On the importance of catalyst-adsorbate 3D interactions for relaxed energy predictions

Alvaro Carbonero, Alexandre Duval, Victor Schmidt, Santiago Miret, Alex Hernandez-Garcia, Yoshua Bengio, David Rolnick

On the importance of catalyst-adsorbate 3D interactions for relaxed energy predictions

Alvaro Carbonero
Mila
alvaro.carbonero@mila.quebec

Alexandre Duval
Mila, CentraleSupélec

Victor Schmidt
Mila, Université de Montréal

Santiago Miret
Intel labs

Alex Hernandez-Garcia
Mila, Université de Montréal

Yoshua Bengio
Mila, Université de Montréal

David Rolnick
Mila, McGill University

Abstract

The use of machine learning for material property prediction and discovery has traditionally centered on graph neural networks that incorporate the geometric configuration of all atoms. However, in practice not all this information may be readily available, e.g. when evaluating the potentially unknown binding of adsorbates to catalyst. In this paper, we investigate whether it is possible to predict a system's relaxed energy in the OC20 dataset while ignoring the relative position of the adsorbate with respect to the electro-catalyst. We consider SchNet, DimeNet++ and FAENet as base architectures and measure the impact of four modifications on model performance: removing edges in the input graph, pooling independent representations, not sharing the backbone weights and using an attention mechanism to propagate non-geometric relative information. We find that while removing binding site information impairs accuracy as expected, modified models are able to predict relaxed energies with remarkably decent MAE. Our work suggests future research directions in accelerated materials discovery where information on reactant configurations can be reduced or altogether omitted.

1 Introduction

Materials discovery is a strong driver of innovation, unlocking new materials with tailored properties that benefit various domains including energy efficiency, transportation systems and electronics [Butler et al., 2018]. Yet, it faces significant hurdles. Characterizing new materials is computationally expensive [Oganov et al., 2019], even when replacing lab experiments by quantum mechanical simulations like the Density Functional Theory (DFT) [Kohn et al., 1996]. Besides, exploration of materials is hindered by the vastness of the search space, which encompasses myriad compositions, atomic arrangements and properties [Pyzer-Knapp et al., 2022].

To overcome these challenges, researchers have turned to Machine Learning (ML) for accelerated materials discovery [Zhang et al., 2023, Miret et al., 2022, Lee et al., 2023, Song et al., 2023] for two primary reasons: First, ML holds the power to quickly model materials' properties (i.e., to evaluate candidates) using Geometric Graph Neural Networks (GNNs). Second, generative ML can automatically propose new and consistent material candidates.

In recent years, ML has emerged as a crucial tool for the discovery of electro-catalysts, which play a key role in promoting renewable energy processes and sustainable chemical production, including the

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

production of ammonia for fertilizers and hydrogen [Zitnick et al., 2020]. The Open Catalyst Project (OCP) [Chanussot et al., 2021] has significantly contributed to this field by releasing an extensive dataset of 1,281,040 DFT relaxations of catalyst-adsorbate pairs, selected from a pool of meanigful candidates[1]. This dataset was specifically designed to train ML models to predict the relaxed energy of 3D adsorbate-catalyst (adslab) systems, a critical property that influences a catalyst's activity and selectivity [Tran and Ulissi, 2018], or its effectiveness for a specific chemical reaction. [2]. OCP has facilitated significant advancements in catalysis discovery, with ML models increasingly bridging the performance gap with DFT simulations, while offering a speed advantage of several orders of magnitude [Musaelian et al., 2022, Gasteiger et al., 2022, Passaro and Zitnick, 2023].

While research has primarily focused on predicting material properties [Wieder et al., 2020], it's equally crucial to efficiently explore the vast search space of potential catalyst candidates [Wei et al., 2019, Zhang et al., 2022]. This involves generating consistent candidates automatically, a challenge due to the intricate process of creating adslab samples. This process, further detailed in Appendix A, includes positioning an adsorbate molecule (e.g., H2O) with a catalyst in 3D space, meaning cutting a surface through the catalyst bulk, selecting the adsorbate's spatial orientation, and sampling a plausible binding site [Chanussot et al., 2021]. These steps, which are time-consuming and challenging to model, determine the input configuration of the adslab sample and significantly influence the relaxed energy prediction [Lan et al., 2022]. An additional challenge is that this process and its relevance to the actual in-lab efficiency of a real material is not yet fully understood by chemists [Deshpande et al., 2020].

In light of these challenges, this paper explores the possibility of predicting the relaxed energy of an adslab without co-locating the adsorbate and the catalyst in the same 3D space. This direction of study is valuable for several reasons: (1) to better understand the role of adsorbate-catalyst geometry in determining the relaxed energy; (2) to reduce reliance on the exact input configuration, which often correlates with local energy minima, (3) to avoid the complexity and computational cost of determining a good input configuration (e.g., finding a binding site and adsorbate orientation). Our primary objective is quantify the loss in accuracy that occurs when the geometric relationship between inputs is unavailable.

To achieve this, we propose four modifications of existing GNN architectures, collectively referred to as *Disconnected GNNs*. All four modifications enable the base architecture to make relaxed energy predictions without explicitly modeling geometric interactions between the adsorbate and the catalyst. We then evaluate the trade-off of omitting these interactions through experiments on the OC20 dataset and suggest future directions to overcome the limits associated with the input adslab configuration.

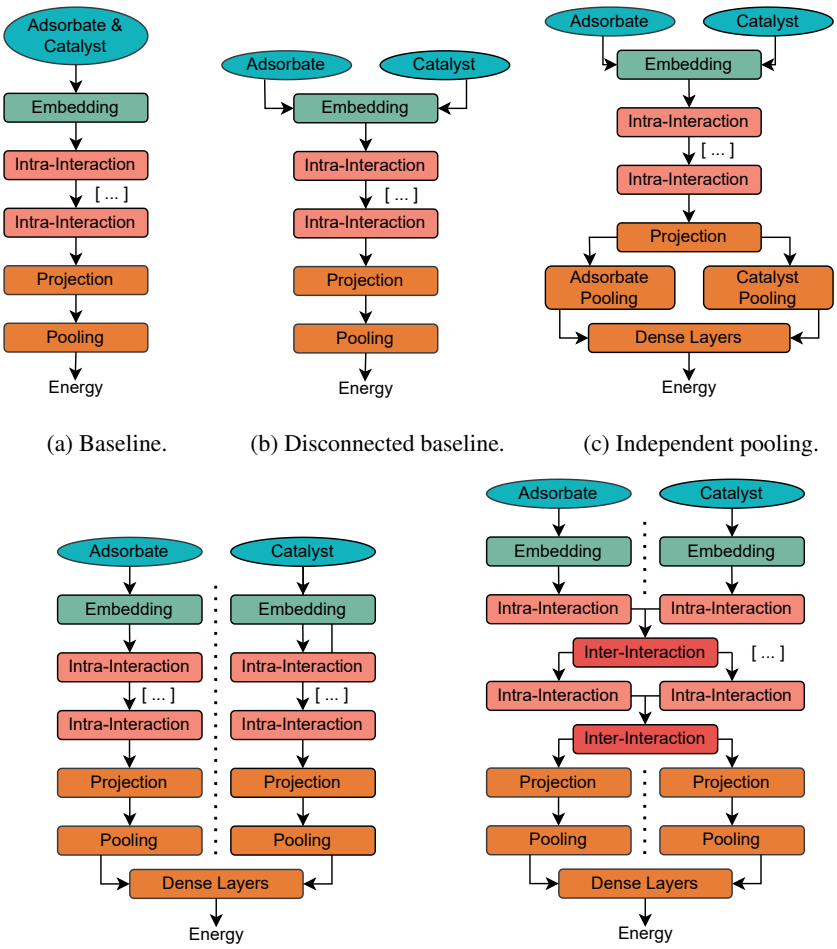
2 Methods

Our methods test the assumption that meaningful results can be obtained in predicting the relaxed energy even when the geometric interactions between the adsorbate and the catalyst are unknown. Specifically, we propose four *Disconnected GNN* models. These models can leverage as their backbone any underlying GNN model that predict properties of 3D atomic systems using the general pipeline detailed under "Refresher" below, where we make targeted changes to the pipeline. In this paper, we investigate the effect of the proposed Disconnected GNN architecture on three backbone models: SchNet [Schütt et al., 2017], DimeNet++ [Klicpera et al., 2020] and FAENet [Duval et al., 2023].

Refresher. Recall that GNNs applied to 3D atomic systems mostly use the following pipeline. (1) **Graph creation**: construct a graph representation of the input point cloud systems (e.g. adslab), represented with atomic numbers and 3D atom positions. (2) **Embedding**: derive node/edge embeddings based on atomic numbers and geometric information (e.g. relative atom positions). (3) **Interaction blocks**: apply a fixed number of message passing layers [Gilmer et al., 2017] to update the node and/or the edge embeddings, using geometric information and preserving data symmetries. (4) **Output block**: project final atom representations into scalar values (e.g. atom-properties). For graph-level prediction, a global pooling (e.g. sum) is performed to aggregate atom predictions.

[1]The adsorbate refers to the molecule involved in the electrochemical reaction that is accelerated through the introduction of a catalyst material, represented by a semi-infinite periodic substructure.

[2]In other words, it quantifies the extent to which the catalyst reduces the energy required for a chemical reaction to occur.



(a) Baseline. (b) Disconnected baseline. (c) Independent pooling.

(d) Independent Backbones. (e) Attention.

Figure 1: Architectures are presented in order of increasing complexity. 1a illustrates the standard property prediction pipeline for Geometric GNNs, with no modifications. In 1b, edges between parts are removed. In 1c, the pooling step is performed separately on atom representations for the adsorbate and the catalyst. Resulting embeddings are concatenated and passed through an MLP. In 1d, each part has its own model with distinct weights. In 1e, an interaction block with artificial edges and weights allows node embeddings between both parts to communicate.

2.1 Disconnected baseline

The baseline Disconnected GNN model does not create any edges between the catalyst and the adsorbate in the graph creation step. By removing all such edges, no relative geometric information will pass from the adsorbate to the catalyst when modelling the system, since GNNs only propagate information through graph edges. Since this modification only relates to the input data representation, this change can be applied to any backbone GNN architecture, like FAENet, SchNet or DimeNet++ for instance. It is illustrated in Figure 1b.

Note that we still use the original OC20 dataset where the adsorbate/catalyst atom positions have been determined after selecting a binding site and an adsorbate orientation. But since GNNs use atom relative positions and not absolute atom positions to preserve translation equivariance, disconnecting the adsorbate-catalyst graph is enough to discard totally the input configuration. Processing such a disconnected graph is equivalent to considering both components independently. Associated results thus inform about the importance of the relative adslab geometric information to each model.

2.2 Independent pooling

This method builds upon the disconnected baseline defined above, also removing all edges between the adsorbate and the catalyst. As illustrated in Figure 1c, there are three additional changes. First, the projection block now outputs a hidden representation $\mathbf{h} \in \mathbb{R}^{\lfloor H/2 \rfloor}$ for each node instead of a scalar quantity, where H is the dimension of the embedding. Second, the resulting node representations are pooled separately for each component, leading to distinct hidden representations for the adsorbate and the catalyst. Lastly, we concatenate these two representations and pass them to an additional dense layer to compute the energy of the system. This final MLP layer gives more expressive power to the model than the disconnected baseline because we explicitly model non-geometric interactions between the adsorbate and the catalyst.

2.3 Independent backbones

This approach, illustrated in Figure 1d, builds upon independent pooling. The main difference is that catalysts and adsorbates are assigned distinct GNN models (hence "independent") instead of sharing backbone weights. The motivation is that the adsorbate and the catalyst have very different sizes and roles. Hence, it might be beneficial to have independent GNNs, where each one is trains for its particular input. Similarly to independent pooling, both models are modified to produce a hidden representation of their respective parts. They are then concatenated and passed through an MLP.

2.4 Attention

This model (see Figure 1e) builds upon the independent model by allowing the nodes of both graphs to communicate during the interaction layers.

We modify the graph creation process to produce a heterogeneous graph where we add the following weighted edges to previous models' edges. For every catalyst node i and adsorbate node j, we create the new edges (i, j) and (j, i) with weights z and $-z$ respectively, where z is the z-axis coordinate of x_i. Besides, in the embedding step, we pass the edge weights through an RBF layer and a linear layer to increase their dimension. Notice that this model can run without locating the adsorbate and catalyst in the same 3D plane since no information regarding the adsorbate's location (e.g. distance, relative position) is given. The motivation behind these weighted edges is to make the model aware of the closeness of a node to the catalyst's surface, while respecting the restrictions of the setup. As noted in [Duval et al., 2022], fixed atoms not near the surface of the catalyst (tag 0 nodes) can be removed without hurting efficiency, indicating that closeness to surface plays a factor.

We now describe how this graph is used. Let F denote a "normal" intra-interaction layer, let h^l_{ads} and h^l_{cat} denote the node embeddings of the adsorbate and the catalyst at interaction layer l, and let e_{ads} and e_{cat} stand for the edges of the adsorbate and catalyst, respectively. To produce h^{l+1}_{ads} and h^{l+1}_{cat}, we first calculate $h'_{\text{ads}} = F(h^l_{\text{ads}}, e_{\text{ads}})$ and $h'_{\text{cat}} = F(h_{\text{cat}}, e_{\text{cat}})$. Subsequently, we calculate $h' = \text{GATConv}(\text{Concat}(h'_{\text{ads}}, h'_{\text{cat}}), e')$ where e' is the set of new weighted edges and GATConv is a convolution layer as described in [Veličković et al., 2017]. Finally,

$$h^{l+1}_{\text{ads}}, h^{l+1}_{\text{cat}} = \text{Norm}(h'_{\text{ads}} + h^l_{\text{ads}}), \text{Norm}(h'_{\text{cat}} + h^l_{\text{cat}}).$$

3 Evaluation

Dataset: We use the OC20 IS2RE [Zitnick et al., 2020] dataset, which involves the direct prediction of the relaxed adsorption energy from the initial atomic structure, i.e. a graph regression task requiring E(3)-invariance. OC20 comes with $450K$ training samples and a predefined train/val/test split. We evaluate models on the four distinct splits of the validation set ($\sim 25K$ samples each): In Domain (ID), Out of Domain Adsorbates (OOD-ads), Out of Domain catalysts (OOD-cat), and Out of Domain Adsorbates and catalysts (OOD-both).

Metrics: We measure accuracy on each validation split through the energy MAE. Running time is measured with the throughput at inference time, i.e. the average number of samples per second that a model can process in its forward pass, on similar GPU types.

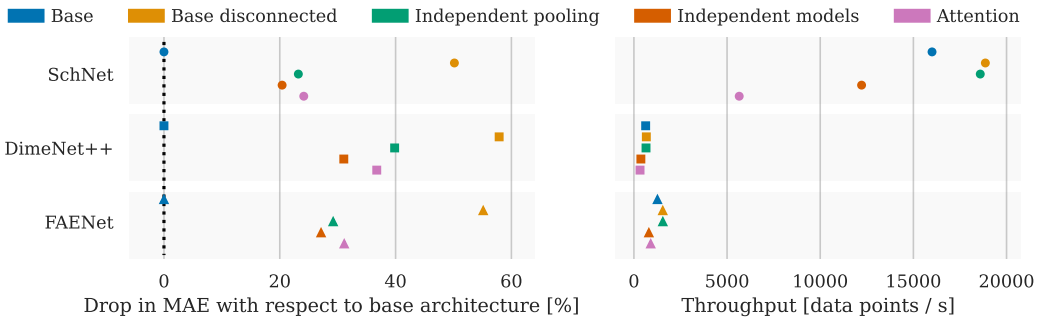


Figure 2: Comparison of MAE performance drops across our four proposed modifications to three base architectures (left), and their associated impact on model throughput (right). We can see the the disconnected baseline consistently achieves the best performance. In terms of throughput, it is on par with other modifications for DimeNet++ and FAENet, while it is slower than all other modifications for SchNet (except for the Attention modification).

Baselines: We compare base results of the original models (considering 3D adsorbate-catalyst interactions) with the four disconnected architectures detailed in Section 2. Note that we apply PhAST components [Duval et al., 2022] on all models to improve performance and inference time.

Results. From Figure 2, we see that the independent model architecture performs best among all proposed disconnected GNNs. However, it presents a performance drop in MAE of 20%, 31%, and 28% with respect to baseline architectures, for Schnet, DimeNet++ and FAENet respectively. A complete numerical table of results is reported in Table 1.

Performance is significantly worse for disconnected GNNs than for GNNs with access to the adslab configuration; however, **these results are promising in several ways**.

Firstly, a drop in performance was expected since we neglect important information by omitting to modelise the adsorbate-catalayst 3D interaction. The important take-away is that our disconnected GNN manages to produce decent energy predictions while being agnostic to the adslab's input configuration. Future work will probably manage to reduce further this performance gap. Note that we discuss some axes of improvements in Appendix C.

Secondly, the OC20 dataset often contains multiple input configuration samples for the same catalyst-adsorbate pair, each with a different relaxed energy target. This cannot be captured by our disconnected GNN, leading to multiple targets for the same sample, which can impair training dynamics. Furthermore, when evaluating on a dataset that does not present this feature, e.g. the 10k split of OC20 IS2RE, disconnected GNNs match (or outperform) baseline models, which is very promising. We present and briefly discuss these results in Appendix D.

The *OC20-Dense* dataset [Lan et al., 2022] was recently released, containing multiple input configurations (e.g. binding site, adsorbate orientation) for each adsorbate-catalyst pair along with the associated relaxed energy. Building upon this dataset, we will be able to train disconnected GNN models to predict the minimum relaxed energy of each adslab over all possible input configurations. The obvious benefit of such approach is that it marginalises over adslab binding sites and adsorbate orientations, avoiding the need to explore all configurations to actually find the relaxed energy global minima of this adsorbate-catalyst pair, which is expected to be the main reaction driver in real life experiments [Lan et al., 2022].

4 Conclusion

In this paper, we have explored predicting an adsorbate-catalyst system's relaxed energy without co-locating them in 3D space. We proposed four disconnected models built on top of existing Geometric GNNs architectures, and showed through experiments that this task was possible with a manageable performance loss. Looking forward, we see the potential of using such disconnected models to find the global energy minima of adslab systems while circumventing the need to consider

all possible input adslab configurations. This sounds particularly promising for generative methods and should contribute to the acceleration of catalysis discovery.

References

Neil W Ashcroft and N David Mermin. *Solid state physics*. Cengage Learning, 2022.

Jacob R Boes, Osman Mamun, Kirsten Winther, and Thomas Bligaard. Graph theory approach to high-throughput surface adsorption structure generation. *The Journal of Physical Chemistry A*, 123(11):2281–2285, 2019.

Keith T Butler, Daniel W Davies, Hugh Cartwright, Olexandr Isayev, and Aron Walsh. Machine learning for molecular and materials science. *Nature*, 559(7715):547–555, 2018.

Lowik Chanussot, Abhishek Das, Siddharth Goyal, Thibaut Lavril, Muhammed Shuaibi, Morgane Riviere, Kevin Tran, Javier Heras-Domingo, Caleb Ho, Weihua Hu, et al. Open catalyst 2020 (oc20) dataset and community challenges. *ACS Catalysis*, 11(10):6059–6072, 2021.

Siddharth Deshpande, Tristan Maxson, and Jeffrey Greeley. Graph theory approach to determine configurations of multidentate and high coverage adsorbates for heterogeneous catalysis. *npj Computational Materials*, 6(1):79, 2020.

Alexandre Duval, Victor Schmidt, Santiago Miret, Yoshua Bengio, Alex Hernández-García, and David Rolnick. Phast: Physics-aware, scalable, and task-specific gnns for accelerated catalyst design. *arXiv preprint arXiv:2211.12020*, 2022.

Alexandre Duval, V. Schmidt, A. Garcia, Santiago Miret, Fragkiskos D. Malliaros, Y. Bengio, and D. Rolnick. Faenet: Frame averaging equivariant gnn for materials modeling. *International Conference on Machine Learning*, 2023. doi: 10.48550/arXiv.2305.05577.

Johannes Gasteiger, Muhammed Shuaibi, Anuroop Sriram, Stephan Günnemann, Zachary Ulissi, C Lawrence Zitnick, and Abhishek Das. Gemnet-oc: developing graph neural networks for large and diverse molecular simulation datasets. *arXiv preprint arXiv:2204.02782*, 2022.

J. Gilmer, S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. *International Conference On Machine Learning*, 2017.

Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, et al. The materials project: A materials genome approach to accelerating materials innovation, apl mater. 2013.

Johannes Klicpera, Shankari Giri, Johannes T Margraf, and Stephan Günnemann. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. *Preprint arXiv:2011.14115*, 2020.

Walter Kohn, Axel D Becke, and Robert G Parr. Density functional theory of electronic structure. *The Journal of Physical Chemistry*, 100(31):12974–12980, 1996.

Georg Kresse and Jürgen Hafner. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Physical Review B*, 49(20):14251, 1994.

Janice Lan, Aini Palizhati, Muhammed Shuaibi, Brandon M. Wood, Brook Wander, Abhishek Das, Matt Uyttendaele, C. Lawrence Zitnick, and Zachary W. Ulissi. Adsorbml: Accelerating adsorption energy calculations with machine learning. *arXiv preprint arXiv: Arxiv-2211.16486*, 2022.

Kin Long Kelvin Lee, Carmelo Gonzales, Marcel Nassar, Matthew Spellings, Mikhail Galkin, and Santiago Miret. Matsciml: A broad, multi-task benchmark for solid-state materials modeling, 2023.

Santiago Miret, Kin Long Kelvin Lee, Carmelo Gonzales, Marcel Nassar, and Matthew Spellings. The open MatSci ML toolkit: A flexible framework for machine learning in materials science. *Preprint arXiv:2210.17484*, 2022.

Albert Musaelian, Simon Batzner, Anders Johansson, Lixin Sun, Cameron J Owen, Mordechai Kornbluth, and Boris Kozinsky. Learning local equivariant representations for large-scale atomistic dynamics. *Preprint arXiv:2204.05249*, 2022.

Artem R Oganov, Chris J Pickard, Qiang Zhu, and Richard J Needs. Structure prediction drives materials discovery. *Nature Reviews Materials*, 4(5):331–348, 2019.

Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L Chevrier, Kristin A Persson, and Gerbrand Ceder. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013.

Saro Passaro and C Lawrence Zitnick. Reducing so (3) convolutions to so (2) for efficient equivariant gnns. *arXiv preprint arXiv:2302.03655*, 2023.

Edward O Pyzer-Knapp, Jed W Pitera, Peter WJ Staar, Seiji Takeda, Teodoro Laino, Daniel P Sanders, James Sexton, John R Smith, and Alessandro Curioni. Accelerating materials discovery using artificial intelligence, high performance computing and robotics. *npj Computational Materials*, 8 (1):84, 2022.

Kristof Schütt, Pieter-Jan Kindermans, Huziel Enoc Sauceda Felix, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in neural information processing systems*, 30, 2017.

Yu Song, Santiago Miret, and Bang Liu. MatSci-NLP: Evaluating scientific language models on materials science language tasks using text-to-schema modeling. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3621–3639, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.acl-long.201. URL `https://aclanthology.org/2023.acl-long.201`.

Kevin Tran and Zachary W Ulissi. Active learning across intermetallics to guide discovery of electrocatalysts for co2 reduction and h2 evolution. *Nature Catalysis*, 1(9):696–703, 2018.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

Jing Wei, Xuan Chu, Xiang-Yu Sun, Kun Xu, Hui-Xiong Deng, Jigen Chen, Zhongming Wei, and Ming Lei. Machine learning in materials science. *InfoMat*, 1(3):338–358, 2019.

Oliver Wieder, Stefan Kohlbacher, Mélaine Kuenemann, Arthur Garon, Pierre Ducrot, Thomas Seidel, and Thierry Langer. A compact review of molecular property prediction with graph neural networks. *Drug Discovery Today: Technologies*, 37:1–12, 2020.

Mengchun Zhang, Maryam Qamar, Taegoo Kang, Yuna Jung, Chenshuang Zhang, Sung-Ho Bae, and Chaoning Zhang. A survey on graph diffusion models: Generative ai in science for molecule, protein and material. *arXiv preprint arXiv:2304.01565*, 2023.

Xu Zhang, Yun Tian, Letian Chen, Xu Hu, and Zhen Zhou. Machine learning: a new paradigm in computational electrocatalysis. *The Journal of Physical Chemistry Letters*, 13(34):7920–7930, 2022.

C Lawrence Zitnick, Lowik Chanussot, Abhishek Das, Siddharth Goyal, Javier Heras-Domingo, Caleb Ho, Weihua Hu, Thibaut Lavril, Aini Palizhati, Morgane Riviere, et al. An introduction to electrocatalyst design using machine learning for renewable energy storage. *Preprint arXiv:2010.09435*, 2020.

A Adslab generation

This section summarises the adslab creation process of the OC20 dataset. Refer to the original paper [Chanussot et al., 2021] for more details.

Adsorbate and catalyst surface selection. The first step is to select the adsorbate and the catalyst that will compose the adslab. For the adsorbate, it extremely simple: it is sampled randomly from a

Baseline / MAE	ID	OOD-ad	OOD-cat	OOD-both	Average	Throughput (samples/s)
SchNet	0.631	0.687	0.625	0.626	**0.642**	16001 ± 2264
Base disconnected-SchNet	0.940	1.037	0.932	0.947	0.964	$\mathbf{18863 \pm 3034}$
Independent pooling-SchNet	0.766	0.847	0.762	0.789	0.791	$\underline{18590 \pm 2230}$
Independent models-SchNet	0.761	0.805	0.764	0.764	$\underline{0.773}$	12220 ± 1593
Attention-SchNet	0.779	0.852	0.762	0.797	0.797	5652 ± 876
DimeNet++	0.577	0.693	0.568	0.621	**0.615**	631 ± 125
Base disconnected-DimeNet++	0.922	1.059	0.914	0.989	0.971	$\mathbf{667 \pm 143}$
Independent pooling-DimeNet++	0.781	0.0.992	0.767	0.901	0.860	$\underline{650 \pm 139}$
Independent models-DimeNet++	0.767	0.884	0.754	0.818	$\underline{0.806}$	374 ± 68
Attention-DimeNet++	0.777	0.952	0.763	0.870	0.841	329 ± 57
FAENet	0.554	0.623	0.546	0.578	**0.575**	1263 ± 766
Base disconnected-FAENet	0.884	0.937	0.878	0.870	0.892	1549 ± 1027
Independent pooling-FAENet	0.698	0.813	0.697	0.766	0.743	$\underline{1553 \pm 1049}$
Independent models-FAENet	0.690	0.782	0.689	0.760	$\underline{0.731}$	$\mathbf{803 \pm 327}$
Attention-FAENet	0.732	0.814	0.715	0.756	0.754	902 ± 565

Table 1: MAE and inference time for various GNNs and their disconnected counterpart on *OC20* IS2RE, averaged over 3 runs. *Average* MAE is computed over all validation splits. Best results are shown in bold, second best underlined. Overall, disconnected models show a lower accuracy than baselines but they are able to predict relaxed energy relatively well despite omitting geometric interaction between parts in the adslab.

set of 82 molecules that are chosen for their utility to renewable energy applications. For the catalyst surface, the process is divided in three stages. First, we choose the number of distinct chemical elements composing it. It can be a unary material (5% chance), a binary material (65%) or a ternary material (30%). These elements are chosen from a set of 55 elements comprising reactive nonmetals, alkali metals, transition metals, etc. Next, a stable bulk material is randomly selected from the 11K samples of the Materials Project [Jain et al., 2013] with the number of elements chosen in the first step. Lastly, all symmetrically distinct surfaces[3] from the material with Miller indices [Ashcroft and Mermin, 2022] less than or equal to 2 are enumerated, including possibilities for different absolute positions of surface plane. From this list of surfaces, one is randomly selected.

Input adslab configuration. The objective of the second step is to place the adsorbate and the catalyst surface in the same 3D plane. Using pymatgen's Voronoi tesselation algorithm [Ong et al., 2013] on surface atoms and adsorbate possible binding sites, Catkit [Boes et al., 2019] enumerates a list of symmetrically distinct adsorption sites along with suggested per-site adsorbate orientations. From this list, an adsorption configuration is randomly selected, yielding the initial adslab structure. This initial structure is then relaxed using Density Functional Theory simulations, performed using the Vienna Ab Initio simulation Package (VASP) [Kresse and Hafner, 1994].

B Full results table

C Axes of improvement

There are two main ways in which Disconnected GNN models can be further improved. The first one is to improve upon dense layers as a way to combine the final hidden representations of the adsorbate and the catalyst. The second one is to find a way to effectively allow the nodes of the adsorbate and catalyst to communicate, and to do this while ommitting the relative position and orientation between the adsorbate and the catalyst. This is attempted in the attention models through the inter-interaction layers. Given that they do not consistently outperform the independent backbone models, it is likely that the way in which attention models use GAT convolution layers hinders the model's prediction. We still hope that this can be fixed since the attention models can outperform baseline models when trained on the 10k dataset, as we described in Appendix D.

[3] a surface is a cut through the bulk operated using Miller indices

Baseline 10k/ MAE	ID	OOD-ad	OOD-cat	OOD-both	Average	Throughput (samples/s)
SchNet	1.151	1.245	1.104	1.213	1.179	18602 ± 4373
Base disconnected-SchNet	1.246	1.236	1.214	1.144	1.210	25487 ± 3693
Independent pooling-SchNet	0.962	1.038	0.935	0.961	<u>0.974</u>	24832 ± 3041
Independent models-SchNet	0.976	0.985	0.946	0.935	**0.961**	<u>16501 ± 3283</u>
Attention-SchNet	0.979	1.072	0.961	0.968	0.995	$\mathbf{7952 \pm 1261}$
DimeNet++	0.851	0.969	0.802	0.875	**0.874**	623 ± 123
Base disconnected-DimeNet++	1.044	1.078	1.026	0.971	1.023	649 ± 140
Independent pooling-DimeNet++	0.908	1.022	0.888	0.928	0.937	752 ± 114
Independent models-DimeNet++	0.889	0.971	0.876	0.883	<u>0.905</u>	<u>366 ± 71</u>
Attention-DimeNet++	0.889	1.017	0.865	0.915	0.921	$\mathbf{330 \pm 60}$
FAENet	1.003	1.017	0.992	1.004	1.004	<u>1417 ± 792</u>
Base disconnected-FAENet	1.104	1.144	1.118	1.061	1.107	1492 ± 1085
Independent pooling-FAENet	1.022	1.009	0.999	0.938	0.992	1626 ± 1123
Independent models-FAENet	0.891	1.020	0.871	1.040	<u>0.955</u>	1089 ± 737
Attention-FAENet	0.894	0.924	0.863	0.872	**0.888**	$\mathbf{916 \pm 567}$

Table 2: MAE and inference time for various GNNs and their disconnected counterpart on *OC20* 10k IS2RE, averaged over 3 runs. *Average* MAE is computed over all validation splits. Disconnected models all show lower accuracy. In every model, every subsurface catalyst atom (tag 0 atoms) is removed as suggested in [Duval et al., 2022] to improve inference time.

D Results in the 10k dataset

Only 89% of the OC20 IS2RE dataset consists of adsorbate-catalyst systems which are unique up to the combination of adsorbate id, bulk id, and bulk cell. That is, to Disconnected GNNs, 11% of the dataset consists of inputs with multiple targets due to distinct binding sites, which impairs training dynamics. When compared to the OC20 10k IS2RE dataset, every single adsorbate-catalyst system has a unique combination of adsorbate id, bulk id, and bulk cell. Thus, evaluating disconnected GNNs on the 10k dataset is a useful way to assess their performance on a suitable data tailored to them. This is corroborated by the results in Table 2.

From Table 2, we see that independent models perform on par or better (for FAENet) than baseline models when trained on the 10k dataset. In particular, we see a performance improvement of 18%, -4%, and 5% for SchNet, DimeNet++, and FAENet respectively. Different than in the all dataset, Attention-FAENet outperforms Independent models-FAENet by 7.5%. These results are a sign that Disconnected GNNs can perform well when tested against the data for which they were designed for. The OC20-Dense dataset [Lan et al., 2022] is thus an exciting development as it will allow for the creation of the datasets that these models need.