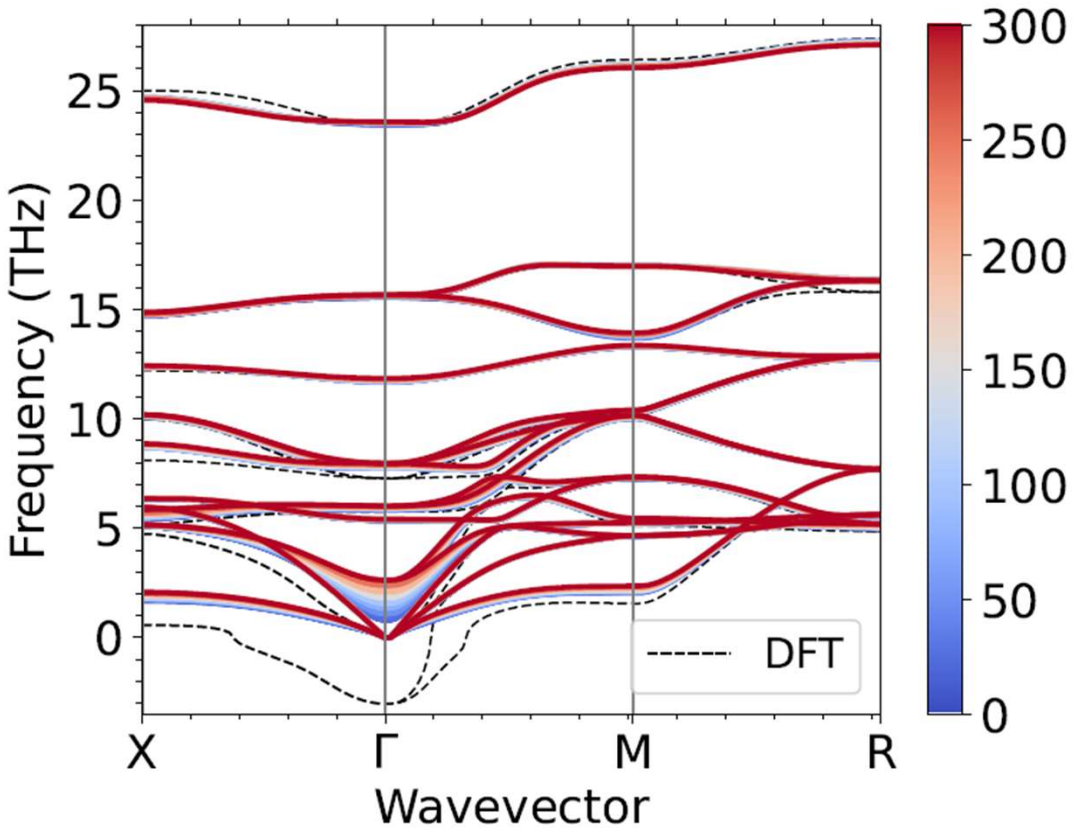
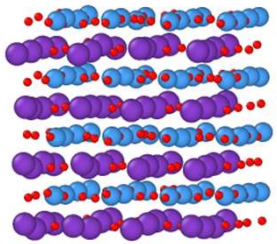



# Machine learning force fields for materials science

Collège de France – 06/06/2023
Ambroise van Roekeghem

# Machine learning force fields/potentials



atomic positions

(ML) FF → Energy, forces, stress

Reference data: energy, forces, stress from Density Functional Theory

Atomic positions are used to build descriptors of the local environments

Machine learning methods can be for instance:

- Neural Networks
- Gaussian processes
- Linear methods

PRL 104, 136403 (2010) PHYSICAL REVIEW LETTERS week ending 2 APRIL 2010

**Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons**

Albert P. Bartók and Mike C. Payne
*Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge, CB3 0HE, United Kingdom*

Risi Kondor
*Center for the Mathematics of Information, California Institute of Technology, MC 305-16, Pasadena, California 91125, USA*

Gábor Csányi
*Engineering Laboratory, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom*
(Received 1 October 2009; published 1 April 2010)

PRL 98, 146401 (2007) PHYSICAL REVIEW LETTERS week ending 6 APRIL 2007

**Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces**

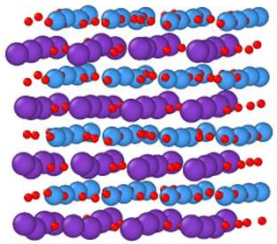
Jörg Behler and Michele Parrinello
*Department of Chemistry and Applied Biosciences, ETH Zurich, USI-Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland*
(Received 27 September 2006; published 2 April 2007)

MULTISCALE MODEL. SIMUL.
Vol. 14, No. 3, pp. 1153–1173


MOMENT TENSOR POTENTIALS: A CLASS OF SYSTEMATICALLY IMPROVABLE INTERATOMIC POTENTIALS*

ALEXANDER V. SHAPEEV†

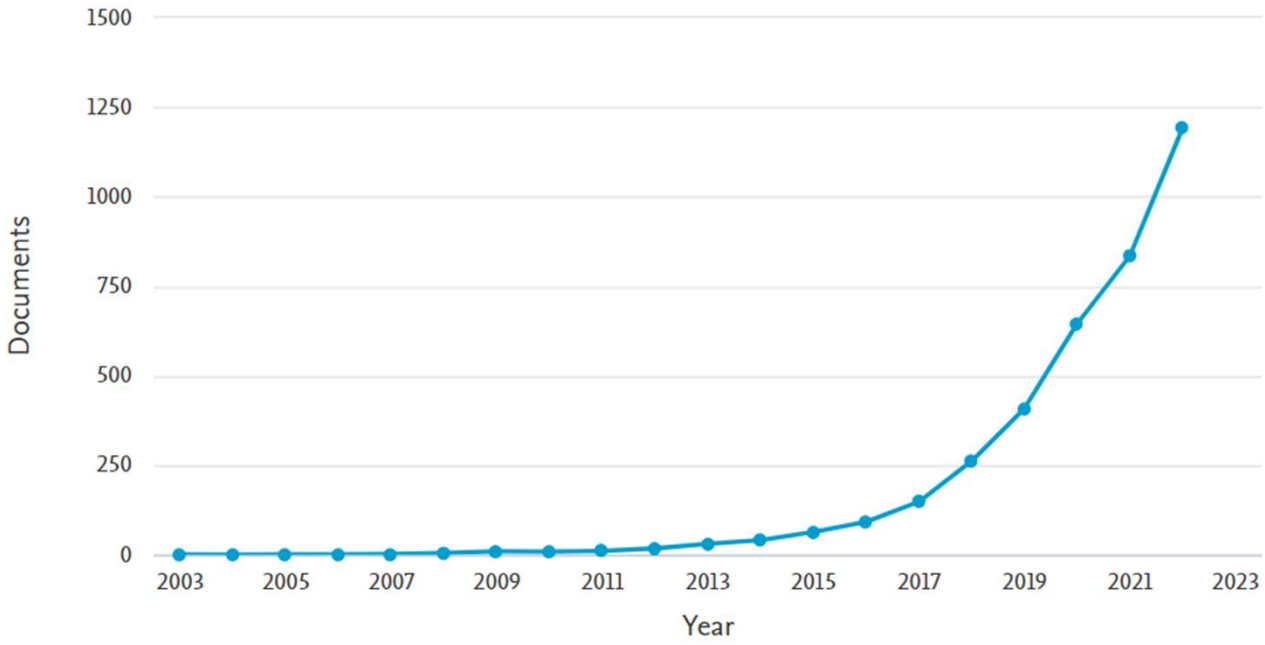
# Machine learning force fields/potentials



atomic positions

(ML) FF → Energy, forces, stress

Documents by year

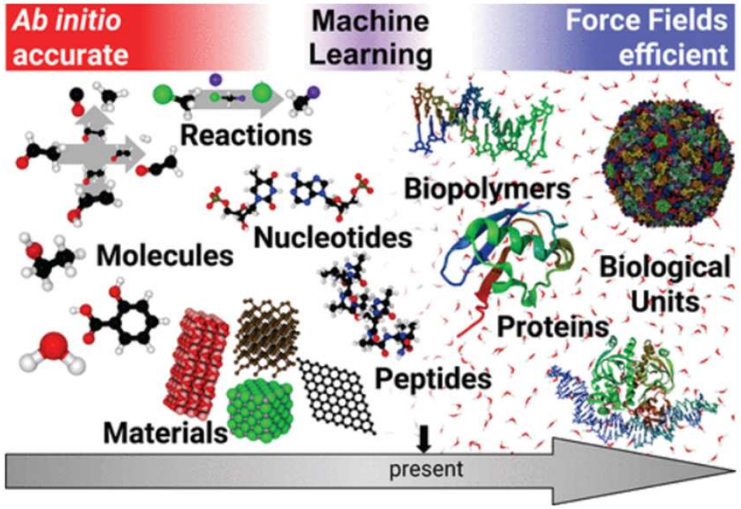



( ALL ( "machine learning force field" ) OR ALL ( "neural network potential" ) OR ALL ( "machine learning potential" ) ) AND ( LIMIT-TO ( SUBJAREA , "PHYS" ) OR LIMIT-TO ( SUBJAREA , "CHEM" ) OR LIMIT-TO ( SUBJAREA , "MATE" ) )

# Why using (machine learning) force fields?

To perform Molecular Dynamics, Monte Carlo simulations, phonon calculations… :

- structure of liquids
- alloys
- extended defects
- surfaces and interfaces
- vibrational properties
- diffusion
- fracture
- macromolecules
- reactions
- etc…




Unke et al., Chem. Rev. 121, 10142 (2021)

Classical force fields are not very accurate quantitatively and not easily parametrized
Density Functional Theory is much slower and typically scales badly with system size

→ Machine learning potentials aim at keeping DFT accuracy at lower cost to access larger systems/longer time scales with ab initio data.

# New paradigm in interatomic potentials

EUROPHYSICS LETTERS 10 June 1994

*Europhys. Lett.*, **26** (8), pp. 583-588 (1994)

## Interatomic Potentials from First-Principles Calculations: the Force-Matching Method.

F. ERCOLESSI (*) and J. B. ADAMS (**)

(*) *Materials Research Laboratory, University of Illinois*
*104 S. Goodwin Ave., Urbana, IL 61801, USA*
*International School for Advanced Studies (SISSA-ISAS)*
*Via Beirut 4, I-34014 Trieste, Italy* (§)
(**) *Department of Materials Science and Engineering, University of Illinois*
*105 S. Goodwin Ave., Urbana, IL 61801, USA*

(received 3 January 1994; accepted in final form 22 April 1994)

PACS. 34.20 – Interatomic and intermolecular potentials and forces.
PACS. 61.50L – Crystal binding.
PACS. 64.70D – Solid-liquid transitions.

**Abstract.** – We present a new scheme to extract numerically «optimal» interatomic potentials from large amounts of data produced by first-principles calculations. The method is based on fitting the potential to *ab initio* atomic forces of many atomic configurations, including surfaces, clusters, liquids and crystals at finite temperature. The extensive data set overcomes the difficulties encountered by traditional fitting approaches when using rich and complex analytic forms, allowing to construct potentials with a degree of accuracy comparable to that obtained by *ab initio* methods. A glue potential for aluminium obtained with this method is presented and discussed.

# There is much more DFT data available, with high reproducibility

Precision of DFT methods is of the order of 1meV/atom, similar to experimental precision

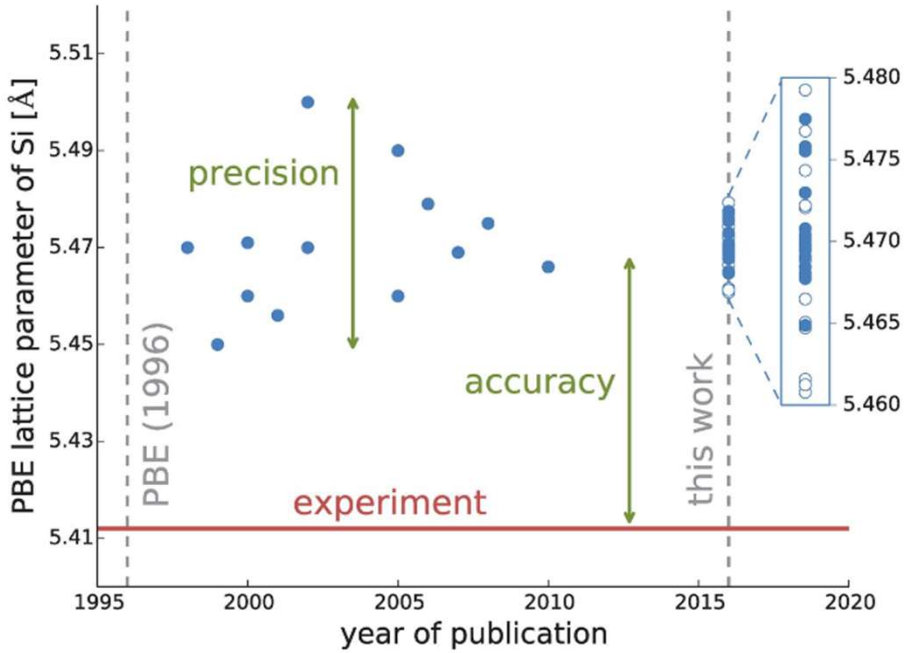



Fig. 1 Historical evolution of the predicted equilibrium lattice parameter for silicon.

Lejaeghere et al., Science 351, 6280 (2016)

Large repositories of DFT data available (hundreds of millions calculations)

MATERIALS PROJECT

AFLOW
Automatic - FLOW for Materials Discovery

NOMAD

OQMD
The Open Quantum Materials Database

But… this does not mean that all data are comparable enough to be in the same training set. At the end many people end up building a dedicated training set for their specific question.

# ML potentials have a more general/less physical form than classical potentials, with many parameters

Typical interatomic potentials use 10s-100s parameters, with the local atomic structure descriptors entering fixed-form functions

ML potentials use 100s-1000s parameters, with possibilities to systematically increase the size of the basis of environment descriptors

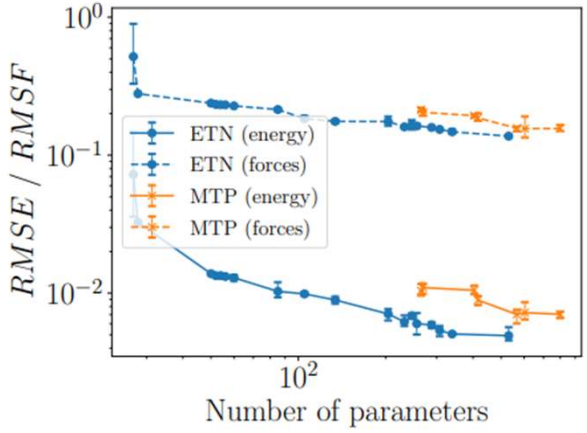
Training set size is of the order of 100 000s – millions of atoms computed

RMSE can be of the order of meVs/atom, close to DFT precision

**Table 1 Fitting and validation errors of the straight NN and PINN models**

| Model | NN architecture | Number of parameters | RMSE of training (meV per atom) | RMSE of validation (meV per atom) |
|---|---|---|---|---|
| NN | 60 × 16 × 16 × 1 | 1265 | 3.36 | 3.85 |
| NN′ | 47 × 18 × 18 × 1 | 1225 | 3.62 | 3.54 |
| PINN | 60 × 15 × 15 × 8 | 1283 | 3.46 | 3.59 |

Purja Pun et al., Nature Comm. 10, 2339 (2019)




Hodapp and Shapeev, arXiv:2304.08226)

# Many ML potentials assume locality

Cutoff: typically of the order of 5-10 Angstroms
Loss function: energies, forces, stress…

$$E = \sum_i E_i.$$

$$f_c(R_{ij}) = \begin{cases} 0.5 \times \left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1\right] & \text{for } R_{ij} \le R_c, \\ 0 & \text{for } R_{ij} > R_c. \end{cases} \quad (3)$$

Radial symmetry functions are constructed as a sum of Gaussians with the parameters $\eta$ and $R_s$,

$$G_i^1 = \sum_{j \neq i}^{\text{all}} e^{-\eta(R_{ij}-R_s)^2} f_c(R_{ij}). \quad (4)$$

Angular terms are constructed for all triplets of atoms by summing the cosine values of the angles $\theta_{ijk} = \frac{\mathbf{R_{ij}} \cdot \mathbf{R_{ik}}}{R_{ij} R_{ik}}$ centered at atom $i$, with $\mathbf{R_{ij}} = \mathbf{R_i} - \mathbf{R_j}$,

$$G_i^2 = 2^{1-\zeta} \sum_{j,k \neq i}^{\text{all}} (1 + \lambda \cos\theta_{ijk})^\zeta \times e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}), \quad (5)$$

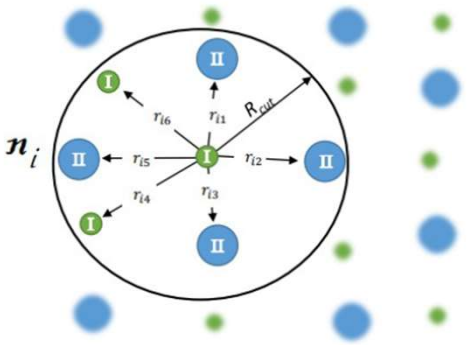
choice of symmetry functions and their parameters is not unique nor does it need to be, and many types of functions can be used, as long as the set of function values is suitable for describing the environment of an atom.

Behler and Parinello, Phys. Rev. Lett. 98, 146401 (2007)

$$M_{\mu,\nu}(n_i) = \sum_j f_\mu(|r_{ij}|, z_i, z_j)\, \underbrace{r_{ij} \otimes \ldots \otimes r_{ij}}_{\nu \text{ times}},$$

Radial: depends on relative distances and types of atoms

Angular: tensor resembling moments of inertia




Gubaev et al., Comput. Mater. Sci 156, 148 (2019)

# Many flavors of possible descriptors

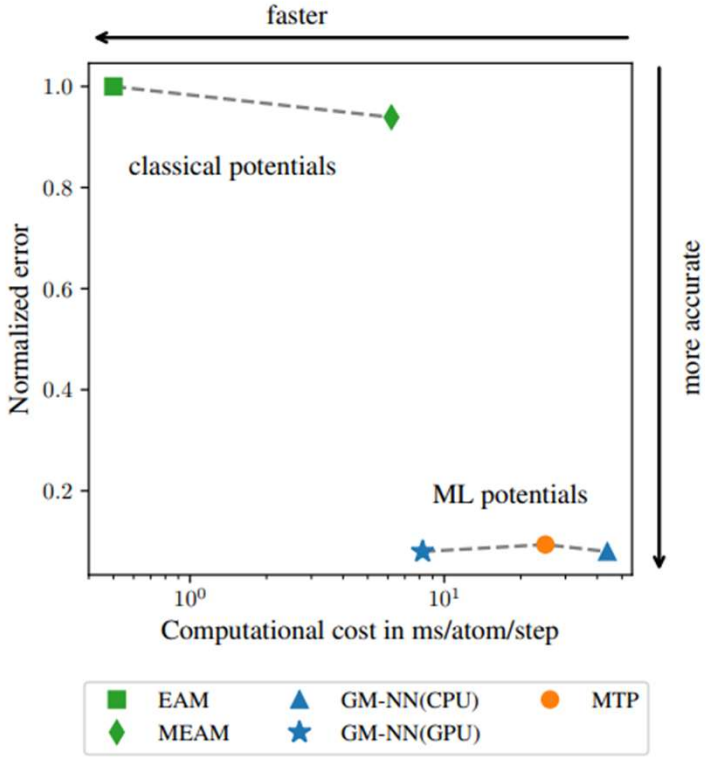
**Table 1. List of Some Typical Descriptors Frequently Used in Machine Learning Potentials**[a]

| descriptor | year | ref |
|---|---|---|
| atom-centered symmetry functions | 2007 | 50,75 |
| bispectrum | 2010 | 57,60 |
| Coulomb matrix | 2012 | 64 |
| SOAP | 2013 | 107 |
| permutation invariant polynomials | 2013 | 69,90 |
| Ewald sum matrix | 2015 | 122 |
| bag of bonds | 2015 | 113 |
| overlap matrix | 2016 | 108 |
| polynomials in MTPs | 2016 | 59 |
| spherical harmonics | 2017 | 109 |
| Chebyshev polynomials | 2017 | 123 |
| many-body tensor representation | 2017 | 115 |
| histogram of internal coordinates | 2017 | 124 |
| FCHL | 2018 | 112,125 |
| weighted symmetry functions | 2018 | 126 |
| smoothed atomic densities | 2019 | 114 |
| orthogonal descriptors | 2019 | 110 |
| long-distance equivariant repres. | 2019 | 127 |

[a]The given year refers to the introduction of the descriptor or, in the case of descriptors which have been used in other contexts before, they refer to the first use in machine learning potentials.

Behler, Chem. Rev. Lett. 121, 10037 (2021)

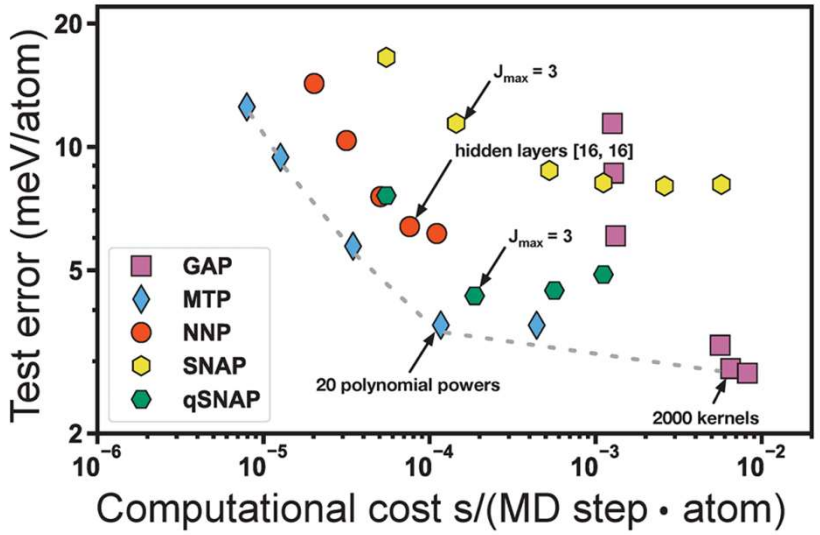
# Computational cost vs accuracy




Gubaev et al. 10.21203/rs.3.rs-2073581/v1

Table S12: The scaling effect on computational cost of different ML-IAPs and classical IAPs. The unit of computational cost is seconds per MD step. Timings were performed by LAMMPS calculations on a single CPU core of Intel i7-6850k 3.6 GHz.

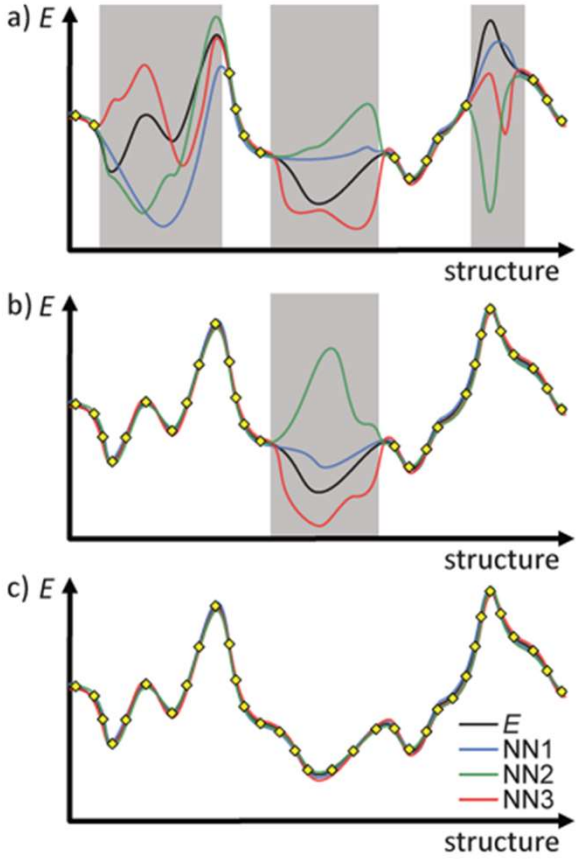
| No. of atoms | 432 | 2000 | 5488 | 11664 |
|---|---|---|---|---|
| GAP | 1.455 | 6.711 | 18.519 | 38.461 |
| MTP | 0.069 | 0.308 | 0.843 | 1.799 |
| NNP | 0.052 | 0.236 | 0.674 | 1.439 |
| SNAP | 0.063 | 0.294 | 0.809 | 1.701 |
| qSNAP | 0.063 | 0.299 | 0.805 | 1.730 |
| EAM[29] | 0.001 | 0.004 | 0.009 | 0.018 |
| MEAM[32] | 0.033 | 0.147 | 0.400 | 0.844 |




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

# Extrapolation and active learning

Ensemble learning:
Training several models




Behler, Chem. Rev. 121, 10037 (2021)

D-optimality criterion

$$\varphi_i^{(p)} = \sum_{v}^{n_v} c_v^{(p)} B_{iv}$$

$$\boldsymbol{B}_i = \left(B_{i1}, B_{i2}, \ldots, B_{in_v}\right)$$

$$\hat{B}_\mu = \begin{bmatrix} B_{11} & \cdots & B_{1n_v} \\ \vdots & \ddots & \vdots \\ B_{N_\mu 1} & \cdots & B_{N_\mu n_v} \end{bmatrix} N_\mu \gg n_v$$

$$\hat{A}_\mu = \begin{bmatrix} B^A_{11} & \cdots & B^A_{1n_v} \\ \vdots & \ddots & \vdots \\ B^A_{n_v 1} & \cdots & B^A_{n_v n_v} \end{bmatrix} \qquad \boldsymbol{B}_i = \sum_{k=1}^{n_v} \gamma_k^{(i)} \boldsymbol{B}_k^A$$

$$\gamma_i = \max_k \left(\left|\gamma_k^{(i)}\right|\right) = \max \left|\boldsymbol{B}_i \cdot \hat{A}_\mu^{-1}\right|$$

Podryabinkin and Shapeev, Comput. Mater. Sci. 140, 171 (2017)
Lysogorskiy, Bochkarev, Mrovec and Drautz, Phys. Rev. Mat. 7, 043801 (2023)

# Extrapolation and active learning

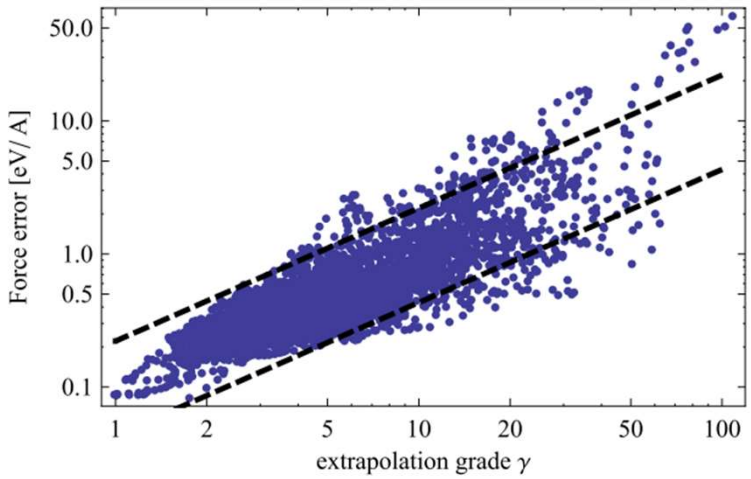



**Fig. 3.** Correlation between the extrapolation grade $\gamma(x)$ and the force error $\Delta f(x) = \left(\frac{1}{N}\sum_{i=1}^{N}|f_i(x) - f_i^{\mathrm{qm}}(x)|^2\right)^{1/2}$. Each point on the graph corresponds to an MD time step. For 95% of configurations the RMS force error is within $[0.04\gamma, 0.22\gamma]$ eV/Å (dashed lines)—this indicates a good correlation between the error and the extrapolation grade.

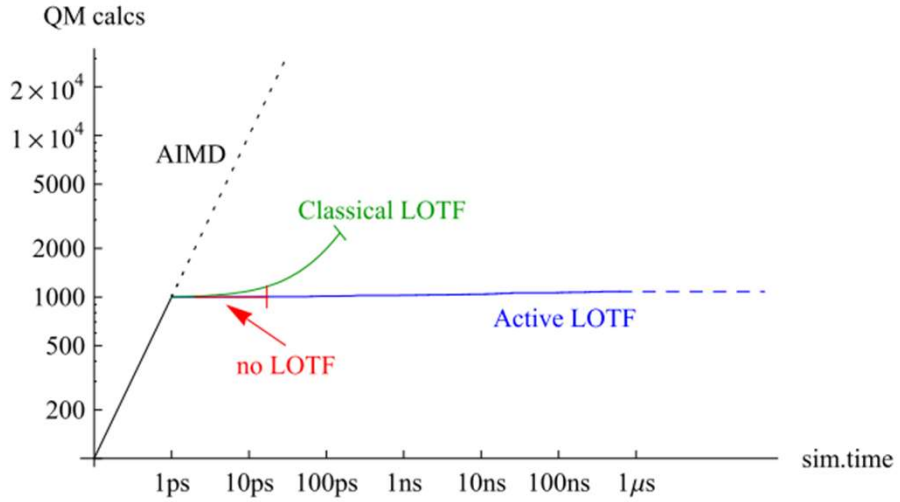
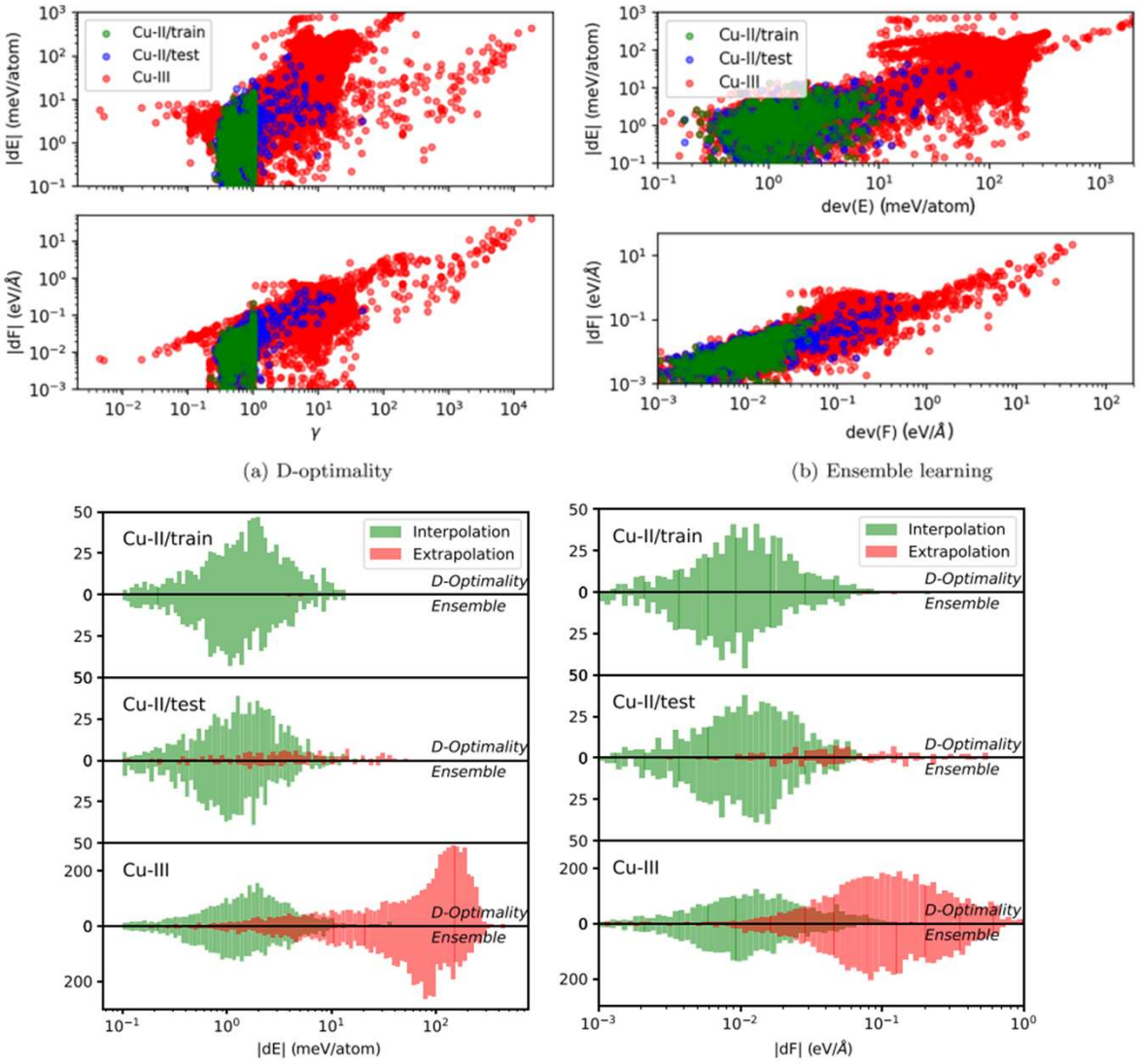



**Fig. 5.** Comparison of *ab initio* molecular dynamics (AIMD) with no-learning MD, classical learning on the fly (LOTF) inspired by [29], and active LOTF. The no-learning and classical LOTF MD are not completely reliable: on average every 15 ps the no-learning MD fails, i.e., escapes into an unphysical region in the phase space. The classical LOTF makes this ten times more reliable (failure time of 150 ps) at the expense of extra 1500 QM calculations. In contrast, the active LOTF makes MD completely reliable (i.e., failures are not observed) at the cost of only 50 QM calculations as measured over the first 0.5 μs of simulation time.

Podryabinkin and Shapeev, Comput. Mater. Sci. 140, 171 (2017)

# Extrapolation and active learning




Lysogorskiy, Bochkarev, Mrovec and Drautz, Phys. Rev. Mat. 7, 043801 (2023)

# Extrapolation and active learning

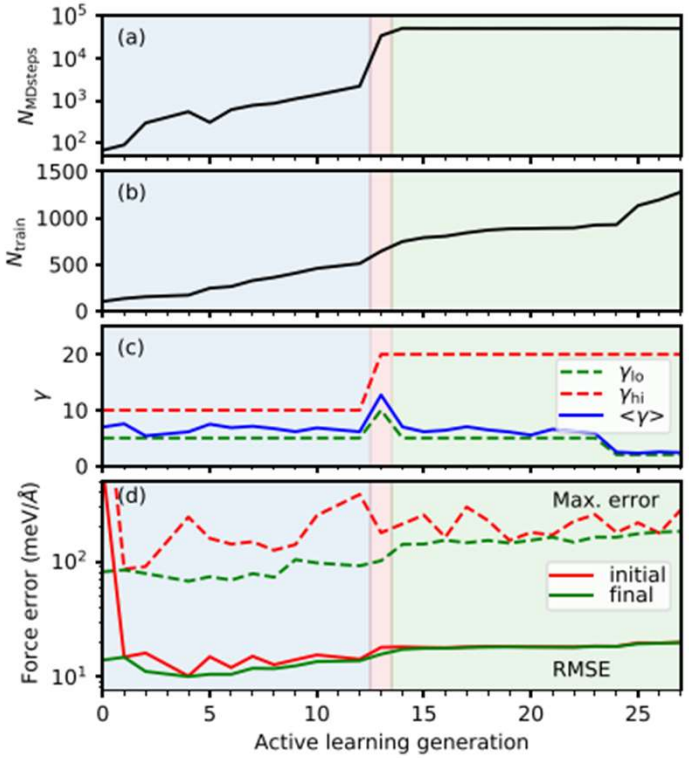



FIG. 7. Analysis of AL procedure for water: (a) number of successful MD steps before extrapolation grade exceeds $\gamma_{hi}$, (b) training set size, (c) extrapolation grade range $[\gamma_{lo}, \gamma_{hi}]$ and average grade $\langle\gamma\rangle$, and (d) initial and final (before and after ACE model retraining) force error metrics (RMSE and max. error). Three highlighted regions on each plot mark MD-exploration regimes; see the text for more details.

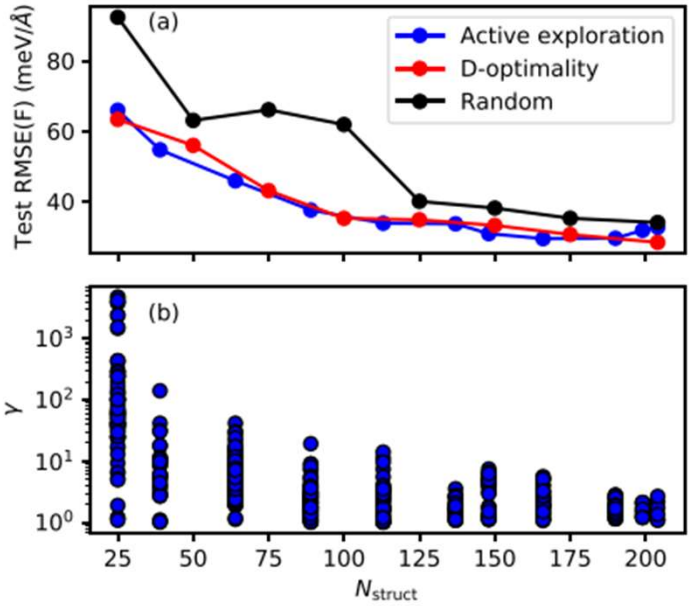



FIG. 9. Comparison of active exploration, D-optimality, and random selection strategies for training of $Li_4$ clusters (a), and variation of the extrapolation grade during active exploration (b); see the text for more details.

Lysogorskiy, Bochkarev, Mrovec and Drautz, Phys. Rev. Mat. 7, 043801 (2023)

# Extrapolation and active learning

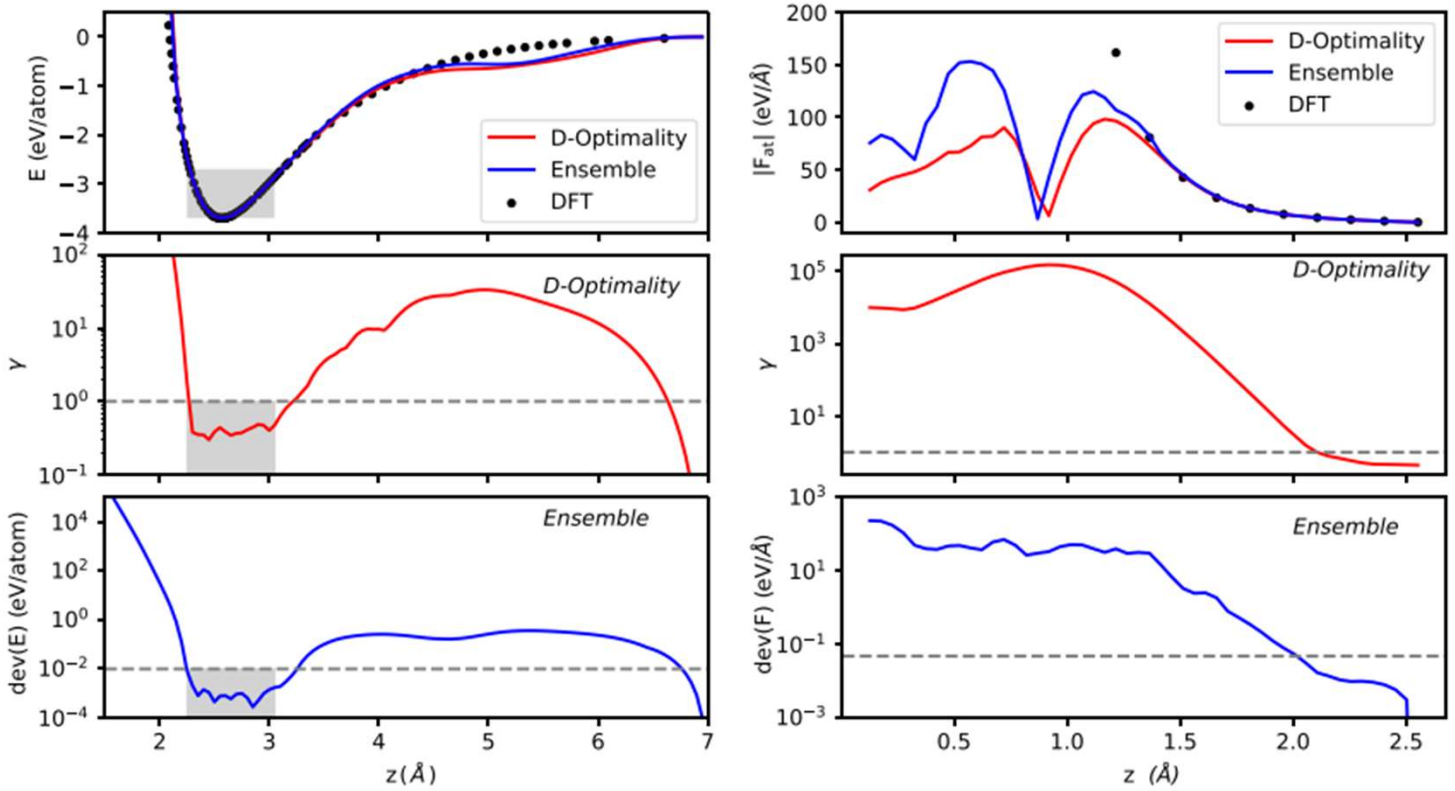



FIG. 13. Extrapolation indicators for homogeneous deformation of fcc Cu (left) and displacement of a single atom in a fcc crystal (right). See the text for details.

Lysogorskiy, Bochkarev, Mrovec and Drautz, Phys. Rev. Mat. 7, 043801 (2023)

# Active learning of local environments

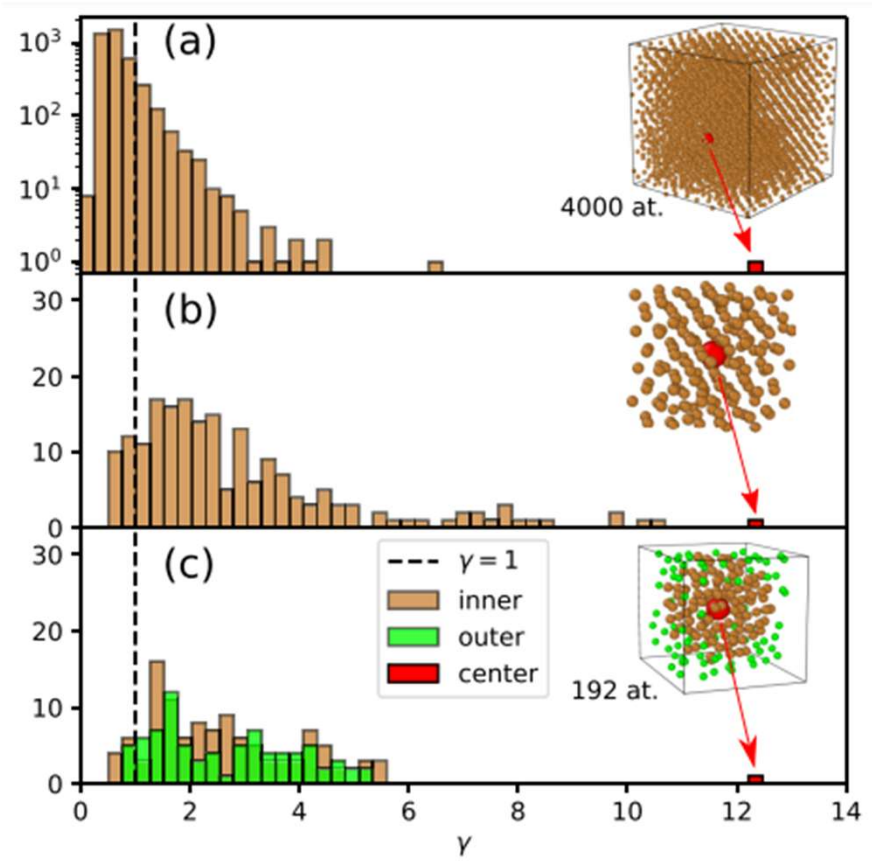



FIG. 10. Distribution of extrapolation grade in a Cu-fcc system: (a) original large cell containing 4000 atoms, (b) an isolated cut-out cluster around the red atom with large $\gamma$; (c) the same cell as in (b) but with periodic conditions and optimized atomic positions of outer-shell atoms (green atoms).

Lysogorskiy, Bochkarev, Mrovec and Drautz, Phys. Rev. Mat. 7, 043801 (2023)

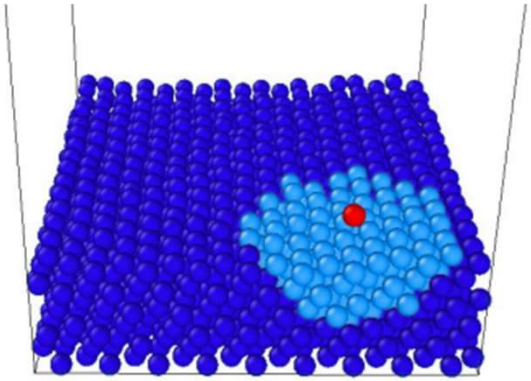


FIG. 5: Large atomic system including candidate extrapolative neighborhoods. The atom with the maximum extrapolation grade and its neighbours are highlighted. These atoms form a spherical neighborhood on which we train the potentials.

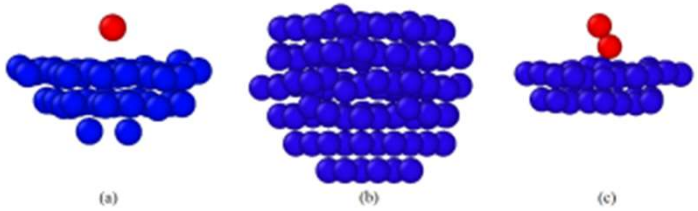
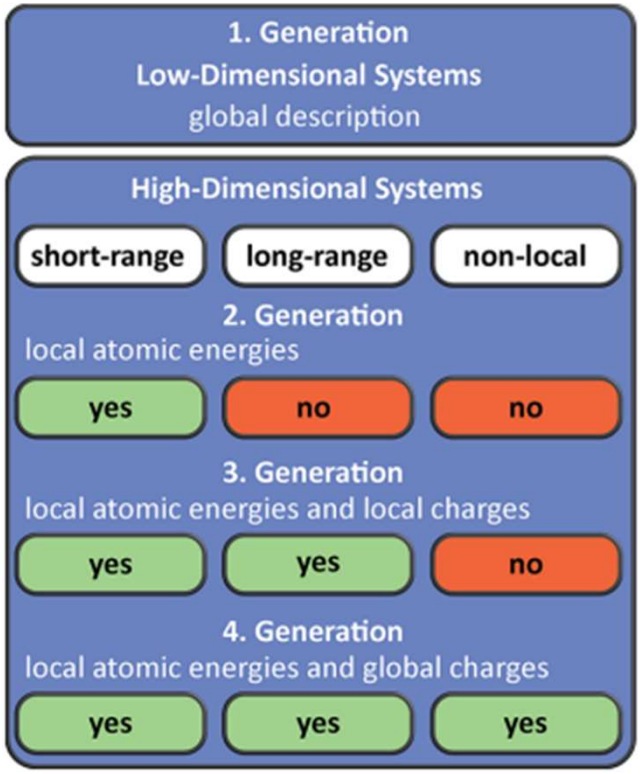
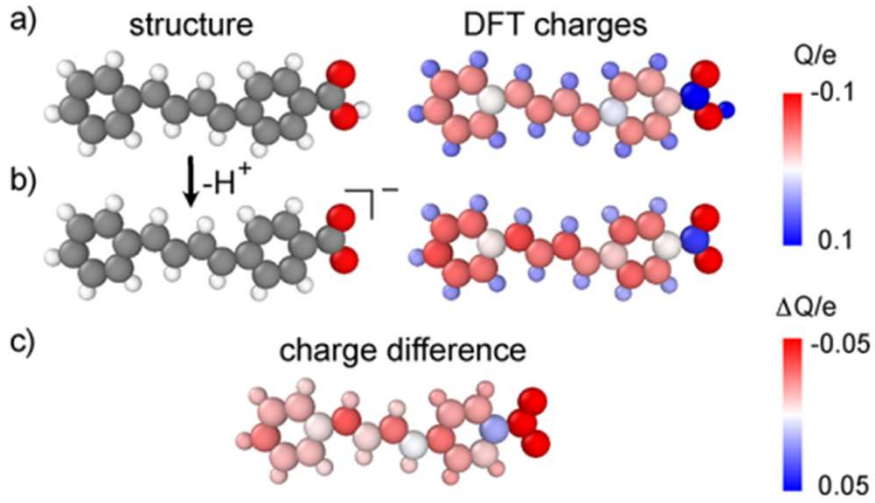


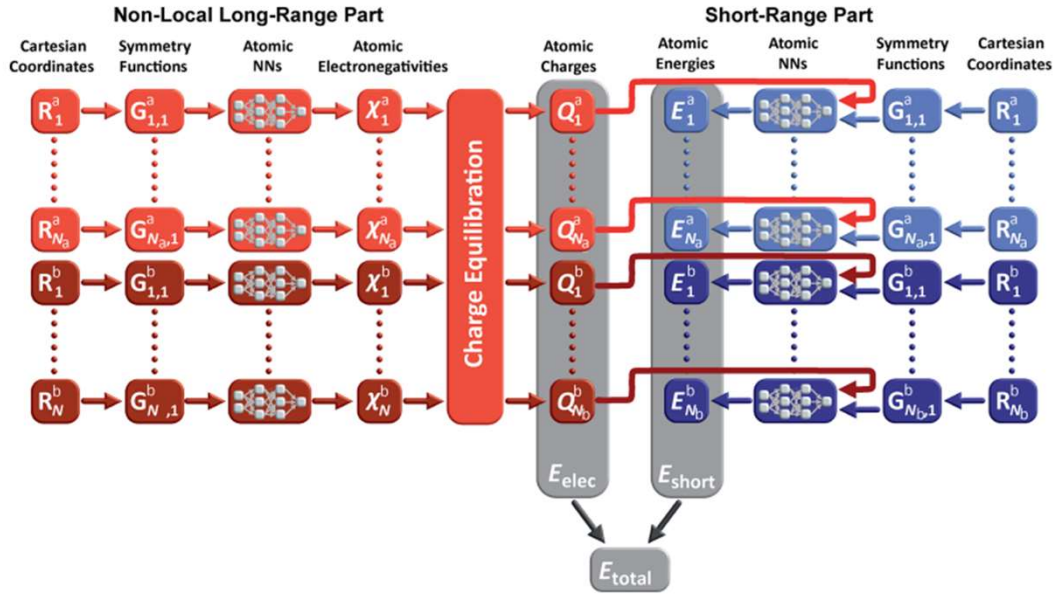

FIG. 6: Examples of atomic neighborhoods selected during active learning. (a) an atom near the substrate; (b) atoms inside the substrate; (c) several atoms on the substrate. The atoms on the substrate and inside it are of different colors, but all of them are copper.

Podryabinkin, Garifullin, Shapeev and Novikov, arXiv:2304.13144 (2023)

# Beyond locality



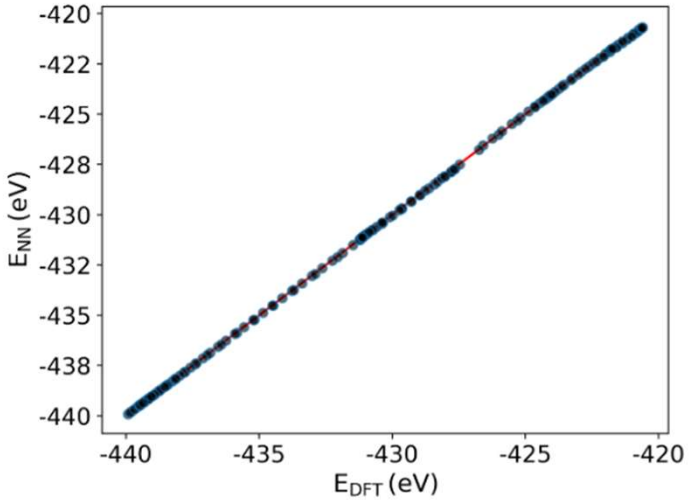



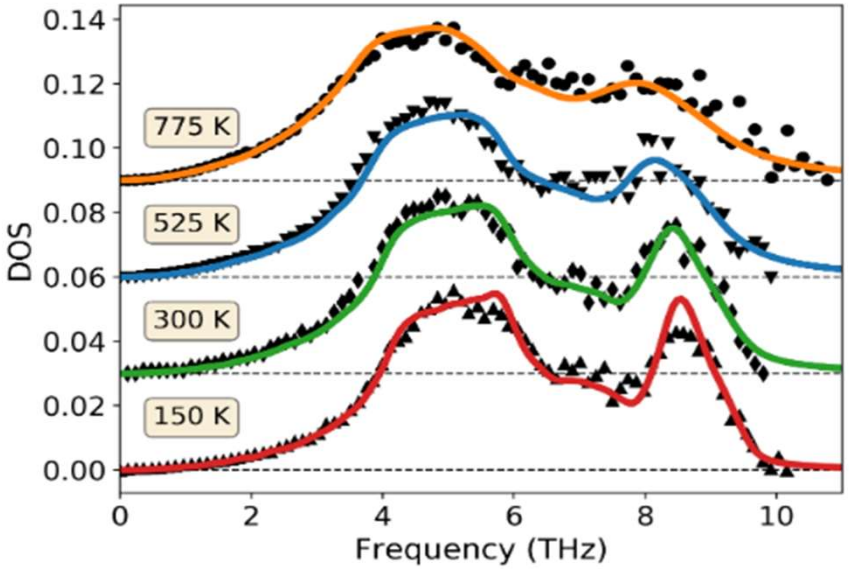

Behler, Chem. Rev. Lett. 121, 10037 (2021)

# Example: Neural network potential

$$D_i^l = \sum_j c^l(Z_i)c^l(Z_j)\exp[-\sigma^l(r_{ij}-\eta^l)^2]f_{\text{cut}}(r_{ij},R_{\text{cut}})$$

$$\mu||E'-E||^2 + \sum_i^{\#\text{atoms}} ||\mathbf{F}_i'-\mathbf{F}_i||^2 \rightarrow \min$$

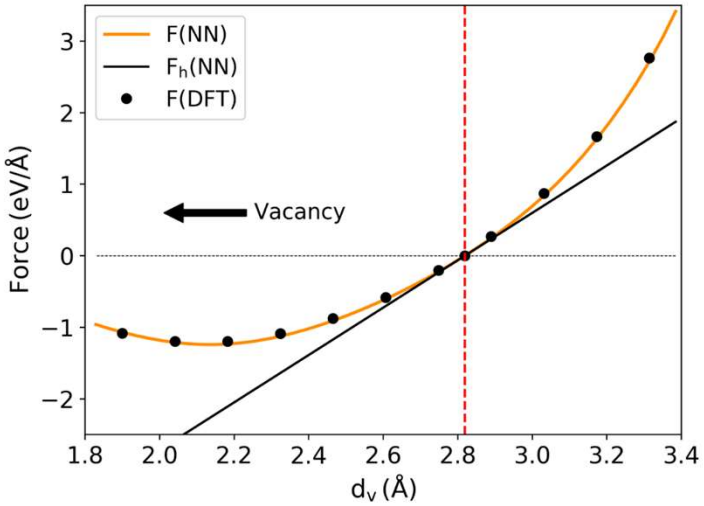


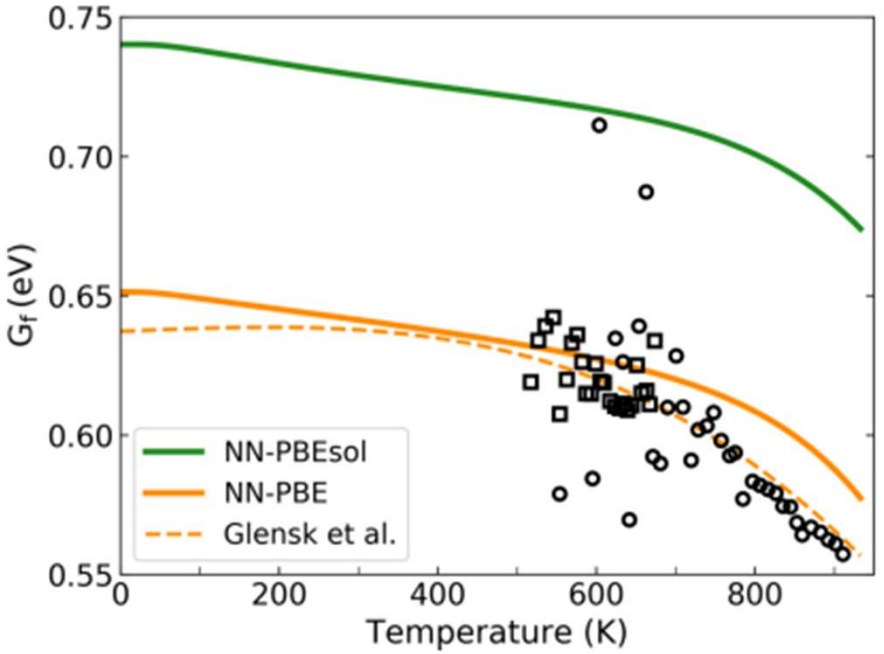



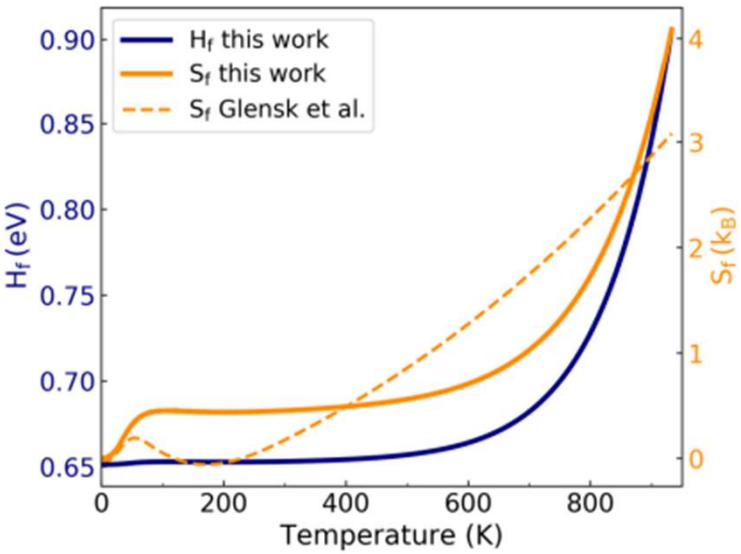

| | | $E$, meV/atom | | $F$, eV/Å | |
|---|---|---|---|---|---|
| | Data set | RMSE | MAE | RMSE | MAE |
| PBEsol | Test | 0.40 | 0.27 | 0.021 | 0.014 |
| | Validation | 0.35 | 0.25 | 0.025 | 0.017 |
| PBE | Test | 0.53 | 0.32 | 0.028 | 0.018 |
| | Validation | 1.4 | 1.3 | 0.035 | 0.023 |

Bochkarev, van Roekeghem, Mossa and Mingo
Physical Review Materials 3, 093803 (2019)

# Example: Neural network potential

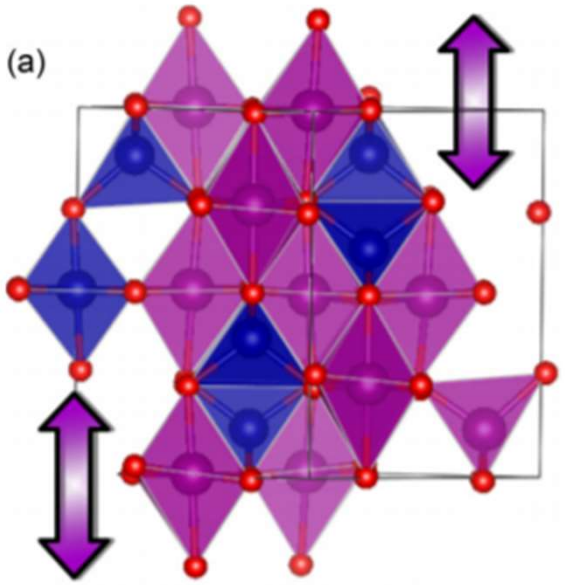


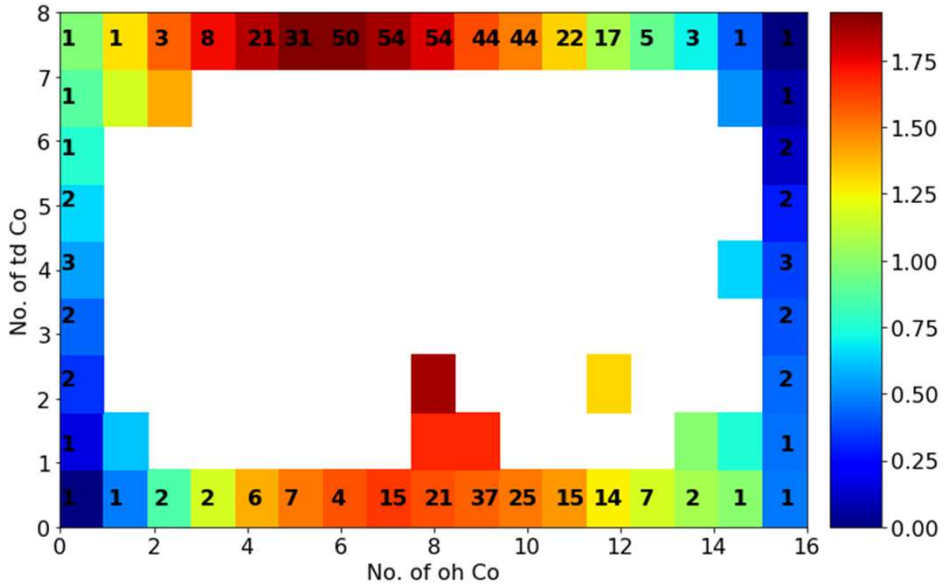



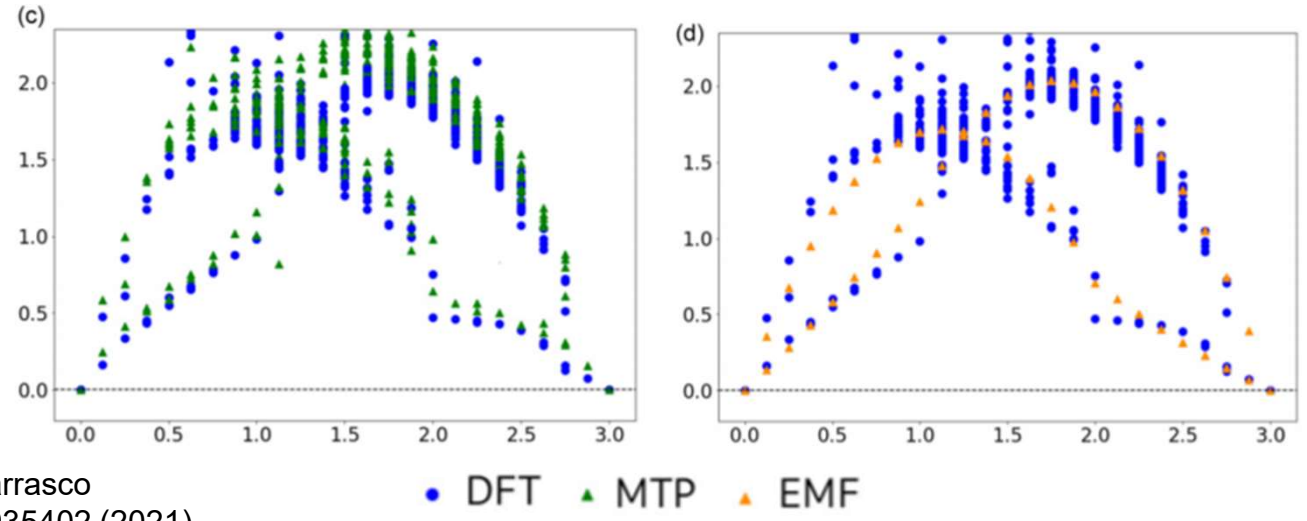

Bochkarev, van Roekeghem, Mossa and Mingo
Physical Review Materials 3, 093803 (2019)

# (CoxMn1-x)3O4 for thermochemical storage



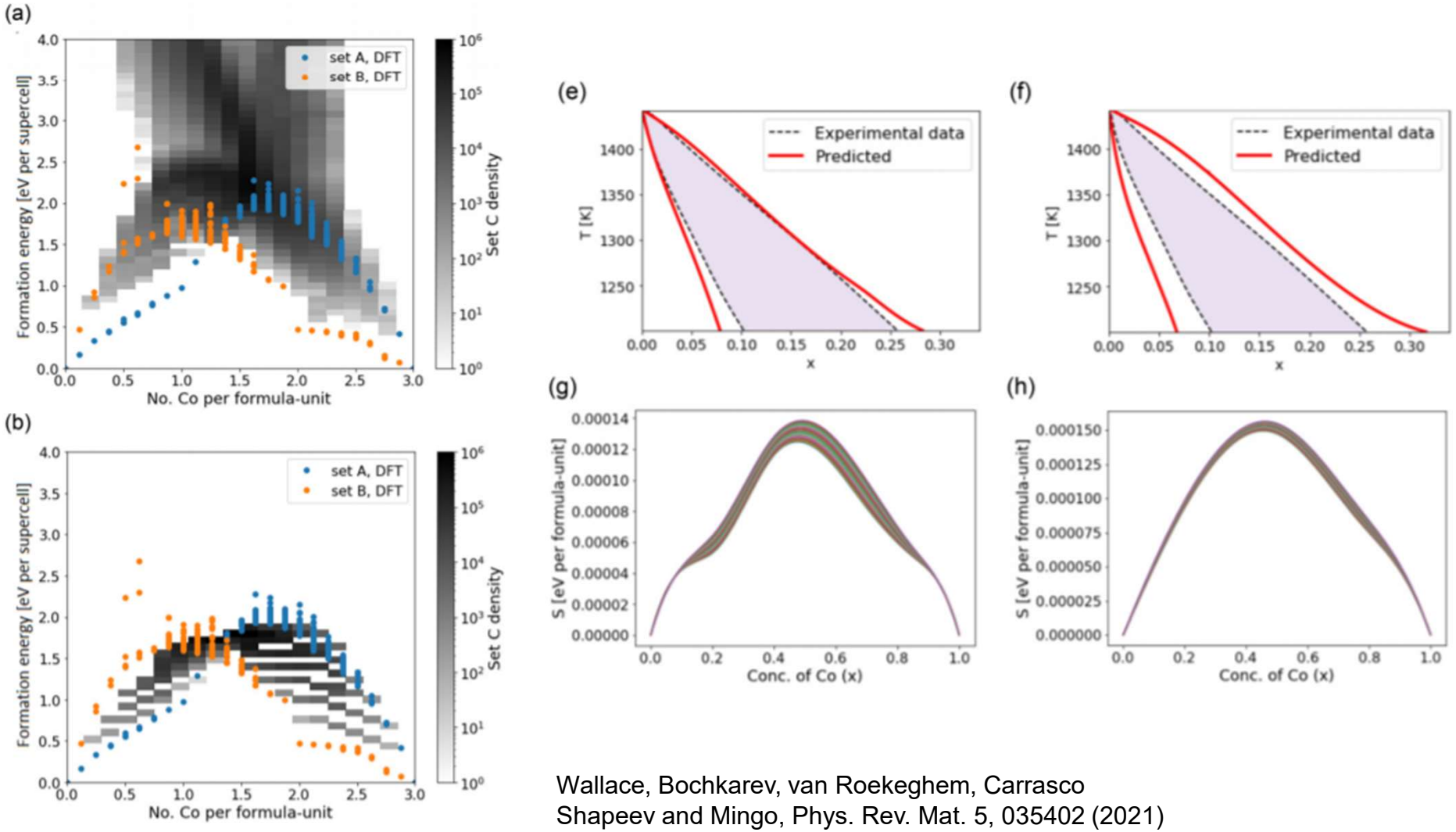




For x=0,5 and 56 atoms about $3*10^6$ possibilities



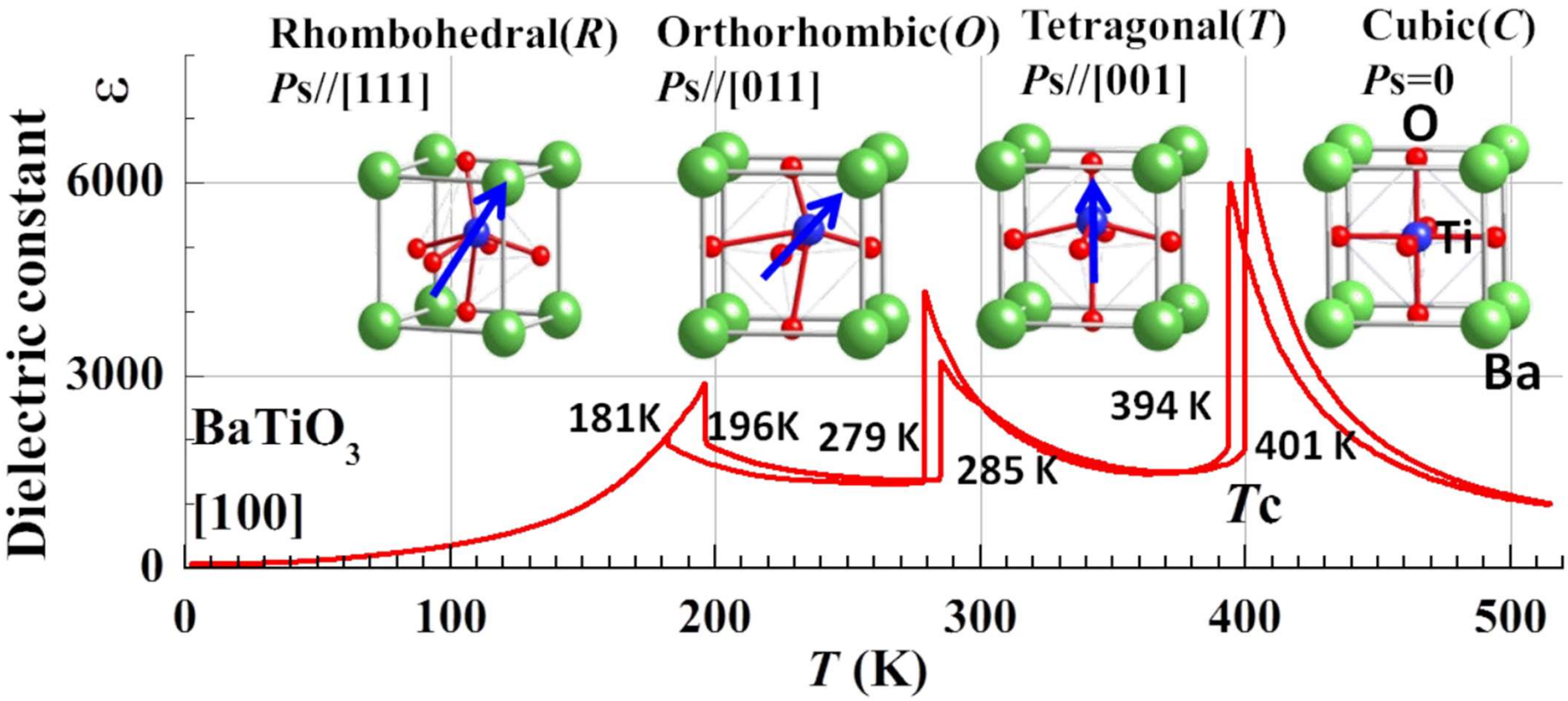

Wallace, Bochkarev, van Roekeghem, Carrasco
Shapeev and Mingo, Phys. Rev. Mat. 5, 035402 (2021)

# Moment Tensor Potentials vs Extended Mean Field



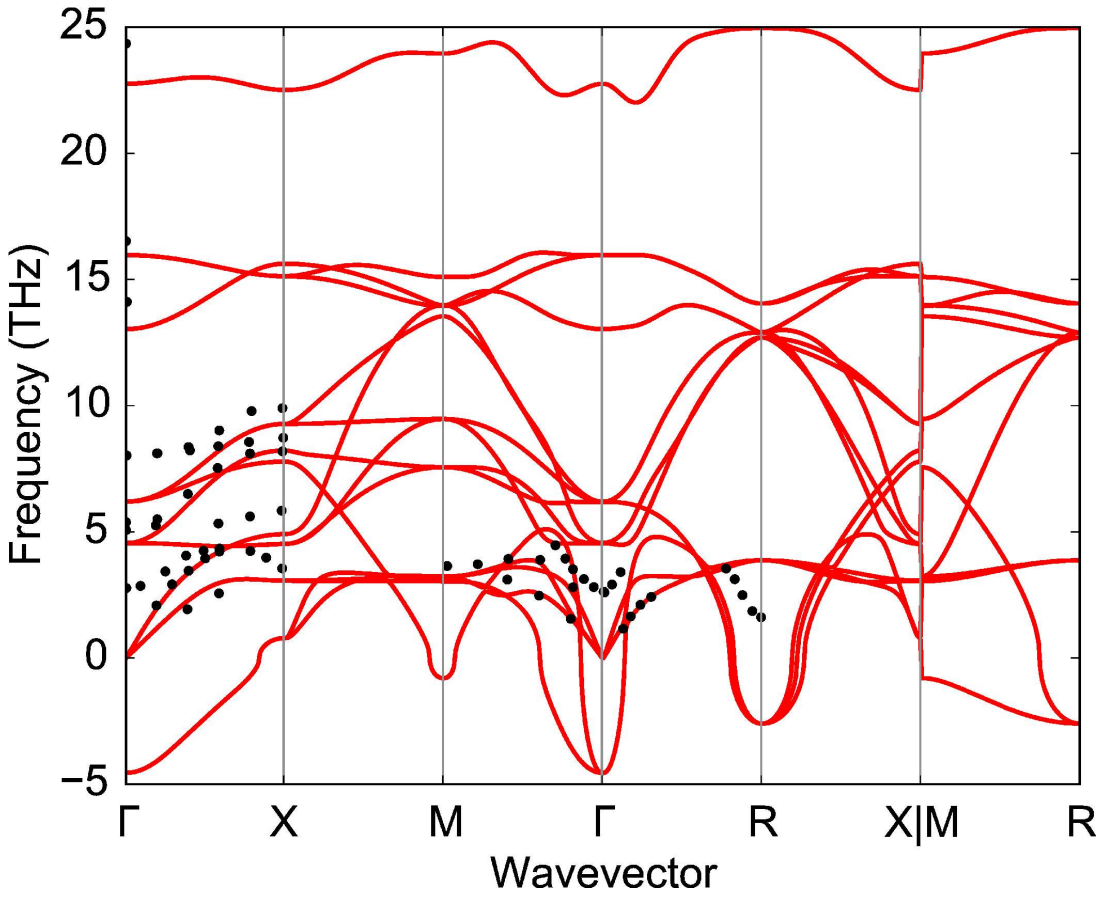

Wallace, Bochkarev, van Roekeghem, Carrasco
Shapeev and Mingo, Phys. Rev. Mat. 5, 035402 (2021)

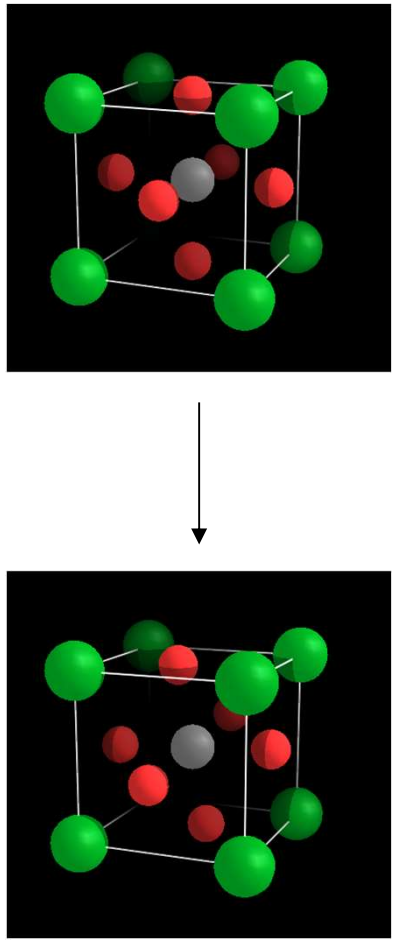
# Distortions in perovskites: the example of BaTiO3




D. Fu and M Itoh, Ferroelectrics Materials – Synthesis and Characterization (2015)

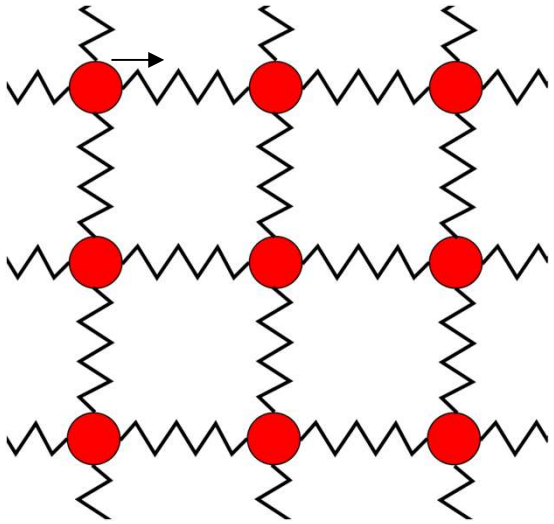
# Distortions in perovskites: the example of SrTiO3



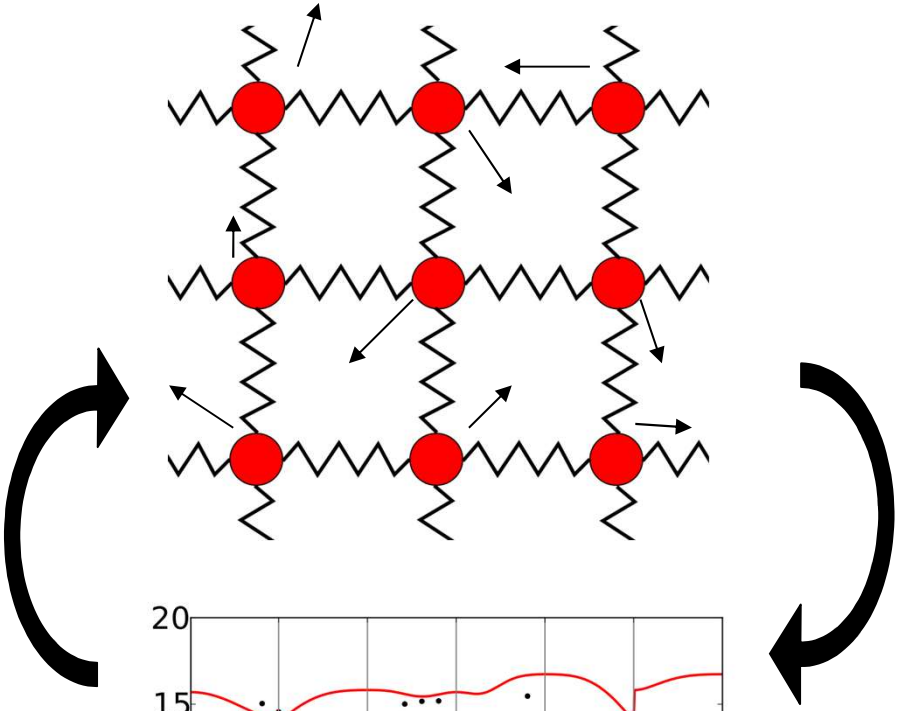



# Finite-temperature phonon calculations: QSCAILD

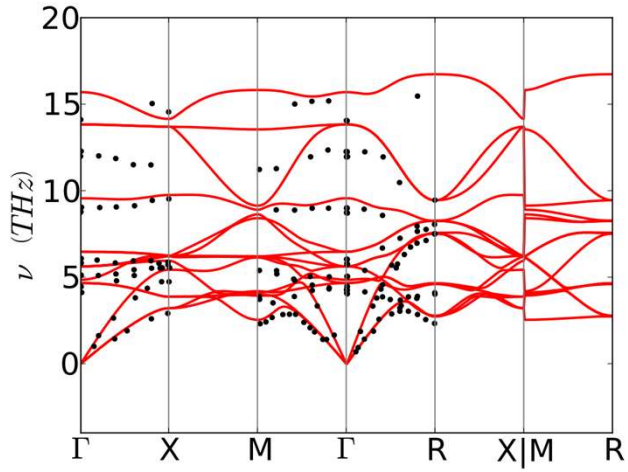
Small displacements



Quantum statistics, finite T



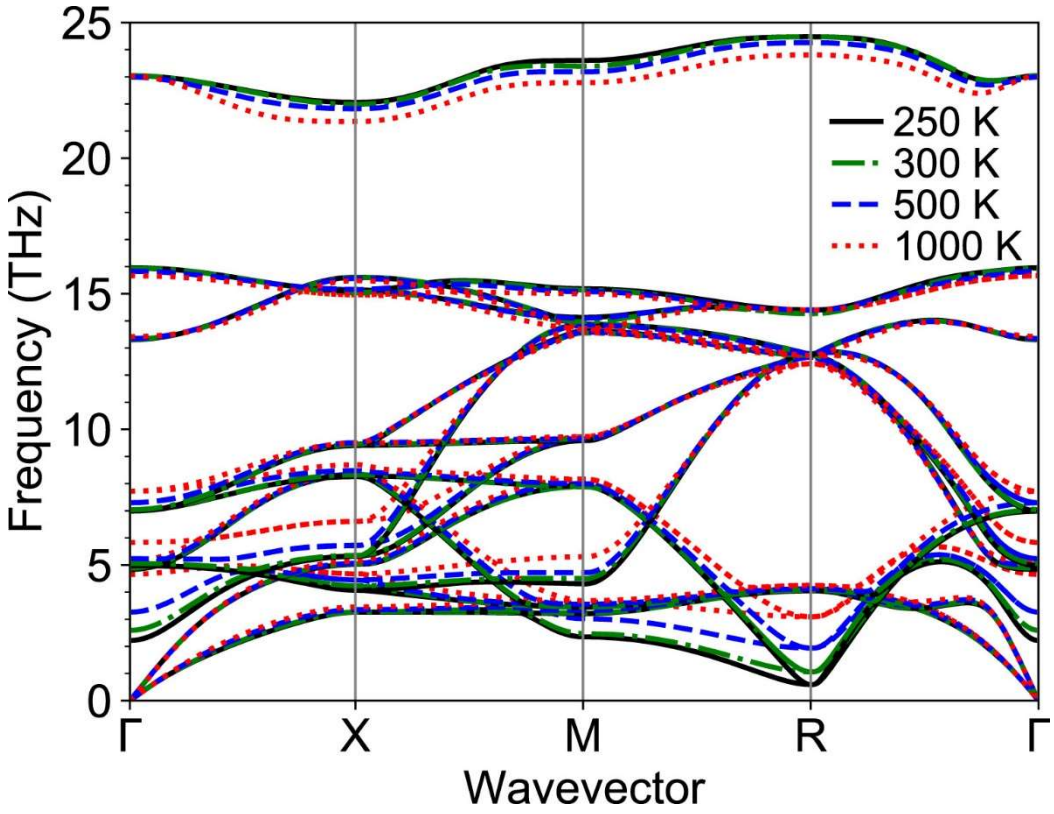
$$F_i^\alpha = \sum_{j\beta} \Phi_{ij}^{\alpha\beta} u_j^\beta + \frac{1}{2} \sum_{jk\beta\gamma} \Psi_{ijk}^{\alpha\beta\gamma} u_j^\beta u_k^\gamma$$

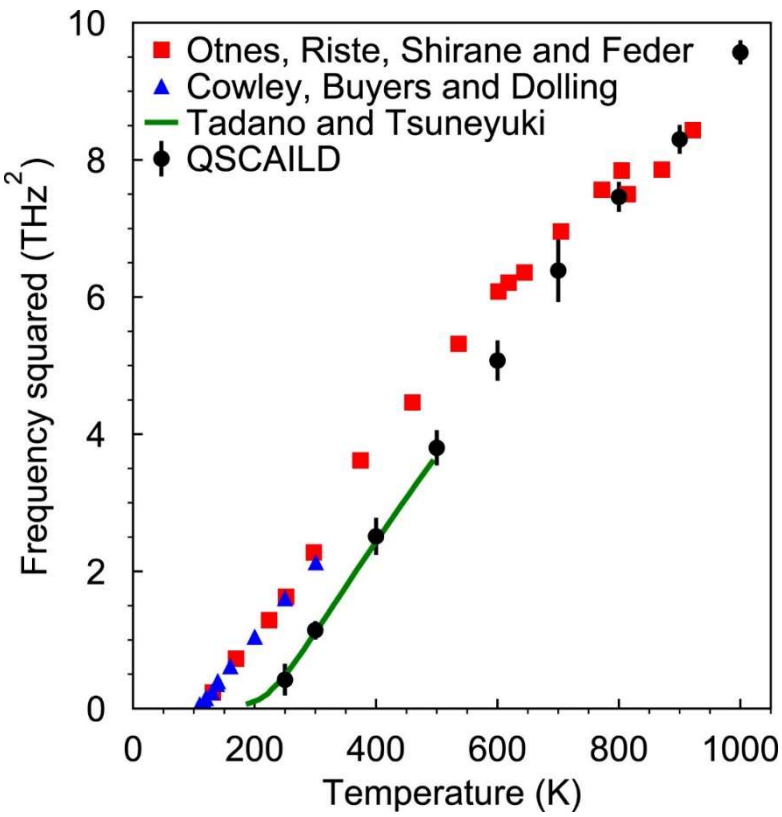



→ 2nd and 3rd order force constants

→ Thermal conductivity

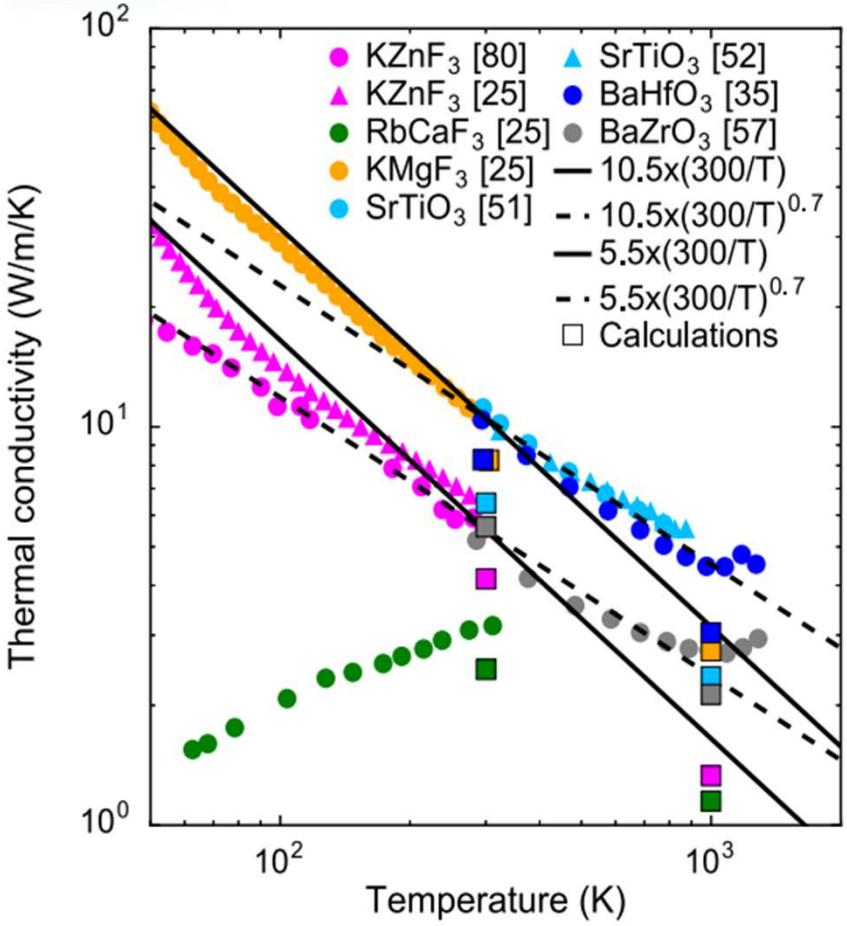
# SrTiO3 at finite temperature



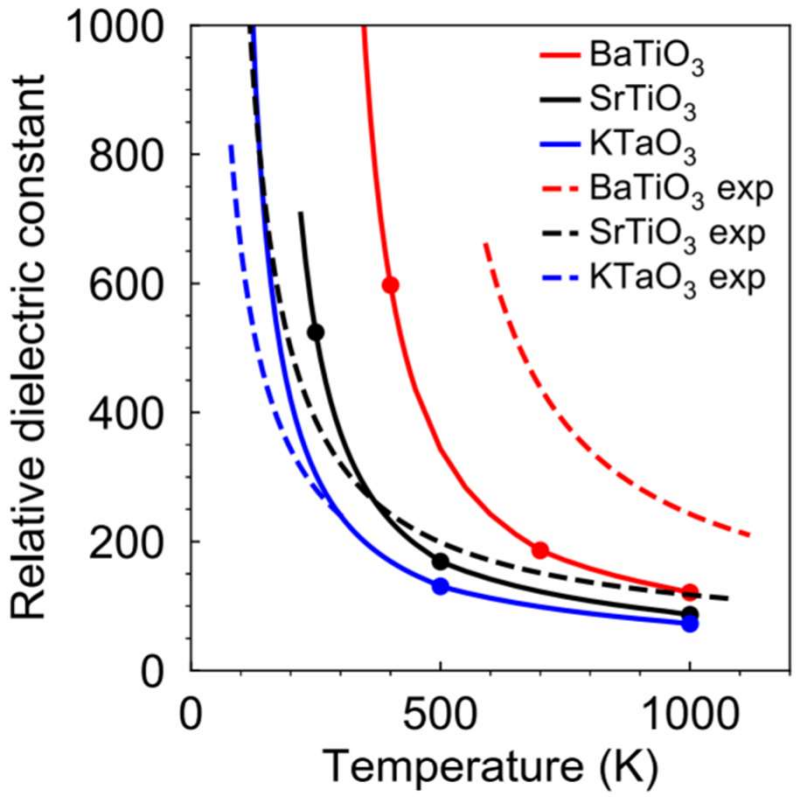

van Roekeghem, Carrete and Mingo, Comp. Phys. Comm. 263, 107945 (2021)

# Thermal conductivities and dielectric constants




$$\kappa \propto T^{-\alpha} \qquad \text{Mean } \alpha = 0.8$$

van Roekeghem, Carrete, Oses, Curtarolo and Mingo
Phys. Rev. X 6, 041061 (2016)




$$\frac{\epsilon_{DC}}{\epsilon_{\infty}} = \prod_j \left( \frac{|\omega_{Lj}|}{\omega_{Tj}} \right)^2$$

van Roekeghem, Carrete, Curtarolo and Mingo
Phys. Rev. Materials 4, 113804 (2020)

# Quantum paraelectricity in KTaO3

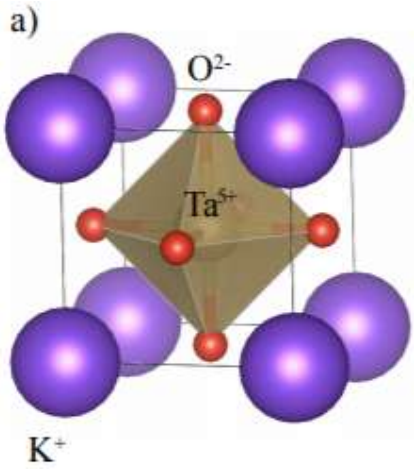
$KTaO_3$ crystallizes in the perovskite structure

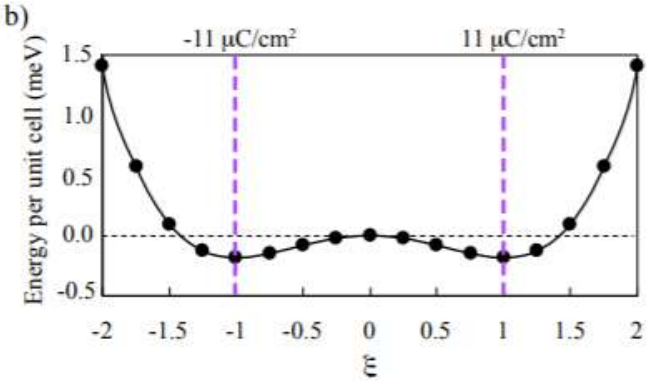
Quantum paraelectric:

existence of a ferroelectric instability

No ferroelectric phase, because quantum fluctuations stabilize unstable phonon branch even at 0 K

Very soft polar branch -> large bandgap but huge dielectric constant



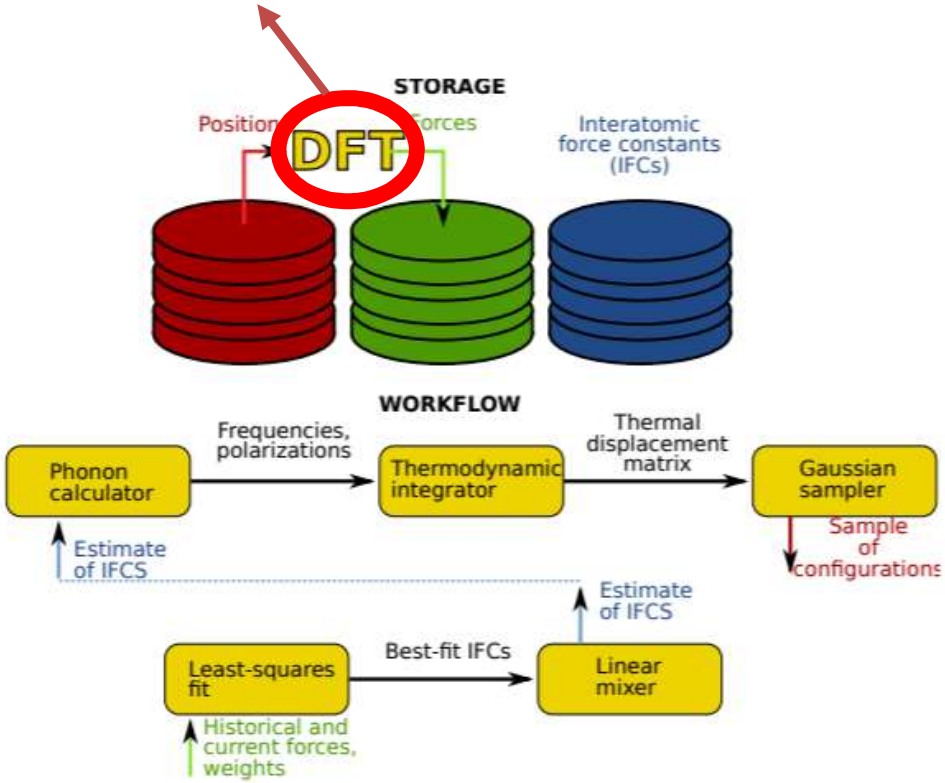




Gattinoni et al, arXiv:2108.10207

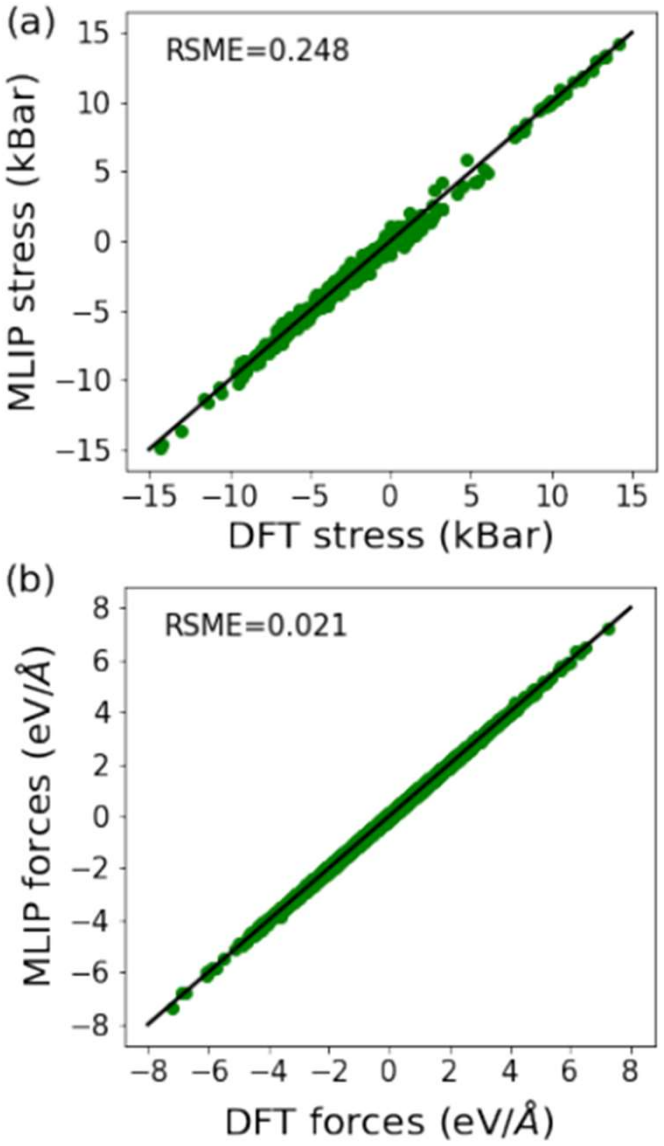
# Speeding up QSCAILD using ML potentials

- Acquiring forces requires many DFT calculations -> slow and expensive
- We have to sample many similar configurations -> ideal to combine with machine learning
- We interface QSCAILD with active learning of Moment Tensor Potentials

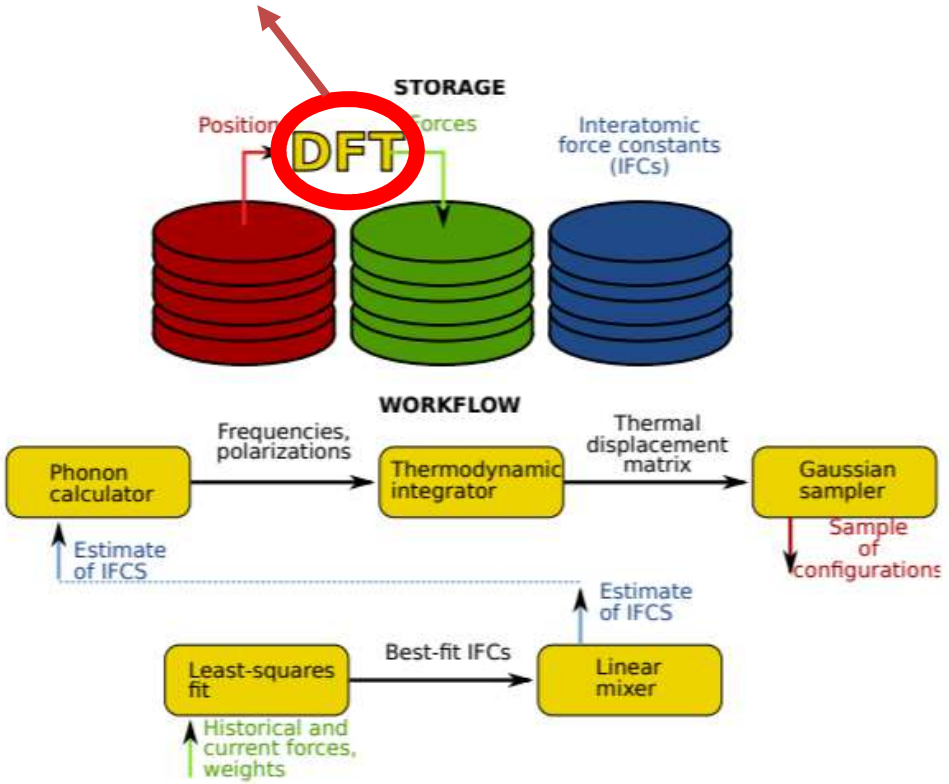
Bottleneck! -> DFT + active learning



# Speeding up QSCAILD using ML potentials

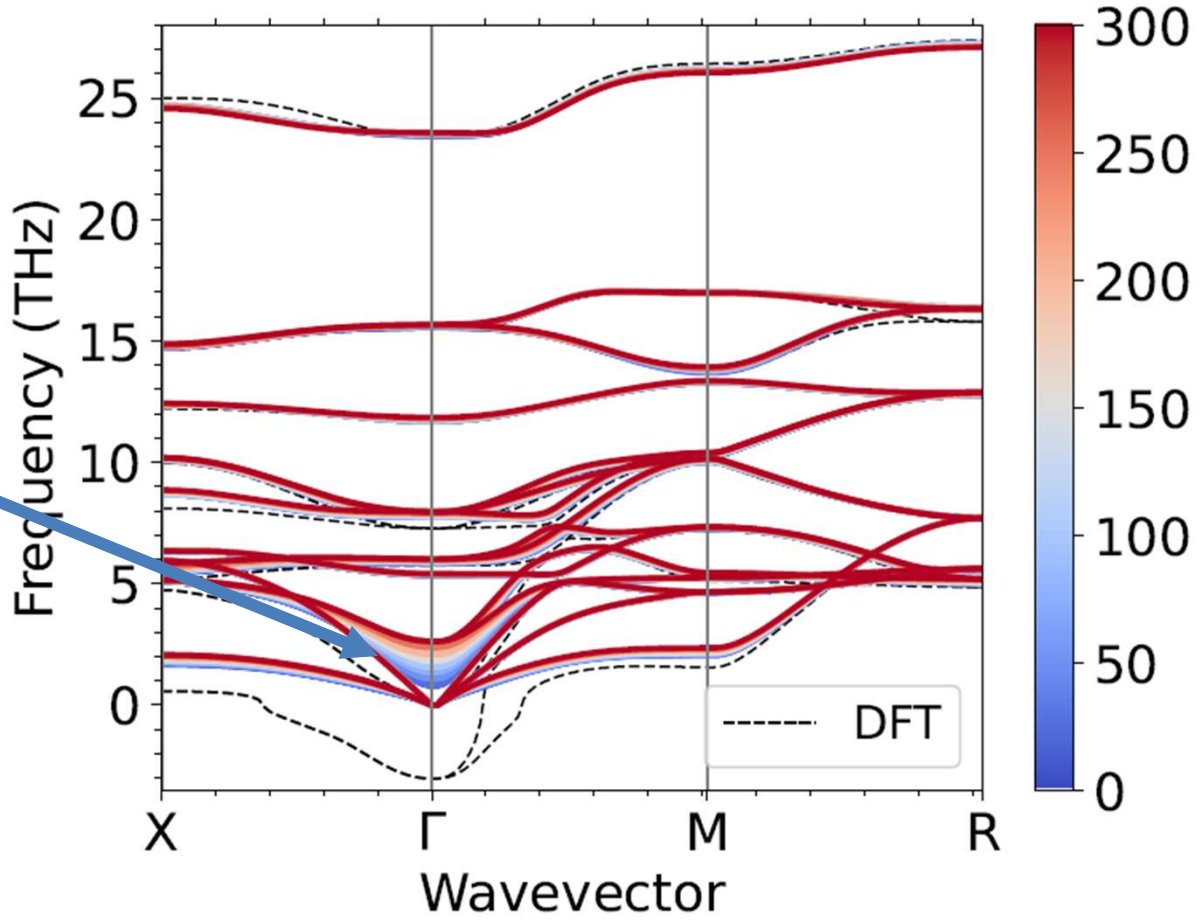



Bottleneck! -> DFT + active learning



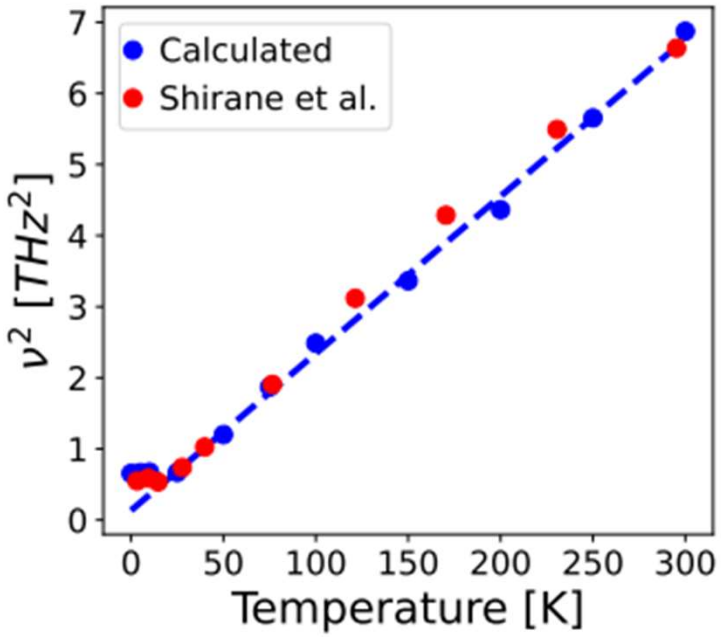
# Results: Temperature dependent phonon spectrum of KTaO3

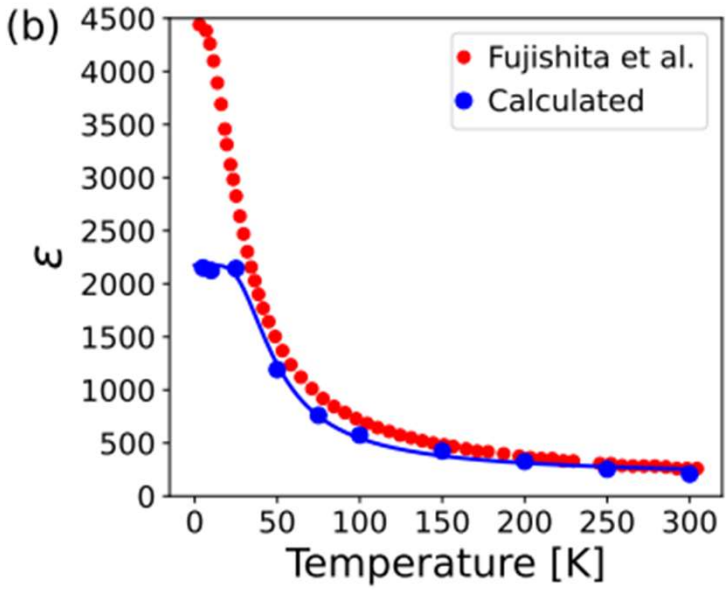
Temperature dependence mostly concentrated on the ferroelectric soft mode




Meier, Mingo and van Roekeghem, arXiv:2206.08296 (2022)

# Phonon frequencies of the soft mode and dielectric constant




Excellent agreement of temperature dependent phonon frequencies with experiment

Small difference in temperature of saturation

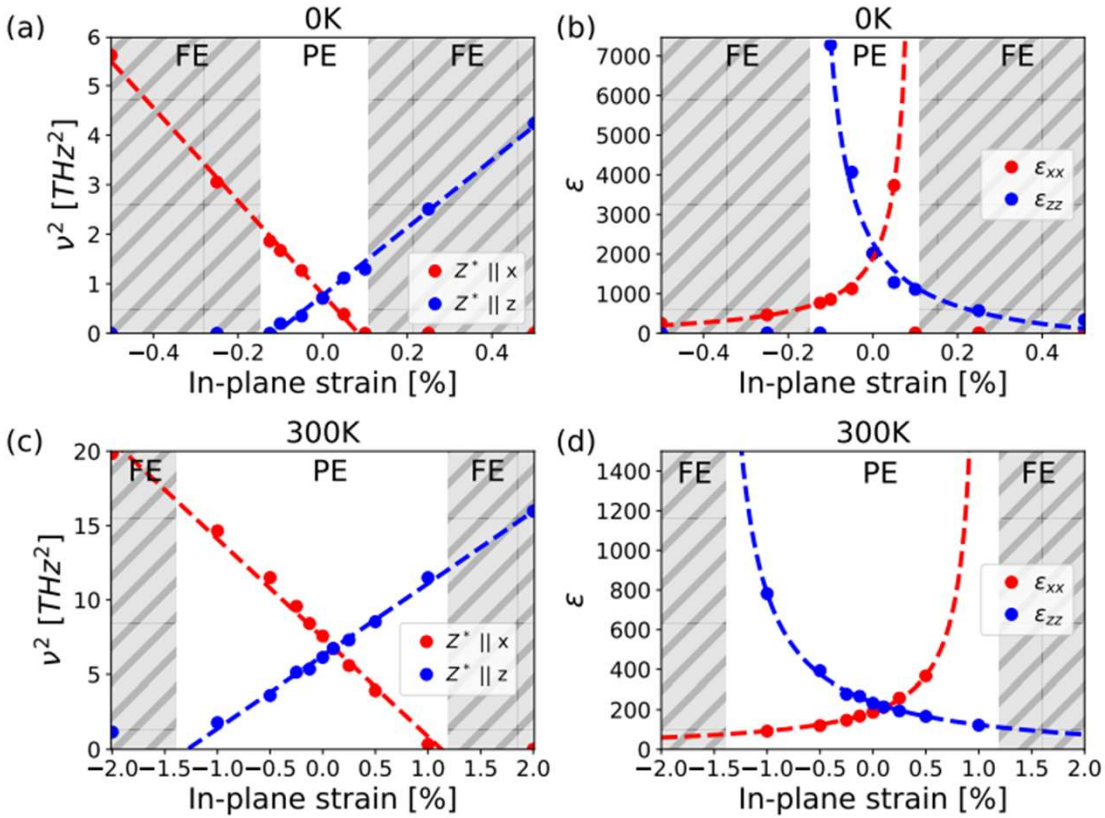
G. Shirane, R. Nathans, and V. J. Minkiewicz
Phys. Rev. **157**, 396 (1967)

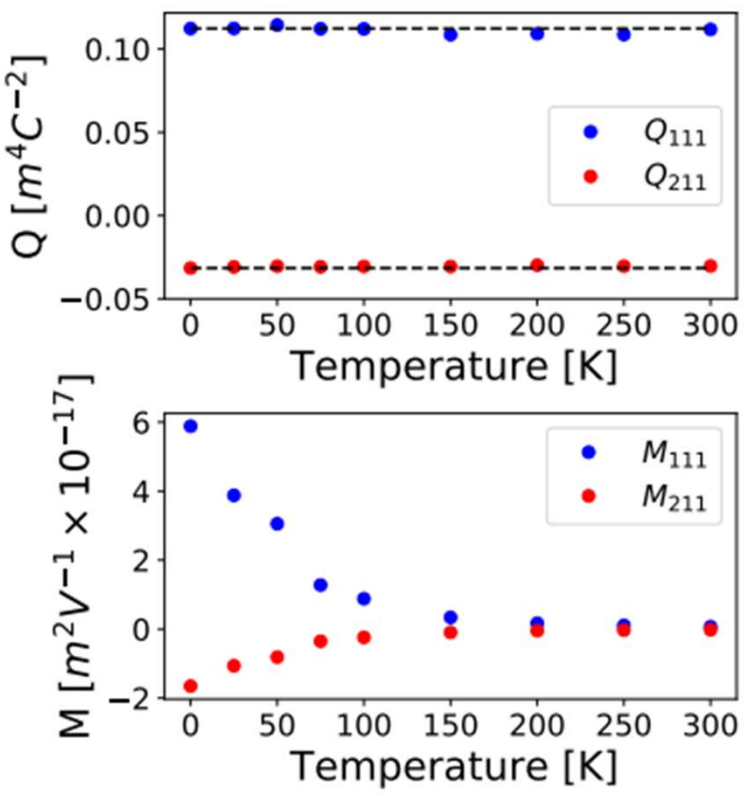
Temperature dependent dielectric constant compared to experiment, excellent agreement up to 25K and observation of Barrett law

J. Barrett, Phys. Rev. 86, 118 (1952)

Meier, Mingo and van Roekeghem, arXiv:2206.08296 (2022)

# Temperature-dependent electrostrictive properties




$$x_i = Q_{ikl} P_k P_l$$
$$x_i = M_{ikl} E_i E_k$$
$$M_{ikk} = Q_{ikk} \chi^2_{kk}$$

Meier, Mingo and van Roekeghem, arXiv:2206.08296 (2022)

# Conclusions

ML interatomic potentials with DFT accuracy open new possibilities.

They are progressively predating previous interatomic potentials.

Selecting and producing data for training is one of the main challenges.

Performance is key, as is user friendliness.

Reliability and fidelity have to be questioned (what is the model?).

Thank you!

And thanks to:

Quintin MEIER
Anton BOCHKAREV
Suzanne WALLACE
Jesús CARRETE
Javier CARRASCO
Stefano CURTAROLO
Natalio MINGO
Alexander SHAPEEV
Stefano MOSSA
Joël M.D. CEA

Commissariat à l'énergie atomique et aux énergies alternatives
17 rue des Martyrs | 38054 Grenoble Cedex
www-liten.cea.fr