

A APPENDIX/SUPPLEMENTAL MATERIAL

A.1 USED SYMBOLS

Table 4: Explanation of the symbols used in the paper.

Used Symbols	Descriptions
$G_0 = \{\mathcal{V}_0, E_0\}$	Clean graph
$\mathcal{V}_t : t = 1 \dots T$	One-hot encoded node- i features
$E_t : t = 1 \dots T$	One-hot encoded edge features
$q(G_t \mid G_{t-1})$	Noise-driven forward diffusion process
$p_\phi(G_{t-1} \mid G_t)$	Reverse diffusion or denoising
n	Number of states
$\mathcal{C}_i : i = 1 \dots B$	i -th block containing nodes
$G_{<k}$	$\mathcal{C}_1 \bigcup \dots \bigcup \mathcal{C}_k$.

A.2 VARIATIONAL OBJECTIVE FOR STRUCTURED GRAPH DIFFUSION

To train the reverse denoising model $p_\theta(G_{t-1} | G_t)$ to approximate the true posterior of the forward process $q_\phi(G_{t-1} | G_t, G_0)$, we derive a variational lower bound (VLB) on the marginal likelihood $\log p_\theta(G_0)$. Starting from the evidence lower bound:

$$\log p_\theta(G_0) = \log \int q_\phi(G_{1:T} \mid G_0) \cdot \frac{p_\theta(G_0, G_{1:T})}{q_\phi(G_{1:T} \mid G_0)} dG_{1:T} \geq \mathbb{E}_{q_\phi(G_{1:T} \mid G_0)} \left[\log \frac{p_\theta(G_0, G_{1:T})}{q_\phi(G_{1:T} \mid G_0)} \right] \quad (3)$$

We decompose the joint distributions as:

$$p_\theta(G_0, G_{1:T}) = p_\theta(G_T) \prod_{t=2}^T p_\theta(G_{t-1} \mid G_t),$$

$$q_\phi(G_{1:T} \mid G_0) = \prod_{t=1}^T q_\phi(G_t \mid G_{t-1}, G_0) \quad (4)$$

Substituting into the ELBO:

$$\log p_\theta(G_0) \geq \mathbb{E}_{q_\phi} \left[\log p_\theta(G_T) + \sum_{t=2}^T \log p_\theta(G_{t-1} \mid G_t) - \sum_{t=1}^T \log q_\phi(G_t \mid G_{t-1}, G_0) \right] \quad (5)$$

This can be rearranged as:

$$\log p_\theta(G_0) \geq \mathbb{E}_{q_\phi} \left[\log p_\theta(G_T) - \log q_\phi(G_1 \mid G_0) + \sum_{t=2}^T \log \frac{p_\theta(G_{t-1} \mid G_t)}{q_\phi(G_t \mid G_{t-1}, G_0)} \right] \quad (6)$$

We reorganize the objective into reconstruction and KI terms:

$$\log p_\theta(G_0) \geq \mathbb{E}_{q_\phi} [\log p_\theta(G_0 \mid G_1)] - \sum_{t=2}^T \mathbb{E}_{q_\phi} [D_{\text{KL}} (q_\phi(G_{t-1} \mid G_t, G_0) \parallel p_\theta(G_{t-1} \mid G_t))] \quad \text{const.} \quad (7)$$

We define the total training objective as:

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{rec}}(\theta) + \sum_{t=2}^T \mathcal{L}_t(\theta) \quad (8)$$

702 where:

703

$$\mathcal{L}_{\text{rec}}(\theta) = -\mathbb{E}_{q_\phi} [\log p_\theta(G_0 \mid G_1)] \quad (9)$$

704

$$\mathcal{L}_t(\theta) = \mathbb{E}_{q_\phi} [D_{\text{KL}}(q_\phi(G_{t-1} \mid G_t, G_0) \parallel p_\theta(G_{t-1} \mid G_t))] \quad (10)$$

705 This variational bound enables efficient training via a hybrid loss that balances data likelihood with
706 forward-reverse consistency across diffusion steps.

707

A.3 PARAMETERIZING FORWARD AND REVERSE TRANSITIONS IN DISCRETE GRAPH 711 DIFFUSION

712 We define the forward and reverse diffusion processes over graphs using a simplified discrete-time
713 formulation, following Zhao et al. Zhao et al. (2024). Our framework focuses on three key distributions:
714 (i) the forward marginal $q(G_t \mid G_0)$, (ii) the backward posterior $q(G_{t-1} \mid G_t, G_0)$, and (iii)
715 the learned reverse process $p_\phi(G_{t-1} \mid G_t)$. This design prioritizes memory efficiency and avoids
716 the complexity introduced by approximations such as those in D3PM Austin et al. (2021).

717 Since the forward process applies noise independently to all nodes and edges (as shown in Eq. (1)),
718 we can model these three distributions by factorizing over individual elements. Let $x \in \mathcal{V}_t \cup \mathcal{E}_t$ be a
719 discrete random variable with one-hot encoding and categorical distribution: $x \sim \text{Cat}(x; \mathbf{p})$, where
720 $\mathbf{p} \in [0, 1]^n$ and $\mathbf{1}^\top \mathbf{p} = 1$. Then the probability of observing a one-hot state x under distribution \mathbf{p}
721 is $x^\top \mathbf{p}$. The corruption step $q(x_t \mid x_{t-1})$ can be expressed using a transition matrix $Q_t \in [0, 1]^{n \times n}$
722 as:

723

$$q(x_t \mid x_{t-1}) = \text{Cat}(x_t; Q_t^\top x_{t-1}) \quad (11)$$

724 Let the composed transition matrix be $\bar{Q}_t = Q_1 Q_2 \cdots Q_t$. Then, the forward marginal becomes:

725

$$q(x_t \mid x_0) = \text{Cat}(x_t; \bar{Q}_t^\top x_0) \quad (12)$$

726 The backward posterior is:

727

$$q(x_{t-1} \mid x_t, x_0) = \text{Cat}\left(x_{t-1}; \frac{Q_t x_t \odot \bar{Q}_{t-1}^\top x_0}{x_t^\top \bar{Q}_t^\top x_0}\right) \quad (13)$$

728 This applies to both node states $x = v_i^t \in \mathcal{V}_t$ and edge types $x = e_{i,j}^t \in \mathcal{E}_t$, with the same
729 formulation. We optionally use shared transition matrices Q_t^V and Q_t^E for all nodes and edges,
730 respectively. See the next section for derivation. To define a uniform and information-less terminal
731 distribution $q(G_T \mid G_0)$, we choose:

732

$$Q_t = \alpha_t I + (1 - \alpha_t) \mathbf{1} \mathbf{m}^\top \quad (14)$$

733 where $\alpha_t \in [0, 1]$ is a time-dependent noise schedule, and $\mathbf{m} \in [0, 1]^n$ is the uniform categorical
734 distribution over n states, such that $m_i = \frac{1}{n}$. For reverse modeling, we use:

735

$$p_\phi(x_{t-1} \mid G_t) = \sum_{x_0} q(x_{t-1} \mid x_t, x_0) \cdot p_\phi(x_0 \mid G_t) \quad (15)$$

736 This formulation enables us to parameterize $p_\phi(x_0 \mid G_t)$ with a neural network and compute
737 $p_\phi(x_{t-1} \mid G_t)$ via marginalization over the clean state space using Eq. (3).

738

A.4 DERIVATION OF $q(x_{t-1} \mid x_t, x_0)$

739 We begin by defining the composite transition matrix over steps s through t as $\bar{Q}_{t|s} =$
740 $Q_s Q_{s+1} \cdots Q_t$. For brevity, we denote $\bar{Q}_t = \bar{Q}_{t|0}$ and $\bar{Q}_{t-1} = \bar{Q}_{t-1|0}$. Our goal is to compute
741 the posterior distribution $q(x_{t-1} \mid x_t, x_0)$, assuming the forward process has the Markov structure
742 given by:

743

$$q(x_t \mid x_{t-1}) = \text{Cat}(x_t; Q_t^\top x_{t-1}) \quad (16)$$

756 From the chain rule of probability:
 757

$$758 q(x_{t-1} | x_t, x_0) = \frac{q(x_t | x_{t-1})q(x_{t-1} | x_0)}{q(x_t | x_0)} \quad (17)$$

760 We expand each term using the forward marginals:
 761

$$762 q(x_{t-1} | x_0) = \text{Cat}(x_{t-1}; \bar{Q}_{t-1}^\top x_0) \quad (18)$$

$$763 q(x_t | x_0) = \text{Cat}(x_t; \bar{Q}_t^\top x_0) \quad (19)$$

765 Thus, the numerator becomes:
 766

$$767 q(x_t | x_{t-1})q(x_{t-1} | x_0) = (x_t^\top Q_t^\top x_{t-1}) \cdot (x_{t-1}^\top \bar{Q}_{t-1}^\top x_0)$$

768 We now marginalize over all possible x_{t-1} to normalize:
 769

$$770 q(x_{t-1} | x_t, x_0) = \frac{Q_t x_t \odot \bar{Q}_{t-1}^\top x_0}{x_t^\top \bar{Q}_t^\top x_0} \quad (20)$$

773 Hence, the posterior is a categorical distribution over x_{t-1} :

$$774 q(x_{t-1} | x_t, x_0) = \text{Cat}\left(x_{t-1}; \frac{Q_t x_t \odot \bar{Q}_{t-1}^\top x_0}{x_t^\top \bar{Q}_t^\top x_0}\right) \quad (20)$$

778 A.5 PROOF OF PERMUTATION INVARIANCE IN BLOCKWISE GRAPH GENERATION

779 To establish that $p_\theta(G)$ is an exchangeable probability distribution over graphs, we aim to prove that
 780 for any permutation matrix P , the model satisfies
 781

$$782 p_\theta(P \star G) = p_\theta(G), \quad (21)$$

783 where $P \star G$ denotes the graph obtained by permuting both node indices and corresponding edge
 784 entries in G .

785 Our generative model factorizes the likelihood of a graph G based on block-wise decomposition
 786 induced by a structural ranking function ψ . Let $\mathcal{B}_1, \dots, \mathcal{B}_{K_B}$ be the node subsets (blocks) ranked
 787 by ψ . The generation is performed sequentially over these blocks:

$$788 p_\theta(G) = \prod_{i=1}^{K_B} p_\theta(G[\mathcal{B}_i] | G[\mathcal{B}_{1:i-1}], G[\mathcal{B}_{1:i-1}] \setminus G[\mathcal{B}_i]). \quad (22)$$

792 **Permutation Equivariance of the Indexing.** Each block \mathcal{B}_i is determined from G using $\psi(G)$,
 793 which is permutation-consistent (Theorem 1). Thus, for any permutation matrix P , we have

$$794 \mathcal{B}_i(P \star G) = P \star \mathcal{B}_i(G). \quad (23)$$

795 Furthermore, indexing operations on graphs are equivariant:

$$796 P \star G[\mathcal{B}_i] = G[P \star \mathcal{B}_i]. \quad (24)$$

798 **Exchangeability of Block Generation.** Consider:

$$799 p_\theta(P \star G) = \prod_{t=1}^{K_B} p_\theta(P \star G[\mathcal{B}_t] | P \star G[\mathcal{B}_{1:t-1}], \\ 800 \quad P \star (G[\mathcal{B}_{1:t-1}] \setminus G[\mathcal{B}_t])) \quad (25)$$

804 Since our model is constructed to be equivariant with respect to permutations, each conditional
 805 satisfies:

$$806 p_\theta(P \star X | P \star Y) = p_\theta(X | Y), \quad (26)$$

807 for arbitrary subgraphs X, Y . Applying this recursively yields:

$$808 p_\theta(P \star G) = \prod_{i=1}^{K_B} p_\theta(G[\mathcal{B}_i] | G[\mathcal{B}_{1:i-1}], G[\mathcal{B}_{1:i-1}] \setminus G[\mathcal{B}_i]) = p_\theta(G). \quad (27)$$

810 **Marginalization of Conditioning Sets.** For further rigor, define the conditional term
 811

$$812 \quad p_\theta(G[\mathcal{B}_i] \mid G[\mathcal{B}_{1:i-1}] \setminus G[\mathcal{B}_i], G[\mathcal{B}_{1:i-1}]) . \quad (28)$$

813 Let $\mathcal{H}_{\mathcal{B}_{1:i-1}}$ denote all other nodes outside $\mathcal{B}_{1:i}$. Then,
 814

$$815 \quad p_\theta(G[\mathcal{B}_i] \mid G[\mathcal{B}_{1:i-1}], G[\mathcal{B}_{1:i-1}] \setminus G[\mathcal{B}_i]) \\ 816 \quad = \int p_\theta(G[\mathcal{B}_i] \mid G[\mathcal{B}_{1:i-1}], \mathcal{H}_{\mathcal{B}_{1:i-1}}) \cdot p(\mathcal{H}_{\mathcal{B}_{1:i-1}}) d\mathcal{H}_{\mathcal{B}_{1:i-1}}. \quad (29)$$

819 Using the fact that the generative model's forward noise and reverse denoising chains are designed
 820 to be permutation equivariant, we have:
 821

$$822 \quad p_\theta(G) = \int p(H_T \mid G) \prod_{t=1}^T p(H_{t-1} \mid H_t) dH_{1:T}. \quad (30)$$

825 Then for any P :

$$827 \quad p_\theta(P \star G) = \int p(H_T \mid P \star G) \prod_{t=1}^T p(H_{t-1} \mid H_t) dH_{1:T}, \quad (31)$$

$$830 \quad = \int p(P \star H_T \mid G) \prod_{t=1}^T p(P \star H_{t-1} \mid P \star H_t) dH_{1:T}, \quad (32)$$

$$833 \quad = \int p(H_T \mid G) \prod_{t=1}^T p(H_{t-1} \mid H_t) dH_{1:T} = p_\theta(G). \quad (33)$$

836 This confirms that $p_\theta(G)$ is invariant under any node permutation P , establishing exchangeability.
 837

839 A.6 UNIFIED TRAINING AND GENERATION FOR BLOCK-WISE STRUCTURED GRAPH 840 DIFFUSION

842 **Algorithm 5** Unified Training and Generation Procedure for Block-wise Structured Graph Diffusion

843 **Require:** Graph G , max diffusion steps T , max hop K_h , block size predictor g_θ , denoising model ℓ_α

844 1: Obtain node ordering ψ from ordering network ϕ (Algorithm 1)

845 2: Partition G into ranked blocks $[\mathcal{C}_1, \dots, \mathcal{C}_{K_B}]$ using ψ

846 3: **for** $i = 1$ to K_B **do**

847 4: $\widehat{\mathcal{C}}_i \leftarrow g_\theta(G_{\leq i-1})$

848 5: $\mathcal{M} \leftarrow \text{mask}(G[\mathcal{C}_{1:i}] \setminus G[\mathcal{C}_{1:i-1}])$

849 6: Sample $t \sim \mathcal{U}(1, T)$

850 7: $\widetilde{G}[\mathcal{C}_i] \leftarrow \mathcal{M} \odot q_t(G[\mathcal{C}_i]) + (1 - \mathcal{M}) \odot G[\mathcal{C}_i]$

851 8: $X \leftarrow f_\theta(\widetilde{G}[\mathcal{C}_i])$

852 9: Compute ℓ_i^{KL} and ℓ_i^{CE} using Eq. (2), Eq. (3)

853 10: **end for**

854 11: Minimize total loss: $\sum_{i=1}^{K_B} \ell_i$

855 12: $G \leftarrow \emptyset$, $i \leftarrow 1$

856 13: Sample $n \sim g_\theta(G)$

857 14: **while** $n > 0$ **do**

858 15: Add block \mathcal{C}_i with n nodes to G

859 16: $\mathcal{M} \leftarrow \text{mask}(G[\mathcal{C}_i] \setminus G[\mathcal{C}_{1:i-1}])$

860 17: $\widetilde{G} \leftarrow \text{Noise}(\mathcal{M})$

861 18: **for** $j = 1$ to T **do**

862 19: $\mathbf{p} \leftarrow f_\theta(\widetilde{G})$

863 20: Sample S from \mathbf{p}

864 21: $\widetilde{G} \leftarrow \mathcal{M} \odot S + (1 - \mathcal{M}) \odot \widetilde{G}$

865 22: **end for**

866 23: $G \leftarrow \widetilde{G}$

867 24: Sample $n \sim g_\theta(G)$

868 25: $i \leftarrow i + 1$

869 26: **end while**

870 27: **return** G

864 A.7 PROOF OF THEOREM 1
865

866 Consider a sequence of transition matrices $\{\mathcal{T}_1, \dots, \mathcal{T}_T\}$, each representing a categorical diffusion
867 step. The matrices should be constructed such that, at long time horizons ($t \rightarrow T$), the resulting dis-
868 tribution converges to a known steady-state distribution $\mu \in D^K$, where D^K is the K -dimensional
869 probability simplex. We define this limiting behavior as:
870

$$872 \lim_{t \rightarrow T} \mathcal{T}_t = \mathbf{1}\mu^\top \quad (34)$$

874 This ensures that every row of the composed matrix approaches μ , making the distribution station-
875 ary. To enforce this convergence in a controllable way, we propose defining each transition matrix
876 \mathcal{T}_t as a convex blend between the identity matrix and the rank-1 matrix $\mathbf{1}\mu^\top$:

$$877 \mathcal{T}_t = \gamma_t \cdot \mathbf{I} + (1 - \gamma_t) \cdot \mathbf{1}\mu^\top, \quad \gamma_t \in [0, 1] \quad (35)$$

879 The accumulated transition from time step s to t , denoted as $\mathcal{T}_{t|s}$, can be recursively written as
880 follows:
881

$$882 \mathcal{T}_{t|s} = \gamma_{t|s} \cdot \mathbf{I} + (1 - \gamma_{t|s}) \cdot \mathbf{1}\mu^\top, \quad (36)$$

883 where the effective decay factor $\gamma_{t|s}$ is the product of all decay terms from step $s + 1$ to t :

$$884 \gamma_{t|s} = \prod_{r=s+1}^t \gamma_r \quad (37)$$

885 This implies:
886

$$887 \gamma_t = \gamma_{t|0} = \gamma_{t|s} \cdot \gamma_s \quad (38)$$

888 With this formulation, we ensure that as $t \rightarrow T$, the accumulated matrix $\mathcal{T}_{t|0}$ becomes fully rank-1,
889 and the variable distribution becomes indistinguishable from the stationary prior μ . This gives the
890 reparameterized posterior for timestep $t - 1$, used in computing the variational loss. We present
891 the argument for node representations; the same reasoning holds for structurally symmetric edges.
892 Suppose that two nodes u and v in a graph G are structurally indistinguishable. Then, there exists a
893 graph automorphism $\pi \in \text{Aut}(G)$ such that:
894

$$895 \pi(u) = v \quad (39)$$

900 Let \mathcal{P}_n denote the set of all node permutation matrices of size $n \times n$. Assume we have a neural
901 function $\psi : G \mapsto \mathbb{R}^{n \times d}$ that is permutation-equivariant, i.e., for any permutation matrix $\pi \in \mathcal{C}_n$,
902 we have:
903

$$904 \psi(\pi \star G) = \pi \star \psi(G) \quad (40)$$

905 Now, apply π as the permutation on nodes. Because π is an automorphism of G , it preserves the
906 graph structure, so $\pi \star G = G$. Thus:
907

$$908 \psi(G) = \psi(\pi \star G) = \pi \star \psi(G) \quad (41)$$

909 This implies:
910

$$911 \psi(G)_u = \psi(G)_v \quad (42)$$

912 In other words, nodes u and v , being symmetric under graph automorphism π , are mapped to iden-
913 tical representations by the function ψ .
914

915 A.8 PROOF OF THEOREM 2
916

917 Let $\sigma \in \text{Auto}(G)$. By definition, $\sigma \star G = G$ (the attributed graph is unchanged by σ). By permutation
918 equivariance of Φ ,

$$919 \Phi(G) = \Phi(\sigma \star G) = \sigma \star \Phi(G). \quad (43)$$

918 Unpacking the rightmost equality component-wise over nodes gives, for every $w \in \mathcal{V}$,
 919

$$920 \quad \Phi(G)_w = (\sigma \star \phi(G))_w = \phi(G)_{\sigma^{-1}(w)}. \quad (44)$$

922 Equivalently, for every w , $\Phi(G)_{\sigma(w)} = \Phi(G)_w$. Now fix any two nodes $u, v \in \mathcal{V}$ in the same
 923 $\text{Aut}(G)$ -orbit. By definition, there exists $\sigma \in \text{Aut}(G)$ with $\sigma(u) = v$. Applying the relation above
 924 with $w = u$ yields
 925

$$926 \quad \Phi(G)_v = \phi(G)_{\sigma(u)} = \Phi(G)_u, \quad (45)$$

928 establishing $\Phi(G)_u = \Phi(G)_v$. The argument for edge embeddings is identical: let $\Phi^{(e)}$ map G
 929 to edge-wise outputs indexed by ordered (or unordered) pairs. Equivariance acts on pairs via $\pi \star$
 930 $(i, j) = (\pi(i), \pi(j))$. For any automorphism σ , $\Phi^{(e)}(G) = \sigma \star \Phi^{(e)}(G)$, hence $\Phi^{(e)}(G)_{(i,j)} =$
 931 $\Phi^{(e)}(G)_{(\sigma(i), \sigma(j))}$. If (u, v) and (u', v') are in the same orbit, choose σ with $\sigma(u) = u'$, $\sigma(v) = v'$
 932 to conclude equality of their edge embeddings. Since the derivation uses only (1) $\sigma \star G = G$ for
 933 $\sigma \in \text{Aut}(G)$ and (2) equivariance of Φ , the result is independent of depth/width/expressivity.
 934

Implication. (1) On features. The statement assumes automorphisms preserve all attributes used by
 935 Φ . If node/edge features break symmetry (e.g., unique IDs), then $\text{Aut}(G)$ shrinks accordingly; the
 936 conclusion applies with respect to that reduced group. (2) Symmetry cannot be broken internally.
 937 The proof formalizes the impossibility of distinguishing nodes within an automorphism orbit by
 938 any permutation-equivariant architecture alone. To separate orbit-mates, one must inject symmetry-
 939 breaking signals (positional encodings, random IDs, anchors, or global tie-breakers); and (3) Group-
 940 theoretic view. The equality $\Phi(G) = \sigma \star \phi(G) \forall \sigma \in \text{Aut}(G)$ means $\Phi(G)$ lies in the fixed-point
 941 subspace of the representation of $\text{Aut}(G)$. Constancy on orbits is exactly the characterization of
 942 such fixed points by Burnside’s lemma/orbit-stabilizer intuition.
 943

944 A.9 PROOF OF THEOREM 3

946 To demonstrate that the learned probability distribution $\mathbb{P}_\phi(G)$ over graphs is exchangeable, we
 947 must verify that for any node permutation matrix $\pi \in \mathcal{C}_n$, the group of node permutations, it holds
 948 that:
 949

$$950 \quad 951 \quad \mathbb{P}_\phi(\pi \star G) = \mathbb{P}_\phi(G) \quad (46)$$

952 Here, $\pi \star G$ denotes the permuted graph, where nodes and their relations (or, edges) are permuted
 953 accordingly: $\pi \star G = (\pi \cdot V, \pi \cdot E \cdot \pi^\top)$. Assume that the generation model produces a graph via a
 954 sequential composition of subgraphs defined by structural neighborhoods or partitions, such that:
 955

$$956 \quad 957 \quad 958 \quad \mathbb{P}_\phi(G) = \prod_{i=1}^K \mathbb{P}_\phi(G_{\leq i} \setminus G_{\leq i-1} \mid G_{\leq i-1}) \quad (47)$$

961 Here, $G_{\leq i}$ denotes the union of the first i block $\mathcal{C}_1 \cdots \mathcal{C}_i$ (e.g., neighborhoods) induced by a binary
 962 mask over nodes. This indexing operation is permutation-equivariant:
 963

$$964 \quad \pi \star (G_{\leq i}) = (\pi \star G)_{\leq i} \quad (48)$$

966 Additionally, suppose that each subset $(\pi \star G)_{\leq i}$ is selected via a deterministic function of the graph
 967 structure (e.g., via neighborhood expansion or hop-based grouping), which is also equivariant under
 968 permutation. Then:
 969

$$970 \quad 971 \quad (\pi \star G)_{\leq i} = \pi \star G_{\leq i} \quad (49)$$

972 Now evaluate the generative model on the permuted graph:
 973

$$\begin{aligned}
 974 \quad \mathbb{P}_\phi(\pi \star G) &= \prod_{i=1}^K \mathbb{P}_\phi((\pi \star G)_{\leq i} \setminus (\pi \star G)_{\leq i-1} \mid (\pi \star G)_{\leq i-1}) \\
 975 \quad &= \prod_{i=1}^K \mathbb{P}_\phi(\pi \star (G_{\leq i} \setminus G_{\leq i-1}) \mid \pi \star G_{\leq i-1}) \\
 976 \quad &= \prod_{i=1}^K \mathbb{P}_\phi(\pi \star \Delta_i \mid \pi \cdot \mathcal{G}_{< i}),
 \end{aligned} \tag{50}$$

983 where $\Delta_i = G_{\leq k} \setminus G_{\leq i-1}$, and $G_{< i} = \mathcal{C}_1 \cup \dots \cup \mathcal{C}_{i-1}$. If the conditional probabilities \mathbb{P}_ϕ are
 984 defined through permutation-invariant functions (e.g., based on multi-set or degree statistics), then
 985 we have:

$$988 \quad \mathbb{P}_\phi(\pi \star \Delta_i \mid \pi \star G_{\leq i}) = \mathbb{P}_\phi(\Delta_i \mid G_{\leq i}) \tag{51}$$

989 Thus,

$$991 \quad \mathbb{P}_\phi(\pi \star G) = \mathbb{P}_\phi(G), \tag{52}$$

992 which confirms the probability distribution modeled by \mathbb{P}_ϕ is invariant under node permutations,
 993 i.e., it is exchangeable.

995 A.10 DERIVING A BLOCK-CAUSAL MATRIX PRODUCT

997 Let $X \in \mathbb{R}^{n \times d}$ and $Y \in \mathbb{R}^{d \times m}$ be two matrices. Define the standard matrix multiplication entry as:
 998

$$999 \quad [XY]_{ij} = \langle \mathbf{x}_i, \mathbf{y}_j \rangle, \tag{53}$$

1000 where \mathbf{x}_i denotes the i -th row of X , and \mathbf{y}_j is the j -th column of Y . Now, in a block-wise AR
 1001 setting, we introduce a function $b : \{1 \dots n\} \mapsto \mathbb{N}$ assigning a block index to each row/column.
 1002 The matrix entry (i, j) should depend only on features from block indices $\leq \max(b(i), b(j))$. To
 1003 ensure this, define a binary mask matrix $\mathcal{M} \in \{0, 1\}^{n \times d}$, where:
 1004

$$1005 \quad \mathcal{M}_{ik} = \begin{cases} 1 & \text{if } b(i) \geq b(k) \\ 0 & \text{otherwise} \end{cases} \tag{54}$$

1006 For a safe computation of the entry \mathcal{Z}_{ij} under this constraint, we define:
 1007

$$1010 \quad \mathcal{Z}_{ij} = \langle \mathbf{x}_i \odot (\boldsymbol{\mu}_i \vee \boldsymbol{\mu}_j), \mathbf{y}_j \rangle \tag{55}$$

1011 Here, $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ are binary indicator vectors selecting valid components, and \odot denotes the
 1012 Hadamard (element-wise) product, while \vee is the element-wise logical OR. We can expand this
 1013 expression as:
 1014

$$1015 \quad \mathcal{Z}_{ij} = \langle \mathbf{x}_i \odot \boldsymbol{\mu}_i, \mathbf{y}_j \rangle + \langle \mathbf{x}_i, \mathbf{y}_j \odot \boldsymbol{\mu}_j \rangle - \langle \mathbf{x}_i \odot \boldsymbol{\mu}_i, \mathbf{y}_j \odot \boldsymbol{\mu}_j \rangle \tag{56}$$

1016 In matrix form, letting \mathbf{Z} be the final output:
 1017

$$1018 \quad \mathbf{Z} = (\mathbf{X} \odot \mathcal{M}) \mathbf{Y} + \mathbf{X} (\mathbf{Y} \odot \mathcal{M}^\top) - (\mathbf{X} \odot \mathcal{M}) (\mathbf{Y} \odot \mathcal{M}^\top) \tag{57}$$

1020 This formulation ensures that information flows only within valid block boundaries, enabling parallel-
 1021 elizable yet causally consistent matrix computation.

1023 A.11 STRUCTURED GRID GRAPHS

1024 This section presents a few more structured artificial grid generated using the proposed PARDIFF
 1025 algorithm:

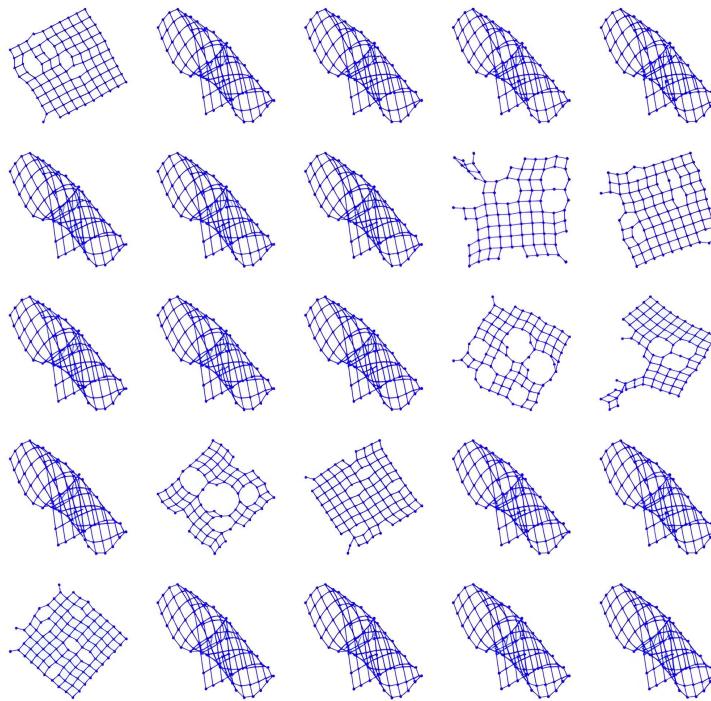


Figure 2: Non-curated structured grid graphs generated by our method, trained with 50 diffusion steps per block. The samples display mostly regular grid-like topology with occasional geometric perturbations, demonstrating the model’s ability to capture both structure and variation without any filtering.

A.11.1 AUTOREGRESSION IN DIFFUSION MODELS: ESSENTIAL OR EXCESS?

Although denoising diffusion models are naturally permutation-invariant, we examine whether incorporating an AR structure offers tangible benefits. Specifically, we investigate whether decomposing the graph generation process into block-wise conditional distributions—based on a structural partial order—can lead to improved quality, efficiency, and stability. To this end, we perform an ablation study by varying the hop radius K_h , which defines the granularity of autoregressive blocks. When $K_h = 0$, the graph is treated as a single undivided structure—this corresponds to pure diffusion without any AR decomposition. Larger values of K_h yield finer block-wise partitions, introducing more AR steps. We also evaluate a variant where diffusion is performed without AR but with a larger number of denoising steps, to control for potential improvements from increased sampling. Across all settings, we report molecule validity, uniqueness, atomic and molecular stability, and the FCD. The results, summarized in Table 5, indicate that autoregressive diffusion significantly enhances generation quality. Notably, PARDiff with $K_h = 3$ achieves the best performance with fewer total diffusion steps compared to non-AR setups. This confirms that AR decomposition provides stronger inductive bias, improved stability, and more efficient training—even in permutation-invariant settings.

Table 5: Ablation on QM9 under different autoregressive granularities K_h . More blocks (higher K_h) improve performance.

K_h	Steps	Blks	Size	Val.	Uni.	Mol-Stab	Atm-Stab	FCD
0	140	1	23.4	93.1	95.7	76.2	97.5	2.15
0	280	1	23.4	94.0	96.2	78.1	97.8	1.84
0	490	1	23.4	94.8	96.6	78.3	98.0	1.69
1	140	4.1	5.7	97.3	96.7	86.8	98.4	1.21
2	140	6.2	3.8	97.5	96.5	87.0	98.6	1.13
3	140	8.0	3.1	97.8	96.9	88.2	98.9	0.96

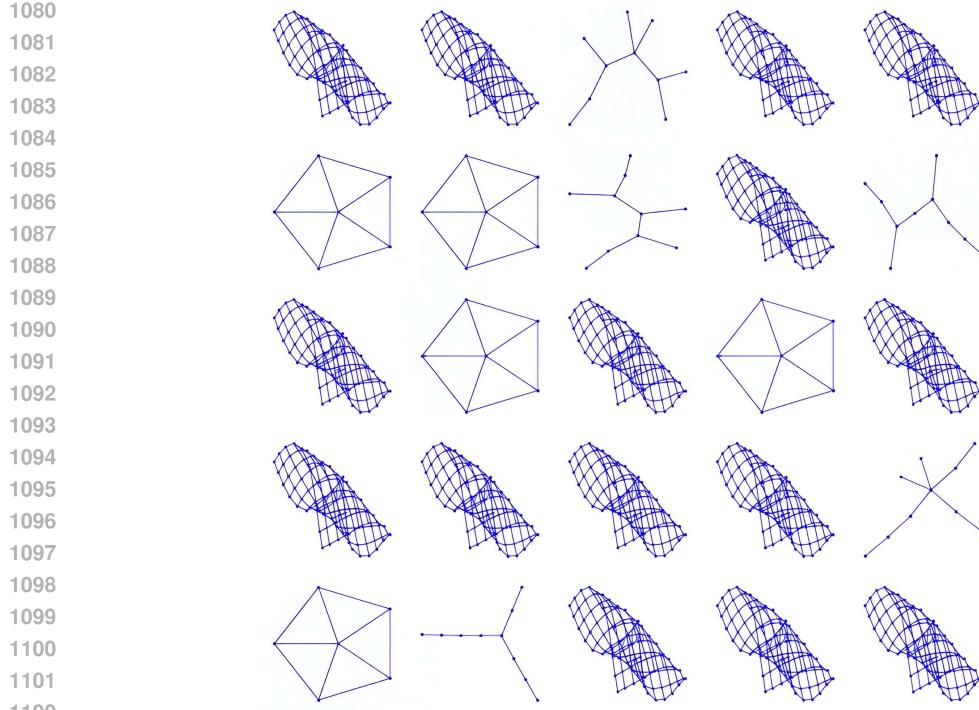


Figure 3: Unfiltered grid-like graphs generated by the eigenvector-enhanced model trained with 50 steps per block.

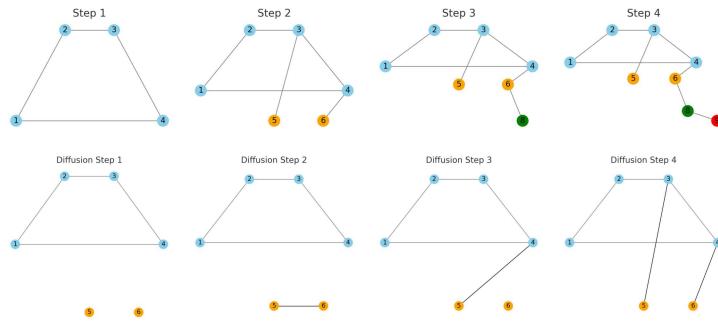


Figure 4: Comparison of autoregressive and diffusion-based graph generation. The top row illustrates autoregressive generation, where nodes and edges are sequentially added in each step. The bottom row shows diffusion-based generation, where the graph is iteratively refined from a noisy initialization toward the target structure.

A.12 ABLATION STUDY

Table 5 investigates the effect of varying autoregressive granularity, controlled by the number of hierarchical blocks K_h , on generation quality in the QM9 dataset. When $K_h = 0$, the model generates the entire graph in a single step, yielding lower performance across all metrics. Increasing the number of diffusion steps improves results incrementally (e.g., FCD drops from 2.15 to 1.69 as steps increase from 140 to 490), but this comes at the cost of significantly higher computational burden, with no structural decomposition. In contrast, introducing even a moderate level of autoregressive structure ($K_h = 1$) immediately boosts performance across all axes—validity, stability, and FCD—indicating that decomposing the graph into substructures introduces useful inductive bias that guides generation more effectively.

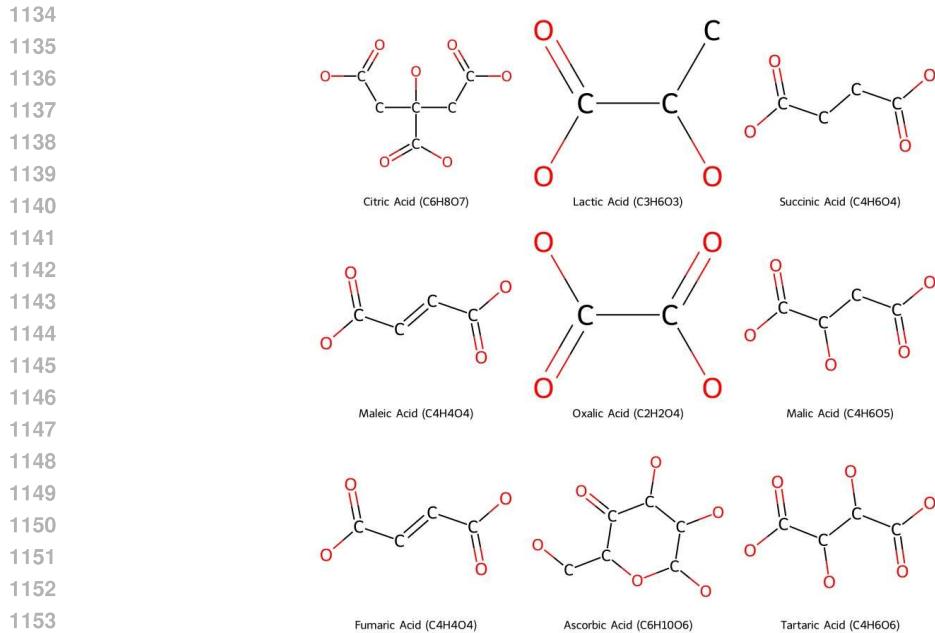


Figure 5: Sample complex molecular structures are generated using PARDIFF.

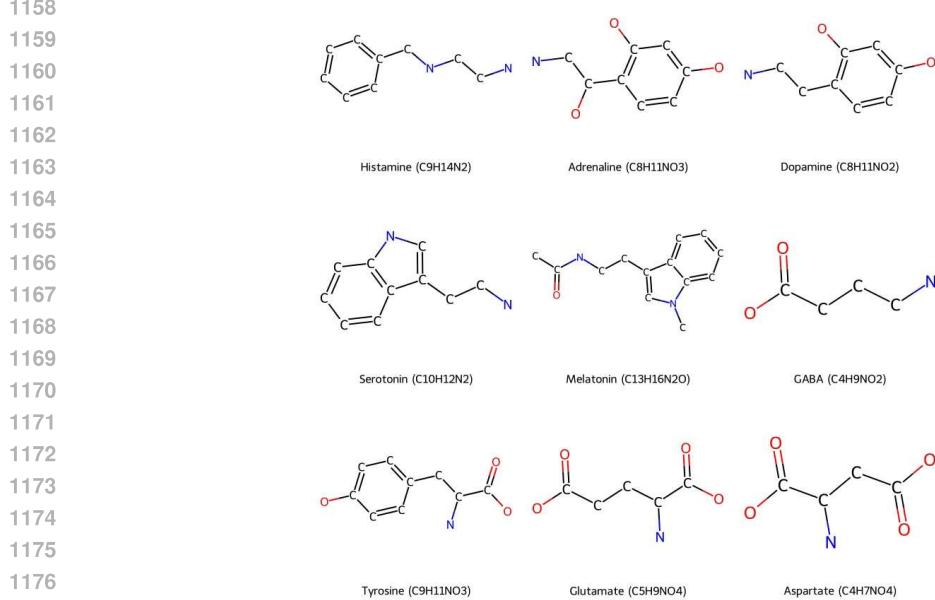


Figure 6: Sample complex molecular structures are generated using PARDIFF.

1181
 1182
 1183
 1184
 1185
 1186
 1187
 1188
 1189
 1190
 1191
 1192
 1193
 1194
 1195
 1196
 1197
 1198
 1199
 1200
 1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218
 1219
 1220
 1221
 1222
 1223
 1224
 1225
 1226
 1227
 1228
 1229
 1230
 1231
 1232
 1233
 1234
 1235
 1236
 1237
 1238
 1239
 1240
 1241
 1242
 1243
 1244
 1245
 1246
 1247
 1248
 1249
 1250
 1251
 1252
 1253
 1254
 1255
 1256
 1257
 1258
 1259
 1260
 1261
 1262
 1263
 1264
 1265
 1266
 1267
 1268
 1269
 1270
 1271
 1272
 1273
 1274
 1275
 1276
 1277
 1278
 1279
 1280
 1281
 1282
 1283
 1284
 1285
 1286
 1287
 1288
 1289
 1290
 1291
 1292
 1293
 1294
 1295
 1296
 1297
 1298
 1299
 1300
 1301
 1302
 1303
 1304
 1305
 1306
 1307
 1308
 1309
 1310
 1311
 1312
 1313
 1314
 1315
 1316
 1317
 1318
 1319
 1320
 1321
 1322
 1323
 1324
 1325
 1326
 1327
 1328
 1329
 1330
 1331
 1332
 1333
 1334
 1335
 1336
 1337
 1338
 1339
 1340
 1341
 1342
 1343
 1344
 1345
 1346
 1347
 1348
 1349
 1350
 1351
 1352
 1353
 1354
 1355
 1356
 1357
 1358
 1359
 1360
 1361
 1362
 1363
 1364
 1365
 1366
 1367
 1368
 1369
 1370
 1371
 1372
 1373
 1374
 1375
 1376
 1377
 1378
 1379
 1380
 1381
 1382
 1383
 1384
 1385
 1386
 1387
 1388
 1389
 1390
 1391
 1392
 1393
 1394
 1395
 1396
 1397
 1398
 1399
 1400
 1401
 1402
 1403
 1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426
 1427
 1428
 1429
 1430
 1431
 1432
 1433
 1434
 1435
 1436
 1437
 1438
 1439
 1440
 1441
 1442
 1443
 1444
 1445
 1446
 1447
 1448
 1449
 1450
 1451
 1452
 1453
 1454
 1455
 1456
 1457
 1458
 1459
 1460
 1461
 1462
 1463
 1464
 1465
 1466
 1467
 1468
 1469
 1470
 1471
 1472
 1473
 1474
 1475
 1476
 1477
 1478
 1479
 1480
 1481
 1482
 1483
 1484
 1485
 1486
 1487
 1488
 1489
 1490
 1491
 1492
 1493
 1494
 1495
 1496
 1497
 1498
 1499
 1500
 1501
 1502
 1503
 1504
 1505
 1506
 1507
 1508
 1509
 1510
 1511
 1512
 1513
 1514
 1515
 1516
 1517
 1518
 1519
 1520
 1521
 1522
 1523
 1524
 1525
 1526
 1527
 1528
 1529
 1530
 1531
 1532
 1533
 1534
 1535
 1536
 1537
 1538
 1539
 1540
 1541
 1542
 1543
 1544
 1545
 1546
 1547
 1548
 1549
 1550
 1551
 1552
 1553
 1554
 1555
 1556
 1557
 1558
 1559
 1560
 1561
 1562
 1563
 1564
 1565
 1566
 1567
 1568
 1569
 1570
 1571
 1572
 1573
 1574
 1575
 1576
 1577
 1578
 1579
 1580
 1581
 1582
 1583
 1584
 1585
 1586
 1587
 1588
 1589
 1590
 1591
 1592
 1593
 1594
 1595
 1596
 1597
 1598
 1599
 1600
 1601
 1602
 1603
 1604
 1605
 1606
 1607
 1608
 1609
 1610
 1611
 1612
 1613
 1614
 1615
 1616
 1617
 1618
 1619
 1620
 1621
 1622
 1623
 1624
 1625
 1626
 1627
 1628
 1629
 1630
 1631
 1632
 1633
 1634
 1635
 1636
 1637
 1638
 1639
 1640
 1641
 1642
 1643
 1644
 1645
 1646
 1647
 1648
 1649
 1650
 1651
 1652
 1653
 1654
 1655
 1656
 1657
 1658
 1659
 1660
 1661
 1662
 1663
 1664
 1665
 1666
 1667
 1668
 1669
 1670
 1671
 1672
 1673
 1674
 1675
 1676
 1677
 1678
 1679
 1680
 1681
 1682
 1683
 1684
 1685
 1686
 1687
 1688
 1689
 1690
 1691
 1692
 1693
 1694
 1695
 1696
 1697
 1698
 1699
 1700
 1701
 1702
 1703
 1704
 1705
 1706
 1707
 1708
 1709
 1710
 1711
 1712
 1713
 1714
 1715
 1716
 1717
 1718
 1719
 1720
 1721
 1722
 1723
 1724
 1725
 1726
 1727
 1728
 1729
 1730
 1731
 1732
 1733
 1734
 1735
 1736
 1737
 1738
 1739
 1740
 1741
 1742
 1743
 1744
 1745
 1746
 1747
 1748
 1749
 1750
 1751
 1752
 1753
 1754
 1755
 1756
 1757
 1758
 1759
 1760
 1761
 1762
 1763
 1764
 1765
 1766
 1767
 1768
 1769
 1770
 1771
 1772
 1773
 1774
 1775
 1776
 1777
 1778
 1779
 1780
 1781
 1782
 1783
 1784
 1785
 1786
 1787
 1788
 1789
 1790
 1791
 1792
 1793
 1794
 1795
 1796
 1797
 1798
 1799
 1800
 1801
 1802
 1803
 1804
 1805
 1806
 1807
 1808
 1809
 1810
 1811
 1812
 1813
 1814
 1815
 1816
 1817
 1818
 1819
 1820
 1821
 1822
 1823
 1824
 1825
 1826
 1827
 1828
 1829
 1830
 1831
 1832
 1833
 1834
 1835
 1836
 1837
 1838
 1839
 1840
 1841
 1842
 1843
 1844
 1845
 1846
 1847
 1848
 1849
 1850
 1851
 1852
 1853
 1854
 1855
 1856
 1857
 1858
 1859
 1860
 1861
 1862
 1863
 1864
 1865
 1866
 1867
 1868
 1869
 1870
 1871
 1872
 1873
 1874
 1875
 1876
 1877
 1878
 1879
 1880
 1881
 1882
 1883
 1884
 1885
 1886
 1887
 1888
 1889
 1890
 1891
 1892
 1893
 1894
 1895
 1896
 1897
 1898
 1899
 1900
 1901
 1902
 1903
 1904
 1905
 1906
 1907
 1908
 1909
 1910
 1911
 1912
 1913
 1914
 1915
 1916
 1917
 1918
 1919
 1920
 1921
 1922
 1923
 1924
 1925
 1926
 1927
 1928
 1929
 1930
 1931
 1932
 1933
 1934
 1935
 1936
 1937
 1938
 1939
 1940
 1941
 1942
 1943
 1944
 1945
 1946
 1947
 1948
 1949
 1950
 1951
 1952
 1953
 1954
 1955
 1956
 1957
 1958
 1959
 1960
 1961
 1962
 1963
 1964
 1965
 1966
 1967
 1968
 1969
 1970
 1971
 1972
 1973
 1974
 1975
 1976
 1977
 1978
 1979
 1980
 1981
 1982
 1983
 1984
 1985
 1986
 1987
 1988
 1989
 1990
 1991
 1992
 1993
 1994
 1995
 1996
 1997
 1998
 1999
 2000
 2001
 2002
 2003
 2004
 2005
 2006
 2007
 2008
 2009
 2010
 2011
 2012
 2013
 2014
 2015
 2016
 2017
 2018
 2019
 2020
 2021
 2022
 2023
 2024
 2025
 2026
 2027
 2028
 2029
 2030
 2031
 2032
 2033
 2034
 2035
 2036
 2037
 2038
 2039
 2040
 2041
 2042
 2043
 2044
 2045
 2046
 2047
 2048
 2049
 2050
 2051
 2052
 2053
 2054
 2055
 2056
 2057
 2058
 2059
 2060
 2061
 2062
 2063
 2064
 2065
 2066
 2067
 2068
 2069
 2070
 2071
 2072
 2073
 2074
 2075
 2076
 2077
 2078
 2079
 2080
 2081
 2082
 2083
 2084
 2085
 2086
 2087
 2088
 2089
 2090
 2091
 2092
 2093
 2094
 2095
 2096
 2097
 2098
 2099
 2100
 2101
 2102
 2103
 2104
 2105
 2106
 2107
 2108
 2109
 2110
 2111
 2112
 2113
 2114
 2115
 2116
 2117
 2118
 2119
 2120
 2121
 2122
 2123
 2124
 2125
 2126
 2127
 2128
 2129
 2130
 2131
 2132
 2133
 2134
 2135
 2136
 2137
 2138
 2139
 2140
 2141
 2142
 2143
 2144
 2145
 2146
 2147
 2148
 2149
 2150
 2151
 2152
 2153
 2154
 2155
 2156
 2157
 2158
 2159
 2160
 2161
 2162
 2163
 2164
 2165
 2166
 2167
 2168
 2169
 2170
 2171
 2172
 2173
 2174
 2175
 2176
 2177
 2178
 2179
 2180
 2181
 2182
 2183
 2184
 2185
 2186
 2187
 2188
 2189
 2190
 2191
 2192
 2193
 2194
 2195
 2196
 2197
 2198
 2199
 2200
 2201
 2202
 2203
 2204
 2205
 2206
 2207
 2208
 2209
 2210
 2211
 2212
 2213
 2214
 2215
 2216
 2217
 2218
 2219
 2220
 2221
 2222
 2223
 2224
 2225
 2226
 2227
 2228
 2229
 2230
 2231
 2232
 2233
 2234
 2235
 2236
 2237
 2238
 2239
 2240
 2241
 2242
 2243
 2244
 2245
 2246
 2247
 2248
 2249
 2250
 2251
 2252
 2253
 2254
 2255
 2256
 2257
 2258
 2259
 2260
 2261
 2262
 2263
 2264
 2265
 2266
 2267
 2268
 2269
 2270
 2271
 2272
 2273
 2274
 2275
 2276
 2277
 2278
 2279
 2280
 2281
 2282
 2283
 2284
 2285
 2286
 2287
 2288
 2289
 2290
 2291
 2292
 2293
 2294
 2295
 2296
 2297
 2298
 2299
 2300
 2301
 2302
 2303
 2304
 2305
 2306
 2307
 2308
 2309
 2310
 2311
 2312
 2313
 2314
 2315
 2316
 2317
 2318
 2319
 2320
 2321
 2322
 2323
 2324
 2325
 2326
 2327
 2328
 2329
 2330
 2331
 2332
 2333
 2334
 2335
 2336
 2337
 2338
 2339
 2340
 2341
 2342
 2343
 2344
 2345
 2346
 2347
 2348
 2349
 2350
 2351
 2352
 2353
 2354
 2355
 2356
 2357
 2358
 2359
 2360
 2361
 2362
 2363
 2364
 2365
 2366
 2367
 2368
 2369
 2370
 2371
 2372
 2373
 2374
 2375
 2376
 2377
 2378
 2379
 2380
 2381
 2382
 2383
 2384
 2385
 2386
 2387
 2388
 2389
 2390
 2391
 2392
 2393
 2394
 2395
 2396
 2397
 2398
 2399
 2400
 2401
 2402
 2403
 2404
 2405
 2406
 2407
 2408
 2409
 2410
 2411
 2412
 2413
 2414
 2415
 2416
 2417
 2418
 2419
 2420
 2421
 2422
 2423
 2424
 2425
 2426
 2427
 2428
 2429<br

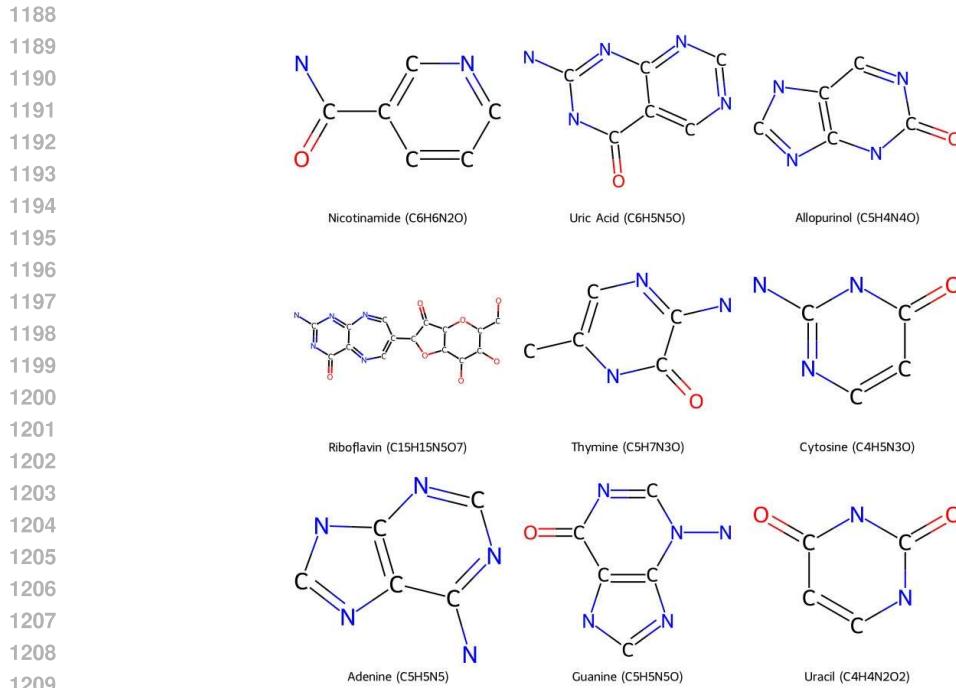


Figure 7: Sample complex molecular structures are generated using PARDIFF.

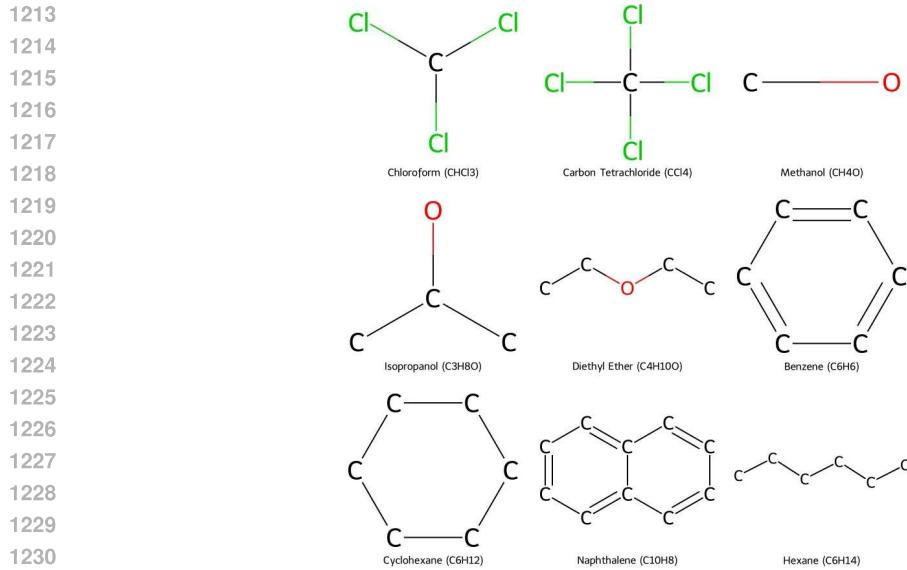


Figure 8: Sample complex molecular structures are generated using PARDIFF.

1235
 1236
 1237
 1238
 1239
 1240
 1241
 1242
 1243
 1244
 1245
 1246
 1247
 1248
 1249
 1250
 1251
 1252
 1253
 1254
 1255
 1256
 1257
 1258
 1259
 1260
 1261
 1262
 1263
 1264
 1265
 1266
 1267
 1268
 1269
 1270
 1271
 1272
 1273
 1274
 1275
 1276
 1277
 1278
 1279
 1280
 1281
 1282
 1283
 1284
 1285
 1286
 1287
 1288
 1289
 1290
 1291
 1292
 1293
 1294
 1295
 1296
 1297
 1298
 1299
 1300
 1301
 1302
 1303
 1304
 1305
 1306
 1307
 1308
 1309
 1310
 1311
 1312
 1313
 1314
 1315
 1316
 1317
 1318
 1319
 1320
 1321
 1322
 1323
 1324
 1325
 1326
 1327
 1328
 1329
 1330
 1331
 1332
 1333
 1334
 1335
 1336
 1337
 1338
 1339
 1340
 1341
 1342
 1343
 1344
 1345
 1346
 1347
 1348
 1349
 1350
 1351
 1352
 1353
 1354
 1355
 1356
 1357
 1358
 1359
 1360
 1361
 1362
 1363
 1364
 1365
 1366
 1367
 1368
 1369
 1370
 1371
 1372
 1373
 1374
 1375
 1376
 1377
 1378
 1379
 1380
 1381
 1382
 1383
 1384
 1385
 1386
 1387
 1388
 1389
 1390
 1391
 1392
 1393
 1394
 1395
 1396
 1397
 1398
 1399
 1400
 1401
 1402
 1403
 1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426
 1427
 1428
 1429
 1430
 1431
 1432
 1433
 1434
 1435
 1436
 1437
 1438
 1439
 1440
 1441
 1442
 1443
 1444
 1445
 1446
 1447
 1448
 1449
 1450
 1451
 1452
 1453
 1454
 1455
 1456
 1457
 1458
 1459
 1460
 1461
 1462
 1463
 1464
 1465
 1466
 1467
 1468
 1469
 1470
 1471
 1472
 1473
 1474
 1475
 1476
 1477
 1478
 1479
 1480
 1481
 1482
 1483
 1484
 1485
 1486
 1487
 1488
 1489
 1490
 1491
 1492
 1493
 1494
 1495
 1496
 1497
 1498
 1499
 1500
 1501
 1502
 1503
 1504
 1505
 1506
 1507
 1508
 1509
 1510
 1511
 1512
 1513
 1514
 1515
 1516
 1517
 1518
 1519
 1520
 1521
 1522
 1523
 1524
 1525
 1526
 1527
 1528
 1529
 1530
 1531
 1532
 1533
 1534
 1535
 1536
 1537
 1538
 1539
 1540
 1541
 1542
 1543
 1544
 1545
 1546
 1547
 1548
 1549
 1550
 1551
 1552
 1553
 1554
 1555
 1556
 1557
 1558
 1559
 1560
 1561
 1562
 1563
 1564
 1565
 1566
 1567
 1568
 1569
 1570
 1571
 1572
 1573
 1574
 1575
 1576
 1577
 1578
 1579
 1580
 1581
 1582
 1583
 1584
 1585
 1586
 1587
 1588
 1589
 1590
 1591
 1592
 1593
 1594
 1595
 1596
 1597
 1598
 1599
 1600
 1601
 1602
 1603
 1604
 1605
 1606
 1607
 1608
 1609
 1610
 1611
 1612
 1613
 1614
 1615
 1616
 1617
 1618
 1619
 1620
 1621
 1622
 1623
 1624
 1625
 1626
 1627
 1628
 1629
 1630
 1631
 1632
 1633
 1634
 1635
 1636
 1637
 1638
 1639
 1640
 1641
 1642
 1643
 1644
 1645
 1646
 1647
 1648
 1649
 1650
 1651
 1652
 1653
 1654
 1655
 1656
 1657
 1658
 1659
 1660
 1661
 1662
 1663
 1664
 1665
 1666
 1667
 1668
 1669
 1670
 1671
 1672
 1673
 1674
 1675
 1676
 1677
 1678
 1679
 1680
 1681
 1682
 1683
 1684
 1685
 1686
 1687
 1688
 1689
 1690
 1691
 1692
 1693
 1694
 1695
 1696
 1697
 1698
 1699
 1700
 1701
 1702
 1703
 1704
 1705
 1706
 1707
 1708
 1709
 1710
 1711
 1712
 1713
 1714
 1715
 1716
 1717
 1718
 1719
 1720
 1721
 1722
 1723
 1724
 1725
 1726
 1727
 1728
 1729
 1730
 1731
 1732
 1733
 1734
 1735
 1736
 1737
 1738
 1739
 1740
 1741
 1742
 1743
 1744
 1745
 1746
 1747
 1748
 1749
 1750
 1751
 1752
 1753
 1754
 1755
 1756
 1757
 1758
 1759
 1760
 1761
 1762
 1763
 1764
 1765
 1766
 1767
 1768
 1769
 1770
 1771
 1772
 1773
 1774
 1775
 1776
 1777
 1778
 1779
 1780
 1781
 1782
 1783
 1784
 1785
 1786
 1787
 1788
 1789
 1790
 1791
 1792
 1793
 1794
 1795
 1796
 1797
 1798
 1799
 1800
 1801
 1802
 1803
 1804
 1805
 1806
 1807
 1808
 1809
 1810
 1811
 1812
 1813
 1814
 1815
 1816
 1817
 1818
 1819
 1820
 1821
 1822
 1823
 1824
 1825
 1826
 1827
 1828
 1829
 1830
 1831
 1832
 1833
 1834
 1835
 1836
 1837
 1838
 1839
 1840
 1841
 1842
 1843
 1844
 1845
 1846
 1847
 1848
 1849
 1850
 1851
 1852
 1853
 1854
 1855
 1856
 1857
 1858
 1859
 1860
 1861
 1862
 1863
 1864
 1865
 1866
 1867
 1868
 1869
 1870
 1871
 1872
 1873
 1874
 1875
 1876
 1877
 1878
 1879
 1880
 1881
 1882
 1883
 1884
 1885
 1886
 1887
 1888
 1889
 1890
 1891
 1892
 1893
 1894
 1895
 1896
 1897
 1898
 1899
 1900
 1901
 1902
 1903
 1904
 1905
 1906
 1907
 1908
 1909
 1910
 1911
 1912
 1913
 1914
 1915
 1916
 1917
 1918
 1919
 1920
 1921
 1922
 1923
 1924
 1925
 1926
 1927
 1928
 1929
 1930
 1931
 1932
 1933
 1934
 1935
 1936
 1937
 1938
 1939
 1940
 1941
 1942
 1943
 1944
 1945
 1946
 1947
 1948
 1949
 1950
 1951
 1952
 1953
 1954
 1955
 1956
 1957
 1958
 1959
 1960
 1961
 1962
 1963
 1964
 1965
 1966
 1967
 1968
 1969
 1970
 1971
 1972
 1973
 1974
 1975
 1976
 1977
 1978
 1979
 1980
 1981
 1982
 1983
 1984
 1985
 1986
 1987
 1988
 1989
 1990
 1991
 1992
 1993
 1994
 1995
 1996
 1997
 1998
 1999
 2000
 2001
 2002
 2003
 2004
 2005
 2006
 2007
 2008
 2009
 2010
 2011
 2012
 2013
 2014
 2015
 2016
 2017
 2018
 2019
 2020
 2021
 2022
 2023
 2024
 2025
 2026
 2027
 2028
 2029
 2030
 2031
 2032
 2033
 2034
 2035
 2036
 2037
 2038
 2039
 2040
 2041
 2042
 2043
 2044
 2045
 2046
 2047
 2048
 2049
 2050
 2051
 2052
 2053
 2054
 2055
 2056
 2057
 2058
 2059
 2060
 2061
 2062
 2063
 2064
 2065
 2066
 2067
 2068
 2069
 2070
 2071
 2072
 2073
 2074
 2075
 2076
 2077
 2078
 2079
 2080
 2081
 2082
 2083
 2084
 2085
 2086
 2087
 2088
 2089
 2090
 2091
 2092
 2093
 2094
 2095
 2096
 2097
 2098
 2099
 2100
 2101
 2102
 2103
 2104
 2105
 2106
 2107
 2108
 2109
 2110
 2111
 2112
 2113
 2114
 2115
 2116
 2117
 2118
 2119
 2120
 2121
 2122
 2123
 2124
 2125
 2126
 2127
 2128
 2129
 2130
 2131
 2132
 2133
 2134
 2135
 2136
 2137
 2138
 2139
 2140
 2141
 2142
 2143
 2144
 2145
 2146
 2147
 2148
 2149
 2150
 2151
 2152
 2153
 2154
 2155
 2156
 2157
 2158
 2159
 2160
 2161
 2162
 2163
 2164
 2165
 2166
 2167
 2168
 2169
 2170
 2171
 2172
 2173
 2174
 2175
 2176
 2177
 2178
 2179
 2180
 2181
 2182
 2183
 2184
 2185
 2186
 2187
 2188
 2189
 2190
 2191
 2192
 2193
 2194
 2195
 2196
 2197
 2198
 2199
 2200
 2201
 2202
 2203
 2204
 2205
 2206
 2207
 2208
 2209
 2210
 2211
 2212
 2213
 2214
 2215
 2216
 2217
 2218
 2219
 2220
 2221
 2222
 2223
 2224
 2225
 2226
 2227
 2228
 2229
 2230
 2231
 2232
 2233
 2234
 2235
 2236
 2237
 2238
 2239
 2240
 2241
 2242
 2243
 2244
 2245
 2246
 2247
 2248
 2249
 2250
 2251
 2252
 2253
 2254
 2255
 2256
 2257
 2258
 2259
 2260
 2261
 2262
 2263
 2264
 2265
 2266
 2267
 2268
 2269
 2270
 2271
 2272
 2273
 2274
 2275
 2276
 2277
 2278
 2279
 2280
 2281
 2282
 2283
 2284
 2285
 2286
 2287
 2288
 2289
 2290
 2291
 2292
 2293
 2294
 2295
 2296
 2297
 2298
 2299
 2300
 2301
 2302
 2303
 2304
 2305
 2306
 2307
 2308
 2309
 2310
 2311
 2312
 2313
 2314
 2315
 2316
 2317
 2318
 2319
 2320
 2321
 2322
 2323
 2324
 2325
 2326
 2327
 2328
 2329
 2330
 2331
 2332
 2333
 2334
 2335
 2336
 2337
 2338
 2339
 2340
 2341
 2342
 2343
 2344
 2345
 2346
 2347
 2348
 2349
 2350
 2351
 2352
 2353
 2354
 2355
 2356
 2357
 2358
 2359
 2360
 2361
 2362
 2363
 2364
 2365
 2366
 2367
 2368
 2369
 2370
 2371
 2372
 2373
 2374
 2375
 2376
 2377
 2378
 2379
 2380
 2381
 2382
 2383
 2384
 2385
 2386
 2387
 2388
 2389
 2390
 2391
 2392
 2393
 2394
 2395
 2396
 2397
 2398
 2399
 2400
 2401
 2402
 2403
 2404
 2405
 2406
 2407
 2408
 2409
 2410
 2411
 2412
 2413
 2414
 2415
 2416
 2417
 2418
 2419
 2420
 2421
 2422
 2423
 2424
 2425
 2426
 2427
 2428
 2429
 2430
 2431
 2432
 2433
 2434
 2435
 2436
 2437
 2438
 2439
 2440
 2441
 2442
 2443
 2444
 2445
 2446
 2447
 2448
 2449
 2450
 2451
 2452
 2453
 2454
 2455
 2456
 2457
 2458
 2459
 2460
 2461
 2462
 2463
 2464
 2465
 2466
 2467
 2468
 2469
 2470
 2471
 2472
 2473
 2474
 2475
 2476
 2477
 2478
 2479
 2480
 2481
 2482
 2483
 2484
 2485
 2486
 2487
 2488

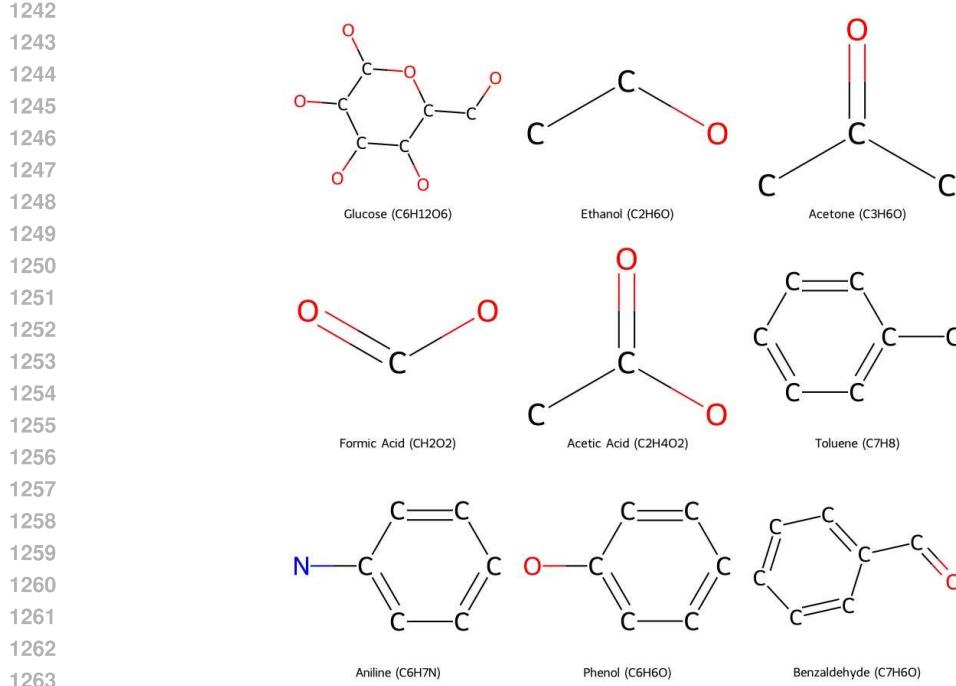


Figure 9: Sample complex molecular structures are generated using PARDIFF.

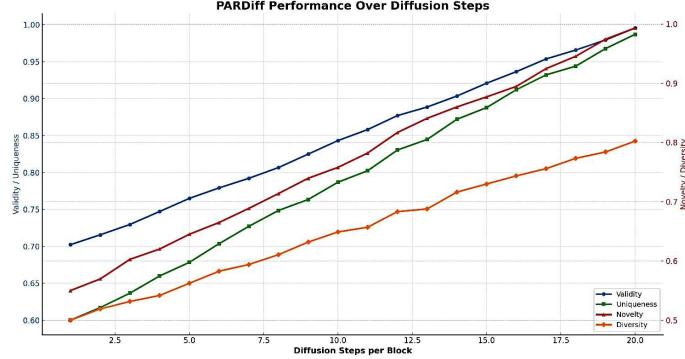


Figure 10: Non-curated structured grid graphs generated by PARDIFF, trained with 50 diffusion steps per block. The samples display mostly regular grid-like topology with occasional geometric perturbations, demonstrating the model’s ability to capture both structure and variation without any filtering.

1283
 1284
 1285 *Proof.* We prove the 1-WL implication by induction over layers and then conclude the orbit claim.
 1286 Let $w \in \mathcal{V}$ is the raw input features of one specific node of the graph $G = (V, E)$. $\mathbf{x}_w \in \mathbb{R}^d$ is
 1287 the input feature vector of node w . For example, in a molecular graph, \mathbf{x}_w might encode atom type,
 1288 charge, etc. Let the MPNN have L layers with following updates:

$$\mathbf{h}_w^{(0)} = \psi(\mathbf{x}_w); \mathbf{h}_w^{(l+1)} = U(\mathbf{h}_w^{(l)}, A(\mathbf{h}_t^{(l)} : t \in \mathcal{N}(w))), \quad (58)$$

1290 where A is a permutation-invariant multiset aggregator and U a shared update; $\Phi(w) = \mathbf{h}_w^{(L)}$. Let
 1291 $c_w^{(k)}$ denote the 1-WL color of node w after k rounds:

$$c_w^{(0)} = \text{Hash}(\mathbf{x}_w); c_w^{(k+1)} = \text{Hash}(c_w^{(k)}, c_t^{(k)} : t \in \mathcal{N}(w)). \quad (59)$$

1292 Induction hypothesis. Suppose for some $k \leq L$; $c_u^{(k)} = c_v^{(k)} \implies \mathbf{h}_u^{(k)} = \mathbf{h}_v^{(k)}$.

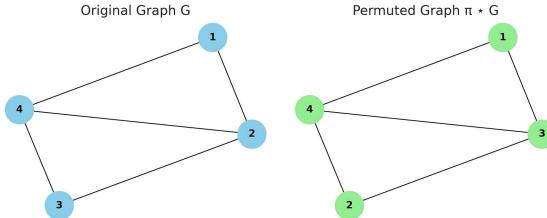


Figure 11: Illustration of permutation-consistency (equivariance, in Theorem 2) in node ranking using a 4 node graph. *Left*: The original graph G , where each node is annotated with its ranking value $\psi(u)$. *Right*: The permuted graph $\pi \star G$, obtained by swapping nodes 2 and 3. The ranking values move consistently with the node labels, showing that $\psi(\pi \star G) = \pi \star \psi(G)$. The structure and relative ordering are preserved under relabeling, demonstrating the permutation-invariance property of Algorithm 1.

1. Base ($k = 0$). If $c_u^{(0)} = c_v^{(0)}$, then \mathbf{x}_u and \mathbf{x}_v are in the same attribute class; since ψ is shared, $\mathbf{h}_u^{(0)} = \psi(\mathbf{x}_u) = \psi(\mathbf{x}_v) = \mathbf{h}_v^{(0)}$.
2. Step ($k \rightarrow k + 1$). Assume $c_u^{(k+1)} = c_v^{(k+1)}$. By 1-WL's update, we must have both

$$c_u^{(k)} = c_v^{(k)}; \{c_t^{(k)} : t \in \mathcal{N}(u)\} = \{c_t^{(k)} : t \in \mathcal{N}(v)\} \quad (60)$$

By the induction hypothesis, $c_u^{(k)} = c_v^{(k)}$ implies $\mathbf{h}_u^{(k)} = \mathbf{h}_v^{(k)}$. Moreover, the multiset equality of neighbor colors implies (again by the hypothesis applied elementwise) that the multisets of neighbor embeddings coincide:

$$\{\mathbf{h}_t^{(k)} : t \in \mathcal{N}(u)\} = \{\mathbf{h}_t^{(k)} : t \in \mathcal{N}(v)\}. \quad (61)$$

Applying the permutation-invariant aggregator A to equal multisets yields equal aggregated messages, and then the shared update U gives:

$$\begin{aligned} \mathbf{h}_u^{(k+1)} &= U(\mathbf{h}_u^{(k)}, A(\mathbf{h}_t^{(k)} : t \in \mathcal{N}(u))) \\ &= U(\mathbf{h}_v^{(k)}, A(\mathbf{h}_t^{(k)} : t \in \mathcal{N}(v))) = \mathbf{h}_v^{(k+1)}. \end{aligned} \quad (62)$$

Thus the claim holds for $k + 1$. By induction, for any L , $u \sim_{\text{WL}} v \implies \mathbf{h}_u^{(L)} = \mathbf{h}_v^{(L)}$, i.e., $\Phi(u) = \Phi(v)$.

Automorphism orbits. 1-WL is permutation-invariant; in particular, it assigns equal colors to nodes in the same automorphism orbit (an automorphism maps neighborhoods bijectively at every radius). Hence the 1-WL partition is a coarsening of the orbit partition, and the argument above shows MPNNs cannot refine beyond the WL partition. Therefore permutation-equivariant MPNNs cannot distinguish nodes within the same orbit, nor any pair that 1-WL fails to separate. This proves the corollary. \square

Corollary 1 makes explicit that the theoretical ceiling for most GNN architectures is the WL color refinement procedure. (1) Automorphism orbits define the hard limit: nodes indistinguishable under symmetry will always collapse to identical embeddings. (2) The WL hierarchy shows the algorithmic limit: even when automorphisms are broken, message-passing can at best refine equivalence classes to the 1-WL partition; and (3) Consequently, higher-order GNNs (e.g., K -WL-GNNs) or symmetry-breaking techniques (e.g., random features, positional encodings, anchor nodes) are required to exceed the expressivity of 1-WL. This bridges group theory (automorphisms), graph theory (orbit partitions), and deep learning (GNN expressivity), offering a unified lens on why standard GNNs fail on hard isomorphism cases such as strongly regular graphs or CAI–FÜRER–IMMERMAN (CFI) graphs Wang et al. (2023) (see Fig. 12).

1350

1351

1352

1353

1354

1355

1356

1357

1358

1359

1360

1361

1362

1363

1364

1365

1366

1367

1368

1369

1370

1371

1372

1373

1374

1375

1376

1377

1378

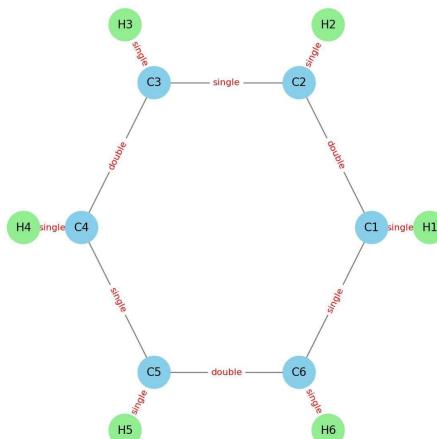
1379

1380

1381

1382

Figure 12: Illustration of orbit equivalence under graph automorphisms. Shown is a 6-cycle graph C_6H_6 (Benzene ring), where nodes sharing the same color belong to the same orbit under the automorphism group $\text{Aut}(G)$. Any permutation-equivariant GNN assigns identical embeddings to nodes within the same orbit, regardless of its depth or capacity. This highlights the fundamental expressivity limitation: GNNs cannot distinguish structurally symmetric nodes without additional symmetry-breaking features or higher-order mechanisms (e.g., K -WL refinements).



1389

1390

1391

1392

1393

1394

1395

1396

1397

1398

1399

1400

1401

1402

1403