Towards ab-initio quality description of porous materials: Developing general Machine-Learned Potentials to simulate physical and adsorption properties of Metal-Organic Frameworks

Yifei Yue [1 2 3] **Saad Aldin Mohamed** [3] **Athuyla S. Palakkal** [3] **N. Duane Loh** [1 2 4] **Jianwen Jiang** [1 3]

[1] *Integrative Sciences and Engineering Programme, National University of Singapore* [2] *Center for Bio-Imaging Sciences, National University of Singapore* [3] *Dept. of Chemical and Biomolecular Engineering, National University of Singapore* [4] *Dept. of Physics, National University of Singapore*. Correspondence to: N.D. Loh duaneloh@u.nus.edu, J. Jiang chejj@u.nus.edu.

1. Introduction

Porous materials, notably Metal-Organic Frameworks (MOFs), are now widely utilized in industrial applications[1]. Design of porous material rely on computational simulations to generate insights into the structure-property relations[1]). Yet, developing accurate and scalable simulations to evaluate the various properties of porous frameworks remains challenging, owing to the limitations of conventional simulation methods; notably, classical force fields (FFs) possess limited accuracy (e.g., in describing hydrogen bonds), whereas ab initio molecular dynamics (AIMD) simulations are prohibitively expensive for large structures. In this regard, advancements in machine-learned potentials (MLPs), and more recently universal MLP (uMLPs)[2], offer improved scalability for ab initio quality simulations that can be generalizable to most elements. This enables accurate and scalable MLP-driven molecular dynamics (MLP-MD) simulations that reliably predict the physical and adsorption properties of MOFs.

Here, we present our recent efforts in developing generalizable MLPs to describe various properties of MOFs[3, 4]. First, we develop a generalizable MLP model for many Zn-based MOFs, and evaluate its prediction performance in computing the physical properties of notable Zn-based MOFs. We show that the trained MLP exhibits good prediction accuracy of physical properties, which affirms it's ability to generalize across many structures[5]. Next, we focus on the application of MLPs to evaluate the water stability and water adsorption properties of STAM (St. Andrew Materials[6]) MOFs, which are especially challenging to describe accurately with classical FFs. Encouragingly, we show that the fine-tuned MLPs can be applied to analyze the hydrolysis mechanism of Cu-paddlewheels, and reliably predict the water adsorption isotherm of STAM MOFs. Overall, we demonstrate the potential of MLPs to reliably simulate various properties of complex porous materials.

2. Results and Discussion

2.1 *Generalizable MLPs for predicting the physical properties of MOFs*

We first investigate how a generalizable MLP model can be developed to reliably predict the physical properties of nearly 3,000 Zn-based MOFs[4]. We begin by retrieving a diverse set of 2,966 Zn-based MOFs from the QMOF database[7]. As shown in **Fig. 1**, these Zn-MOFs contain a diverse set of inorganic secondary building units (SBUs) and linker groups. However, we observe the prevalence of common inorganic and organic SBUs, including the tetrahedral Zn node and the imidazolate linker in ZIF-1 (**Fig. 1a**), and the relevant Zn_4O cluster and the BPDC linker in IRMOF-10 (**Fig. 1b**). These observations indicate that the majority of Zn-MOFs contain similar chemical moieties (in **Fig. 1d-e**) that may be potentially described by one generalizable MLP.

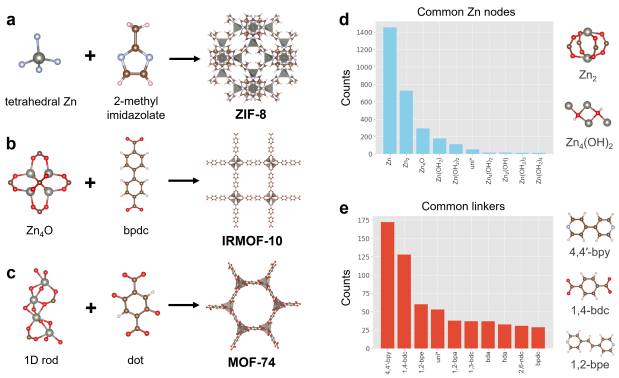

Fig. 1: **Building blocks of Zn-based MOFs share similar moieties.** (a) Tetrahedral Zn node and 2-methyl-imidazolate linker in ZIF-8, (b) Zn_4O cluster and BPDC linker in IRMOF-10, (c) 1D Zn rod and DOT linker in MOF-74. Distribution of common (d) Zn nodes and (e) linkers detected among 2,966 Zn-MOFs in the QMOF database[7]. Unidentified (uni*) indicates few structures whose nodes and linkers are undetected.

Next, we trained an generalizable MLP model using a curated ab initio dataset. The ab initio dataset was generated via DFT calculations of the pertubed structures (to mimic structures under vibrations) of these Zn-MOFs[4]. Subsequently, we leveraged the equvariant `nequIP`[2] model to learn interatomic potentials. **Fig. 2** presents our validation of the accuracy of the MLP in predicting the various physical properties, namely phonon densities of state (DOS), of notable Zn-MOFs. Notably, we affirm that the trained MLP exhibits reasonable accuracy in its prediction of phonon DOS and thermal properties (inset in **Fig. 2**). Although we find that the MLP tends to generalize poorly for a few Zn-MOFs containing rare chemical moieties, it displays greater accuracy in the prediction of phonon DOS compared to classical FFs [4].

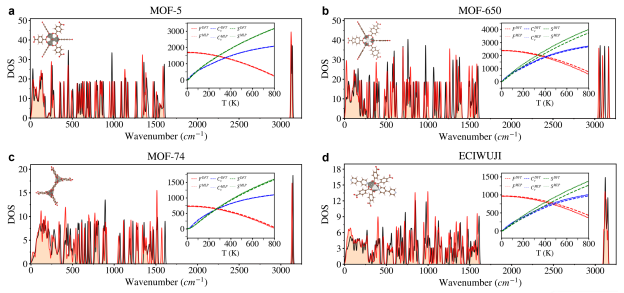

Fig. 2: **Validation of phonon DOS for representative structures containing common inorganic Zn nodes**. Comparison of phonon DOS predicted by the MLP (`nequIP`[2] at $l = 2$) and DFT calculations (**PBE+D3**) for structures containing prevalent Zn building blocks: Zn_4O cluster in (a) MOF-5 and (b) MOF-650, 1D rod in (c) MOF-74, Zn paddlewheel in (d) ECIWUJI. The inorganic Cu paddlewheels are illustrated in the top left, and the inserted subpanels show the comparison between thermal properties as predicted by DFT (--) and MLP (···), including the free energy F, heat capacity C_v and entropy S.

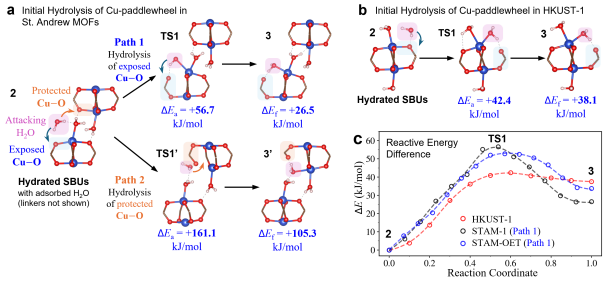

Fig. 3: **Hydrolysis mechanism of Cu-paddlewheels in STAMs and HKUST-1.** (a) Proposed initial hydrolysis step of hydrated (labeled 2) to defective Cu paddlewheel with one hydrolyzed Cu-O bond (labeled 3) in STAMs. Two hydrolysis paths (Path 1 and Path 2) are proposed based on the Cu-O bond that an approaching H_2O can potentially hydrolyze. The transition state (TS1 or TS1') and final state (3 or 3') are labeled with associated activation energy barrier (ΔE_a) and final energy change (ΔE_f), respectively. (c) Same initial hydrolysis step in HKUST-1. (d) ΔE profiles for Path 1 in initial hydrolysis (from 2 to 3) in STAMs and HKUST-1.

2.2 Predicting the water stability and adsorption properties of STAM MOFs

Having demonstrated the efficacy of MLP in reliably predicting the physical properties of many MOFs, we now examine how MLP can be applied to probe the hydrolytic properties (namely water stability and water adsorption) of STAM MOFs[6, 3]. Although improving the water stability and adsorption properties of MOFs is of significant experimental interest, accurately capturing these properties with classical force fields remains challenging due to complex water-MOF interactions. To train MLPs, we employ an active-learning (AL) strategy to iteratively fine-tune MLPs to simulate hydrolytic properties in five STAM MOFs containing various functionalized linkers[3].

Next, we apply the fine-tuned MLP to probe the hydrolysis mechanism of Cu-paddlewheel in STAMs and HKUST-1 (**Fig 3**). To date, a mechanistic understanding of Cu-paddlewheel decomposition MOFs remains unexplored via ab initio calculations. We propose that the hydrolysis of the Cu-paddlewheel units proceeds via two paths in STAM-1 (in **Fig 3a**): an approaching H_2O can either attack the (1) exposed Cu-O bond or (2) the protected Cu-O bond. By applying MLP-driven CI-NEB calculations, the activation energy barrier ΔE_a associated with Path 1 was evaluated to be 56.7 kJ/mol (TS1 in **Fig 3a**), below than 161.1 kJ/mol in Path 2 (TS1'). Importantly, the ΔE_a of Cu-O hydrolysis in HKUST-1 is significantly lower than in STAMs in both paths (**Fig 3c**). This can be attributed to the greater steric hindrance from the stacking Cu-paddlewheels in STAMs, which explains the superior water-stability of STAM MOFs[6] observed in experiments.

Besides water stability, it is important to elucidate the effect of functional groups on water adsorption in STAMs. We then employed MLP-driven grand canonical monte-carlo (GCMC) simulations to evaluate the water adsorption properties of STAMs at 298 K (**Fig. 4**). Crucially, MLP-GCMC simulation reliably reproduces the experimental water adsorption isotherm in STAM-1 and STAM-OET[6] (**Fig. 4**). Moreover, we find that linker functionalization with hydrophilic groups (e.g., amide) generally enhances water uptake (**Fig. 4b**), as the hydrophilic groups act as water seeding sites that favorably enhance water adsorption in the hexagonal pores (**Fig. 4c**).

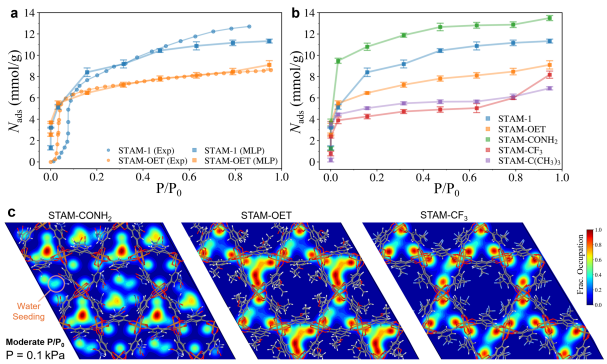

Fig. 4: **Water adsorption in STAM MOFs from MLP-GCMC simulations.** (a) Water adsorption isotherms from MLP-GCMC simulations and experiments in pristine STAM-1[6] and STAM-OET[8] at 298 K. (b) Water adsorption isotherms from MLP-GCMC simulations in pristine STAMs with different functional groups at 298 K. (c) Fractional occupations of adsorption sites in STAM MOFs at 0.1 kPa.

In summary, we demonstrate how generalizable MLPs can be developed to accurately describe physical and adsorption properties of MOFs. We expect that MLPs can be extended to simulate complex processes (e.g., ion transport) in porous materials.

Acknowledgments

The authors acknowledge financial support provided by NUS HPC (NUSREC-HPC-00001 and CFP01-CF-077) and the Center of Biological Imaging Sciences (CBIS). We thank Bai Chang and Wang Junhong for HPC support.

References

[1] Peter G Boyd, Arunraj Chidambaram, Enrique García-Díez, Christopher P Ireland, Thomas D Daff, Richard Bounds, Andrzej Gładysiak, Pascal Schouwink, Seyed Mohamad Moosavi, M Mercedes Maroto-Valer, and et al. Data-driven design of metal–organic frameworks for wet flue gas co2 capture. *Nature*, 576(7786):253–256, 2019.

[2] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.*, 13(1):2453, 2022.

[3] Yifei Yue, Athulya S Palakkal, N Duane Loh, and Jianwen Jiang. Are hemilabile metal–organic frameworks overlooked as promising water-stable adsorbents? elucidating their physical and hydrolytic properties using machine-learned potentials. *J. Am. Chem. Soc.*, 147(40):36180–36196, 2025.

[4] Yifei Yue, Saad Aldin Mohamed, N Duane Loh, and Jianwen Jiang. Toward a generalizable machine-learned potential for metal–organic frameworks. *ACS Nano*, 19(1):933–949, 2025.

[5] Alin Marin Elena, Prathami Divakar Kamath, Théo Jaffrelot Inizan, Andrew S Rosen, Federica Zanca, and Kristin A Persson. Machine learned potential for high-throughput phonon calculations of metal—organic frameworks. *npj Comput. Mater.*, 11(1):125, 2025.

[6] M Infas H Mohideen, Bo Xiao, Paul S Wheatley, Alistair C McKinlay, Yang Li, Alexandra MZ Slawin, David W Aldous, Naomi F Cessford, Tina Düren, and Xuebo Zhao. Protecting group and switchable pore-discriminating adsorption properties of a hydrophilic–hydrophobic metal–organic framework. *Nat. Chem.*, 3(4):304–310, 2011.

[7] Andrew S Rosen, Shaelyn M Iyer, Debmalya Ray, Zhenpeng Yao, Alan Aspuru-Guzik, Laura Gagliardi, Justin M Notestein, and Randall Q Snurr. Machine learning the quantum-chemical properties of metal–organic frameworks for accelerated materials discovery. *Matter*, 4(5):1578–1597, 2021.

[8] Lauren N McHugh, Matthew J McPherson, Laura J McCormick, Samuel A Morris, Paul S Wheatley, Simon J Teat, David McKay, Daniel M Dawson, Charlotte EF Sansome, and Sharon E Ashbrook. Hydrolytic stability in hemilabile metal–organic frameworks. *Nat. Chem.*, 10(11):1096–1102, 2018.