WallpaperNet: A $p6mm$-Equivariant Graph Neural Network for Molecule Adsorption on Graphene

Rostislav Fedorov
HITS gGmbH
Heidelberg, Germany
rostislav.fedorov@h-its.org

Ganna (Anya) Gryn'ova
University of Birmingham
Birmingham, UK
g.grynova@bham.ac.uk

Abstract

We present WallpaperNet, a $p6mm$-group-equivariant graph neural network (GNN) for modeling adsorption of small molecules on graphene. Unlike conventional approaches that freeze surface atoms and operate on large system sizes, our method focuses exclusively on the adsorbate atoms, which allows fast AI-guided design of molecule-graphene composite systems. We encode the symmetry of the underlying hexagonal lattice through Wyckoff-Anchor vector encoding and D_6-equivariant attention mechanism.

1 Introduction

Adsorption of small molecules on graphene, graphite or layered graphene plays a big role in sensing, catalysis, electronics, and energy storage.[1–3] In a common simulation scenario [4–8] the graphene/graphite atoms are frozen, while the small molecule is being relaxed. The ratio between atoms that are actively moving and atoms of the surface, that are frozen, and build a hexagonal grid is usually around 25/300. The purpose of this work is to develop an efficient $p6mm$-Equivariant GNN, that would operate only on the atoms of the small molecule, but take into account the underlying hexagonal grid. This would allow for efficient AI-guided design of molecule-graphene composite systems. Figure 1 (right) illustrates the problem of $p6mm$-equivariance.

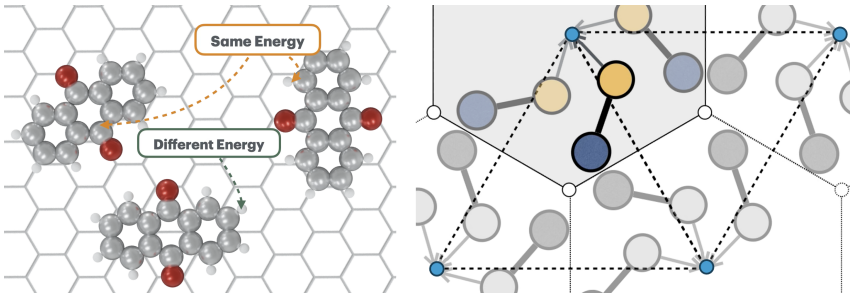

Figure 1: Left: a molecule at two different co-variant positions, yielding the same adsorption energy and a molecule at a non-equivariant position, yielding a different adsorption energy. Right: illustration of the Wyckoff-Anchor vector for the yellow atom (paired with the blue atom) and co-variantly rotated and reflected versions of the same atom atom-pair.

To address these challenges, we propose a $p6mm$-group-equivariant GNN architecture. To handle translational equivariance we introduce a Wyckoff-Anchor vector encoding, that is covariant under the action of the finite group $p6mm$, but can be extended to any Wallpaper Group. The main idea of

39th Conference on Neural Information Processing Systems (NeurIPS 2025).

the architecture is based on two cornerstones: the E(n)-GNN [9] and its extension for a transformer [10]; and E(n)-Equivariant Steerable CNNs with arbitrary $G \leq O(2)$ steerable kernels [11]. The architecture is designed to operate on the small molecule atoms only, while taking into account the underlying hexagonal grid.

Geometric deep learning enables the principled incorporation of symmetry constraints into machine-learning models [12, 13]. To the best of our knowledge, the only work, that extends a finite-group equivariance to the notion of graphs is done by Zhdanov and Cesa [14] and this is the first work on the finite group equivariant attention on graphs. A convolution on a 2d hexagonal image grid was proposed in [15]. Also originally envisioned for the $p6mm$ group equivariance, this approach can be extended to other finite group of interest, given the fact that restraining to a certain space group is a common approach in computational chemistry. [16, 17]. A short background theory section on group theory and equivariant GNNs provided in the Appendix A.

2 Finite group equivariance learning

To enforce a $p6mm$-equivariance of the hexagonal grid, we need to solve two problems: translational and rotational equivariance. We consider a freestanding, infinite graphene monolayer modeled by the wallpaper group $G = p6mm$.

2.1 Wyckoff-Anchor Vector Encoding

Let $\Lambda \subset \mathbb{R}^2$ be the 2D Bravais lattice[18] and let H denote the in-plane point group. In crystallographic notation $H = 6\mathrm{mm}$ and in group-theoretic notation $H \cong C_{6v} \cong D_6$.[18] [1] We write $\rho : H \to O(2)$ for the standard 2×2 orthogonal representation acting on in-plane vectors. For any lattice translation $t \in \Lambda$ we set $\rho(t) = I_2$. Let $A = \{a^{(1)}, \ldots, a^{(m)}\} \subset \mathbb{R}^2$ be a finite set of highest-symmetry Wyckoff positions (anchors) of the crystal; for graphene we may take the hexagon centers. By construction A is G-invariant: for every $g \in G$, $gA = A$ up to permutation. For an adsorbate atom with in-plane position $x_i \in \mathbb{R}^2$,we define its unique nearest anchor (with periodic boundary conditions)

$$a(x_i) \;:=\; \arg\min_{a \in A} \min_{t \in \Lambda} \|x_i - (a + t)\|, \tag{1}$$

which is unique inside the Wigner–Seitz cell. We then define the *Wyckoff–Anchor vector embedding* as the 2D pointer

$$v_{\mathrm{wa}}(x_i) \;:=\; a(x_i) - x_i \;\in\; \mathbb{R}^2. \tag{2}$$

Equivariance. For any $g \in G$ and position x, v_{wa} satisfies

$$v_{\mathrm{wa}}(gx) \;=\; \rho(g)\, v_{\mathrm{wa}}(x). \tag{3}$$

Proof. Since g is an isometry, $\|gx - g(a + t)\| = \|x - (a + t)\|$. Because $gA = A$ and $g\Lambda = \Lambda$, the set of candidates is unchanged, hence $a(gx) = ga(x)$. Therefore, $v_{\mathrm{wa}}(gx) = a(gx) - gx = ga(x) - gx = \rho(g)\,[a(x) - x] = \rho(g)\, v_{\mathrm{wa}}(x)$. For pure translations $t \in \Lambda$, $a(x + t) = a(x) + t$ so $v_{\mathrm{wa}}(x + t) = v_{\mathrm{wa}}(x)$.

Pairwise positional coupling. Given interatomic displacements $r_{ij} = x_j - x_i$, we use the rotation-aware scalar couplings

$$s_{ij} \;=\; \langle v_{\mathrm{wa}}(x_i),\, r_{ij} \rangle, \qquad \|v_{\mathrm{wa}}(x_i)\|, \tag{4}$$

which are H-invariant and translation-invariant, and enter the message function as additional scalar channels. Figure 1 (right) illustrates the Wyckoff-Anchor vector.

2.2 Finite-rotation invariant block

Let $z = x + iy \in \mathbb{C}$ be the complexified in-plane vector and a fix $n \in \mathbb{N}$. Define

$$\phi_n : \mathbb{R}^2 \to \mathbb{R}^2, \qquad \phi_n(x, y) \;=\; \big(\Re(z^n),\, \Im(z^n)\big). \tag{5}$$

[1] The 2D, in-plane point group is $6\mathrm{mm} \equiv C_{6v} \simeq D_6$. If one includes the mirror with respect to the plane (out-of-plane symmetry), the 3D Schoenflies notation D_{6h} is sometimes used. Here we work in 2D and use D_6 for the dihedral group of order 12.

Invariance under C_n. For any rotation R_α, $\phi_n(R_\alpha z) = R_{n\alpha}\,\phi_n(z)$. Hence, for $\alpha = \frac{2\pi k}{n}$ (i.e., $R_\alpha \in C_n$), $R_{n\alpha} = I$ and ϕ_n is C_n-invariant. (proof in the Appendix A)

Behavior under reflections. For any reflection σ through a line through the origin, $z \mapsto \bar{z}$ up to a phase; thus $z^n \mapsto \overline{z^n}$. Consequently, $\Re(z^n)$ is reflection-even and $\Im(z^n)$ is reflection-odd, so $(\Re(z^n))$ stays the same and $(\Im(z^n))$ flips the sign under a reflection of D_n.

Injectivity on the orbit space. Writing $z = re^{i\theta}$, we have $z^n = r^n e^{in\theta}$. Thus, ϕ_n induces an injective map on the quotient $(\mathbb{R}^2\backslash\{0\})/C_n$: two vectors have the same image iff their radii coincide and their angles differ by a multiple of $2\pi/n$ (i.e., they are in the same C_n orbit). In practice, passing ϕ_n through an MLP together with fully SO(2) invariant features (e.g., atom types) lets the network learn features that are either injective on C_n-orbits or deliberately periodic when injectivity is not needed (e.g., when atoms are far from the sheet along z).

2.3 Finite-rotation equivariant block

Let $r_1, \ldots, r_m \in \mathbb{R}^2$ be non-colinear input vectors (e.g., edge displacements, anchor pointers), and let α_i be D_n/C_n-invariant scalars (learned from node / edge features). Then

$$v \;=\; \sum_{i=1}^{m} \alpha_i\, r_i \tag{6}$$

is C_n–equivariant: for any $g \in C_n$, $v \mapsto \sum_i \alpha_i\, \rho(g) r_i = \rho(g) \sum_i \alpha_i r_i$. This construction is used to form Q, K, V vectors for attention while keeping finite-rotation symmetry.

2.4 Regular-representation equivariant block

To increase expressivity while preserving equivariance, we lift features into the regular representation $\mathbb{R}[G]$ with $G \in \{C_n, D_n\}$. Each channel corresponds to a group element and transforms by channel permutation under G. Hence pointwise nonlinearities are automatically G-equivariant. Stacking linear maps

$$V \xrightarrow{L_1} \mathbb{R}[G] \xrightarrow{\sigma} \mathbb{R}[G] \xrightarrow{L_2} V$$

with pointwise σ yields a *regular-rep MLP* that is guaranteed to be G-equivariant and universal for G-equivariant maps. The derivation for the learnable linear map between different representations of different finite groups is very well documented and provided in [11].

2.5 WallpaperNet - architecture overview

We represent the molecular adsorbate + sheet as a graph $G = (V, E)$ with node features h_i, positions $x_i = (x_i^{(x)}, x_i^{(y)}, x_i^{(z)})$, and edges (i, j) of the molecule. The core layer implements three ingredients:

(i) Invariant scalar message channels. For each edge (i, j), form $r_{ij} = x_j - x_i$, its squared norm $\|r_{ij}\|^2$, the C_6 embedding $\phi_6(r_{ij}^{xy}) = (\Re[(x + iy)^6], \Im[(x + iy)^6])$, the anchor couplings $s_{ij} = \langle v_{\mathrm{wa}}(x_i), r_{ij}^{xy} \rangle$ and $\|v_{\mathrm{wa}}(x_i)\|$. These feed an MLP to produce messages m_{ij}, which are pooled to node updates $m_i = \sum_j m_{ij}$ and combined with a residual to update h_i.

(ii) Equivariant vector aggregation. A learned scalar weight $w_{ij} = \psi(m_{ij})$ multiplies the displacement r_{ij} to produce edge vectors $f_{ij} = w_{ij}\, r_{ij}$. Summing over neighbors yields an equivariant node vector $\sum_j f_{ij}$. We treat the z component with a separate scalar MLP.

(iii) Multi-head radial attention (equivariant). Using only invariant scalars m_{ij}, we form per-head weights w_q, w_k, w_v. Queries/keys/values are made equivariant by scaling r_{ij} radially, e.g., $q_{ij}^{(h)} = w_q^{(h)}(m_{ij})\, r_{ij}$, etc. Attention logits are invariant $\langle \widehat{Q}_i^{(h)}, \widehat{K}_{ij}^{(h)} \rangle$ (unit-normalized), softmaxed over j, and values are aggregated equivariantly. A scalar head-mixing reweights heads at the end. We pass the in-plane (x, y) part through a small regular-rep network to increase expressivity while preserving D_6–equivariance, that lifts to the regular representation, applies pointwise nonlinearities (equivariant in $\mathbb{R}[D_6]$ because the group action is a channel permutation), and projects back.

Table 1: Force prediction performance

Component	MAE Test / Train [eV / Å]	R^2 Test / Train
X-axis	0.053±0.001 / 0.040±0.001	0.90±0.01 / 0.96±0.1
Y-axis	0.051±0.001 / 0.040±0.001	0.90±0.01 / 0.96±0.1
Z-axis	0.036±0.002 / 0.025±0.002	0.81±0.2 / 0.94±0.1

3 Experimental Setup

We evaluate the performance of the proposed architecture on a dataset of small molecules adsorbed on graphene, with a focus on predicting binding energies (invariant property) and atomic forces (equivariant property). The dataset is generated using GFN2-xTb, accuracy was verified by benchmarking against DFT. Overall, we have 15087 systems, each with on average 300 atoms of graphene and 25 atoms of the adsorbed molecule. The average length of the relaxation process is 201 steps, with a maximum of 714 steps. More details about the experimental setup and data generation and regression plots can be found in the Appendix B.

Force vector prediction. The model is trained on both the initial position of the geometry optimization, the train/validation/test split is 80/10/10. For the force evaluation setup, we took the first relaxation step of every system. The force vector in the sets set has an average magnitude of 0.37 eV / Å and a maximum of 4.52 eV / Å and minimum of 0.0005 eV / Å. The train set has an average of 0.31 eV / Å, minimum of 0.003 eV / Å and the maximum of 4.1 eV / Å.

Table 1 summarizes the force prediction performance of our model. Overall, we can learn the force field with a MAE of 0.047 eV / Å. The x/y prediction should rotate equivariantly to the $p6mm$, while z component should stay invariant. The values reported as averages across three runs. Interestingly, the z-component has the smallest MAE but the lowest R^2. This is likely due to the fact that the z-component has a smaller variance, as the molecules are adsorbed on the surface, and thus the forces in the z-direction are generally smaller. A further investigation of the difference between prediction of equivariant and invariant parts of the vector is needed.

Binding energy prediction. The binding energy of the dataset spans between 0.01 eV to 1.51 eV. We evaluate performance of the model in two different scenarios, from the fully relaxed configuration and from initial position. In both cases train/test/validation split remains the same. The results are presented in the Table 2.

Table 2: Binding energy prediction performance

Property	MAE Train / Test [eV]	R^2 Train / Test
Binding energy (fully relaxed)	0.020±0.001 / 0.009±0.003	0.987±0.002 / 0.989±0.001
Binding energy (initial position)	0.064±0.005 / 0.052±0.003	0.78±0.01 / 0.902±0.03

As we can see from the results, prediction from the initial positions allows for a fast pre-screening of potential binding sites, and knowing the relaxed position allows us to predict near-perfectly the binding energy. We showcase the ablation study of different components of the architecture as well as a performance comparison with E(n)-GNN in the Appendix C.

4 Conclusion

In this work, we have presented a novel approach for predicting the binding energy of small molecules on graphene surfaces using graph neural networks. Our method leverages the unique symmetry properties of the graphene lattice, allowing for efficient and accurate predictions. We have demonstrated the effectiveness of our approach for a real computational chemistry problem. Extension of this method to other finite groups is a promising direction for future research.

References

[1] Lingmei Kong, Axel Enders, Talat S Rahman, and Peter A Dowben. Molecular adsorption on graphene. 26(44):443001. ISSN 0953-8984, 1361-648X. doi: 10.1088/0953-8984/26/44/443001. URL https://iopscience.iop.org/article/10.1088/0953-8984/26/44/443001.

[2] Guo Hong, Qi-Hui Wu, Jianguo Ren, Chundong Wang, Wenjun Zhang, and Shuit-Tong Lee. Recent progress in organic molecule/graphene interfaces. 8(4):388–402. ISSN 17480132. doi: 10.1016/j.nantod.2013.07.003. URL https://linkinghub.elsevier.com/retrieve/pii/S1748013213000704.

[3] Gamze Ersan, Onur G. Apul, Francois Perreault, and Tanju Karanfil. Adsorption of organic contaminants by graphene nanosheets: A review. 126:385–398. ISSN 00431354. doi: 10.1016/j.watres.2017.08.010. URL https://linkinghub.elsevier.com/retrieve/pii/S0043135417306632.

[4] Tsan-Yao Chen, Yanhui Zhang, Liang-Ching Hsu, Alice Hu, Yu Zhuang, Chia-Ming Fan, Cheng-Yu Wang, Tsui-Yun Chung, Cheng-Si Tsao, and Haw-Yeu Chuang. Crystal shape controlled H2 storage rate in nanoporous carbon composite with ultra-fine Pt nanoparticle. 7 (1):42438. ISSN 2045-2322. doi: 10.1038/srep42438. URL https://www.nature.com/articles/srep42438.

[5] Elisa Londero, Emma K Karlson, Marcus Landahl, Dimitri Ostrovskii, Jonatan D Rydberg, and Elsebeth Schröder. Desorption of n-alkanes from graphene: A van der Waals density functional study. 24(42):424212. ISSN 0953-8984, 1361-648X. doi: 10.1088/0953-8984/24/42/424212. URL https://iopscience.iop.org/article/10.1088/0953-8984/24/42/424212.

[6] Eunice Cunha, Maria Fernanda Proença, Maria Goreti Pereira, Maria José Fernandes, Robert J. Young, Karol Strutyński, Manuel Melle-Franco, Mariam Gonzalez-Debs, Paulo E. Lopes, and Maria Da Conceição Paiva. Water Dispersible Few-Layer Graphene Stabilized by a Novel Pyrene Derivative at Micromolar Concentration. 8(9):675. ISSN 2079-4991. doi: 10.3390/nano8090675. URL https://www.mdpi.com/2079-4991/8/9/675.

[7] Young-Joon Song, Charlotte Gallenkamp, Genís Lleopart, Vera Krewald, and Roser Valentí. Influence of graphene on the electronic and magnetic properties of an iron(III) porphyrin chloride complex. 26(41):26370–26376. ISSN 1463-9076, 1463-9084. doi: 10.1039/D4CP01551G. URL https://xlink.rsc.org/?DOI=D4CP01551G.

[8] Shane R. Carlson, Otto Schullian, Maximilian R. Becker, and Roland R. Netz. Modeling Water Interactions with Graphene and Graphite via Force Fields Consistent with Experimental Contact Angles. 15(24):6325–6333. ISSN 1948-7185, 1948-7185. doi: 10.1021/acs.jpclett.4c01143. URL https://pubs.acs.org/doi/10.1021/acs.jpclett.4c01143.

[9] Victor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) Equivariant Graph Neural Networks. URL https://arxiv.org/abs/2102.09844.

[10] Phil Wang Phillip Wang. En-transformer. URL https://github.com/lucidrains/En-transformer.

[11] Gabriele Cesa, Leon Lang, and Maurice Weiler. A program to build E(N)-equivariant steerable CNNs. In *International Conference on Learning Representations (ICLR)*. URL https://openreview.net/forum?id=WE4qe9xlnQw.

[12] Haitz Sáez de Ocáriz Borde and Michael Bronstein. Mathematical Foundations of Geometric Deep Learning. URL https://arxiv.org/abs/2508.02723.

[13] Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges.

[14] Maksim Zhdanov, Nico Hoffmann, and Gabriele Cesa. Implicit Convolutional Kernels for Steerable CNNs. URL https://arxiv.org/abs/2212.06096.

[15] Emiel Hoogeboom, Jorn W. T. Peters, Taco S. Cohen, and Max Welling. HexaConv. URL https://arxiv.org/abs/1803.02108.

[16] Sam Cox and Andrew D. White. Symmetric Molecular Dynamics. 18(7):4077–4081. ISSN 1549-9618, 1549-9626. doi: 10.1021/acs.jctc.2c00401. URL https://pubs.acs.org/doi/10.1021/acs.jctc.2c00401.

[17] Andriy Anishkin, Adina L. Milac, and H. Robert Guy. Symmetry-restrained molecular dynamics simulations improve homology models of potassium channels. 78(4):932–949. ISSN 0887-3585, 1097-0134. doi: 10.1002/prot.22618. URL https://onlinelibrary.wiley.com/doi/10.1002/prot.22618.

[18] Mildred S. Dresselhaus, Gene Dresselhaus, and Ado Jorio, editors. *Group Theory*. SpringerLink Bücher. Springer Berlin Heidelberg. ISBN 978-3-540-32897-1 978-3-540-32899-5. doi: 10.1007/978-3-540-32899-5.

[19] Carolyn Pratt Brock, T Hahn, H Wondratschek, U Müller, U Shmueli, E Prince, A Authier, V Kopskỳ, D Litvin, E Arnold, et al. International tables for crystallography volume A: Space-group symmetry.

[20] Alexandre Duval, Simon V. Mathis, Chaitanya K. Joshi, Victor Schmidt, Santiago Miret, Fragkiskos D. Malliaros, Taco Cohen, Pietro Liò, Yoshua Bengio, and Michael Bronstein. A Hitchhiker's Guide to Geometric GNNs for 3D Atomic Systems. URL https://arxiv.org/abs/2312.07511.

[21] Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International Conference on Machine Learning*, pages 2990–2999. PMLR.

[22] Maurice Weiler, Mario Geiger, Max Welling, Wouter Boomsma, and Taco Cohen. 3D Steerable CNNs: Learning Rotationally Equivariant Features in Volumetric Data. URL https://arxiv.org/abs/1807.02547.

[23] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, Eric D Hermes, Paul C Jennings, Peter Bjerre Jensen, James Kermode, John R Kitchin, Esben Leonhard Kolsbjerg, Joseph Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jón Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob Schiøtz, Ole Schütt, Mikkel Strange, Kristian S Thygesen, Tejs Vegge, Lasse Vilhelmsen, Michael Walter, Zhenhua Zeng, and Karsten W Jacobsen. The atomic simulation environment—a Python library for working with atoms. 29(27):273002. ISSN 0953-8984, 1361-648X. doi: 10.1088/1361-648X/aa680e. URL https://iopscience.iop.org/article/10.1088/1361-648X/aa680e.

[24] Greg Landrum, Paolo Tosco, Brian Kelley, Ricardo Rodriguez, David Cosgrove, Riccardo Vianello, sriniker, Peter Gedeck, Gareth Jones, Eisuke Kawashima, NadineSchneider, Dan Nealschneider, Andrew Dalke, Matt Swain, Brian Cole, Samo Turk, Aleksandr Savelev, useprefix=true family=cdd, prefix=tadhurst, Alain Vaucher, Maciej Wójcikowski, Ichiru Take, Rachel Walker, Vincent F. Scalfani, Hussein Faara, Kazuya Ujihara, Daniel Probst, Juuso Lehtivarjo, useprefix=true family=godin, prefix=guillaume, Axel Pahl, and Niels Maeder. Rdkit/rdkit: 2025_03_2 (Q1 2025) Release. URL https://zenodo.org/doi/10.5281/zenodo.15286010.

[25] Rostislav Fedorov, Anastasiia Nihei, and Ganna Gryn'ova. Multi-Solvent Graph Neural Network for Reduction Potential Prediction across the Chemical Space. URL https://chemrxiv.org/engage/chemrxiv/article-details/685aa5243ba0887c335198f4.

[26] Christoph Bannwarth, Sebastian Ehlert, and Stefan Grimme. GFN2-xTB—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. 15(3):1652–1671. ISSN 1549-9626. doi: 10.1021/acs.jctc.8b01176. URL http://dx.doi.org/10.1021/acs.jctc.8b01176.

[27] B. Hourahine, B. Aradi, V. Blum, F. Bonafé, A. Buccheri, C. Camacho, C. Cevallos, M. Y. Deshaye, T. Dumitrică, A. Dominguez, S. Ehlert, M. Elstner, T. Van Der Heide, J. Hermann, S. Irle, J. J. Kranz, C. Köhler, T. Kowalczyk, T. Kubař, I. S. Lee, V. Lutsker, R. J. Maurer, S. K. Min, I. Mitchell, C. Negre, T. A. Niehaus, A. M. N. Niklasson, A. J. Page, A. Pecchia, G. Penazzi, M. P. Persson, J. Řezáč, C. G. Sánchez, M. Sternberg, M. Stöhr, F. Stuckenberg, A. Tkatchenko, V. W.-z. Yu, and T. Frauenheim. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. 152(12):124101. ISSN 0021-9606, 1089-7690. doi: 10.1063/1.5143190. URL `https://pubs.aip.org/jcp/article/152/12/124101/953756/DFTB-a-software-package-for-efficient-approximate`.

[28] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An Imperative Style, High-Performance Deep Learning Library. URL `http://arxiv.org/abs/1912.01703`.

[29] Matthias Fey and Jan Eric Lenssen. Fast Graph Representation Learning with PyTorch Geometric. URL `http://arxiv.org/abs/1903.02428`.

[30] Maurice Weiler and Gabriele Cesa. General $E(2)$-Equivariant Steerable CNNs. URL `http://arxiv.org/abs/1911.08251`.

A Technical appendices and Supplementary Material

Wallpaper groups and graphene. Graphene is a two-dimensional hexagonal lattice with wallpaper group *p6mm*. The in-plane point group is $6\mathrm{mm}$ (Schoenflies $C_{6v} \cong D_6$), acting by planar rotations and reflections [18, 19]. Wallpaper groups combine a planar Bravais lattice with a finite point group acting on that lattice. Throughout, we assume an ideal, infinite monolayer and focus on in-plane symmetry.

Wyckoff positions and anchors. *Wyckoff positions* are sets of points in a crystal that share the same site symmetry (stabilizer subgroup) [19]. For *p6mm*, high-symmetry choices include hexagon centers, vertices, and edge midpoints. We refer to a chosen finite set of such positions within a unit cell as *anchors*. Under any symmetry operation of *p6mm*, the set of anchors maps to itself (up to permutation).

Equivariance and invariant/equivariant features. Equivariant message passing is now standard for 3D molecules [9, 20]. A function f from geometric data to predictions is *equivariant* to a group G if $f(g \cdot x) = \rho(g)f(x)$ for all $g \in G$ and an output representation ρ. Invariant predictions have $f(g \cdot x) = f(x)$. Group-equivariant neural networks exploit this structure for sample efficiency and inductive bias [11, 13, 21, 22].

Invariance to the cyclic subgroup C_n and reflections. Let R_θ denote rotation by angle θ, acting as $R_\theta \cdot z = e^{i\theta} z$. Then for any $\theta = k\gamma = \frac{2\pi k}{n}$ with $k \in \mathbb{Z}$ we have $e^{in\theta} = e^{i2\pi k} = 1$, so

$$\phi_n(R_{k\gamma} \cdot (x_{ij}, y_{ij})) = \big(\mathrm{Re}(e^{i\frac{2\pi k}{n}*n} z^n),\, \mathrm{Im}(e^{i\frac{2\pi k}{n}*n} z^n)\big) = \phi_n(x_{ij}, y_{ij}). \tag{7}$$

For a reflection, $z \mapsto \bar{z}$ (up to a phase), so $z^n \mapsto \overline{z^n}$, yielding $\Re(z^n)$ even and $\Im(z^n)$ odd.

B Data generation details

All molecular preprocessing and adsorption model setup were performed using the ASE [23] and the RDKit [24]. The initial molecular geometries were taken from [25]. All structures comprise one organic adsorbate anchored on a three-layer graphene surface. The graphene sheets of pristine monolayer graphene were stacked in AAA order. The molecular placement and PBC cell construction were performed using our molecule-placing algorithm. The graphene atoms remained frozen, while atoms of the absorbed molecule were relaxed using extended tight-binding model GFN2-xTB [26]

using Γ-point only sampling with the DFTB+ software [27] until the maximum force acting on each atom was below 0.01 eV Å. The xTB protocol was chosen, based on the result of a benchmarking study with DFT of 300 systems. The linearly fitted binding energies from GFN2-xTB to DFT (PBE-D3/PAW) yielded a mean absolute error of 0.19 eV and R^2 of 0.86. The exact simulation protocol goes beyond the scope of this paper and will be published elsewhere.

C Experimental Details and Regression Plots

Model and experimental details. The model is implemented using PyTorch [28] and PyTorch Geometric [29]. The regular-representation equivariant block was implemented using escnn library [11, 30]. Figure 2 shows the architectural overview of the message passing block. We used $T = 4$ rounds of message passing. After T steps, node embeddings consisting of the scalar and vector

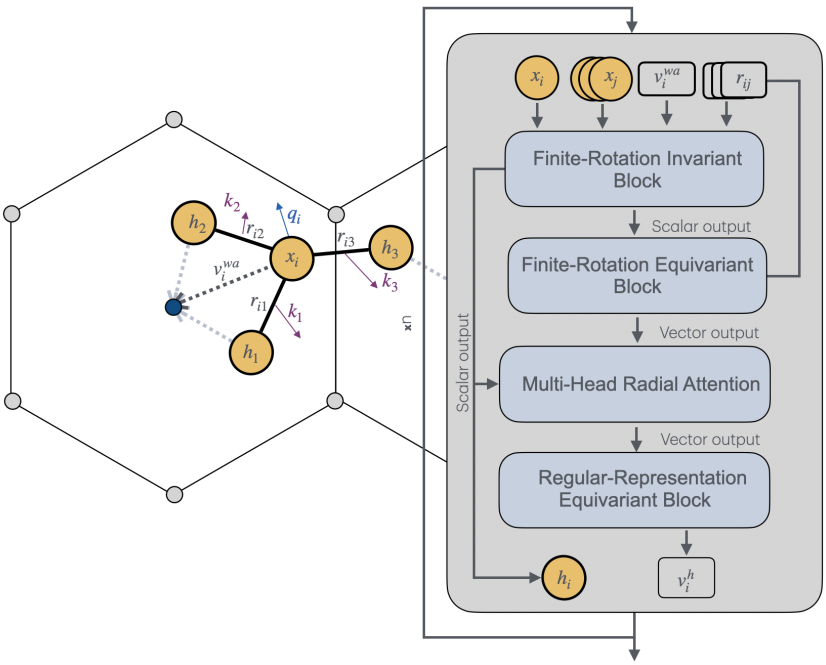

Figure 2: Architectural overview of the message passing block.

outputs were used as:

- Graph-level (binding energy, invariant): a permutation-invariant readout on the scalar features of the node followed by an MLP to predict a scalar.
- Node-level (force vector, equivariant): vectorial feature of the node.

The Wyckoff-Anchor vector encoding was computed using the ASE library, by getting the atomic coordinates modulo the unit cell vectors and finding the nearest Wyckoff position. A small numerical jitter of 10^{-6} Å was added to the atomic positions in case of a perfect overlap with the Wyckoff position or in case of two anchors being equidistant from the atom.

We used AdamW (batch size 64) with a learning-rate scheduler stepped once per epoch using the average training loss. The best model was checkpointed on the lowest validation loss. We used MAE as a per-node loss function for vector prediction and MSE for the adsorption energies.

The figures 3 and 4 show the regression plots for the binding energy and force prediction, respectively on the test set of the best run (lowest MAE).

Table 3: Ablation study on force vector prediction: MAE increased by 5-7% vs. baseline

Variant	Component	MAE Test / Train [eV / Å]	R^2 Test / Train
w/o Finite-rotation C_6 block	X-axis	0.057±0.001 / 0.043±0.001	0.86±0.01 / 0.95±0.10
	Y-axis	0.055±0.001 / 0.043±0.001	0.86±0.01 / 0.95±0.10
	Z-axis	0.039±0.002 / 0.027±0.002	0.76±0.20 / 0.93±0.10
w/o WA Vector Encoding	X-axis	0.059±0.001 / 0.042±0.001	0.84±0.01 / 0.92±0.10
	Y-axis	0.058±0.001 / 0.042±0.001	0.84±0.01 / 0.92±0.10
	Z-axis	0.042±0.002 / 0.026±0.002	0.73±0.20 / 0.90±0.10

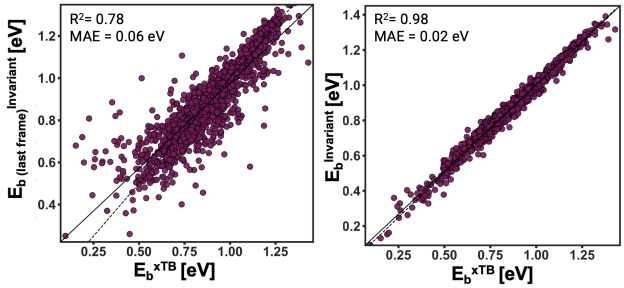

Figure 3: Binding energy regression plot on the test set. Left: from the relaxed position, Right: from the initial position.

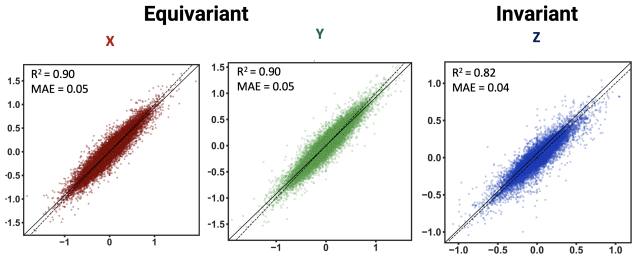

Figure 4: Force regression plot on the test set. (x, y and z parts)

Ablation study on force vector prediction. To showcase the importance of different components of the architecture, we performed an ablation study, where we removed different components (Finite-rotation C_6 block, Wyckoff-Anchor Vector Encoding) of the architecture and evaluated the performance on the force vector prediction from the initial position. The results are summarized in the Table 3. In both cases, we see a measurable drop in performance, by 5-7% in MAE. Interestingly, the WA-vector encoding seems to have a smaller effect on the performance, which is the only translational equivariant component of the architecture. This aspect requires further investigation. The C_6 block provides an additional benefit in the prediction accuracy, but the final geometrical rotational constrains comes from the D_6 regular-representation network, therefore excluding the C_6 block does not completely remove the rotational equivariance of the architecture.

Full comparison with E(n)-GNN. We compare the performance of our model with the E(n)-GNN architecture on the binding energy and force vector prediction tasks. The results are summarized in Table 4. Our model shows similar results to E(n)-GNN in both tasks, demonstrating the effectiveness of the proposed architecture. We note that E(n)-GNN is being applied to the full simulation cell of 300 atoms, while our model only sees the adsorbate atoms (on average 25 atoms). We kept the hyperparameters, the training protocol and the model size similar for both models to ensure a fair comparison, by removing the symmetry-constraining layers and increasing the width of the E-(n)GNN layers (keeping both networks at 300 K learnable parameters). The adsorption energy prediction performance is summarized in Table 5.

Table 4: E(n)-GNN force prediction performance

Component	MAE Test / Train [eV / Å]	R^2 Test / Train
X-axis	0.055±0.001 / 0.039±0.001	0.90±0.01 / 0.96±0.1
Y-axis	0.049±0.001 / 0.042±0.001	0.90±0.01 / 0.96±0.1
Z-axis	0.036±0.002 / 0.026±0.002	0.80±0.2 / 0.94±0.1

Table 5: E(n)-GNN binding energy prediction performance

Property	MAE Train / Test [eV]	R^2 Train / Test
Binding energy (fully relaxed)	0.019±0.001 / 0.008±0.003	0.988±0.002 / 0.988±0.001
Binding energy (initial position)	0.065±0.003 / 0.051±0.003	0.78±0.01 / 0.904±0.03

D Acknowledgments

We thank Lydia Elisa Koerber for the proofreading and feedback on the manuscript.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: We introduced a novel model for the Wallpaper group $p6mm$ and demonstrated its effectiveness in predicting material properties.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [No]

   Justification: Given it's a short workshop paper, and work in progress, we didn't have space to discuss limitations in detail.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: The paper provides a proof for its theoretical results or refers to all relevant literature.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [No]

   Justification: Since it's a work in progress, we do not share the dataset and code yet.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [No]

Justification: Since its a work in progress, we do not share the dataset and code yet.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [No]

   Justification: Since it's a work in progress, we do not share the dataset and code yet.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: In our very small experimental session, we provide standard deviation for all results.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).

- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [No]

   Justification: Since it's a work in progress, and we don't share the code or the dataset, we don't see the reason for providing the information about the computational resources.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research conducted in the paper conforms with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [No]

    Justification: The paper does not discuss any potential societal impacts, since it's a short paper workshop submission.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.

- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [No]

    Justification: We don't release any data or models that have a high risk for misuse.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [NA]

    Justification: We don't use any existing assets.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.

- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [No]

    Justification: Since its a work in progress, we do not share the dataset and code yet.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: The paper does not involve LLMs as any important, original, or non-standard components.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.