ADSORPTION-SITE-GUIDED MESSAGE PASSING FOR CATALYTIC SYSTEM ENERGY PREDICTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Graph neural networks (GNNs) have been widely used for energy prediction in catalytic systems. However, existing methods primarily focus on modeling geometric relationships among atoms, while adsorption-site information has not been sufficiently incorporated into message passing. To address this limitation, we propose an adsorption-site-guided message-passing framework that encodes the distances between atoms and adsorption-site using radial basis functions (RBFs) and integrates the resulting site features into node or directed-edge representations. A gating mechanism then adaptively modulates the contribution of site information based on the current structural representations, enabling effective integration of adsorption-site information with the backbone representations. The proposed framework is integrated into SchNet, DimeNet++, and PaiNN and evaluated on the Initial Structure to Relaxed Energy (IS2RE) task of the Open Catalyst 2020 (OC20) dataset. Experiments across three data regimes—10k, 100k, and the full training set—show that the proposed framework improves upon the corresponding baselines across most models and evaluation metrics, with more pronounced gains in the low-data regime. Ablation studies on SchNet further demonstrate that adsorption-site information and the gating mechanism play complementary roles, and their combination yields consistent performance improvements.These results demonstrate that the proposed framework can effectively exploit structural information associated with adsorption-site and is broadly applicable across different graph neural network architectures.

1 INTRODUCTION

Accurate and efficient prediction of surface adsorption properties is important for catalyst design and high-throughput screening. Density functional theory (DFT) can accurately compute the energies and electronic-structure properties of catalytic systems, but its high computational cost limits large-scale screening of candidate materials and adsorption configurations(Vogt & Weckhuysen, 2022; Calle-Vallejo et al., 2015a; Kohn & Sham, 1965; Hohenberg & Kohn, 1964; Lan et al., 2023; Jung et al., 2023). Machine learning provides a computationally efficient approach by learning structure–energy relationships from atomic configurations, enabling rapid energy prediction for catalytic systems(Pablo-García et al., 2023; Pederson et al., 2022).

With the emergence of large-scale datasets for catalysis such as OC20(Chanussot et al., 2021), GNNs have become increasingly important for energy prediction in catalytic systems. Representative models such as SchNet(Schütt et al., 2017), DimeNet++(Gasteiger et al., 2020a;b), and PaiNN(Schütt et al., 2021) learn structural representations from interatomic distances, angles, and other 3D geometric information, achieving strong predictive performance. However, existing methods primarily focus on modeling general interatomic geometric relationships(Mou et al., 2023; Li et al., 2023), while adsorption-site information remains underutilized. In catalytic systems, adsorption behavior is strongly influenced by local interactions among the adsorbate, the adsorption-site, and neighboring surface atoms(Calle-Vallejo et al., 2015a;b). Previous work has explored local-environment pooling to aggregate node representations in the local neighborhood of the adsorbate, thereby reducing the influence of irrelevant global structural information(Li et al., 2023). However, such approaches mainly exploit local information at the graph-level readout stage, without incorporating the spatial relationships between atoms or edges and adsorption-site across multiple message-passing layers.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

This raises an important question: how can adsorption-site information be incorporated into message passing while adaptively modulating its contribution based on the current structural representations?

To address this limitation, we propose an adsorption-site-guided message-passing framework for the Initial Structure to Relaxed Energy (IS2RE) task on the Open Catalyst 2020 (OC20). The framework encodes the spatial relationships between atoms or edges and adsorption-site and employs a gating mechanism to adaptively modulate the contribution of site information, enabling adsorption-site information to be incorporated into feature updates across multiple message-passing layers. We instantiate the proposed framework on three representative geometric GNN backbones, SchNet, DimeNet++, and PaiNN, to evaluate its applicability across different architectures. Experimental results show that the proposed framework generally improves performance across different training-set sizes and backbone models, with more pronounced gains in the low-data regime.

The main contributions of this work are as follows:

1. We develop an adsorption-site-guided message-passing framework that incorporates local information associated with adsorption-site into multiple message-passing layers and adaptively modulates its contribution through a gating mechanism.
2. We adapt the proposed framework to three representative geometric GNN backbones, SchNet, DimeNet++ and PaiNN; consistent performance gains across architectures and training-set sizes demonstrate its broad applicability.
3. Ablation studies reveal a complementary effect between adsorption-site information and the gating mechanism, with their joint use yielding more consistent performance gains, particularly in the low-data regime.

2 Related Work

2.1 Graph Neural Networks for Catalytic Energy Prediction

The OC20 dataset(Chanussot et al., 2021) and the Open Catalyst Challenge(Das et al., 2021) provide unified benchmarks for large-scale training and evaluation in catalysis, facilitating the development of GNNs for materials modeling. SchNet(Schütt et al., 2017) primarily models atomic structures through interatomic distances, DimeNet++(Gasteiger et al., 2020a;b) incorporates directional and angular information, and PaiNN(Schütt et al., 2021) extends message passing to rotationally equivariant representations, enabling more effective modeling of directional information in 3D structures. Other representative methods, including ForceNet(Hu et al., 2021), SpinConv(Shuaibi et al., 2021), GemNet(Gasteiger et al., 2021), GemNet-OC(Gasteiger et al., 2022), 3D-Graphormer(Shi et al., 2022), SCN(Zitnick et al., 2022; Passaro & Zitnick, 2023), FAENet(Duval et al., 2023), PhAST(Duval et al., 2024), and Equiformer(Liao & Smidt, 2023; Liao et al., 2024), have further advanced 3D atomic modeling from different perspectives, including force prediction, equivariant representations, directional modeling, attention mechanisms, and computational efficiency.

Beyond architectural improvements, recent studies have increasingly explored models tailored to catalytic systems and large-scale pretraining. GAME-Net(Pablo-García et al., 2023) and AdsorbML(Lan et al., 2023) target adsorption energy prediction and adsorption configuration screening, respectively, while JMP(Shoghi et al., 2024), OMat24(Barros-Luque et al., 2026), eSEN(Fu et al., 2025), and UMA(Wood et al., 2025) further explore pretraining across diverse chemical systems and universal atomistic models. Overall, existing methods have demonstrated the effectiveness of GNN-based approaches for energy prediction in catalytic systems. However, their message-passing mechanisms still primarily model general 3D geometric relationships among atoms. As a result, adsorption-site information is captured only implicitly through the global structural representation, while the spatial relationships between atoms or edges and adsorption-site have not yet been fully incorporated into message passing.

2.2 Adsorption-Site Priors and Local Modeling

Prior studies have shown that adsorption behavior is closely associated with the adsorbate, the adsorption-site and its surrounding atomic environment(Calle-Vallejo et al., 2015a;b). Li et al. proposed a local-environment pooling method that aggregates atomic representations in the neighbor-

hood of the adsorbate to reduce the influence of irrelevant global structural information(Li et al., 2023). Chen et al. proposed AdsMT, which uses positional encoding to represent the relative height of surface atoms and employs cross-attention to capture relationships between the adsorbate and surface atoms(Chen et al., 2025). These studies suggest that local structural and spatial information associated with adsorption can complement global structural representations for energy prediction.

However, existing methods mainly exploit such information through local readout, positional encoding, or attention-based modeling, while the spatial relationships between atoms or edges and adsorption-site have not yet been directly incorporated into message passing. Although gating and attention mechanisms can adaptively control information propagation and aggregation in graphs(Li et al., 2015; Veličković et al., 2018), in existing applications to catalytic systems, their weights are typically learned from hidden representations, with limited integration of task-specific priors such as adsorption-site information(Li et al., 2023; Chen et al., 2025). Therefore, a key question addressed in this work is how to directly incorporate the spatial relationships between atoms or edges and adsorption-site into message passing while adaptively modulating the contribution of adsorption-site information.

3 Adsorption-Site-Guided Message Passing Framework

3.1 Problem Definition

We consider the IS2RE task for catalytic systems, which aims to predict the relaxed energy directly from the initial atomic structure without performing structural relaxation.

Specifically, an initial catalytic structure containing n atoms is represented as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{A}, \boldsymbol{z}, \boldsymbol{R})$, where $\mathcal{V} = \{\boldsymbol{v}_i\}_{i=1}^{n}$ denotes the node set, and each node corresponds to an atom. $\boldsymbol{z} = (z_1, \cdots, z_n)$ contains the atomic numbers of all atoms, where z_i denotes the atomic number of atom i and identifies its chemical element. $\boldsymbol{R} = (\boldsymbol{r}_1, \cdots, \boldsymbol{r}_n)^{\mathsf{T}} \in \mathbb{R}^{n\times 3}$ denotes the atomic coordinates, where $\boldsymbol{r}_i \in \mathbb{R}^3$ is the position of atom i. $\mathcal{A} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the edge set, where two atoms are connected if their interatomic distance is smaller than a predefined cutoff. The geometric GNNs considered in this work initialize atomic representations from atomic numbers and construct geometric features from atomic coordinates. Message passing then aggregates information from local atomic neighborhoods to learn representations for energy prediction.

In the standard IS2RE formulation, a model learns a mapping from the initial atomic structure to the energy of the corresponding relaxed structure, which can be written as $\hat{E} = F_\theta(\mathcal{G})$, where $F_\theta(\cdot)$ denotes an energy prediction model parameterized by θ, and $\hat{E}$ denotes the predicted energy of the relaxed structure. Building on this formulation, we introduce the adsorption site as an additional local structural prior. Specifically, $\boldsymbol{r}_{\mathrm{ads}} \in \mathbb{R}^3$ represents the Cartesian coordinate of the adsorption site in the initial adsorption configuration. In this work, $\boldsymbol{r}_{\mathrm{ads}}$ is directly obtained from the adsorption-site metadata provided by the OC20 dataset, which stores the adsorption configurations before DFT relaxation.The prediction problem considered in this work is formulated as $\hat{E} = F_\theta(\mathcal{G}, \boldsymbol{r}_{\mathrm{ads}})$. Compared with the standard IS2RE formulation, the prediction target remains unchanged, while the adsorption-site coordinate is introduced as an additional local structural prior to facilitate modeling of adsorption-related local environments.

3.2 Motivation

To illustrate the importance of modeling the local atomic environment around adsorption sites for the IS2RE task, we consider NO_2 adsorption on the VSSe surface as a representative example (Peng et al., 2025) and analyze the charge density difference and atomic force distributions, as shown in Figure 1. The adsorption induced charge redistribution is primarily localized near the interface between the NO_2 molecule and the VSSe surface. This localized redistribution is evident in Figure 1(b, e), particularly within the red dashed regions. Meanwhile, the atomic forces obtained from DFT single-point calculations exhibit pronounced spatial heterogeneity. In Figure 1(c, f), the blue arrows indicate the atomic forces, with their lengths and directions representing the force magnitudes and directions, respectively. Larger force magnitudes are concentrated around the adsorbate and the adsorption interface within the red dashed regions, while atoms farther from the adsorption region experience relatively smaller forces.

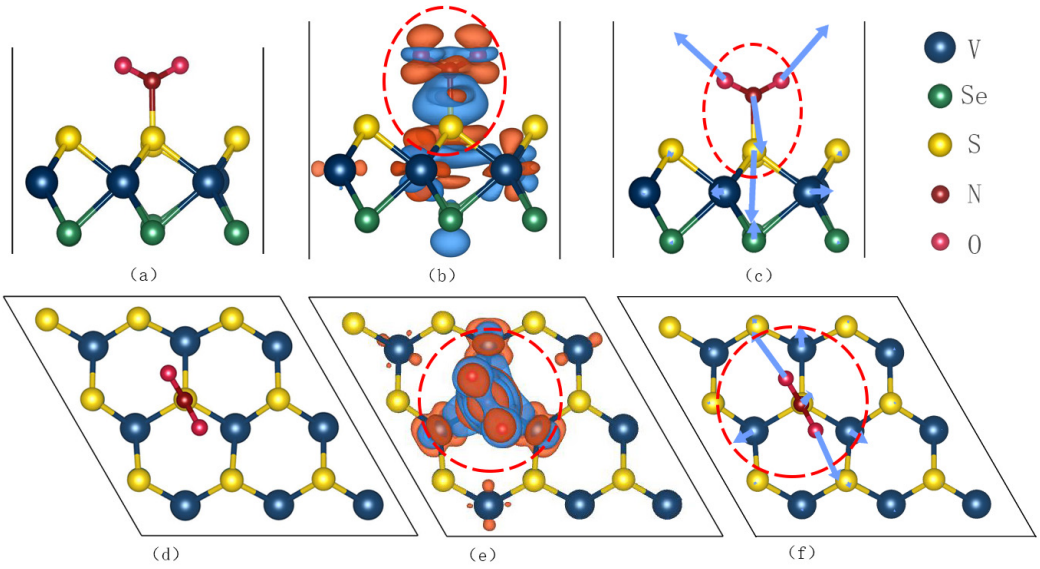

Figure 1: Physical motivation for adsorption-site-guided modeling. (a, d) Side and top views of NO_2 adsorption on the VSSe surface, respectively; (b, e) charge density difference distributions, with red and blue isosurfaces indicating electron accumulation and depletion, respectively; (c, f) atomic force vectors obtained from DFT single-point calculations.

These observations suggest that, in this example, charge redistribution and force responses are spatially localized around the adsorption region, indicating that the adsorption site and its surrounding environment contain task-relevant local structural information for energy prediction. Motivated by this spatial locality, we treat the adsorption site as a local structural prior for GNN-based energy prediction. Incorporating this prior into the backbone representations allows the model to explicitly leverage adsorption-related local structural information, reducing its reliance on inferring such task-relevant patterns solely from the training data and the original geometric representations.

3.3 Framework Overview

Motivated by the above observations, we propose an adsorption-site-guided message-passing framework that can be integrated into different geometric GNNs, as illustrated in Figure 2. The framework treats the adsorption-site as a local structural prior and incorporates the spatial relationships between atoms or edges and the adsorption-site into the feature updates of the backbone model. The framework follows a unified pipeline of relative positional encoding, site information fusion, and adaptive gating, and is adapted according to the feature representations and message-passing mechanisms of different backbone models.

To formalize this process, we denote $\boldsymbol{X}^{(l)}$ the intermediate feature representation at the l-th message-passing layer by $\boldsymbol{X}^{(l)}$, with $\boldsymbol{x}_i^{(l)}$ representing the feature associated with atom i. We first construct a relative positional encoding based on the spatial relationship between atoms or edges and the adsorption-site:

$$\boldsymbol{p}_{i,\mathrm{ads}} = \mathcal{E}_{\mathrm{pos}}\left(\mathcal{R}\left(\boldsymbol{r}_i, \boldsymbol{r}_{\mathrm{ads}}\right)\right), \tag{1}$$

where $\mathcal{R}(\cdot)$ represents the relative spatial relationship between them, $\mathcal{E}_{\mathrm{pos}}(\cdot)$ denotes the positional encoding function. This operation maps the spatial relationship associated with the adsorption-site to a positional representation $\boldsymbol{p}_{i,\mathrm{ads}}$, allowing it to continuously contribute to subsequent feature updates. For different backbone models, the specific form of $\mathcal{R}(\cdot)$ can be adapted to node, edge, or other geometry-based representations.

Next, the relative positional encoding $\boldsymbol{p}_{i,\mathrm{ads}}$ is fused with the current intermediate features $\boldsymbol{X}^{(l)}$:

$$\boldsymbol{f}_i^{(l)} = \mathcal{F}_{\mathrm{pos}}\left(\boldsymbol{x}_i^{(l)}, \boldsymbol{p}_{i,\mathrm{ads}}\right), \tag{2}$$

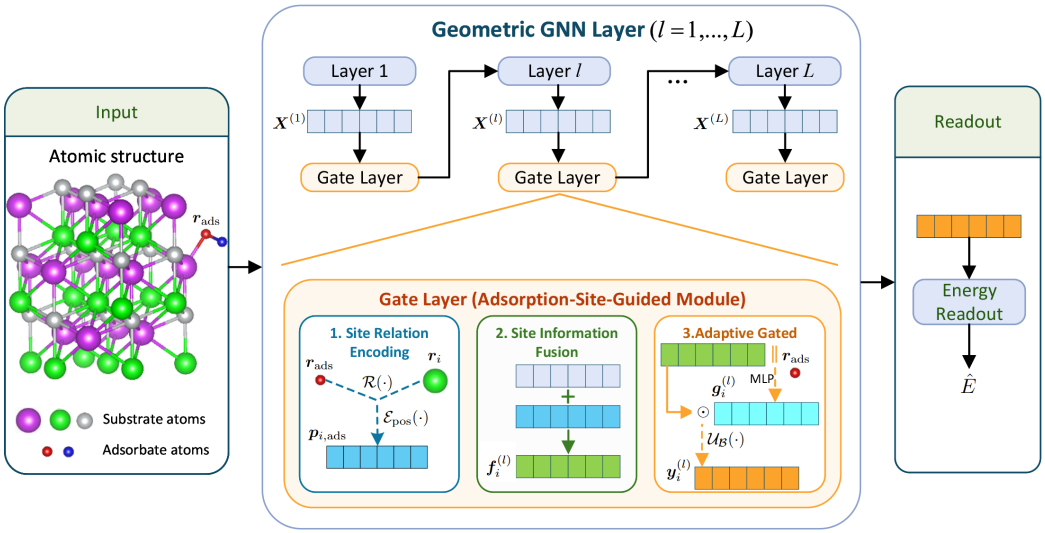

Figure 2: Overall architecture of the adsorption-site-guided message-passing framework.

where $\mathcal{F}_{\text{pos}}(\cdot)$ denotes the site information fusion operation, $\boldsymbol{f}_i^{(l)}$ denotes the feature representation after incorporating adsorption-site information. This fusion enables joint modeling of the backbone geometric representations and local structural information associated with the adsorption-site.

$$\boldsymbol{g}_i^{(l)} = \sigma\Big(\text{MLP}\left(\boldsymbol{f}_i^{(l)} \parallel \boldsymbol{r}_{\text{ads}}\right)\Big), \tag{3}$$

where $\parallel$ denotes feature concatenation, $\text{MLP}(\cdot)$ denotes a multilayer perceptron, and $\sigma(\cdot)$ is the sigmoid function. $\boldsymbol{g}_i^{(l)}$ denotes the channel-wise gating weights generated jointly from the site-enhanced representation $\boldsymbol{f}_i^{(l)}$ and the adsorption-site information. The sigmoid function constrains each gating weight to (0,1), allowing the model to adaptively modulate different feature channels according to the adsorption-site information. The generated gating weights are then applied channel-wise to the site-enhanced feature representation through element-wise multiplication, followed by a backbone-specific update:

$$\boldsymbol{y}_i^{(l)} = \mathcal{U}_{\mathcal{B}}\left(\boldsymbol{f}_i^{(l)}, \boldsymbol{g}_i^{(l)} \odot \boldsymbol{f}_i^{(l)}\right), \tag{4}$$

where $\odot$ denotes element-wise multiplication, $\mathcal{U}_{\mathcal{B}}(\cdot)$ denotes the feature update function for backbone model $\mathcal{B}$, and $\boldsymbol{y}_i^{(l)}$ denotes the updated feature representation. Depending on the backbone architecture, the gated feature is either used directly to update the representation or incorporated through a residual connection. The resulting representation is passed to subsequent message-passing layers, allowing the adsorption-site prior to contribute to feature updates across multiple layers.

3.4 Architecture-Specific Adaptation

Within the unified framework, adsorption-site information is integrated according to the feature representations and message-passing mechanisms of different backbone architectures. We instantiate the framework on SchNet, DimeNet++, and PaiNN, and use SchNet as a representative backbone to illustrate the adaptation in detail.

SchNet primarily uses scalar node features as its intermediate representations. Accordingly, the adsorption-site-guided module is used to update the node features produced by the interaction blocks, as illustrated in Figure 3. We first compute the relative distance between each atom and the adsorption-site:

$$d_{i,\text{ads}} = \|\boldsymbol{r}_i - \boldsymbol{r}_{\text{ads}}\|_2 . \tag{5}$$

Next, $d_{i,\text{ads}}$ is expanded using RBFs and then passed through a linear layer followed by a SiLU activation, producing a site representation with the same dimensionality as the node hidden features:

$$\boldsymbol{p}_{i,\text{ads}} = \text{SiLU}\left(\boldsymbol{W}_p \ \text{RBF}(d_{i,\text{ads}}) + \boldsymbol{b}_p\right). \tag{6}$$

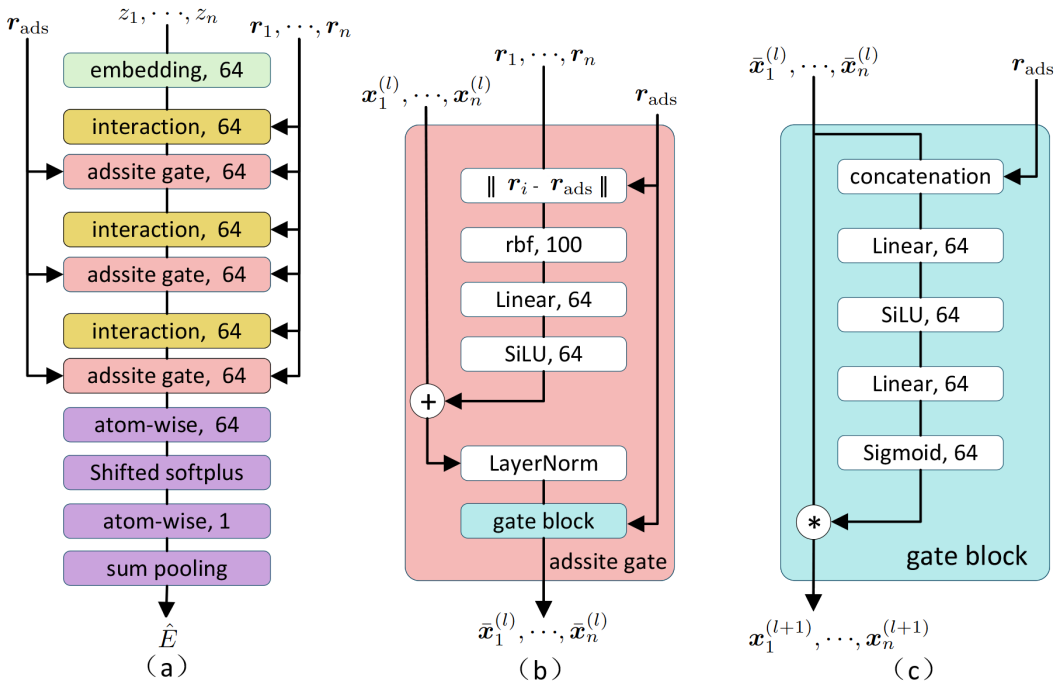

Figure 3: Implementation of the adsorption-site-guided message passing framework in SchNet. (a) Overall integration pipeline in SchNet; (b) adsorption-site relative positional encoding and site information fusion module(SchNet-site); (c) adaptive gating and fusion module(SchNet-gate).

The site representation $\boldsymbol{p}_{i,\text{ads}}$ is then fused with the node features $\boldsymbol{x}_i^{(l)}$ through a residual connection, followed by layer normalization to obtain the site-enhanced representation:

$$\bar{\boldsymbol{x}}_i^{(l)} = \text{LayerNorm}\left(\boldsymbol{x}_i^{(l)} + \boldsymbol{p}_{i,\text{ads}}\right). \tag{7}$$

The fused node representation is then passed to the gating module, where channel-wise gating weights are generated from the current node representation and the site information. The gating weights are directly applied to the fused representation through element-wise multiplication, and the resulting gated representation is passed to the subsequent interaction blocks for further message passing. This allows adsorption-site information to contribute to node feature updates across multiple layers.

The adaptations to DimeNet++ and PaiNN follow the same principles of relative positional encoding, site information fusion, and adaptive gating, with implementation details adjusted to their respective architectures. In DimeNet++, adsorption-site information is incorporated into directed-edge representations, whereas in PaiNN, it is fused with the RBF representations of edge distances before subsequent message passing. Detailed formulations and implementation details are provided in Appendix A.

4 Experiment and Results

4.1 Experimental Setup

4.1.1 Dataset, Data Splits and Evaluation Metrics

We conduct experiments on the OC20 IS2RE task. The models take atomic structures and their corresponding adsorption-site coordinates as input. Following the predefined OC20 data splits, we train the models separately on the 10k, 100k, and all training sets to evaluate their performance under different training data scales. Regardless of the training set size, all models are evaluated on the same four validation subsets: `val_id`, `val_ood_ads`, `val_ood_cat`, and `val_ood_both`.

Here, `val_id` corresponds to the in-domain setting, while the other three subsets evaluate out-of-distribution generalization to unseen adsorbates, unseen catalysts, and both unseen adsorbates and catalysts, respectively.

We evaluate model performance using the official OC20 IS2RE metrics, Mean Absolute Error (MAE) and Energy within Threshold (EwT), and additionally report Mean Squared Error (MSE) as a complementary error metric. Lower MAE and MSE indicate smaller prediction errors, whereas higher EwT indicates better performance. EwT is defined as the fraction of samples whose absolute energy prediction error is within 0.02 eV.

4.1.2 Training Settings and Baselines

For the PaiNN experiments at the 10k and 100k training set sizes, both the baseline PaiNN and its adsorption-site-guided variant are retrained using the same backbone configuration and training protocol. For the remaining baselines, we load publicly available pretrained weights(Chanussot et al., 2021) and re-evaluate the corresponding models on the same validation subsets; the resulting MAE on `val_id` matches the published OC20 results. Complete training configurations and hyperparameters are provided in the YAML configuration files released with the source code.

4.2 Performance and Ablation Analysis

4.2.1 Overall Performance Comparison

To evaluate the effectiveness of the adsorption-site-guided message passing framework under different backbone models and training set sizes, we apply it to SchNet, DimeNet++, and PaiNN and compare each variant with its corresponding baseline across the 10k, 100k, and All training set. Table 1 reports MAE, MSE, and EwT on `val_id` and Aggregated OOD, where Aggregated OOD comprises `val_ood_ads`, `val_ood_cat`, and `val_ood_both`. The complete results for the three individual OOD validation subsets are provided in Appendix B.

Table 1: Comparison of baseline and adsorption-site-guided models on OC20 IS2RE.

Train set	Model name	`val_id`			Aggregated OOD		
		MAE↓	MSE↓	EwT↑	MAE↓	MSE↓	EwT↑
10k	SchNet	1.0581	2.3178	0.0139	1.0529	2.1721	0.0141
	SchNet-Ours	**0.9712**	**1.8912**	**0.0171**	**0.9393**	**1.7836**	**0.0160**
	DimeNet++	0.8837	1.6569	0.0190	0.8927	1.6174	0.0175
	DimeNet++-Ours	**0.7836**	**1.3747**	**0.0245**	**0.7438**	**1.2091**	**0.0229**
	PaiNN	0.9457	1.7572	0.0157	0.9135	1.6399	0.0161
	PaiNN-Ours	**0.9063**	**1.6520**	**0.0170**	**0.8882**	**1.5642**	**0.0170**
100k	SchNet	0.7137	1.1694	0.0256	0.7400	1.2208	0.0224
	SchNet-Ours	**0.6752**	**1.0915**	**0.0292**	**0.6973**	**1.1484**	**0.0242**
	DimeNet++	0.6388	1.0107	0.0355	0.6925	1.1034	0.0275
	DimeNet++-Ours	**0.6215**	**0.9755**	**0.0390**	**0.6769**	**1.0320**	0.0270
	PaiNN	0.6736	1.0359	0.0238	0.6757	1.0441	0.0256
	PaiNN-Ours	**0.6535**	**1.0111**	**0.0309**	**0.6644**	**1.0085**	**0.0273**
all	SchNet	0.6458	0.9968	0.0295	0.6726	1.0459	0.0254
	SchNet-Ours	**0.5998**	**0.8904**	**0.0339**	**0.6484**	**0.9580**	**0.0269**
	DimeNet++	0.5639	0.8345	0.0423	0.6413	1.004	0.0314
	DimeNet++-Ours	**0.5575**	**0.8166**	0.0421	**0.6120**	**0.8935**	0.0314
	PaiNN	0.5729	0.8286	0.0366	0.6102	0.9055	0.0290
	PaiNN-Ours	**0.5547**	**0.7976**	**0.0394**	**0.6015**	**0.8681**	**0.0311**

Note. For Aggregated OOD, MAE, MSE, and EwT are computed over all samples pooled from the three OOD validation subsets. Bold values indicate improved performance over the corresponding baseline.

At the 10k training set sizes, all three adsorption-site-guided models outperform their corresponding baselines across all evaluation metrics, with particularly pronounced improvements for SchNet and DimeNet++. For example, SchNet reduces MAE and MSE on `val_id` by 8.2% and 18.4%, respectively. DimeNet++ reduces MAE and MSE on Aggregated OOD by 16.7% and 25.2%, respectively, while increasing EwT by 30.9%. PaiNN shows relatively smaller gains but still outperforms its baseline across all metrics. These results suggest that, in the low-data regime, local structural information associated with adsorption-site can complement the geometric representations of the backbone models.

At the 100k and All training set sizes, the performance gains generally become smaller, but the adsorption-site-guided variants of all three backbone models still achieve lower MAE and MSE than their corresponding baselines. At the All training set size, SchNet reduces MAE and MSE on `val_id` by 7.1% and 10.7%, respectively, while DimeNet++ reduces MAE and MSE on Aggregated OOD by 4.6% and 11.0%, respectively. PaiNN also achieves lower MAE and MSE and higher EwT on both `val_id` and Aggregated OOD. Overall, the framework delivers consistent performance gains across different backbone models and training-set sizes, with more pronounced improvements in the low-data regime.

4.2.2 Ablation Study

To assess the individual and combined contributions of the adsorption-site information module and the gating module, we conduct ablation experiments using SchNet as the backbone. The adsorption-site information module consists of two components: relative positional encoding and site information fusion. We compare four variants: the SchNet baseline, SchNet with only the adsorption-site information module (SchNet-site), SchNet with only the gating module (SchNet-gate), and the full model incorporating both modules (SchNet-site+gate). All variants use the same backbone configuration and training settings. The results are reported in Table 2, where the best-performing results are shown in bold and the second-best results are underlined.

Table 2: Ablation results for SchNet.

Train set	Site	Gate	`val_id`			Aggregated OOD		
			MAE↓	MSE↓	EwT↑	MAE↓	MSE↓	EwT↑
10k	–	–	1.0581	2.3178	0.0139	1.0529	2.1721	0.0141
	✓	–	1.0860	2.4505	<u>0.0152</u>	1.0400	2.3479	0.0147
	–	✓	<u>1.0041</u>	<u>2.0037</u>	0.0151	<u>0.9864</u>	<u>1.9444</u>	<u>0.0150</u>
	✓	✓	**0.9712**	**1.8912**	**0.0171**	**0.9393**	**1.7836**	**0.0160**
100k	–	–	<u>0.7137</u>	<u>1.1694</u>	0.0256	<u>0.7400</u>	<u>1.2208</u>	<u>0.0224</u>
	✓	–	0.7470	1.3121	<u>0.0263</u>	0.8442	1.4532	0.0201
	–	✓	0.7648	1.3142	0.0239	0.8526	1.5258	0.0187
	✓	✓	**0.6752**	**1.0915**	**0.0292**	**0.6973**	**1.1484**	**0.0242**
all	–	–	0.6458	0.9968	0.0295	0.6726	1.0459	0.0254
	✓	–	<u>0.6027</u>	<u>0.9048</u>	<u>0.0311</u>	<u>0.6559</u>	<u>0.9777</u>	<u>0.0266</u>
	–	✓	0.6400	1.0105	0.0303	0.7221	1.1599	0.0231
	✓	✓	**0.5998**	**0.8904**	**0.0339**	**0.6484**	**0.9580**	**0.0269**

Overall, SchNet-site+gate achieves the most consistent performance improvements across the 10k, 100k, and All training set sizes and obtains the best results across all metrics on both `val_id` and Aggregated OOD. In contrast, the two single-module variants show greater performance variation across training set sizes. SchNet-gate provides some improvements at the 10k training set size but shows degradation on several metrics at 100k and All. SchNet-site shows inconsistent performance gains at 10k and 100k, while achieving more pronounced improvements at the All training-set size. These results indicate that neither SchNet-site nor SchNet-gate alone provides consistent gains, whereas combining the adsorption-site information module with the gating module, as in SchNet-site+gate, yields more consistent performance improvements across training set sizes. This suggests that the two modules play complementary roles.

In addition, to further investigate how different backbone models respond internally to adsorption-site related information, Appendix C presents atom-level response visualizations of SchNet and DimeNet++ on representative adsorption structure.

5 DISCUSSION

The experimental results show that the adsorption-site-guided message-passing framework generally improves predictive performance across different backbone models and training-set sizes, with more pronounced gains in the low-data regime (10k). This suggests that, in the low-data regime, local structural information associated with adsorption-site can provide an additional structural prior, helping the backbone models capture local structural patterns that may be difficult to learn from limited samples. As the amount of training data increases, the backbone models can learn richer structural representations, which may explain why the relative gains from the adsorption-site prior tend to decrease.

Because backbone models differ in their feature representations and message-passing mechanisms, adsorption-site information is incorporated and propagated differently across architectures. This is consistent with the distinct atom-level response patterns observed for SchNet and DimeNet++ in Appendix C. The ablation results further show that neither SchNet-site nor SchNet-gate provides consistent performance gains, whereas SchNet-site+gate yields more consistent improvements across training set sizes, suggesting that the adsorption-site information module and the gating module play complementary roles. The current framework requires pre-specified adsorption-site coordinates, whose accuracy may affect the construction of adsorption-site related features. Future work could explore automatic adsorption-site identification and joint modeling to reduce reliance on pre-specified site information.

6 CONCLUSION

We propose an adsorption-site-guided message-passing framework that incorporates adsorption-site priors into the multi-layer message-passing process of GNNs by encoding local spatial relationships between atoms or edges and adsorption-site and adaptively modulating the contribution of site information through a gating mechanism. The framework is implemented and evaluated on SchNet, DimeNet++, and PaiNN, achieving overall performance improvements across different training set sizes on the OC20 IS2RE task. Ablation studies further demonstrate the complementary roles of adsorption-site information and the gating mechanism. Overall, the proposed framework provides an effective approach for enhancing geometric representation learning in catalytic systems by leveraging local structural priors associated with adsorption-site.

AI USE STATEMENT

In this work, we used generative AI tools to assist with language editing, improving the clarity and readability of the manuscript, refining the presentation and structure of technical descriptions, and preparing LaTeX expressions, mathematical notation, and table formatting. Generative AI tools were also used to assist with literature searching and identifying potentially relevant references. In some cases, they provided suggestions on how to present the methodology and experimental results more clearly.

Generative AI was not used to generate or modify the experimental datasets, conduct model training or evaluation, or produce the numerical results reported in this paper. The research problem, proposed framework, experimental design, implementation, experiments, analysis of quantitative results, and scientific conclusions were developed and verified by the authors. Mathematical expressions suggested or reformatted with the assistance of generative AI were checked by the authors to ensure that they accurately reflected the proposed method.

All AI-assisted text, technical descriptions, references, mathematical expressions, and formatting were reviewed and, where necessary, revised by the authors. We take full responsibility for the final content of this work, including all text, claims, results, and artifacts produced with the aid of generative AI.

Reproducibility statement

We have taken several steps to support the reproducibility of our experiments. The OC20 IS2RE dataset, data splits, evaluation metrics, and experimental settings are described in Section 4. The proposed adsorption-site-guided message-passing framework is described in Section 3.3. Section 3.4 presents its architecture-specific adaptations, using SchNet as a representative backbone, while detailed implementations for DimeNet++ and PaiNN are provided in Appendix A. Training configurations and model settings for the 10k, 100k, and full training sets are reported in the experimental setup. In addition, complete results on the individual OOD validation subsets, including `val_ood_ads`, `val_ood_cat`, and `val_ood_both`, are provided in the appendix to complement the aggregated results reported in the main paper. These descriptions are intended to provide the information necessary to reproduce the reported experiments and evaluate the proposed framework under the same settings.

Acknowledgments

This work was supported by the National Natural Science Foundation of China(No's. 62566025, 62561029), Jiangxi Provincial Natural Science Foundation of China(20242BAB26025),the Jiangxi Provincial Cultivation Program for Academic and Technical Leaders of Major Subjects(20243BCE51080), and the Institutional Scientific Research Project of Jiangxi Vocational College of Mechanical & Electrical Technology(JD260108).

References

Luis Barros-Luque, Muhammed Shuaibi, Xiang Fu, Brandon M Wood, Misko Dzamba, Meng Gao, Ammar Rizvi, Matt Uyttendaele, C Lawrence Zitnick, and Zachary W Ulissi. The open materials 2024 (omat24) inorganic materials dataset and models. *Nature Computational Science*, pp. 1–11, 2026. ISSN 2662-8457.

Federico Calle-Vallejo, David Loffreda, Marc TM Koper, and Philippe Sautet. Introducing structural sensitivity into adsorption–energy scaling relations by means of coordination numbers. *Nature chemistry*, 7(5):403–410, 2015a. ISSN 1755-4330.

Federico Calle-Vallejo, Jakub Tymoczko, Viktor Colic, Quang Huy Vu, Marcus D Pohl, Karina Morgenstern, David Loffreda, Philippe Sautet, Wolfgang Schuhmann, and Aliaksandr S Bandarenka. Finding optimal surface sites on heterogeneous catalysts by counting nearest neighbors. *Science*, 350(6257):185–189, 2015b. ISSN 0036-8075.

Lowik Chanussot, Abhishek Das, Siddharth Goyal, Thibaut Lavril, Muhammed Shuaibi, Morgane Riviere, Kevin Tran, Javier Heras-Domingo, Caleb Ho, and Weihua Hu. Open catalyst 2020 (oc20) dataset and community challenges. *Acs Catalysis*, 11(10):6059–6072, 2021. ISSN 2155-5435.

Junwu Chen, Xu Huang, Cheng Hua, Yulian He, and Philippe Schwaller. A multi-modal transformer for predicting global minimum adsorption energy. *Nature Communications*, 16(1):3232, 2025. ISSN 2041-1723.

Abhishek Das, Muhammed Shuaibi, Aini Palizhati, Siddharth Goyal, Aditya Grover, Adeesh Kolluru, Janice Lan, Ammar Rizvi, Anuroop Sriram, and Brandon M Wood. The open catalyst challenge 2021: Competition report. In *NeurIPS (Competition and Demos)*, pp. 29–40, 2021.

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. *Journal of Machine Learning Research*, 25(106):1–26, 2024. ISSN 1533-7928.

Alexandre Agm Duval, Victor Schmidt, Alex Hernández-García, Santiago Miret, Fragkiskos D Malliaros, Yoshua Bengio, and David Rolnick. Faenet: Frame averaging equivariant gnn for materials modeling. In *International Conference on Machine Learning*, pp. 9013–9033. PMLR, 2023. ISBN 2640-3498.

Xiang Fu, Brandon M Wood, Luis Barroso-Luque, Daniel S Levine, Meng Gao, Misko Dzamba, and C Lawrence Zitnick. Learning smooth and expressive interatomic potentials for physical property prediction. *arXiv preprint arXiv:2502.12147*, 2025.

Johannes Gasteiger, Shankari Giri, Johannes T. Margraf, and Stephan Günnemann. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. In *Machine Learning for Molecules Workshop, NeurIPS*, 2020a. URL https://ml4molecules.github.io/papers2020/ML4Molecules_2020_paper_35.pdf. Earlier version: arXiv:2011.14115.

Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs. In *International Conference on Learning Representations*, 2020b. URL https://openreview.net/forum?id=B1eWbxStPH.

Johannes Gasteiger, Florian Becker, and Stephan Günnemann. Gemnet: Universal directional graph neural networks for molecules. *Advances in Neural Information Processing Systems*, 34:6790–6802, 2021.

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.

Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.

Weihua Hu, Muhammed Shuaibi, Abhishek Das, Siddharth Goyal, Anuroop Sriram, Jure Leskovec, Devi Parikh, and C Lawrence Zitnick. Forcenet: A graph neural network for large-scale quantum calculations. *arXiv preprint arXiv:2103.01436*, 2021.

Hyunwook Jung, Lena Sauerland, Sina Stocker, Karsten Reuter, and Johannes T Margraf. Machine-learning driven global optimization of surface adsorbate geometries. *npj Computational Materials*, 9(1):114, 2023. ISSN 2057-3960.

Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A):A1133, 1965.

Janice Lan, Aini Palizhati, Muhammed Shuaibi, Brandon M Wood, Brook Wander, Abhishek Das, Matt Uyttendaele, C Lawrence Zitnick, and Zachary W Ulissi. Adsorbml: a leap in efficiency for adsorption energy calculations using generalizable machine learning potentials. *npj Computational Materials*, 9(1):172, 2023. ISSN 2057-3960.

Xinyu Li, Raymond Chiong, Zhongyi Hu, and Alister J Page. A graph neural network model with local environment pooling for predicting adsorption energies. *Computational and Theoretical Chemistry*, 1226:114161, 2023. ISSN 2210-271X.

Yujia Li, Richard Zemel, Marc Brockschmidt, and Daniel Tarlow. Gated graph sequence neural networks. In *Proceedings of ICLR'16*, 2015.

Yi-Lun Liao and Tess Smidt. Equiformer: Equivariant graph attention transformer for 3d atomistic graphs. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=KwmPfARgOTD.

Yi-Lun Liao, Brandon M Wood, Abhishek Das, and Tess Smidt. Equiformerv2: Improved equivariant transformer for scaling to higher-degree representations. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=mCOBKZmrzD.

Tianyou Mou, Hemanth Somarajan Pillai, Siwen Wang, Mingyu Wan, Xue Han, Neil M Schweitzer, Fanglin Che, and Hongliang Xin. Bridging the complexity gap in computational heterogeneous catalysis with machine learning. *Nature Catalysis*, 6(2):122–136, 2023. ISSN 2520-1158.

Sergio Pablo-García, Santiago Morandi, Rodrigo A Vargas-Hernández, Kjell Jorner, Žarko Ivković, Núria López, and Alán Aspuru-Guzik. Fast evaluation of the adsorption energy of organic molecules on metals via graph neural networks. *Nature Computational Science*, 3(5):433–442, 2023. ISSN 2662-8457.

Saro Passaro and C Lawrence Zitnick. Reducing so(3) convolutions to so(2) for efficient equivariant gnns. In *International conference on machine learning*, pp. 27420–27438. PMLR, 2023. ISBN 2640-3498.

Ryan Pederson, Bhupalee Kalita, and Kieron Burke. Machine learning and density functional theory. *Nature Reviews Physics*, 4(6):357–358, 2022. ISSN 2522-5820. doi: 10.1038/s42254-022-00470-2.

Zejiang Peng, Tong Chen, Liyun Dai, Mengqiu Long, and Jiancheng Sun. Machine learning-based approach for rapid prediction and analysis of gas-sensing properties of janus materials. *Langmuir*, 42(1):962–973, 2025. ISSN 0743-7463.

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.

Kristof Schütt, Oliver Unke, and Michael Gastegger. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International conference on machine learning*, pp. 9377–9388. PMLR, 2021. ISBN 2640-3498.

Yu Shi, Shuxin Zheng, Guolin Ke, Yifei Shen, Jiacheng You, Jiyan He, Shengjie Luo, Chang Liu, Di He, and Tie-Yan Liu. Benchmarking graphormer on large-scale molecular modeling datasets. *arXiv preprint arXiv:2203.04810*, 2022.

Nima Shoghi, Adeesh Kolluru, John Kitchin, Zachary Ulissi, Larry Zitnick, and Brandon Wood. From molecules to materials: Pre-training large generalizable models for atomic property prediction. In *International Conference on Learning Representations*, volume 2024, pp. 17183–17212, 2024.

Muhammed Shuaibi, Adeesh Kolluru, Abhishek Das, Aditya Grover, Anuroop Sriram, Zachary Ulissi, and C Lawrence Zitnick. Rotation invariant graph neural networks using spin convolutions. *arXiv preprint arXiv:2106.09575*, 2021.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=rJXMpikCZ.

Charlotte Vogt and Bert M Weckhuysen. The concept of active site in heterogeneous catalysis. *Nature Reviews Chemistry*, 6(2):89–111, 2022. ISSN 2397-3358.

Brandon M Wood, Misko Dzamba, Xiang Fu, Meng Gao, Muhammed Shuaibi, Luis Barroso-Luque, Kareem Abdelmaqsoud, Vahe Gharakhanyan, John R Kitchin, and Daniel S Levine. Uma: A family of universal models for atoms. *arXiv preprint arXiv:2506.23971*, 2025.

Larry Zitnick, Abhishek Das, Adeesh Kolluru, Janice Lan, Muhammed Shuaibi, Anuroop Sriram, Zachary Ulissi, and Brandon Wood. Spherical channels for modeling atomic interactions. *Advances in Neural Information Processing Systems*, 35:8054–8067, 2022.

A Architecture-Specific Adaptations for DimeNet++ and PaiNN

Building on the unified framework, we adapt the integration of adsorption-site information into each backbone according to its feature representations and message passing mechanism. The adaptations for DimeNet++ and PaiNN are described below, while the SchNet implementation is presented in Section 3.4

A.1 ADAPTATION FOR DIMENET++

DimeNet++ primarily performs message passing through directed-edge representations. Accordingly, adsorption-site information is incorporated into the directed-edge representations updated by the interaction blocks. Similar to the node-level implementation in SchNet, we first encode the distances between atoms and the adsorption-site using RBFs to obtain atom-level site representations. For each directed edge $j \to i$, the site representations of its source j and target i are concatenated and projected to obtain an edge-level site feature:

$$\boldsymbol{p}_{ji} = \mathcal{F}_{\text{edge}}\left([\boldsymbol{p}_{i,\text{ads}} \parallel \boldsymbol{p}_{j,\text{ads}}]\right), \tag{8}$$

where $\parallel$ denotes feature concatenation, $\mathcal{F}_{\text{edge}}(\cdot)$ denotes the mapping to edge level site features. To restrict the influence of site information to a local neighborhood, we define an edge relevance score based on the distances of the two endpoint atoms from the adsorption-site:

$$\rho_{ji} = \max\left(\rho_i,\, \rho_j\right), \tag{9}$$

where ρ_i and ρ_j are computed using a smooth distance cutoff function. When both endpoints of an edge lie outside the cutoff region around the adsorption-site, $\rho_{ji} = 0$, so that site information does not contribute to the update of that edge.

Let $\boldsymbol{x}_{ji}^{(l)}$ denote the directed-edge representation output by the l-th interaction block. The edge-level site feature is weighted by the edge relevance ρ_{ji} and then fused with the original edge representation:

$$\bar{\boldsymbol{x}}_{ji}^{(l)} = \text{LayerNorm}\left(\boldsymbol{x}_{ji}^{(l)} + \rho_{ji}\boldsymbol{p}_{ji}\right). \tag{10}$$

Next, the fused edge representation $\bar{\boldsymbol{x}}_{ji}^{(l)}$ is concatenated with the edge-level site feature $\boldsymbol{p}_{ji}$, and the resulting representation is passed through a two-layer MLP followed by a sigmoid function to generate channel-wise gating weights $\boldsymbol{g}_{ji}^{(l)}$. These weights adaptively modulate the contribution of different feature channels. Finally, the original edge representation is updated through a residual connection:

$$\boldsymbol{x}_{ji}^{(l+1)} = \boldsymbol{x}_{ji}^{(l)} + \alpha\rho_{ji}\,\boldsymbol{g}_{ji}^{(l)} \odot \left(\bar{\boldsymbol{x}}_{ji}^{(l)} - \boldsymbol{x}_{ji}^{(l)}\right), \tag{11}$$

where $\boldsymbol{x}_{ji}^{(l+1)}$ denotes the edge representation after the adsorption-site-guided gated update, and α is a learnable residual update coefficient.

A.2 ADAPTATION FOR PAINN

For PaiNN, following its coupled scalar–vector message-passing scheme, we apply adsorption-site information fusion and gating to the RBF representations of edge distances. Specifically, edge-level site features are constructed from the spatial relationships between the two endpoint atoms and the adsorption-site, and are incorporated into the original RBF representations through channel-wise gating and a residual connection. The updated RBF representations are then used in subsequent message passing, allowing adsorption-site information to further influence the updates of both scalar and vector features in PaiNN.

B DETAILED RESULTS ON INDIVIDUAL OOD SUBSETS

The main text uses Aggregated OOD for an overall evaluation of the three OOD validation subsets. To further investigate model performance under different types of distribution shift, Table 3 reports the results for SchNet, DimeNet++, and PaiNN, together with their adsorption-site- guided variants, on the three individual validation subsets: `val_ood_cat`, `val_ood_ads` and `val_ood_both`.

Overall, the results on the individual OOD subsets follow the same trend as the Aggregated OOD results reported in the main text. At the 10k training set size, all three adsorption-site-guided models outperform their corresponding baselines in terms of MAE, MSE, and EwT across all three OOD subsets, with particularly pronounced improvements for DimeNet++. For example, for DimeNet++ on `val_ood_both`, MAE decreases from 0.8780 to 0.7063 and MSE from 1.4992 to 1.0296, while EwT increases from 0.0158 to 0.0226. These results further suggest that, in the low data regime, local structural information associated with adsorption-site can complement the geometric representations of the backbone models across different OOD settings.

Table 3: Comparison of baseline and adsorption-site-guided models on the three OOD validation subsets.

Train set	Model name	val_ood_cat			val_ood_ads			val_ood_both		
		MAE↓	MSE↓	EwT↑	MAE↓	MSE↓	EwT↑	MAE↓	MSE↓	EwT↑
10k	SchNet	1.052	2.264	0.0153	1.0902	2.36	0.0133	1.0167	1.8927	0.0136
	SchNet-Ours	**0.9398**	**1.7225**	**0.0168**	**0.9814**	**2.0624**	**0.0159**	**0.8913**	**1.5445**	**0.0162**
	DimeNet++	0.8425	1.4988	0.0204	0.9576	1.8542	0.0161	0.8780	1.4992	0.0158
	DimeNet++-Ours	**0.7484**	**1.2498**	**0.0241**	**0.7768**	**1.3479**	**0.0219**	**0.7063**	**1.0296**	**0.0226**
	PaiNN	0.9066	1.6127	0.0164	0.9565	1.8356	0.0156	0.8774	1.4715	0.0163
	PaiNN-Ours	**0.8701**	**1.5095**	**0.0173**	**0.9303**	**1.7719**	**0.0163**	**0.8642**	**1.4112**	**0.0176**
100k	SchNet	0.6907	**1.0921**	0.0267	0.8145	1.5104	0.0196	0.7148	**1.06**	0.0211
	SchNet-Ours	**0.6707**	1.1505	**0.0311**	**0.7391**	**1.2295**	**0.0222**	**0.6857**	1.0713	**0.0218**
	DimeNet++	0.6236	0.9981	0.0357	0.7583	1.2501	**0.0214**	0.6955	1.062	**0.0253**
	DimeNet++-Ours	**0.6028**	**0.9137**	**0.0362**	**0.7499**	**1.2117**	0.0207	**0.678**	**0.9707**	0.0242
	PaiNN	0.6662	1.0296	0.0281	0.7072	1.1772	0.0244	0.6538	0.9256	0.0242
	PaiNN-Ours	**0.6422**	**0.9795**	**0.0297**	**0.7016**	**1.1388**	**0.0262**	**0.6496**	**0.9074**	**0.0259**
all	SchNet	0.6478	1.0444	0.0301	0.7074	1.1263	0.0222	0.6627	0.967	0.0238
	SchNet-Ours	**0.5984**	**0.9021**	**0.0328**	**0.7032**	**1.0920**	**0.0230**	**0.6402**	**0.8755**	**0.0251**
	DimeNet++	0.5617	0.8855	**0.044**	0.7127	1.1275	0.0247	0.6493	0.999	0.0256
	DimeNet++-Ours	**0.5467**	**0.7929**	0.0401	**0.6799**	**1.0603**	**0.0263**	**0.6092**	**0.8273**	**0.0278**
	PaiNN	0.5864	0.8976	0.0326	**0.6478**	1.0075	0.0262	0.5962	0.8117	0.0283
	PaiNN-Ours	**0.5585**	**0.8156**	**0.0375**	0.6511	**0.9814**	**0.0264**	**0.5948**	**0.8073**	**0.0295**

At the 100k and All training-set sizes, the gains on the individual OOD subsets are generally smaller than those at 10k, although most model–metric combinations still improve over their corresponding baselines. A few metrics show slight degradation. At the 100k training-set size, DimeNet++ shows lower EwT on `val_ood_ads` and `val_ood_both`, while at the All training-set size, PaiNN shows a slight increase in MAE on `val_ood_ads`. Overall, the results across the three individual OOD subsets are consistent with the overall improvement trend observed for Aggregated OOD in the main text.

C Visualization of Adsorption-Site-Guided Responses

To further analyze the internal response patterns of the adsorption-site-guided variants of SchNet and DimeNet++, we use NO_2 adsorption on the VSSe surface as a representative example and visualize their atom level responses, as shown in Figure 4. The atom-level responses are derived from the atomic representations before graph level aggregation, and their spatial distributions are visualized using a color map. Both models exhibit stronger responses around NO_2, while the responses of the VSSe substrate atoms are generally weaker. To further reveal the response distribution within the substrate, the NO_2 molecule is omitted from the visualization and the color scale is rescaled in Figure 4(c–d) and Figure 4(g–h). These adjustments are made solely for visualization purposes. Within the substrate, both models exhibit non-uniform spatial responses with distinct distribution patterns. SchNet shows relatively weaker responses in atomic layers near the adsorption-site and stronger responses in more distant layers, whereas DimeNet++ exhibits stronger responses near the adsorption-site that gradually decay with distance. These differences may be related to the different information aggregation and message-passing mechanisms of the two backbone models. The atom-level responses are derived from internal model features and are intended only for comparing spatial response patterns within the models; they should not be interpreted as direct physical contributions of individual atoms to the system energy.

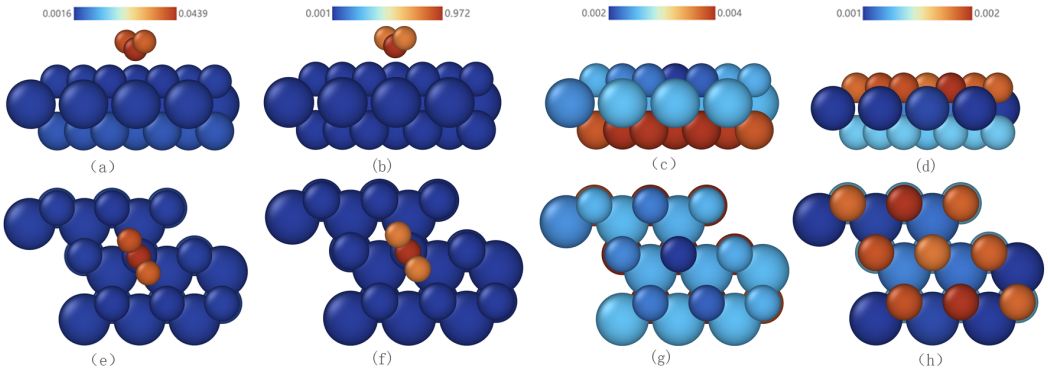

Figure 4: Atom-level response distributions of adsorption-site-guided SchNet and DimeNet++. (a,e,c,g) SchNet; (b,f,d,h) DimeNet++. (a–d) side views; (e–h) Top views; (c–d, g–h) VSSe substrate responses. Colors indicate atom-level response values.