Predicting hydrogen storage in nanoporous materials using meta-learning

Yangzesheng Sun
University of Minnesota
sun00032@umn.edu

Robert F. DeJaco
University of Minnesota
dejac001@umn.edu

J. Ilja Siepmann
University of Minnesota
siepmann@umn.edu

Abstract

Discovering the optimal materials and conditions for hydrogen storage is crucial to the development of fuel cell vehicles. In this work, we present a method towards this goal by combining meta-learning and high-throughput molecular simulations. Our meta-learning model efficiently utilizes the datasets generated by high-throughput simulations and applies the physical domain knowledge of hydrogen storage. Comparing to training many small models each on a material, meta-learning showed more accurate prediction, higher data efficiency and improved generalization capability. Our meta-learning approach not only accelerates the workflow of computational materials discovery, but also introduces a practical application of meta-learning and few-shot learning in the physical sciences.

1 Introduction

Hydrogen fuel cell vehicles combine the advantages of both traditional engines and electric motors. [1–3] In fuel cell vehicles currently on the market, hydrogen is stored as a compressed gas at 700 times the atmospheric pressure. [4] Such high pressures require a strong fuel tank and pose safety threats after a severe traffic accident. [5, 6] Recently, there has been growing interests in storing hydrogen using nanoporous materials. By using a nanoporous material to adsorb hydrogen in the fuel tank, a storage capacity comparable to commercial compressed hydrogen tanks can be achieved at much lower pressures. [7] This allows for more flexibility in the shape of the fuel tank in a vehicle and improves its safety. Typically, the amount of molecules adsorbed in a nanoporous material is a function of the thermodynamic condition, i.e., temperature and pressure. Extensive research has been done in the physical science fields to model the adsorption as closed-form functions. [8]

One can use regression models to predict the hydrogen adsorption as a function of pressure and temperature, but finding the optimal materials to store hydrogen requires traversing enormous amounts of possible material structures. Although carrying out millions of experiments is evidently unfeasible under current technology, high-throughput molecular simulations can be utilized to obtain hydrogen storage data for large numbers of nanoporous materials. [9, 10] For each material, usually a dataset of hydrogen adsorption with varying temperatures and pressures is collected, so the high-throughput simulation results in hundreds or thousands of small datasets.

A common practice in materials discovery is to apply a simple model on each of the "base" datasets and run those models separately to make predictions, namely solving many small regression tasks independently. However, such a workflow suffers from several shortcomings. A simple model with few parameters imposes a strong inductive bias (assumption) which limits its capacity, and a complex model has greater risk in overfitting. Because each model is independently trained, similarity or other relationship among materials cannot be utilized.

Therefore, we leveraged meta-learning combined with high-throughput simulation to address the aforementioned problems. By virtue of being trained on the meta-dataset of many base tasks, a

Second Workshop on Machine Learning and the Physical Sciences (NeurIPS 2019), Vancouver, Canada.

meta-learning model can efficiently adapt to new tasks especially with limited data. [11–15] Here, we describe a meta-learning method to predict the hydrogen adsorption as a function of temperature and pressure given a few example data points for any nanoporous material. Comparing with training many small models, the advantages of our meta-learning approach include (1) lower prediction errors due to larger model capacity, (2) lower sensitivity to lack of data which enables few-shot learning, and (3) improved generalization in extrapolation capability.

2 Related Work

Machine learning has been used to aid the computational discovery of nanoporous materials. [16, 17] The majority of existing research focused on predicting the property at the same thermodynamic state for all materials, which requires training a different model when the condition of interest is changed. A recent work [18] trained a neural network operating on both thermodynamic states and the material structure to predict hydrogen storage, while we argue that its performance is intrinsically sensitive to the quality of material structure representation in its input.

Our meta-learning model was constructed as an encoder-decoder network. A common meta-learning approach is to generate an efficient initialization for adaptation or transfer leaning, [13] and model-based meta-learning models have been developed on various architectures. [19, 14] Encoder-decoder networks for meta-learning were also reported, [20–22] while they mainly used artificial datasets for regression problems.

3 Methods

Dataset generation In this work, we performed high-throughput Monte Carlo simulations [23] for 211 all-silica zeolites (a subset of nanoporous materials with the chemical formula SiO_2 [24]) which are of interest in hydrogen storage. The hydrogen adsorption in each zeolite material was simulated at 8 temperatures ranging from 77.0 K to 275.9 K and 8 pressures ranging from 0.10 MPa to 40.34 MPa. A base task for the meta-learning method consisted of 32 randomly selected state points, and 16 tasks were sampled from all data available in each zeolite. This produced an augmented meta-dataset with 3376 base tasks, where 160 zeolite materials were randomly selected whose tasks constitute the meta-training set.

Neural network for adsorption modeling First, we justify the usage of neural networks (NNs) in predicting adsorption of a material as the traditional physical science wisdom can coincide with the basic element of an NN. Equation 1 shows an adsorption model commonly used in materials science and chemical engineering derived from physics domain knowledge. [8]

$$y = f(p, T; A, B, n) = \frac{A\mathrm{e}^{\frac{B}{T}} p^n}{1 + A\mathrm{e}^{\frac{B}{T}} p^n}$$

where y is the fraction the material occupied ($y = 0$ means the material is empty, $y = 1$ means full), p is the pressure, T is the temperature, and A, B, n are parameters. Rearranging the equation obtains

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

which is exactly the sigmoid function with $\mathbf{x} = (\ln p, 1/T)$ as the feature vector. It is straightforward to extend the logistic regression into a neural network with sigmoid activation, and in this case the physical interpretation is that the material has multiple types of adsorption sites to occupy.

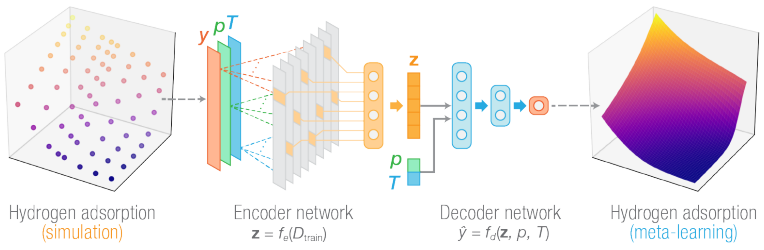

Figure 1: The meta-learning network architecture.

Meta-learning In a meta-learning algorithm, the error of a meta-learner over a distribution of datasets $p(D)$ is minimized. [13] Here, a dataset $D = \{\mathbf{x}, y\}$ contains the features $\mathbf{x} = (\ln p, 1/T)$ and the hydrogen adsorption y. The meta-learning model has the entire training set comprising multiple variable-output pairs in its input, given as

$$y = f(D_{\text{train}}, \mathbf{x})$$

Such a formulation is different from common multi-task learning models whose outputs for many tasks are predicted simultaneously $[\{y_i, \cdots, y_n\} = f(\mathbf{x})]$. For those models, more output units need to be added to predict a new material. In contrast, predicting a new material in meta-learning only requires a different D_{train} sample. The encoder-decoder structure is chosen as the meta-learning model because it is able to generate latent representations of each dataset (Figure 1). In the physical sciences, it is especially common to use a small but physically reasonable model because its parameters contain physical information about the system of interest. Regarding those parameters as lower-dimensional representation of the system, it echos the encoder-decoder structure of our meta-learning model as the model also gives a representation of a base dataset. The encoder $f_e(\cdot)$ consists of a point-wise feedforward network followed by max-pooling to make the latent representation invariant to the permutation of samples in the base dataset. [25] The decoder $f_d(\cdot)$ is a feedforward network predicting the adsorption whose input includes the latent variables $\mathbf{z}$ for a specific task (material) and the thermodynamic state $\mathbf{x} = (\ln p, 1/T)$,

$$\mathbf{z} = f_e(D_{\text{train}}), \quad \hat{y} = f_d(\mathbf{z}, \mathbf{x})$$

Regularization of the neural network is implemented by penalizing the correlation among dimensions in the latent space. The task-specific loss is formulated in a semi-supervised manner as to both reconstruct the training examples and predict test examples,

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

where the base training set D_{train} is a subset of D and $\text{cov}[\cdot]$ denotes the covariance matrix. λ is the hyperparameter controlling the strength of regularization.

4 Results and Discussion

Table 1: Reconstruction accuracy of hydrogen storage data from the meta-learning network and individually trained models. $\text{MLP}(\cdot)$ denotes a multi-layer neural network with number of hidden units listed in parentheses.

Model	Mean MSE ($\times 10^{-2}$)	Minimum MSE ($\times 10^{-2}$)	Maximum MSE ($\times 10^{-2}$)
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	**0.014**	**0.004**	0.365

Our meta-learning model was compared with the tradition in materials science that one model is fit to each material of interest. Such "curve fitting" is essentially a reconstruction task since the objective is to obtain a closed-form representation containing the material's physical properties. In our meta-learning model, this means the input $D_{\text{train}} = D$ and the network works as an autoencoder. Table 1 shows the geometric average, minimum, and maximum mean square errors (MSE) on the reconstruction of hydrogen adsorption in all 211 materials. The meta-learning network outperformed the entirely physics-informed "logistic regression" and the small NN with 8 hidden units with regard to all metrics listed. It also did not overfit the meta-training set.

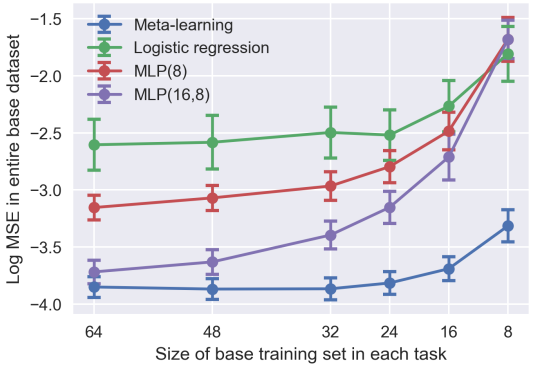

Figure 2: Few-shot performance of the meta-learning network and independently training multiple models. Error bars refer to the standard deviation of MSE on all 211 materials scaled by 1/3 for better visualization. The same subsampling of 8–48 thermodynamic states was used for all materials.

Training a relatively large NN on each material indeed result in similar or lower MSE than the meta-learning network, however such NNs are dramatically overparametrized on datasets with only 64 samples and it takes even more space to store the weights of all NNs than the original data.

Meta-learning is known to excel in few-shot learning problems where there is very limited data for a new task. [13] Few-shot learning could be particularly useful in physical sciences since conducting actual experiments can be very expensive. To investigate the few-shot learning capability of our meta-learning network, the size of D_{train} was decreased from 64 to 8 by random sampling subsets while the MSE was always evaluated on the entire dataset D. Figure 2 shows the MSE attained by each type of model when reducing the base-training set size in each material. Compared to training individual NNs, our meta-learning network is significantly less sensitive to the reduction of base-training set. On the contrary, training a large neural network on each material suffered from a dramatic increase in error with a decreasing training set.

Finally, we applied our meta-learning model to find the optimal condition of hydrogen storage for each material. We consider the fuel tank of a hydrogen-powered car to be filled up to a pressure p_1 and depleted to another pressure p_0 when it is empty. We used $p_0 = 0.271$ MPa and $p_1 = 3$ MPa based on reducing the infrastructure for fuel stations and the materials requirements for the tank. The temperature of the fuel tank is kept constant, so the "optimal condition" refers to the temperature at which the tank can provide maximum amount of hydrogen from filled to depleted, denoted as T^*. A unique T^* can be calculated for each material using the meta-learning network, however some of them may fall outside of the temperature range of our high-throughput simulation dataset, and this requires the network to extrapolate beyond the data distribution on which it was trained. For example, the zeolite material with the largest hydrogen storage capacity is predicted by meta-learning to have $T^* \approx 70$ K, while the lowest temperature in the high-throughput simulation is 77 K (Figure 3). Although the capacity monotonically decreases with temperature in the high-throughput simulation data, the network still predicts a maximum. This prediction is successfully validated by additional molecular simulations performed at temperatures below 77 K, as these simulation results are in good agreement with the prediction given by an ensemble of 8 independently trained meta-learning networks. Not surprisingly, training a neural network on data only for this zeolite led to very poor extrapolation. It highlights the benefit of our meta-learning approach that the neural network gives more accurate predictions by encompassing the information learned on multiple materials, and is especially valuable in the example here because molecular simulations become more difficult at lower temperatures.

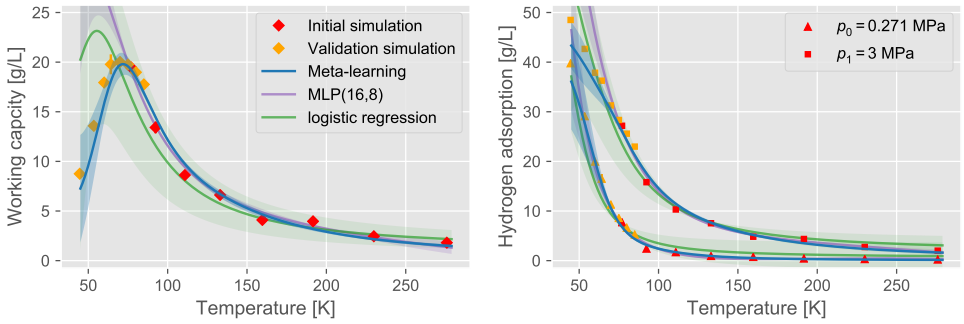

Figure 3: Extrapolation of the meta-learning network to predict the hydrogen storage in a high-capacity zeolite. The left column shows the storage capacity in the fuel tanks, and the right column shows the hydrogen adsorption at p_1 (upper branch, squares) and p_0 (lower branch, triangles). The baseline of training an ensemble of 8 $\mathrm{MLP}(16, 8)$ models or logistic regression (the domain knowledge model for adsorption) on the data for this zeolite is also shown.

5 Conclusion

We applied meta-learning to computational materials discovery by leveraging its benefits on robustness to limited data and improved generalization. While it could be very expensive to perform high-throughput simulations in the joint space of material structures and thermodynamic conditions,

meta-learning provides a route towards overcoming this difficulty by making the maximum use of available data. Apart from the decoder-encoder networks used in this work, other meta-learning methods can also be applied in conjunction with high-throughput simulations in the physical sciences, such as model-agnostic meta-learning [13] and memory augmented neural networks. [14] With regression in meta-learning usually benchmarked on artificially created data, data obtained from high-throughput simulations may serve as a novel real-world application of meta-learning methods.

References

[1] M. Z. Jacobson, W. G. Colella, and D. M. Golden. Cleaning the air and improving health with hydrogen fuel-cell vehicles. *Science*, 308:1901–1905, 2005.

[2] L. Schlapbach. Hydrogen-fuelled vehicles. *Nature*, 460:809–811, 2009.

[3] C.E. Thomas. Fuel cell and battery electric vehicles compared. *Int. J. Hydrogen Energy*, 34:6005–6020, 2009.

[4] A. Yamashita, M. Kondo, S. Goto, and N. Ogami. Development of high-pressure hydrogen storage system for the toyota "mirai". In *SAE 2015 World Congress & Exhibition*, 2015.

[5] D.K. Ross. Hydrogen storage: The major technological barrier to the development of hydrogen fuel cell cars. *Vacuum*, 80:1084–1089, 2006.

[6] M. Felderhoff, C. Weidenthaler, R. von Helmolt, and U. Eberle. Hydrogen storage: The remaining scientific and technological challenges. *Phys. Chem. Chem. Phys.*, 9:2643–2653, 2007.

[7] L. J. Murray, M. Dincă, and J. R. Long. Hydrogen storage in metal–organic frameworks. *Chem. Soc. Rev.*, 38:1294–1314, 2009.

[8] K.Y. Foo and B.H. Hameed. Insights into the modeling of adsorption isotherm systems. *Chem. Eng. J.*, 156:2–10, 2010.

[9] N. S. Bobbitt, J. Chen, and R. Q. Snurr. High-throughput screening of metal–organic frameworks for hydrogen storage at cryogenic temperature. *J. Phys. Chem. C*, 120:27328–27341, 2016.

[10] Alauddin Ahmed, Saona Seth, Justin Purewal, Antek G. Wong-Foy, Mike Veenstra, Adam J. Matzger, and Donald J. Siegel. Exceptional hydrogen storage achieved by screening nearly half a million metal–organic frameworks. *Nat. Commun.*, 10:1568, 2019.

[11] R. Vilalta and Y.f Drissi. A perspective view and survey of meta-learning. *Artificial Intelligence Review*, 18:77–95, 2002.

[12] M. Andrychowicz, M. Denil, S. Gómez, M. W. Hoffman, D. Pfau, T. Schaul, B. Shillingford, and N. de Freitas. Learning to learn by gradient descent by gradient descent. In *Advances in Neural Information Processing Systems 29*, pages 3981–3989, 2016.

[13] C. Finn, P. Abbeel, and S. Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pages 1126–1135, 2017.

[14] A. Santoro, S.y Bartunov, M. Botvinick, D. Wierstra, and T. Lillicrap. Meta-learning with memory-augmented neural networks. In *Proceedings of The 33rd International Conference on Machine Learning*, volume 48, pages 1842–1850, 2016.

[15] S. Ravi and H. Larochelle. Optimization as a model for few-shot learning. In *5th International Conference on Learning Representations*, Toulon, France, 2017.

[16] A. W. Thornton, C. M. Simon, J. Kim, O. Kwon, K. S. Deeg, K. Konstas, S. J. Pas, M. R. Hill, D. A. Winkler, M. Haranczyk, and B. Smit. Materials genome in action: Identifying the performance limits of physical hydrogen storage. *Chem. Mater.*, 29:2844–2854, 2017.

[17] B. J. Bucior, N. S. Bobbitt, T. Islamoglu, S. Goswami, A. Gopalan, T. Yildirim, O. K. Farha, N. Bagheri, and R. Q. Snurr. Energy-based descriptors to rapidly predict hydrogen storage in metal–organic frameworks. *Mol. Syst. Des. Eng.*, 4:162–174, 2019.

[18] G. Anderson, B. Schweitzer, R. Anderson, and D. A. Gómez-Gualdrón. Attainable volumetric targets for adsorption-based hydrogen storage in porous crystals: Molecular simulation and machine learning. *J. Phys. Chem. C*, 123:120–130, 2019.

[19] F. Sung, Y. Yang, L. Zhang, T. Xiang, P. H.S. Torr, and T. M. Hospedales. Learning to compare: Relation network for few-shot learning. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[20] T. Wu, J. Peurifoy, I. L. Chuang, and M. Tegmark. Meta-learning autoencoders for few-shot prediction, 2018. arXiv:1807.09912. arXiv.org e-Print archive. https://arxiv.org/abs/1807.09912.

[21] Andrei A. Rusu, Dushyant Rao, Jakub Sygnowski, Oriol Vinyals, Razvan Pascanu, Simon Osindero, and Raia Hadsell. Meta-learning with latent embedding optimization, 2018. arXiv:1807.05960. arXiv.org e-Print archive. https://arxiv.org/abs/1807.05960.

[22] Marta Garnelo, Jonathan Schwarz, Dan Rosenbaum, Fabio Viola, Danilo J. Rezende, S. M. Ali Eslami, and Yee Whye Teh. Neural processes, 2018. arXiv:1807.01622. arXiv.org e-Print archive. https://arxiv.org/abs/1807.01622.

[23] P. Bai, M. Y. Jeon, L. Ren, C. Knight, M. Deem, M. W.and Tsapatsis, and J. I. Siepmann. Discovery of optimal zeolites for challenging separations and chemical transformations using predictive materials modeling. *Nat. Commun.*, 6:5912, 2015.

[24] Ch. Baerlocher and L.B. McCusker. Database of zeolite structures. http://www.iza-structure.org/databases/, (accessed Apr 25, 2019).

[25] C. R. Qi, H. Su, K. Mo, and L. J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.