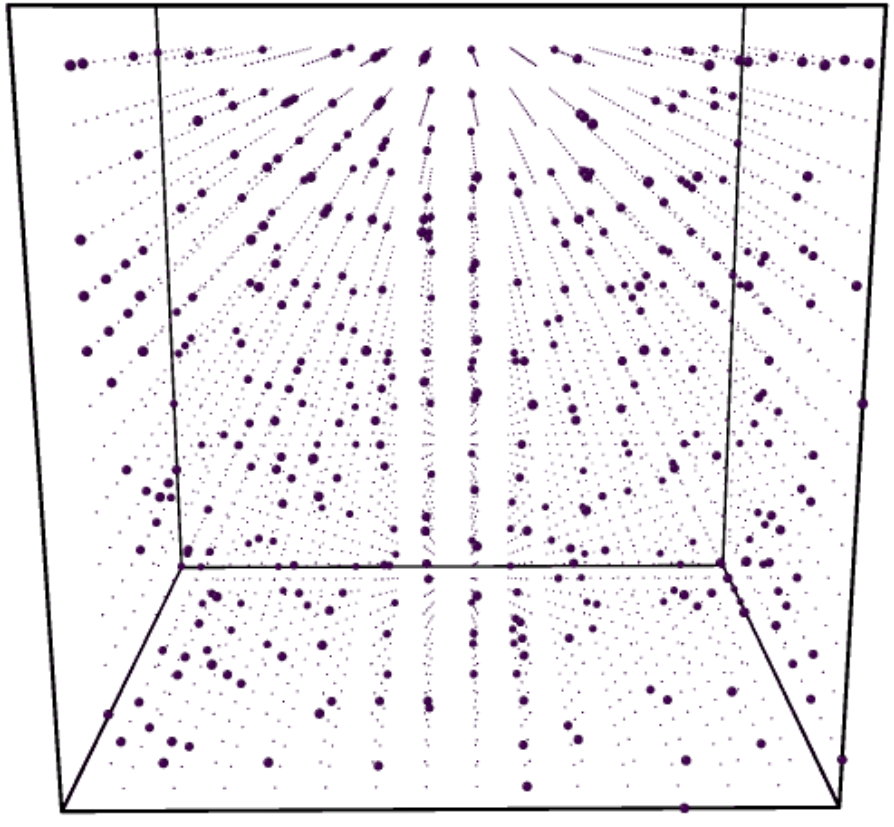




Graphene flake cracking

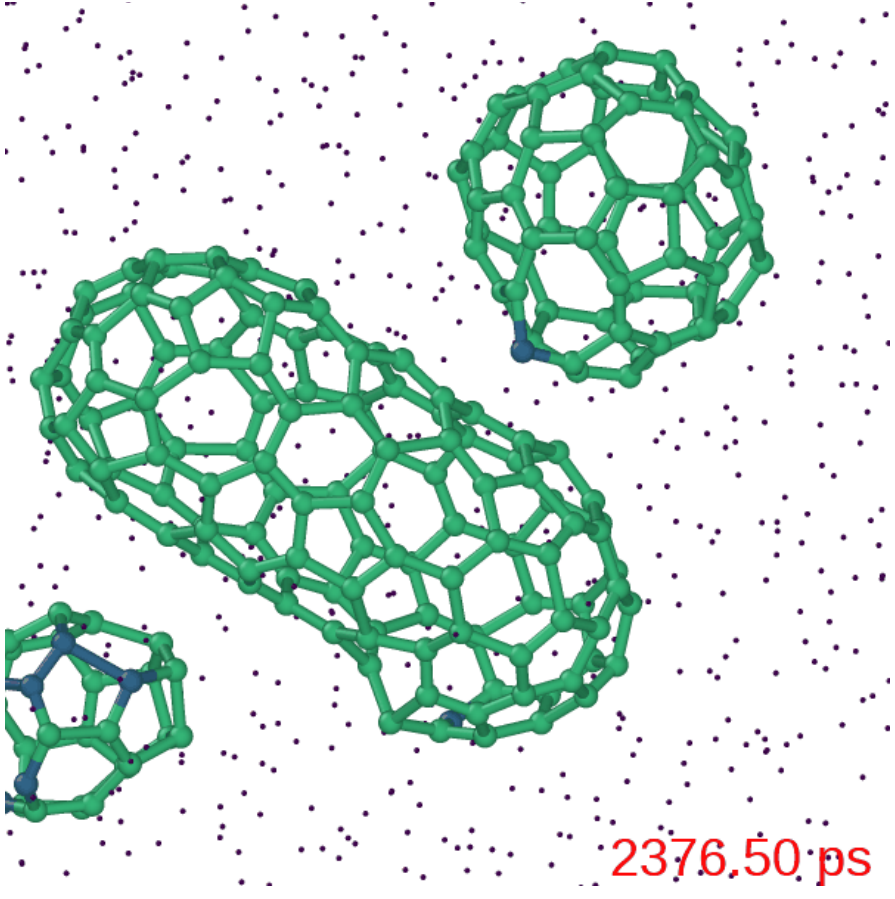


Carbon in argon atmosphere



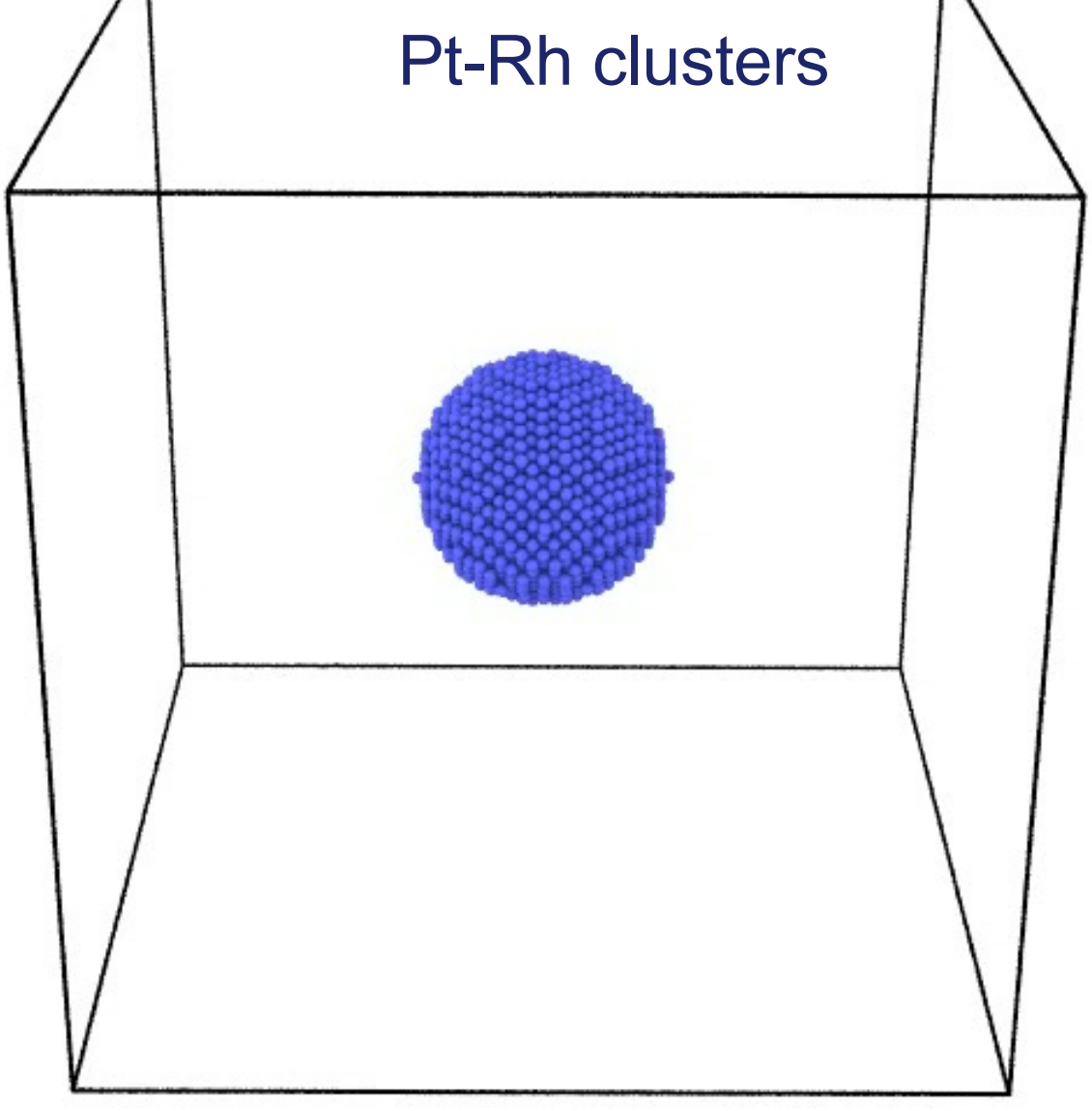
With Romain Perriot (LANL),
Simulation setup:
Pineau et al., J. Chem. Phys. 129 (2008) 024708

Merging of buckyballs



Nanoclusters

Pt-Rh clusters



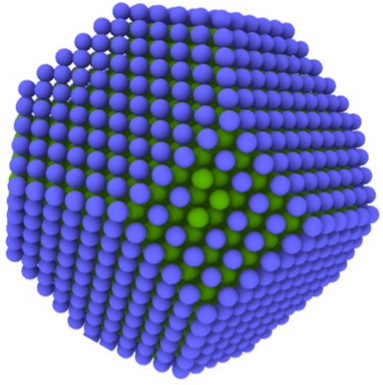
Initial:

- Pt core
- Rh shell

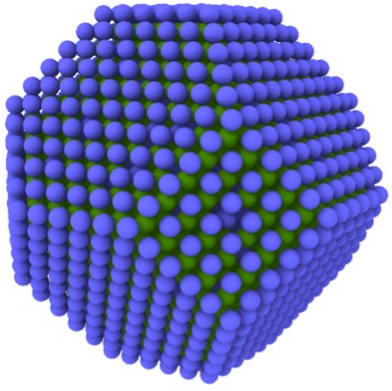
OVITO

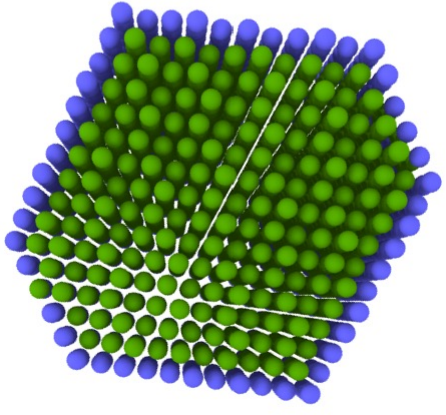
www.ovito.org

Pt-Rh clusters

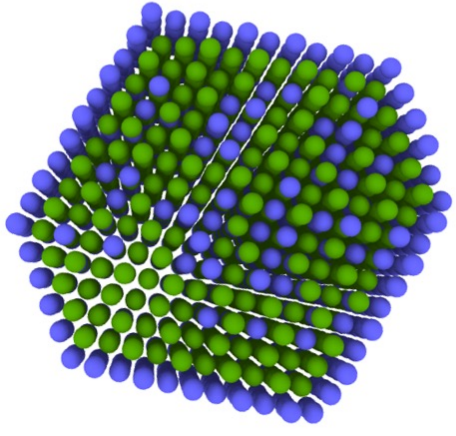


← 30% Pt



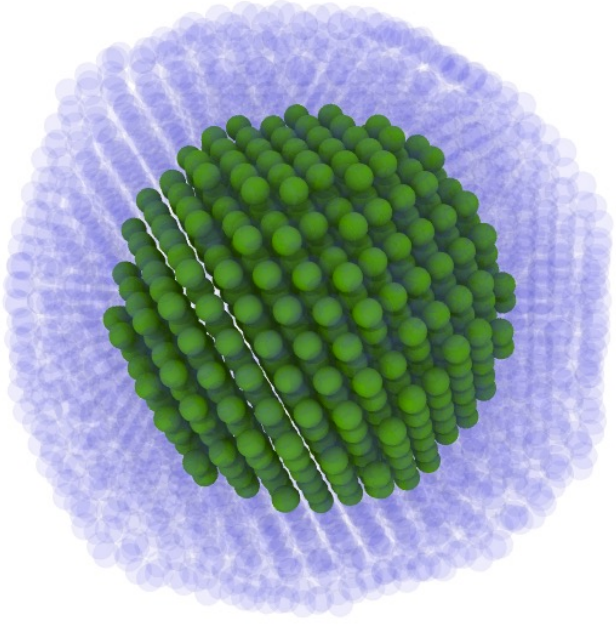


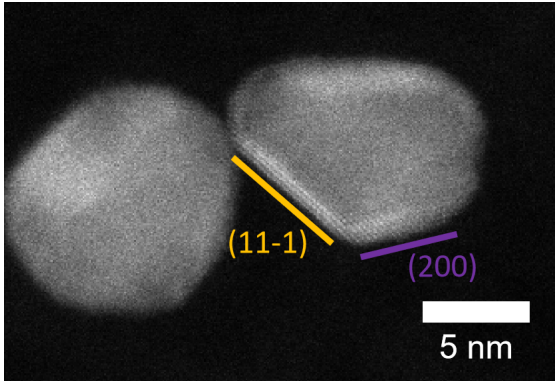
45% Pt →

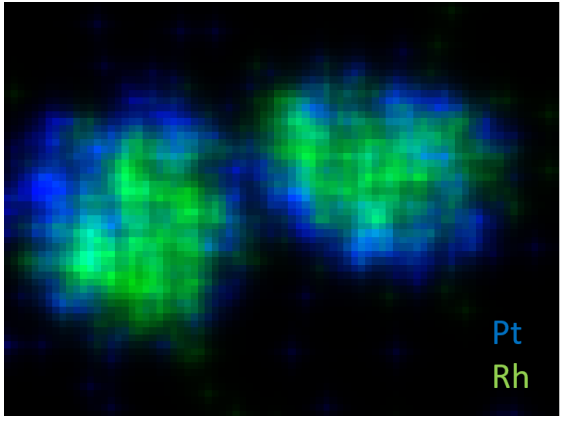


Pt-Rh clusters

- MD simulation at 1000 K for 2 ns



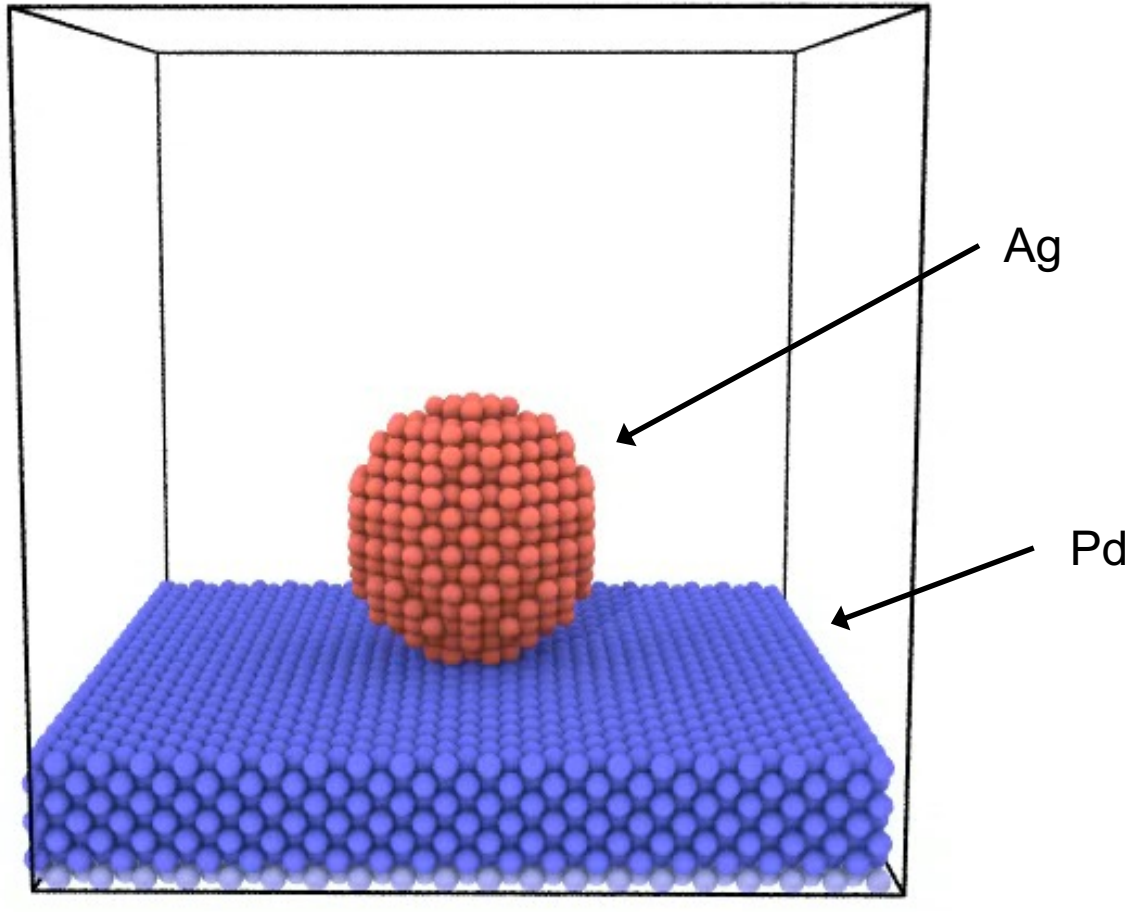




➜ Rh core (meta-)stable

Liang et al., arXiv:2303.07465

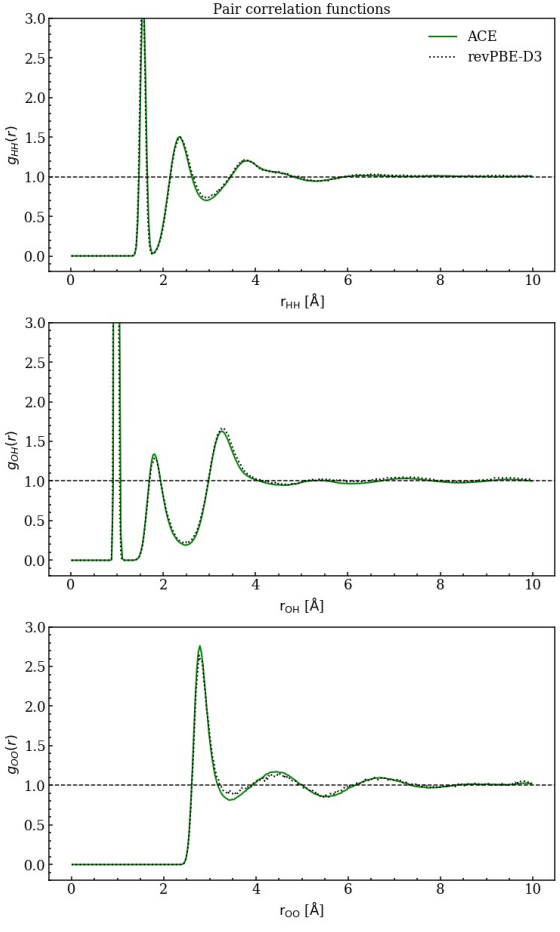
Ag-Pd

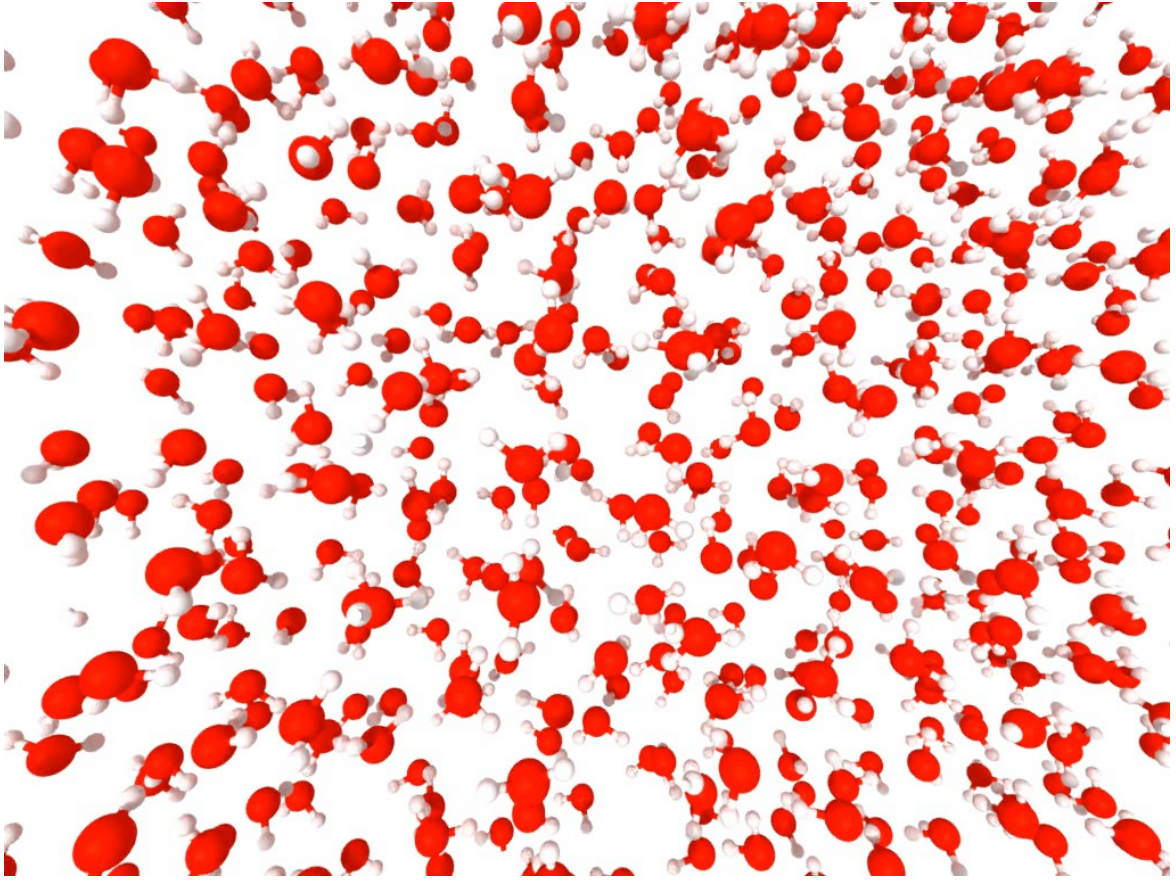


T = 900K

Water

Water

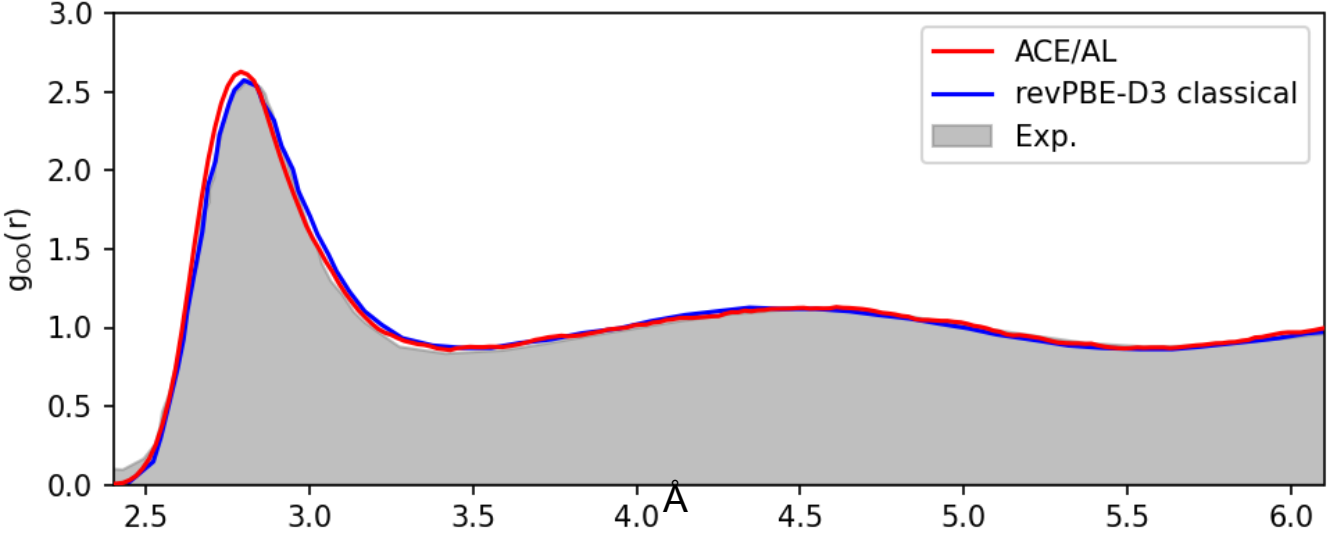




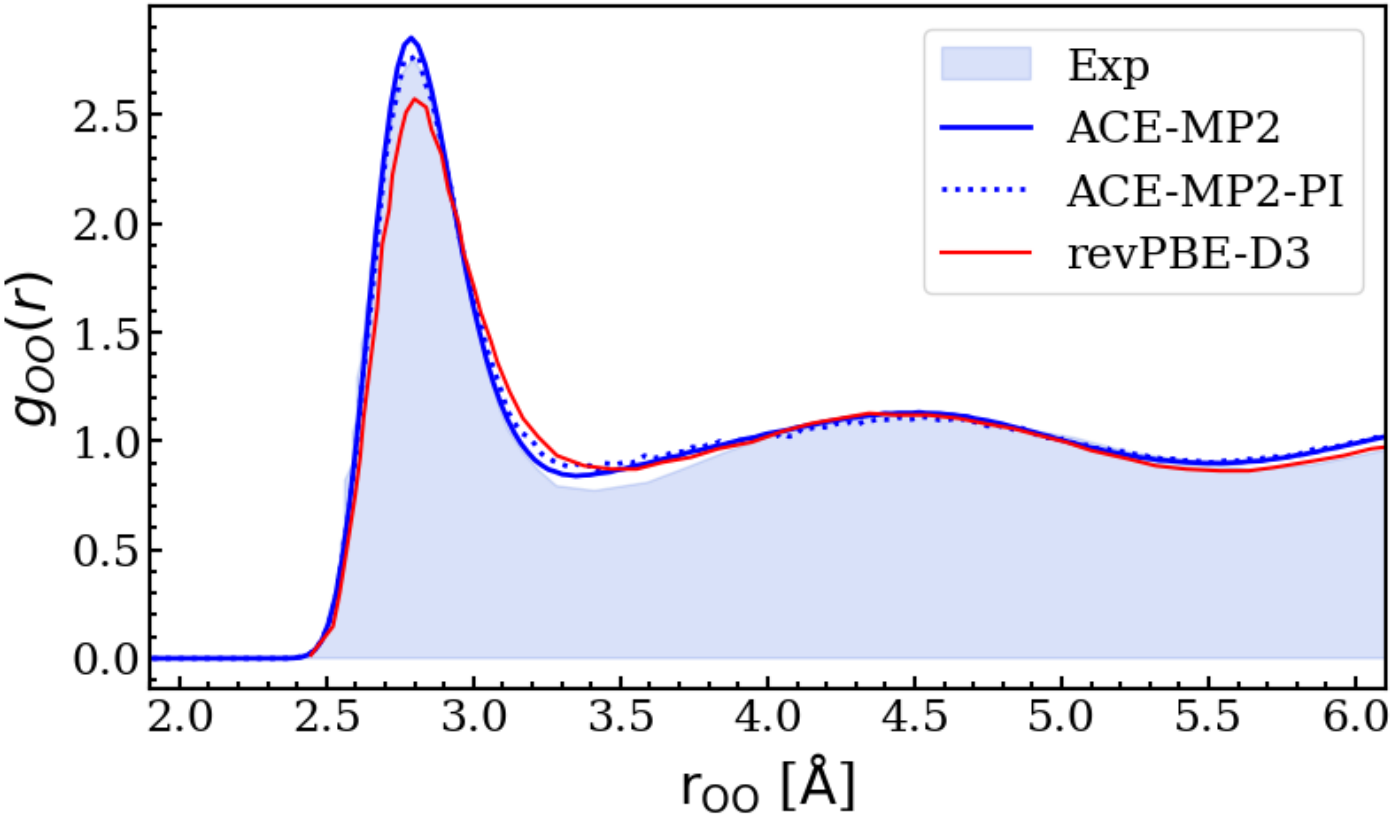
DFT reference data: Cooper, Kästner, Urban, Artrith, npj Comput. Mater. 6 (2020) 54

Water – Active exploration

- Three generations active learning
- In total 311 DFT computations with 64 water molecules each
- Energy RMSE = 6.07 meV/atom, force RMSE = 55.91 meV/Å



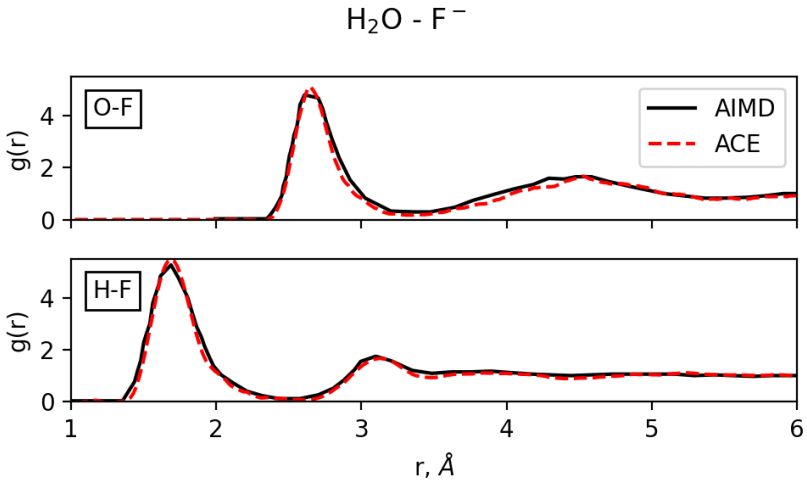
Water – DFT and MP2

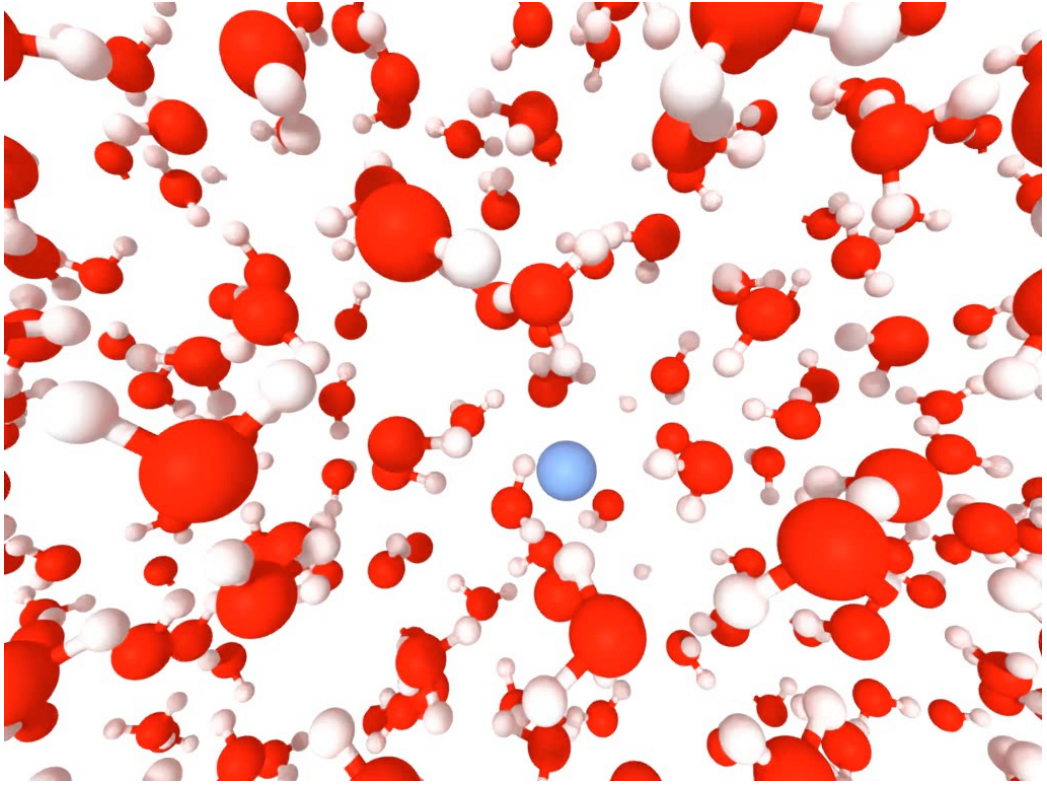


MP2 reference data: Daru, Forbert, Behler, Marx, Phys. Rev. Lett. 129 (2022) 226001

Fluorine in Water

- Fit to reference DFT data
- Energy RMSE = 0.26 meV/at
- Force RMSE = 34.52 meV/ Å

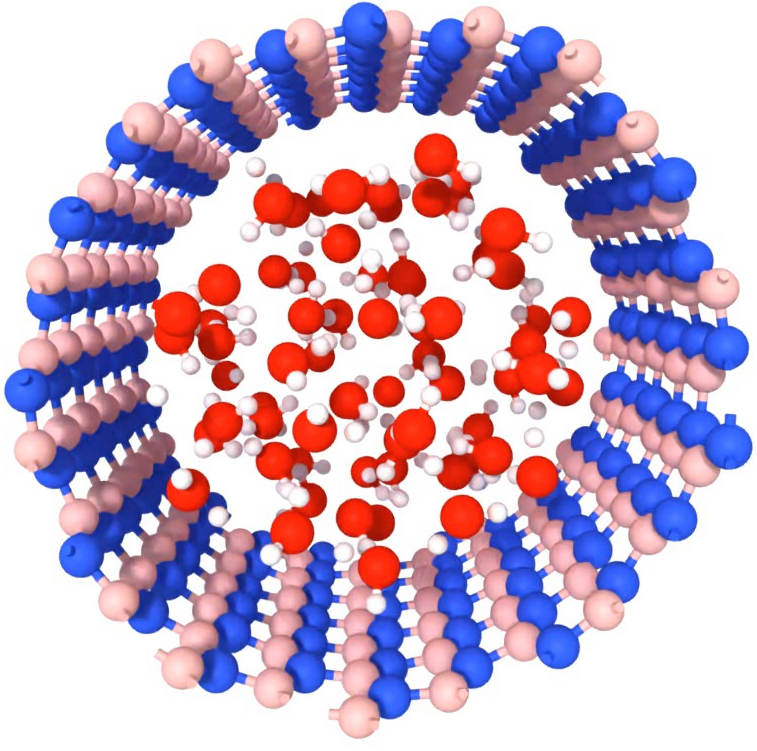




DFT reference data: Schran, Thiemann, Rowe, Müller, Marsalek, Michaelides, PNAS 118 (2021) e2110077118

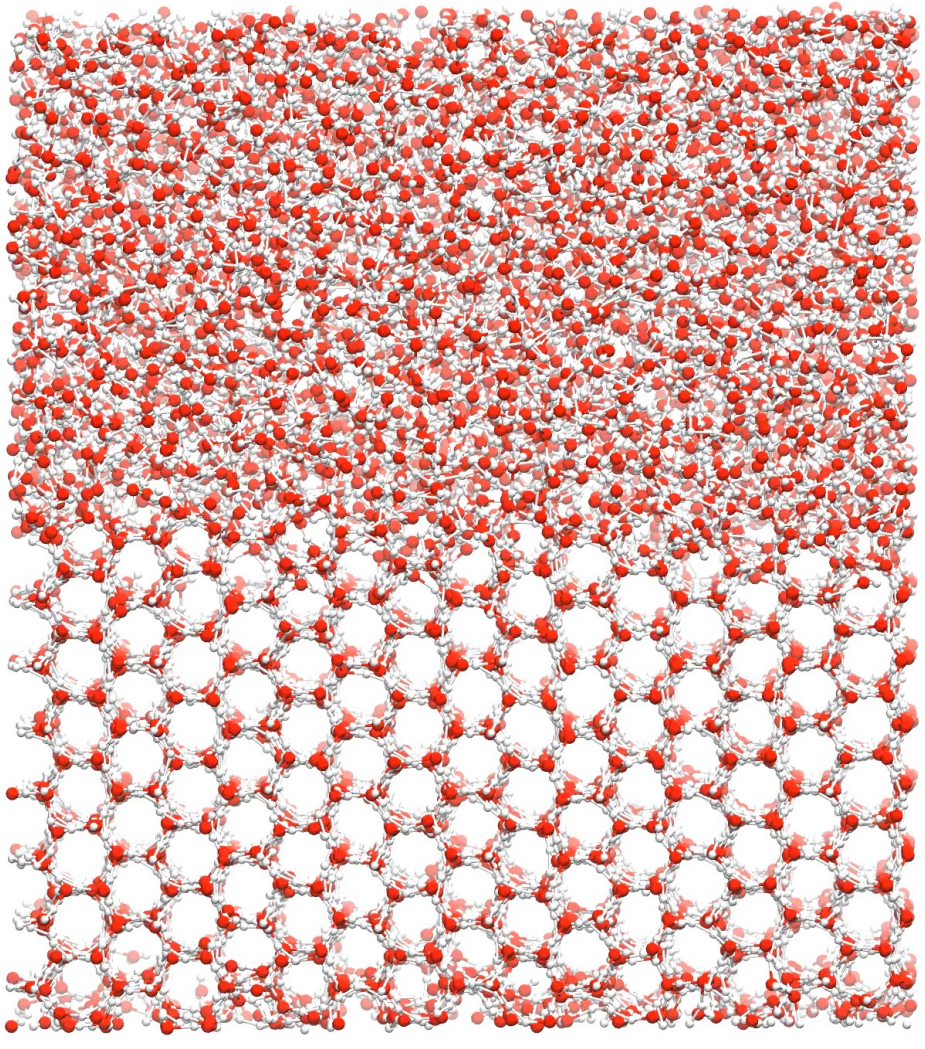
Water in BN tube

- Fit to reference DFT data



DFT reference data: Schran, Thiemann, Rowe, Müller, Marsalek, Michaelides, PNAS 118 (2021) e2110077118

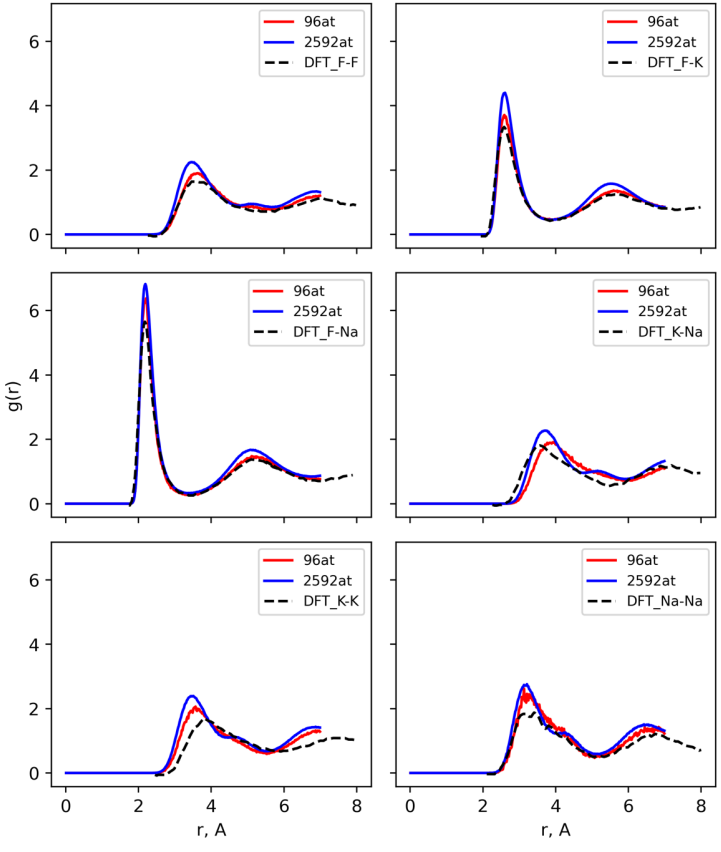
Melting

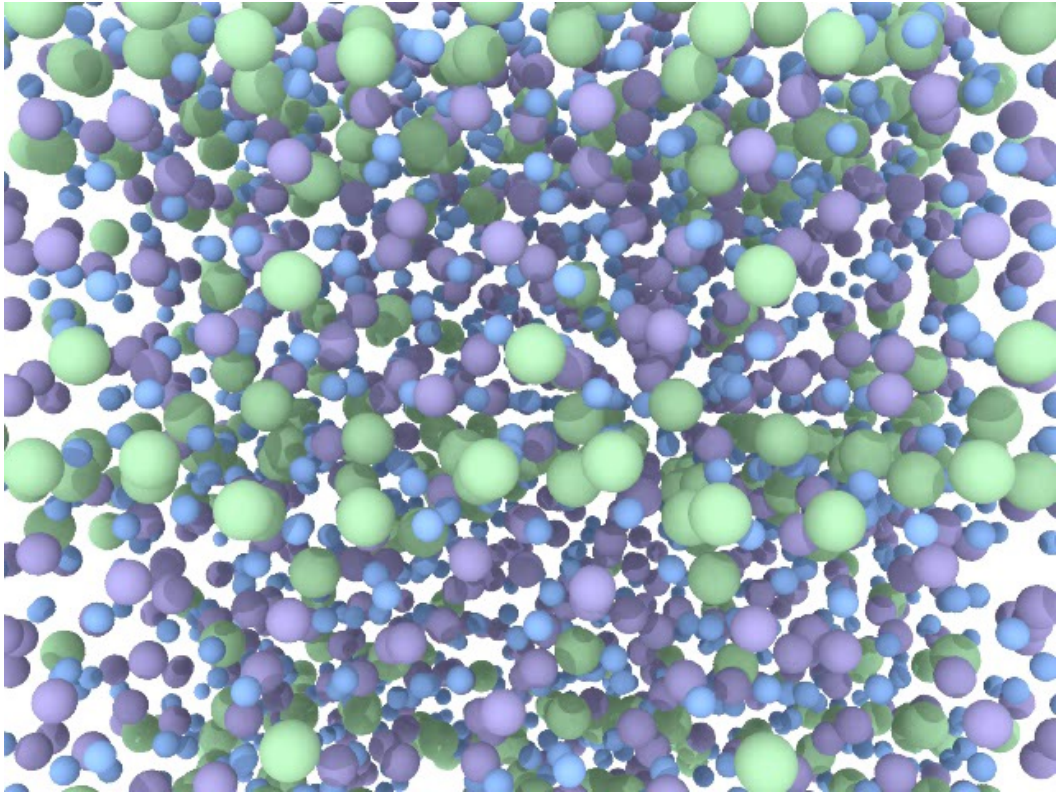


Molten salt

2KF-NaF

- Fit to DFT, RMSE 0.3 meV/at, 10 meV/Å

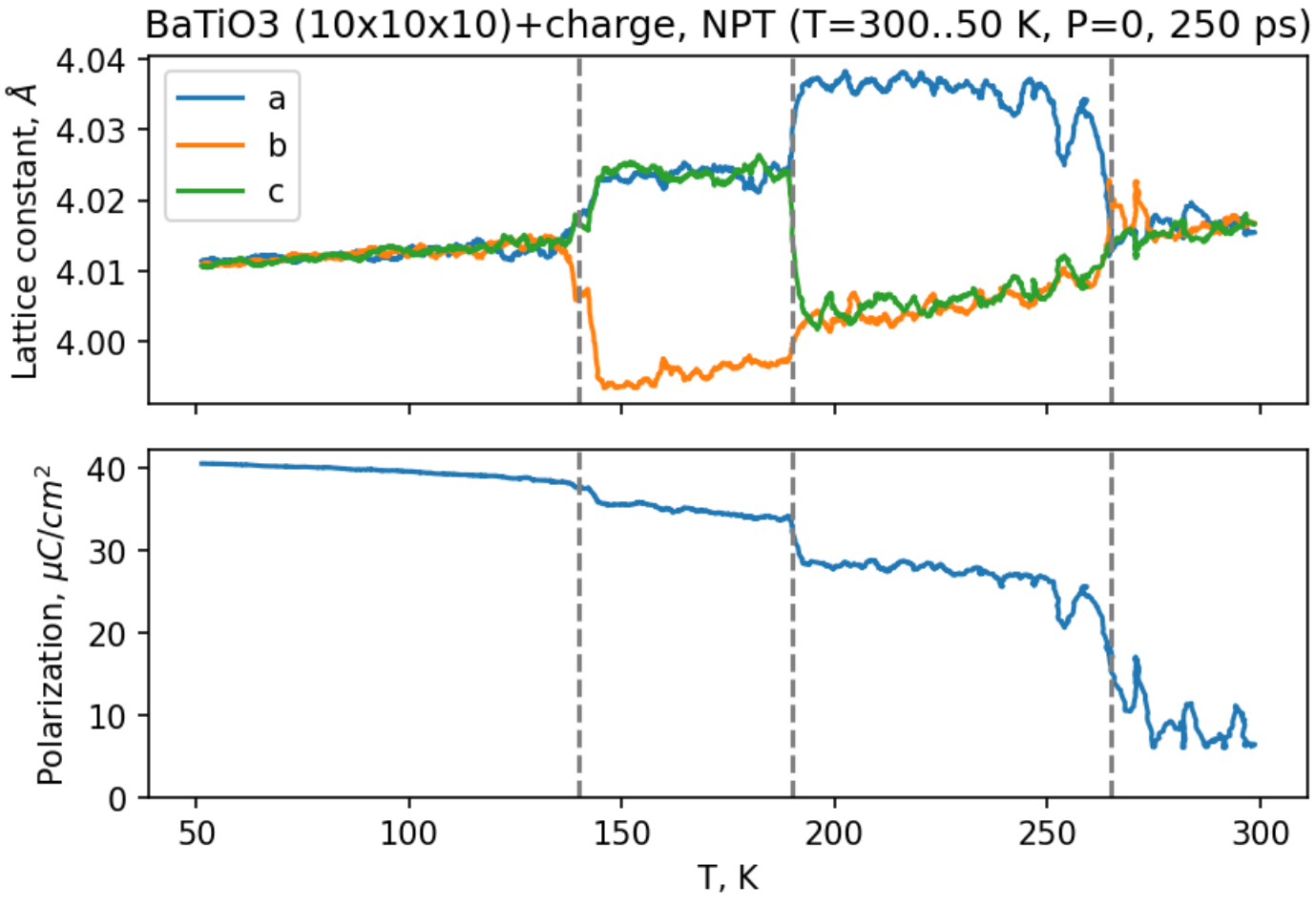




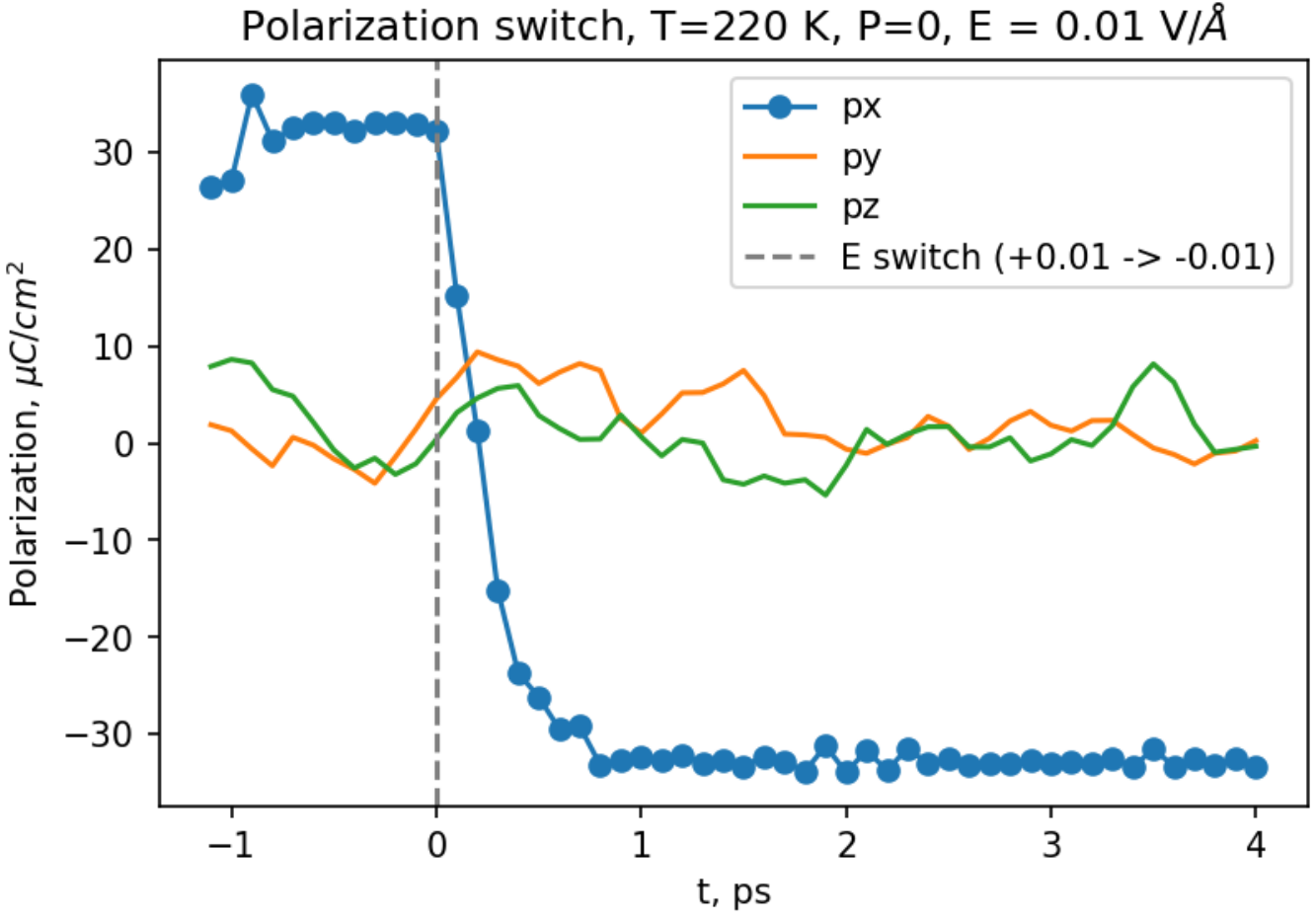
DFT reference data: Winner, Williams, Scarlat, Asta, J. Mol. Liq. 335 (2021) 116351

Ferroelectrics

BTO phase diagram



BTO polarization switching



Extensions

Charges, magnetism, tensors, messages

- Include further degrees of freedom (charge, magnetism, …)
- Expand vectorial or tensorial properties

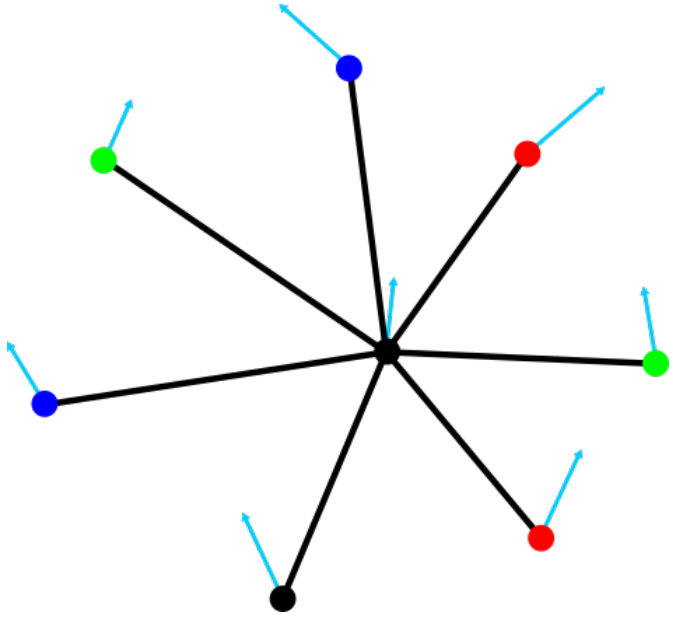
- Extend single particle basis functions

$$\phi_v(\boldsymbol{r}) \;\rightarrow\; \phi_v(\boldsymbol{r}, \boldsymbol{m})$$

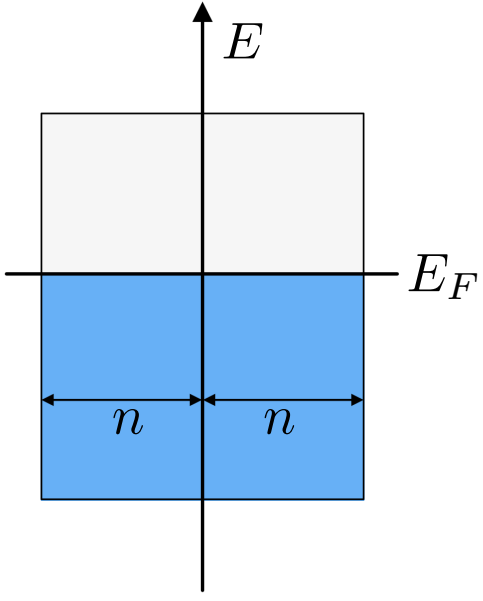
➜ Atomic cluster expansion unchanged

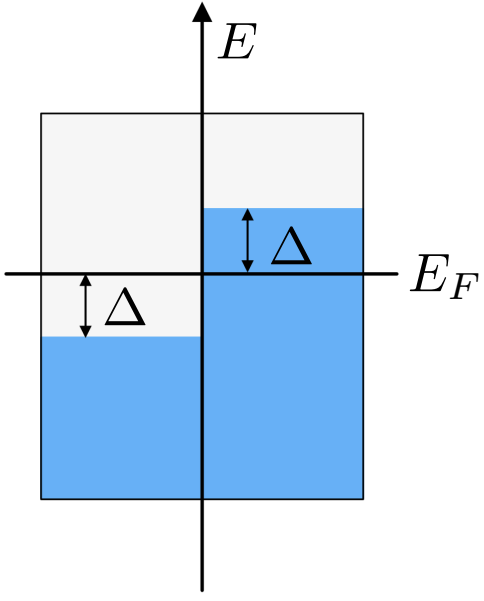
$$E_i(\boldsymbol{\sigma}) = \sum_{\mathbf{v}} \tilde{c}_{\mathbf{v}} \boldsymbol{A}_{\mathbf{v}}$$

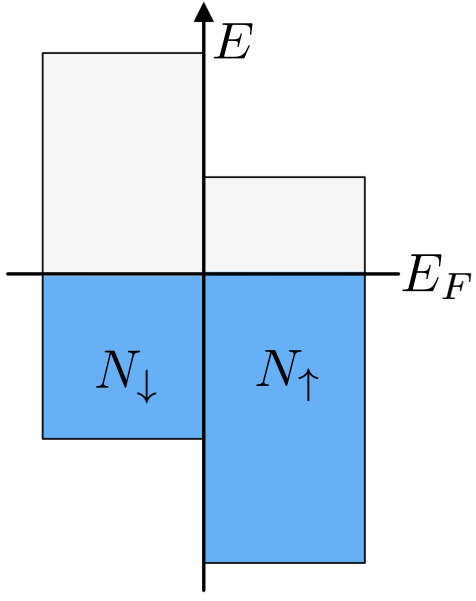
- But more parameters



Stoner ferromagnetism







$$\delta N = n\Delta \qquad m = N_\uparrow - N_\downarrow = 2(n\Delta)$$

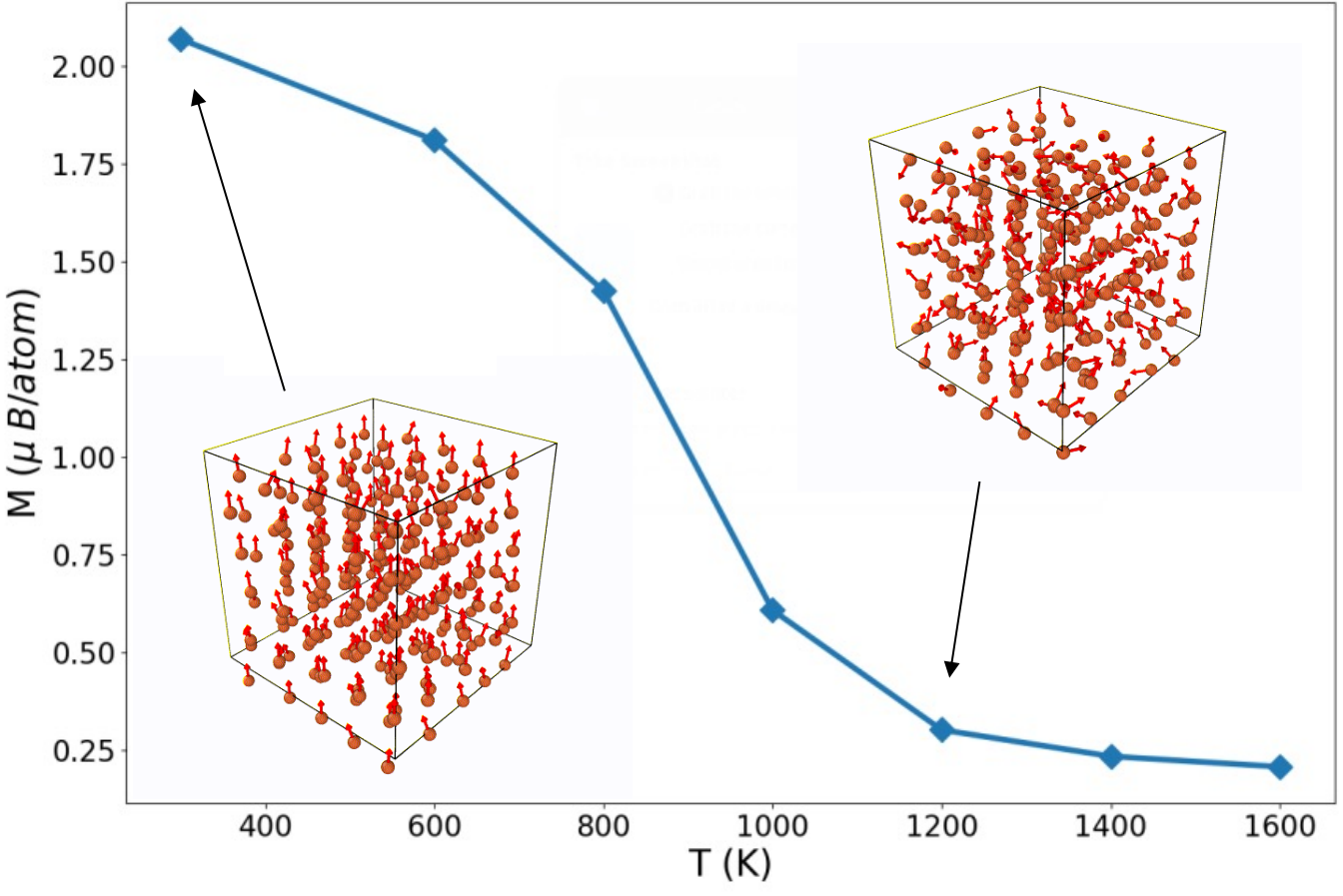
$$\delta E_{kin} = (n\Delta)\Delta \qquad E_X = -\frac{1}{4} I m^2 = -I(n\Delta)^2$$

Magnetism weakens bonds and lowers atomic energies.

$$\delta E = \delta E_{kin} + E_X = n\Delta^2(1 - In)$$

Magnetism in iron

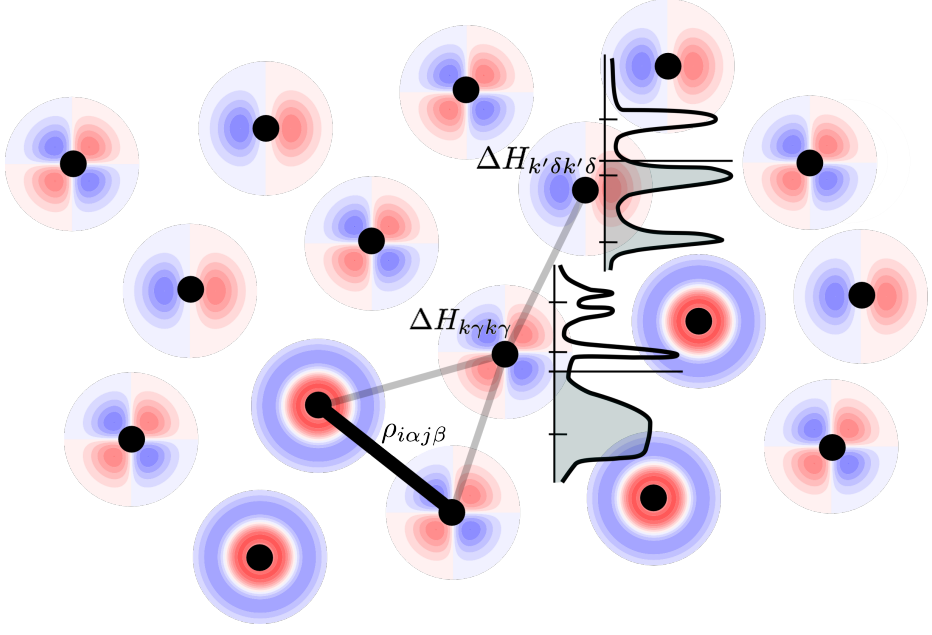
- Hamiltonian Monte Carlo combining atomic and spin dynamics
- Including longitudinal fluctuations

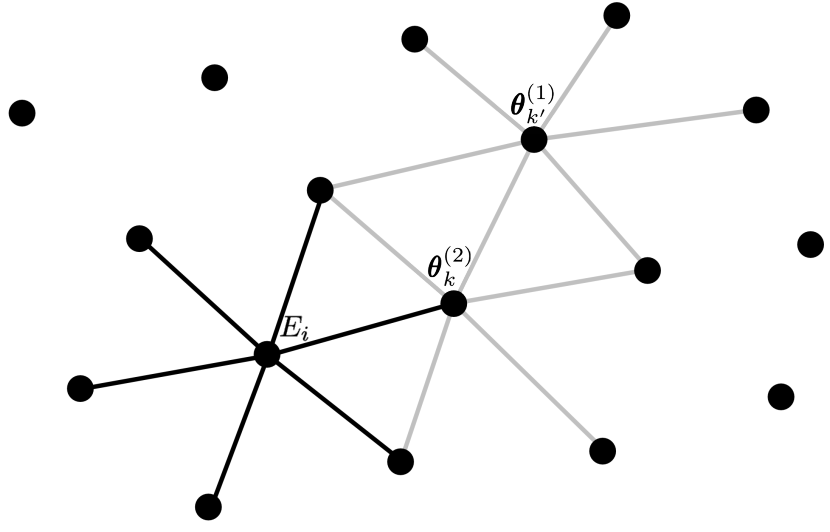


Drautz, PRB (2019)

Equivariant Message Passing Networks

- ACE can be used to generalize message passing networks
- Example: semi-local interactions
- Different flavors: mlACE, multi ACE, MACE





Bochkarev, et al, Phys. Rev. Res. Lett. (2022), Batatia, et al, arXiv:2205.06643/arXiv:2206.07697

Software

PACE
LAMMPS

PACEmaker (TF)
parameterization

PACE-al
active learning

workflow

phase diagrams

ACE.jl
parameterization

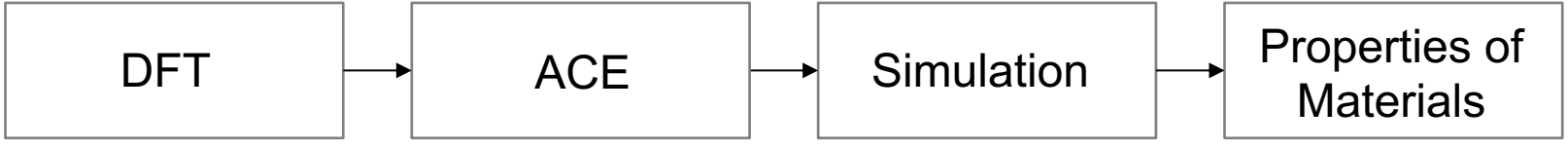
FitSNAP
parameterization

github.com/ICAMS, pyiron.org, calphy.org,
github.com/ACEsuit/ACE.jl, github.com/FitSNAP/FitSNAP

Lysogorskiy et al., npj Comput. Mater. 7 (2021) 97
Bochkarev et al., Phys. Rev. Mat. 6 (2022) 103804
Janssen et al., Comput. Mat. Sci. 163 (2019) 24
Menon, Lysogorskiy, Rogal, Drautz, Phys. Rev. Mat. 5 (2021) 103801

Conclusions

- **Robust workflows from DFT to properties**



- **Tutorial**
  http://pyiron.org/potentials-workshop-2022/intro.html

POTENTIALS