



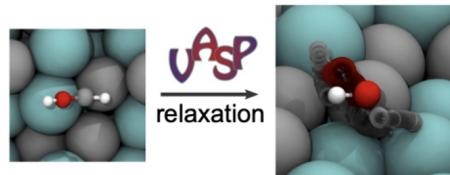
# Hands-on 9 : MLP02 - Universal MLPs and revisiting hands-on 2, 3, 4

## ▼ Recap

### Time scale limitation of DFT

#### Challenges of Theoretical Modeling

##### Huge computational cost of DFT calculations



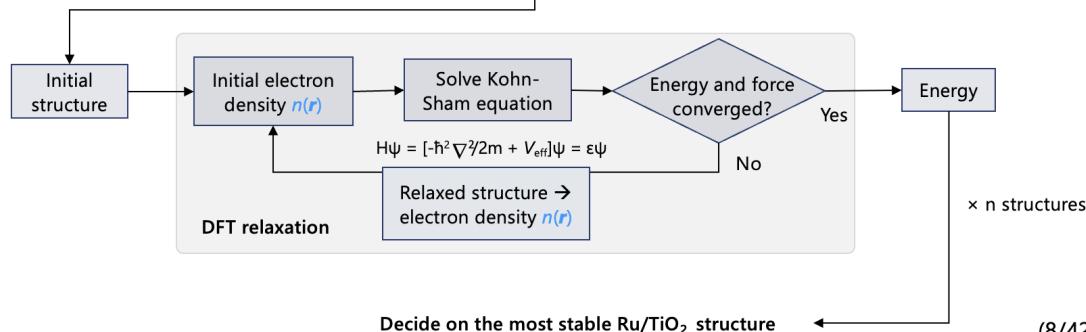
- ✓ The DFT computation scales  $O(n^3)$  with the number of electrons in the system.
- ✓ The DFT computations of 100 atoms approximately take ~1 hour per relaxation on 12 core CPUs.  
→ few hours~ day for get energy of one DFT structure

Ulissi et al. ACS Catal., 11, 10, 6059 (2021).



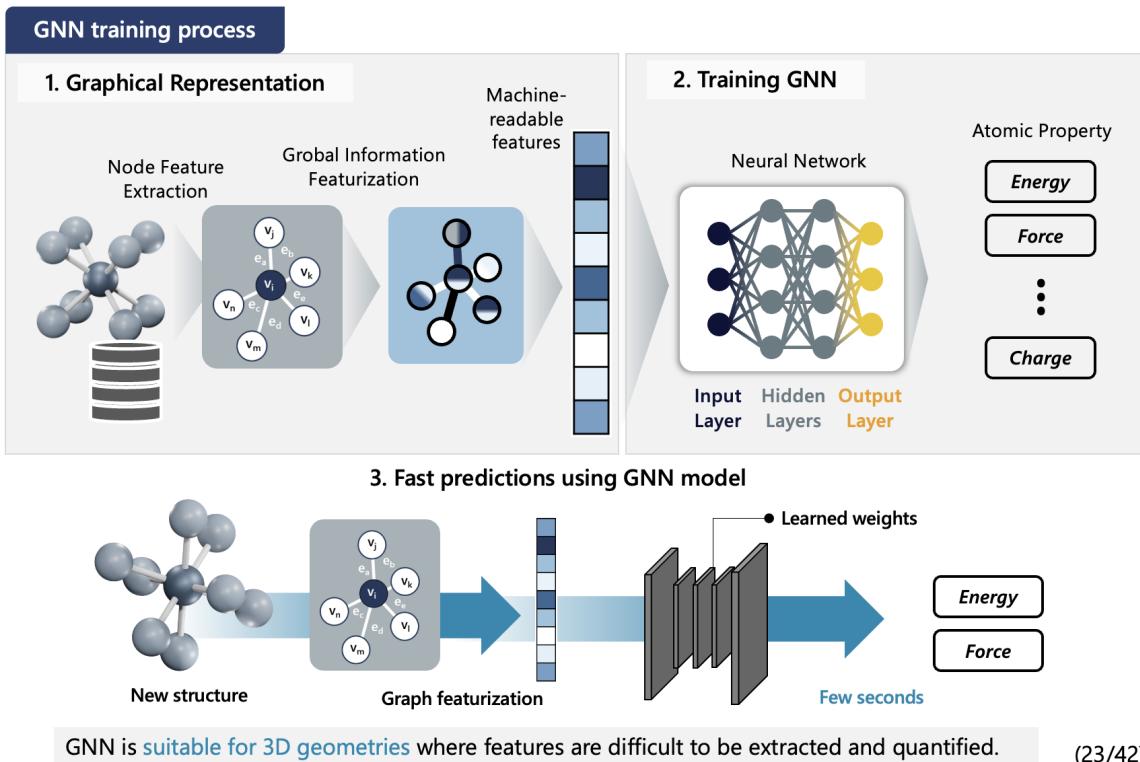
"I want the most stable Ru SAC structures on TiO<sub>2</sub> surface."

"Heuristically" determined initial structures of Ru/TiO<sub>2</sub>



### Machine learning Potential (MLP) using graph neural network

# Graph Neural Network (GNN)

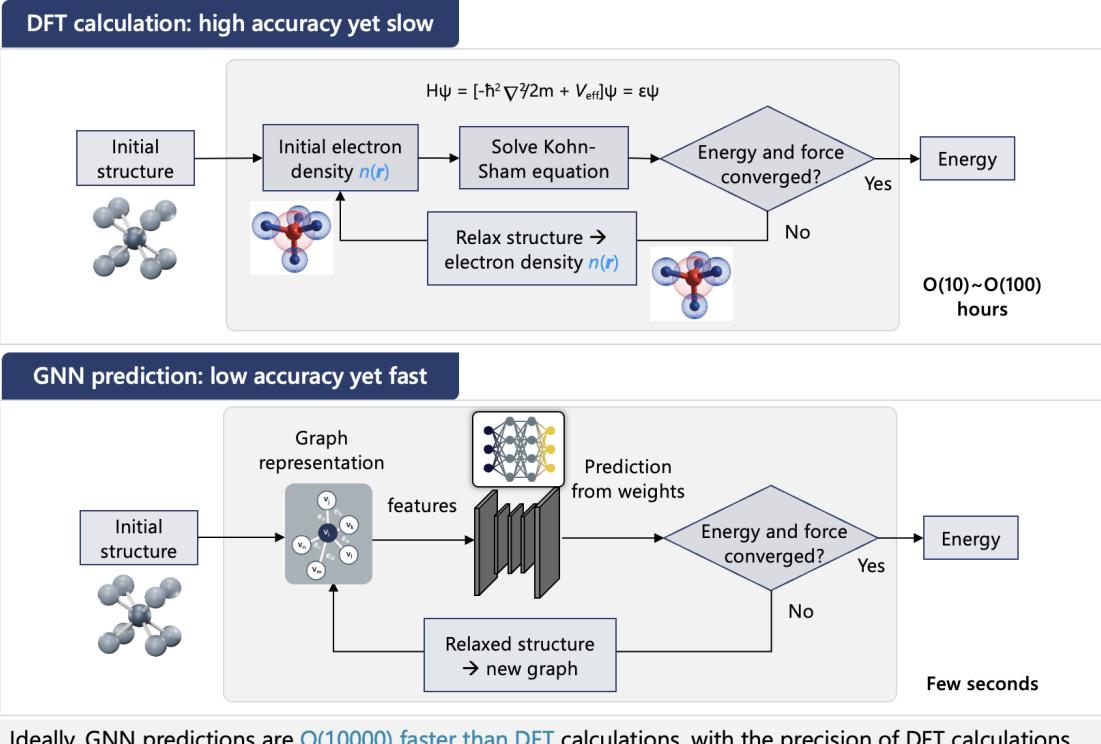


<https://www.sciencedirect.com/science/article/pii/S138589472404244X#da005>

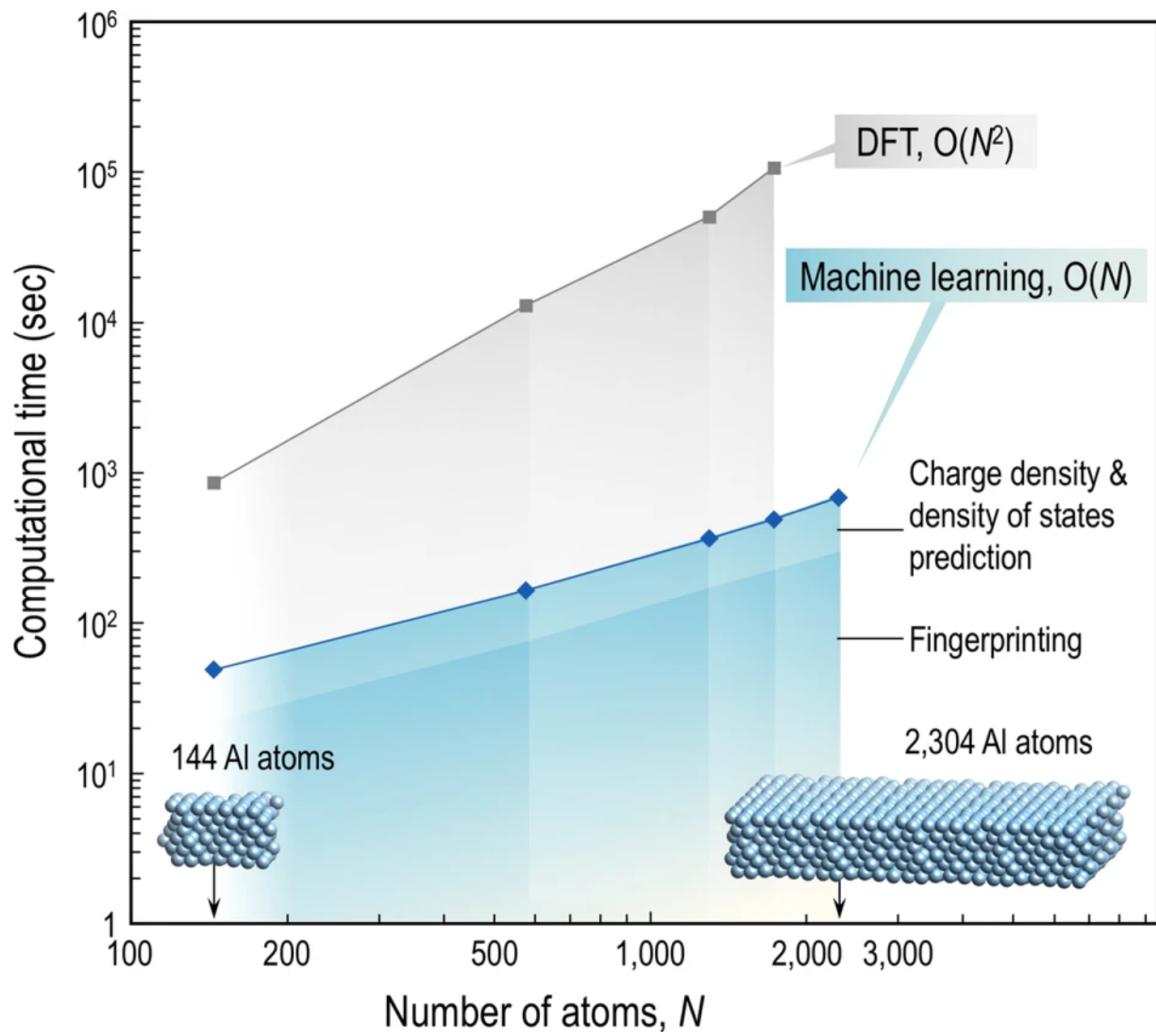
## Why we use MLP?

Computational cost of MLP compared to DFT

# GNN can Ideally Replace DFT Calculations



Length scale and computational cost of MLP compared to DFT

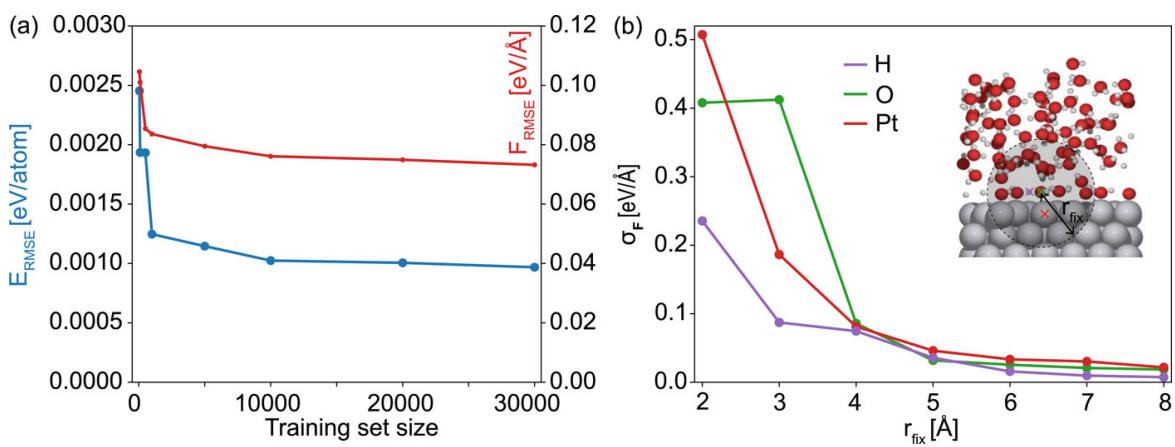
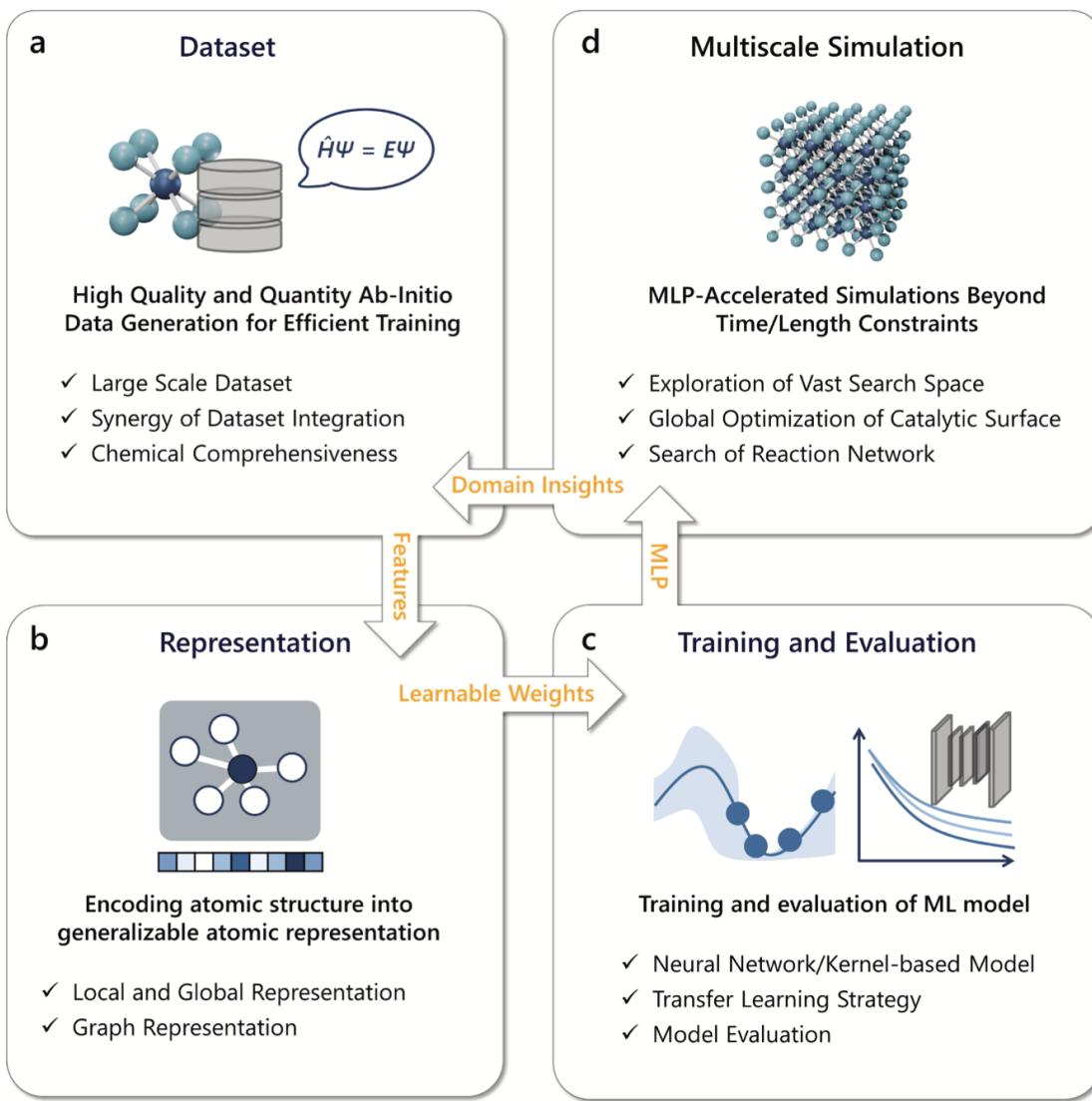


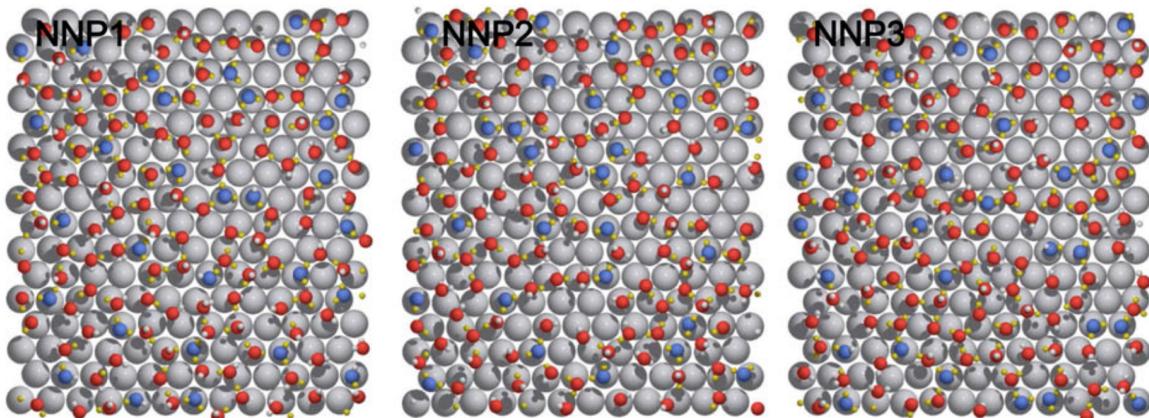
Chandrasekaran *et al.* *npj. Comp. Mater.*, 5, 22, (2019).

**This is because the computational requirements of MLP are lower than those of DFT**

- DFT computational cost =  $O(N^2)$
- MLP computational cost =  $O(N)$

## Typical workflow of application of MLP

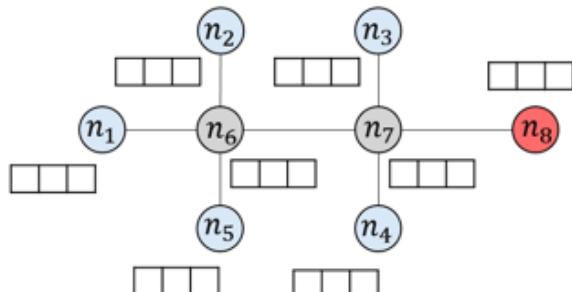
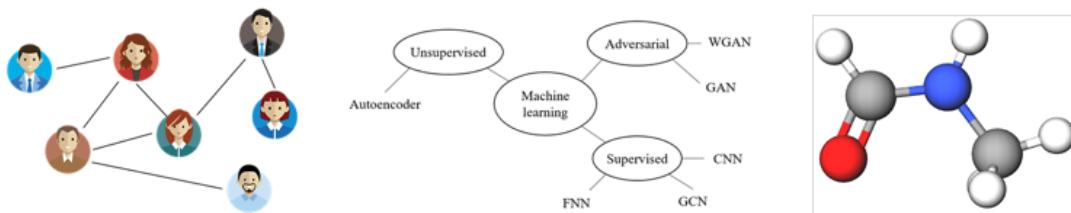




Karsten W. Jacobsen et al. J. Chem. Phys. 155, 224701 (2021)

## Architectures of MLPs (GNNS)

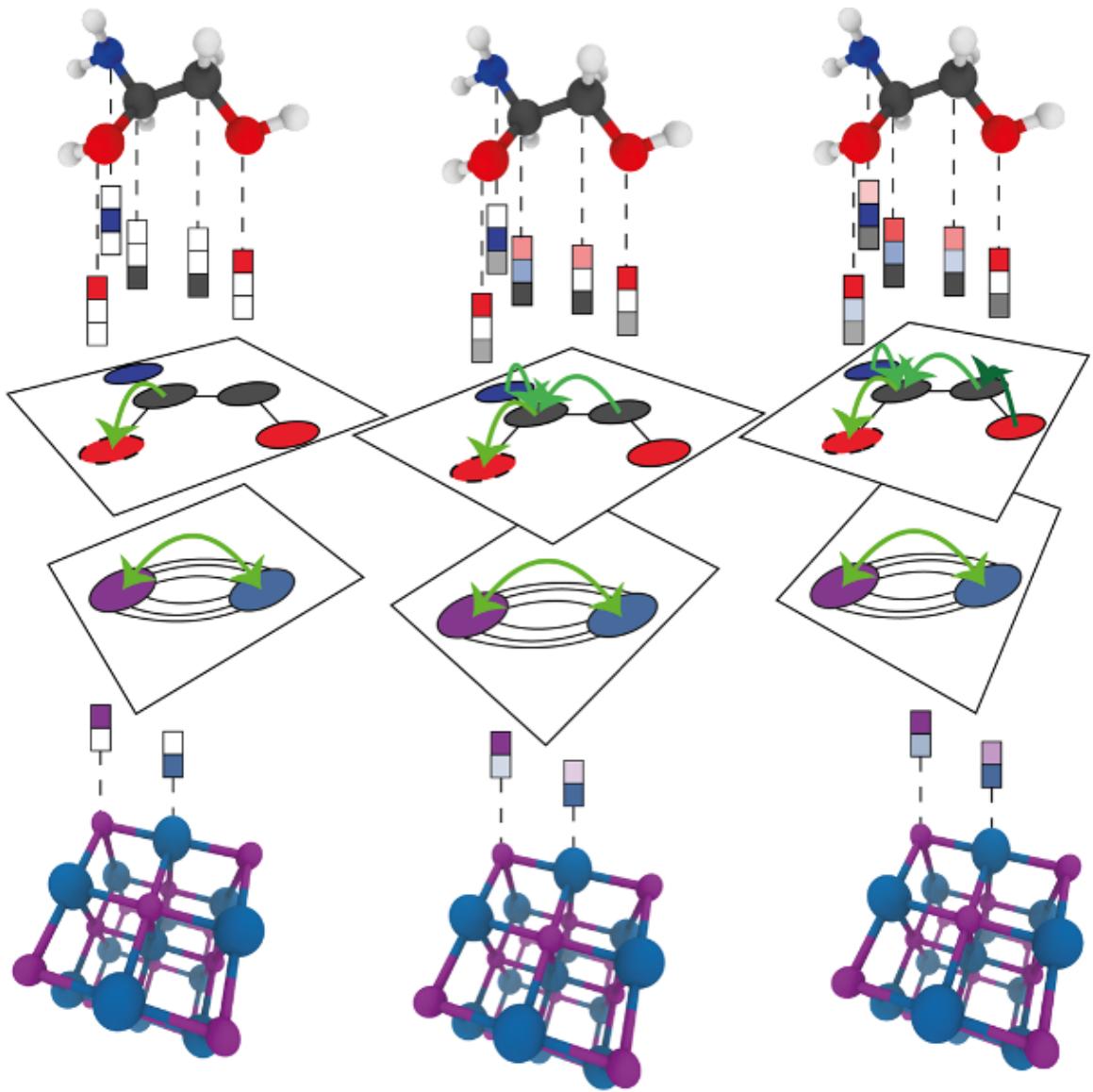
### GNN



$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

<https://untitledtblog.tistory.com/152>

Message passing GNN



Graph-level	Attributes	Description
nodes	atom-type	type of atoms (one-hot)
	chirality	R or S (one-hot or null)
	degree	number of covalent bonds (one-hot)
	radical	number of radical electrons (integer)
	hybridization	sp, sp <sup>2</sup> , sp <sup>3</sup> ... (one-hot)
	aromaticity	part of an aromatic system (binary)
edges	charge	formal charge (integer)
	bond-type	single, double, ... (one-hot)
	conjugation	is conjugated (binary)
	ring	bond is part of a ring (binary)
graph	stereo	None, Any, Z, E (one-hot)
	weight	average atomic weight (float)
	bonds	average bonds per atom (float)

Friederich Comm. Mat. 3, 93 (2022)

## CGCNN

---

PHYSICAL REVIEW LETTERS 120, 145301 (2018)

---

### Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties

Tian Xie and Jeffrey C. Grossman  
*Department of Materials Science and Engineering, Massachusetts Institute of Technology,  
Cambridge, Massachusetts 02139, USA*

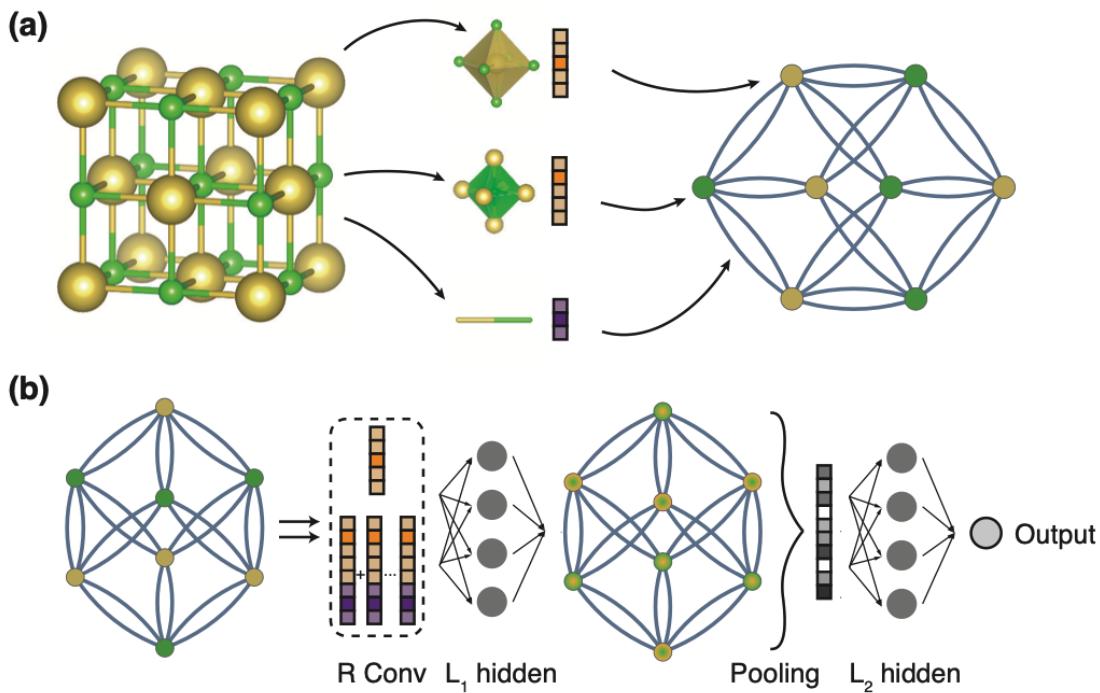


TABLE I. Summary of the prediction performance of seven different properties on test sets.

Property	# of train data	Unit	MAE <sub>model</sub>	MAE <sub>DFT</sub>
Formation energy	28046	eV/atom	0.039	0.081–0.136[18]
Absolute energy	28046	eV/atom	0.072	–
Band gap	16458	eV	0.388	0.6[23]
Fermi energy	28046	eV	0.363	–
Bulk moduli	2041	log(GPa)	0.054	0.050[24]
Shear moduli	2041	log(GPa)	0.087	0.069[24]
Poisson ratio	2041	–	0.030	–

<https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.120.145301>

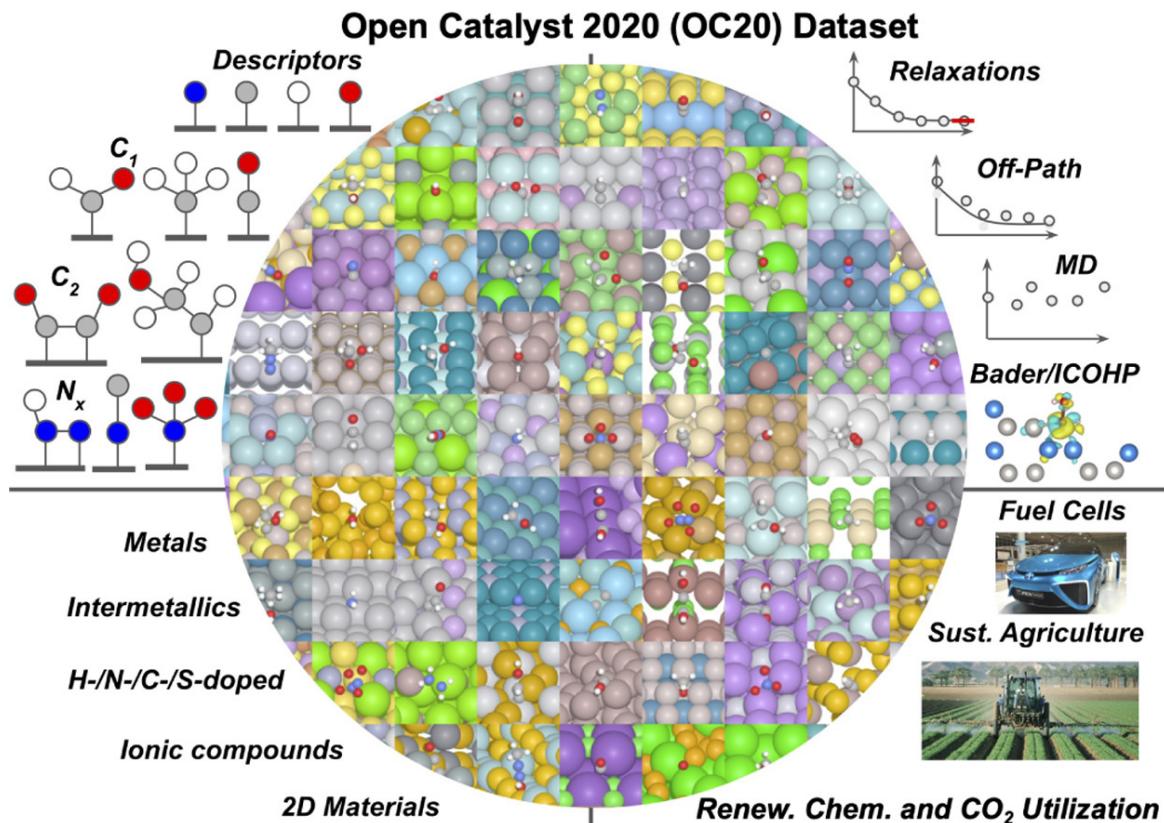
## Advent of large dataset

### Advent of Large dataset

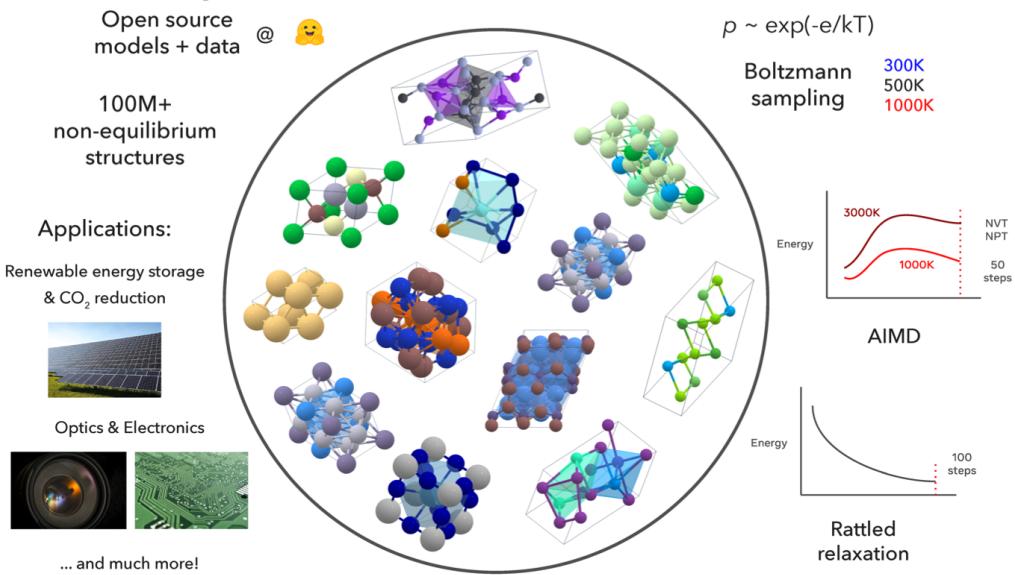
Ab initio datasets of inorganic materials in computational catalysis.

Dataset	Property	Dataset size	Level of theory	Chemical Composition	Accessibility and contribution
Materials Project [34]	Atomic structure, band structure, magnetic properties, and others	$1.5 \times 10^5$	PBE, HSE06, and r <sup>2</sup> SCAN	Inorganic Crystals and molecules	Open source, open contribution,
OQMD [35]	Atomic structure, and band structure band gap	$1 \times 10^6$	PBE	Inorganic Crystals	Open source
NOMAD [36]	Atomic structure, and band structure	$3 \times 10^6$	PBE, HSE06, HLE17, and others	Inorganic Crystals	Open source, open contribution
AFLOW [37]	Atomic structure, band structure, magnetic properties and phonon spectra	$3.5 \times 10^6$	PBE	Inorganic Crystals	Open source
OC20 [12]	Atomic structure, Bader charge, total energy, and force,	$2.5 \times 10^8$	RPBE	molecular adsorptions on inorganic crystal slabs	Open source, community challenges
OC22 [13]	Atomic structure, total energy, and force	$1 \times 10^7$	PBE	molecular adsorptions on metal oxide slabs	Open source, community challenges
OpenDAC [38]	Atomic structure, total energy and force	$3.8 \times 10^7$	PBE-D3	CO <sub>2</sub> and H <sub>2</sub> O adsorptions on MOFs	Open source, community challenges
Catalysis-Hub [39]	Atomic structure, total energy, and reaction energy	$1.3 \times 10^5$	PBE, RPBE, PW91 and others	Molecular adsorptions on inorganic crystals	Open source, open contribution,
PFP [40]	Atomic structure, and total energy	$9 \times 10^6$	PBE and PBE + U	Molecules, transition metal slabs, clusters	Partially open
GNoMe [29]	Atomic structure, total energy, and decomposition energy	$3.8 \times 10^5$	PBE and r <sup>2</sup> SCAN	Inorganic Crystal	Partially open

## Open Catalyst Project

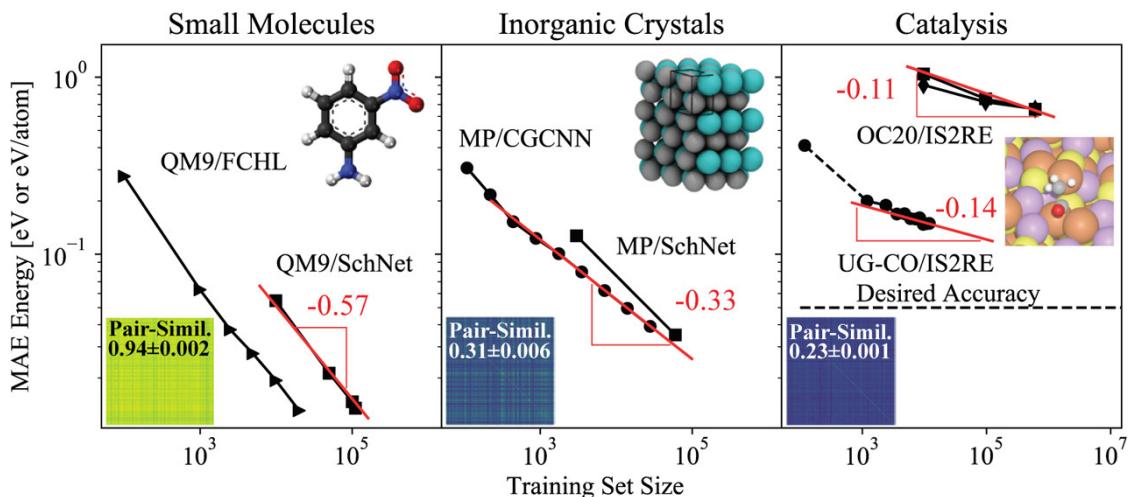


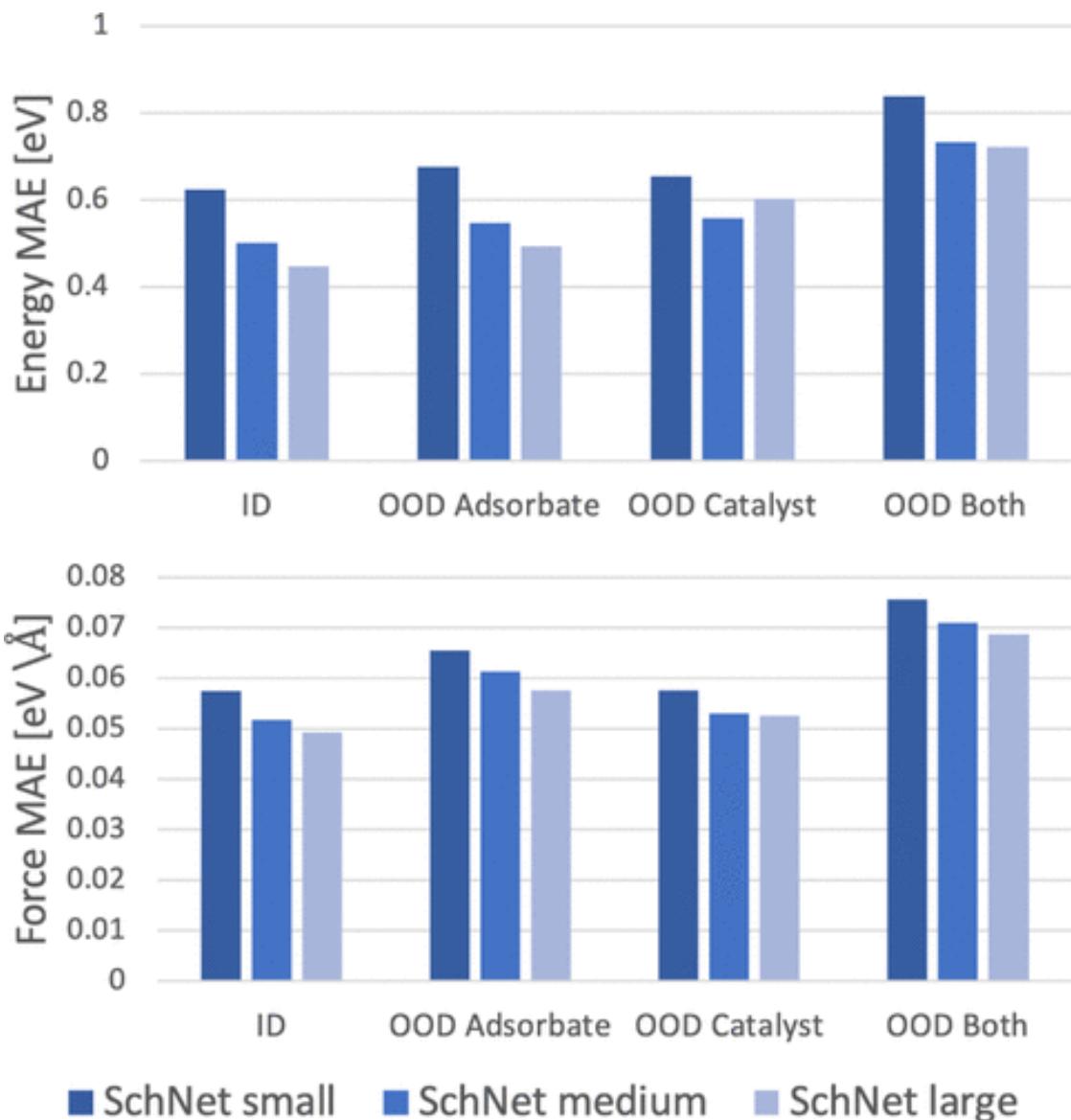
# Open Materials 2024 (OMat24)



<https://arxiv.org/pdf/2410.12771>

## ▼ Scale of training dataset in MLP





*ACS Catal.* 2021, 11, 10, 6059-6072

## Scale matters: GNoME model

### Article

# Scaling deep learning for materials discovery

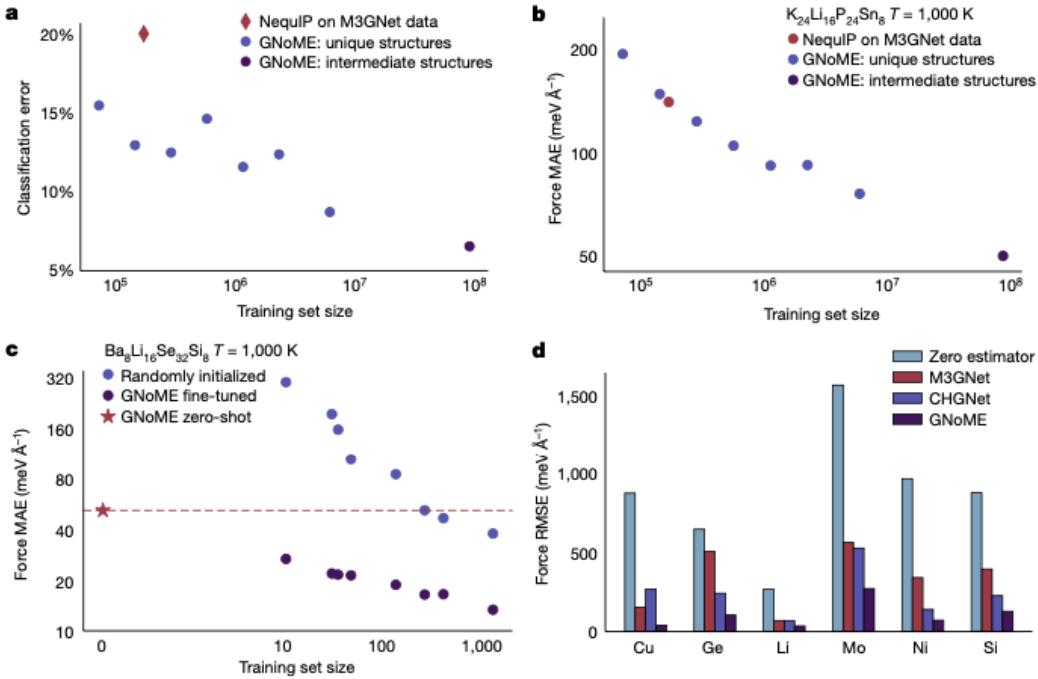
<https://doi.org/10.1038/s41586-023-06735-9>

Amil Merchant<sup>1,3</sup>, Simon Batzner<sup>1,3</sup>, Samuel S. Schoenholz<sup>1,3</sup>, Muratahan Aykol<sup>1</sup>,

Received: 8 May 2023

Gwoon Cheon<sup>2</sup> & Ekin Dogus Cubuk<sup>1,3</sup>

Advanced 10 October 2023



<https://www.nature.com/articles/s41586-023-06735-9>

## ▼ 'Universal' MLP

How researchers are making universal potential?

- Huge scale of database
- Rich informations in representation
- Computational efficiency in application
- Easy to use in ASE, LAMMPS and other packages

## ▼ Joint Multi-domain Pre-training (JMP)

pre-training strategy that leverages ~120M atomic systems from various datasets (OC20, OC22, ANI-1x, Transition-1x)

state-of-the-art on 34 out of 40 tasks

# FROM MOLECULES TO MATERIALS: PRE-TRAINING LARGE GENERALIZABLE MODELS FOR ATOMIC PROPERTY PREDICTION

Nima Shoghi<sup>\*1</sup> Adeesh Kolluru<sup>2</sup> John R. Kitchin<sup>2</sup>  
Zachary W. Ulissi<sup>1</sup> C. Lawrence Zitnick<sup>1</sup> Brandon M. Wood<sup>1</sup>

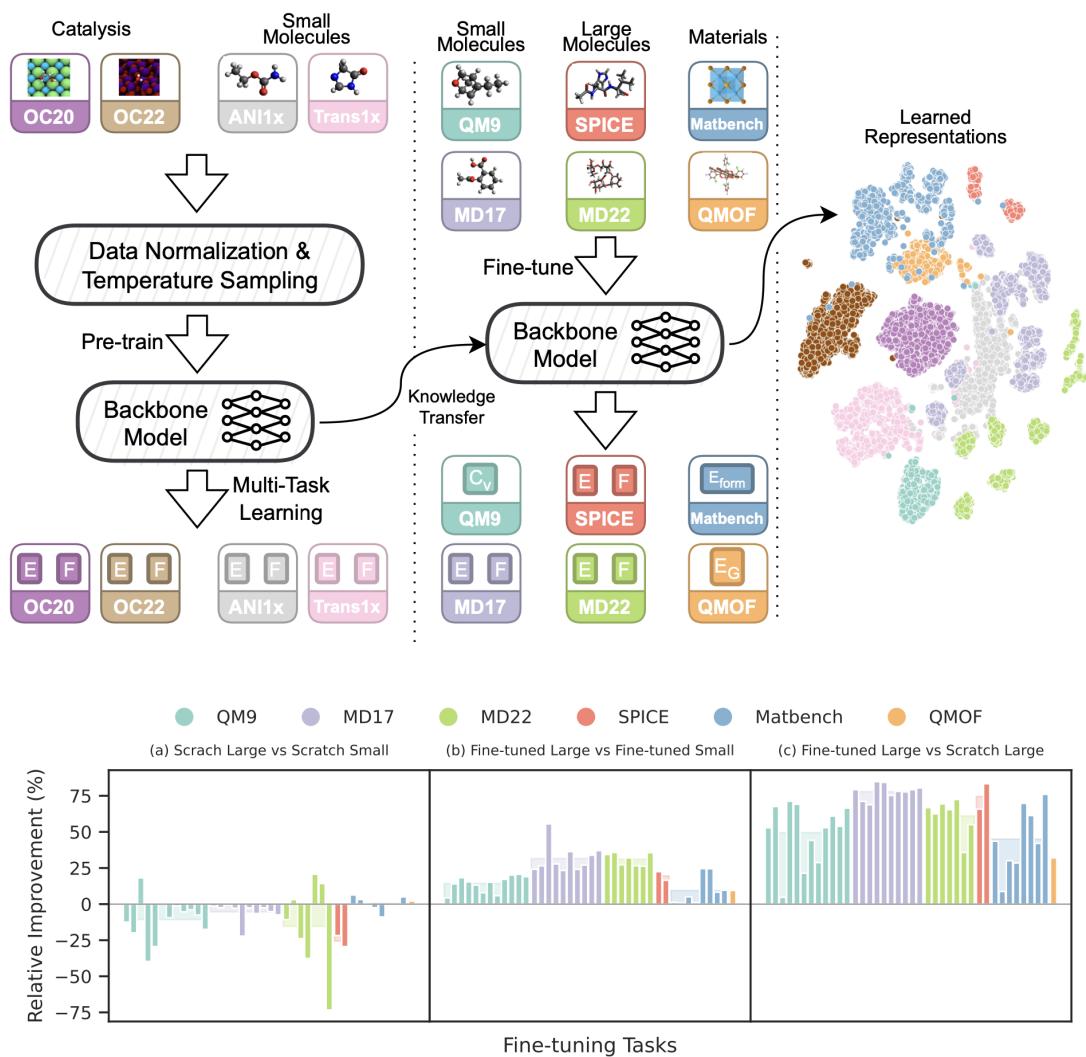
<sup>1</sup>Fundamental AI Research (FAIR) at Meta

<sup>2</sup>Carnegie Mellon University

\*Work done while at FAIR

Correspondence to: ns@nima.sh, bmwood@meta.com

## ABSTRACT

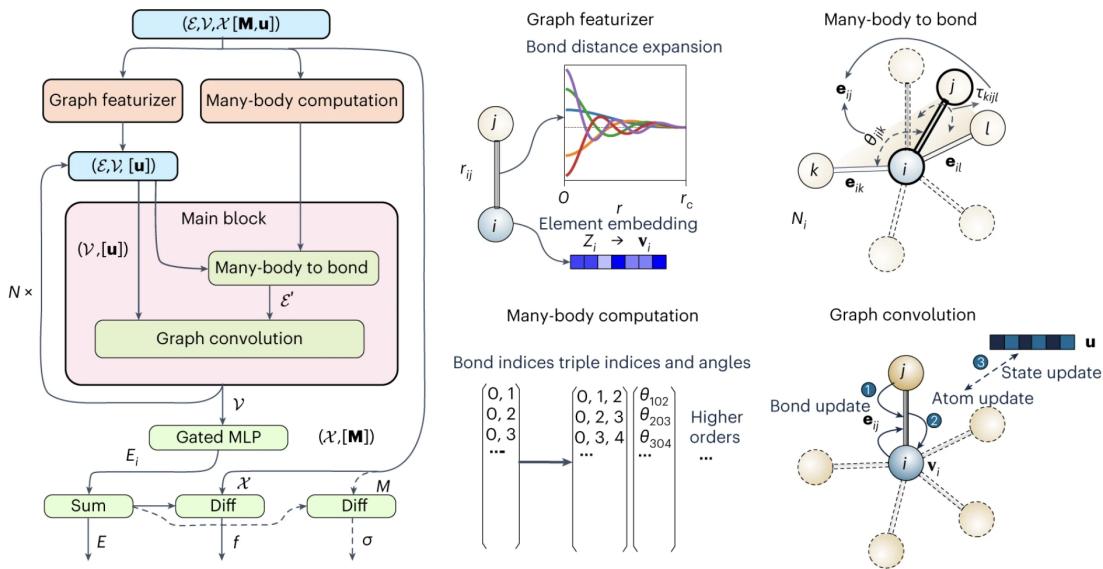


<https://arxiv.org/pdf/2310.16802.pdf>

## ▼ M3GNET: rich information in edge feature

three-body spherical harmonics features are used to update bond features

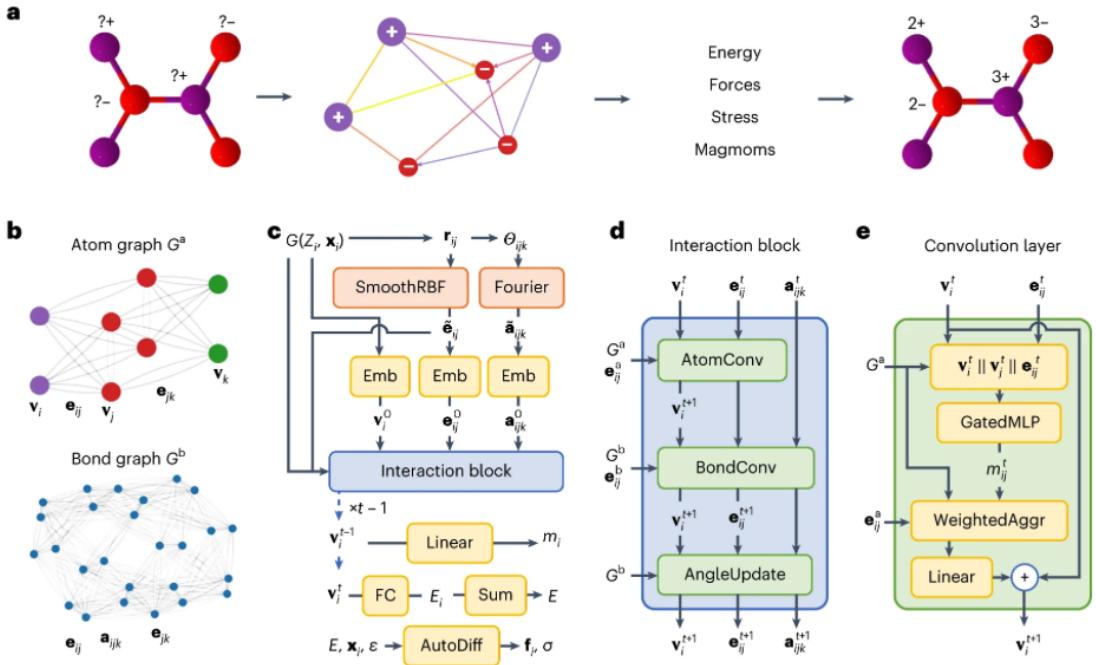
- **Angle-Based Interactions:** angles between atoms and neighbors.
- **Smooth Basis Encoding:** Uses spherical Bessel and harmonics functions to represent distances and angles smoothly.
- **Enhanced Bond Updates:** Bond features are updated with three-body terms for capturing complex atomic interactions.



<https://www.nature.com/articles/s43588-022-00349-3>

## ▼ CHGNet: rich information in node feature

CHGNet incorporates **magnetic moments as node features** in its graph neural network, allowing each node (representing an atom) to encode charge information.



```

from ase.io.trajectory import Trajectory
from pymatgen.io.ase import AseAtomsAdaptor
from chgnet.utils import solve_charge_by_mag

traj = Trajectory("md_out.traj")
mag = traj[-1].get_magnetic_moments()

# get the non-charge-decorated structure
structure = AseAtomsAdaptor.get_structure(traj[-1])
print(structure)

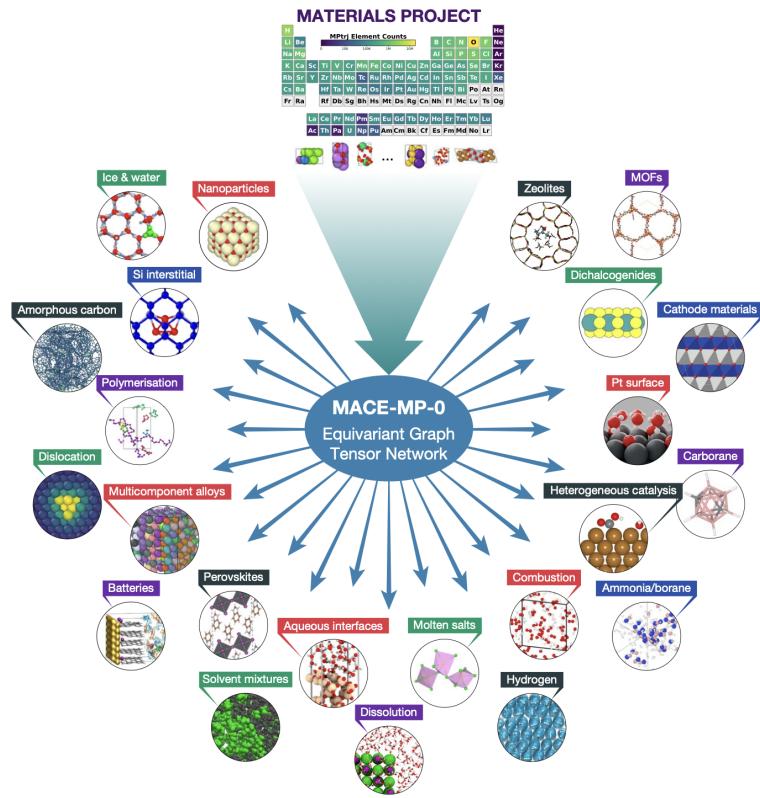
# get the charge-decorated structure
struct_with_chg = solve_charge_by_mag(structure)
print(struct_with_chg)

```

<https://github.com/CederGroupHub/chgnet>

*Nature Machine Intelligence* volume 5, pages 1031–1041 (2023)

## ▼ MACE-MP: rich higher order information efficiently



<https://arxiv.org/pdf/2401.00096>

---

## MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields

---

**Ilyes Batatia**  
 Engineering Laboratory,  
 University of Cambridge  
 Cambridge, CB2 1PZ UK  
 Department of Chemistry,  
 ENS Paris-Saclay, Université Paris-Saclay  
 91190 Gif-sur-Yvette, France  
 ilyes.batatia@ens-paris-saclay.fr

**Gregor N. C. Simm**  
 Engineering Laboratory,  
 University of Cambridge  
 Cambridge, CB2 1PZ UK

**Dávid Péter Kovács**  
 Engineering Laboratory,  
 University of Cambridge  
 Cambridge, CB2 1PZ UK

**Christoph Ortner**  
 Department of Mathematics  
 University of British Columbia  
 Vancouver, BC, Canada V6T 1Z2

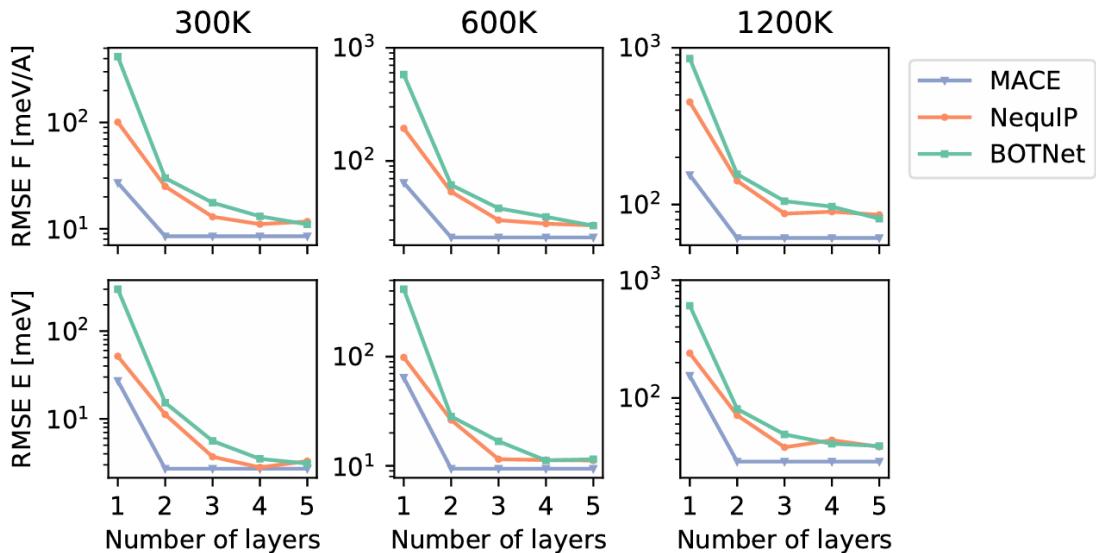
**Gábor Csányi**  
 Engineering Laboratory,  
 University of Cambridge  
 Cambridge, CB2 1PZ UK

<https://doi.org/10.48550/arXiv.2206.07697>

## Characteristics of MACE

- MPNN Interatomic Potentials
- Equivariant Graph Neural Networks

## Performance of MACE



- Uses four-body interactions, capturing complex atomic relationships in just two layers.
- **Equivariant Design:** Maintains rotational and reflection symmetry, enhancing accuracy.

**State-of-the-Art Accuracy:** Outperforms other models in accuracy on benchmarks like rMD17 and 3BPA.

## Install MACE

### Installation from PyPI

This is the recommended way to install MACE.

```
pip install --upgrade pip  
pip install mace-torch
```



- MACE-MP : **Materials Project Force Fields, trained on bulk dataset**

#### Example usage in ASE

```
from mace.calculators import mace_mp
from ase import build

atoms = build.molecule('H2O')
calc = mace_mp(model="medium", dispersion=False, default_dtype="float32", device='cuda')
atoms.calc = calc
print(atoms.get_potential_energy())
```

- MACE-OFF : Transferable Organic Force Fields, trained on organic molecule dataset

#### Example usage in ASE

```
from mace.calculators import mace_off
from ase import build

atoms = build.molecule('H2O')
calc = mace_off(model="medium", device='cuda')
atoms.calc = calc
print(atoms.get_potential_energy())
```

## ▼ SevenNet: Scale up efficiently

efficient parallelization based on NequIP architecture

*large-scale MD simulations, offering researchers a powerful tool to explore complex material systems with high accuracy and efficiency.*

***J. Chem. Theory Comput. 2024, 20, 11, 4857-4868.***

**SevenNet Calculator for ASE**

ASE (Atomic Simulation Environment) is a set of tools and Python modules for atomistic simulations. SevenNet-0 and SevenNet-trained potentials can be used with ASE for its use in python.

For pre-trained models,

```
from sevenn.sevennet_calculator import SevenNetCalculator
sevennet_0_cal = SevenNetCalculator("7net-0", device='cpu') # 7net-0, SevenNet-0, 7net-0_22
```

For user trained models,

```
from sevenn.sevennet_calculator import SevenNetCalculator
checkpoint_path = ### PATH TO CHECKPOINT ####
sevennet_cal = SevenNetCalculator(checkpoint_path, device='cpu')
```

After the PyTorch installation, run

```
pip install sevenn
```

To download the latest version of SevenNet(not stable!), run

```
pip install https://github.com/MDIL-SNU/SevenNet.git
```

<https://github.com/MDIL-SNU/SevenNet>

## ▼ Hands-on

# Task 01 : Revisiting Handson 2,3,4

## Handson 2

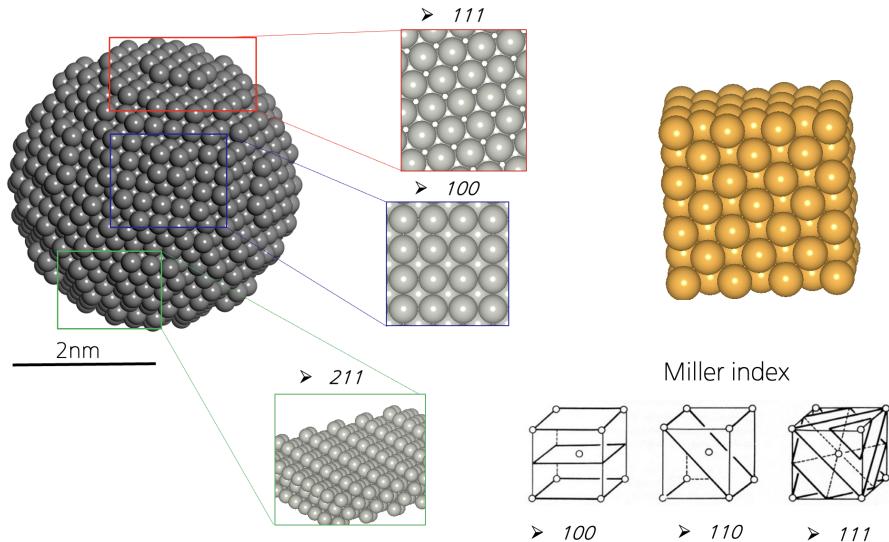
### 3. Testing different surface facets (fcc100, fcc110, fcc111):

```
In [12]: import time
import numpy as np
from ase.build import fcc100, fcc110, fcc111
from ase.constraints import FixAtoms
from gpaw import GPAW, PW

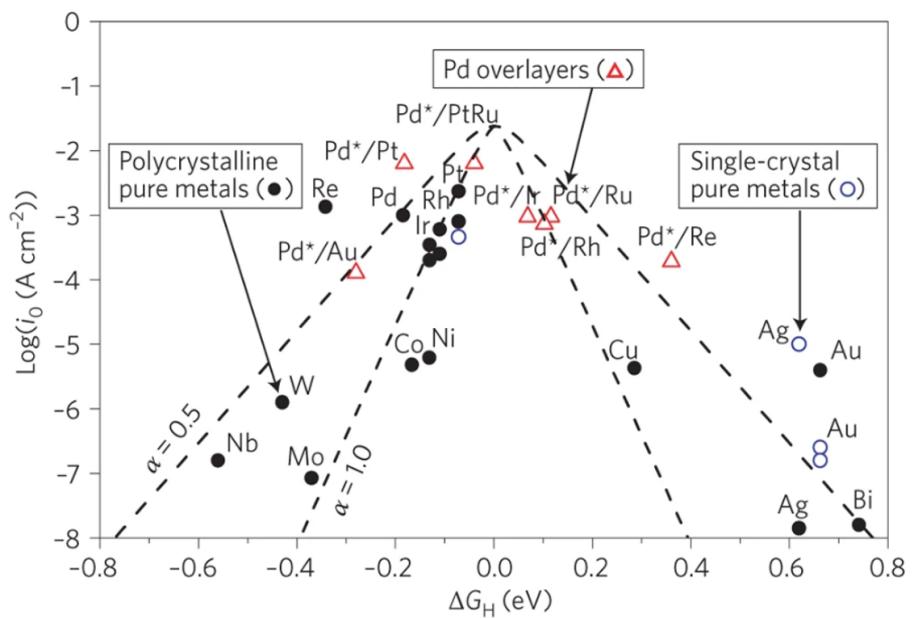
# Parameters
a0 = 4 # lattice constant
k = 4 # KPOINTS
N_layer = 4 # number of layers
size = 1 # supercell size
vac = 7.5 # vacuum size

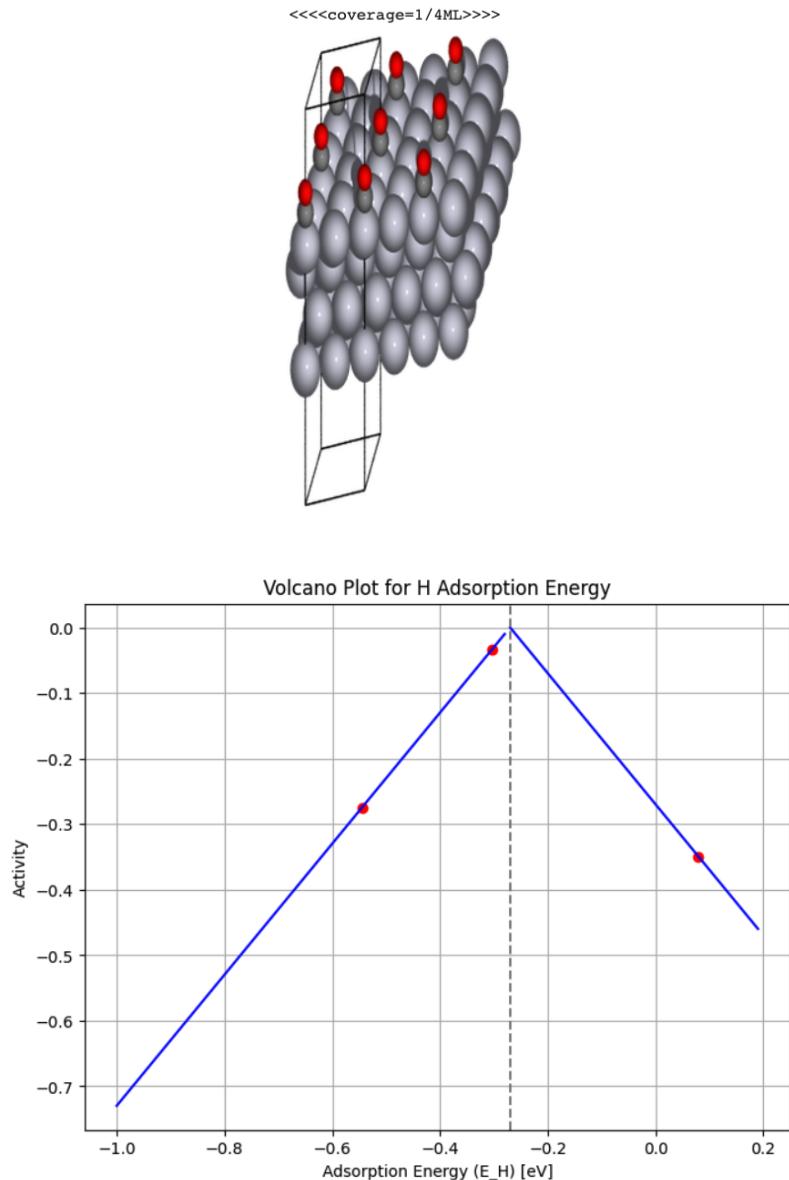
# List of FCC surfaces to test
fcc_surfaces = [('fcc100', fcc100), ('fcc110', fcc110), ('fcc111', fcc111)]

# Iterate over different FCC orientations
for surface_name, surface_function in fcc_surfaces:
    surf = surface_function('Al', (size, size, N_layer), a=a0, vacuum=vac)
    # Calculate the median z-position
```

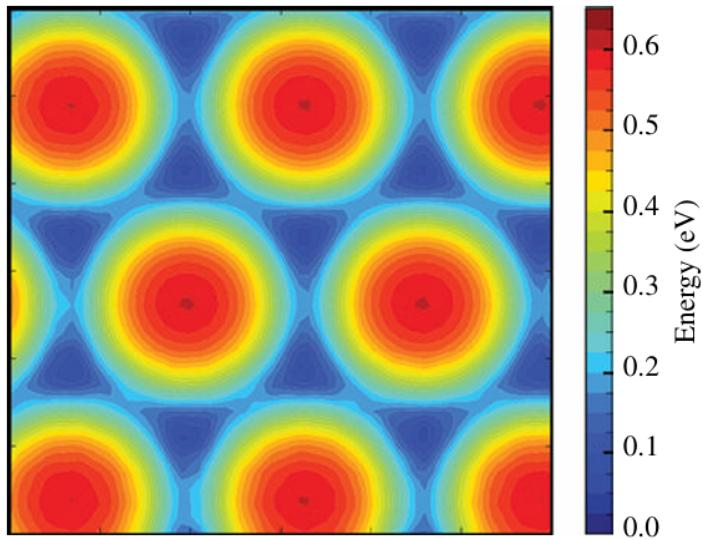


## Handson 3 and 4





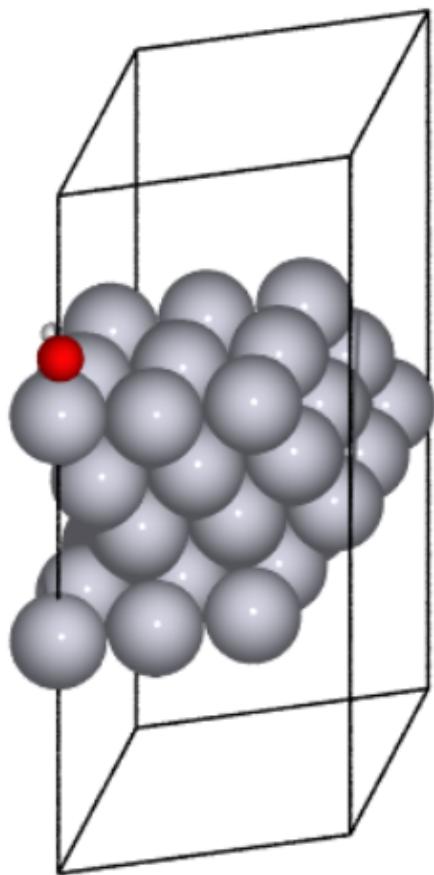
## Task 02 : PES of H diffustion on metal surface



**FIGURE 2.7** PES of H diffusion on Cu(111). The potential energy is plotted over part of the Cu surface area. H adsorbed in the threefold position has been chosen as the reference.

Let's utilize the fast computation speed of MLP to visualize the PES of H diffusion on the Pt(111) surface!

## Task 03 : Finding a optimal adsorption configuration



Let's use a grid search to identify the most stable adsorption site and geometric adsorption structure for OH on the Pt surface.

## Assignment

- Reproduce Hand-on 2,3,4 using MACE-MP pretrained model and compare the result with DFT calculations (using GPAW).
- Plot PES of any transition metal surface using MACE-MP model.
- Find optimal OH biding sites and binding energies for Pt, Pd, Au and Ni.