Supporting Information
of
A Transfer Learning Study of Gas Adsorption in Metal-Organic Frameworks

Ruimin Ma¹, Yamil J. Colón^{2*}, Tengfei Luo^{1,2*}

¹Department of Aerospace and Mechanical Engineering, ²Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, Indiana 46556, United States

* Corresponding emails: yclon@nd.edu, tluo@nd.edu

Figure S1. Transfer learning from H₂ adsorption at 100 bar and 243 K to H₂ adsorption at 100bar and 77 K: R^2 for (a) direct learning (R^2_{DL}), (b) transfer learning (R^2_{TL}), and (c) improvement from transfer learning ($R^2_{TL} - R^2_{DL}$) of all 1,000 target cases. Data in panels (a-c) are sorted in by $R^2_{TL} - R^2_{DL}$ in the descending order, there are 736 cases that $R^2_{TL} - R^2_{DL} > 0$. (d) R^2_{TL} plotted against R^2_{DL} with their respective distributions also shown. (e) $R^2_{TL} - R^2_{DL}$ as a function of R^2_{DL} , where the inset shows R^2_{TL} vs. R^2_{DL} for the $R^2_{TL} - R^2_{DL} < 0$ cases.

Figure S2. Transfer learning from H₂ adsorption at 100 bar and 243 K to H₂ adsorption at 100 bar and 200 K: R^2 for (a) direct learning (R^2_{DL}), (b) transfer learning (R^2_{TL}), and (c) improvement from transfer learning ($R^2_{TL} - R^2_{DL}$) of all 1,000 target cases. Data in panels (a-c) are sorted in by $R^2_{TL} - R^2_{DL}$ in the descending order, there are 956 cases that $R^2_{TL} - R^2_{DL} > 0$. (d) R^2_{TL} plotted against R^2_{DL} with their respective distributions also shown. (e) $R^2_{TL} - R^2_{DL}$ as a function of R^2_{DL} , where the inset shows R^2_{TL} vs. R^2_{DL} for the $R^2_{TL} - R^2_{DL} < 0$ cases.

Figure S3. Transfer learning from H_2 adsorption at 100 bar and 243 K to CH_4 adsorption at 65 bar and 298 K:

R^2 for (a) direct learning (R^2_{DL}), (b) transfer learning (R^2_{TL}), and (c) improvement from transfer learning ($R^2_{TL} - R^2_{DL}$) of all 1,000 target cases. Data in panels (a-c) are sorted in by $R^2_{TL} - R^2_{DL}$ in the descending order, there are 735 cases that $R^2_{TL} - R^2_{DL} > 0$. (d) R^2_{TL} plotted against R^2_{DL} with their respective distributions also shown. (e) $R^2_{TL} - R^2_{DL}$ as a function of R^2_{DL} , where the inset shows R^2_{TL} vs. R^2_{DL} for the $R^2_{TL} - R^2_{DL} < 0$ cases.

Figure S4. Transfer learning from H_2 adsorption at 100 bar and 243 K to Xe/Kr mixture adsorption at 1 bar and 298 K: R^2 for (a) direct learning (R^2_{DL}), (b) transfer learning (R^2_{TL}), and (c) improvement from transfer learning ($R^2_{TL} - R^2_{DL}$) of all 1,000 target cases. Data in panels (a-c) are sorted in by $R^2_{TL} - R^2_{DL}$ in the descending order, there are 449 cases that $R^2_{TL} - R^2_{DL} > 0$. (d) R^2_{TL} plotted against R^2_{DL} . (e) $R^2_{TL} - R^2_{DL}$ as a function of R^2_{DL} .

Figure S5. Pair plot between descriptors (left) and the Pearson correlation coefficient map of descriptors (right).

Figure S6. Transfer learning from H₂ adsorption at 100 bar and 243 K to H₂ adsorption at 100bar and 77 K:

MSE for (a) direct learning (MSE_{DL}), (b) transfer learning (MSE_{TL}), and (c) improvement from transfer learning ($MSE_{DL} - MSE_{TL}$) of all 1,000 target cases. Data in panels (a-c) are sorted in by $MSE_{DL} - MSE_{TL}$ in the descending order, there are 908 cases that $MSE_{DL} - MSE_{TL} > 0$.

Table S1. Quantitative relationships between descriptors and labels using random forest, and measurement of corresponding descriptor importance scores. The most dominant descriptor for predicting each label are marked in bold face.

Case	R_{test}^2	Descriptor Importance Score				
		Void fraction	Vol S. A.	Grav. S. A.	Pore limiting diameter	Largest cavity diameter
H ₂ @100 bar/243K	0.998	0.952	0.023	0.022	0.002	0.001
H ₂ @100 bar/130K	0.997	0.926	0.017	0.054	0.002	0.001
CH ₄ @100 bar/298K	0.994	0.776	0.019	0.199	0.004	0.002
Xe/Kr@5bar/298K	0.620	0.257	0.182	0.202	0.202	0.157