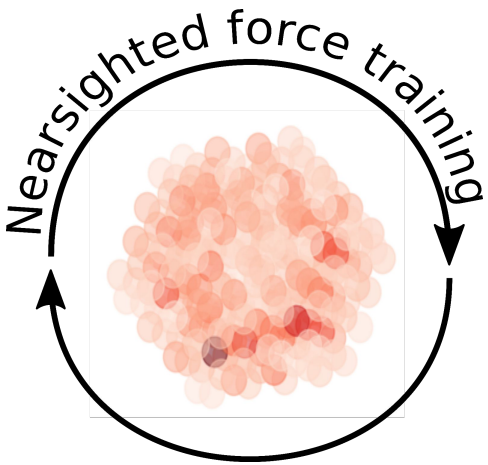
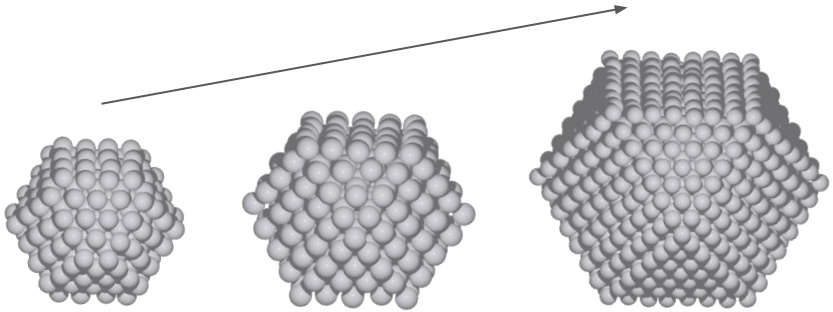
A Nearsighted Force-Training (NFT) Approach for Machine Learning of Large Atomic Structures

Cheng Zeng, Andrew Peterson

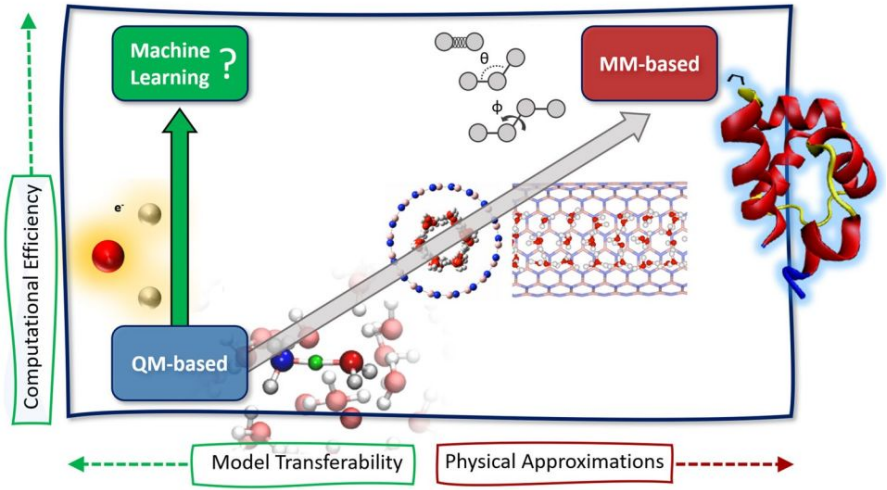




Catalyst Design Lab

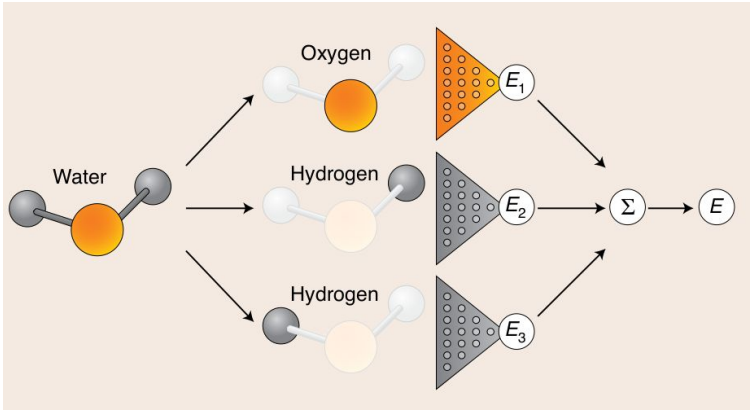
Machine Learning (ML) Interatomic Potentials

ML models as potential solutions to overcome limitations of QM-based methods



T. Morawietz *et. al., J. Comput. Aided Mol. Des.,* **35** (2021)

Behler-Parrinello (BP) neural network potentials



P. Friederich *et. al., Nat. Mater.,* **20** (2021)

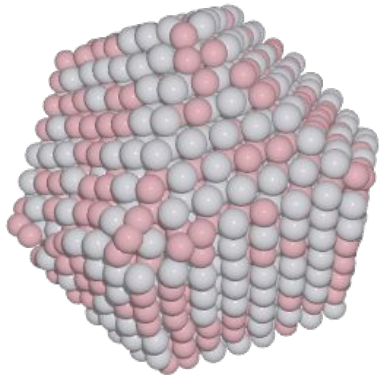
J. Behler *et.al., Phys. Rev. Lett.,* **98**, 146401 (2007)

Challenges in ML potentials

Normally a large amount of data are needed

- Active learning to minimize the number of ab initio calculations

How to generate ***small reference data*** when ML predictions fail on ***large structures***?

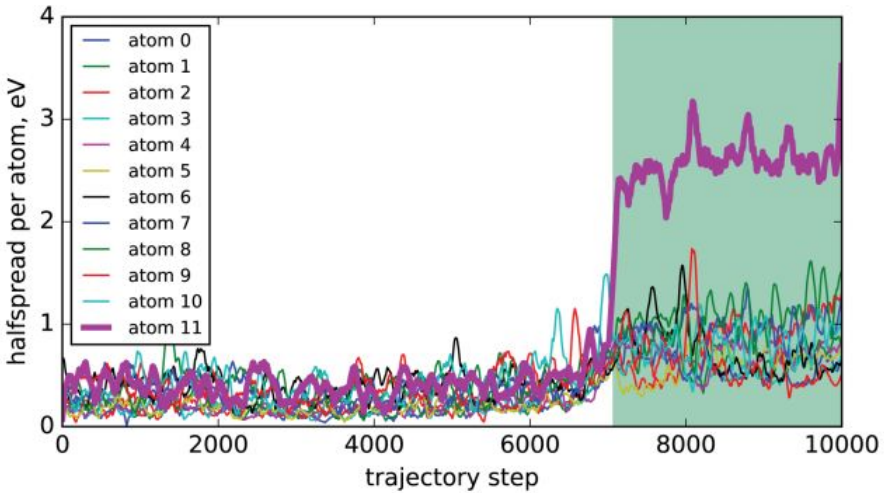


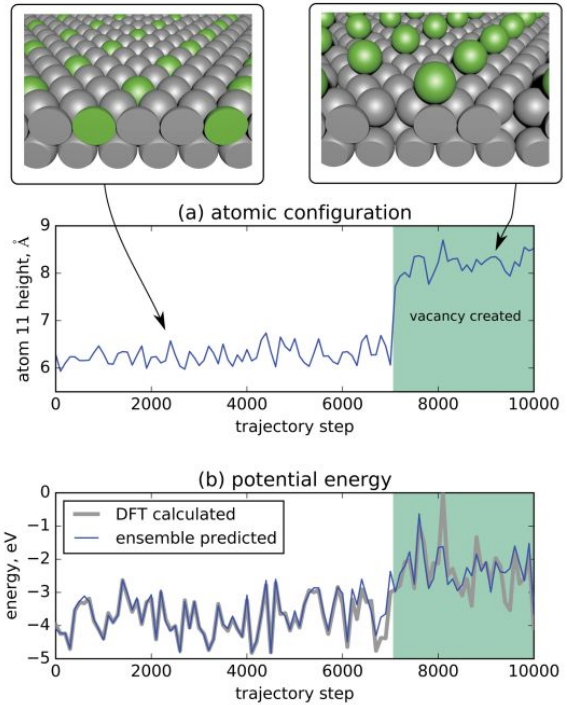
Crucial to explore and exploit ***relevant*** configuration space

Localizing Uncertainty to Atoms

Ensemble models can isolate prediction errors to atoms

Largest error from the atom moving onto the surface





A. A. Peterson *et. al., Phys. Chem. Chem. Phys.,* **19** (2017)

Nearsightedness of Finite-ranged ML Potentials

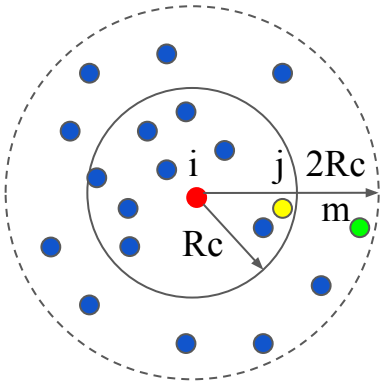
Atomic energy is represented by its local chemical environment

Atomic energy locality, Rc

$$E_i = E\left(\{\vec{R}_{ij}\}\right), \text{ where } |R_{ij}| < R_c$$

Force locality, 2Rc

$$f_i = -\frac{\partial E}{\partial R_i} = -\sum_j^N \frac{\partial E_j}{\partial R_i} = -\sum_j^{R_{ij}<R_c} \frac{\partial E_j}{\partial R_i} = -\sum_j^{R_{ij}<R_c} \frac{\partial E\left(\{R_{jm}\}\right)}{\partial R_i}, \text{ where } R_{jm} < R_c$$



C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

Nearsightedness of Electronic Matter

Nearsightedness principle exist in typical atomic systems[1]

Local electronic properties mostly depend on nearby atoms and electrons[2]

Quality of ML potentials rely on

- **Degree of nearsightedness** of electronic structure methods
- **Balance of nearsightedness** between ML models and electronic structure methods

[1] W. Kohn*, Phys. Rev. Lett.,* **76,** 17 (1996)
[2] E. Prodan *et. al., Proc. Natl. Acad. Sci.,* **102,** 33 (2005)

Why Forces?

More details of potential energy surfaces

- 1 energy *versus* 3N forces
- More useful in many simulations

Insights into nearsightedness of ab initio methods

- per-atom properties by design
- No well-defined atomic energies

C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

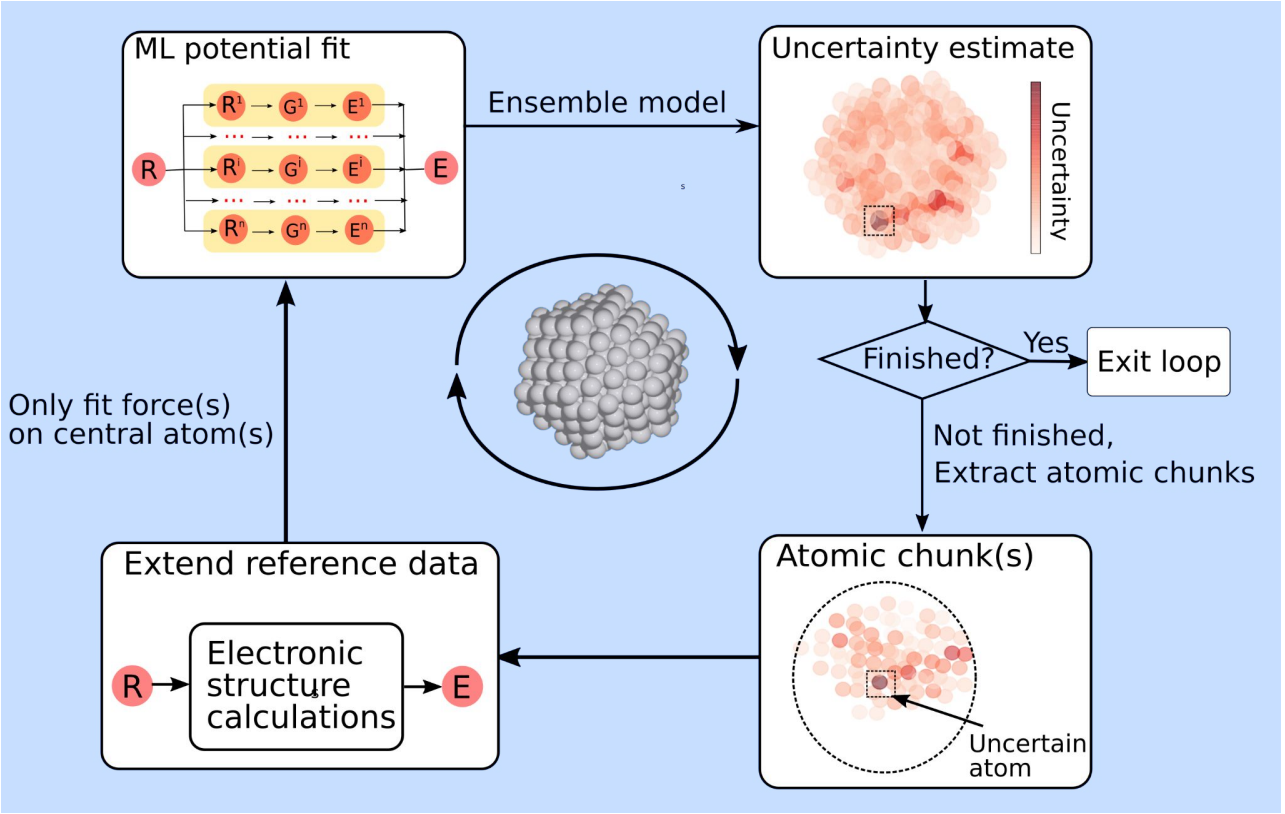
NFT Active Learning

Atomic uncertainty

$$\delta_i = 2.58\sigma_f = 2.58\sqrt{\frac{\sum_{j=1}^{M}\left\|\mathbf{f}_i^{(j)} - \overline{\mathbf{f}}_i\right\|^2}{M-1}}$$

Structure uncertainty

$$\delta = \max_i(\{\delta_i\})$$

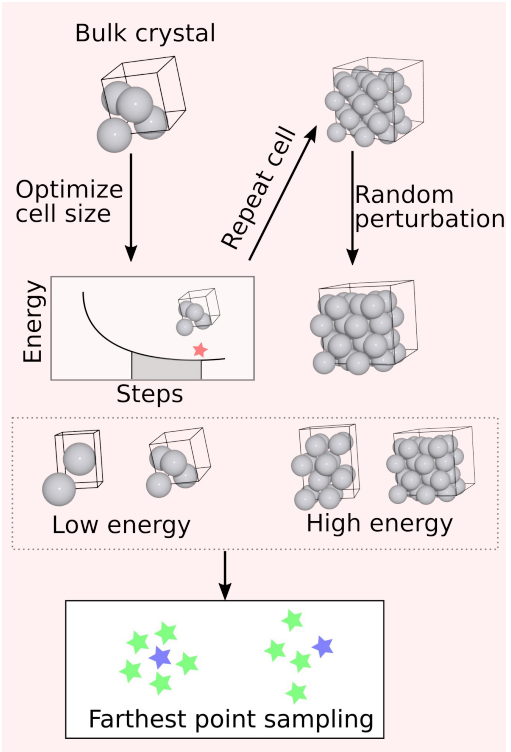


C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

Test System & Initialization

- **Atomic system.** Pt_{260} cuboctahedron nanoparticle. Atomic positions randomly displaced, resulting atomic forces in the range of [0.1, 5.1] eV/Å
- **Parent calculators**
  - **DFT**, a long-ranged calculator, conducted in GPAW[1]
  - EMT, a nearsighted calculator
- **Initial training structures.** 20 bulk cells with 2-16 atoms selected out of 30 structures
- **Machine learning models**
  - 10-member BP-NN ensemble models, trained with *Amp*[2]
  - Gaussian symmetry functions as descriptors/features
  - Ensemble average as ML predictions for energy and forces

Initialization



[1] J. Enkovaara *et. al.,* J. Phys. Condens. Matter., **22** (2010)
[2] A. Khorshidi *et. al.,* Comput. Phys. Commun., **207** (2016)

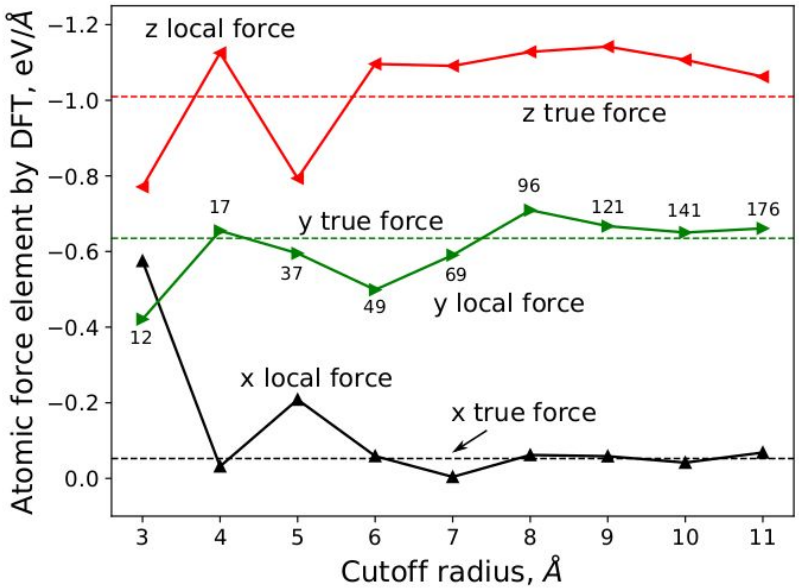
C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

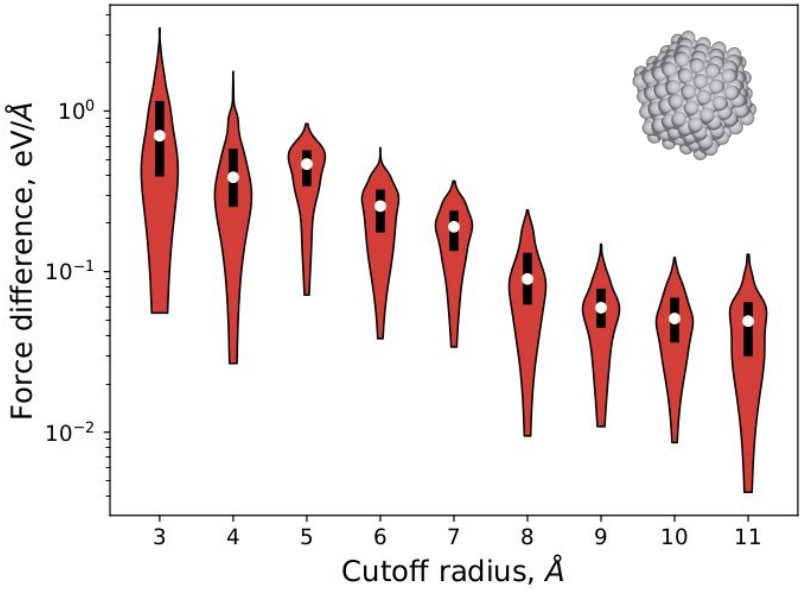
DFT Force Locality

Cutoff of 8 Å offers a good balance between accuracy and efficiency

- Maximum force difference of **0.24 eV/Å**
- Average force difference of **0.10 eV/Å**

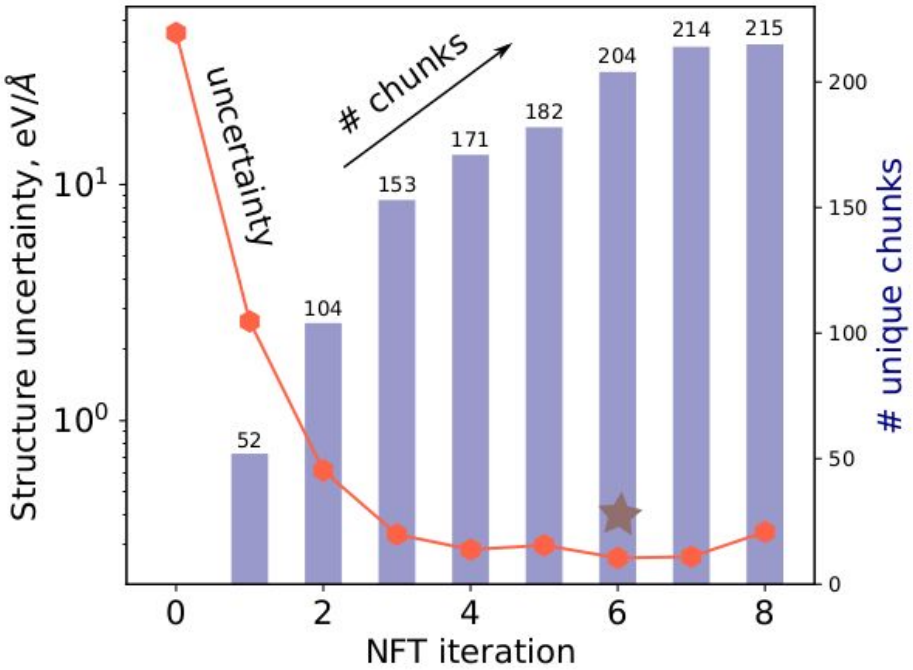
GPAW
Plane-waves mode
Cutoff=450 eV
PBE XC-functional
Gamma point





C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

Nearsighted Force Training



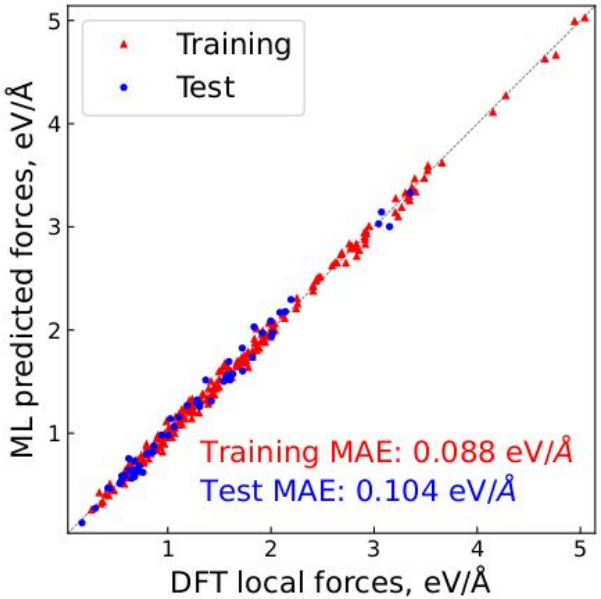
Learning on-the-fly

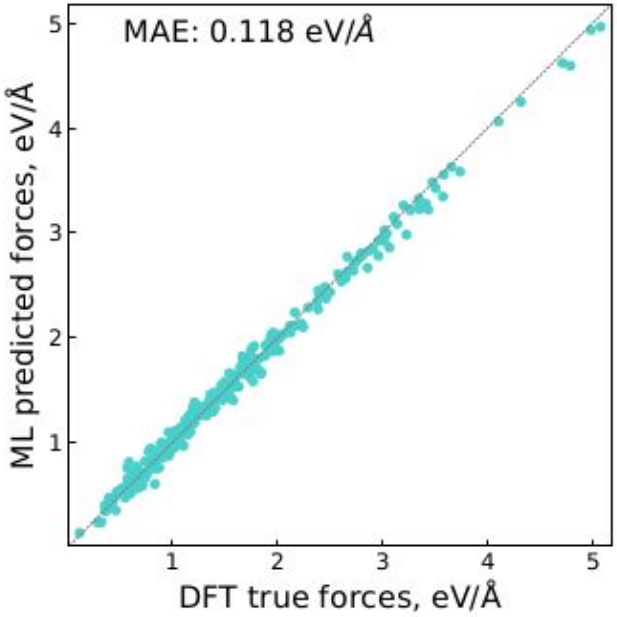
- Rate of addition of chunks diminishes
- Lowest uncertainty of 0.26 eV/Å obtained with six NFT iterations

C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

ML Forces *versus* DFT Forces

Low training and test MAEs

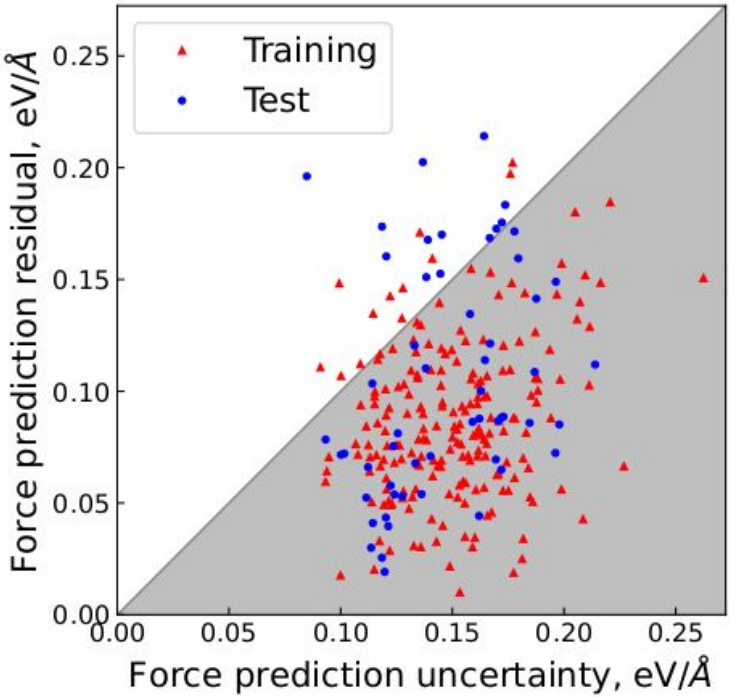




True fit is close to the best possible fit

C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

Prediction Accuracy *versus* Atomic Uncertainty

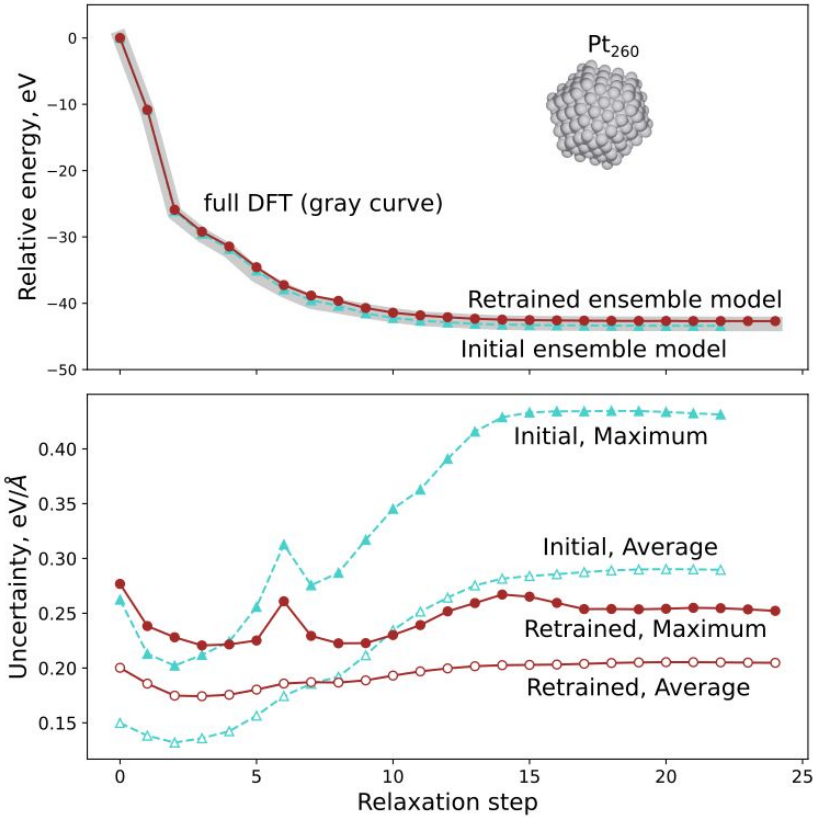


Errors controlled by uncertainties

- Most points below the parity line
- Structure uncertainty (0.26 eV/Å) provides an upper bound for the maximum prediction residual (0.21 eV/Å)

C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

Structure Optimization

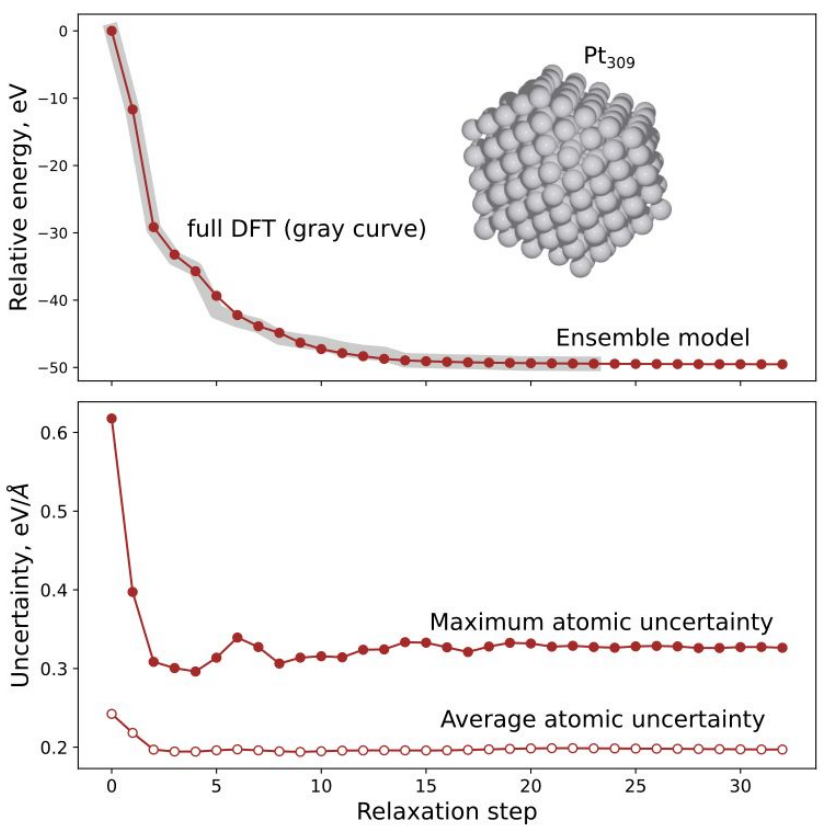


Addressing uncertainty in relaxation

- Initial model loses confidence after a few steps
- Uncertainty over the relaxation trajectory can be systematically addressed by adding new chunks to the fitting database

C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

Transferability to Larger Systems

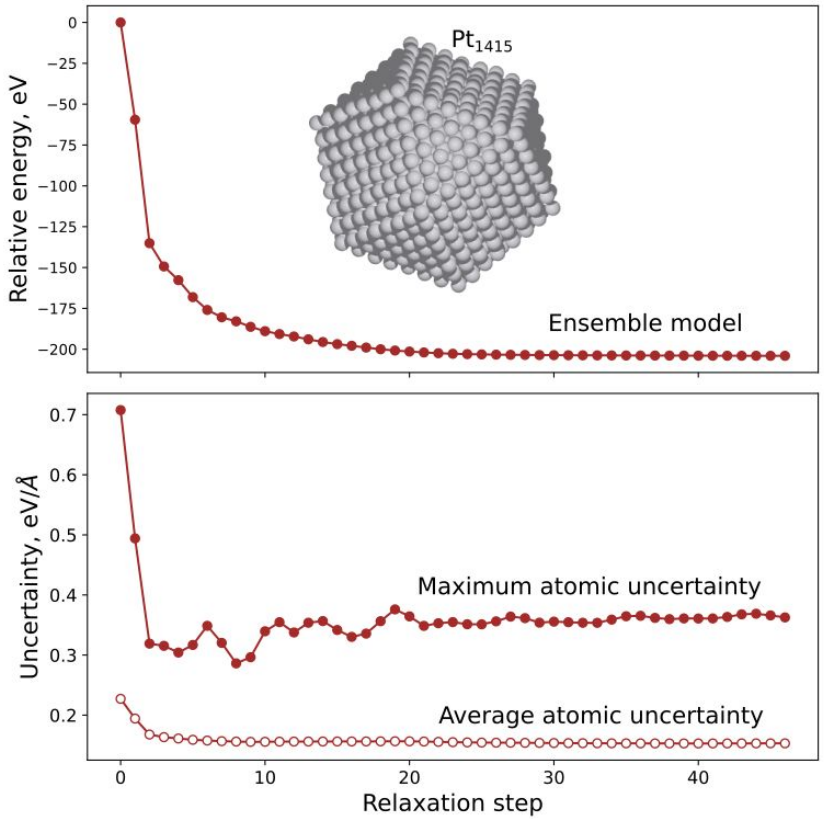


From uncertain to certain

- Relaxes to a familiar region after two steps
- MAE of 0.13 eV/Å for both initial and ML-relaxed structures

C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

Transferability to Larger Systems



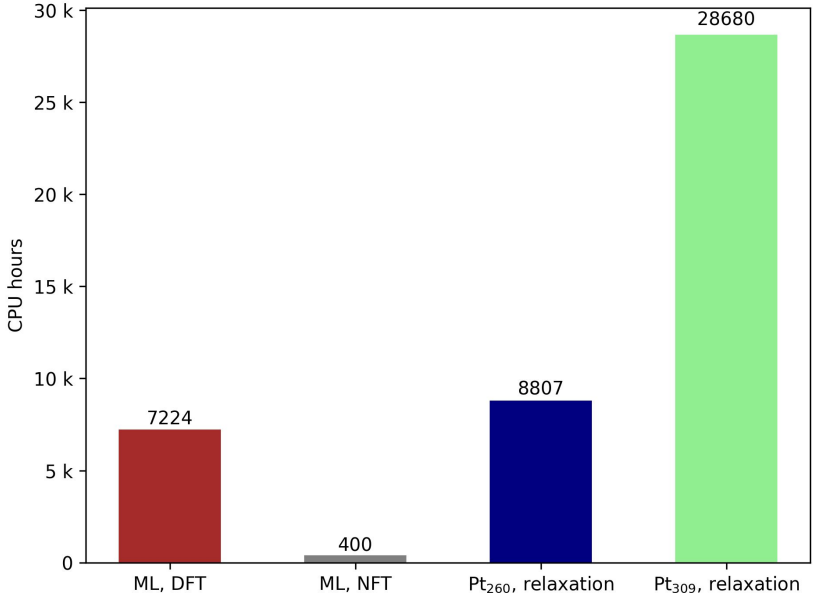
Beyond pure DFT with certainty

- Largest nanoparticle that has ever been studied in a pure DFT calculation*
- ML relaxed structure should be close to the true one by DFT, as indicated by uncertainties

*L. Li *et. al., J. Phys*. Chem. Lett., **4** (2013)

Computational Time, Scalability and Parallelizability

Computational time



Scalability and Parallelizability

- Sub O(N) scaling as only a fraction of chunks need to be evaluated by DFT
- Embarrassingly parallel as DFT jobs for chunks can be submitted individually

C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

Conclusions

We have developed a robust nearsighted force-training approach*

It allows for exploring configuration space in an active learning scheme

- Less familiar local chemical environments identified by per-atom uncertainties
- Small chunks carved out and evaluated by *ab initio* methods
- Addressing uncertainty by retraining only forces on central atoms of chunks

It exploits the nearsightedness of the parent calculator

- With a cutoff radius of 8 Å for atomic chunks, the upper-bound for the ability of ML calculators in replicating DFT true forces is 0.1 eV/Å of MAE
- Nearly perfect fit for a true nearsighted calculator, such as EMT

It offers significant computation savings for large structures

- O(N) scaling at worst
- Easy to parallelize

*Now available in *Amp*: https://bitbucket.org/andrewpeterson/amp/src/master/

Acknowledgements

Andrew Peterson

Xi Chen

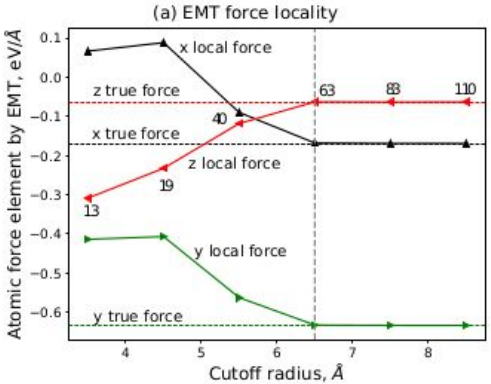
Mayank Agrawal

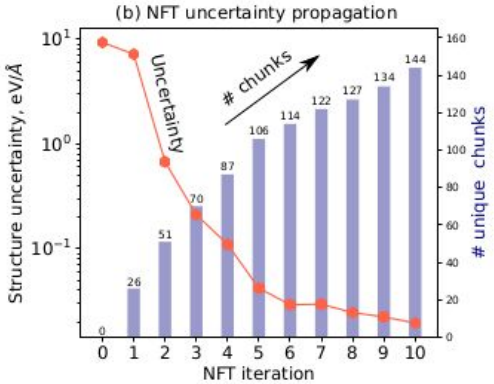
Catalyst
DESIGN LAB

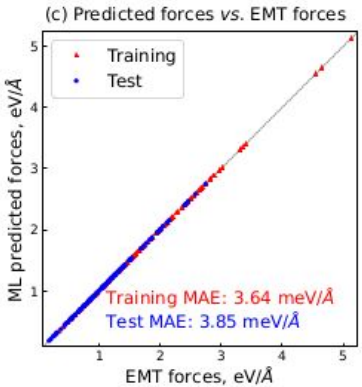
U.S. DEPARTMENT OF ENERGY

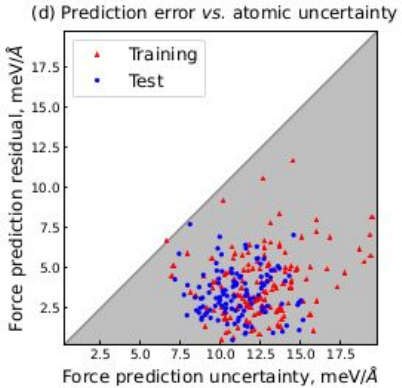
BROWN

EMT—Nearsighted Force Training





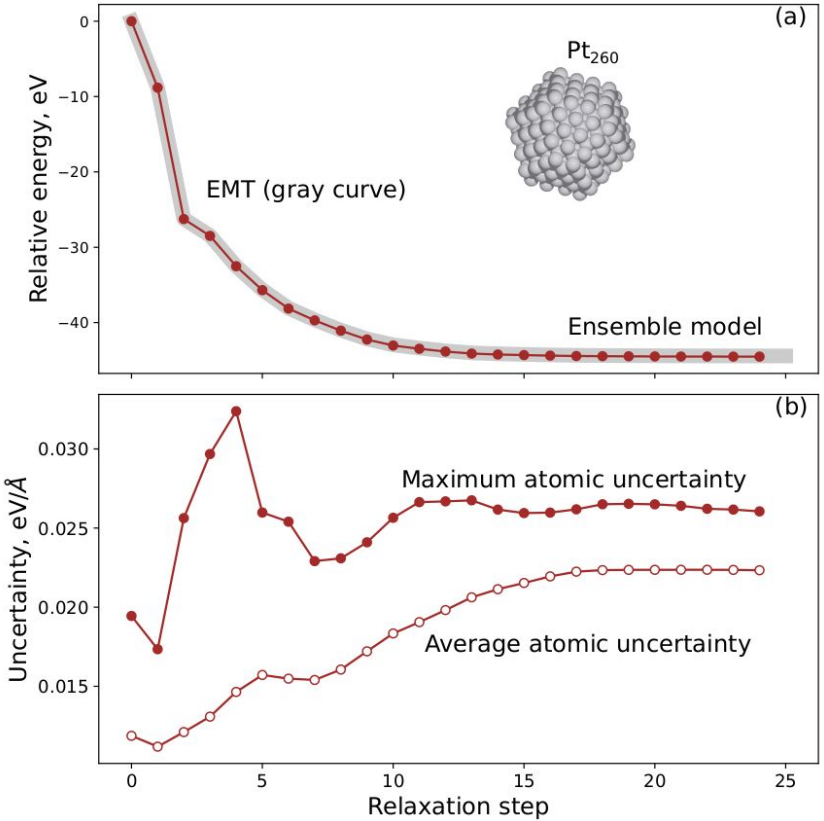




- True nearsighted calculator
- Much less chunks needed
- Much lower uncertainty

C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

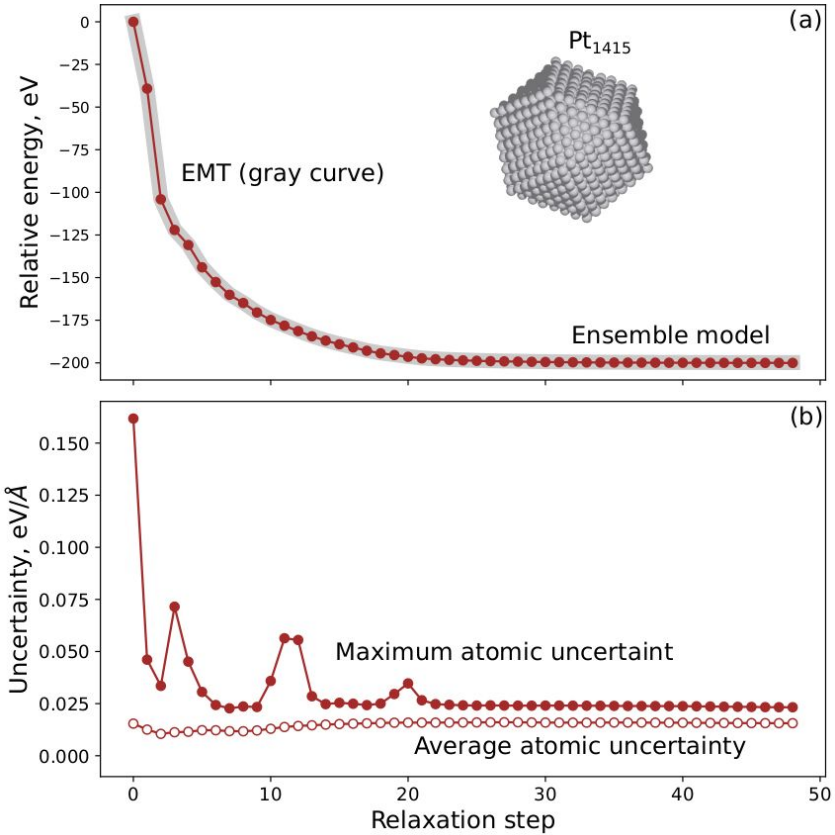
EMT—Structure Optimization



- Nearly identical results
- Controlled by uncertainty

C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

EMT—Transferability



- Perfect fit despite in extrapolation region
- Despite frequently entering less confident regions, it eventually relaxes to a structure which is almost identical to the true one by EMT.

C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

DFT—Computational Methods

Machine learning models

- 20 initial images with 2-16 atoms
- 20 input features, and a cutoff of 6.5 Å.
- Hidden layers of [5, 5]

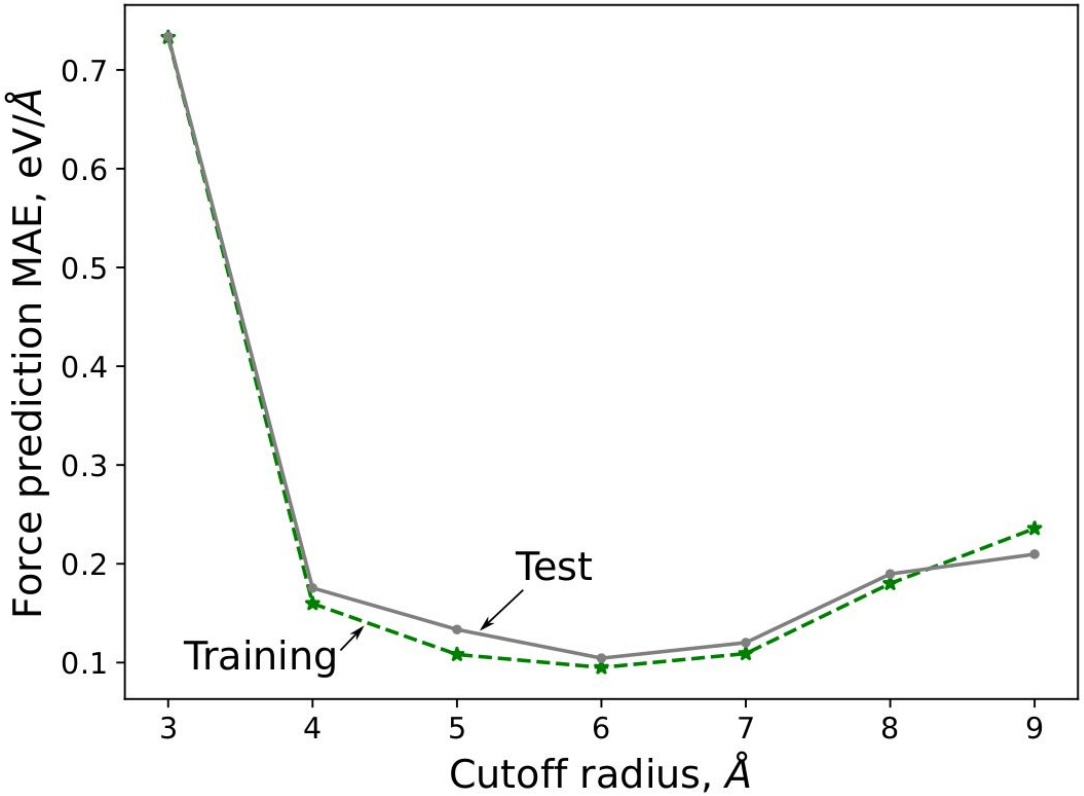
DFT settings

- GPAW plane-wave mode with a cutoff of 450 eV [2]
- PBE exchange-correlation functional
- Fermi-Dirac smearing of 0.1 eV
- K-point mesh
  - gamma point for clusters
  - at least 30/l for small bulk cells

[1] A. Khorshidi *et. al.,* Comput. Phys. Commun., **207** (2016)

[2] J. Enkovaara *et. al.,* J. Phys. Condens. Matter., **22** (2010)

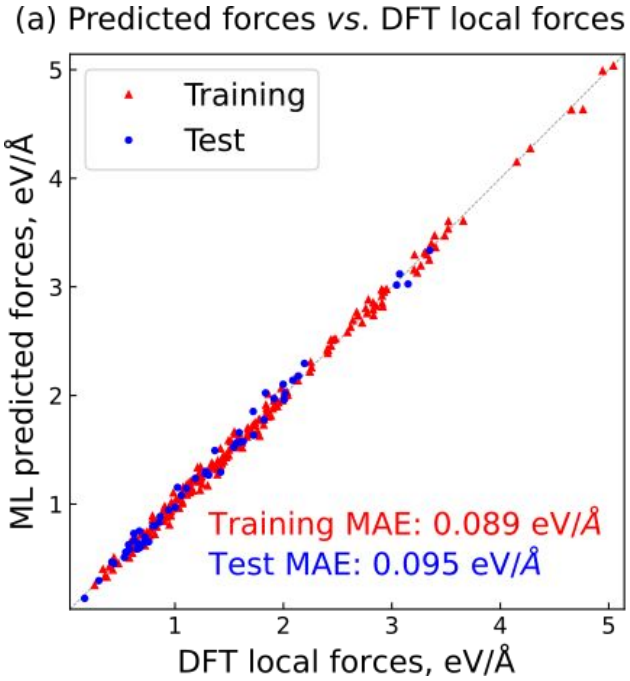
Model Cutoff: DFT-based Reference Data

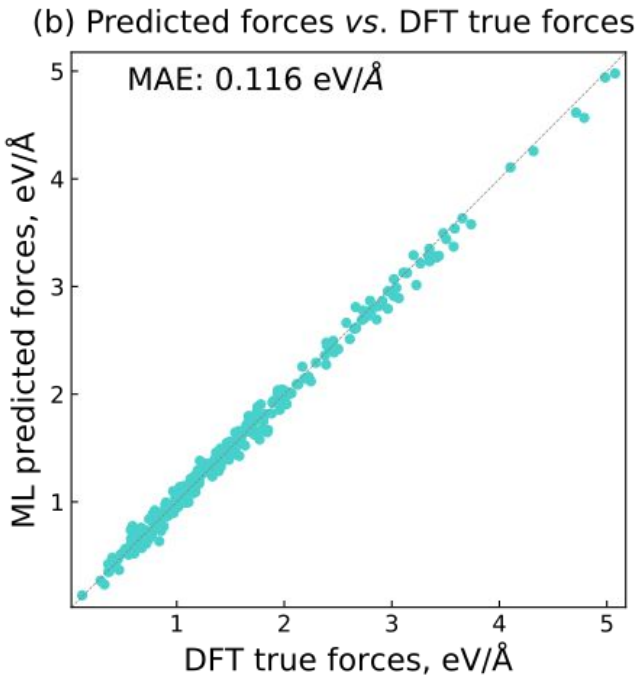


- Lowest Force MAE achieved with a ML model cutoff of 6-7 Å

Retrained Model: ML forces *versus* DFT forces

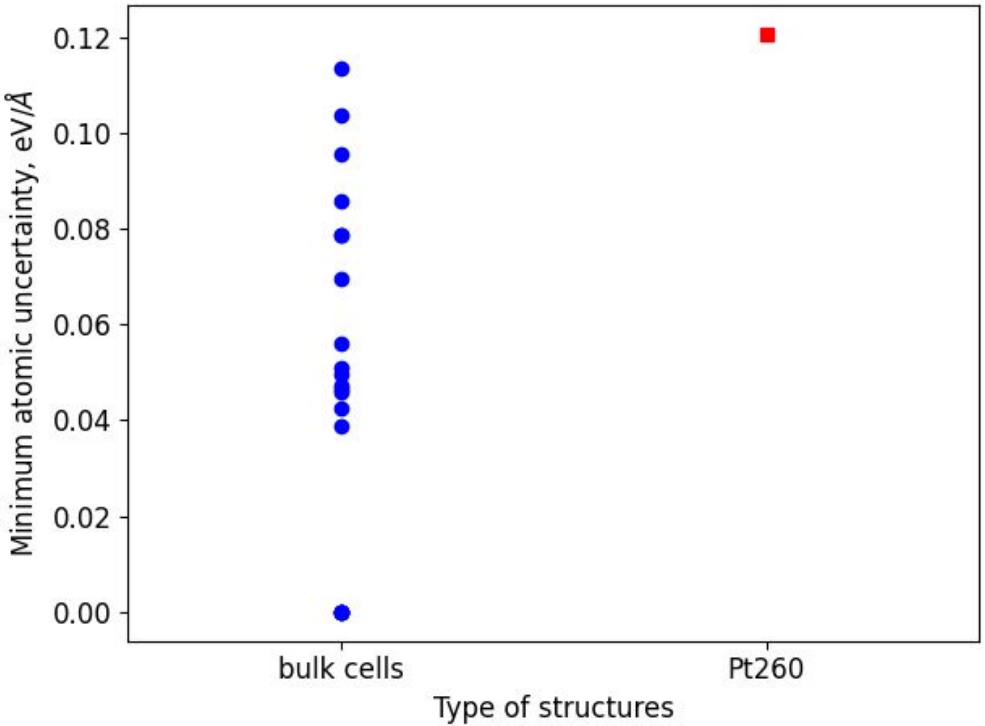
Almost identical to the initial model





C. Zeng, X. Chen and A. A. Peterson, *J. Chem. Phys.* (2022) (accepted)

DFT—Why no low-uncertainty atoms



- No true bulk atoms in the ‘rattled ’ Pt260 nanoparticle