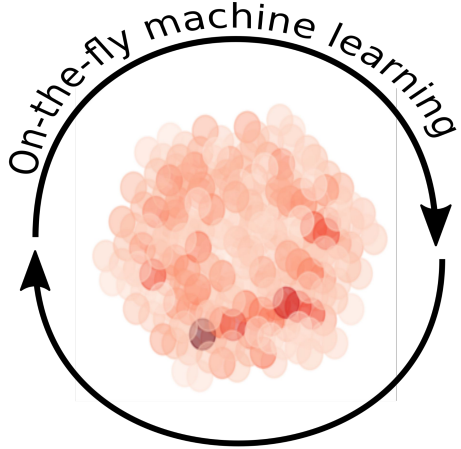
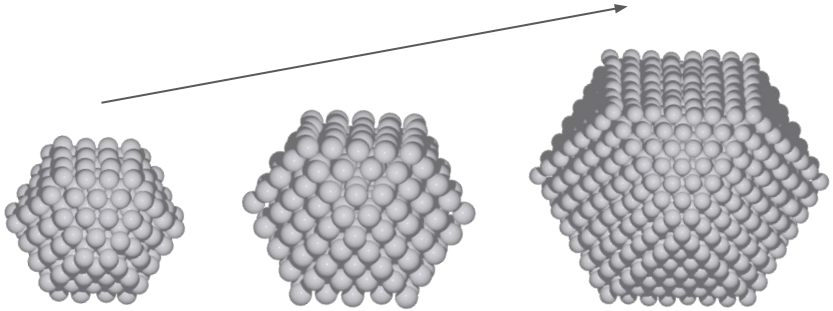
BROWN

Catalyst Design Lab

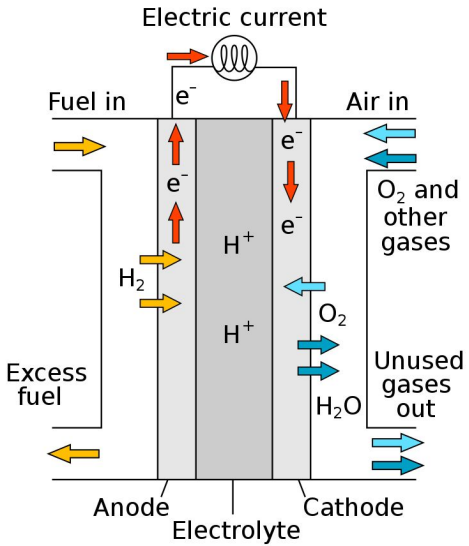
On-the-fly Machine Learning for Large Atomic Structures

Cheng Zeng





Fuel Cell & Oxygen Reduction Reaction (ORR)



Advantages

- CO_2 emission free
- High power density

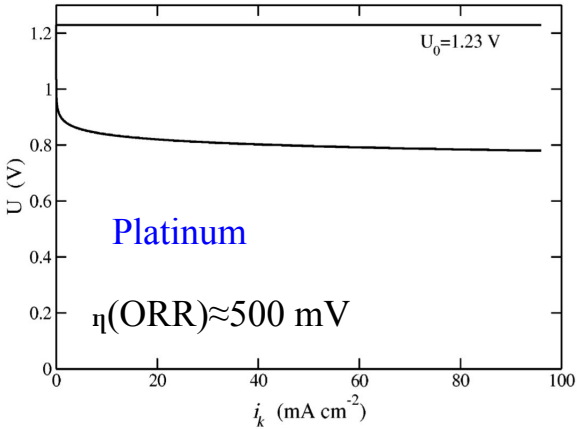
Disadvantages

- Sluggish ORR kinetics

Solution?

Pt alloys as ORR catalysts

Platinum is considered to be the best elemental catalyst for ORR.



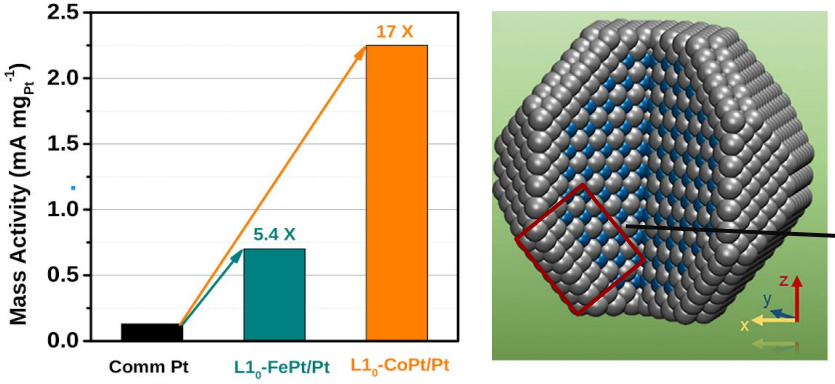
Anode: $2H_2 \longrightarrow 4H^+ + 4e^-$ (HOR)

Cathode: $O_2 + 4H^+ + 4e^- \longrightarrow 2H_2O$ (ORR)

J. K. Nørskov et al., *J. Phys. Chem. B*, **108** (2004)

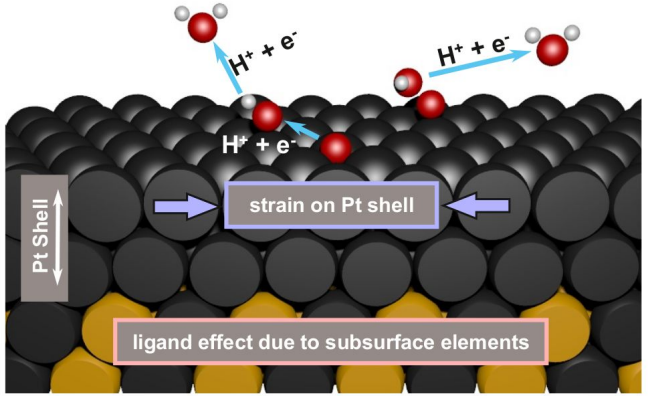
Pt Alloy Nanoparticles for Enhanced ORR

CoPt core-shell nanoparticle surpasses DOE targets for both ORR activity and stability



J. Li et al., *Joule*, **3**: 124-135 (2019)

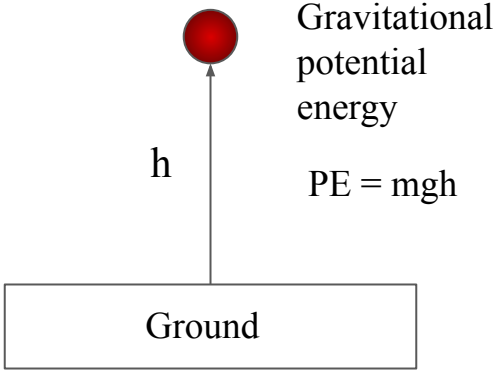
Slab models to understand physical origins



S. Sharma et al., *J. Chem. Phys.*, **150** (2019)

- The entire nanoparticle has on the order of 10,000 atoms.
- Quantum mechanical (QM) calculations for slabs/surfaces are limited to ~500 atoms.
- Potential energies are the central properties of atomic systems in QM calculations.

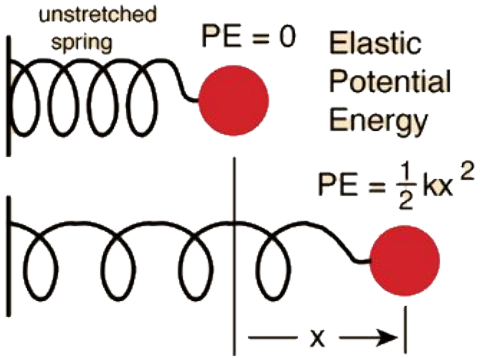
Potential Energy (PE)



Gravitational potential energy

PE = mgh

Force of gravity: f = - mg



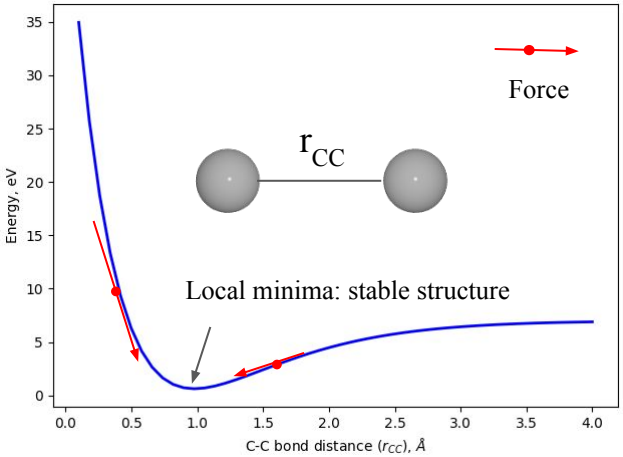
Force of spring: f = - kx

- Potential energies depend only on relative positions.
- Forces drive the state to lower potential energies (more stable states).

Potential Energy Surfaces

Potential energy surfaces (PES) provide a numerical framework to describe chemistry. Each structure (geometry) is associated with a unique energy. Since geometry changes are smooth, the energy landscape created is also smooth. In this way, chemistry becomes topology.

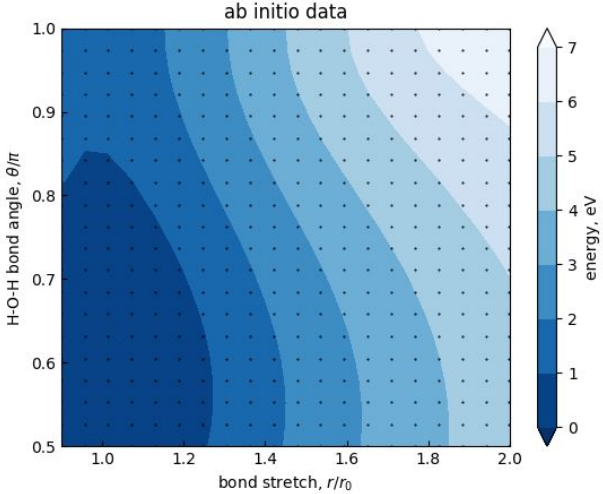
The simplest example: 1D PES

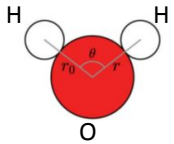


$$E = E(\{\vec{R}\})$$

$$\mathbf{F}_i = -\left(\frac{\partial E}{\partial \mathbf{R}_i}\right)_{j \neq i}$$

Single water molecule: 2D PES





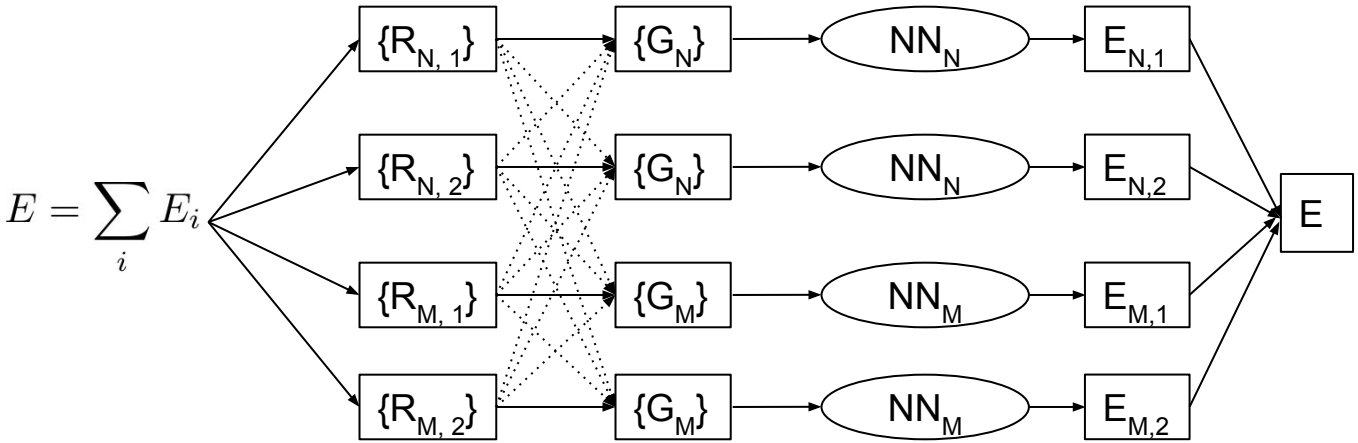
- Forces are by design atomic properties whereas energy is a structure property.
- Forces indicate how close the structure is to a local minima.

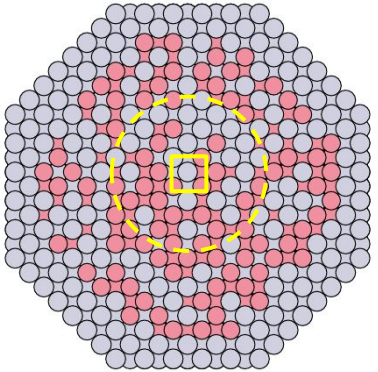
Atom-centered Neural Network Potential

Neural network (NN) math fitting has gained its momentum in the last two decades because it can fit any arbitrary function, which makes it ideal for fitting the PES whose function form is unknown.

$$E = E(\{\vec{R}\})$$

Atom-centered neural network

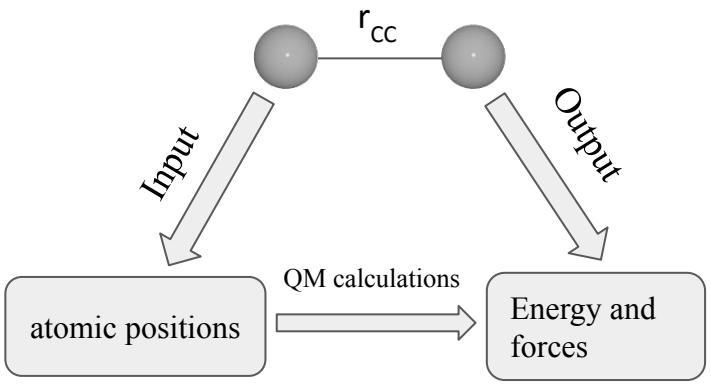




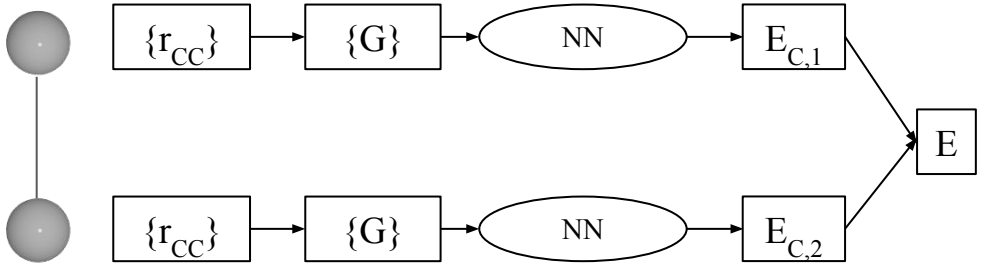
- Each atomic energy comes from the atom interacting with its neighbors (local chemical environment).
- Local chemical environments are the genes of NN potentials.

How Atom-centered NN Potential Works?

Example training data



Journey of a structure in atom-centered NN



Forces: $\mathbf{F}_i = -\left(\frac{\partial E}{\partial \mathbf{R}_i}\right)_{j \neq i}$

Loss function

$$Loss = \frac{1}{2}\sum_{j=1}^{M}\left\{\left(E_j/N_j - \widehat{E}_j/N_j\right)^2 + \frac{\alpha}{3N_j}\sum_{k=1}^{3}\sum_{i=1}^{N_j}\left(F_{ik} - \widehat{F}_{ik}\right)^2\right\}$$

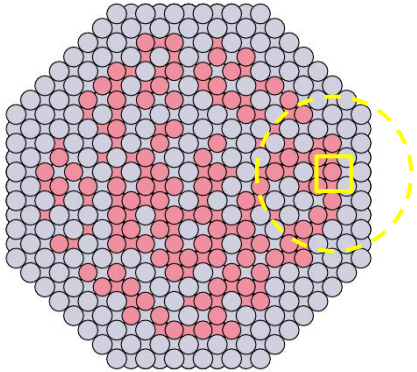
Challenges in NN potentials

Normally a large amount of training data are needed.

- On-the-fly learning to minimize the number of expensive quantum mechanical calculations

How to generate ***small atomic structures*** when ML predictions ***fail on a large structure***?

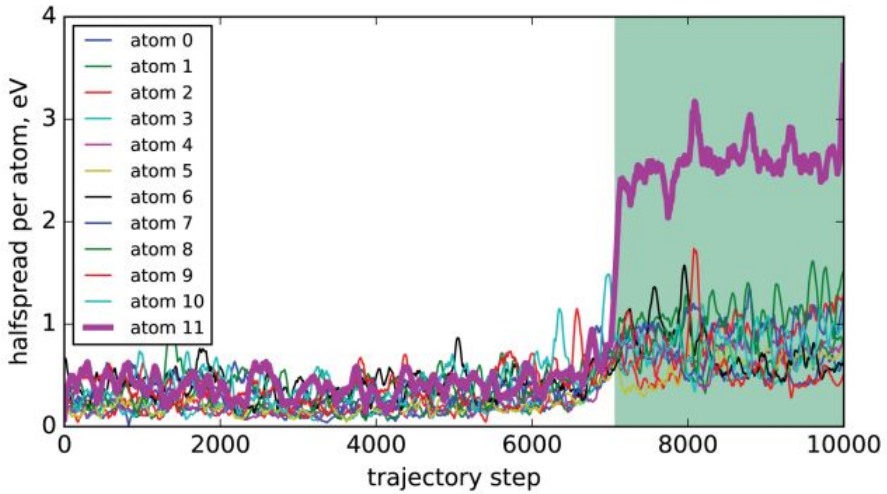
- Per-atom uncertainty quantification
- Effective range of atomic interactions

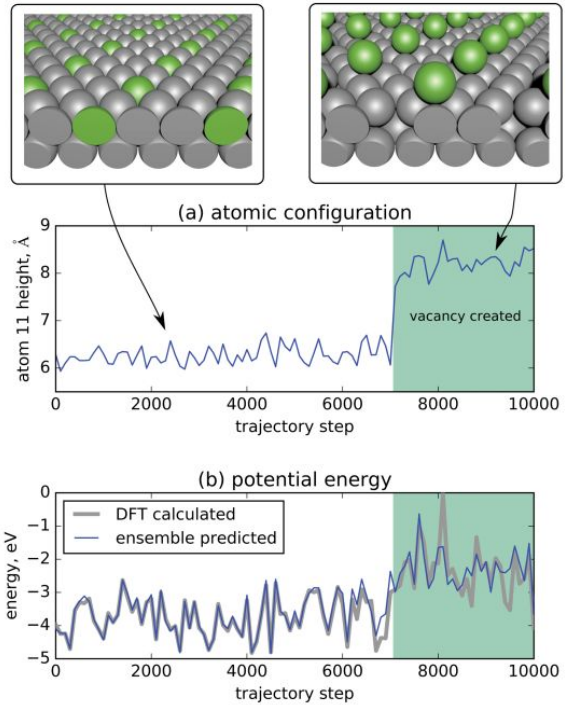


Localizing Prediction Uncertainty to Atoms

Ensemble models can isolate prediction errors to atoms.

Largest error from the atom moving onto the surface

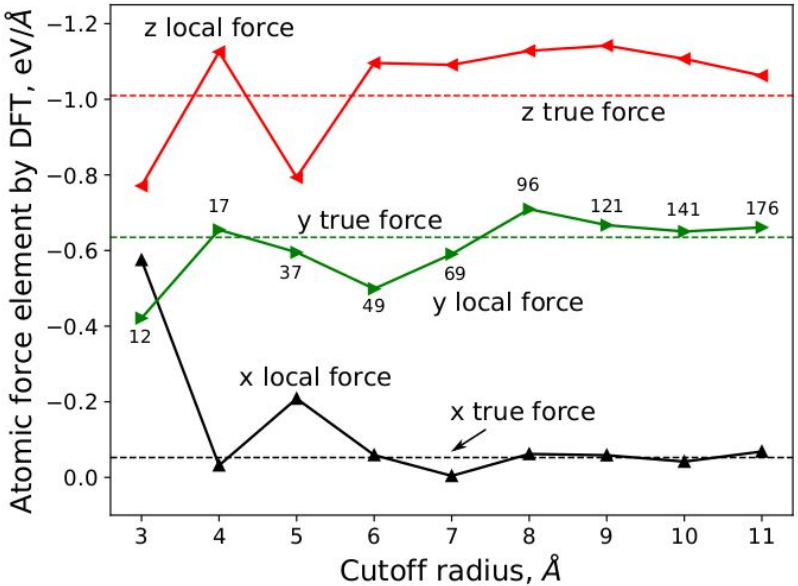


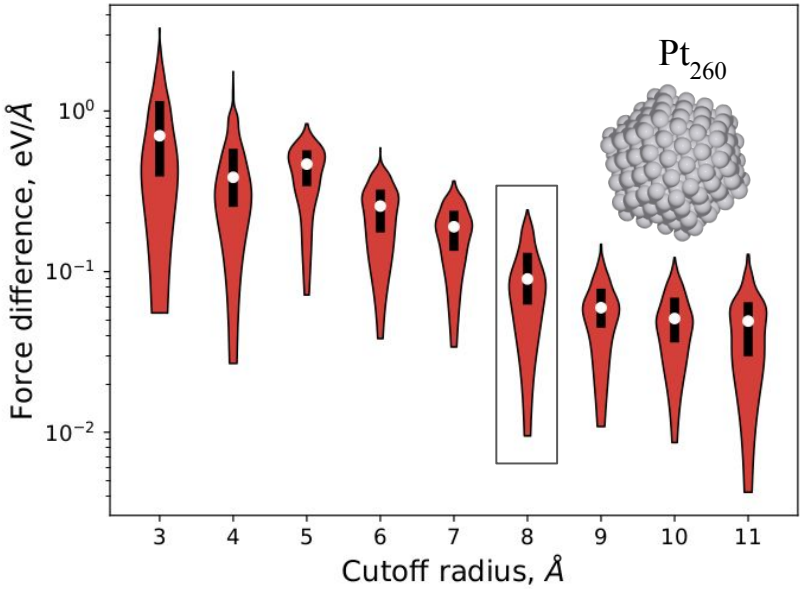


A. A. Peterson *et al., Phys. Chem. Chem. Phys.,* **19** (2017)

Effective Range of Atomic Interactions

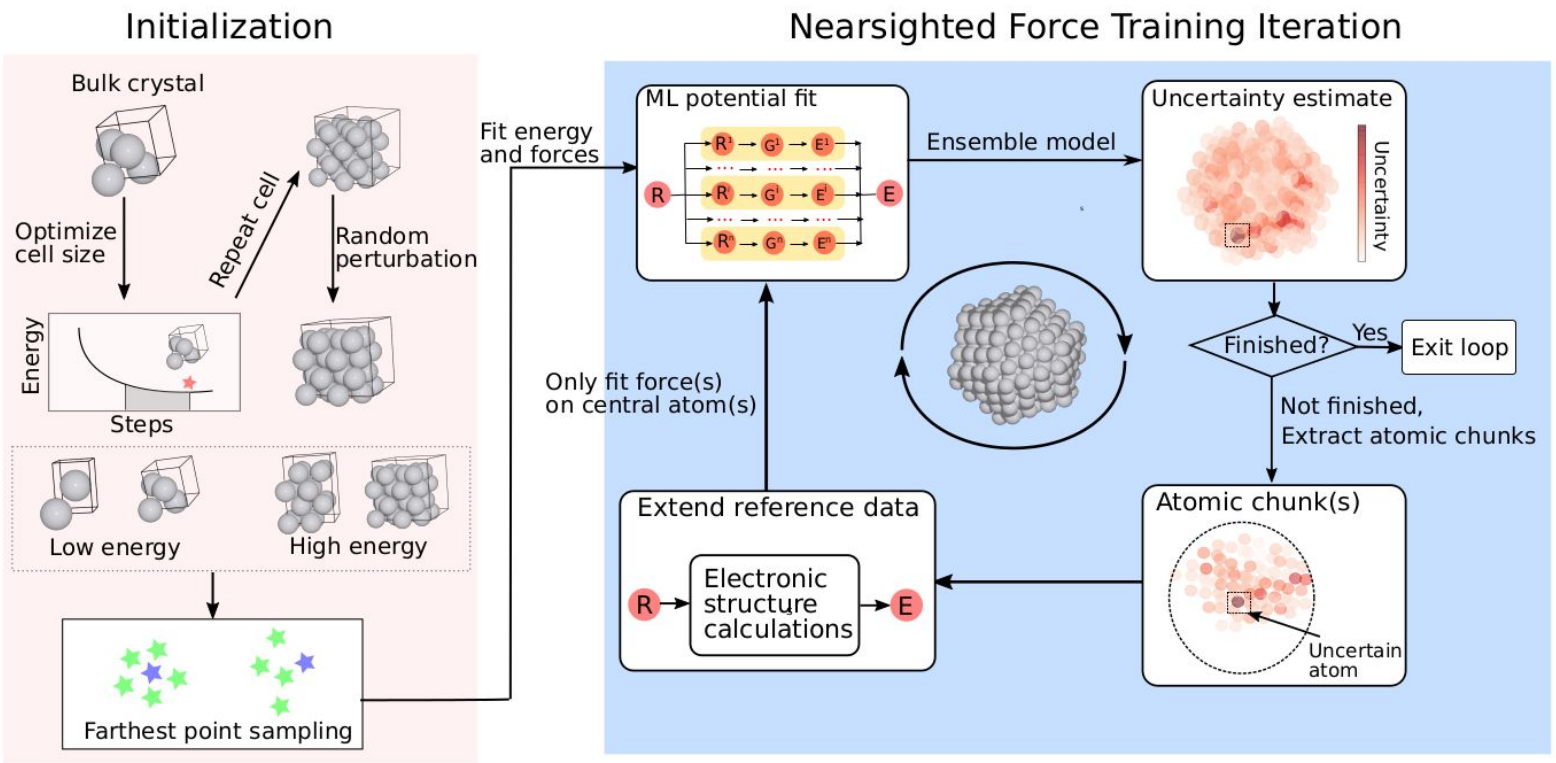
Cutoff of 8 Å offers a good balance between accuracy and efficiency





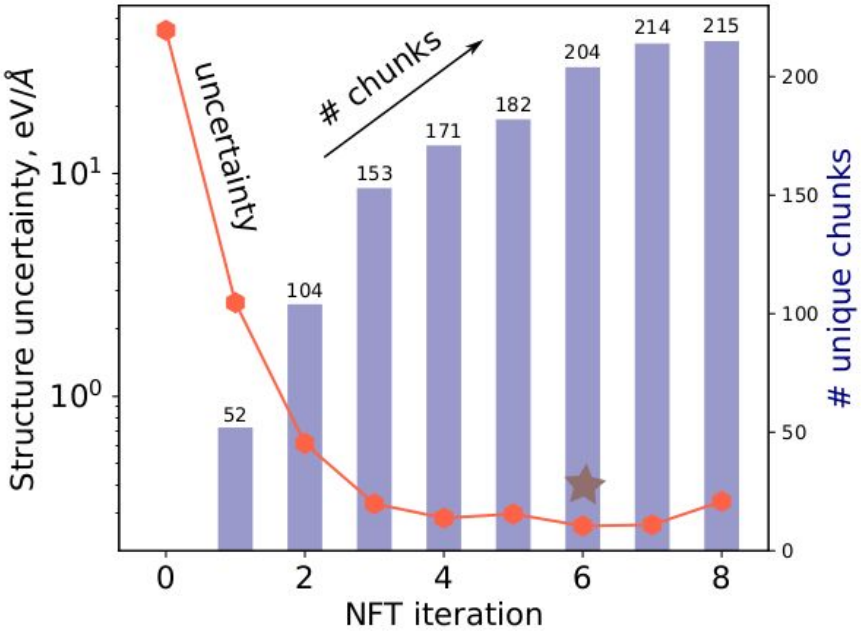
C. Zeng et al., *J. Chem. Phys.* **156** (2022)

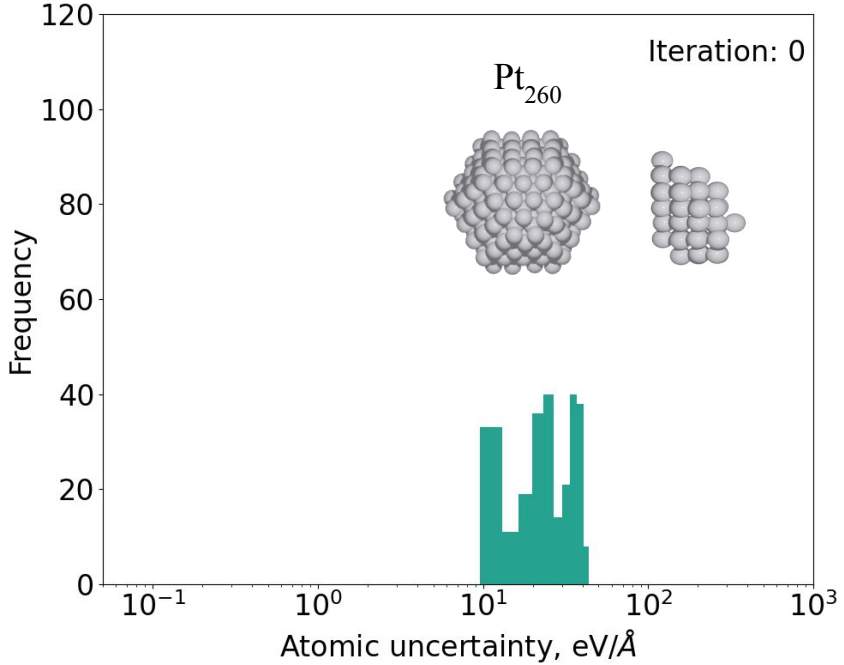
On-the-fly Machine Learning



C. Zeng et al., *J. Chem. Phys.* **156** (2022)

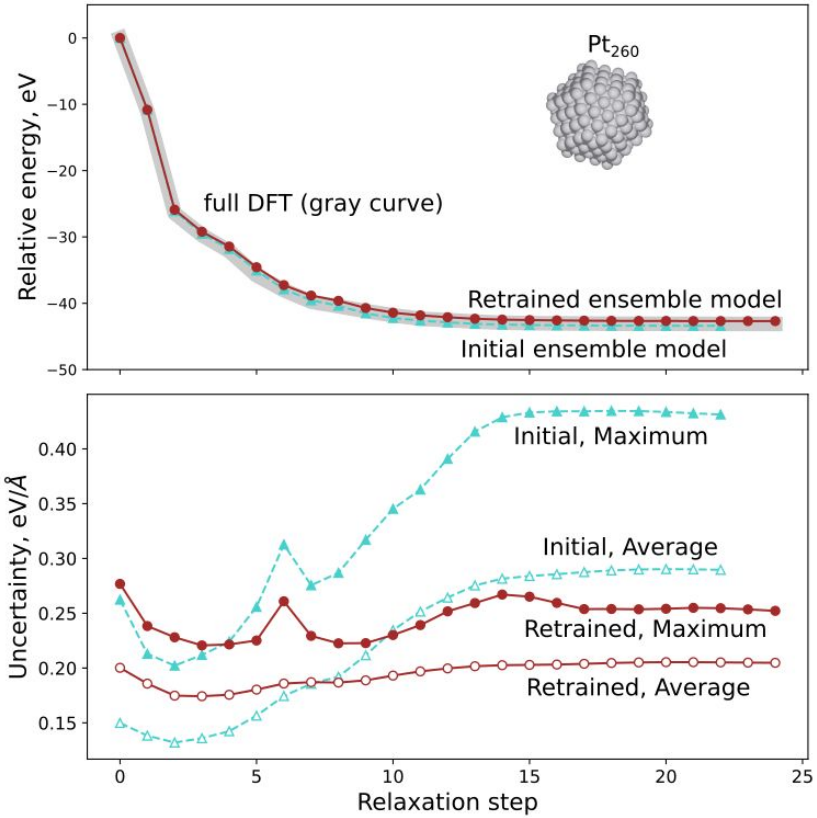
Benchmark on a Pt Nanoparticle with 260 Atoms





C. Zeng et al., *J. Chem. Phys.* **156** (2022)

Finding the Stable Structure of Pt_{260}



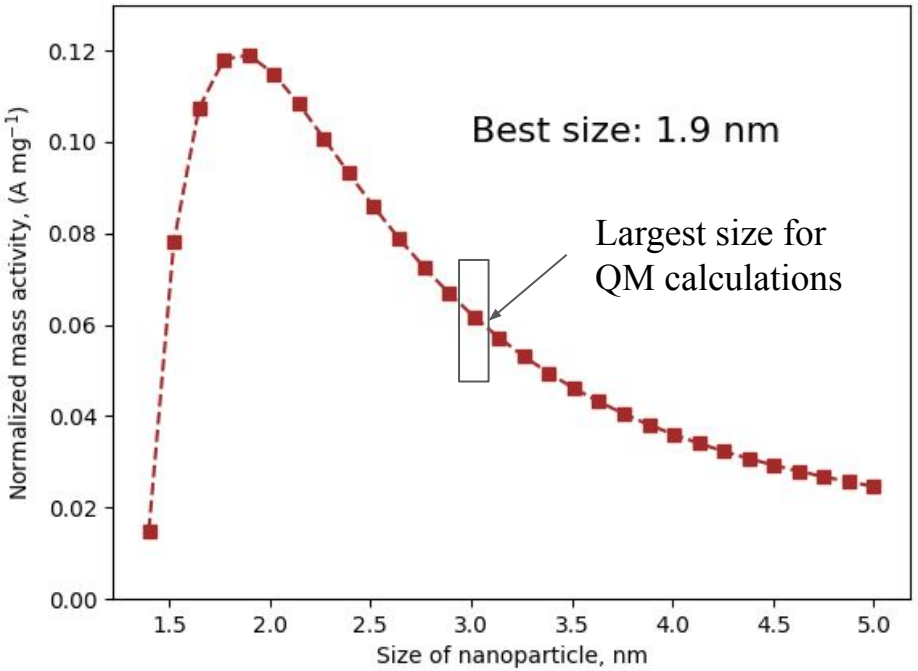
Addressing uncertainty in relaxation

- Initial model loses confidence after a few relaxation steps.
- Uncertainty can be systematically addressed by fitting new chunks identified by the ensemble model.

C. Zeng et al., *J. Chem. Phys.* **156** (2022)

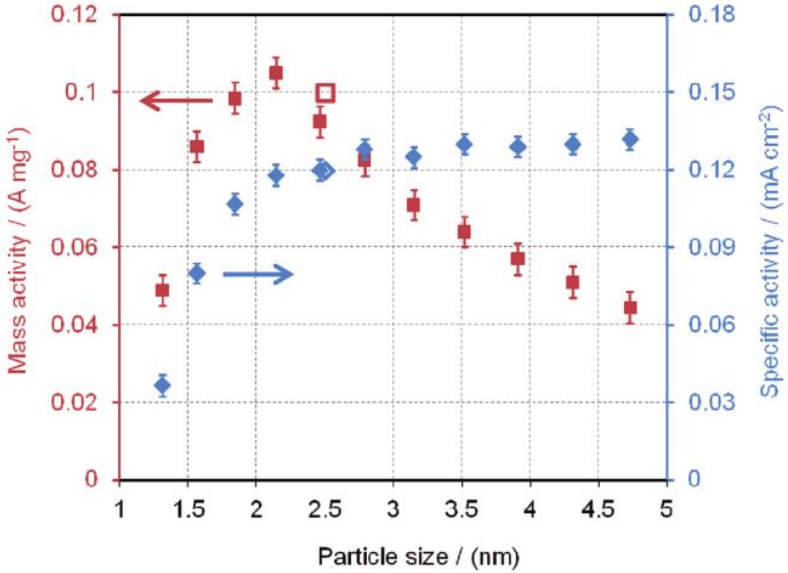
Trends in ORR Activity for Pt Nanoparticles

ML predicted ORR mass activity



C. Zeng and A. A. Peterson (manuscript under preparation)

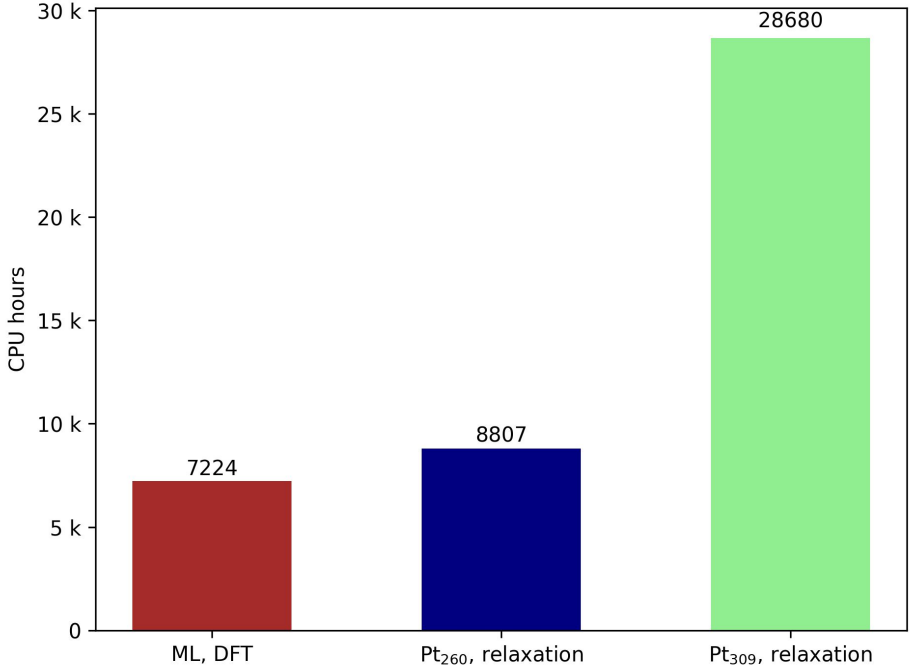
Experiment, maximum activity at 2.2 nm



M. Shao et al., *Nano. Lett.*, **11**: 3714-3719 (2011)

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 QM.
- 'Embarrassingly' parallel since QM jobs for chunks can be submitted individually.

C. Zeng et al., *J. Chem. Phys.* **156** (2022)

Conclusions

We have developed a robust learning on-the-fly approach*

It allows for exploring PES in a self-guided manner

- 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 offers significant computation savings for large atomic structures

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

*Now available in the open-source machine learning package '*AMP*': https://bitbucket.org/andrewpeterson/amp/src/master/