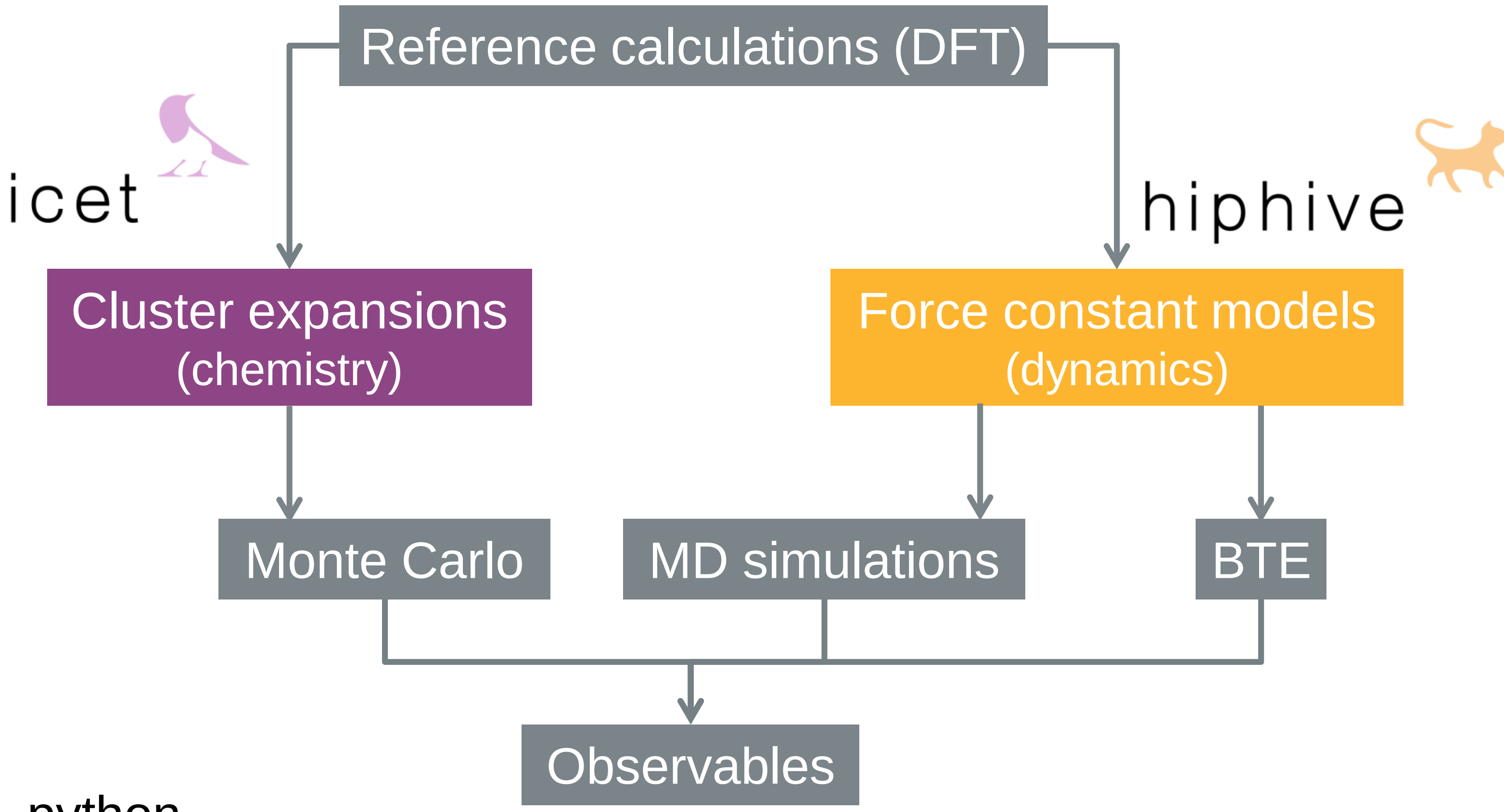
Tools for atomic-scale model construction
Cluster expansions and force constants

Erik Fransson

Department of Physics
Chalmers University of Technology

Tools for atomic-scale model construction

$$\mathcal{Z} = \mathcal{Z}_{\mathrm{conf}}\mathcal{Z}_{\mathrm{vib}}\mathcal{Z}_{\mathrm{elec}}\cdots$$



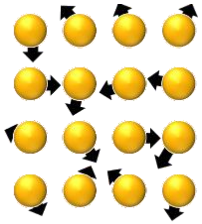
python

Force constant expansions

hiphive

Taylor expansion of the potential energy

$$E = E_0 + \sum_i \Phi_i u_i + \frac{1}{2!} \sum_{ij} \Phi_{ij} u_i u_j + \frac{1}{3!} \sum_{ijk} \Phi_{ijk} u_i u_j u_k \ldots$$



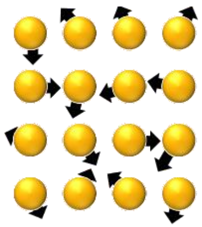
u_i Displacement from static equilibrium position

Force constant expansions

hiphive

Taylor expansion of the potential energy

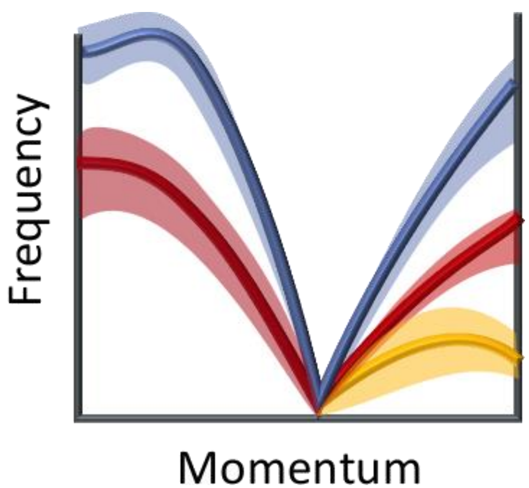
$$E = E_0 + \sum_i \Phi_i u_i + \frac{1}{2!} \sum_{ij} \Phi_{ij} u_i u_j + \frac{1}{3!} \sum_{ijk} \Phi_{ijk} u_i u_j u_k \ldots$$



u_i Displacement from static equilibrium position

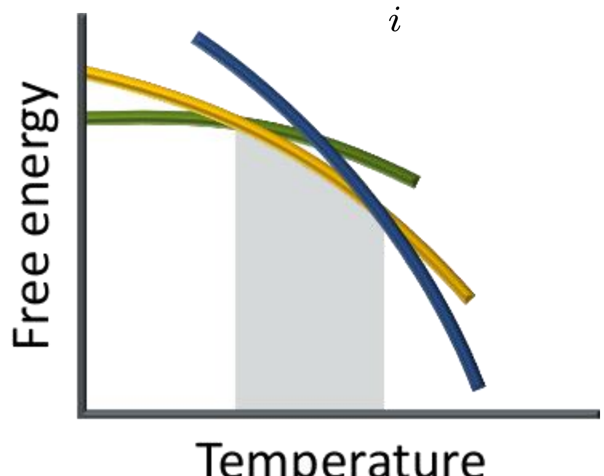
Phonon dispersions

$$\Phi_{ij} \rightarrow D_{\alpha\beta}(\boldsymbol{q}) \rightarrow \omega_k(\boldsymbol{q})$$



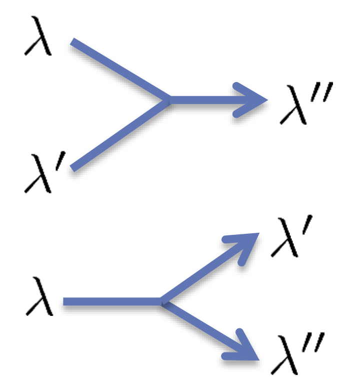
Free energies

$$F_{vib} = k_B T \sum_i \ln \omega_i$$



Transport

$$\kappa_l = \frac{1}{2V} \sum_{j\boldsymbol{q}} \lambda_{j\boldsymbol{q}}(T) v_{j\boldsymbol{q}} c_{j\boldsymbol{q}}(T)$$



CHALMERS
UNIVERSITY OF TECHNOLOGY

Force constant extraction

hiphive

$$E = E_0 + \sum_i \Phi_i u_i + \frac{1}{2!}\sum_{ij} \Phi_{ij} u_i u_j + \frac{1}{3!}\sum_{ijk} \Phi_{ijk} u_i u_j u_k \ldots$$

"Direct" approach: systematic enumeration

$$\Phi_{ij} = \frac{\partial^2 E}{\partial u_i \partial u_j} \approx -\frac{F_i(u_j) - F_i(0)}{u_j}$$

Harmonic approximation → phonopy

Third-order FCs → phono3py, almaBTE

→ Poor scaling with system size and order

CHALMERS
UNIVERSITY OF TECHNOLOGY

Force constant extraction

hiphive

$$E = E_0 + \sum_i \Phi_i u_i + \frac{1}{2!} \sum_{ij} \Phi_{ij} u_i u_j + \frac{1}{3!} \sum_{ijk} \Phi_{ijk} u_i u_j u_k \ldots$$

"Direct" approach: systematic enumeration

Harmonic approximation → phonopy

Third-order FCs → phono3py, almaBTE

→ Poor scaling with system size and order

$$\Phi_{ij} = \frac{\partial^2 E}{\partial u_i \partial u_j}$$

$$\approx -\frac{F_i(u_j) - F_i(0)}{u_j}$$

"Regression" approach: fit to snapshots

→ TDEP & ALAMODE

Hellman *et al.* PRB 2011, Tadano *et al.* JPCM 2014

→ Compressive sensing

Zhou, Ozolins *et al.* PRL 2014

$$\Phi_{ij} = \Phi_{ij}(\boldsymbol{x})$$

$$\min(\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{f}\|_2^2)$$

hiphive

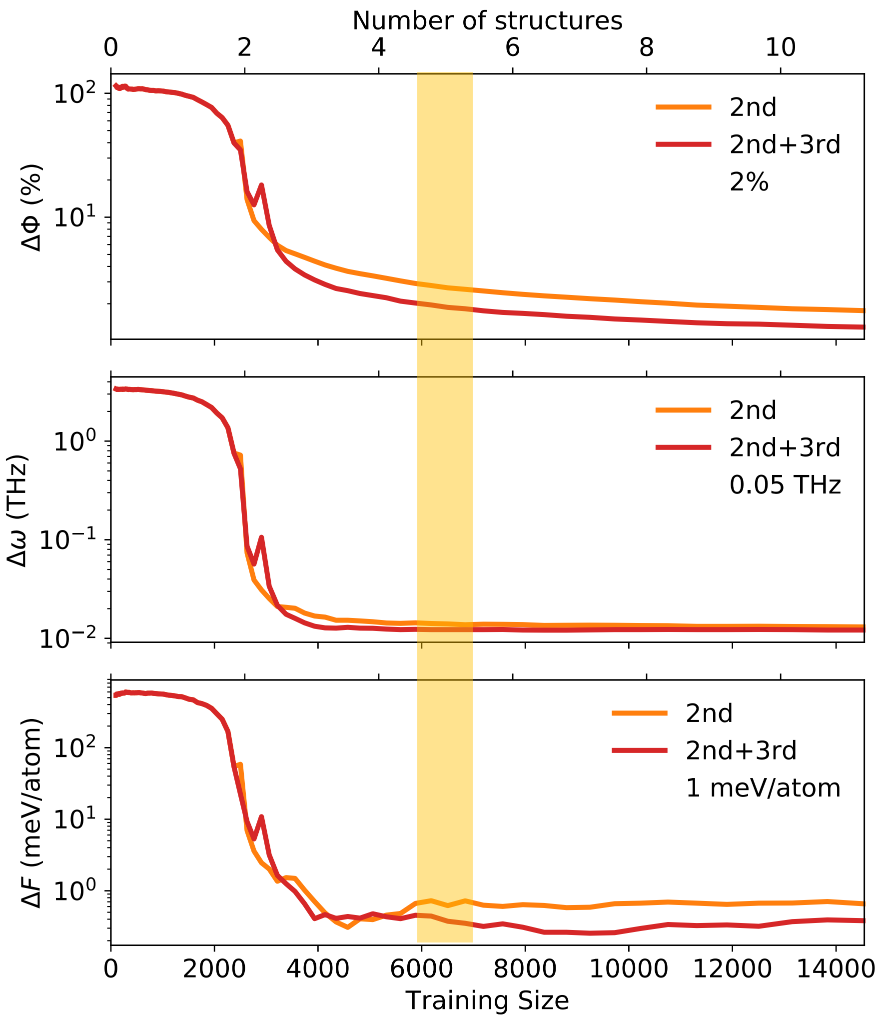
Generalizes regression approach
and makes it easily accessible

CHALMERS
UNIVERSITY OF TECHNOLOGY

Efficiency 2nd-order: Vacancy in BCC-Ta

hiphive

Convergence with training size

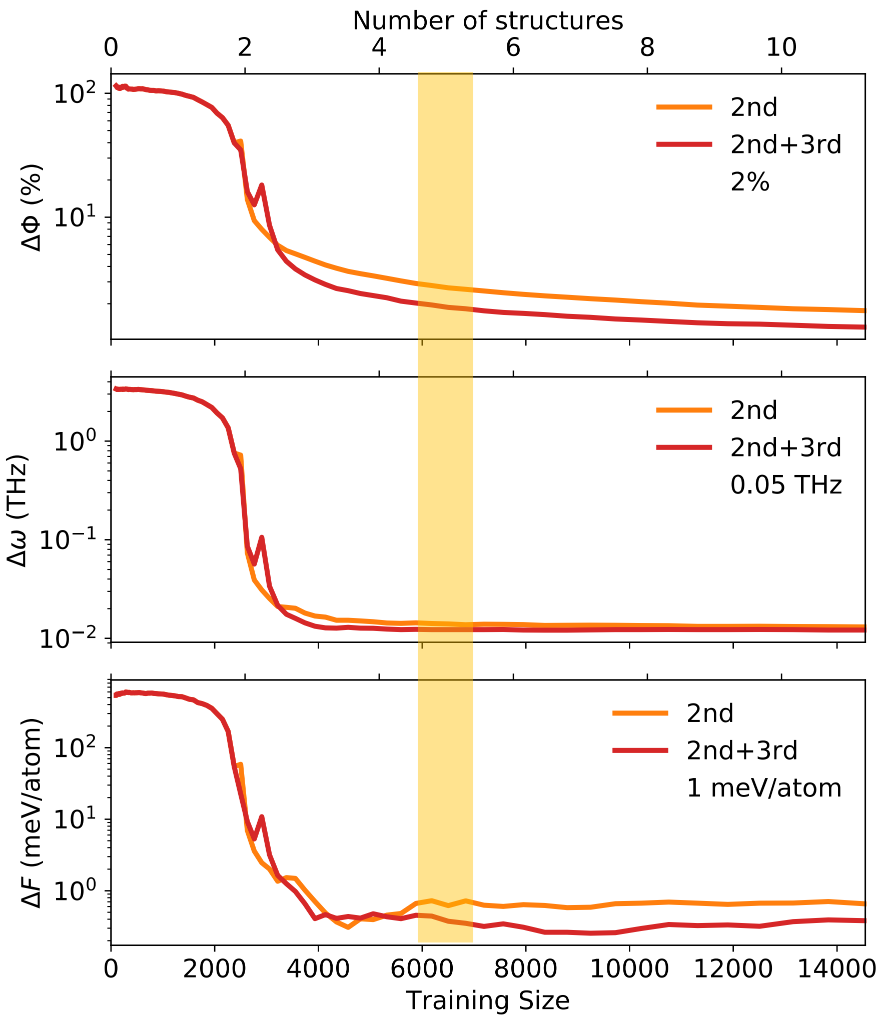


CHALMERS
UNIVERSITY OF TECHNOLOGY

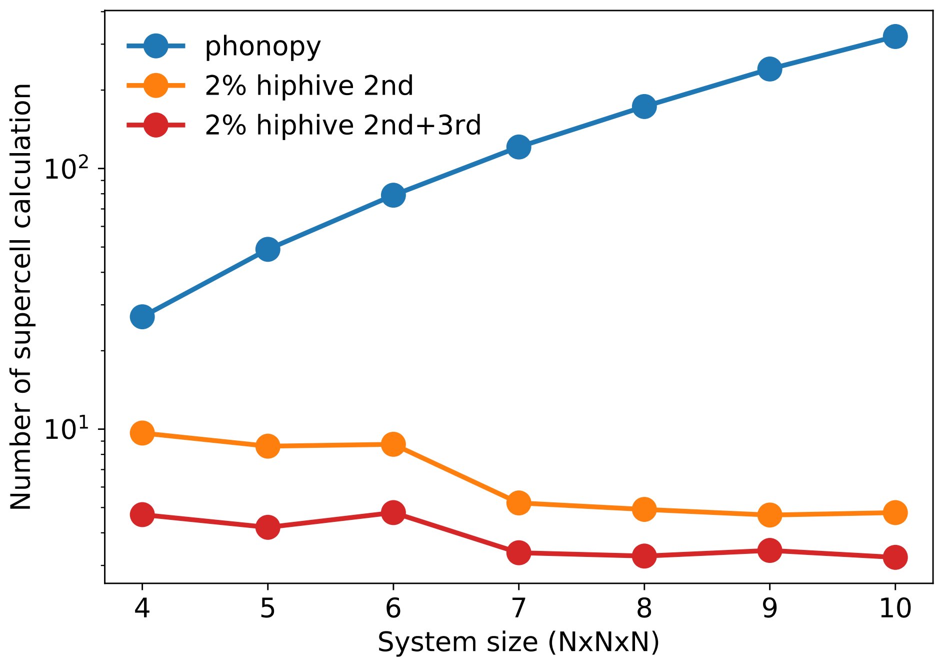
Efficiency 2nd-order: Vacancy in BCC-Ta

hiphive

Convergence with training size



Scaling compared to direct-approach



Favorable size-scaling;
especially interesting for defects

CHALMERS
UNIVERSITY OF TECHNOLOGY

Efficiency 3rd-order: Silicon

Thermal conductivity of Silicon via Boltzmann transport

$$\kappa_l = \frac{1}{2V} \sum_{j\boldsymbol{q}} \lambda_{j\boldsymbol{q}}(T) v_{j\boldsymbol{q}} c_{j\boldsymbol{q}}(T)$$

$$\lambda_{j\boldsymbol{q}}(T) = \tau_{j\boldsymbol{q}}(T) v_{j\boldsymbol{q}}$$

Lifetime ($\tau_{j\boldsymbol{q}}$) — Group velocity ($v_{j\boldsymbol{q}}$)

Requires third-order force constants

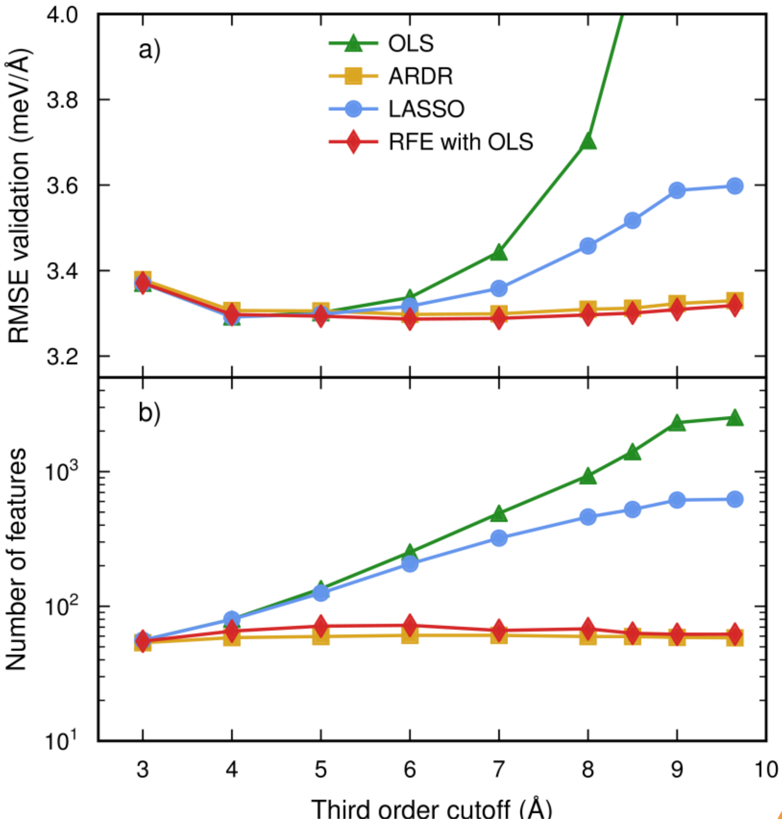
Efficiency 3rd-order: Silicon

hiphive

Thermal conductivity of Silicon via Boltzmann transport

Third-order cutoff selection

4Å



scikit learn

CHALMERS
UNIVERSITY OF TECHNOLOGY

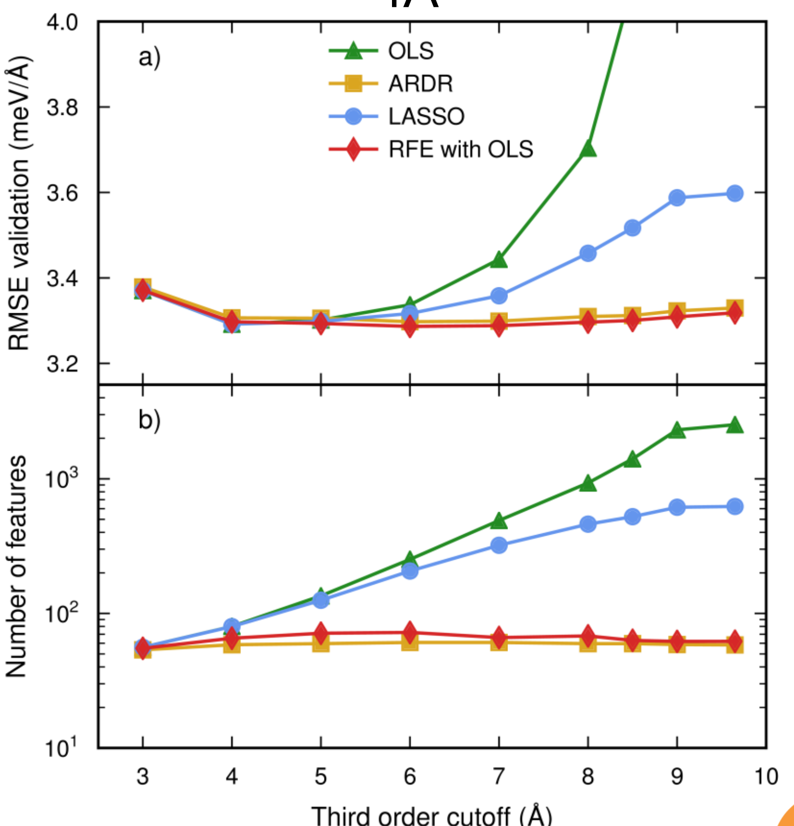
Efficiency 3rd-order: Silicon

hiphive

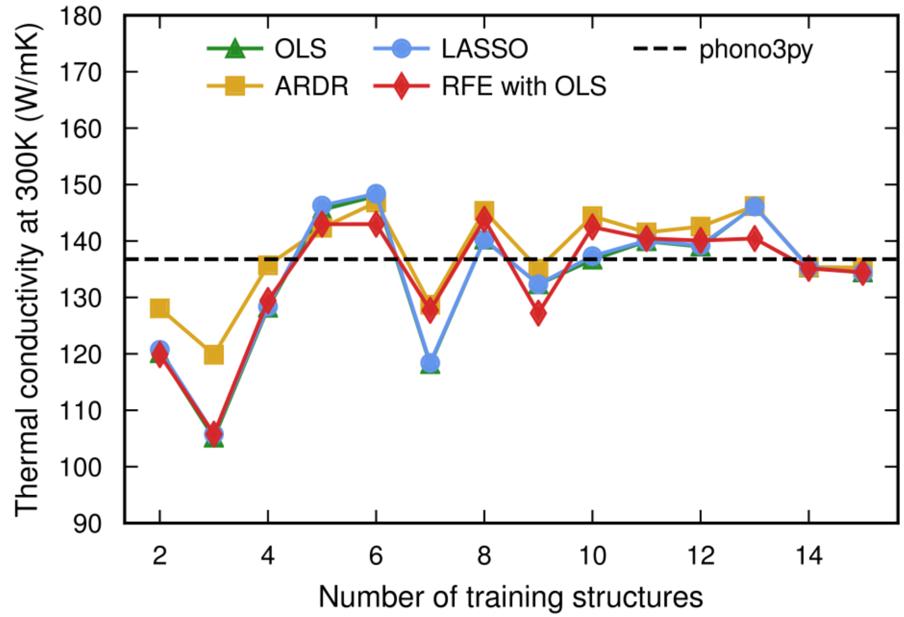
Thermal conductivity of Silicon via Boltzmann transport

Third-order cutoff selection

4Å



Thermal conductivity convergence



phono3py calculation 50-800 structures

CHALMERS
UNIVERSITY OF TECHNOLOGY

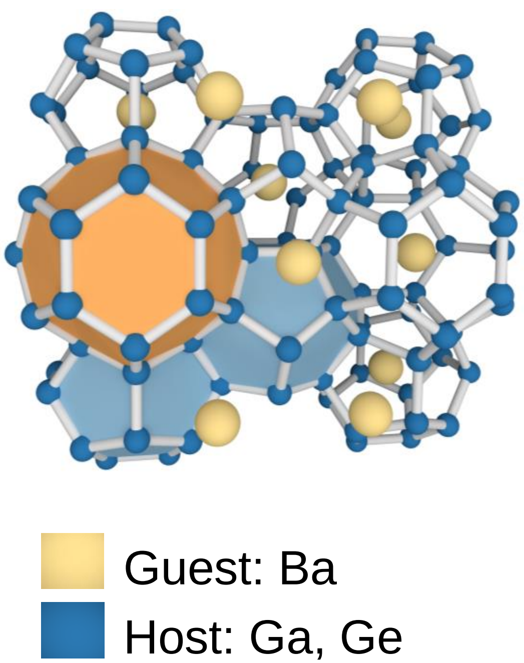
scikit learn

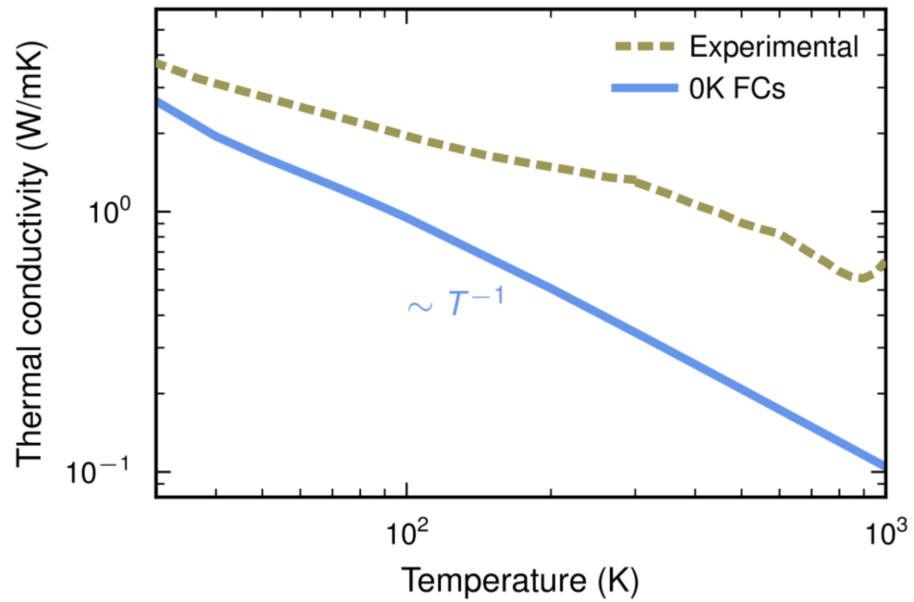
Anharmonicity in clathrates

hiphive

Inorganic clathrate $Ba_8Ga_{16}Ge_{30}$

- Anharmonic Ba (rattler) modes





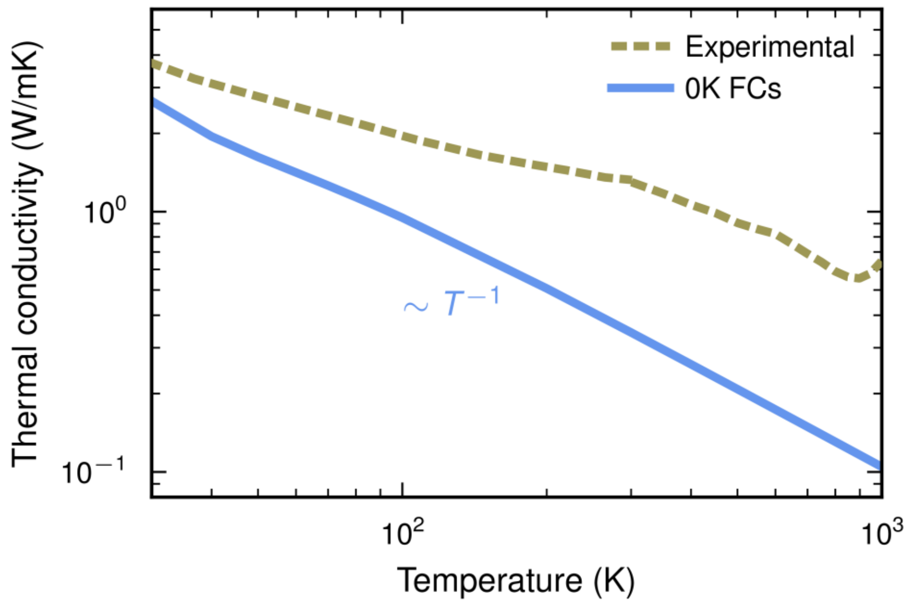
CHALMERS
UNIVERSITY OF TECHNOLOGY

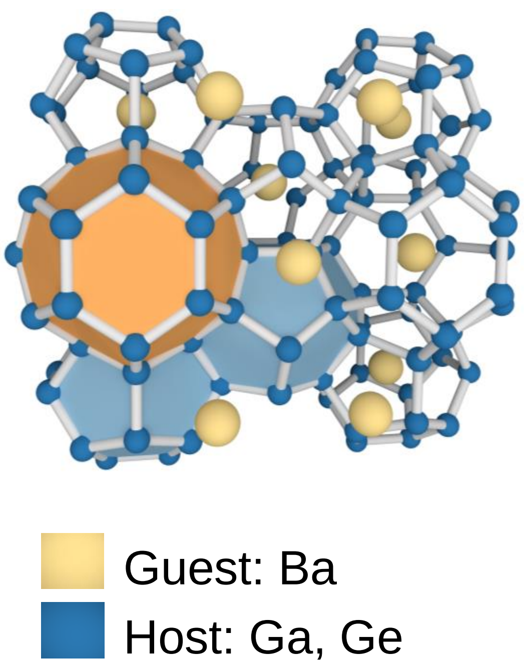
Anharmonicity in clathrates

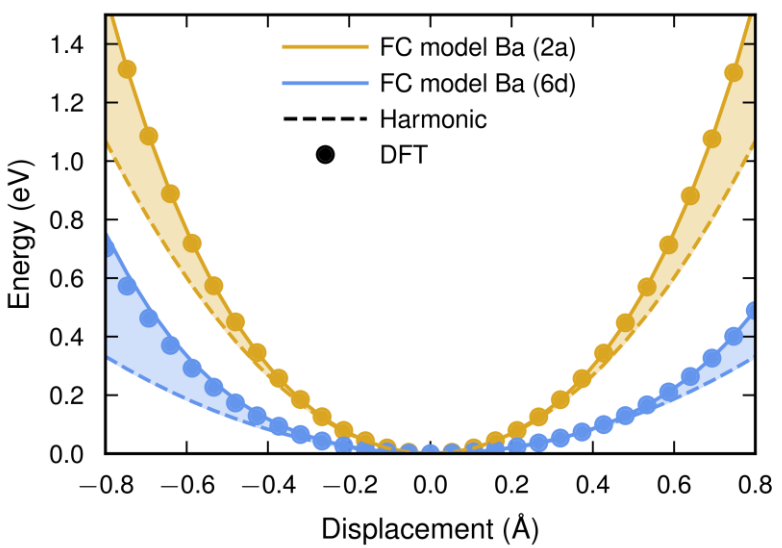
hiphive

Inorganic clathrate $\mathrm{Ba_8Ga_{16}Ge_{30}}$

- Anharmonic Ba (rattler) modes
- Fourth-order model (6000 parameters)
- Reproduces anharmonicity of Ba







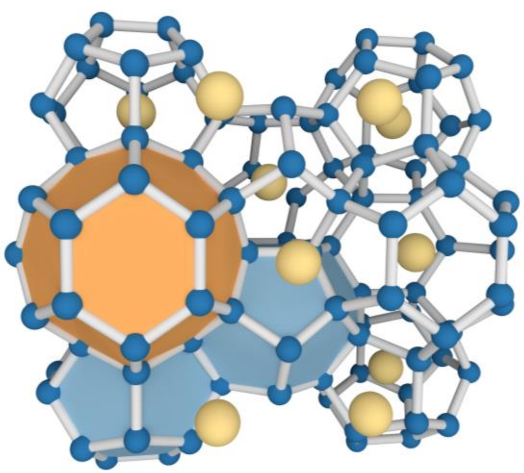
CHALMERS
UNIVERSITY OF TECHNOLOGY

Anharmonicity in clathrates

hiphive

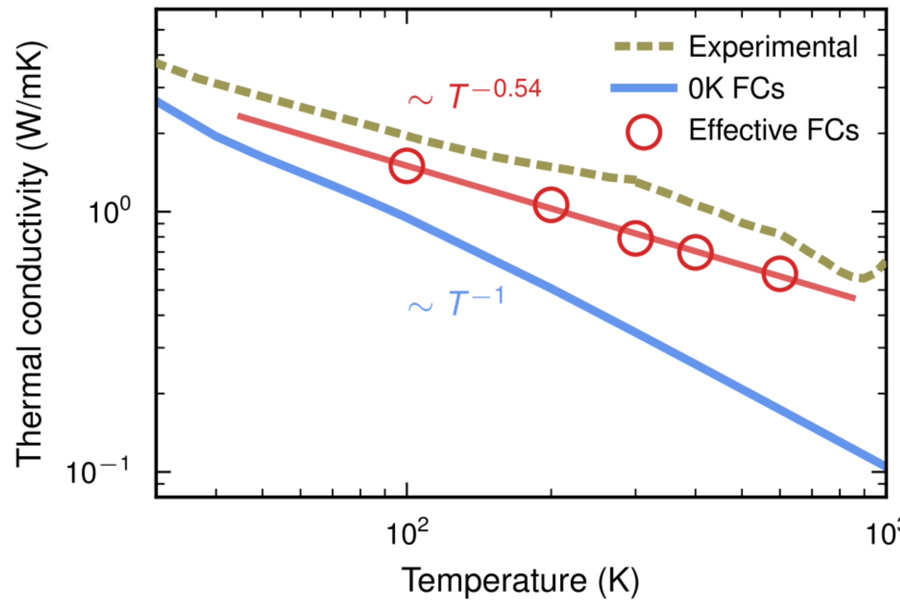
Inorganic clathrate $Ba_8Ga_{16}Ge_{30}$

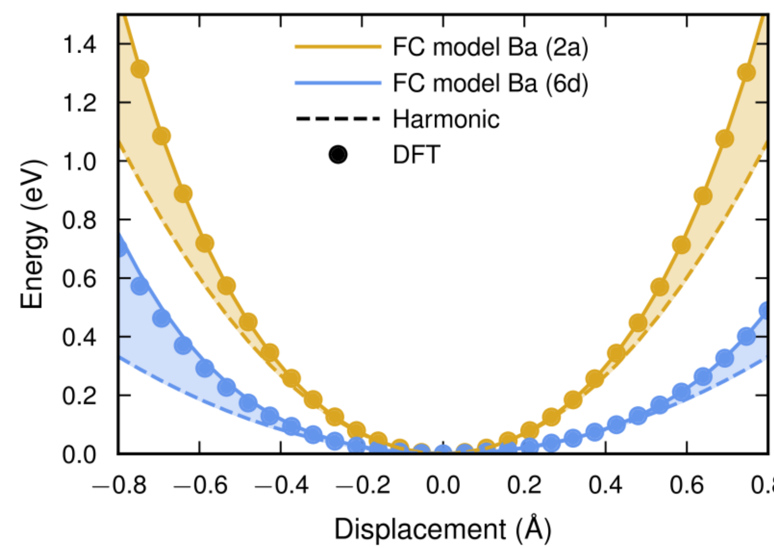
- Anharmonic Ba (rattler) modes
- Fourth-order model (6000 parameters)
- Reproduces anharmonicity of Ba
- Reproduces thermal conductivity



Guest: Ba

Host: Ga, Ge





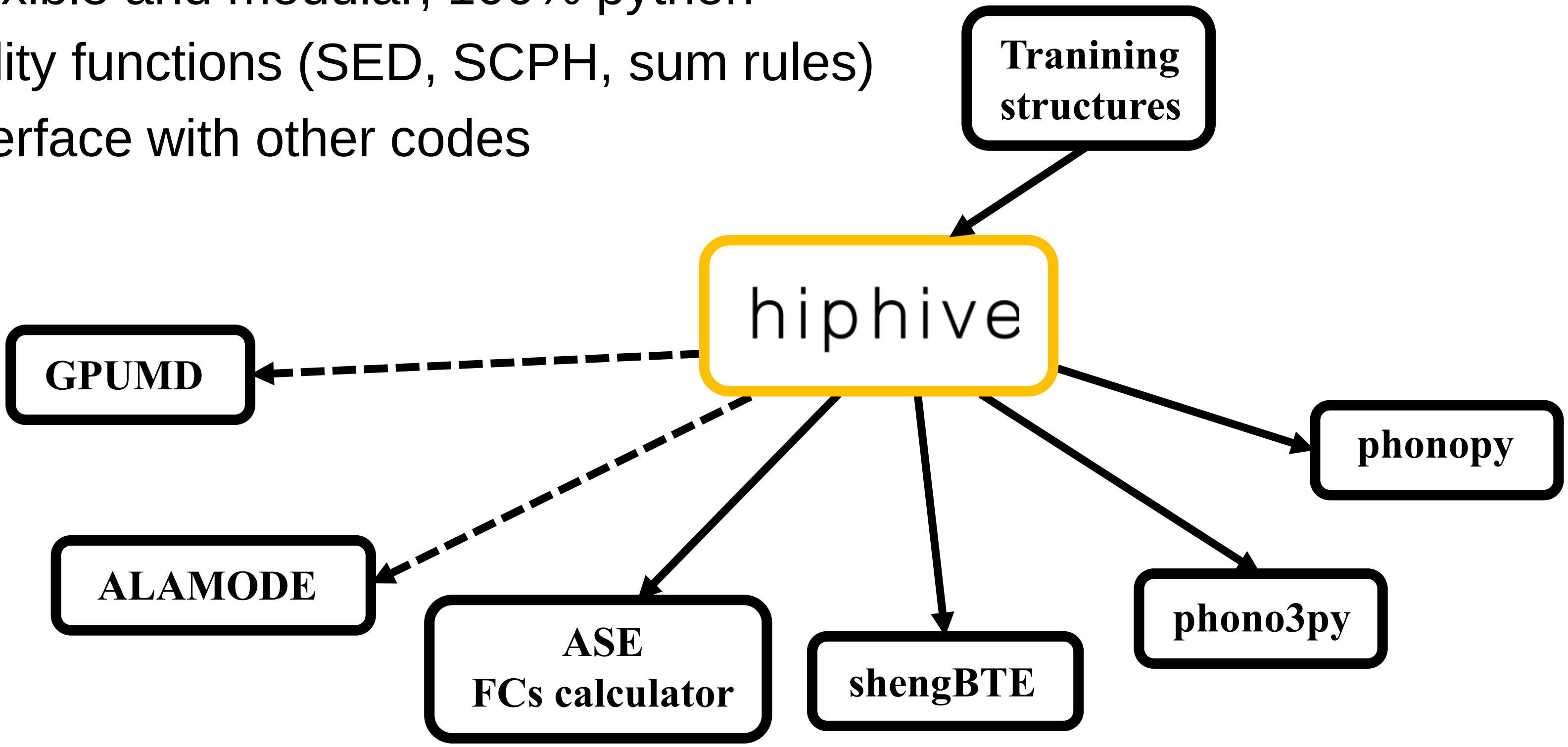
CHALMERS
UNIVERSITY OF TECHNOLOGY

hiphive

hiphive

Force constants extraction tool

- Flexible and modular, 100% python
- Utility functions (SED, SCPH, sum rules)
- Interface with other codes



CHALMERS
UNIVERSITY OF TECHNOLOGY

Eriksson *et al.*, Adv. Theory Simul. **2019**, 1800184 icet

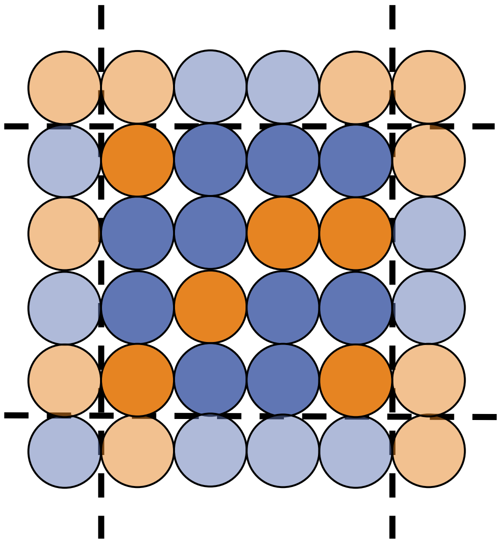
Ångqvist *et al.*, Adv. Theory Simul. **2019**, 1900015

Cluster Expansions

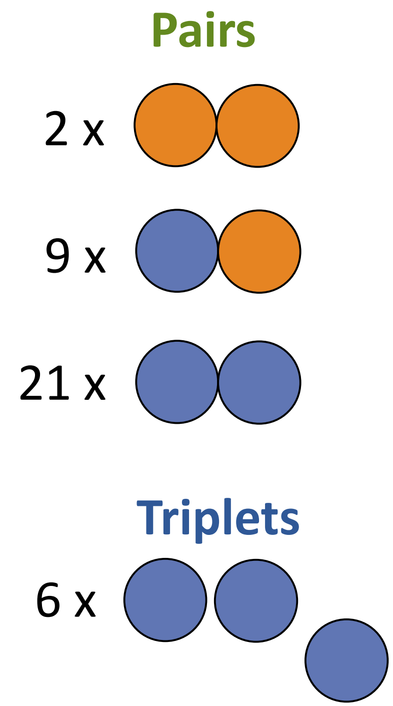
icet

Expansion of the total energy in site occupations

$$E(\boldsymbol{\sigma}) = E_0 + \sum_i J_i \sigma_i + \sum_{ij} J_{ij} \sigma_i \sigma_j + \sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k + \ldots$$



Decompose into “clusters”



$\boldsymbol{\sigma}$ Occupation vector

Need efficient means to extract effective cluster interactions (ECIs) J_α

→ Again a linear problem

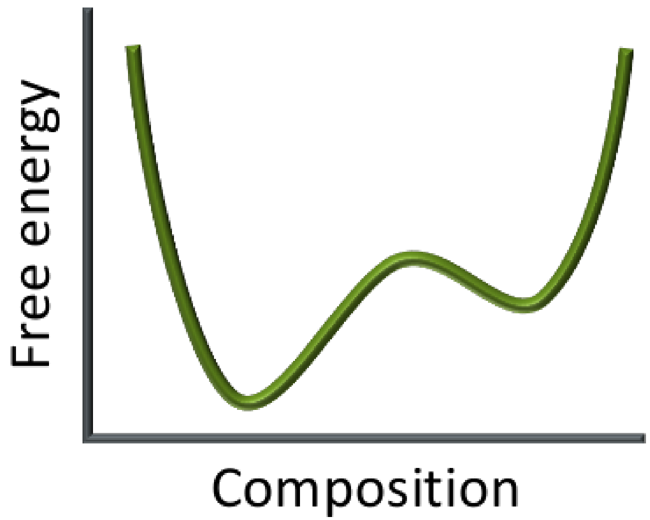
Cluster Expansions

icet

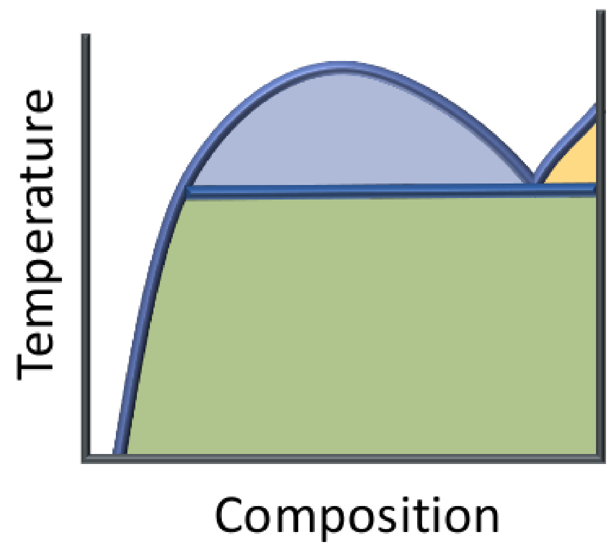
Expansion of the total energy in site occupations

$$E(\boldsymbol{\sigma}) = E_0 + \sum_i J_i \sigma_i + \sum_{ij} J_{ij} \sigma_i \sigma_j + \sum_{ijk} J_{ijk} \sigma_i \sigma_j \sigma_k + \ldots$$

Free energies



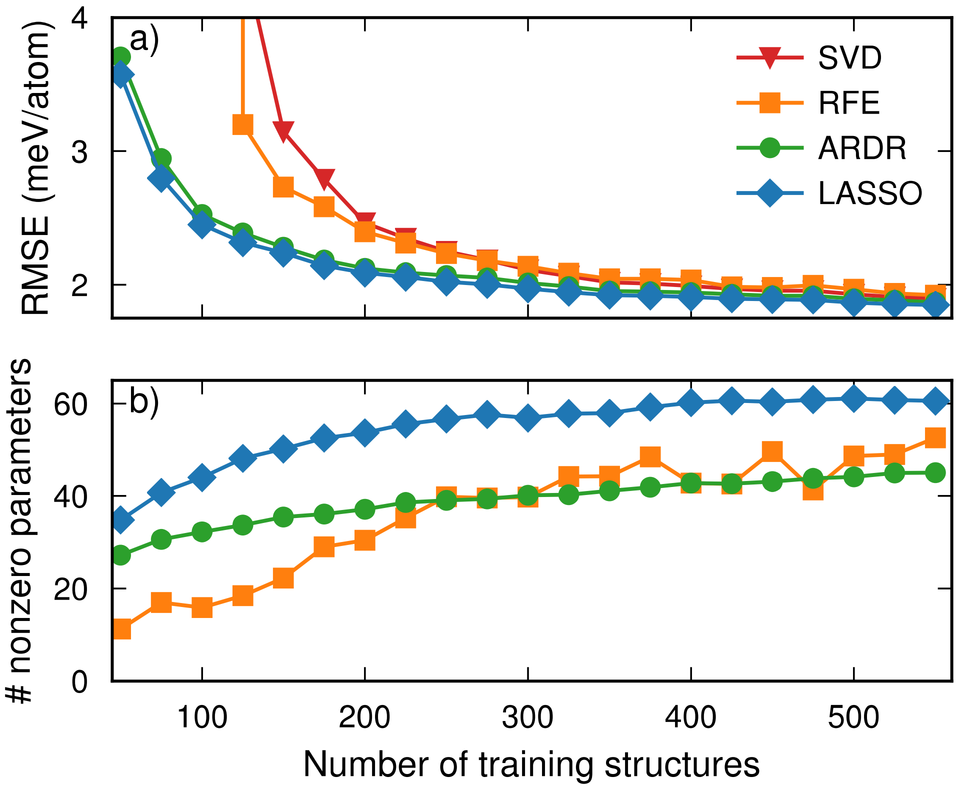
Phase diagrams



CHALMERS
UNIVERSITY OF TECHNOLOGY

CE construction for a Ag-Pd alloy

icet



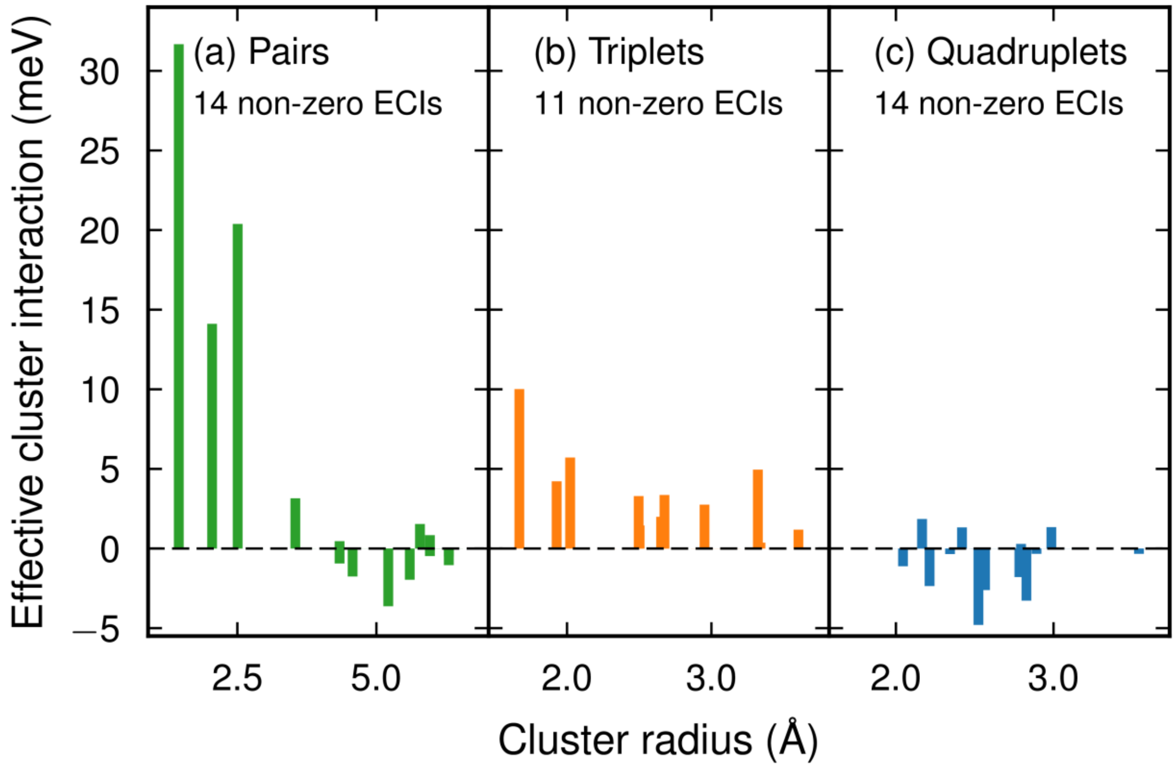
zerolet + singlet
+ 24 pairs
+ 20 triplets
+ 35 quadruplets
= 81 parameters

ARDR: fast convergence and sparse solution
LASSO: more false-positives
RFE: requires more structures

Ångqvist *et al.*, Adv. Theory Simul. **2019**, 1900015

CE construction for a Ag-Pd alloy

icet



zerolet + singlet
+ 24 pairs
+ 20 triplets
+ 35 quadruplets
= 81 parameters

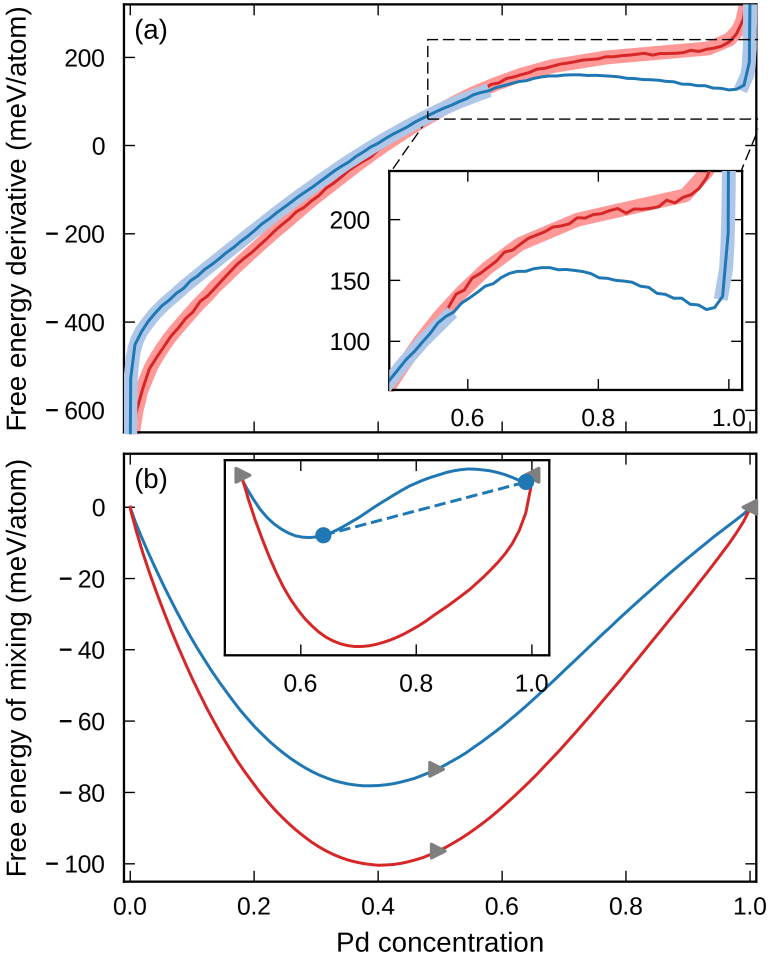
CHALMERS
UNIVERSITY OF TECHNOLOGY

Ångqvist *et al.*, Adv. Theory Simul. **2019**, 1900015

Ag-Pd alloy: CE sampling

icet

Monte Carlo simulations
at 800 K and 400 K



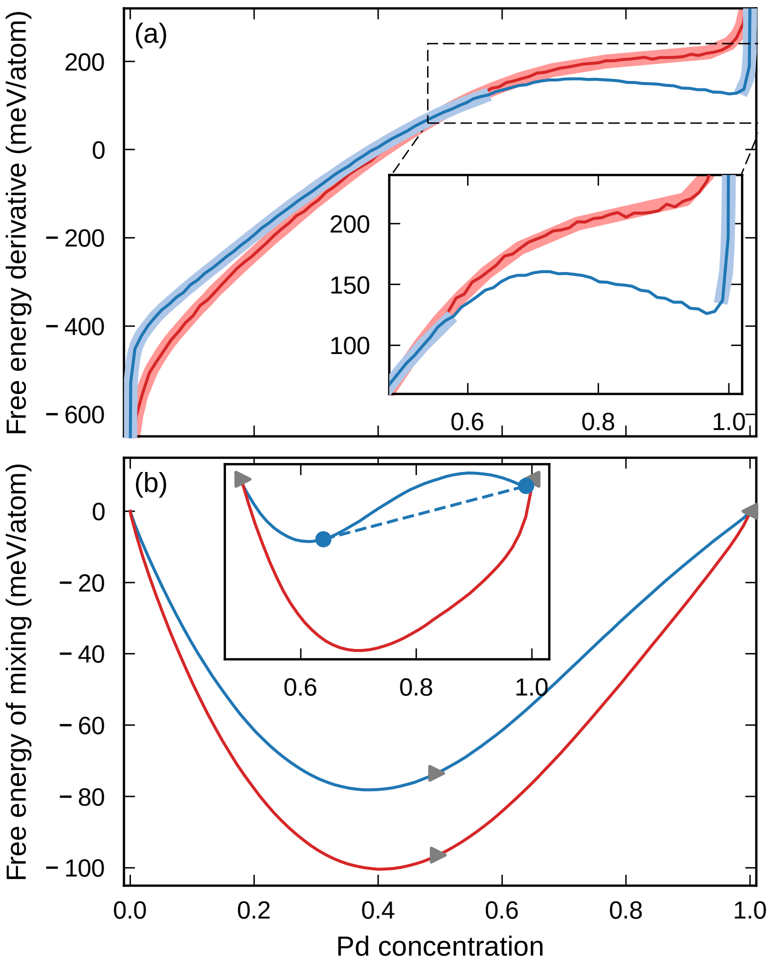
CHALMERS
UNIVERSITY OF TECHNOLOGY

Ångqvist *et al.*, Adv. Theory Simul. **2019**, 1900015

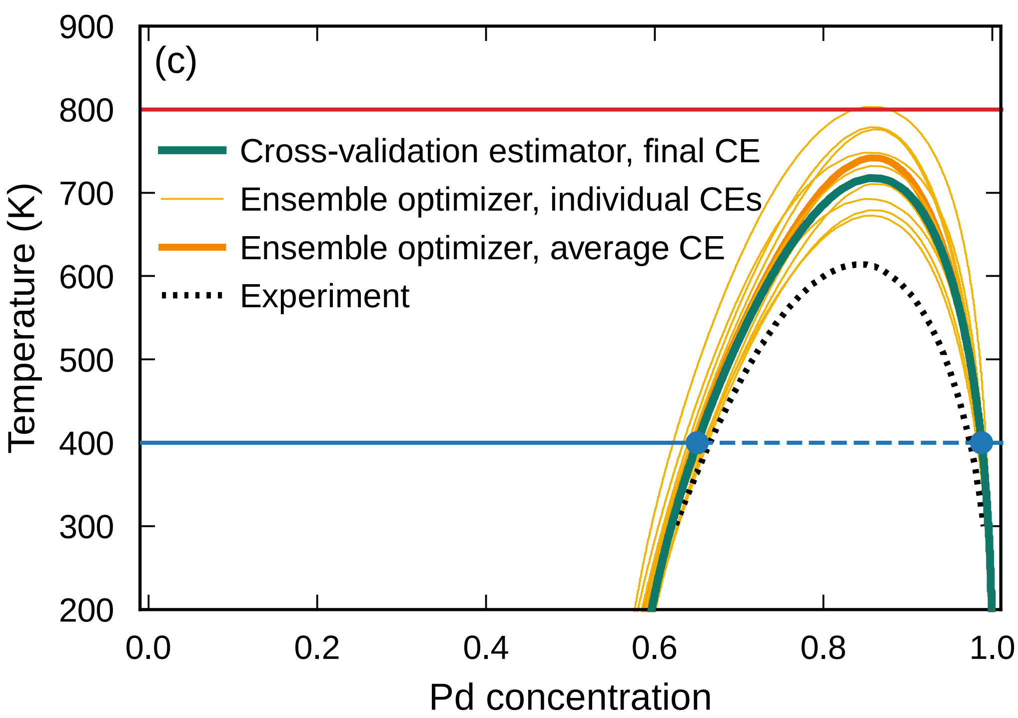
Ag-Pd alloy: CE sampling

icet

Monte Carlo simulations
at 800 K and 400 K



Phase diagram



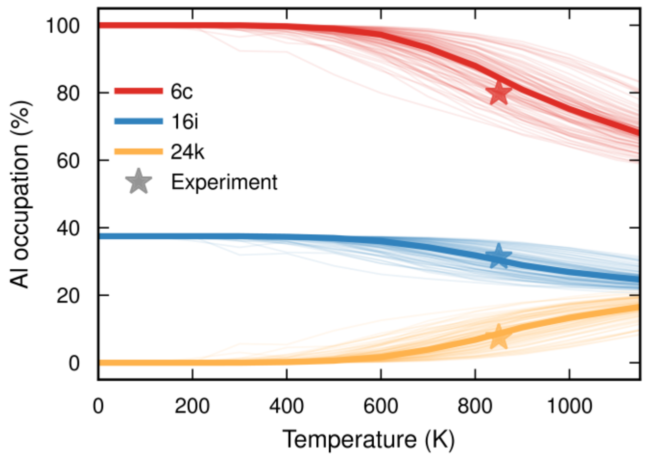
CHALMERS
UNIVERSITY OF TECHNOLOGY

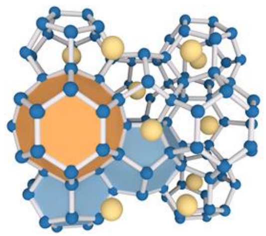
Ångqvist *et al.*, Adv. Theory Simul. **2019**, 1900015

Clathrates

- Chemical order → thermoelectric performance

Ångqvist *et al.* **2017**

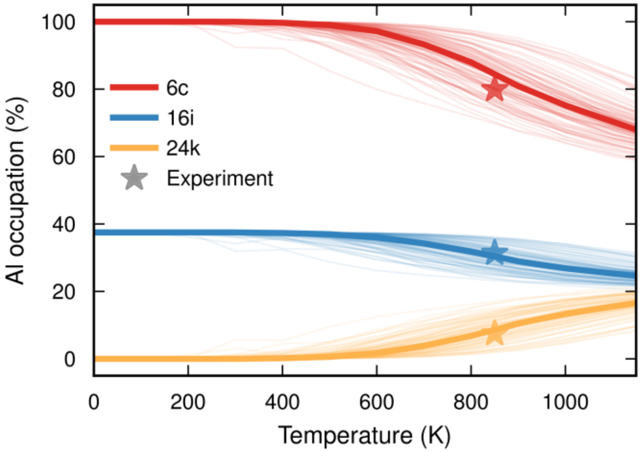


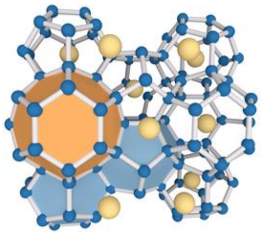


Clathrates

- Chemical order → thermoelectric performance

Ångqvist *et al.* **2017**

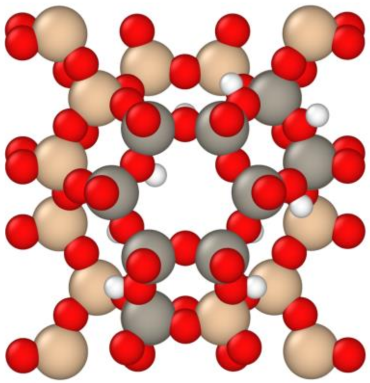




Zeolites

- Aluminium distribution on framework
- Löwenstein's rule violated

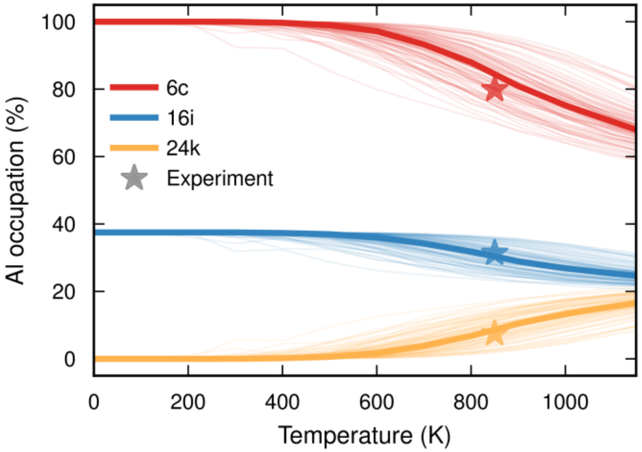
Magnus Fant 2019

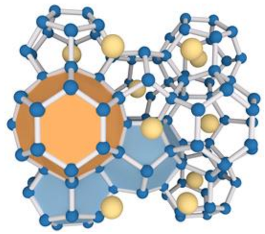


Clathrates

- Chemical order → thermoelectric performance

Ångqvist *et al.* **2017**

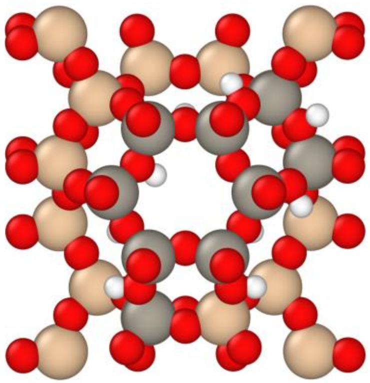




Zeolites

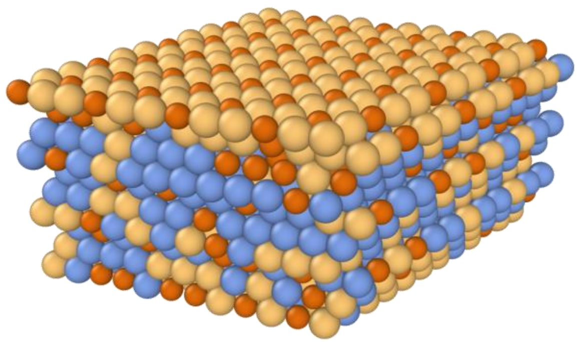
- Aluminium distribution on framework
- Löwenstein's rule violated

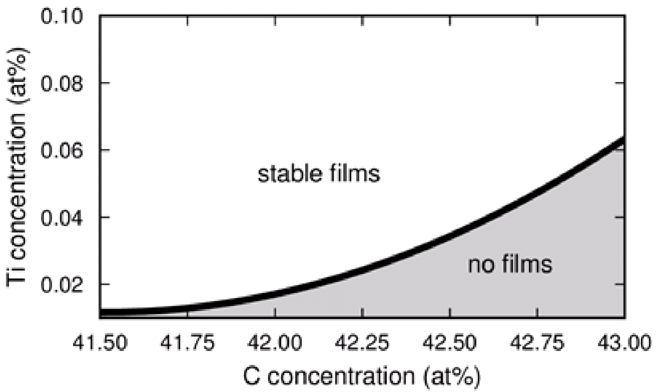
Magnus Fant 2019



Interface systems

- Surface segregation in PdAuCu (Pernilla Tanner 2019)
- Stability of WC-Co interfacial phases (Martin Gren 2019)





CHALMERS
UNIVERSITY OF TECHNOLOGY

icet features

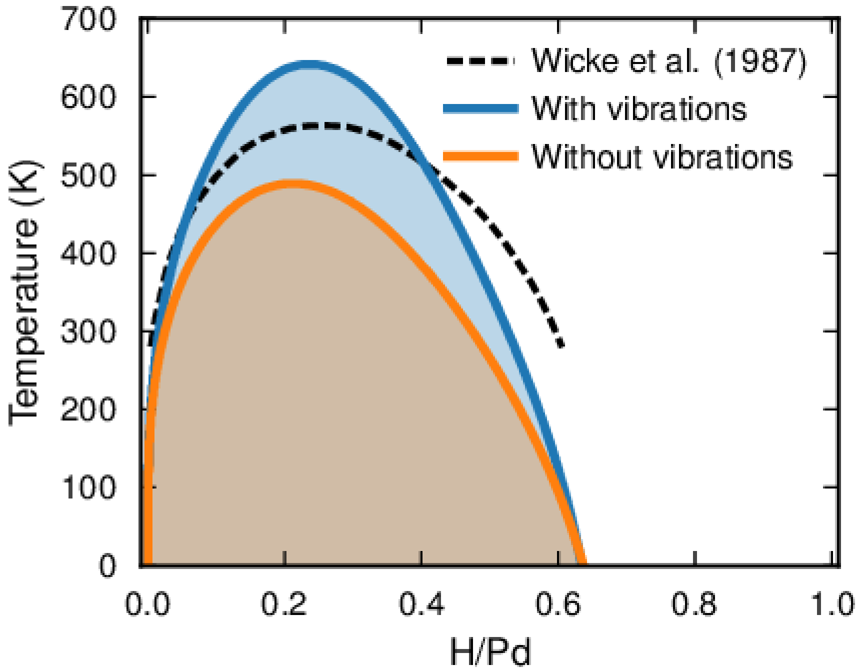
- **Structure generation**
  - Enumeration
  - SQS (Special Quasi-random Structures)
  - Ground-states via MIP (Mixed Integer programming)
- **Training via scikit-learn**
  - LASSO, ARDR, RFE, …
- **Monte Carlo sampling**
  - Canonical ensemble
  - Semi-Grand canonical ensemble
  - Variance-constrained semi-grand canonical ensemble
  - Simulated annealing
  - Wang-Landau sampling

Future work: Configurational + Vibrational

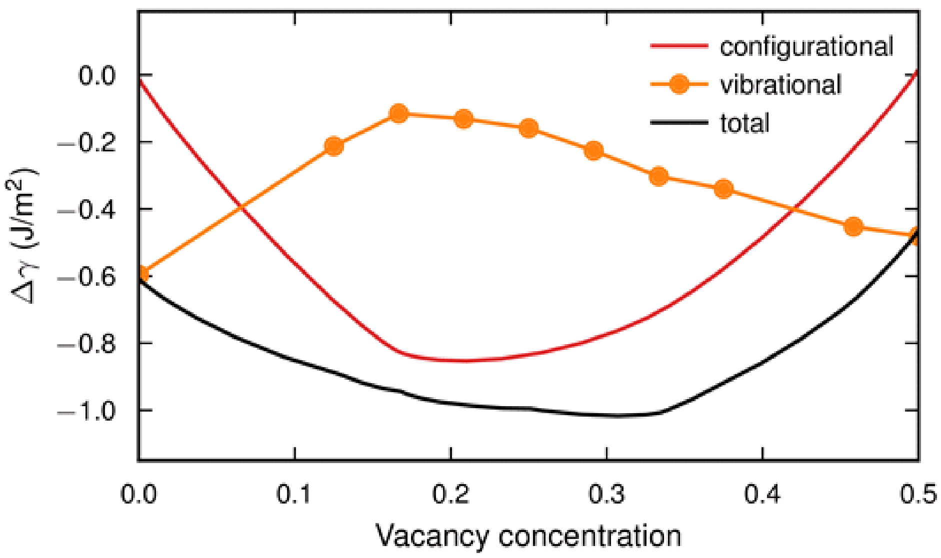
$$\mathcal{Z} = \mathcal{Z}_{\mathrm{conf}}\mathcal{Z}_{\mathrm{vib}}$$

Coupling between configurational and vibrational dofs
→ Very computational expensive

Alloys



WC phases



CHALMERS
UNIVERSITY OF TECHNOLOGY

Tutorials

icet tutorial Thursday 1pm

```
# setup
cs = ClusterSpace(atoms, [6.0, 5.0], ['Ag', 'Pd'])
sc = StructureContainer(cs)
for structure in training_structures:
    sc.add_structure(structure)

# training
opt = Optimizer(sc.get_fit_data(), fit_method='ardr')
opt.train()
ce = ClusterExpansion(cs, opt.parameters)

# sample
calc = ClusterExpansionCalculator(ce, atoms)
mc = CanonicalEnsemble(atoms, calc, temperature=600)
mc.run(1000)
```

hiphive tutorial Thursday 3pm

```
# setup
cs = ClusterSpace(atoms, [8.0, 6.0])
sc = StructureContainer(cs)
for structure in training_structures:
    sc.add_structure(structure)

# training
opt = Optimizer(sc.get_fit_data(), fit_method='rfe')
opt.train()
fcp = ForceConstantPotential(cs, opt.parameters)

# write to phono3py format
fcs = fcp.get_force_constants(supercell)
fcs.write_to_phonopy('fc2.hdf5')
fcs.write_to_phono3py('fc3.hdf5')
```

The end

icet

- Mattias Ångqvist
- William Muñoz
- Magnus Rahm
- Erik Fransson
- Joakim Brorsson
- Céline Durniak
- Piotr Rozyczko
- Thomas Rod
- Paul Erhart

hiphive

- Fredrik Eriksson
- Erik Fransson
- Paul Erhart

EUROPEAN SPALLATION SOURCE

https://materialsmodeling.org/software

CHALMERS UNIVERSITY OF TECHNOLOGY

Knut och Alice Wallenbergs Stiftelse

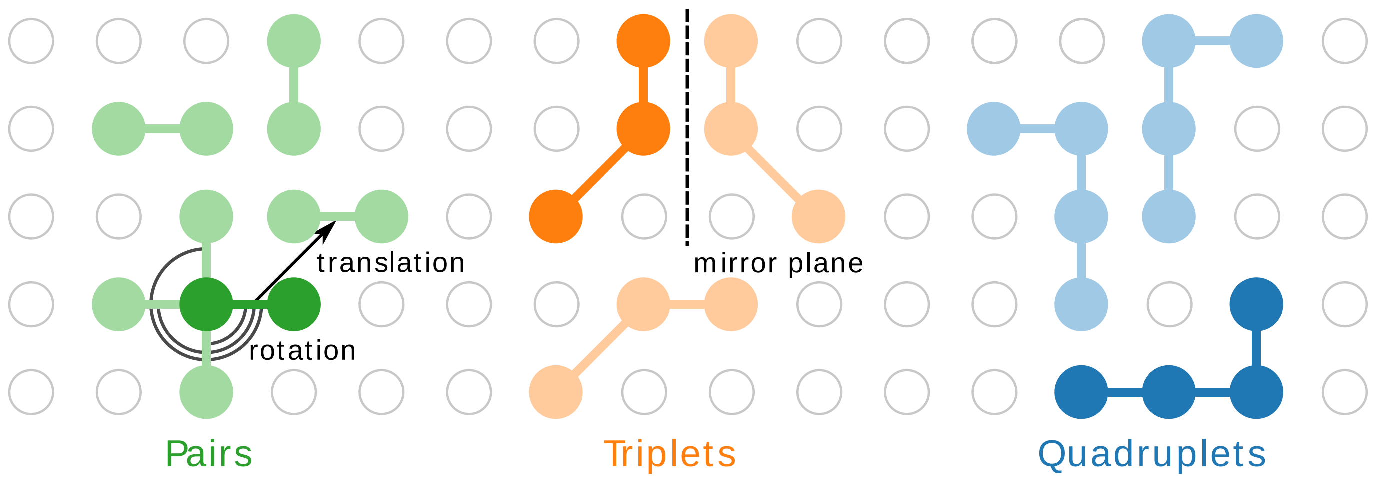
Vetenskapsrådet

Extra

Cluster decomposition

General procedure

1. Generate list of clusters
2. Obtain symmetry operations
3. Categorize clusters into orbits
4. Identify independent parameters
5. Apply sum rules

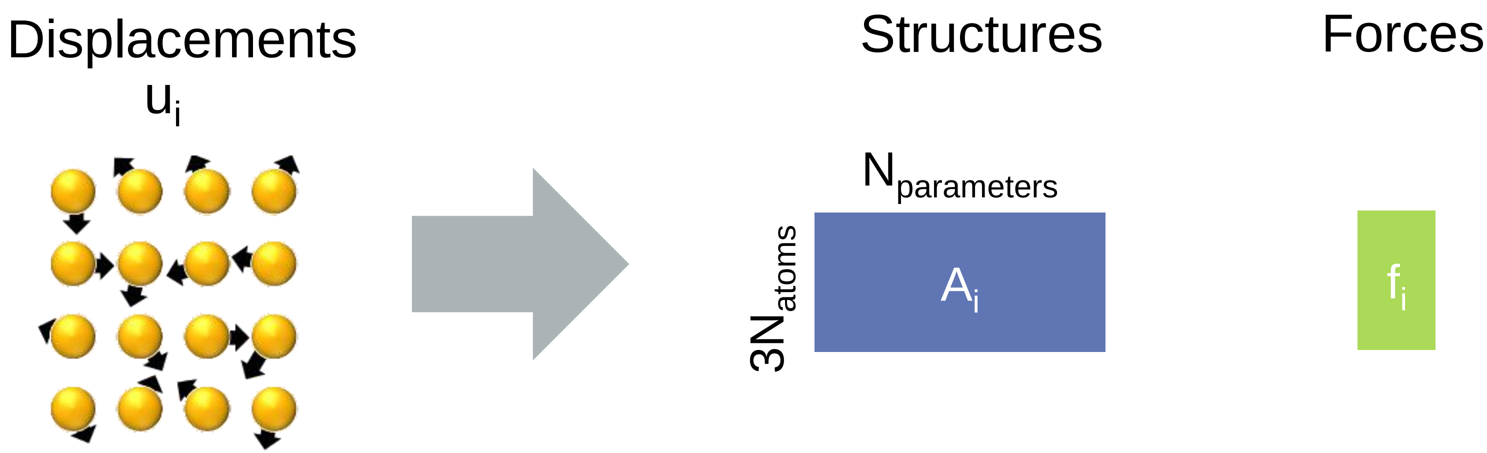


Structure decomposition

1. Generate structure e.g., by applying random displacements or superposing normal modes

2. Convert displacements u_i into “cluster vectors”
   → each structure yields a matrix A_i with $N_{parameters}$ columns and $3N_{atoms}$ rows

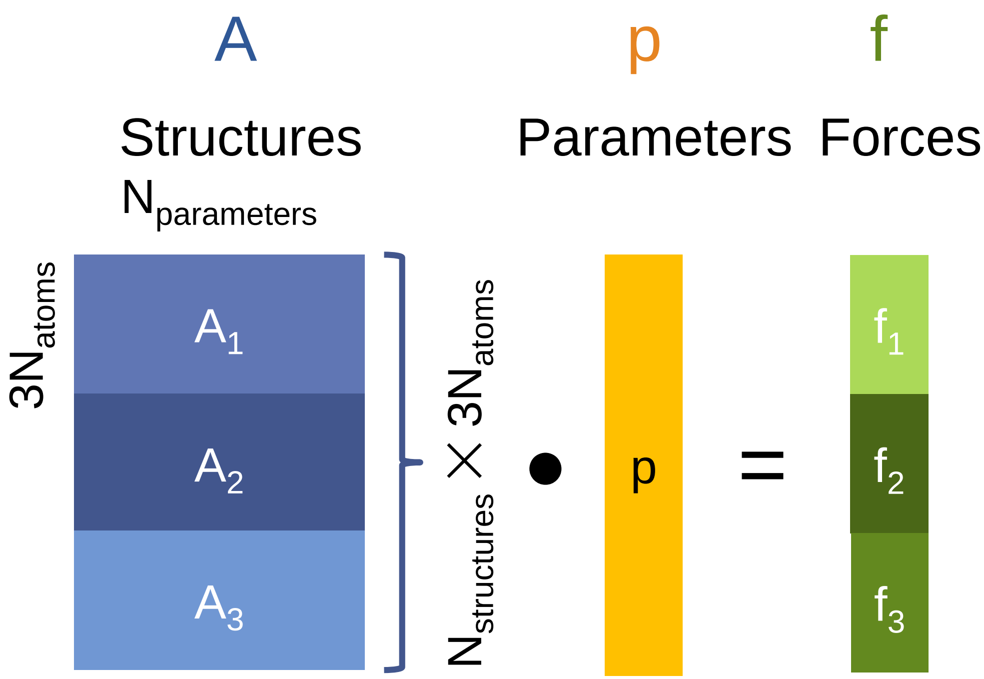
3. Each structure comes with a target force vector



Parameter optimization

hiphive

1. Compile multiple structures into one (large) fit matrix A
2. Solve the linear problem Ap=f



How?

Least-squares

$$||\boldsymbol{Ap} - \boldsymbol{f}||_2 < \epsilon$$

Ridge-regression/LASSO/

$$||\boldsymbol{Ap} - \boldsymbol{f}||_2^2 + \alpha||\boldsymbol{p}||_1 + \beta||\boldsymbol{p}||_2^2$$

Automatic relevance determination regression (ARDR)

Recursive feature elimination (RFE)

+ Cross-validation