



Science and
Technology
Facilities Council

Scientific Computing

Machine Learning Interatomic Potentials Meet Applications

Alin M Elena

computational materials and molecular sciences



Data-driven Materials
and Molecular Science

July 4, 2025



CCP5
Collaborative
Computational Project
Computer Simulation of Condensed Phases

CoSeC

Computational Science Centre
for Research Communities



PSDI
PHYSICAL SCIENCES
DATA INFRASTRUCTURE



Science and
Technology
Facilities Council

Ada Lovelace Centre



Royal Observatory
Edinburgh UK ATC

Boulby Underground
Laboratory

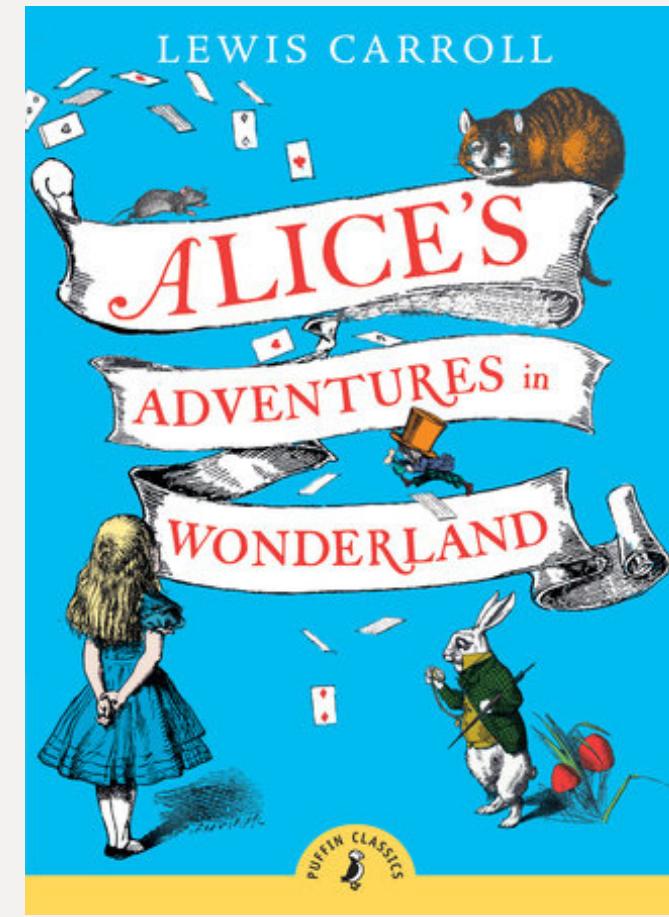
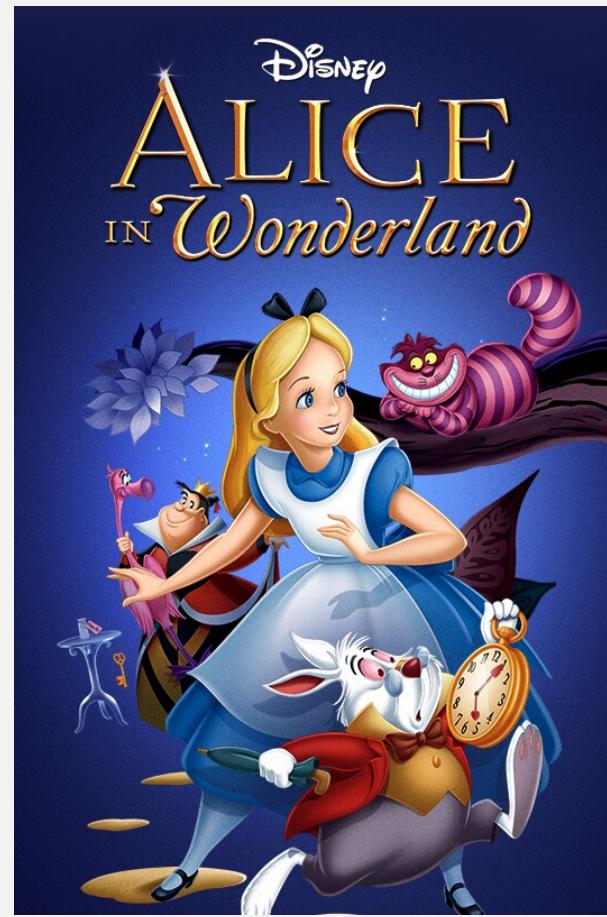
Rutherford Appleton
Laboratory

Daresbury
Laboratory

Polaris House
Swindon

Chilbolton
Observatory





agenda



introduction

Theoretical era

Simulations era

ML era

metal organic frameworks

2,5-Diiodothiophene (pressure)

zeolites

performance





introduction

Theoretical era

Simulations era

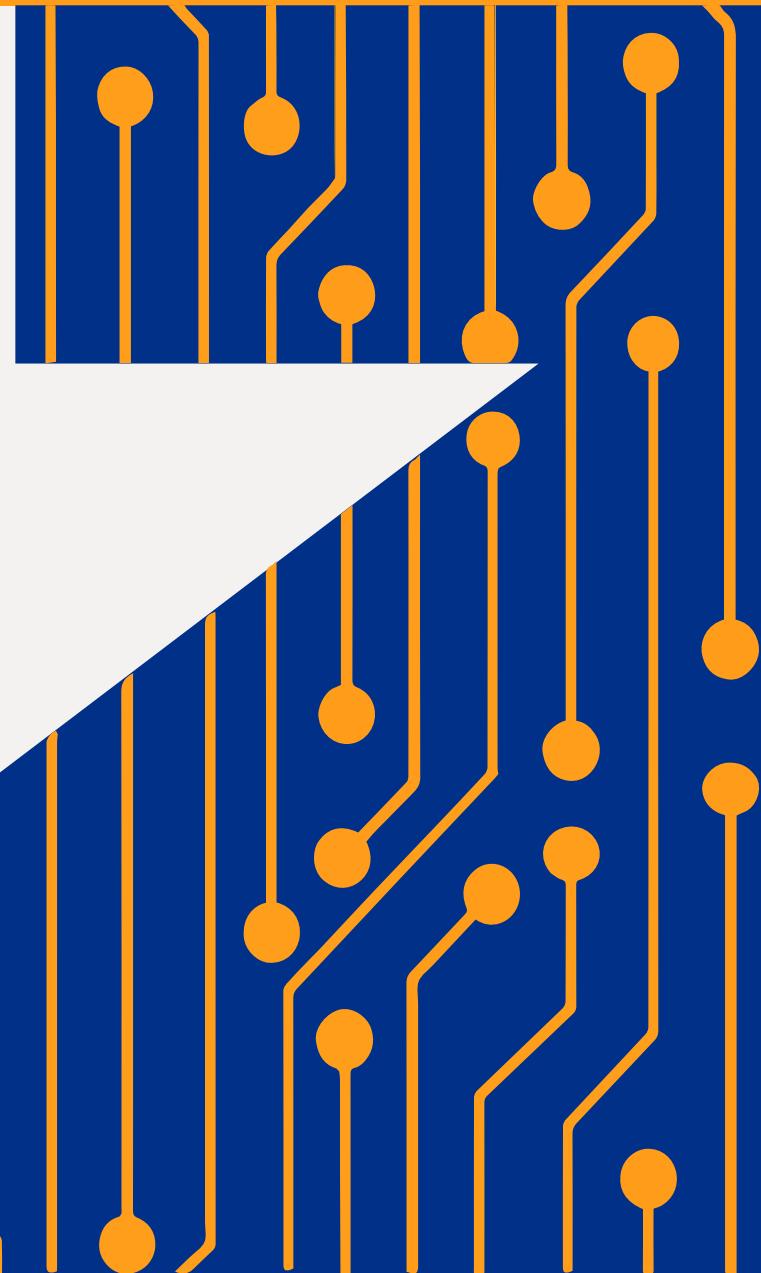
ML era

metal organic frameworks

2,5-Diiodothiophene (pressure)

zeolites

performance





acknowledgments

Prathami Kamath, Andrew Rosen, Théo J Inizan and
Kristin Persson

Jamal Abdul Nasir, Alexey Sokol, and Richard Catlow UCL

Jeff Armstrong and Adam Jackson

Ilyes Batatia and Gábor Csányi

Elliott Kasoar, Jacob Wilkins

UC Berkeley



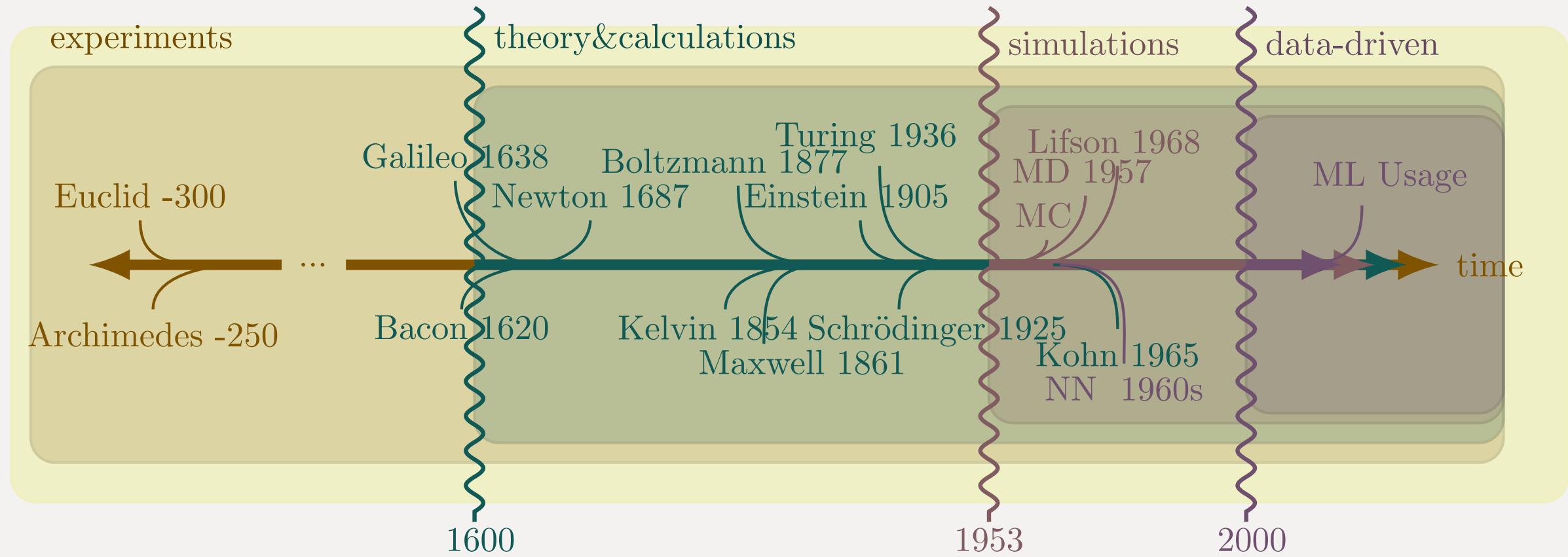


Disclaimer





Science Paradigms





or the Wolfram version



structural (antiquity)

e.g. geometrical elements

explicit time
not considered

static facts deduced
by reasoning

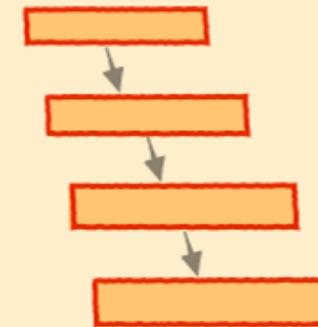


mathematical (1600s)

e.g. differential equations

time as mathematical
coordinate

find behavior at any
time from formula



computational (1980s)

e.g. cellular automata

time as progress
of computation

determine future only
by running program



multicomputational (2020s)

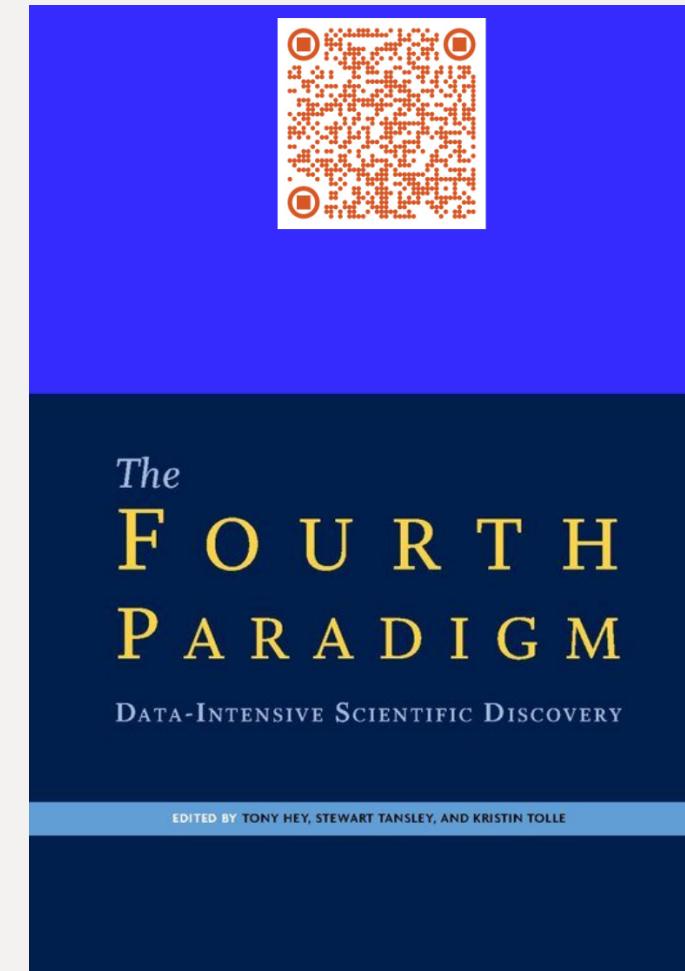
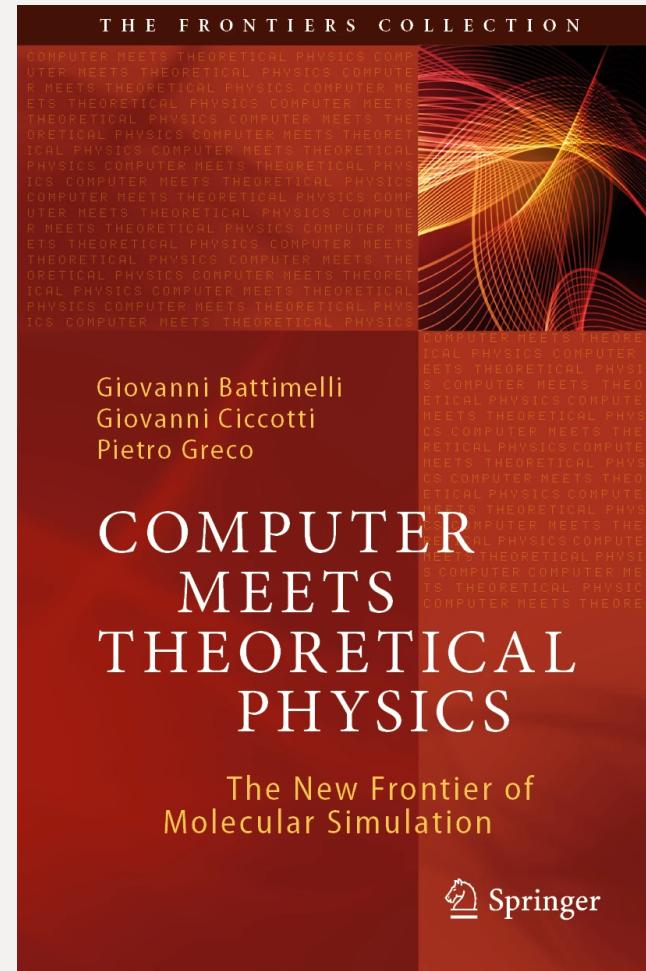
e.g. multiway systems

many computational
threads of time

need model of observer
to determine state



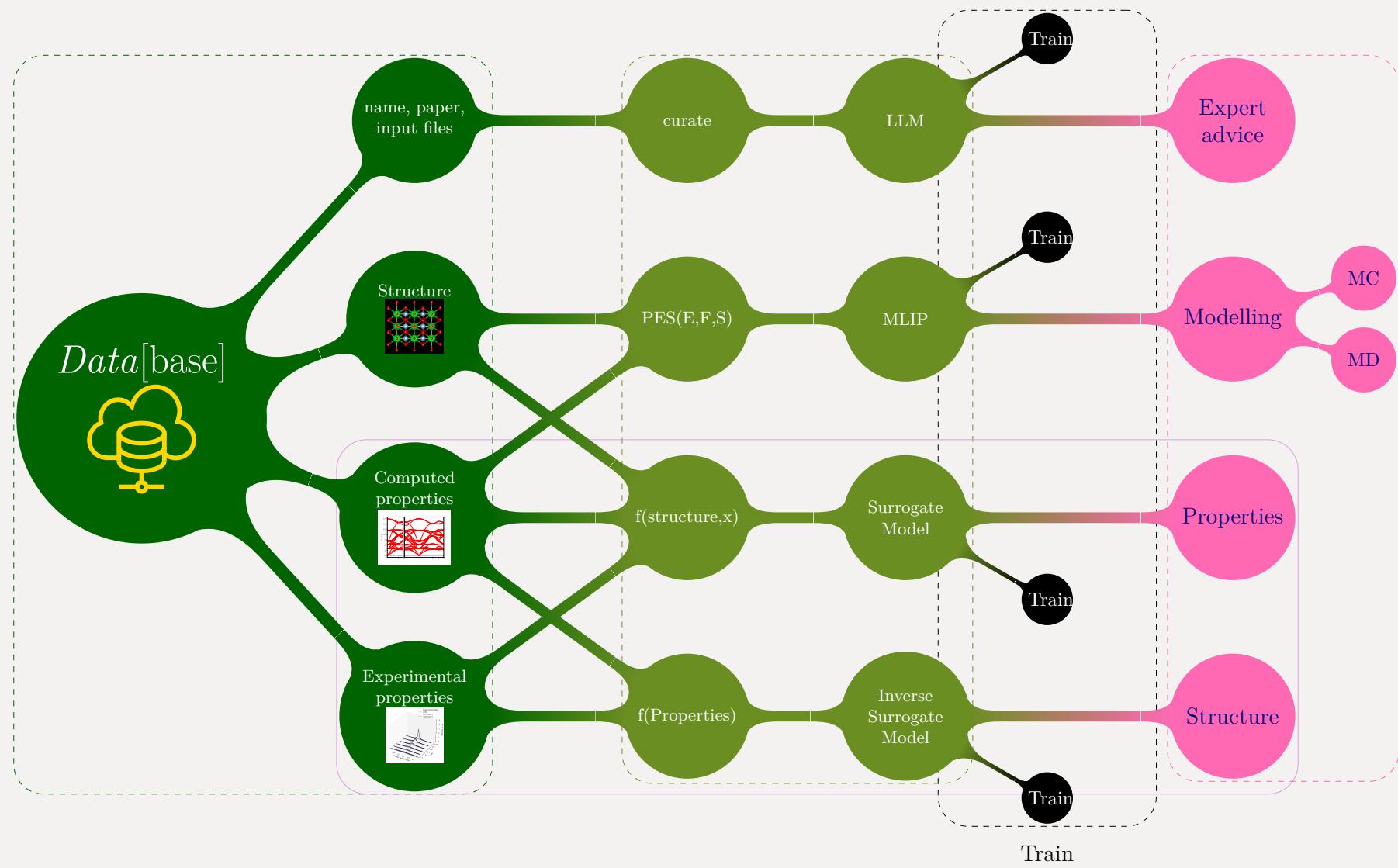
history



[qr to Jim Gray's original presentation](#)

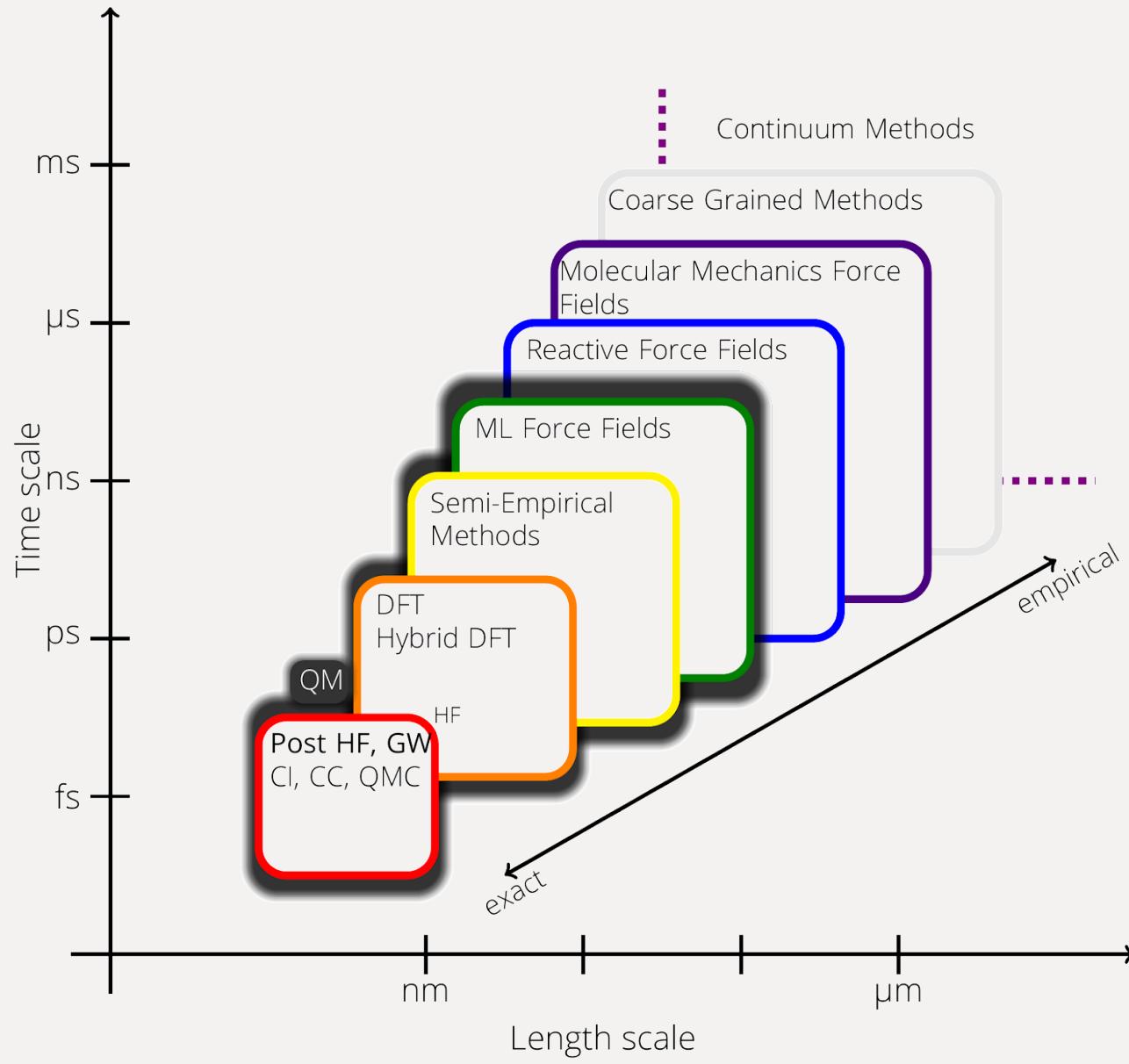


AI & materials modelling





the big picture





introduction

Theoretical era

Simulations era

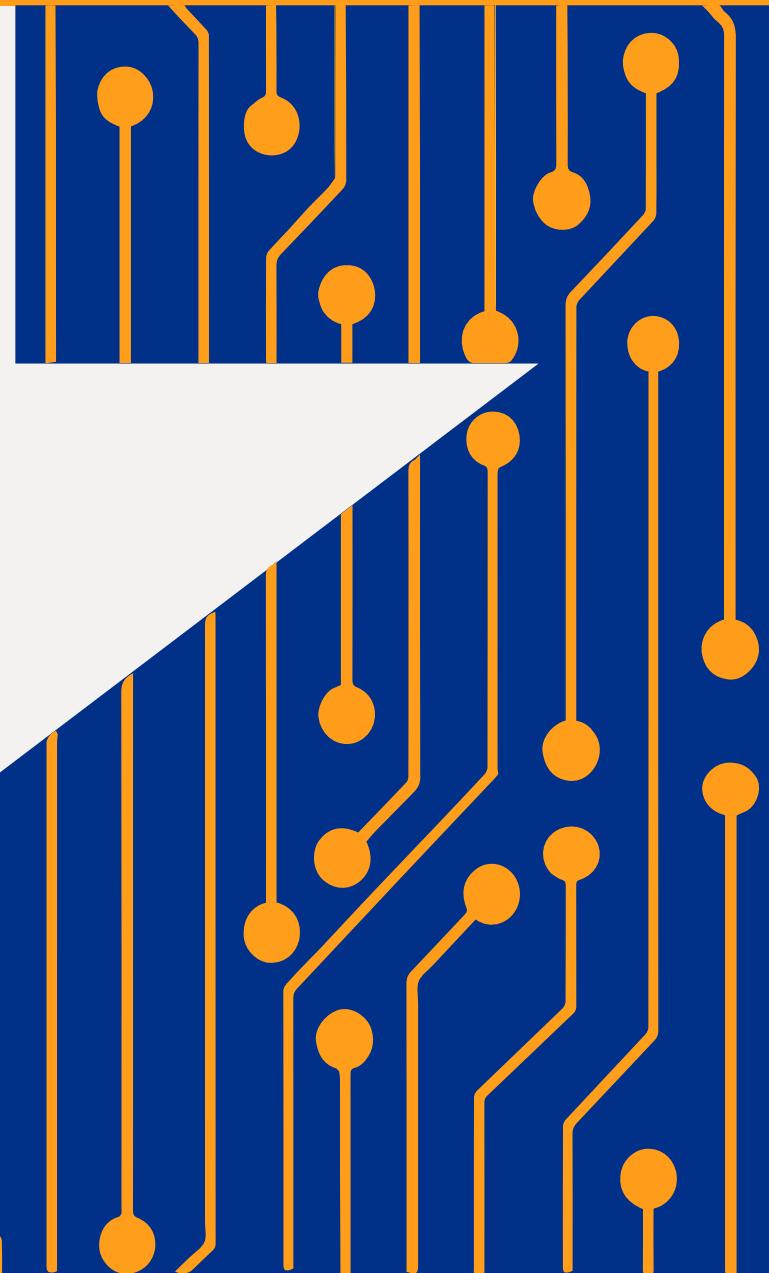
ML era

metal organic frameworks

2,5-Diiodothiophene (pressure)

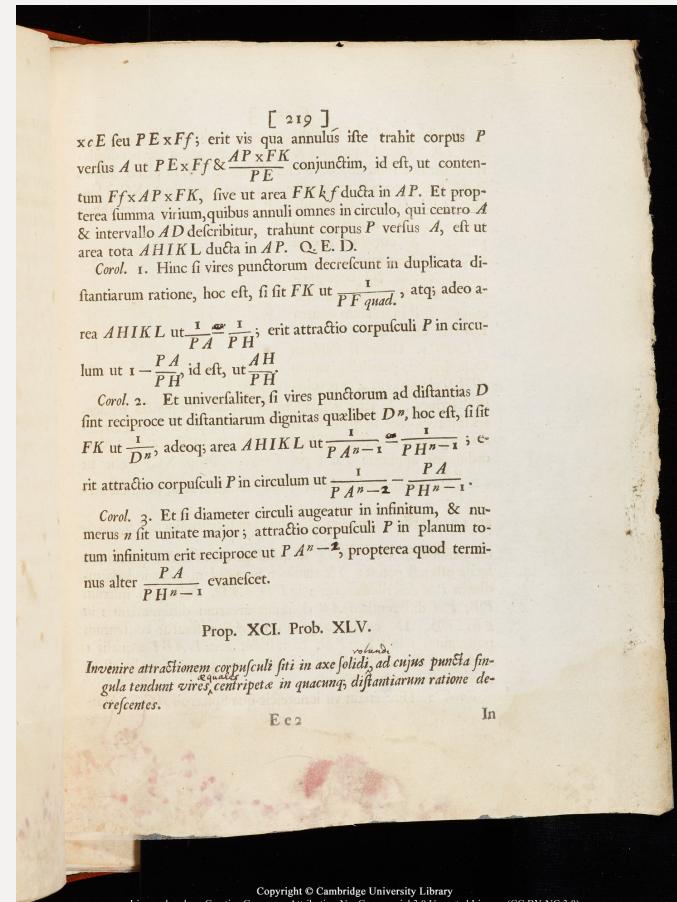
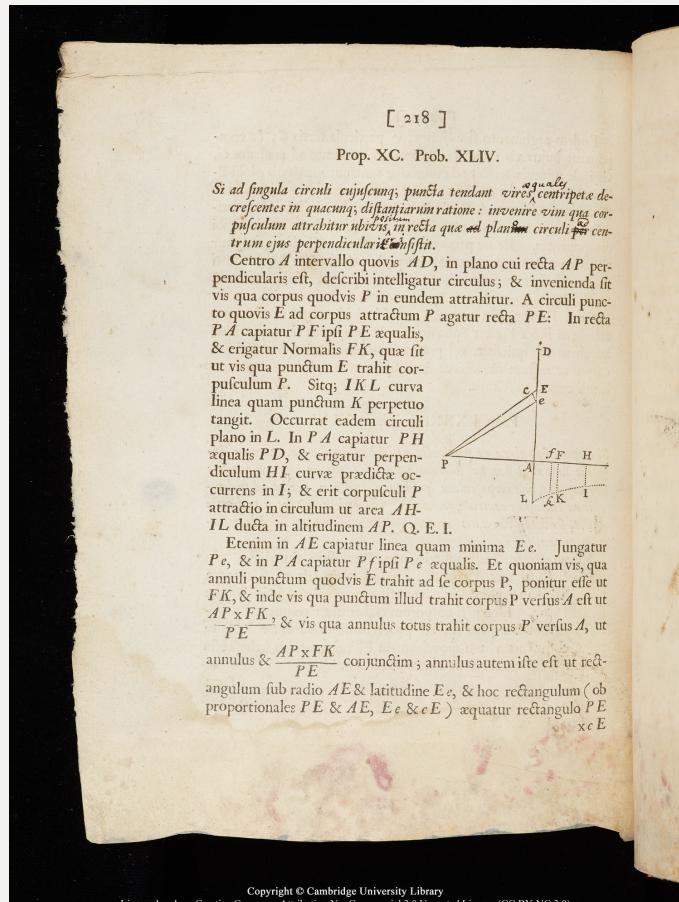
zeolites

performance





Newton: Philosophiæ naturalis principia mathematica 1687



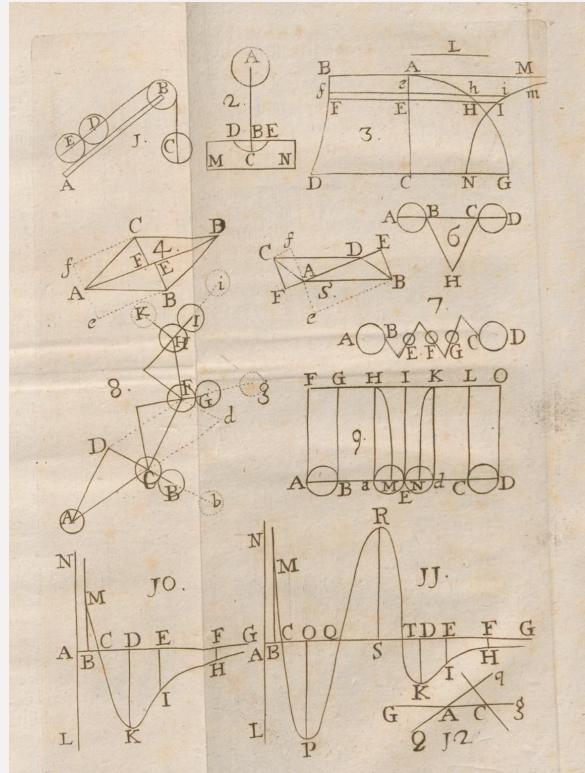
$$\text{attraction (corpuscles): } W = \frac{1}{D^n}$$



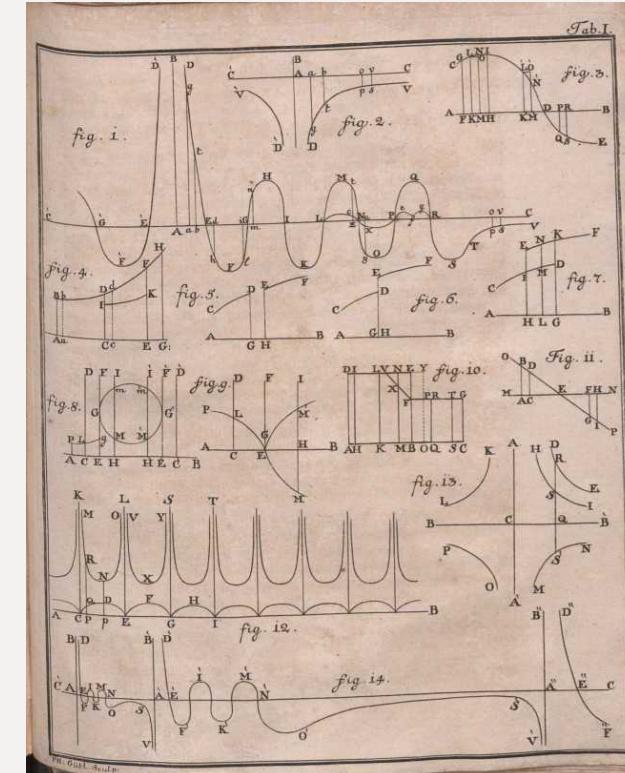


Rogerio Boscovich

De viribus vivis dissertatio 1745



Philosophiae naturalis theoria 1758



- ▶ attraction and repulsion
- ▶ Boscovich's demon, these days Laplace's demon is more common



Capillarity & gasses 19th century

- ▶ Clairault - Theorie de la Figure de la Terre - 1743; capillarity,
- ▶ Laplace - 1805 Book X, capillarity again
- ▶ Gauss - 1830 again capillarity, mutual attractive force between molecules and attractive between molecules in liquid and points in the wall

$$-\frac{a}{r^n} \quad n > 5$$

- ▶ Maxwell - theory of gases attractive force $n = 5$
- ▶ Boltzman - $n = 4$
- ▶ Sutherland - $n = 3$



Mie, Van der Waals, Grüneisen

- Mie 1903 - Zur kinetischen Theorie der einatomigen Körper

$$(10) \quad \Phi = \Phi_1 - \Phi_2 .$$

$$n \cdot \frac{1}{2} \cdot \int_0^{\infty} \Phi_2(r) \cdot 4\pi r^2 dr = A, \quad \Phi_1(r) = \frac{a}{r^{\nu}} .$$

$$-\Pi(f) = \frac{Af}{f} - \frac{f}{1} \dots \dots \dots (1) .$$

- Van der Waals 1908 - Lehrbuch der Thermodynamik

$$- \frac{a}{r^x} + \frac{b}{r^y} ,$$

- Grüneisen 1912 - Theorie des festen Zustandes einatomiger Elemente

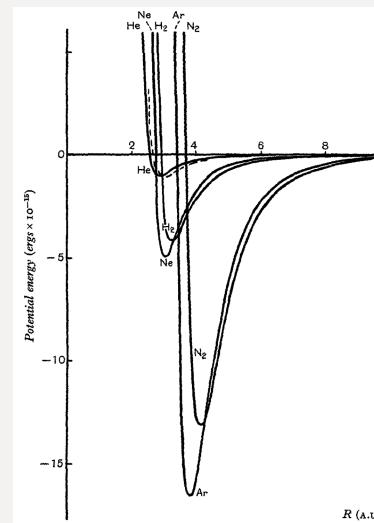


Lennard-Jones

- On the Determination of Molecular Fields II. From the Equation of State of a Gas (1924)

$$f(r) = \frac{\lambda_n}{r^n} - \frac{\lambda_m}{r^m}, \quad (2.05)$$

- Cohesion (1931)





► 1933 Bernal, Fowler

Now n cannot be very different from its value

for the rare gases¹⁹ ≈ 12 we have therefore

► 2020 Wang, Frenkel

$$\begin{aligned}\phi(r) &\equiv \varepsilon\alpha(1, 1; r_c) \left[\left(\frac{\sigma}{r}\right)^2 - 1 \right] \left[\left(\frac{r_c}{r}\right)^2 - 1 \right]^2 && \text{for } r \leq r_c \\ &= 0 && \text{for } r > r_c\end{aligned}\quad (3)$$



beyond 12-6

1932 Born Mayer

Dies kommt daher, daß London die Abstoßungskräfte vernachlässigt hat. Das gesamte Potential $\Phi(r)$ eines Molekülkristalls wird die Form haben:

$$\Phi(r) = -\frac{C}{r^6} + b e^{-\frac{r}{\rho}}. \quad (26)$$

$$\Phi(r) = A(r) + B(r), \quad (2)$$

in which $A(r)$, the attractive potential, and $B(r)$, the repulsive potential, are given by the following equations:

$$A(r) = -\alpha e^2/r - C/r^6 - D/r^8, \quad (3)$$

$$B(r) = B_{+-}(r) + [B_{--}(r) + B_{++}(r)]/2, \quad (4)$$

1933 Huggins Mayer

- A. Neon: I. $E(r) = be^{-r/\rho} - \mu r^{-6}$ (equation (26)).
II. $E(r) = \lambda r^{-12} - \mu r^{-6}$ (constants from Table II).
- B. Argon: I. $E(r) = be^{-r/\rho} - \mu r^{-6}$ (equation (27)).
II. $E(r) = \lambda r^{-13.5} - \mu r^{-6}$ (constants interpolated from Table II).

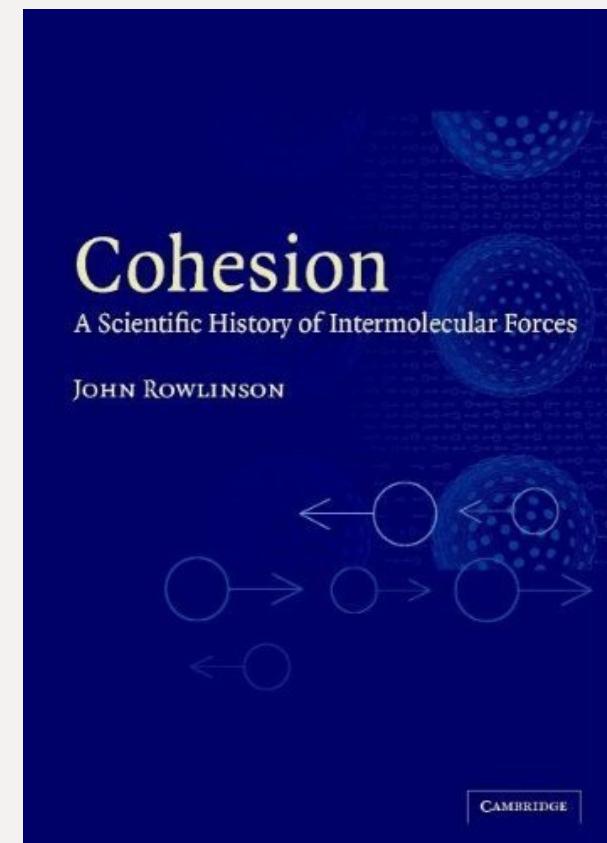
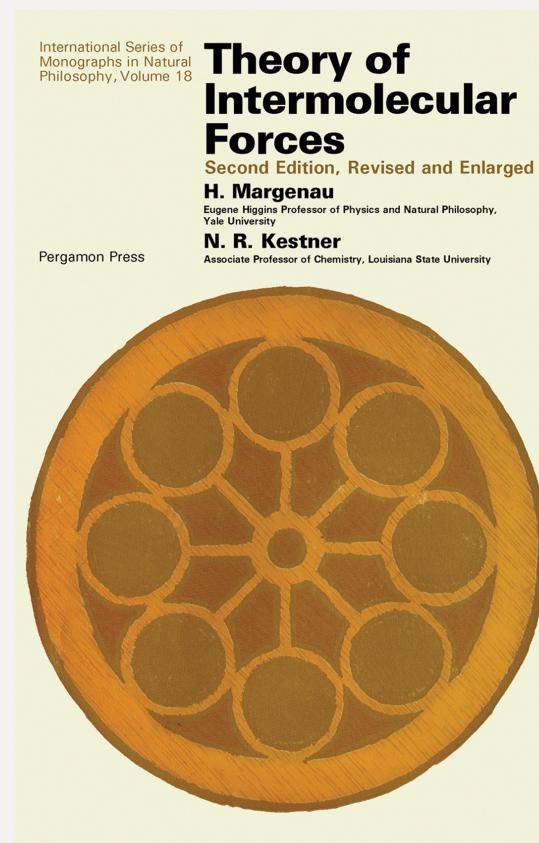
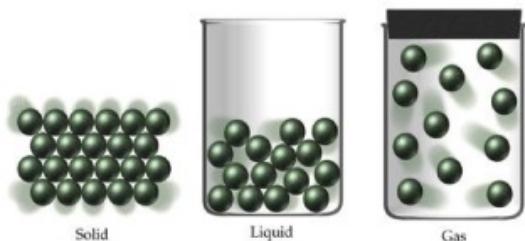
1938 Buckingham



readings

Molecular theory of gases and liquids

Joseph O. Hirschfelder,
Charles F. Curtiss
and R. Byron Bird





introduction

Theoretical era

Simulations era

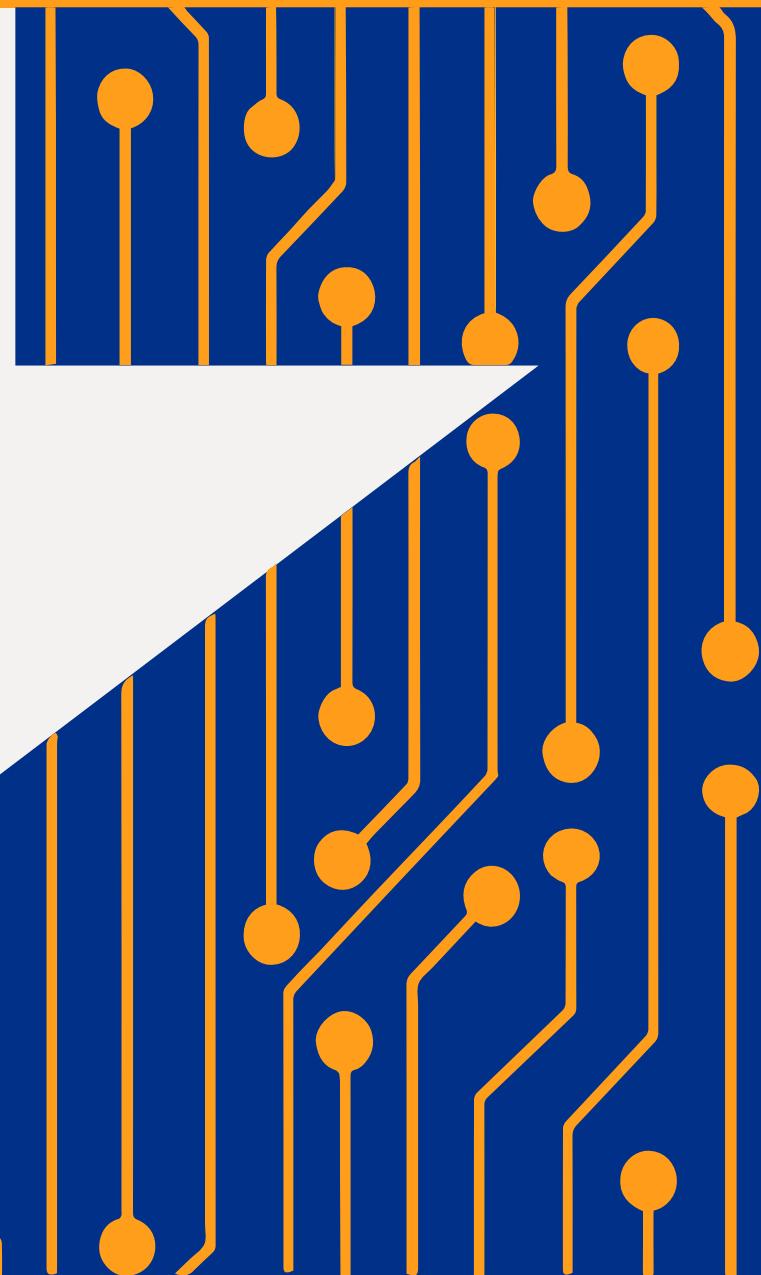
ML era

metal organic frameworks

2,5-Diiodothiophene (pressure)

zeolites

performance





seminal papers

THE JOURNAL OF CHEMICAL PHYSICS

VOLUME 21, NUMBER 6

JUNE, 1953

Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

AND

EDWARD TELLER,* *Department of Physics, University of Chicago, Chicago, Illinois*
 (Received March 6, 1953)

Phase Transition for a Hard Sphere System

B. J. ALDER AND T. E. WAINWRIGHT

University of California Radiation Laboratory, Livermore, California
 (Received August 12, 1957)

THE JOURNAL OF CHEMICAL PHYSICS

VOLUME 31, NUMBER 2

AUGUST, 1959

Studies in Molecular Dynamics. I. General Method*

B. J. ALDER AND T. E. WAINWRIGHT

Lawrence Radiation Laboratory, University of California, Livermore, California
 (Received February 19, 1959)





PHYSICAL REVIEW

VOLUME 120, NUMBER 4

NOVEMBER 15, 1960

Dynamics of Radiation Damage*

J. B. GIBSON, A. N. GOLAND,[†] M. MILGRAM, AND G. H. VINEYARD

Brookhaven National Laboratory, Upton, New York

PHYSICAL REVIEW

VOLUME 136, NUMBER 2A

19 OCTOBER 1964

Correlations in the Motion of Atoms in Liquid Argon*

A. RAHMAN

Argonne National Laboratory, Argonne, Illinois



Lifson&Warshal 1968

THE JOURNAL OF CHEMICAL PHYSICS VOLUME 49, NUMBER 11 1 DECEMBER 1968

Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and *n*-Alkane Molecules

S. LIFSON AND A. WARSHEL

Department of Chemical Physics, Weizmann Institute of Science, Rehovot, Israel

$$\begin{aligned}
 V(\mathbf{s}) = & \frac{1}{2} \sum_i K_b (b_i - b_0)^2 + \frac{1}{2} \sum_i K_\theta (\theta_i - \theta_0)^2 \\
 & + \frac{1}{2} \sum_{i,\sigma} K_a (a_i^\sigma - a_0) + \frac{1}{2} \sum_{i,\sigma} K_\gamma (\gamma_i^\sigma - \gamma_0) \\
 & + \frac{1}{2} \sum_i K_\delta (\delta_i - \delta_0)^2 + \frac{1}{2} \sum_i K_\phi (1 + \cos 3\phi_i) \\
 & + \sum_{i,j} V_{nb}(\mathbf{r}_{ij}), \quad (27)
 \end{aligned}$$

$$\begin{aligned}
 V_{(6-\text{exp})}(\mathbf{r}_{ij}) = & A \exp(-B\mathbf{r}_{ij}) - \frac{C}{\mathbf{r}_{ij}^6} \\
 \equiv & \exp\left(\frac{\mathbf{r}^0 - \mathbf{r}_{ij}}{\rho}\right) - \frac{C}{\mathbf{r}_{ij}^6}, \quad (28)
 \end{aligned}$$



many body: Finis-Sinclair 1983

$$U_{metal} = \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N V_{ij}(r_{ij}) + \sum_{i=1}^N F(\rho_i) , \quad \rho_i = \sum_{j=1, j \neq i}^N \rho_{ij}(r_{ij}) .$$

$$V_{ij}(r_{ij}) = \begin{cases} (r_{ij} - c)^2(c_0 + c_1 r_{ij} + c_2 r_{ij}^2) & r_{ij} < c \\ 0 & r_{ij} > c \end{cases}$$

$$\rho_{ij}(r_{ij}) = \begin{cases} (r_{ij} - d)^2 + \beta \frac{(r_{ij} - d)^3}{d} & r_{ij} < d \\ 0 & r_{ij} > d \end{cases}$$

$$F(\rho_i) = -A\sqrt{\rho_i} ,$$



PHYSICAL REVIEW B

VOLUME 31, NUMBER 8

15 APRIL 1985

Computer simulation of local order in condensed phases of silicon

Frank H. Stillinger and Thomas A. Weber

AT&T Bell Laboratories, Murray Hill, New Jersey 07974

(Received 7 November 1984)

PHYSICAL REVIEW B

VOLUME 37, NUMBER 12

15 APRIL 1988-II

New empirical approach for the structure and energy of covalent systems

J. Tersoff

IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598

(Received 28 August 1987)



tersoff

$$E = \frac{1}{2} \sum_i \sum_{j \neq i} V_{ij}$$

$$V_{ij} = f_C(r_{ij} + \delta) [f_R(r_{ij} + \delta) + b_{ij} f_A(r_{ij} + \delta)]$$

$$f_C(r) = \begin{cases} 1 & r < R - D \\ \frac{1}{2} - \frac{1}{2} \sin\left(\frac{\pi}{2} \frac{r-R}{D}\right) & R - D < r < R + D \\ 0 & r > R + D \end{cases}$$

$$f_R(r) = A \exp(-\lambda_1 r)$$

$$f_A(r) = -B \exp(-\lambda_2 r)$$

$$b_{ij} = (1 + \beta^n \zeta_{ij}^n)^{-\frac{1}{2n}}$$

$$\zeta_{ij} = \sum_{k \neq i, j} f_C(r_{ik} + \delta) g[\theta_{ijk}(r_{ij}, r_{ik})] \exp[\lambda_3 m (r_{ij} - r_{ik})^m]$$

$$g(\theta) = \gamma_{ijk} \left(1 + \frac{c^2}{d^2} - \frac{c^2}{[d^2 + (\cos \theta - \cos \theta_0)^2]} \right)$$





...





Almost perfect but

- ▶ accuracy is an issue
- ▶ fitting can be not trivial, mixed theoretical and experimental data
- ▶ transferability
- ▶ reactivity possible but restricted



introduction

Theoretical era

Simulations era

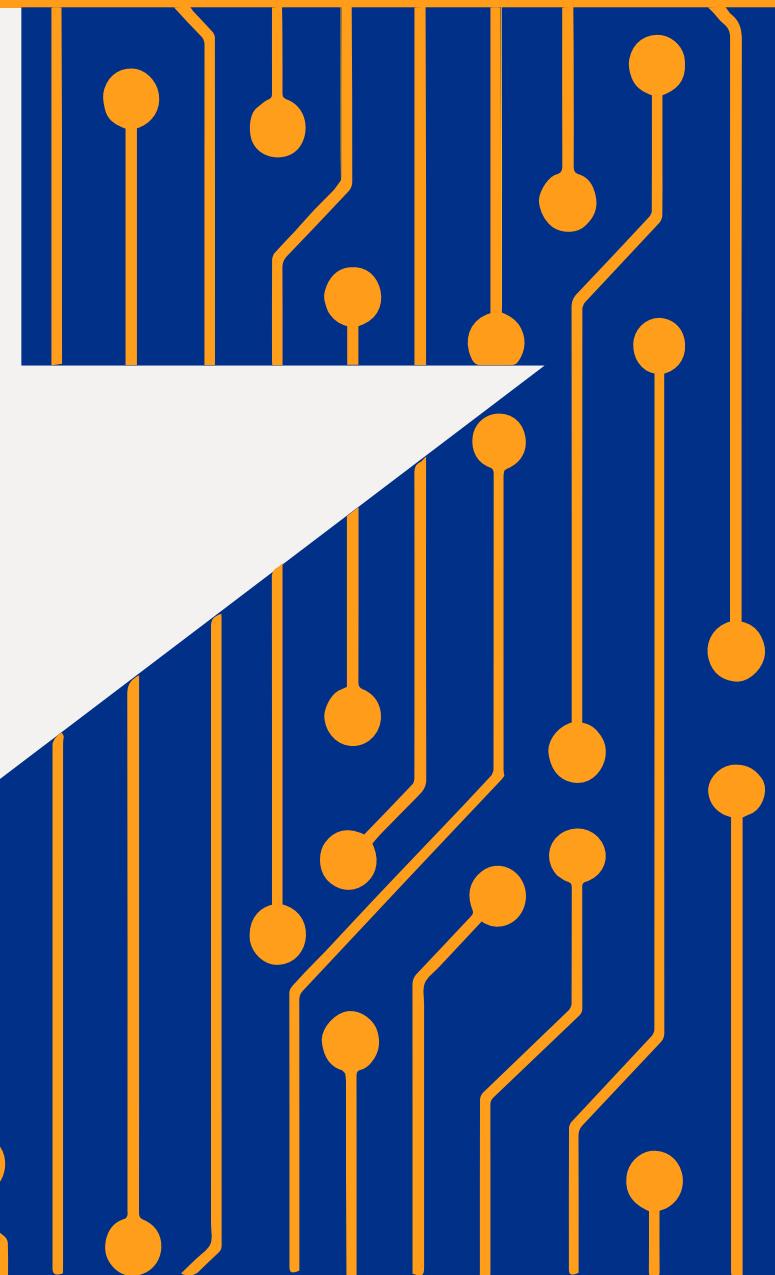
ML era

metal organic frameworks

2,5-Diiodothiophene (pressure)

zeolites

performance





Neural network models of potential energy surfaces

Thomas B. Blank, Steven D. Brown, August W. Calhoun,^{a)} and Douglas J. Doren

Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716

(Received 6 February 1995; accepted 7 June 1995)

VOLUME 93, NUMBER 17

PHYSICAL REVIEW LETTERS

22 OCTOBER 2004

“Learn on the Fly”: A Hybrid Classical and Quantum-Mechanical Molecular Dynamics Simulation

Gabor Csányi,¹ T. Albaret,³ M. C. Payne,¹ and A. De Vita^{2,3}

¹*Cavendish Laboratory, Madingley Road, Cambridge, CB3 0HE, United Kingdom*

²*Physics Department, King’s College London, Strand, London WC2R 2LS, United Kingdom*

³*INFM-DEMOCRITOS National Simulation Center and Center of Excellence for Nanostructured Materials (CENMAT),*

University of Trieste, Trieste, Italy

(Received 18 May 2004; published 19 October 2004)

PRL 98, 146401 (2007)

PHYSICAL REVIEW LETTERS

week ending
6 APRIL 2007

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

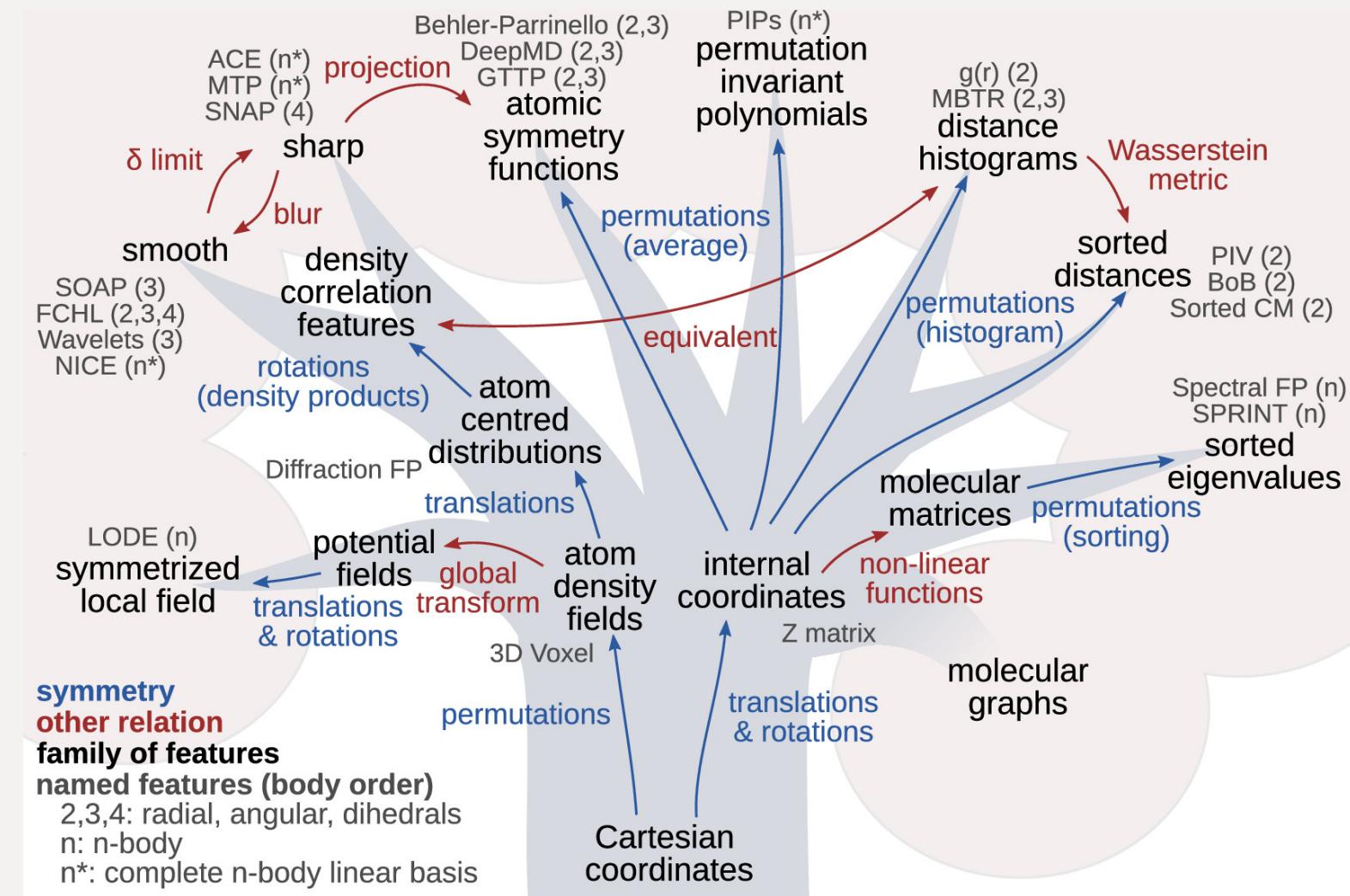
Jörg Behler and Michele Parrinello

Department of Chemistry and Applied Biosciences, ETH Zurich, USI-Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland





potentials





...

**Chemical
Science**

EDGE ARTICLE

 CrossMark
click for updates

Cite this: *Chem. Sci.*, 2017, 8, 3192



View Article Online
[View Journal](#) | [View Issue](#)

ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost†

J. S. Smith,^a O. Isayev^{*b} and A. E. Roitberg^{*a}

JCTC Journal of Chemical Theory and Computation

 **Cite This:** *J. Chem. Theory Comput.* 2019, 15, 3793–3809

[pubs.acs.org/JCTC](#)

Article

From Molecular Fragments to the Bulk: Development of a Neural Network Potential for MOF-5

Marco Eckhoff^{*ID} and Jörg Behler^{*ID}

Universität Göttingen, Institut für Physikalische Chemie, Theoretische Chemie, Tammannstraße 6, D-37077 Göttingen, Germany



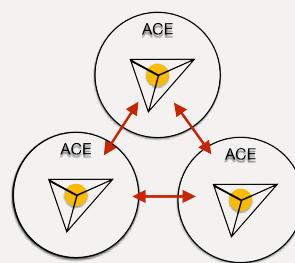
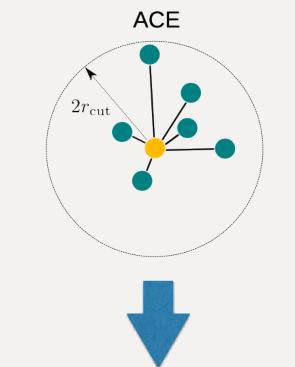
what are they?

- ▶ many body interatomic potentials
- ▶ use descriptors for local environments
- ▶ in general short range
- ▶ long range models are developed
- ▶ use gaussian regression or some variation of neural networks
- ▶ achieve DFT&beyond accuracy
- ▶ at least 5-6 orders of magnitude cheaper than DFT
- ▶ incrementally improvable
- ▶ depend on good quality data generation
- ▶ trained against, energies, forces, [stress]



Batatia et al, NeurIPS 2022 or J. Chem. Phys. 159, 044118 (2023)

MACE: message passing network using Atomic Cluster Expansion layers



$$h_{i,k00}^{(0)} = \sum_z W_{kz} \delta_{zz_i} \quad (1)$$

$$\bar{h}_{i,kl_2m_2}^{(s)} = \sum_{\bar{k}} W_{\bar{k}\bar{k}l_2}^{(s)} h_{i,\bar{k}l_2m_2}^{(s)} \quad (2)$$

$$j_0^n(r_{ij}) = \sqrt{\frac{2}{r_{cut}}} \sin\left(n\pi \frac{r_{ij}}{r_{cut}}\right) f_{cut}(r_{ij}) \quad (3)$$

$$R_{k\eta_1l_1l_2l_3}^{(s)}(r_{ij}) = \text{MLP}(\{j_0^n(r_{ij})\}_n) \quad (4)$$

$$\phi_{ij,k\eta_1l_3m_3}^{(s)} = \sum_{l_1l_2m_1m_2} C_{\eta_1,l_1m_1l_2m_2}^{l_3m_3} R_{k\eta_1l_1l_2l_3}^{(s)}(r_{ij}) \times Y_{l_1}^{m_1}(\hat{r}_{ij}) \bar{h}_{j,kl_2m_2}^{(s)} \quad (5)$$

$$A_{i,kl_3m_3}^{(s)} = \sum_{\bar{k},\eta_1} W_{\bar{k}\bar{k}\eta_1l_3}^{(s)} \sum_{j \in \mathcal{N}(i)} \phi_{ij,\bar{k}\eta_1l_3m_3}^{(s)} \quad (6)$$

$$\mathbf{A}_{i,klm}^{(s),\nu} = \prod_{\xi=1}^{\nu} A_{i,kl\xi m_\xi}^{(s)} \quad (7)$$

$$\mathbf{B}_{i,\eta_\nu kLM}^{(s),\nu} = \sum_{lm} C_{\eta_\nu l m}^{LM} \mathbf{A}_{i,klm}^{(s),\nu} \quad (8)$$

$$m_{i,kLM}^{(s)} = \sum_{\nu} \sum_{\eta_\nu} W_{z_i \eta_\nu k L M}^{(s),\nu} \mathbf{B}_{i,\eta_\nu kLM}^{(s),\nu} \quad (9)$$

atom features:
element embedding
(cf: ACE tensor decomposition!)

Bessel functions

R radial basis: splines

ϕ one particle basis

A_k atomic basis
(perm invariance)

\mathbf{A}_k many-particle basis

\mathbf{B}_k ACE basis
(rot equivariance)

sum pool

node feature update

Only nonlinear activations in radial basis and final readout

Sparse polynomial “symmetric tensor network”

$$(10)$$



Foundation models

Model	CPS ↑	Acc ↑	F1 ↑	DAF ↑	Prec ↑	MAE ↓	R² ↑	K _{SRME} ↓	RMSD ↓	Training Set	Params	Targets
eSEN-30M-OAM	0.888	0.977	0.925	6.069	0.928	0.018	0.866	0.170	0.061	6.6M (113M) OMat24+MPtrj+sAlex	30.2M	EFS _G
ORB v3	0.861	0.971	0.905	5.912	0.904	0.024	0.821	0.210	0.075	6.47M (133M) MPtrj+Alex+OMat24	25.5M	EFS _G
SevenNet-MF-ompa	0.845	0.969	0.901	5.825	0.890	0.021	0.867	0.317	0.064	6.6M (113M) OMat24+sAlex+MPtrj	25.7M	EFS _G
GRACE-2L-OAM	0.837	0.963	0.880	5.774	0.883	0.023	0.862	0.294	0.067	6.6M (113M) OMat24+sAlex+MPtrj	12.6M	EFS _G
AlphaNet-v1-OMA	0.804	0.971	0.909	5.777	0.883	0.020	0.868	0.644	0.032	6.6M (113M) OMat24+sAlex+MPtrj	4.65M	EFS _G
DPA-3.1-3M-FT	0.802	0.963	0.884	5.667	0.866	0.023	0.869	0.469	0.069	163M OpenLAM	3.27M	EFS _G
MatterSim v1 5M	0.767	0.959	0.862	5.852	0.895	0.024	0.863	0.574	0.073	17M MatterSim	4.55M	EFS _G
GNoME	NaN	0.955	0.829	5.523	0.844	0.035	0.785	n/a	n/a	6M (89M) GNoME	16.2M	EFS _G
eSEN-30M-MP	0.797	0.946	0.831	5.260	0.804	0.033	0.822	0.340	0.075	146k (1.58M) MPtrj	30.1M	EFS _G
MACE-MPA-0	0.795	0.954	0.852	5.582	0.853	0.028	0.842	0.412	0.073	3.37M (12M) MPtrj+sAlex	9.06M	EFS _G
GRACE-1L-OAM	0.761	0.944	0.824	5.255	0.803	0.031	0.842	0.516	0.072	6.6M (113M) OMat24+sAlex+MPtrj	3.45M	EFS _G
DPA-3.1-MPtrj	0.718	0.936	0.803	5.024	0.768	0.037	0.812	0.650	0.080	146k (1.58M) MPtrj	4.81M	EFS _G
SevenNet-l3i5	0.714	0.920	0.760	4.629	0.708	0.044	0.776	0.550	0.085	146k (1.58M) MPtrj	1.17M	EFS _G
MatRIS v0.5.0 MPtrj	0.681	0.938	0.809	5.049	0.772	0.037	0.803	0.861	0.077	146k (1.58M) MPtrj	5.83M	EFS _{G,M}
GRACE-2L-MPtrj	0.681	0.896	0.691	4.163	0.636	0.052	0.741	0.525	0.090	146k (1.58M) MPtrj	15.3M	EFS _G
MACE-MP-0	0.644	0.878	0.669	3.777	0.577	0.057	0.697	0.647	0.091	146k (1.58M) MPtrj	4.69M	EFS _G
eqV2 M	0.558	0.975	0.917	6.047	0.924	0.020	0.848	1.771	0.069	3.37M (102M) OMat24+MPtrj	86.6M	EFS _D
ORB v2	0.529	0.965	0.880	6.041	0.924	0.028	0.824	1.732	0.097	3.25M (32.1M) MPtrj+Alex	25.2M	EFS _D
eqV2 S DeNS	0.522	0.941	0.815	5.042	0.771	0.036	0.788	1.676	0.076	146k (1.58M) MPtrj	31.2M	EFS _D
ORB v2 MPtrj	0.470	0.922	0.765	4.702	0.719	0.045	0.756	1.725	0.101	146k (1.58M) MPtrj	25.2M	EFS _D
M3GNet	0.428	0.813	0.569	2.882	0.441	0.075	0.585	1.412	0.112	62.8k (188k) MPF	228k	EFS _G
CHGNet	0.400	0.851	0.613	3.361	0.514	0.063	0.689	1.717	0.095	146k (1.58M) MPtrj	413k	EFS _{G,M}

Materials Discovery

Model	CPS ↑	Acc ↑	F1 ↑	DAF ↑	Prec ↑	MAE ↓	R² ↑	K _{SRME} ↓	RMSD ↓	Training Set	Params	Targets
eSEN-30M-OAM	0.888	0.977	0.925	6.069	0.928	0.018	0.866	0.170	0.061	6.6M (113M) OMat24+MPtrj+sAlex	30.2M	EFS _G
ORB v3	0.861	0.971	0.905	5.912	0.904	0.024	0.821	0.210	0.075	6.47M (133M) MPtrj+Alex+OMat24	25.5M	EFS _G
GRACE-2L-OAM	0.837	0.969	0.901	5.825	0.890	0.021	0.867	0.317	0.064	6.6M (113M) OMat24+sAlex+MPtrj	12.6M	EFS _G
SevenNet-MF-ompa	0.845	0.969	0.901	5.825	0.890	0.021	0.867	0.317	0.064	6.6M (113M) OMat24+sAlex+MPtrj	25.7M	EFS _G
eSEN-30M-MP	0.797	0.946	0.831	5.260	0.804	0.033	0.822	0.340	0.075	146k (1.58M) MPtrj	30.1M	EFS _G
MACE-MPA-0	0.795	0.954	0.852	5.582	0.853	0.028	0.842	0.412	0.073	3.37M (12M) MPtrj+sAlex	9.06M	EFS _G
DPA-3.1-3M-FT	0.802	0.963	0.884	5.667	0.866	0.023	0.869	0.469	0.069	163M OpenLAM	3.27M	EFS _G
GRACE-1L-OAM	0.761	0.944	0.824	5.255	0.803	0.031	0.842	0.516	0.072	6.6M (113M) OMat24+sAlex+MPtrj	3.45M	EFS _G
GRACE-2L-MPtrj	0.681	0.896	0.691	4.163	0.636	0.052	0.741	0.525	0.090	146k (1.58M) MPtrj	15.3M	EFS _G
SevenNet-l3i5	0.714	0.920	0.760	4.629	0.708	0.044	0.776	0.550	0.085	146k (1.58M) MPtrj	1.17M	EFS _G
MatterSim v1 5M	0.767	0.959	0.862	5.852	0.895	0.024	0.863	0.574	0.073	17M MatterSim	4.55M	EFS _G
AlphaNet-v1-OMA	0.804	0.971	0.909	5.777	0.883	0.020	0.868	0.644	0.032	6.6M (113M) OMat24+sAlex+MPtrj	4.65M	EFS _G
MACE-MP-0	0.644	0.878	0.669	3.777	0.577	0.057	0.697	0.647	0.091	146k (1.58M) MPtrj	4.69M	EFS _G
DPA-3.1-MPtrj	0.718	0.936	0.803	5.024	0.768	0.037	0.812	0.650	0.080	146k (1.58M) MPtrj	4.81M	EFS _G
MatRIS v0.5.0 MPtrj	0.681	0.938	0.809	5.049	0.772	0.037	0.803	0.861	0.077	146k (1.58M) MPtrj	5.83M	EFS _{G,M}
M3GNet	0.428	0.813	0.569	2.882	0.441	0.075	0.585	1.412	0.112	62.8k (188k) MPF	228k	EFS _G
eqV2 S DeNS	0.522	0.941	0.815	5.042	0.771	0.036	0.788	1.676	0.076	146k (1.58M) MPtrj	31.2M	EFS _D
CHGNet	0.400	0.851	0.613	3.361	0.514	0.063	0.689	1.717	0.095	146k (1.58M) MPtrj	413k	EFS _{G,M}
ORB v2 MPtrj	0.470	0.922	0.765	4.702	0.719	0.045	0.756	1.725	0.101	146k (1.58M) MPtrj	25.2M	EFS _D
ORB v2	0.529	0.965	0.880	6.041	0.924	0.028	0.824	1.732	0.097	3.25M (32.1M) MPtrj+Alex	25.2M	EFS _D
eqV2 M	0.558	0.975	0.917	6.047	0.924	0.020	0.848	1.771	0.069	3.37M (102M) OMat24+MPtrj	86.6M	EFS _D
GNoME	NaN	0.955	0.829	5.523	0.844	0.035	0.785	n/a	n/a	6M (89M) GNoME	16.2M	EFS _G

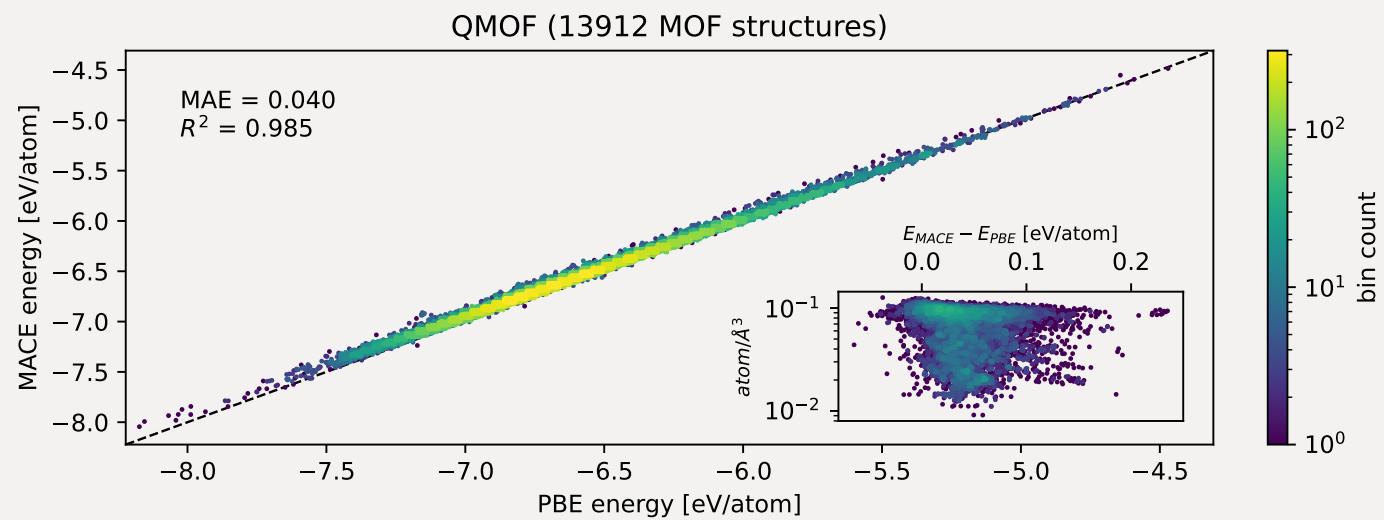
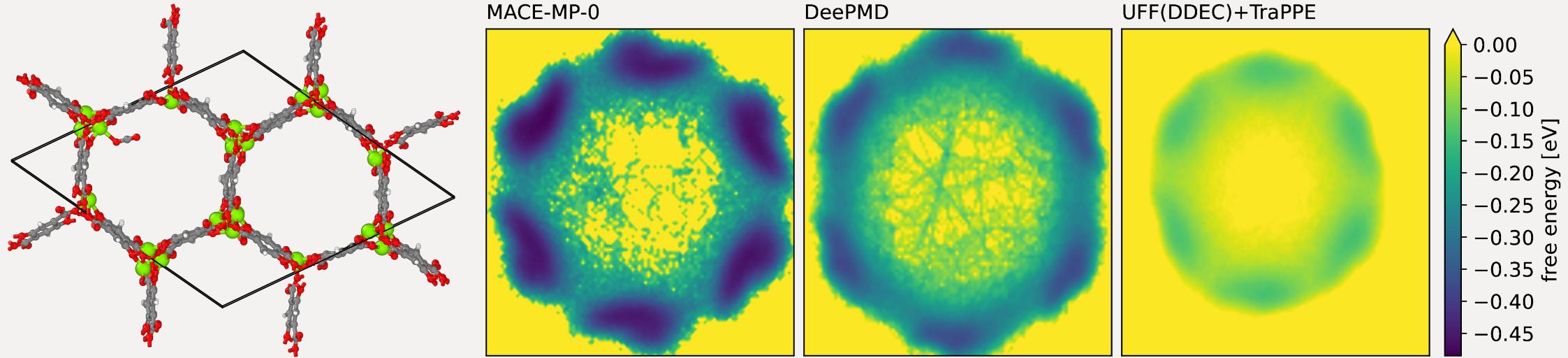
Physics/Chemistry





foundation models: MACE

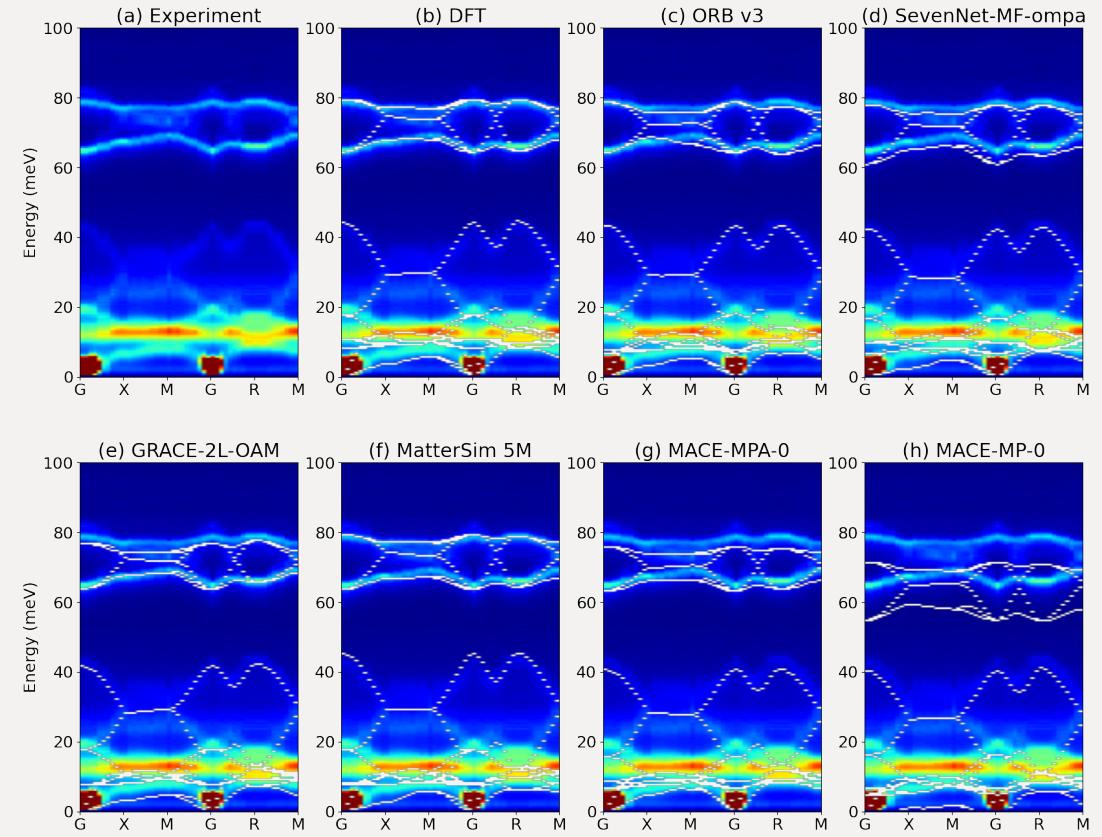
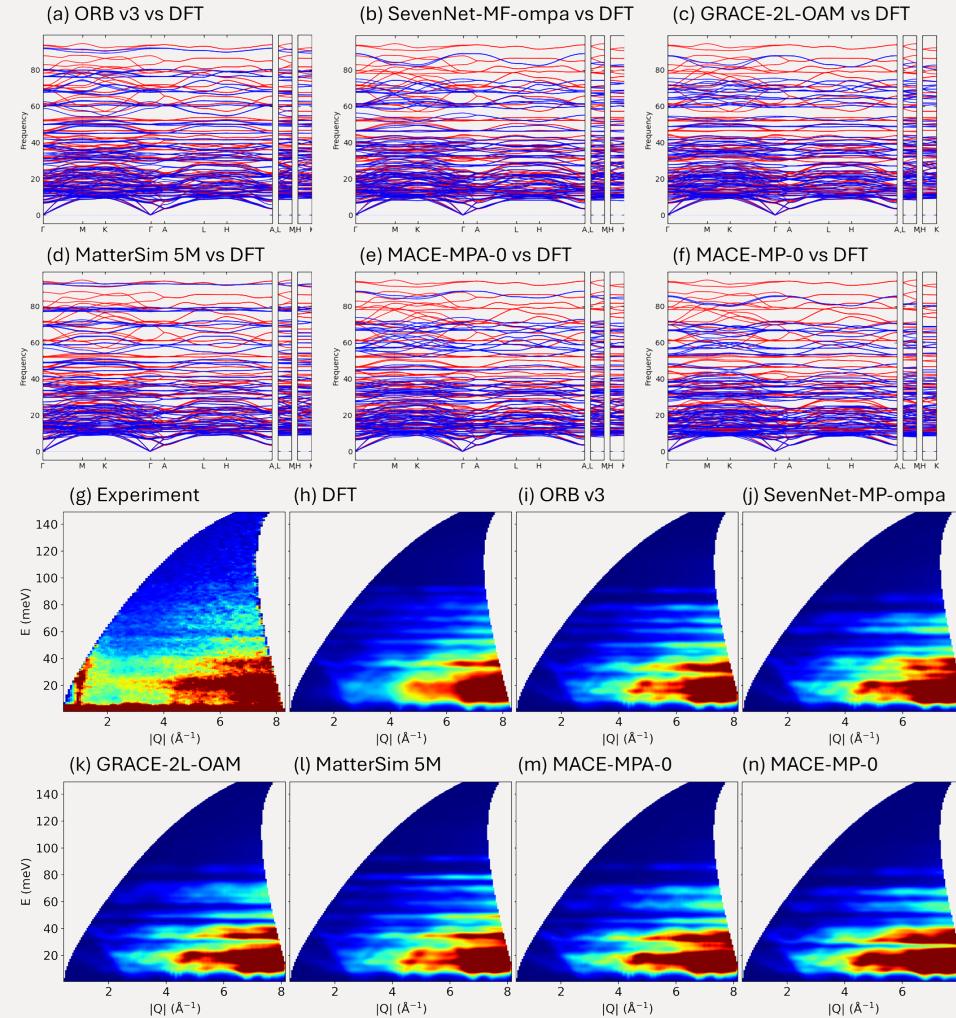
Batatia et al, <https://arxiv.org/abs/2401.00096> (2023)





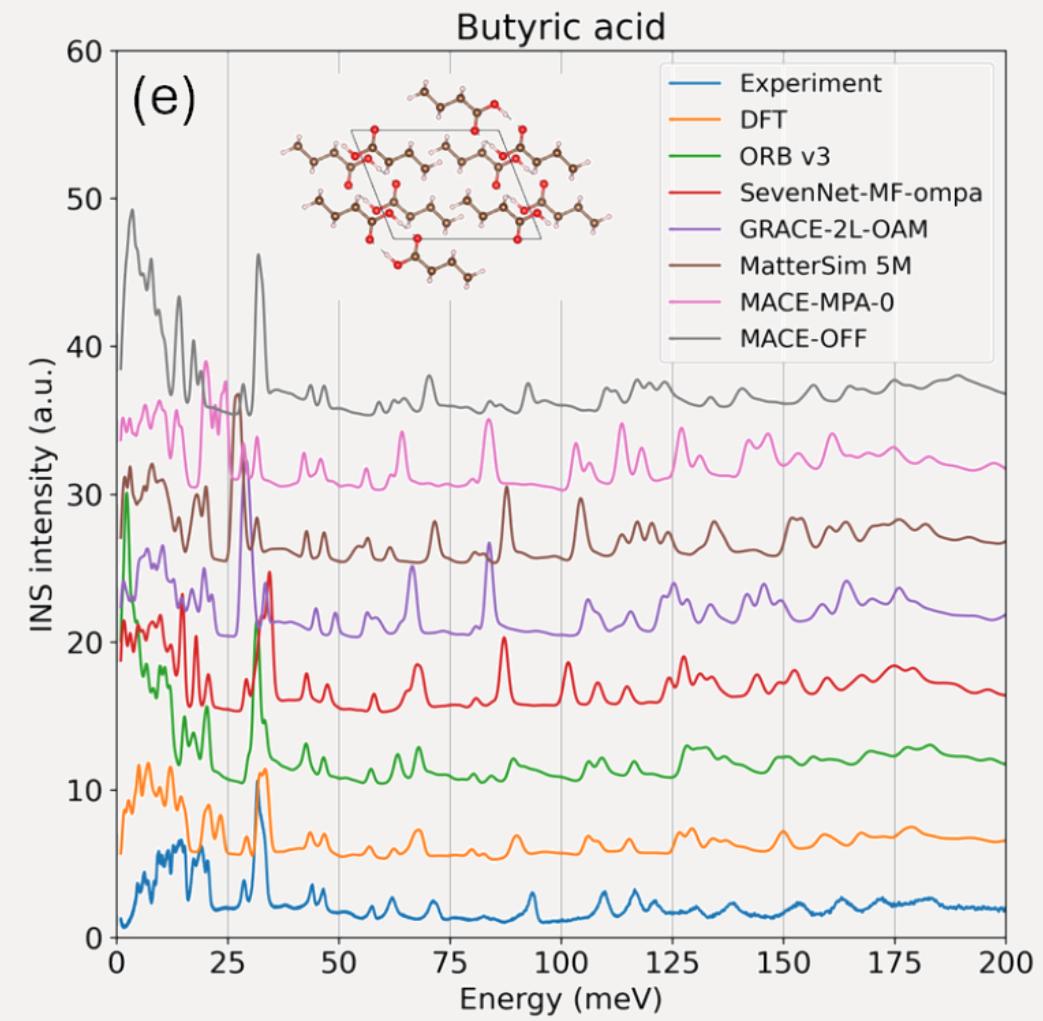
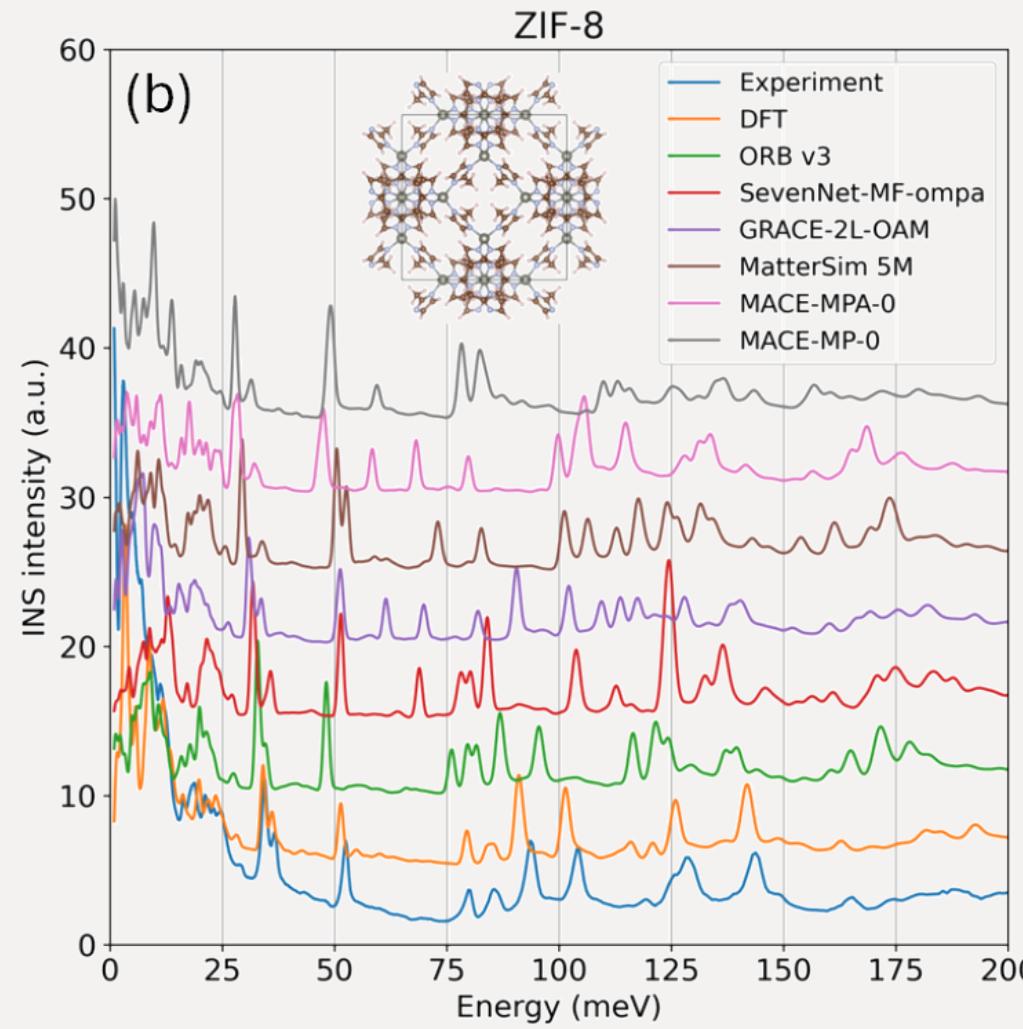
what about experiment?

Han, B and Cheng Y, arXiv:2506.01860v1





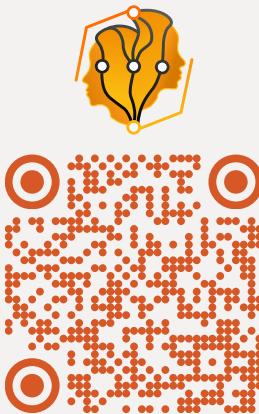
not so simple





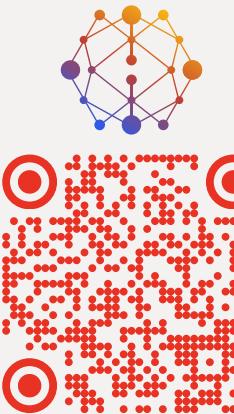
software

janus_core



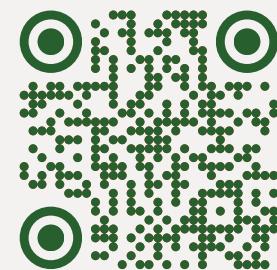
github.com/stfc/janus-core

aiida-mlip



github.com/stfc/aiida-mlip

pack-mm



github.com/ddmms/pack-mm



introduction

Theoretical era

Simulations era

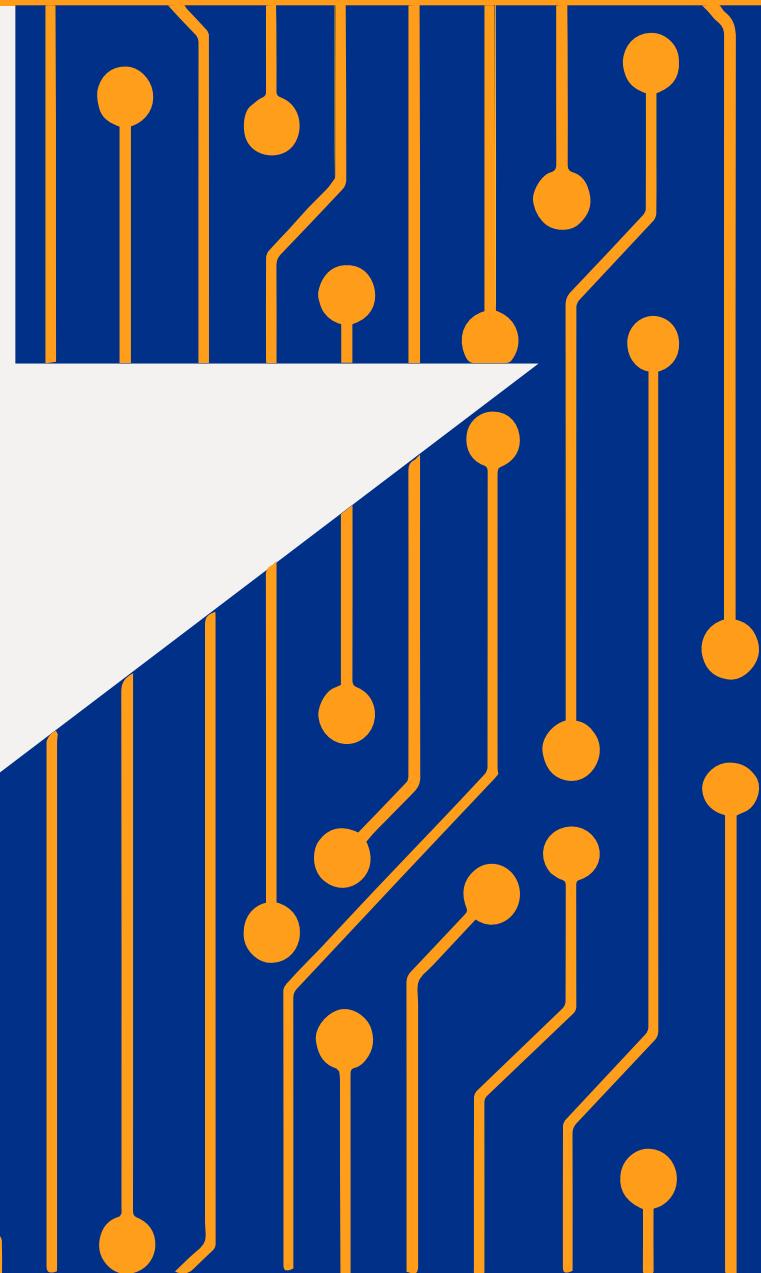
ML era

metal organic frameworks

2,5-Diiodothiophene (pressure)

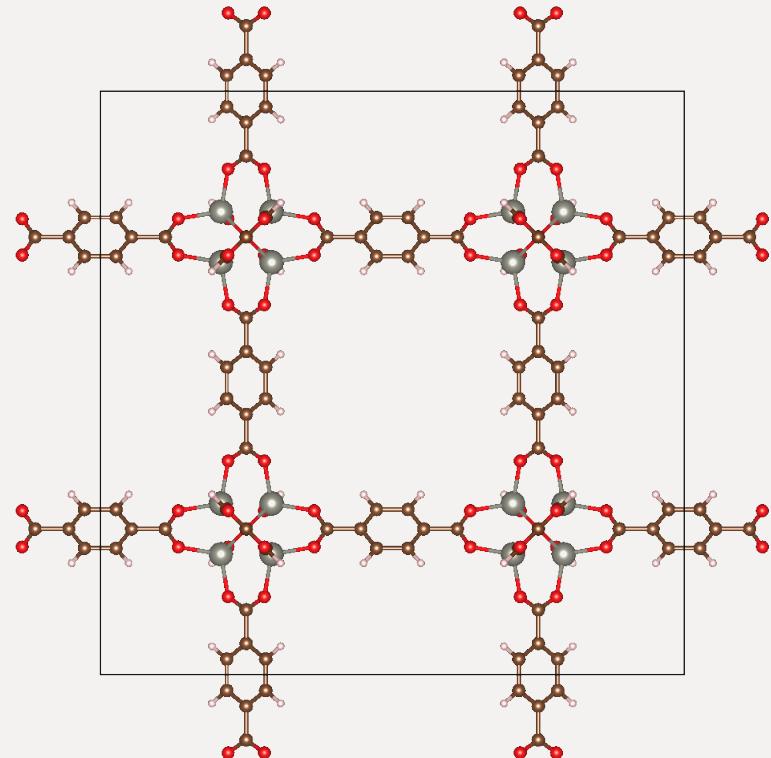
zeolites

performance

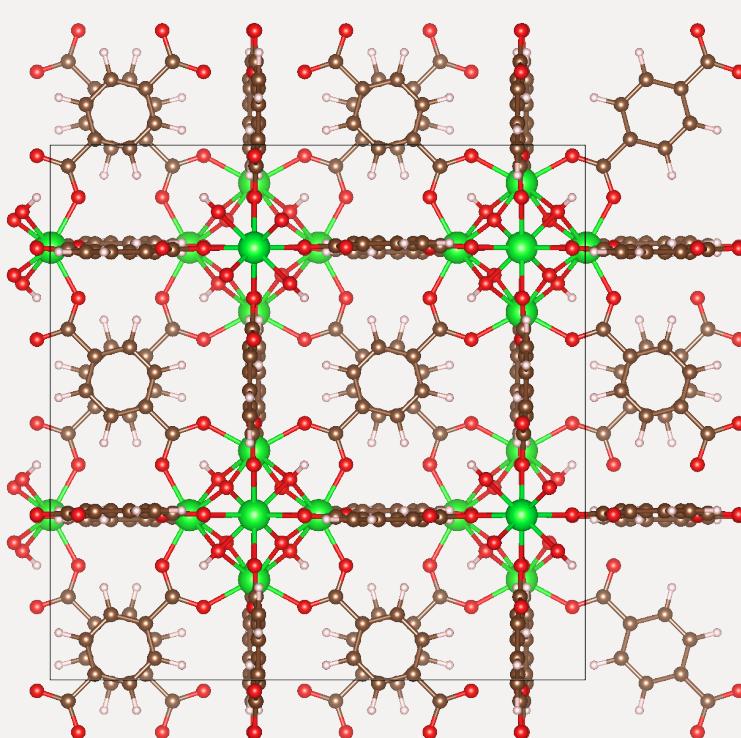




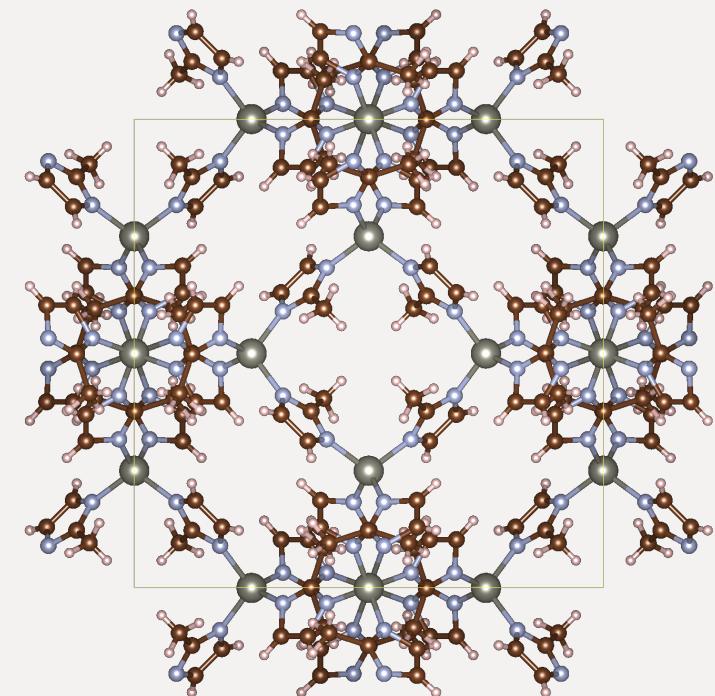
intro



MOF-5



UiO-66



ZIF-8

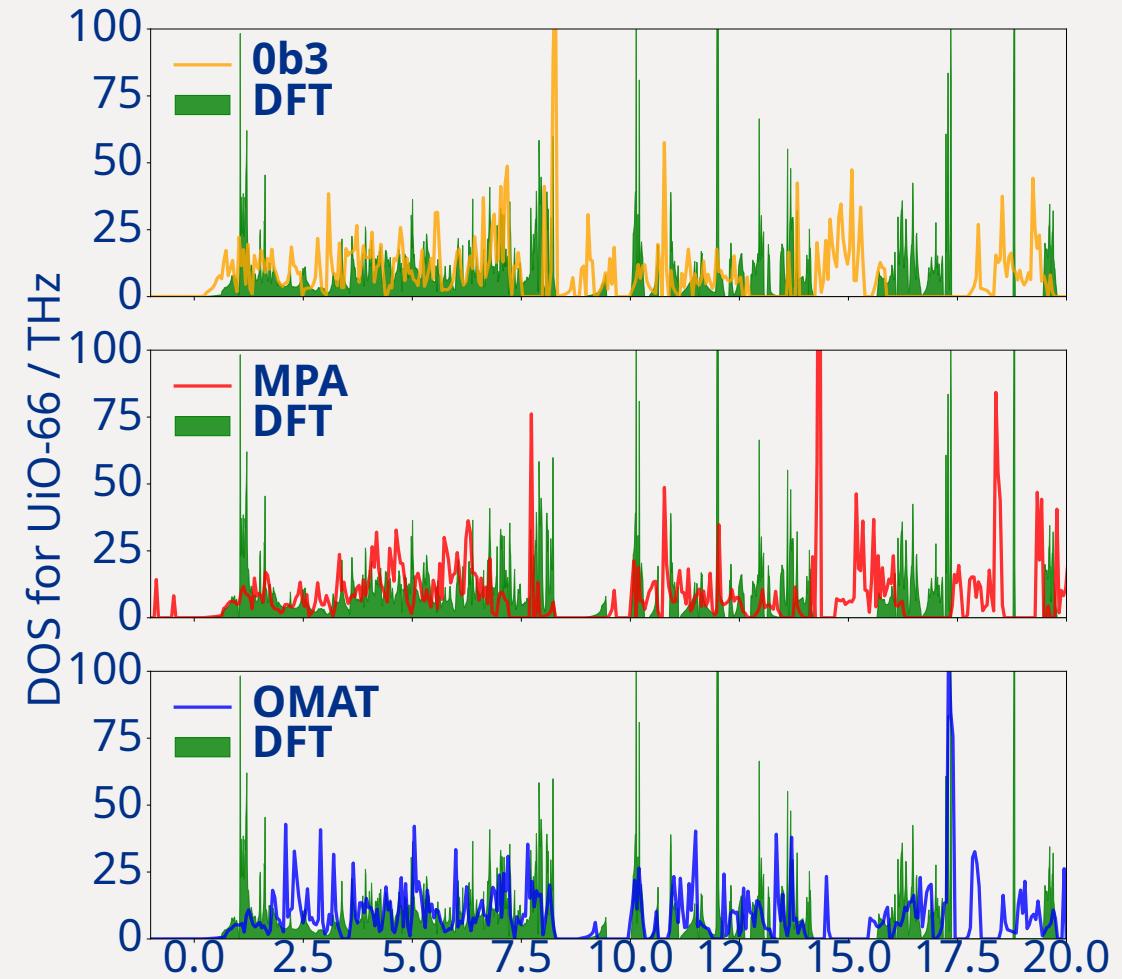
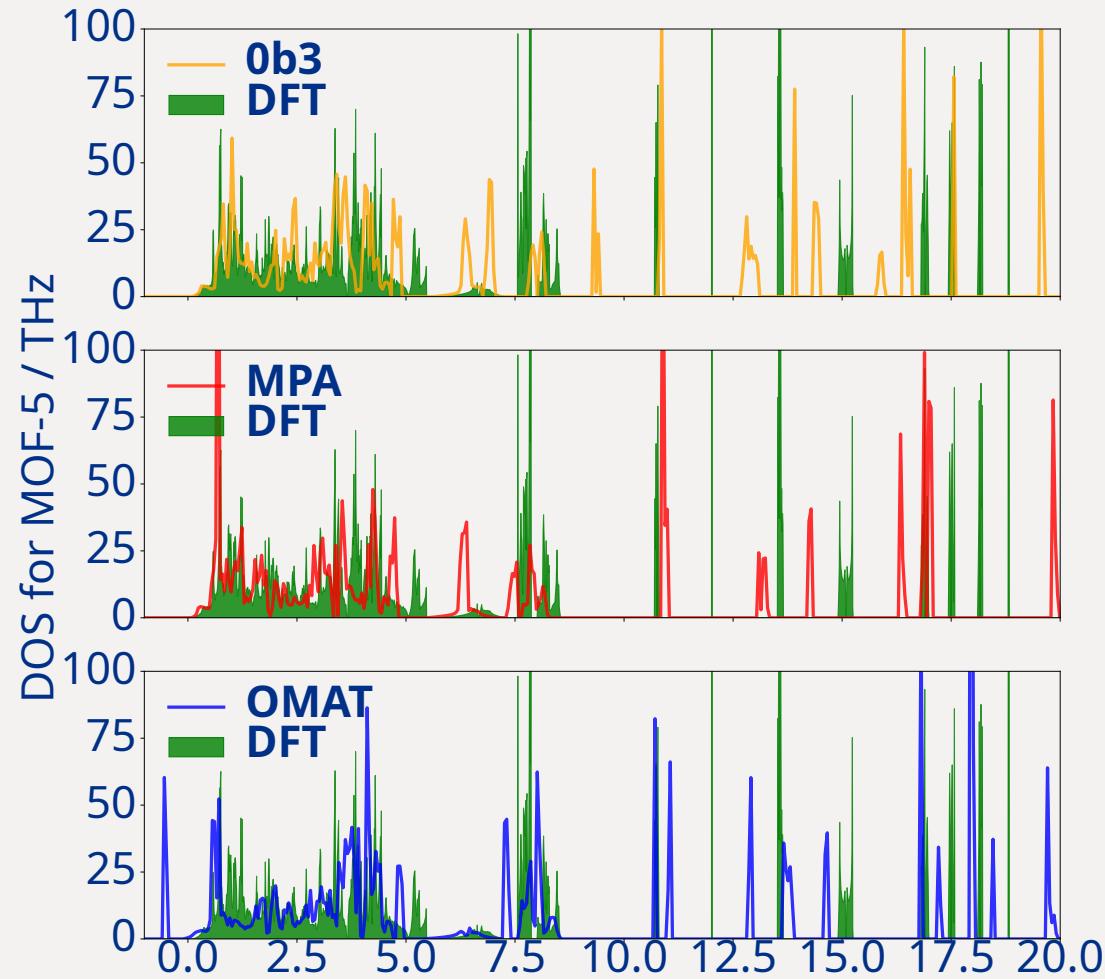


cell parameters

Method	MOF	Group	Space	Unit Cell Angles (°)
			Unit Cell Parameters (Å)	
DFT	MOF-5	Fm-3m	[18.433, 18.433, 18.433]	[60.000, 60.000, 60.000]
	UiO-66	F-43m	[14.803, 14.803, 14.803]	[60.000, 60.000, 60.000]
GFN-1xTB	MOF-5	R-3m	[18.225, 18.225, 18.225]	[59.204, 59.204, 59.204]
	UiO-66	F-43m	[14.560, 14.560, 14.560]	[59.996, 59.997, 59.995]
MACE-MP-0	MOF-5	Fm-3m	[18.412, 18.412, 18.412]	[60.000, 60.000, 60.000]
	UiO-66	F-43m	[14.713, 14.713, 14.713]	[60.000, 60.000, 60.000]



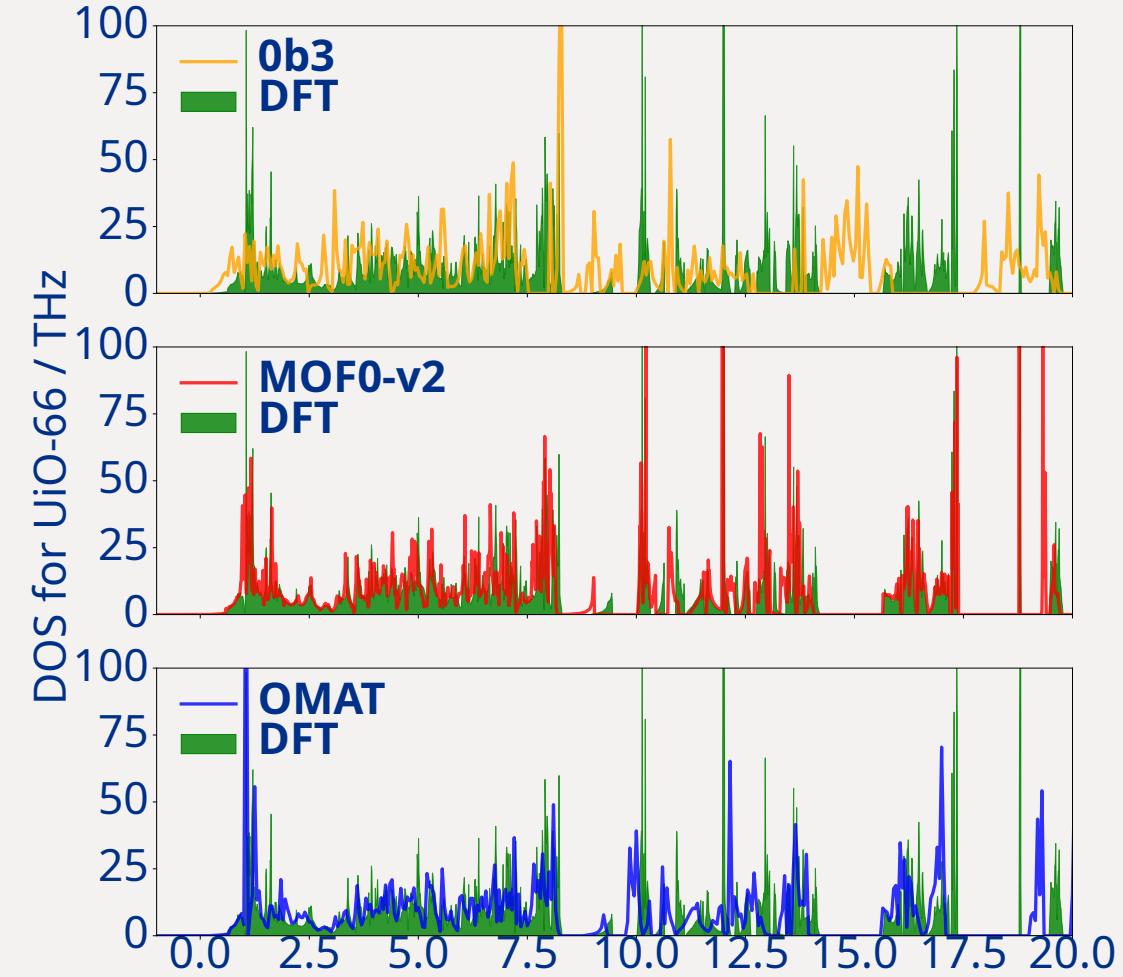
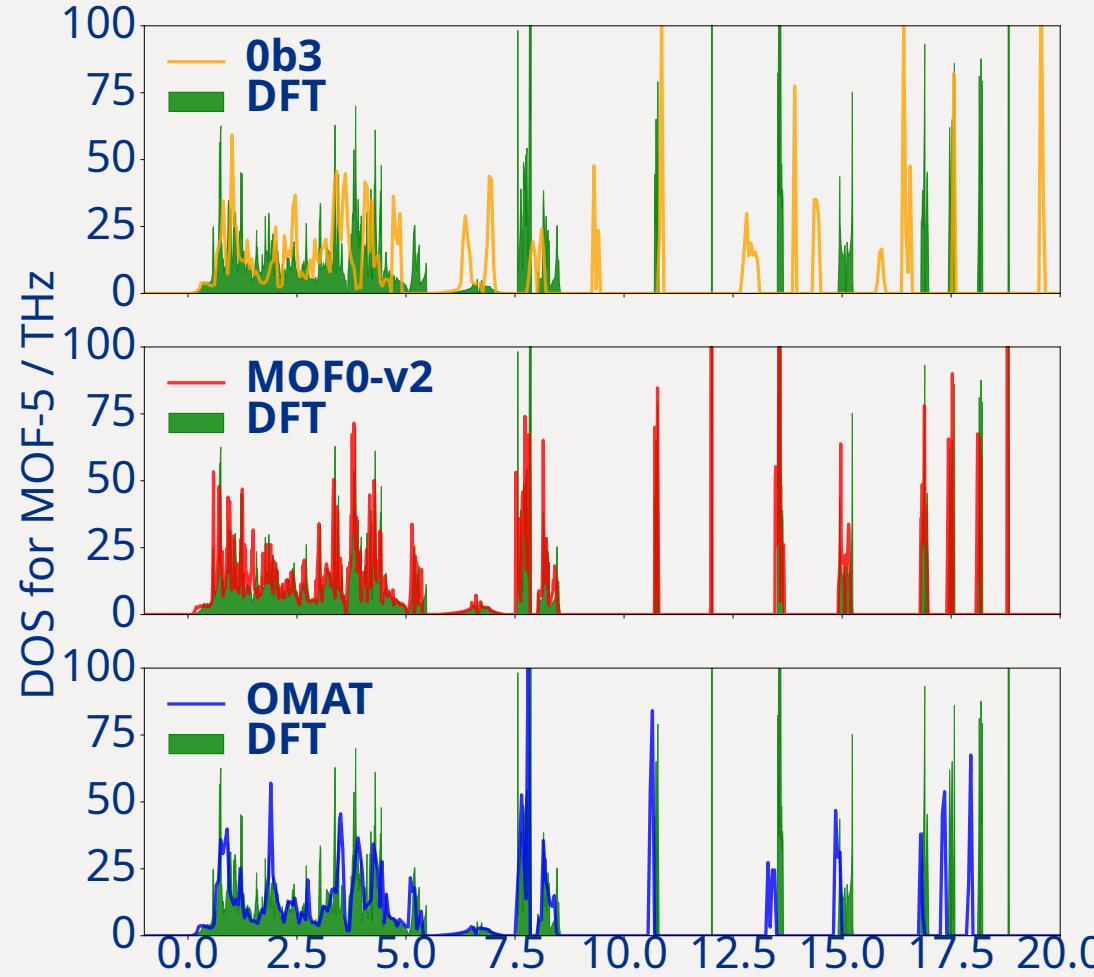
but... density of states





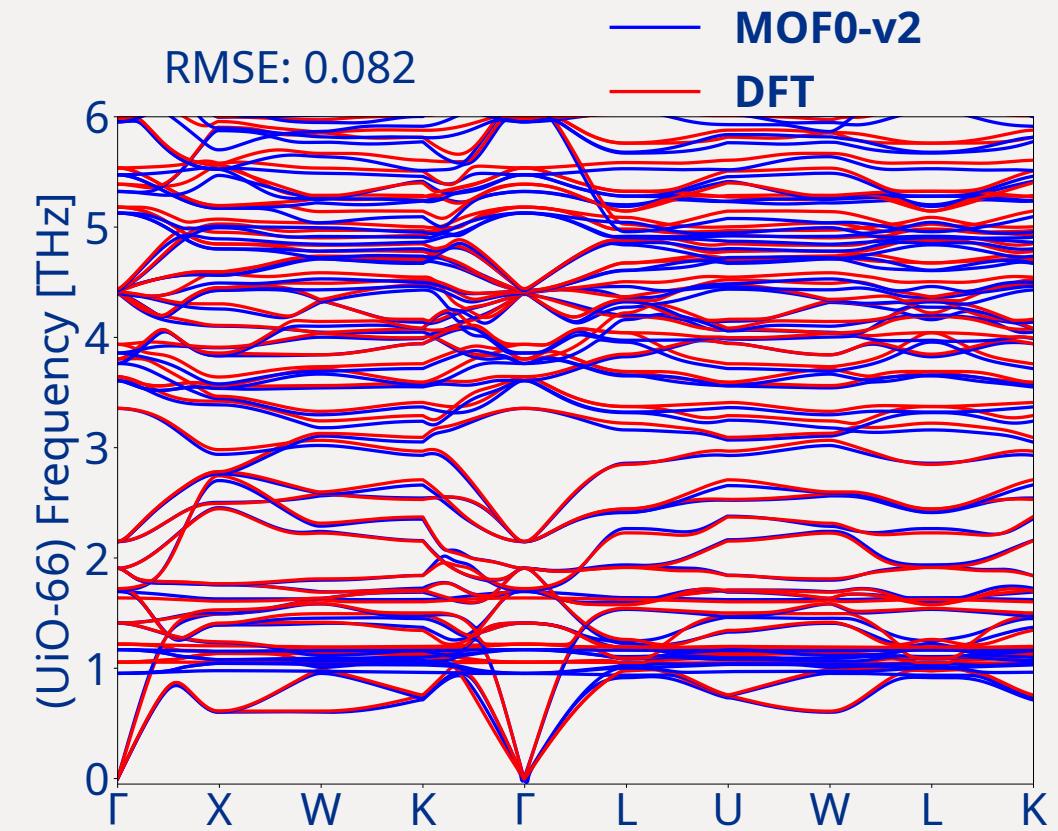
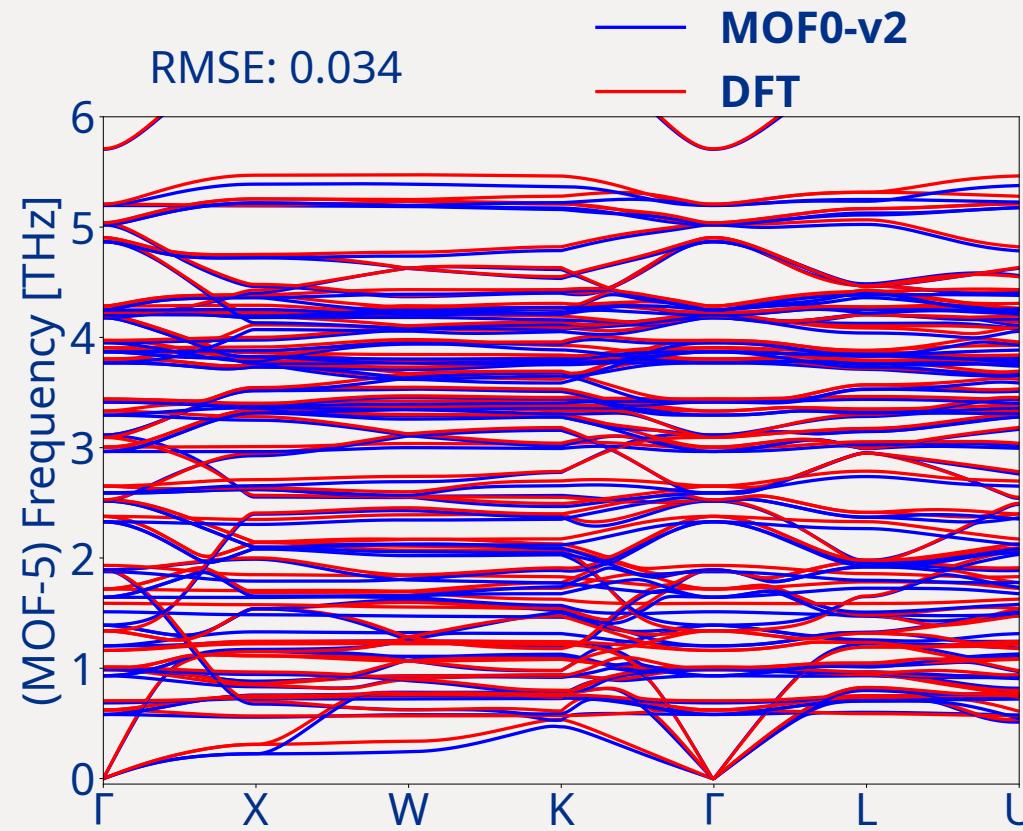
fine tuning

Elena, A.M., Kamath, P.D. et. al, npj Comput Mater 11, 125 (2025). <https://doi.org/10.1038/s41524-025-01611-8>



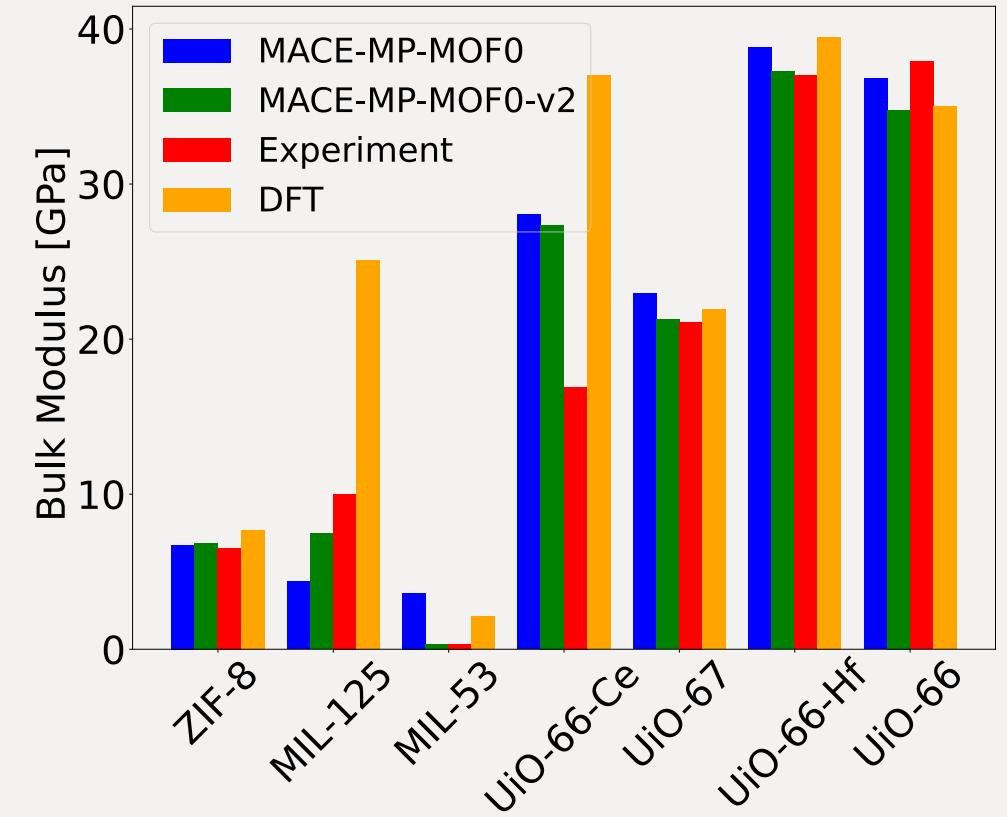
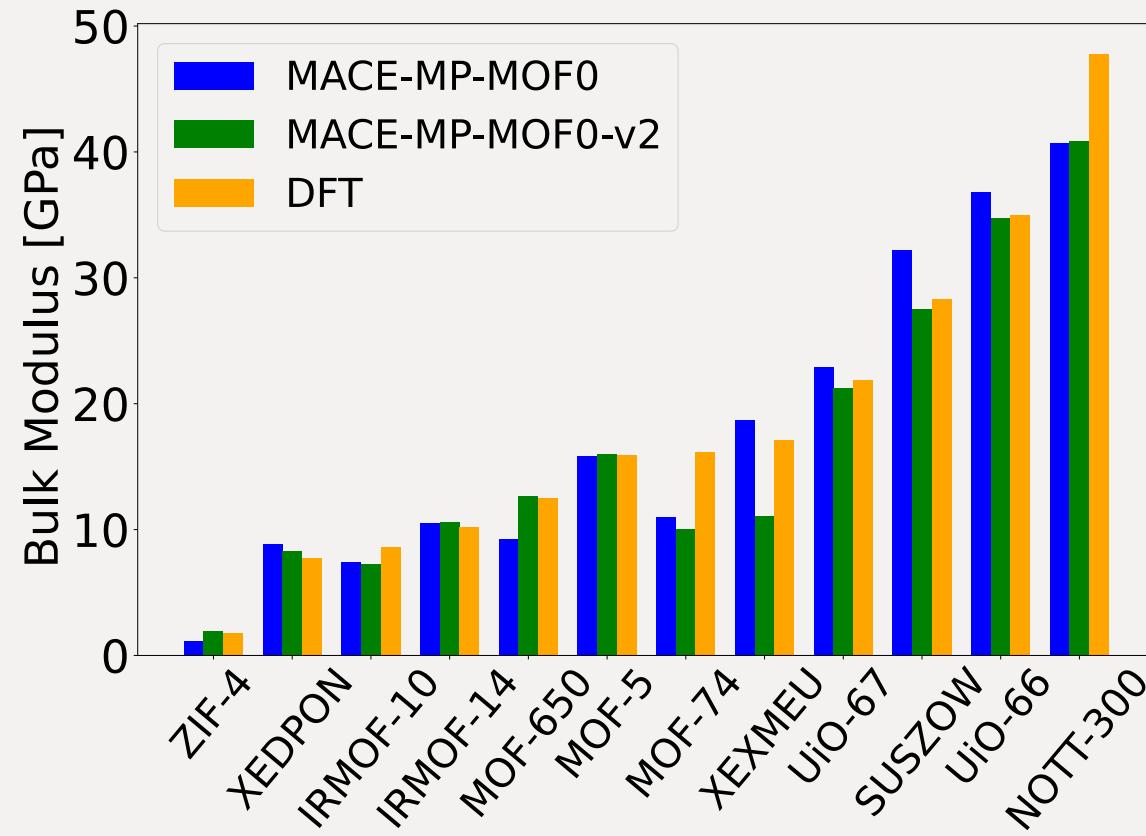


better bands





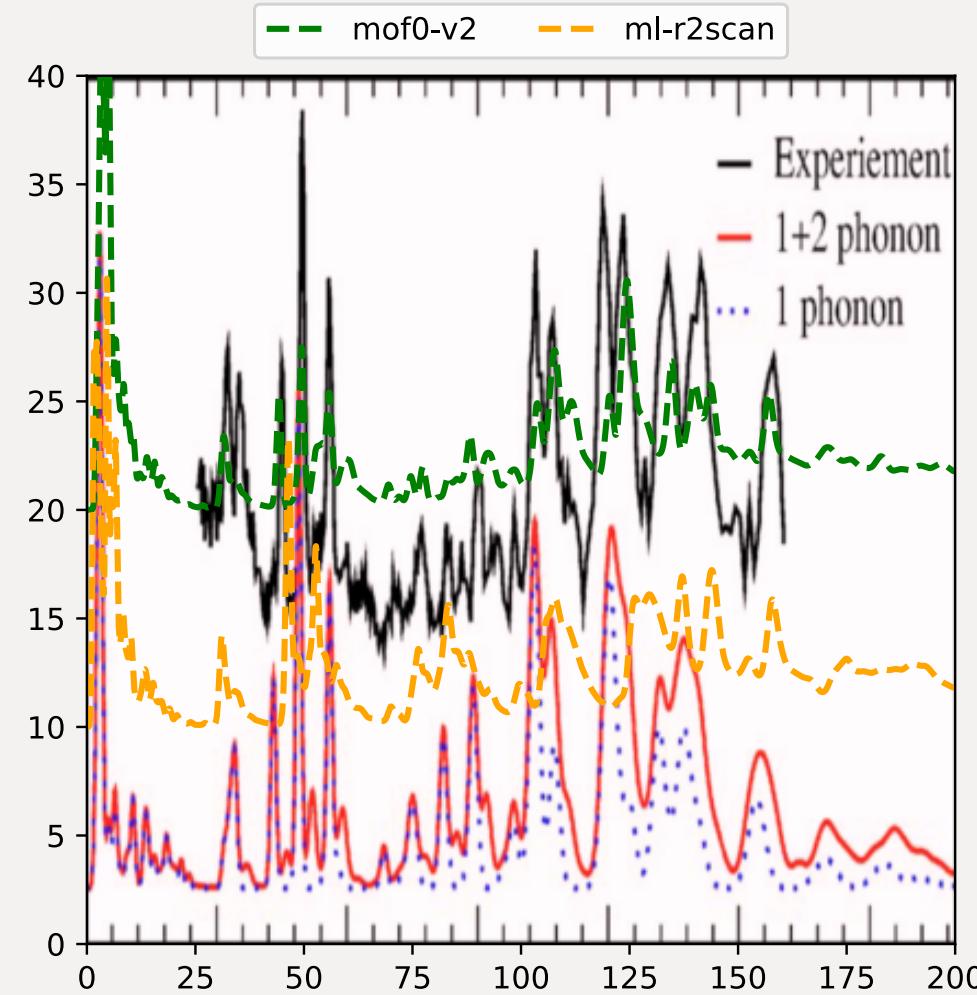
bulk modulus





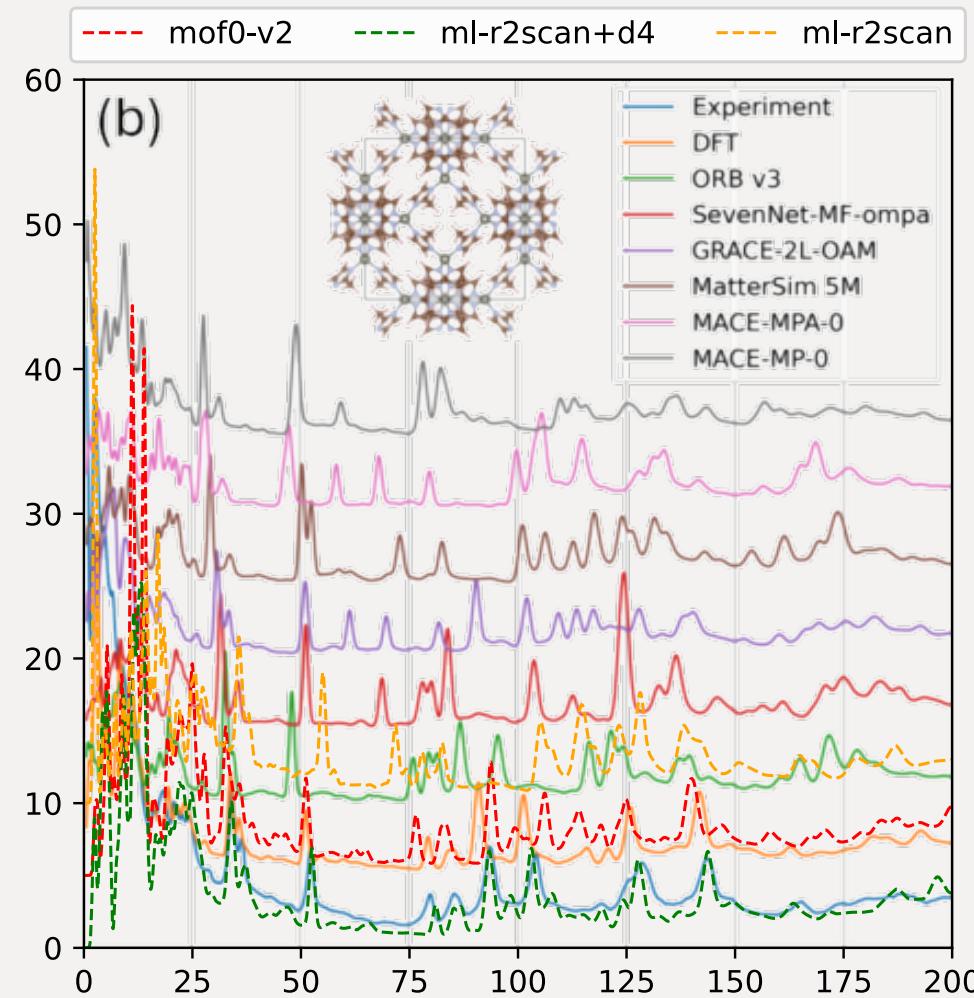
MOF-5 INS

Zhou, W et al, Phys. Rev. B 74, 180301(R), 2006





what about zif-8?





introduction

Theoretical era

Simulations era

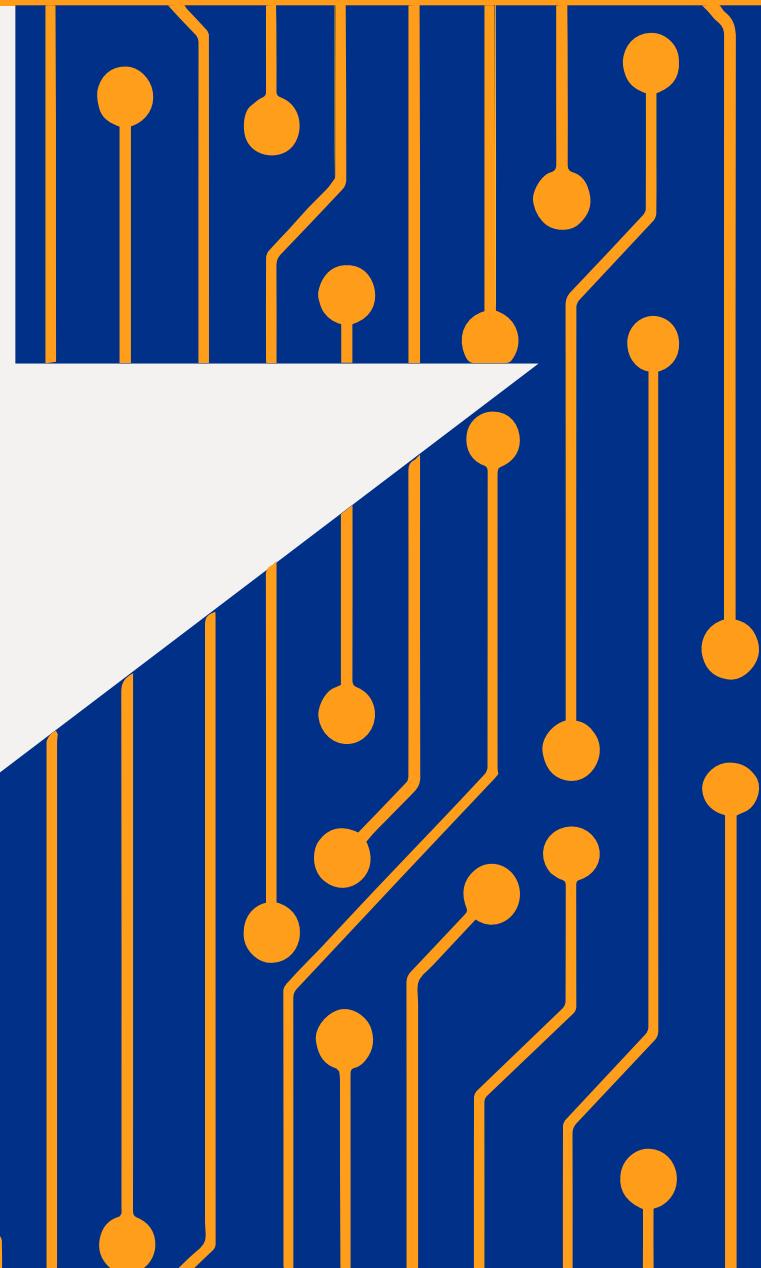
ML era

metal organic frameworks

2,5-Diiodothiophene (pressure)

zeolites

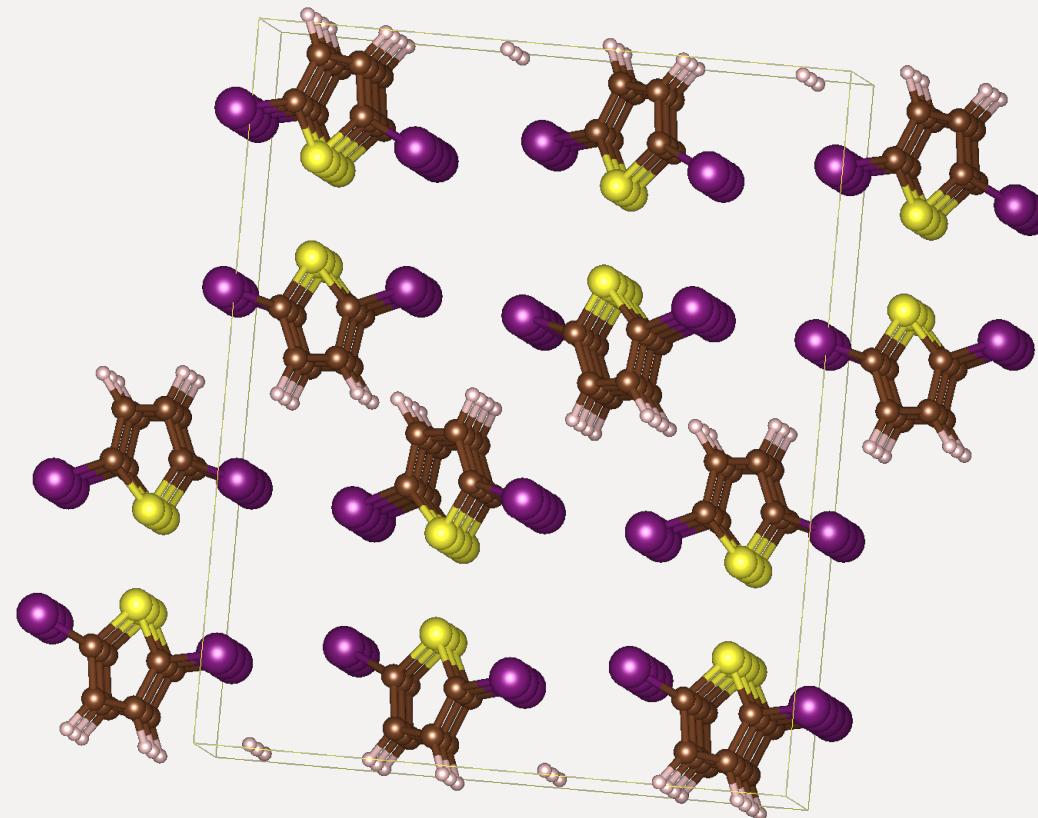
performance





2,5-Diiodothiophene

Parker S, et al., J. Phys. Chem. C 2017, 121, 23, 12636-12642



geometry



pressure 0 GPa

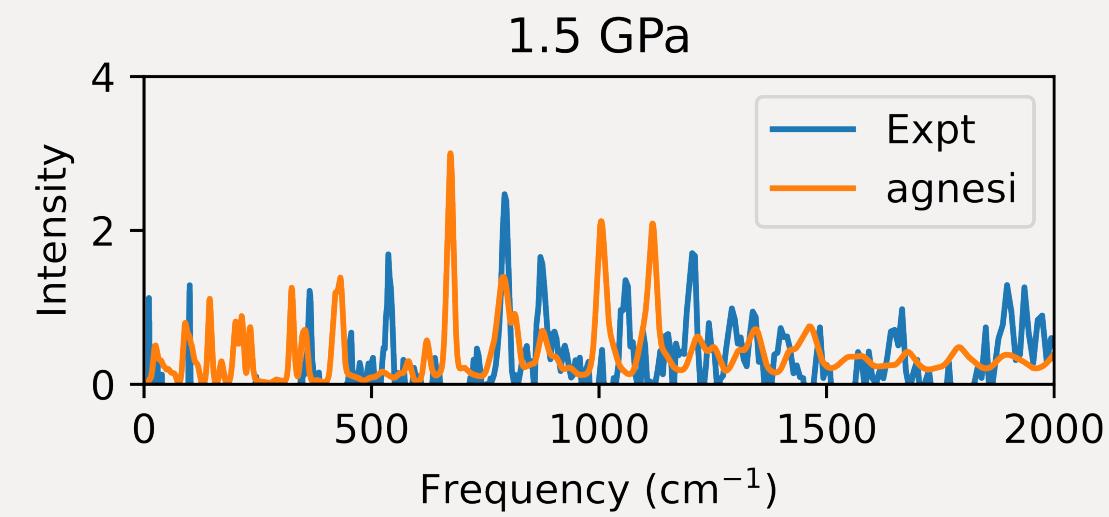
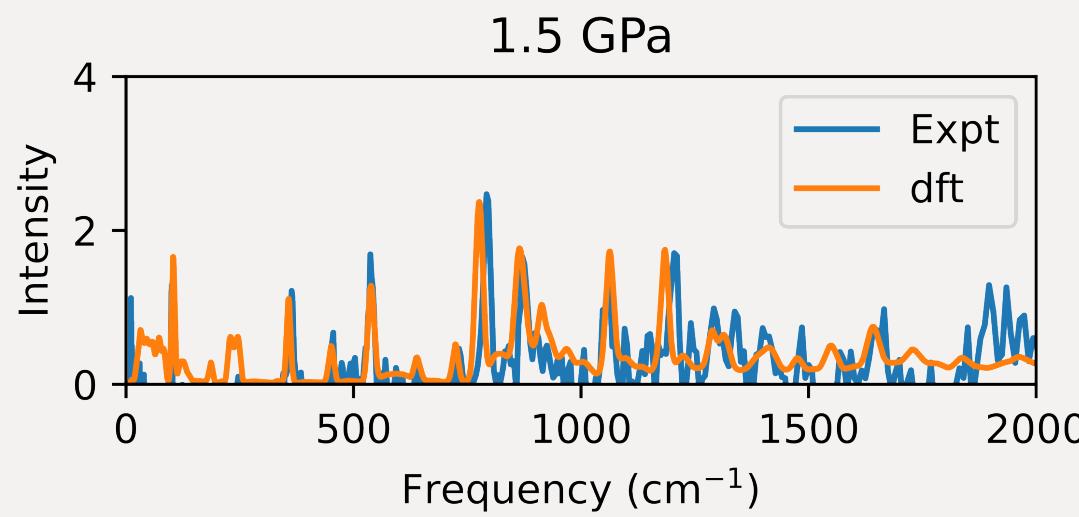
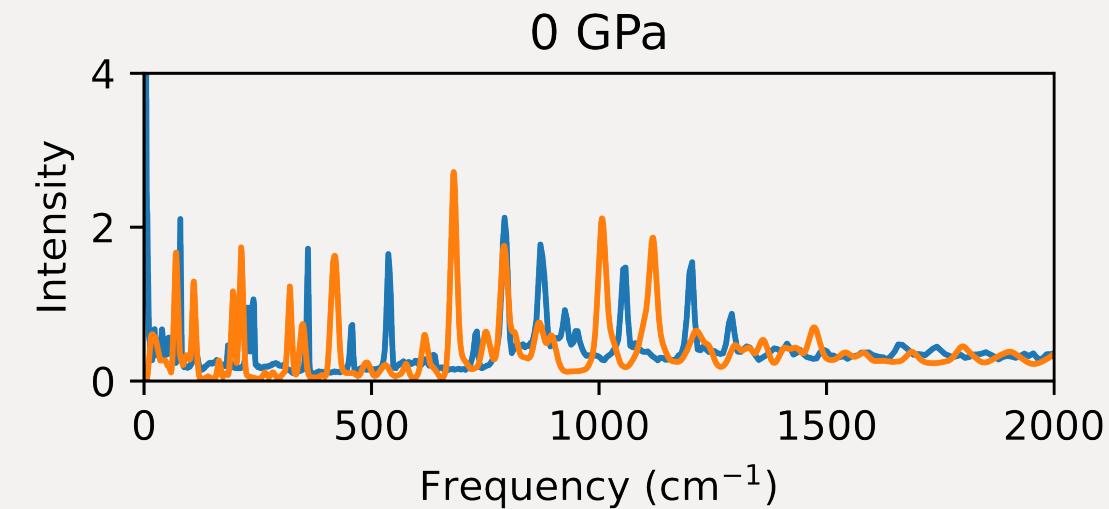
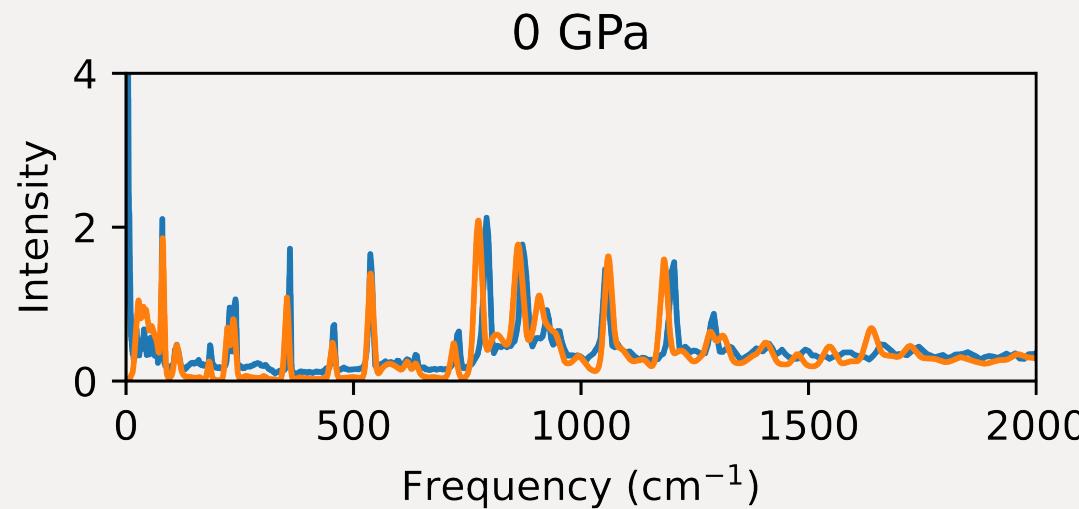
cell	a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
exp	14.944	18.294	5.254	90.0	90.0	90.0
DFT	14.895	18.123	5.136	90.0	90.0	90.0
ML base	12.705	18.236	5.702	90.0	90.0	90.0
ML FT	14.915	18.114	5.129	90.0	90.0	90.0

pressure 1.5 GPa

cell	a [Å]	b [Å]	c [Å]	α [°]	β [°]	γ [°]
exp						
DFT	14.310	17.322	5.008	90.0	90.0	90.0
ML base	11.872	17.329	5.723	90.0	90.0	90.0
ML FT	14.358	17.335	4.989	90.0	90.0	90.0



INS spectra

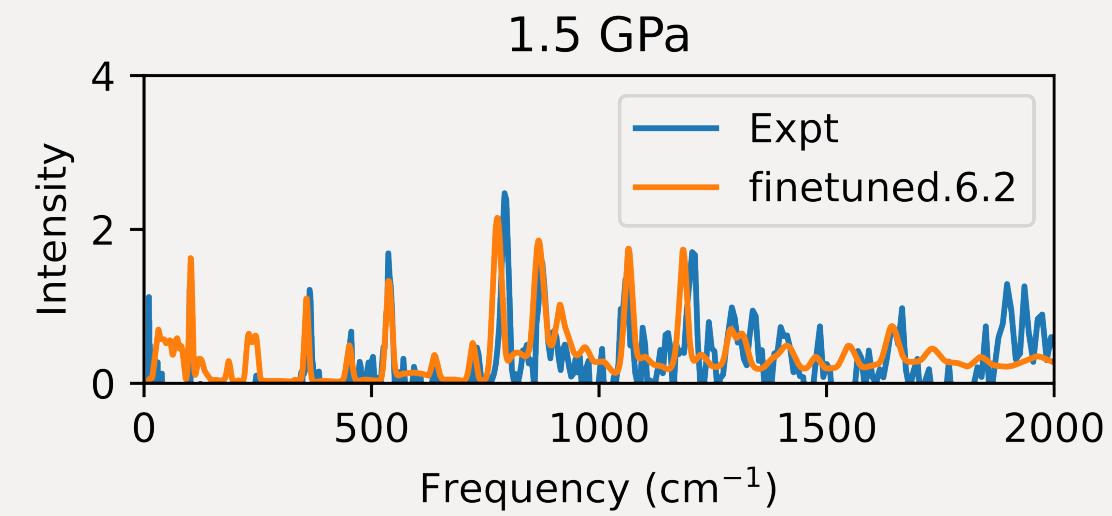
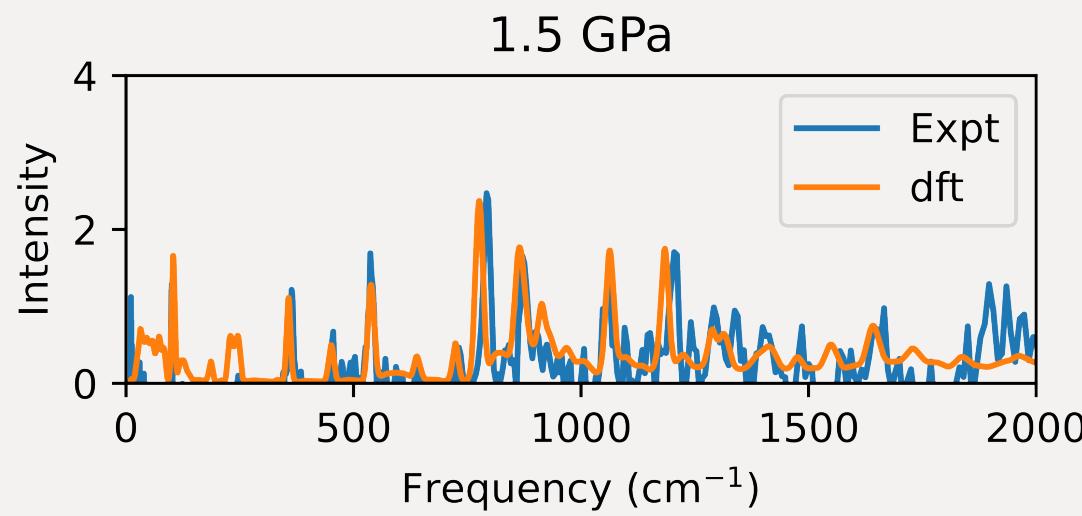
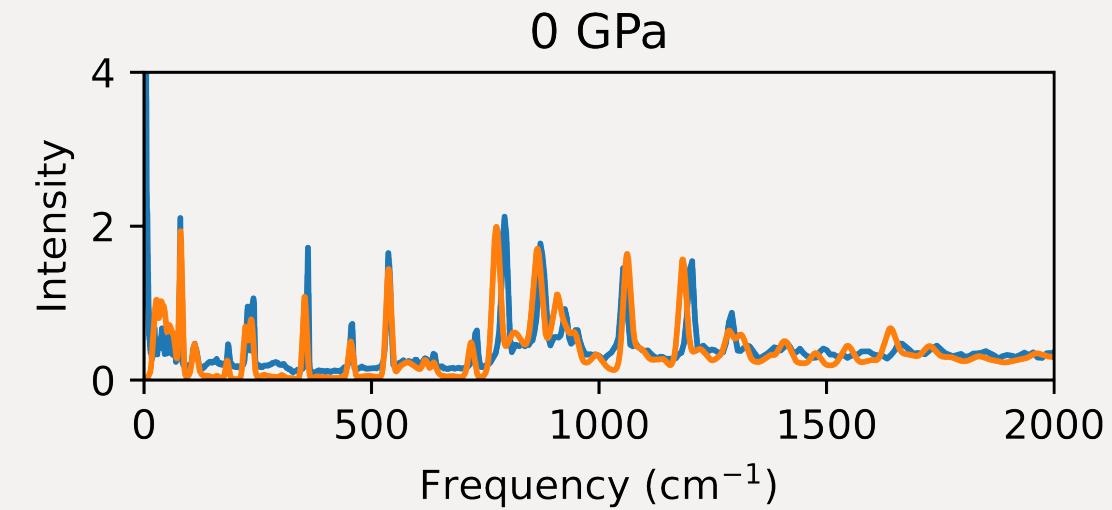
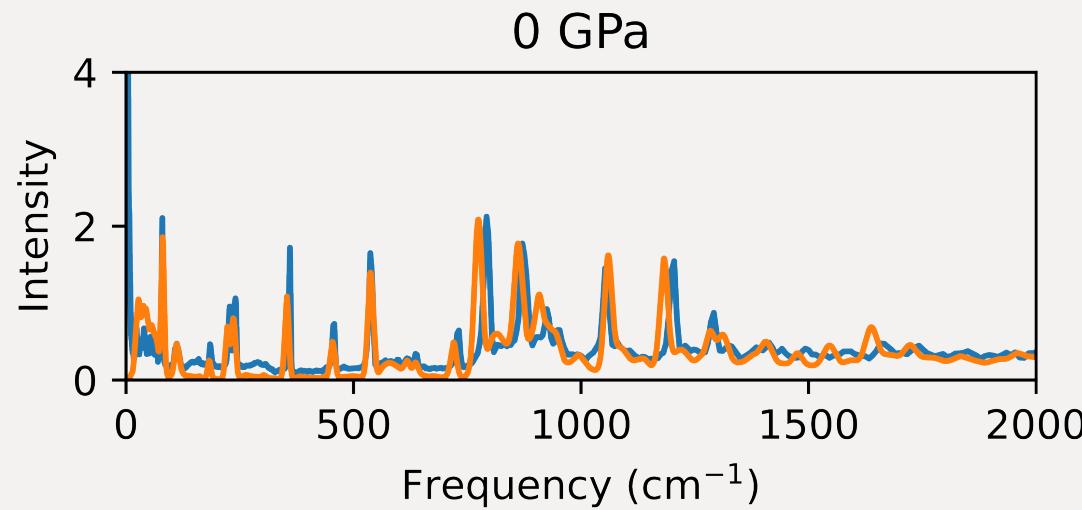


DFT

MACE-MP-0b3



fine tuned



DFT

MACE-MP-0 ft



introduction

Theoretical era

Simulations era

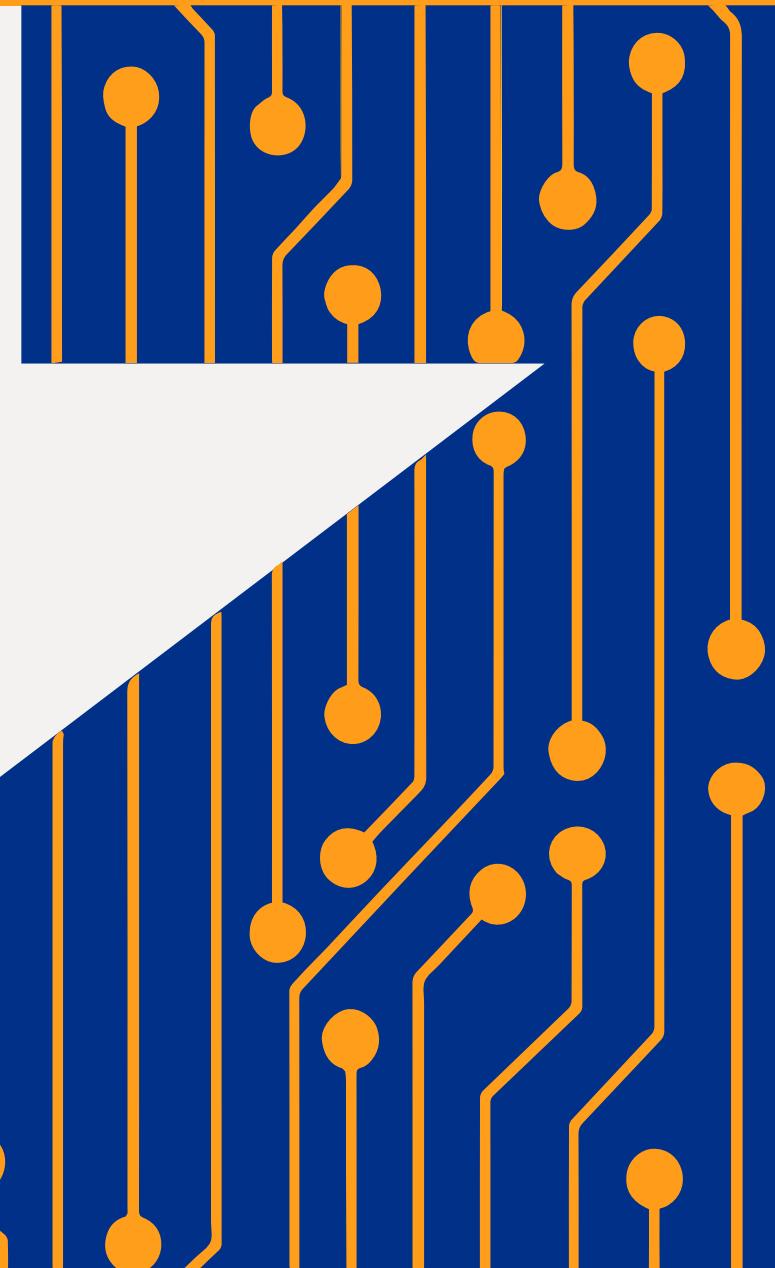
ML era

metal organic frameworks

2,5-Diiodothiophene (pressure)

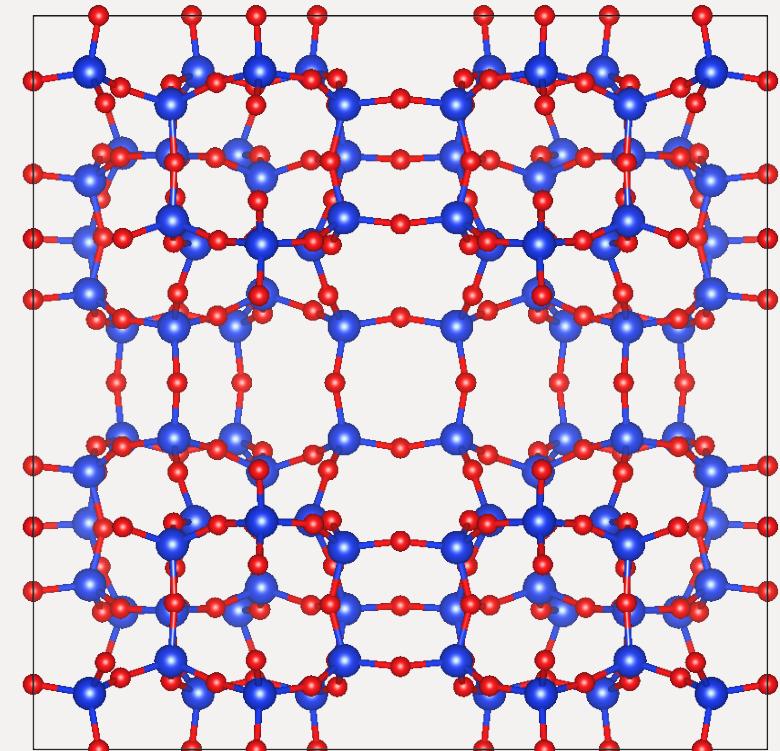
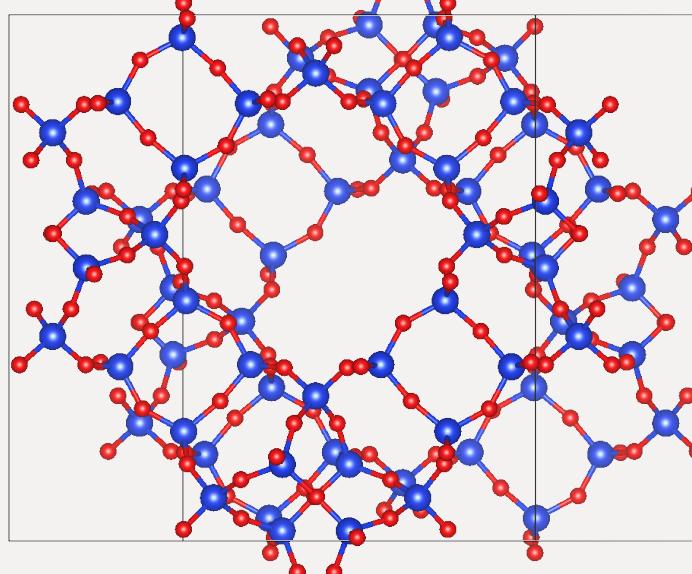
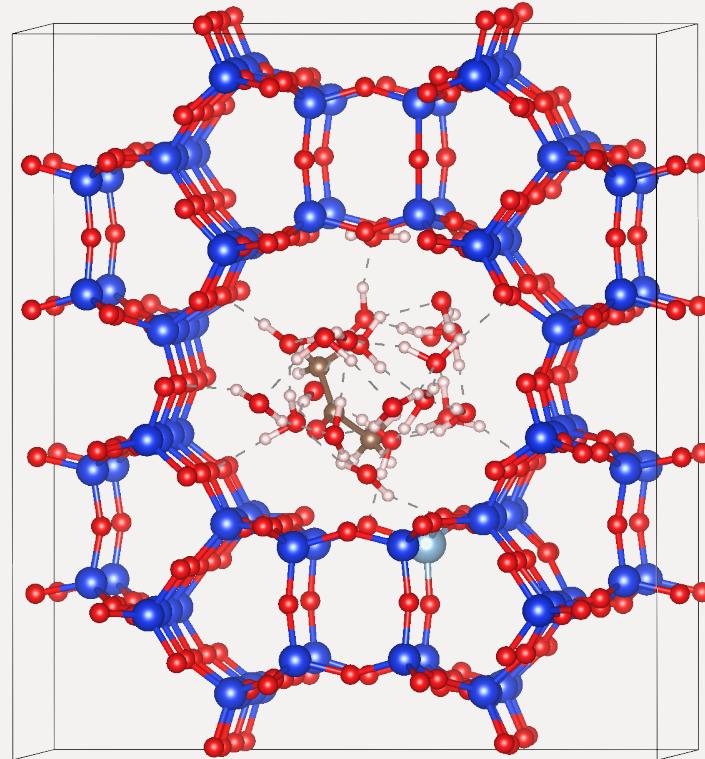
zeolites

performance

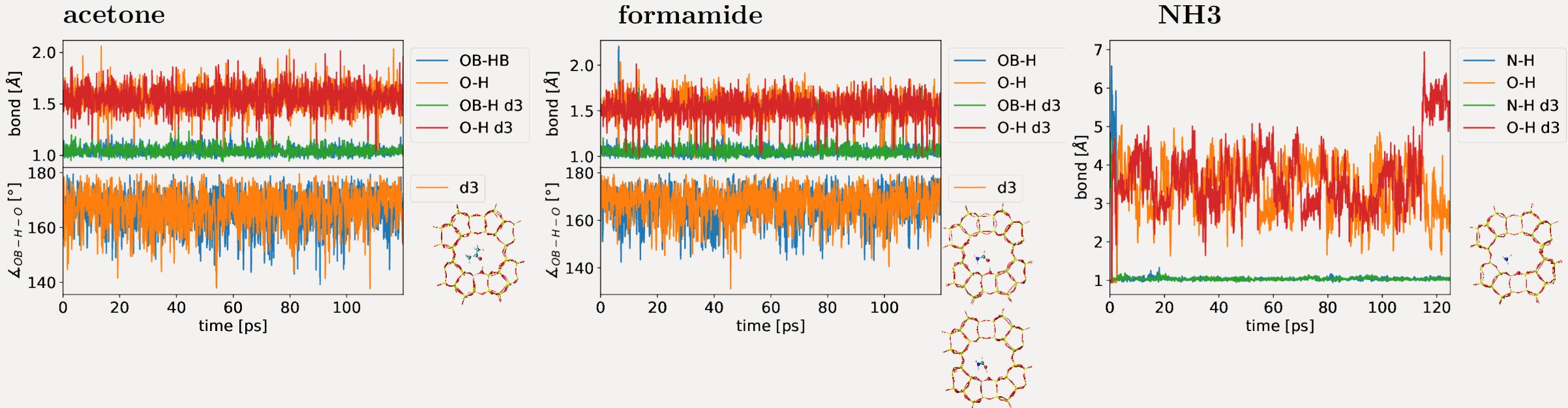




intro

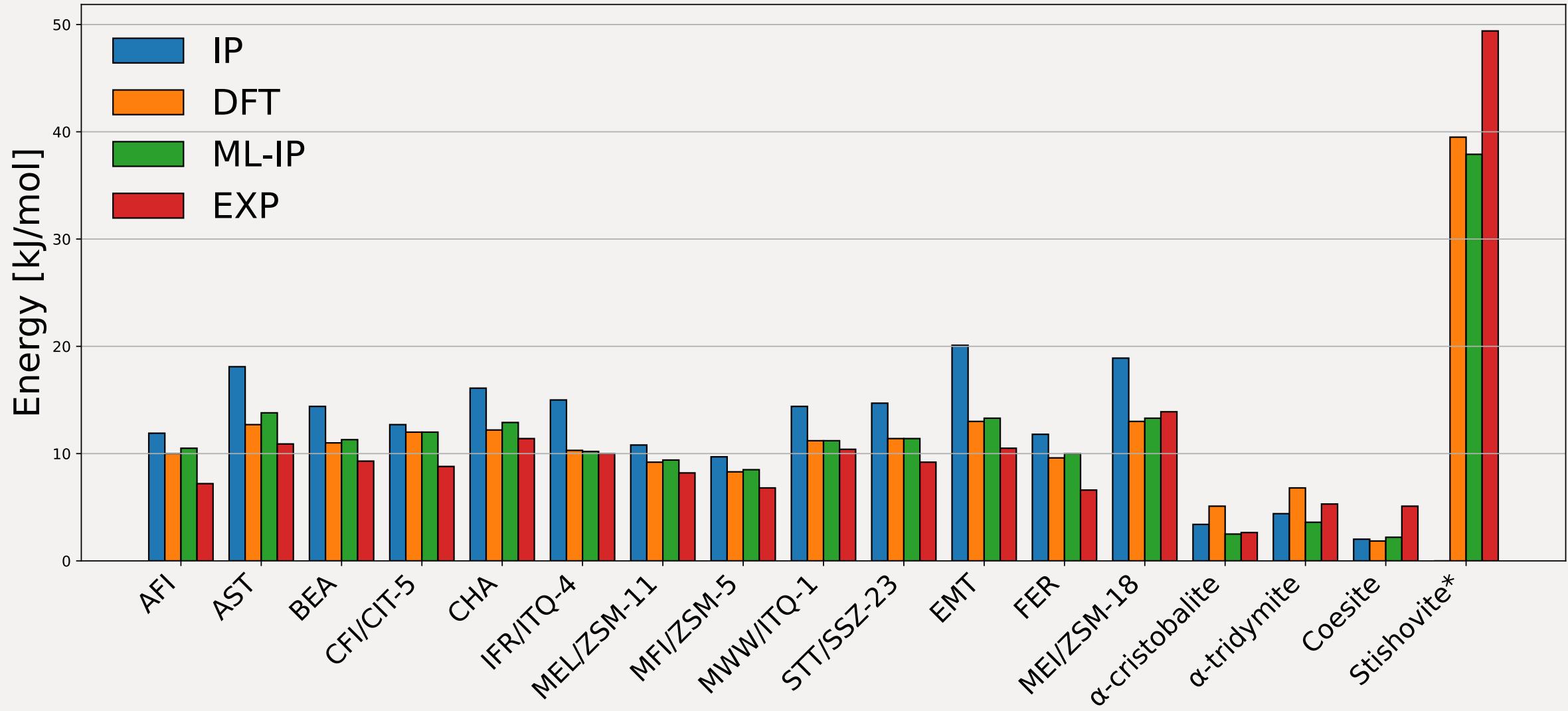


<https://arxiv.org/abs/2411.00436>



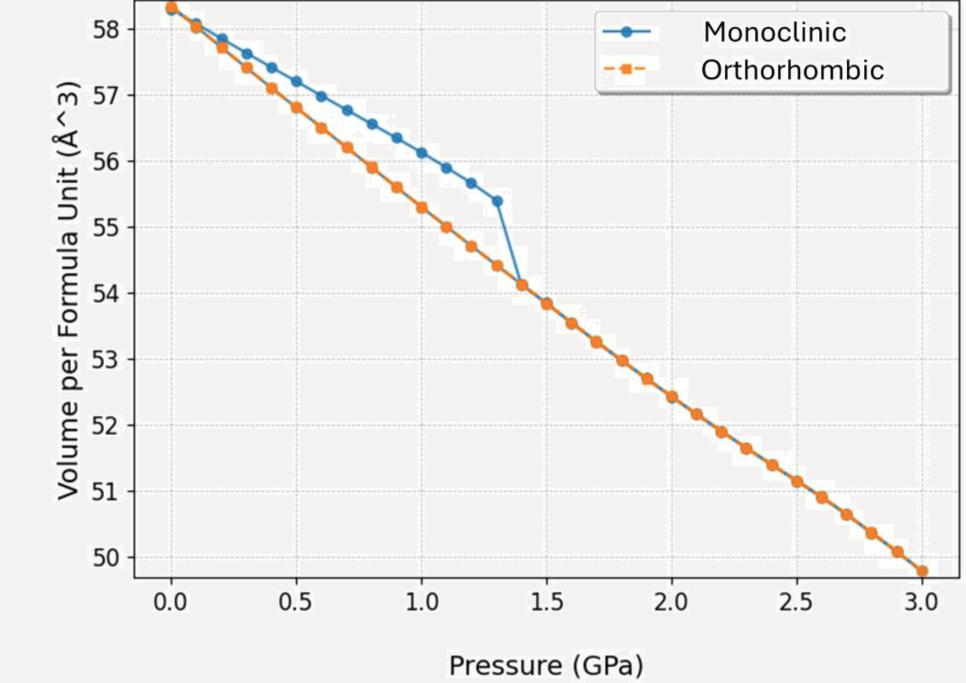
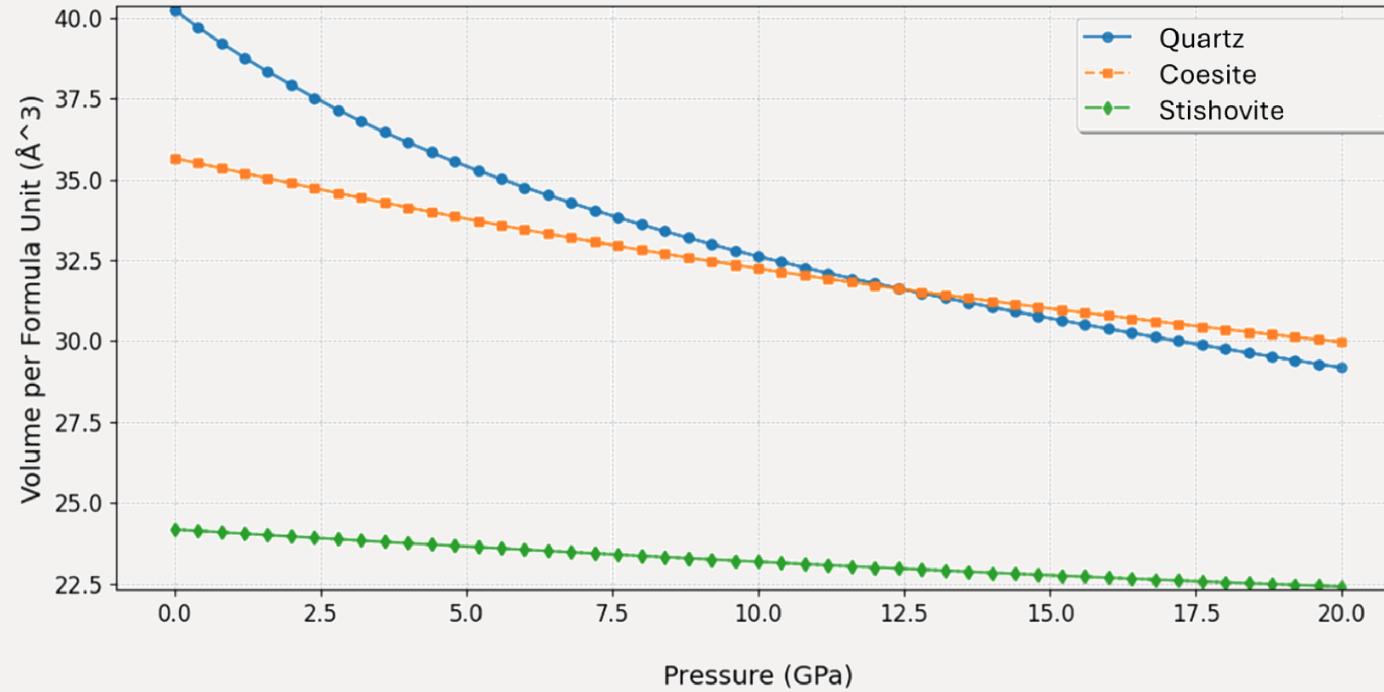


lattice energy





under pressure



ZSM-5/MFI



introduction

Theoretical era

Simulations era

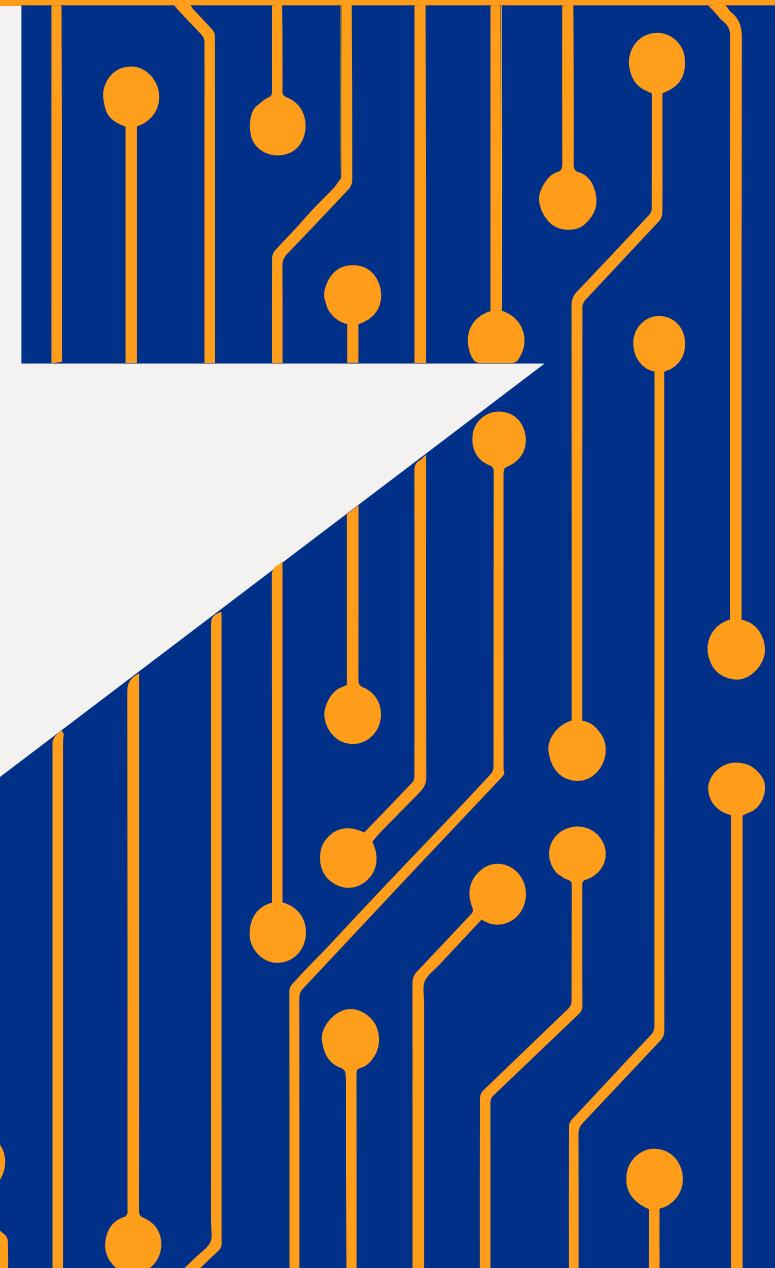
ML era

metal organic frameworks

2,5-Diiodothiophene (pressure)

zeolites

performance



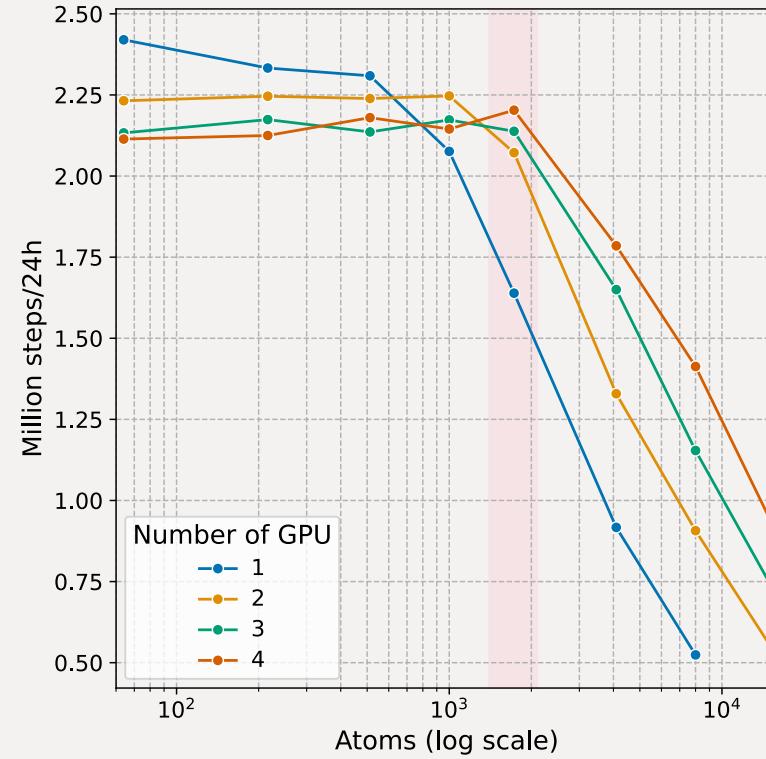


Implementations

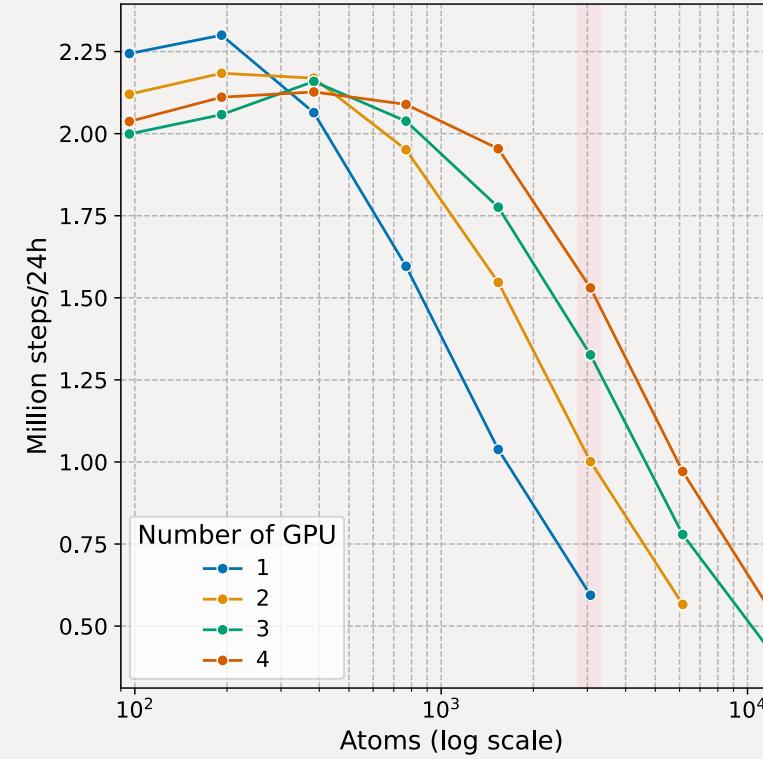
- ▶ python (MACE-pytorch and jax) - CPU, GPU
- ▶ lammps (mliap) GPU
- ▶ symmetrix (available in lammps) - CPU, GPU
- ▶ MACE_MP models only, medium if not specified
- ▶ NaCl 64 - 16000, 11-56Å
- ▶ water 96 - 12288 9.85-49.7Å
- ▶ UiO-66 456-12312 20.75-62.28Å
- ▶ CPU on archer2 - dual AMD EPYC 7742 64-core 2.25GHz
- ▶ GPU on scarf NVIDIA A100



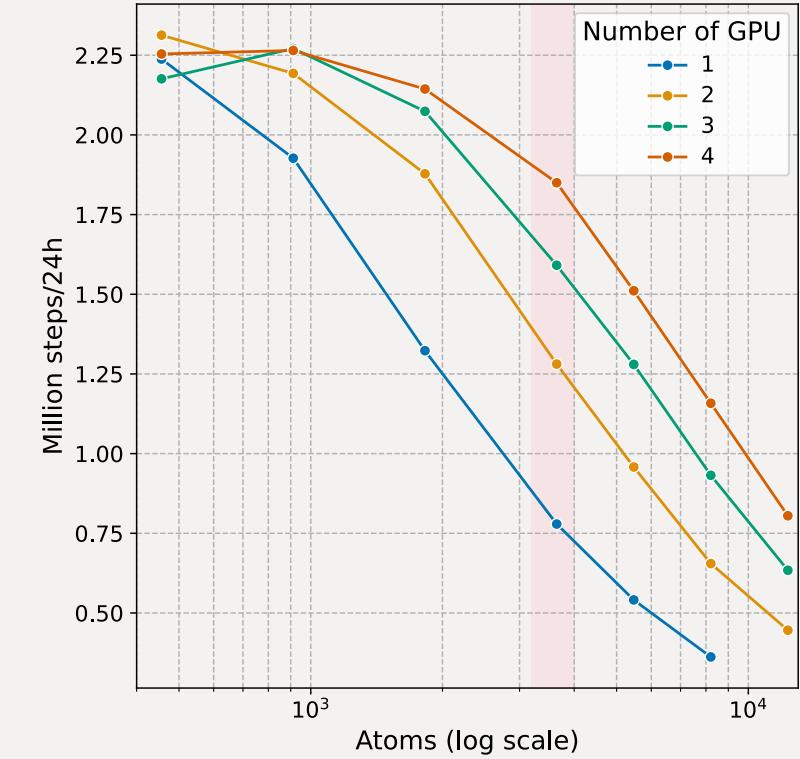
"ns" per day - MLIAP GPU



NaCl - 33.8 Å - 1728 atoms



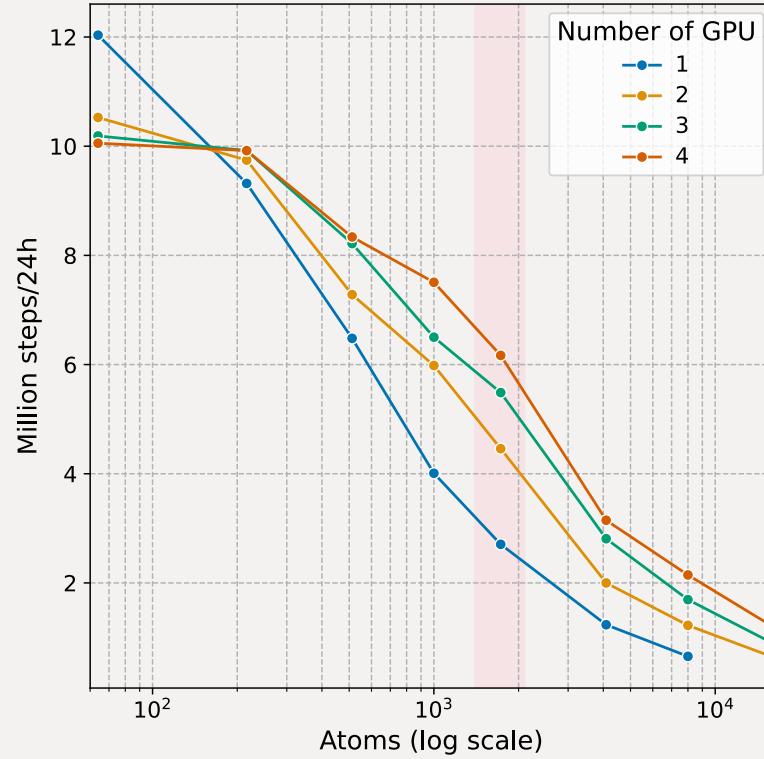
Water - 31.3 Å - 3072 atoms



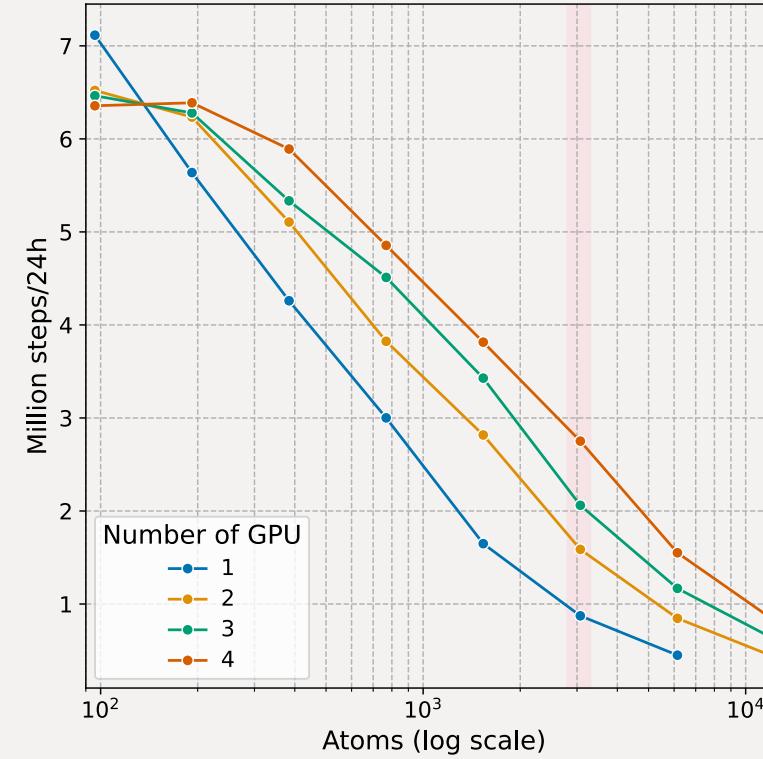
UiO-66 - 41.5 Å - 3648 atoms



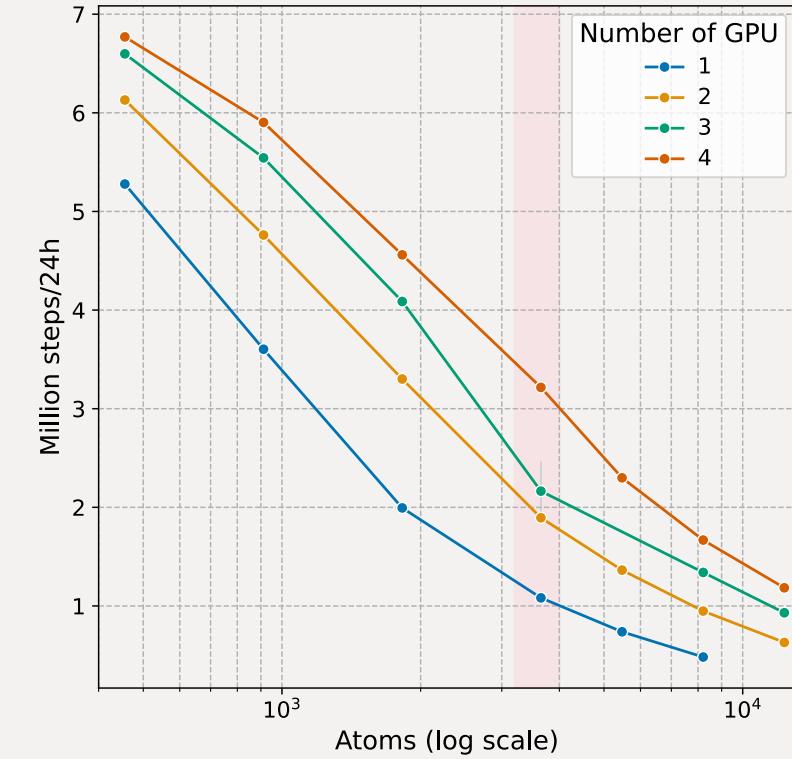
"ns" per day - symmetrix GPU



NaCl - 33.8 Å - 1728 atoms



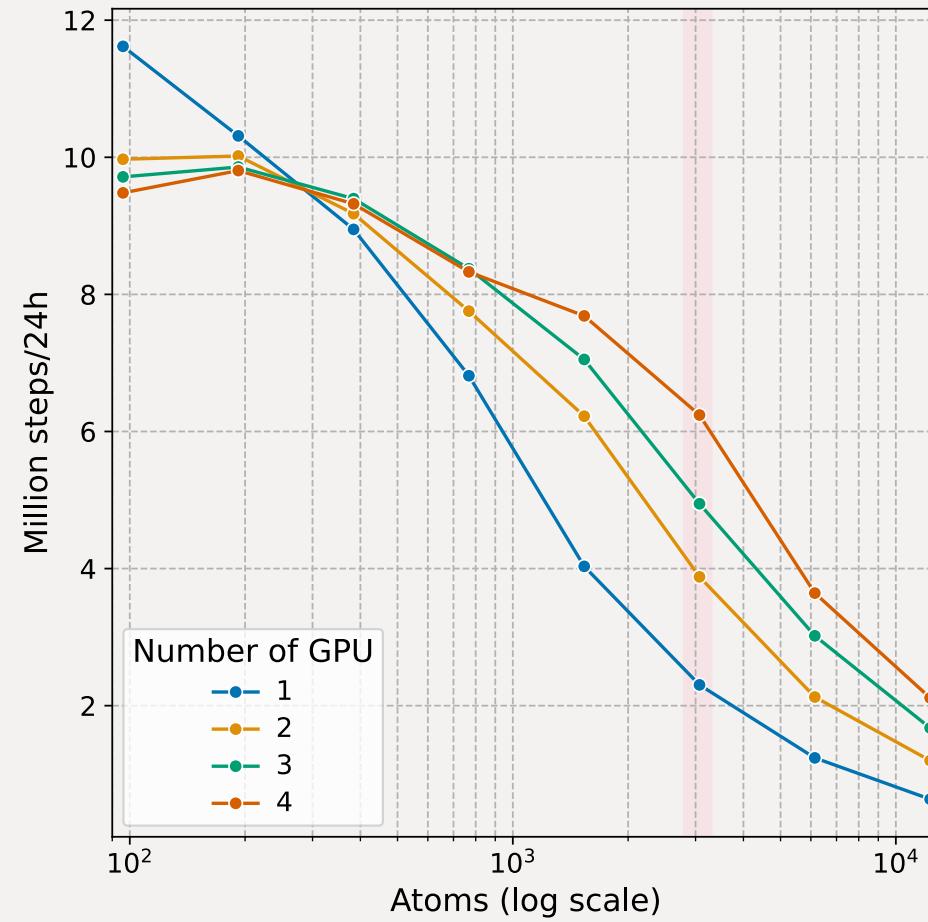
Water - 31.3 Å - 3072 atoms



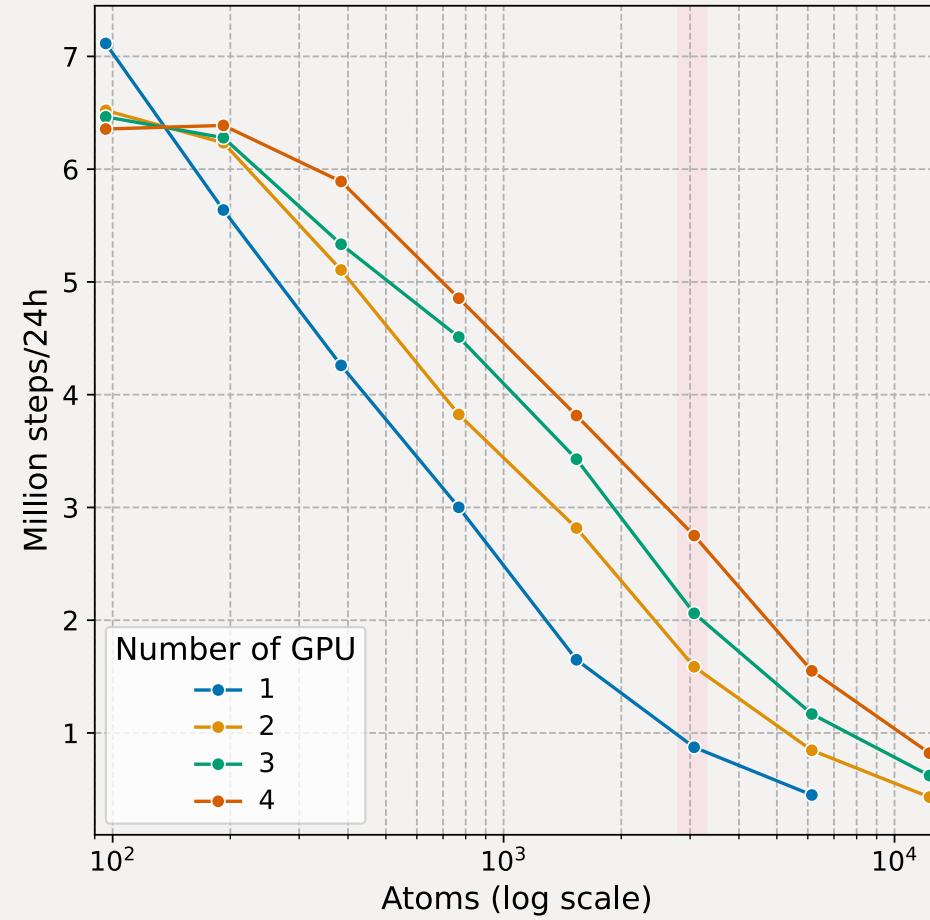
UiO-66 - 41.5 Å - 3648 atoms



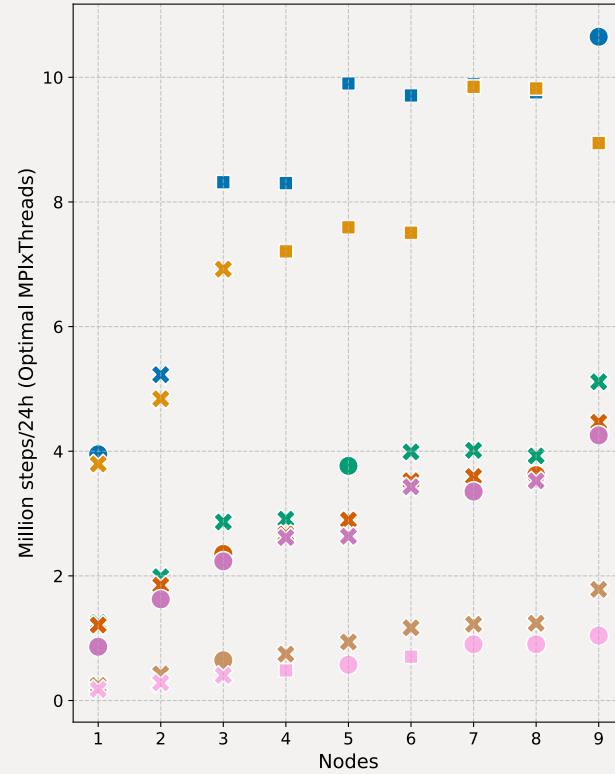
size of the model



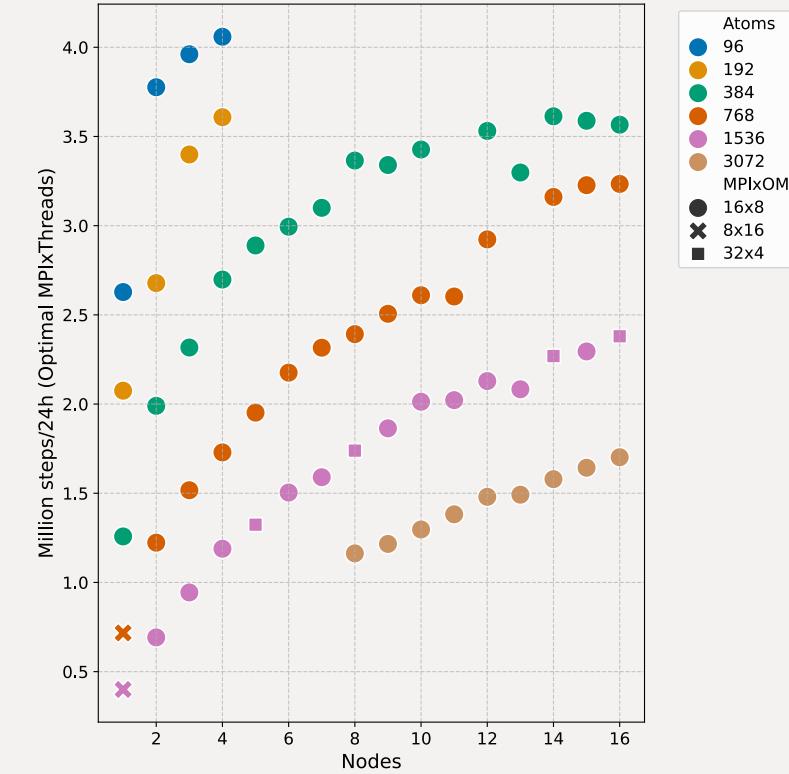
Water - 31.3 Å - small



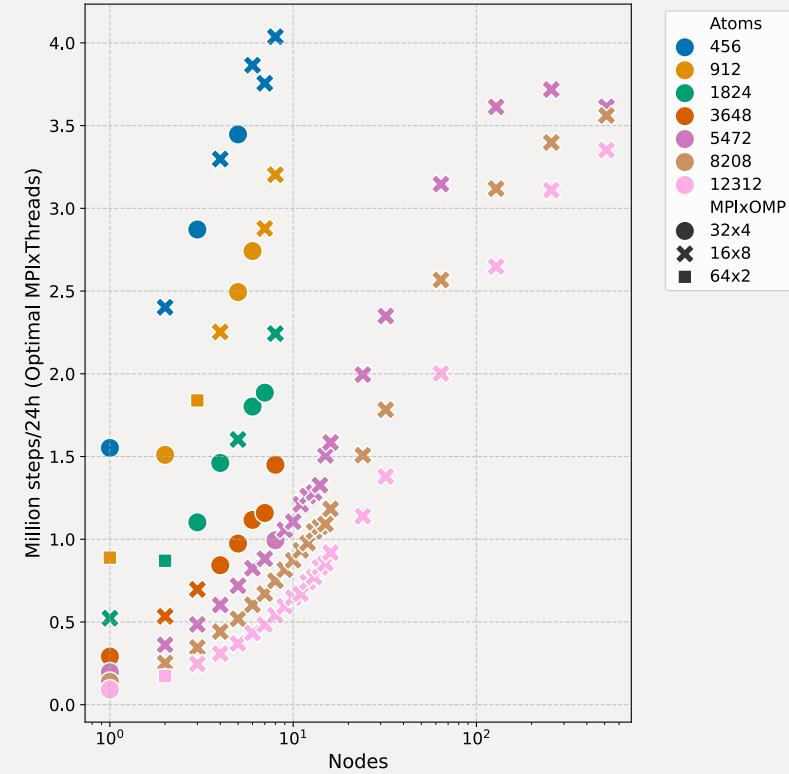
Water - 31.3 Å medium



NaCl - 33.8 Å - 1728 atoms



Water - 31.3 Å - 3072 atoms



UiO-66 - 41.5 Å - 3648 atoms

conclusions

- ▶ DFT or beyond accuracy at a fraction of cost
- ▶ work well outside of the training
- ▶ out of the box MLIPs can be good
- ▶ things are not perfect but we can fix them by fine tuning
- ▶ sizes and speeds enable achieved can enable





Science and
Technology
Facilities Council

Scientific Computing

Questions?





Science and
Technology
Facilities Council

Scientific Computing

Thank you

scd.stfc.ac.uk