# Predicting hydrogen storage in nanoporous materials using meta-learning

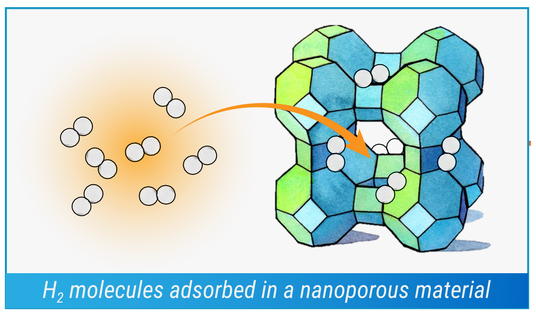
Yangzesheng Sun, Robert F. DeJaco and J. Ilja Siepmann

**Department of Chemistry & Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis MN, USA**

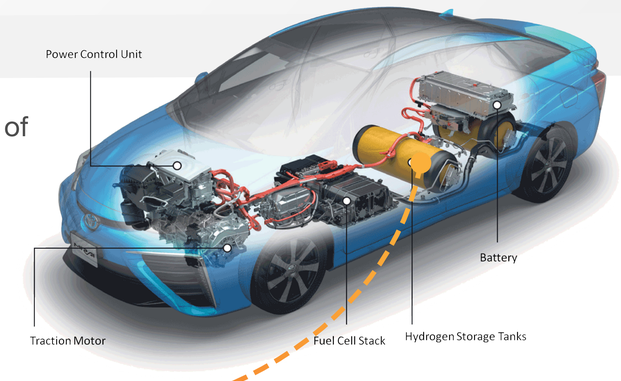
## Introduction

**Hydrogen vehicles** combine advantages of traditional engines and electric motors[1]

$H_2$ is compressed at **70 MPa** in production models[2]

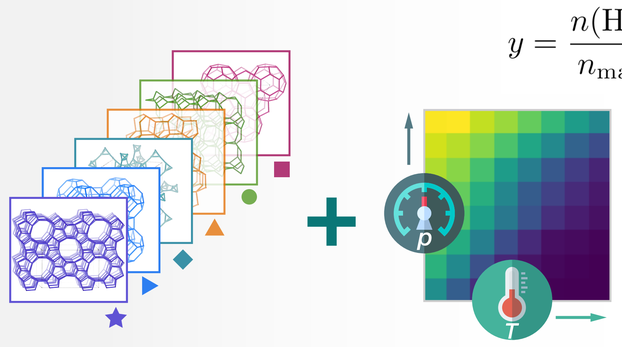


*$H_2$ molecules adsorbed in a nanoporous material*




Adsorptive storage of $H_2$: [3]
**Much lower** pressures (< 10 MPa)
**Better flexibility** of fuel tank
Requires **cryogenic temperatures** (77–200 K)

Adsorption depends on **material structure** and **thermodynamic state**[4]




$$y = \frac{n(\mathrm{H_2})}{n_{\max}} = \frac{A\mathrm{e}^{\frac{B}{T}}p^m}{1 + A\mathrm{e}^{\frac{B}{T}}p^m} = \frac{1}{1 + \mathrm{e}^{-(m\ln p + B\cdot\frac{1}{T} + \ln A)}}$$

Sips adsorption isotherm[4] (equation of state / EOS)

Logistic regression

**Physical model** echoes **machine learning**

## Methods

TRADITIONAL WORKFLOW

**Base learning on each material**

$D_{\blacksquare} = \{n, p, T\}$ ⇄ $f(p,T;\theta_{\blacksquare})$

$D_{\bullet} = \{n, p, T\}$ ⇄ $f(p,T;\theta_{\bullet})$

$D_{\blacktriangle} = \{n, p, T\}$ ⇄ $f(p,T;\theta_{\blacktriangle})$

$D_{\bigstar} = \{n, p, T\}$ ⇄ $f(p,T;\theta_{\bigstar})$

$x_{\text{test}} = (p, T)$ → $n(\mathrm{H_2})$

Molecular simulations generate **many small datasets**
$H_2$ adsorption for **each material modeled independently**[5]
Predicting for a new material **cannot use previous information**
Difficult with limited data

OUR CONTRIBUTION

**Meta-learning on all materials**

$D_{\blacksquare} = \{n, p, T\}$

$D_{\bullet} = \{n, p, T\}$

$D_{\blacktriangle} = \{n, p, T\}$

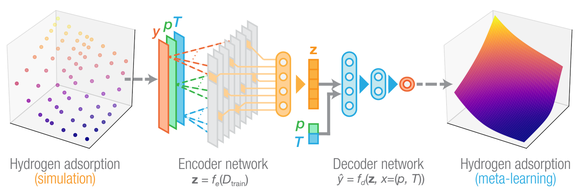
$D_{\bigstar} = \{n, p, T\}$

$f(p,T,D;\theta)$

$x_{\text{test}} = (p, T)$ → $n(\mathrm{H_2})$

$H_2$ adsorption for all materials **given by the same meta-learner**
Predicting for a new material can **benefit from all simulation data**

### Model architecture




**Encoder:**
compresses the example adsorption data into a **fingerprint representation**

**Decoder:**
**predicts adsorption** using the fingerprint and the query state point

### Training

$$y = f(D_{\text{train}}, \mathbf{x}) \longleftarrow L_\theta(D) = \sum_{(\mathbf{x}_i, y_i)\in D} [y_i - f_d(f_e(D_{\text{train}}), \mathbf{x_i})]^2 + \lambda \sum_{i\neq j} \mathrm{cov}[f_e(D_{\text{train}})]_{ij}^2$$

Trained over **a distribution of tasks** (base datasets)[6]
Use **a subset of base dataset** (one material, multiple states) as example data
No inner loop, can be further improved by adding adaptation steps

Does not use features about the material structure
Simulation vs. experiment: **same material, different data**

## References

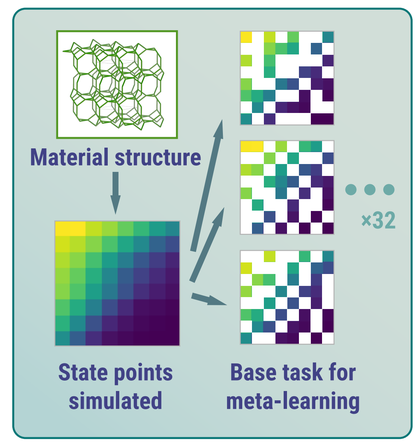
[1] Jacobson, M. Z.; Colella, W. G.; Golden, D. M. Cleaning The Air and Improving Health with Hydrogen Fuel-cell Vehicles. *Science*, **2005**, *308*, 1901–1905.
[2] Yamashita, A.; Kondo, M.; Goto, S.; Ogami, N. Development of High-pressure Hydrogen Storage System for The Toyota "Mirai". SAE 2015 World Congress & Exhibition, **2015**.
[3] Murray, L. J.; Dincă, M.; Long, J. R. Hydrogen Storage in Metal–Organic Frameworks. *Chem. Soc. Rev.* **2009**, *38*, 1294–1314.
[4] Foo, K.; Hameed, B. Insights Into The Modeling of Adsorption Isotherm Systems. *Chem. Eng. J.* **2010**, *156*, 2–10.
[5] Haldoupis, E.; Nair, S.; Sholl, D. S. Finding MOFs for Highly Selective $CO_2/N_2$ Adsorption Using Materials Screening Based on Efficient Assignment of Atomic Point Charges. *J. Am. Chem. Soc.* **2012,** *134,* 4313–4323.
[6] Finn, C.; Abbeel, P.; Levine, S. Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. Proceedings of the 34th International Conference on Machine Learning. **2017**, 1126–1135.
[7] Bai, P.; Jeon, M. Y.; Ren, L.; Knight, C.; Deem, M. W.; Tsapatsis, M.; Siepmann, J. I. Discovery of Optimal Zeolites for Challenging Separations and Chemical Transformations Using Predictive Materials Modeling. *Nat. Commun.* **2015**, *6*, 5912.
[8] Durbin, D.; Malardier-Jugroot, C. Review of Hydrogen Storage Techniques for On-board Vehicle Applications. *Int. J. of Hydrogen Energy* **2013**, *38*, 14595–14617.
[9] Icons credit: www.flaticon.com

## Dataset

Meta-learning regression **usually benchmarked on artificial datasets**
Molecular simulations can provide **a concrete problem**

Gibbs Ensemble Monte Carlo simulations[7] for **211 all-silica zeolites** (porous $SiO_2$ materials with different topological structures)
8 Temperatures: **77.0 K – 275.9 K**
8 Pressures: **0.10 MPa – 40.34 MPa**

64 state points for a material are **subsampled to create training examples** for meta-learning
Meta-training: 75% of zeolites




## Results

### Reconstructing simulation data

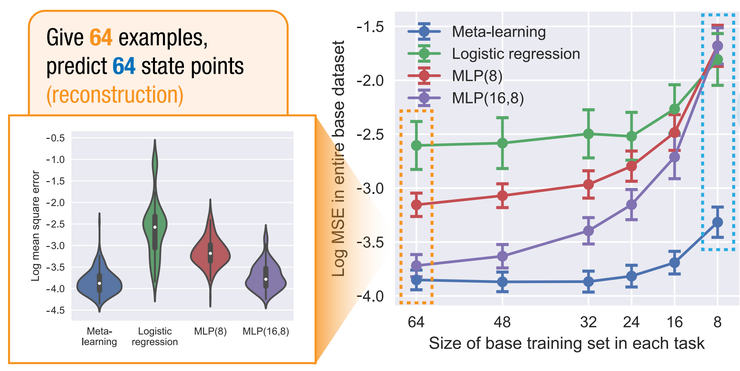
| Model | Geometric Mean MSE ($\times10^{-2}$) | Minimum MSE ($\times10^{-2}$) | Maximum MSE ($\times10^{-2}$) |
|---|---|---|---|
| Logistic regression | 0.249 | 0.014 | 9.858 |
| MLP(8) | 0.070 | 0.006 | 0.705 |
| MLP(16, 8) | 0.019 | 0.006 | **0.184** |
| Meta-learning (train/test) | **0.014/0.014** | **0.004/0.005** | 0.080/0.365 |

**Representations**/model parameters **contain physical properties**

Better **reconstruction**: more accurate **property representation**

### Few-shot learning

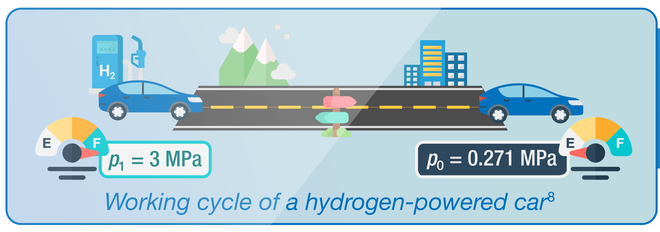
Give **64** examples, predict **64** state points (reconstruction)




Give **8** examples, predict **64** state points (few-shot)

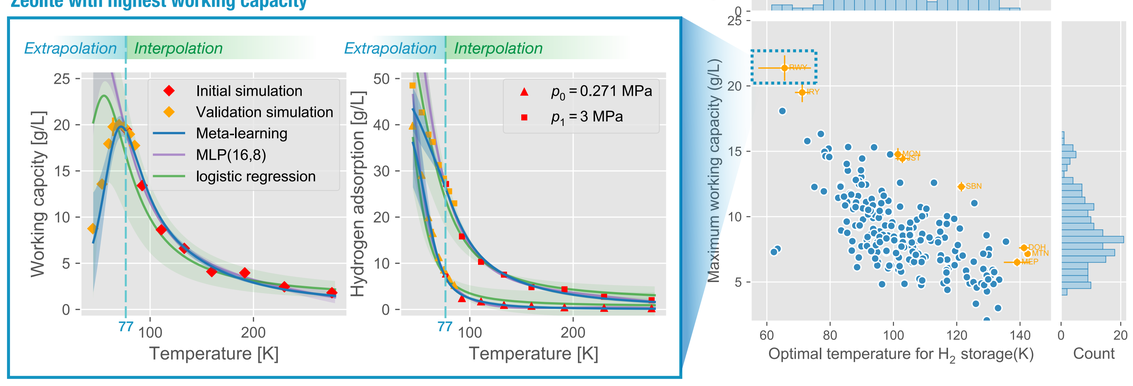
The **same subsampling of state points** were used for all materials
Variation by different subsampling << variation by material
Meta-learning **significantly outperforms physical model & MLP** in few-shot prediction

### Hydrogen storage in vehicles




**Working capacity:**
$n_{\text{work}} = n(\mathrm{H_2}, p_1) - n(\mathrm{H_2}, p_0)$
Each material has a temperature which attains the largest capacity

Zeolite with highest working capacity




Lowest temperature in training set: **77 K**
Meta-learning **achieves good agreement with additional simulations** performed at < 77 K
Extrapolation is valuable: **simulations are more difficult at lower temperatures**

## Conclusions

Meta-learning provides a route towards **efficiently investigating the joint space of material structures and thermodynamic states** for materials discovery problems

Meta-learning achieves **higher few-shot performance and extrapolation ability** compared with independently modeling each material

High-throughput molecular simulations can serve as **real-world regression applications for meta-learning**

## Acknowledgment

*This research is primarily supported by the U.S. Department of Energy (DOE), Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under Award DE-FG02-17ER16362. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. Additional computer resources were provided by the Minnesota Supercomputing Institute at the University of Minnesota.*

U.S. DEPARTMENT OF ENERGY | Office of Science