ALGEBRAIC KNOT DEEP GRAPH LEARNING FOR STRUCTURAL GENERALIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

As graph-structured data grow increasingly complex, the ability to reliably distinguish non-isomorphic graphs becomes essential, yet the difficulty of achieving it is often underestimated. Strongly regular graphs illustrate this challenge sharply: non-isomorphic graphs with the same parameters may share identical degrees and identical common-neighbor counts for every adjacent and every non-adjacent pair of nodes. As a result, they are indistinguishable not only to Message Passing Neural Networks (MPNNs), the dominant class of Graph Neural Networks (GNNs), but also to more expressive architectures matching the 3-dimensional Weisfeiler–Leman (3-WL) test. To address this limitation, knot theory offers a conceptually different structural perspective, since chord diagrams represent knot crossing relations through interlacement graphs, providing a connection between knot-theoretic structure and graph adjacency. Motivated by this phenomenon, we introduce **KnotState**, which characterizes induced subgraphs through the rank–nullity structure of their adjacency matrices over $\mathbb{F}_2$. We then design **KSGL**, a graph learning framework that integrates multi-order local and global KnotState information with message passing. Specifically, we establish an order-dependent discriminative hierarchy for KSGL and introduce the KnotState separation order, showing that structural distinctions obtained at a given order are preserved at all higher orders. Further, we find that KSGL contains its underlying MPNN as a special case and can distinguish graph pairs that MPNNs, and in some cases 3-WL, leave unresolved. Experiments on challenging non-isomorphic graph benchmarks show that increasing the KnotState order progressively resolves previously indistinguishable graph pairs. We also demonstrate the utility of KnotState on zeolite property prediction under both interpolation and unseen-topology generalization. To the best of our knowledge, KSGL is the first GNN framework to systematically translate knot-theoretic state information into multi-order local–global graph representations with theoretical guarantees, establishing a new connection between knot theory and expressive graph learning. Code is available at `https://anonymous.4open.science/r/KSGL-0DBD/`

1 INTRODUCTION

Learning on graphs depends critically on structural information retained by the underlying representation. Message passing is a foundational mechanism in graph neural networks (GNNs), enabling nodes to exchange information along graph edges. Message passing neural networks (MPNNs) use this mechanism to learn node representations by iteratively aggregating information from neighboring nodes (Gilmer et al., 2017; Zhang et al., 2025). However, these standard MPNNs do not provide unrestricted structural resolution (Scarselli et al., 2009; Veličković et al., 2018; Zhou et al., 2020), and their discriminative capability is bounded by the Weisfeiler–Leman (WL) test (Xu et al., 2019; Morris et al., 2019). This limitation is particularly pronounced on regular graphs, where identical local features may conceal substantial differences in global organization (Bodnar et al., 2021). These observations shift the focus from *how messages are aggregated* to *which structural distinctions are actually retained*.

Existing approaches can be broadly divided into three lines of work. One line changes or enriches the basic computational unit through higher-order representations (Maron et al., 2019), permutation-sensitive constructions (Murphy et al., 2019), identity-aware message passing (You et al., 2021), and

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

structural identifiers in Subgraph GNNs (Huang et al., 2023). The second line exposes richer local structures through rooted subgraphs (Zhang & Li, 2021; Bevilacqua et al., 2022), subgraph-based structural representations (Frasca et al., 2022; Southern et al., 2025), multi-hop neighborhoods (Feng et al., 2022), cycle-aware refinement (Paolino et al., 2024), edge-centered ego structures (Alvarez-Gonzalez et al., 2024), and induced k-hop neighborhoods (Kim & Oh, 2025). The third line introduces graph-wide structural information through spectral encodings (Zhang et al., 2024), positional encodings (Flora et al., 2026), expressive pooling (Lachi et al., 2025), and global interaction mechanisms (Stoll et al., 2025; Ma et al., 2026). Recent logical analyses further characterize the expressive boundaries of these model classes (Xu, 2025; Besta et al., 2026).

However, distinguishing non-isomorphic graphs remains a challenging problem whose difficulty is often overlooked. Knot theory motivates an alternative perspective on structural characterization through algebraic invariants (Rolfsen, 1976; Lickorish, 1997; Adams, 2004). Recent advances in Knot Data Analysis (KDA) turn knot-theoretic quantities into localized quantitative descriptors for data analysis (Shen et al., 2024; Song et al., 2025). Fundamentally, chord diagrams provide a direct combinatorial connection between knots and graphs: chords correspond to vertices, while their intersections define graph adjacency (Chmutov et al., 2012). The resulting interlacement graph represents crossing relations through a binary adjacency matrix (Arratia et al., 2004a), while the interlace polynomial organizes vertex subsets through the ranks and nullities of principal adjacency submatrices (Arratia et al., 2004b). This connection suggests that induced subgraphs can be described not only by their explicit topology, but also by the algebraic dependency structure carried by their adjacency matrices.

We propose KnotState Deep Graph Learning (**KSGL**), a local–global KnotState framework for learning structural representations that improve graph discrimination and support generalization across graph families. Motivated by prior results on structural augmentation in Subgraph GNNs (Zhang & Li, 2021; Bevilacqua et al., 2022) and topological descriptors for graph learning (Horn et al., 2022; Shen et al., 2024), KSGL constructs higher-order induced structures both at the subgraph and graph. The resulting local and global KnotState representations are jointly incorporated into node features before message passing. To the best of our knowledge, KSGL is the first graph deep learning framework to introduce knot-theoretic state representations for joint local–global graph characterization. Our main contributions can be summarized as follows:

- We establish the order-dependent discriminative hierarchy of KSGL. We prove that KnotState representations are nested with respect to the maximum order K, so that every structural distinction obtained at order K is preserved at higher orders. We further introduce the KnotState separation order $\kappa_R(G, H)$ and prove that a graph pair with finite separation order remains distinguishable for every $K \geq \kappa_R(G, H)$.
- We characterize the discriminative advantage of KSGL over standard message passing, and prove that the underlying MPNN is contained as a special case of KSGL, while KSGL can strictly exceed its discriminative power. Moreover, KSGL is partially more discriminative than 3-WL, since graph pairs that remain indistinguishable to 3-WL can be separated by KSGL whenever their KnotState separation order is finite. Experiments on strongly regular graph families and BREC validate this advantage.
- We demonstrate the practical effectiveness of KSGL on multiple real-world graph learning tasks. On zeolite property prediction, we evaluate KSGL under both interpolation and topology-generalization settings, showing that KnotState representations provide useful structural information beyond synthetic expressivity benchmarks.

2 Related Works

Structural Augmentation of GNNs. Structural information can be introduced by modifying graph connectivity or by augmenting the representations supplied to a GNN. Graph rewiring changes the propagation structure itself. Locality-aware rewiring, for example, modifies connectivity while preserving sparsity and local structure (Barbero et al., 2024). Other approaches leave the graph unchanged and augment node or edge representations. Edge-level ego-network encodings expose local topology beyond ordinary message passing (Alvarez-Gonzalez et al., 2024), while spectral invariant models incorporate eigenspace information into node representations (Zhang et al., 2024).

Recent studies also use walk-based centrality as structural encoding (Southern et al., 2025), generalized distances and positional encodings in Graph Transformers (Stoll et al., 2025), and truncated spectral or walk-based positional encodings (Flora et al., 2026). Expressive pooling offers another way to retain structural distinctions at graph readout (Lachi et al., 2025). These methods enrich the structural signal available to the network, but generally encode neighborhoods, distances, spectra, or positional information rather than algebraic invariants of induced adjacency matrices.

Higher-order Graph Learning. Higher-order graph learning extends node-wise message passing through subgraphs, richer neighborhoods, or more expressive refinement schemes. Loopy WL and its neural counterpart incorporate cycle-related information beyond ordinary WL refinement (Paolino et al., 2024). Edge-level ego-network encodings increase expressivity through local subgraph structure (Alvarez-Gonzalez et al., 2024), while spectral invariant GNNs provide a different structural route beyond standard MPNNs (Zhang et al., 2024). WLKS applies Weisfeiler-Leman kernels to induced k-hop neighborhoods and combines information across multiple hop levels (Kim & Oh, 2025). HyMN reduces the computational burden of Subgraph-GNNs by selecting subgraphs through walk-based centrality while retaining stronger structural encodings (Southern et al., 2025). Recent logical analyses provide a finer characterization of the expressivity of higher-order and Subgraph-GNN architectures (Xu, 2025), and broader analyses of higher-order GNNs emphasize the trade-off between polyadic structural information and computational cost (Besta et al., 2026). Existing higher-order methods largely represent structure through explicit subgraphs, tuples, neighborhoods, or structural encodings, and induced-subgraph algebraic dependency statistics remain less explored.

Discriminative Power of GNNs for Graph Substructures. Most recent studies tend to move from asymptotic WL bounds toward finer empirical and structural measures of GNN expressivity. For example, BREC evaluates realized expressiveness on 800 non-isomorphic, 1-WL-indistinguishable graphs and contains instances with difficulty extending to higher-order WL tests (Wang & Zhang, 2024). Spectral invariant GNNs have been analyzed through a detailed hierarchy relative to Subgraph-GNNs and 3-WL (Zhang et al., 2024), while Loopy GNNs characterize expressivity through cycle and cactus-graph counting (Paolino et al., 2024). Edge-level ego-network encodings explicitly improve discrimination on strongly regular graphs (Alvarez-Gonzalez et al., 2024). Expressive pooling shows that graph-level aggregation itself can increase realized discrimination on BREC (Lachi et al., 2025), and recent Graph Transformers have also been evaluated directly on realized expressivity benchmarks (Stoll et al., 2025; Ma et al., 2026). Logical expressivity provides another characterization of which graph patterns different GNN architectures can represent (Xu, 2025), while recent work on positional encodings studies how truncation changes their WL-level discriminative power (Flora et al., 2026). These studies mainly measure graph-pair separation, motif counting, or WL-equivalent expressivity. By contrast, algebraic dependencies within the adjacency matrices of induced subgraphs remain comparatively underexplored.

3 Preliminaries on Knots

Traversing a regular knot diagram encounters each crossing twice. Recording these encounters in cyclic order on a circle and joining the two points corresponding to each crossing yields a *chord diagram*, with over/under information omitted (Chmutov et al., 2012). Two chords are *interlaced* if their endpoints alternate around the circle. The *interlacement graph* has one vertex per chord and an edge between each interlaced pair (Arratia et al., 2004a).

For a chord diagram D, its binary interlacement matrix M_D is the adjacency matrix of its interlacement graph. Selecting a subset S of chords gives the principal submatrix $M_D[S]$, corresponding to the induced subgraph. The binary adjacency-nullity descriptor introduced below applies to arbitrary simple undirected graphs, including those that admit no chord-interlacement representation.

Let $G = (V, E)$ be a finite simple undirected graph with $|V| = n$ and adjacency matrix A_G. For $S \subseteq V$, let $G[S]$ be its induced subgraph and $A_G[S]$ be the corresponding principal adjacency matrix. All ranks and kernels below are over $\mathbb{F}_2$, with arithmetic modulo 2.

Definition 3.1 (KnotState in Graph)**.** For $S \subseteq V$, the *KnotState* of the induced vertex set S is defined by $\nu_2(S) = \dim_{\mathbb{F}_2} \ker A_G[S] = |S| - \operatorname{rank}_{\mathbb{F}_2} A_G[S]$, where $x \in \ker_{\mathbb{F}_2} A_G[S]$ represents a binary adjacency dependency within the induced subgraph $G[S]$.

Definition 3.2 (Global and local KnotState distributions)**.** For an induced order s, the *global KnotState distribution* is $H_G^s(r) = \#\{S \subseteq V : |S| = s,\ \nu_2(S) = r\}$. For a root $u \in V$, let $B_R(u) = \{v \in V : d_G(u, v) \leq R\}$ be its radius-R neighborhood. The corresponding *local KnotState distribution* is $H_{G,u}^{s,R}(r) = \#\{S \subseteq B_R(u) : u \in S,\ |S| = s,\ \nu_2(S) = r\}$.

The counts allow disconnected induced subgraphs and remain preserved under graph isomorphisms, with roots mapped accordingly. The counts also recover basic size and degree statistics: $H_G^1(1) = n$, $H_G^2(0) = |E|$, and $H_{G,u}^{2,1}(0) = \deg_G(u)$. The complete family of global KnotState counts contains the same information as the two-variable interlace polynomial (Arratia et al., 2004b). (See Appendix A for more detailed discussion.)

4 Background on Structural Discriminative Power of Graph Neural Networks

We now proceed to a brief overview of theoretical results on discriminative power of Graph Neural Networks (GNNs).

Structural Discriminative Power of MPNNs A standard Message Passing Neural Network (MPNN) updates each node representation by combining its current state with information from its neighbors (Gilmer et al., 2017):

$$\mathbf{h}_u^{(t+1)} = U_t\left(\mathbf{h}_u^{(t)}, \sum_{v \in \mathcal{N}(u)} M_t(\mathbf{h}_u^{(t)}, \mathbf{h}_v^{(t)}, \mathbf{e}_{uv})\right), \tag{1}$$

where U_t and M_t represent two learnable functions with weights shared across all nodes. After T layers, the node representations are aggregated into a graph representation $\mathbf{h}_G$. The structural discriminative power of MPNNs is closely related to the 1-dimensional Weisfeiler–Leman test (1-WL).

Theorem 4.1 (Discriminative power of MPNNs (Xu et al., 2019; Morris et al., 2019))**.** *Standard neighborhood-aggregation MPNNs cannot distinguish graph pairs that are indistinguishable by 1-WL. With injective aggregation and update functions, this bound can be attained.*

However, graph structures that induce the same refinement remain indistinguishable.

Proposition 4.1 (Regular Graph (Zhang & Li, 2021))**.** *Let G and H be two d-regular graphs with the same number of vertices, constant and identical initial node features, and identical edge features. Then every standard MPNN produces the same graph representation for G and H.*

Discriminative Power of Subgraph MPNNs Subgraph MPNNs replace an aggregated neighborhood by an explicit graph associated with each node. Typical constructions include ego-subgraphs, node-marked graphs, and related node-based subgraph representations (Zhang & Li, 2021; Bevilacqua et al., 2022; Frasca et al., 2022). Because edges inside G_u are explicitly available to the encoder, Subgraph MPNNs can distinguish structural configurations that ordinary message passing collapses.

Proposition 4.2. *There exist non-isomorphic graphs that are indistinguishable by standard MPNNs but distinguishable by a node-based Subgraph GNN (Zhang & Li, 2021).*

For example, Nested GNNs are strictly more discriminative than 1-WL and can distinguish broad classes of regular graphs that ordinary message passing cannot distinguish (Zhang & Li, 2021). However, the additional subgraph context does not remove all structural ambiguity.

Theorem 4.2 (Upper bound of node-based Subgraph GNNs (Frasca et al., 2022))**.** *The node-based Subgraph GNN class analyzed by Frasca et al. (2022) is upper bounded in graph discriminative power by 3-WL. Consequently, graph pairs indistinguishable by 3-WL cannot be distinguished by models in this class.*

5 KnotState Deep Graph Learning

The limitations of MPNNs and Subgraph MPNNs motivate a *representation that can retain structural relations beyond neighborhood aggregation*. A key observation is that two non-isomorphic

graphs may share the same neighborhood-based representations, while exhibiting different structural relationships within their induced subgraphs. Therefore, it is conjectured that explicitly encoding such induced structural relations can improve the discriminative power on graphs whose local neighborhoods are highly homogeneous.

A naive solution is to explicitly encode all induced subgraphs. However, the number of vertex subsets grows combinatorially with the graph size, making such a construction impractical. Fortunately, the KnotState defined in Section 3 summarizes an induced subgraph through its binary adjacency dependency without requiring the induced subgraph itself to be retained as a separate neural input.

Given these observations, a global–local KnotState augmentation strategy is adopted, leading to *KnotState Deep Graph Learning (KSGL)*. For an induced size s, let $\mathbf{g}_s(G) = \left[H_G^s(0), \ldots, H_G^s(s)\right]$, $\mathbf{l}_{u,s}^R(G) = \left[H_{G,u}^{s,R}(0), \ldots, H_{G,u}^{s,R}(s)\right]$. The global vector is shared by all nodes in the same graph, while the local vector depends on u. For a maximum KnotState order K, KSGL assigns node u the structural representation $\mathbf{k}_u^{K,R} = \bigoplus_{s=1}^{K} \left(\mathbf{l}_{u,s}^R(G) \oplus \mathbf{g}_s(G)\right)$, where $\oplus$ denotes concatenation.

The resulting model associates every node with both its local KnotState information and the global KnotState information of the graph. Let $\mathbf{x}_u$ denote the original node attribute. The initial node representation is $\mathbf{h}_u^{(0)} = \mathrm{MLP}_{\mathrm{in}}\left(\mathbf{x}_u \oplus \rho\big(\mathbf{k}_u^{K,R}\big)\right)$, where $\rho(c) = \log(1+c)$ is applied elementwise. When node attributes are unavailable, $\mathbf{x}_u$ is replaced by a constant feature. KSGL then applies a shared GNN to the original graph. In the implementation, GIN (Xu et al., 2019) is used:

$$\forall u \in V, \qquad \mathbf{h}_u^{(\ell+1)} = \mathrm{MLP}_\ell\left((1+\varepsilon_\ell)\mathbf{h}_u^{(\ell)} + \sum_{v \in \mathcal{N}(u)} \mathbf{h}_v^{(\ell)}\right), \tag{2}$$

where ε_ℓ is a learnable scalar initialized to zero that controls the contribution of the node's own representation at layer ℓ. After $\mathcal{L}$ iterations, a graph representation is obtained by

$$\mathbf{h}_G = R_{\mathrm{graph}}\left(\left\{\mathbf{h}_u^{(\mathcal{L})} \mid u \in V\right\}\right). \tag{3}$$

Intuitively, KSGL improves structural resolution by augmenting each node with multi-order KnotState information, while leaving the underlying message-passing backbone unchanged. Hence, the additional structural encoding does not sacrifice the representation class of the original MPNN. See Figure 1 for an illustration of the KSGL encoding process, which extracts local and global KnotState representations from rooted and graph-wide induced subsets, respectively.

Proposition 5.1. *For any fixed K and R, KSGL contains its underlying MPNN as a special case. Proof is in Appendix B.1.*

Proposition 5.1 follows by suppressing the KnotState coordinates in the input encoder (see Appendix B.1 for more details). Hence, any additional distinction obtained by KSGL is introduced on top of the information available to the underlying MPNN. We next study how the structural information available to KSGL changes with the maximum KnotState order K. Since KSGL concatenates all orders up to K, increasing K preserves the existing representation while appending a new higher-order component.

Proposition 5.2 (Order hierarchy)**.** *For fixed radius R, we have $k_u^{K+1,R} = k_u^{K,R} \oplus l_{u,K+1}^R(G) \oplus g_{K+1}(G)$. Consequently, every KnotState component available at order K remains contained at every order $K' > K$. Proof is in Appendix B.2.*

As such, increasing K can introduce new structural distinctions but cannot remove a distinction already encoded at a lower order. To compare two graphs at a fixed induced order, we collect the joint local–global KnotState vectors over all roots, denoted as

$$\Phi_s^R(G) = \left\{\!\left\{ l_{u,s}^R(G) \oplus g_s(G) : u \in V(G) \right\}\!\right\},$$

where $\{\!\{\cdot\}\!\}$ is a multiset. Since the local and global components occupy separate coordinate blocks, the joint signature preserves a distinction present in either component. The corresponding global–local refinement property is proved in Appendix B.3. More importantly, the order-wise signature further allows us to identify the first structural scale at which a graph pair becomes distinguishable.

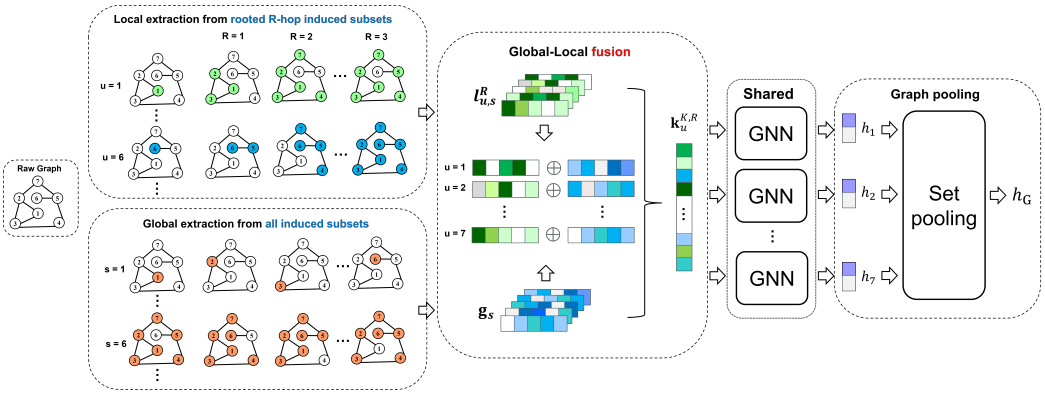


Figure 1: Architecture of KSGL: extract local KnotState statistics from rooted R-hop induced subsets and global KnotState statistics from all induced subsets, fuse them into node representations, and apply a shared GNN followed by graph pooling.

Definition 5.1 (KnotState Separation Order)**.** For two graphs G and H, their KnotState separation order is defined as

$$\kappa_R(G, H) = \min\left\{s \geq 1 : \Phi_s^R(G) \neq \Phi_s^R(H)\right\},$$

with $\kappa_R(G, H) = \infty$ if no such finite order exists.

Hence, $\kappa_R(G, H)$ is the smallest induced order at which KnotState information exposes a structural distinction between G and H. Specifically, the order hierarchy of Proposition 5.2 above implies that such a distinction, once exposed, cannot disappear when higher-order KnotState information is added.

Theorem 5.1 (Order-persistent separation)**.** *If $\kappa_R(G, H) < \infty$, then the KSGL structural representation distinguishes G and H at $K = \kappa_R(G, H)$, and the distinction remains available for every $K \geq \kappa_R(G, H)$. Proof is in Appendix B.4.*

Theorem 5.1 establishes a monotone hierarchy of structural resolution with respect to K (see Appendix B.4 for more detailed discussion). Compared to Proposition 4.1, for pairs of regular graphs, a finite KnotState separation order yields a distinction that is unavailable to the underlying MPNN, with strongly regular graphs importantly considered in our experiments. The same result also provides pairwise separations beyond the 3-WL boundary discussed in Theorem 4.2.

Corollary 5.1. *Let G and H be a non-isomorphic graph pair that remains indistinguishable to 3-WL. If $\kappa_R(G, H) < \infty$, then the KSGL structural representation distinguishes G and H for every $K \geq \kappa_R(G, H)$. Proof is in Appendix B.5.*

Hence, we can conclude that finite KnotState separation order can provide graph-pair distinctions beyond 3-WL.

6 Experiments

This section validates our theoretical results and evaluates KSGL's capabilities empirically. Specifically, we focus on the following questions:

Q1: How does KSGL's discriminative power evolve across KnotState orders?
Q2: How far can KSGL's discriminative power be extended on challenging non-isomorphic graph benchmarks?
Q3: How does KSGL perform compared to GNNs, and SymGNN on prediction benchmarks?

Table 1: Evolution of KSGL's cumulative pair-separation accuracy (%) across selected KnotState orders on strongly regular graph families.

Family	$K \leq 1$	$K \leq 4$	$K \leq 6$	$K \leq 10$	Family	$K \leq 1$	$K \leq 4$	$K \leq 6$	$K \leq 10$
$\mathrm{SRG}(16, 6, 2, 2)$	0.0000	100.0000	100.0000	100.0000	$\mathrm{SRG}(36, 14, 4, 6)$	0.0000	98.0819	99.9938	99.9938
$\mathrm{SRG}(25, 12, 5, 6)$	0.0000	100.0000	100.0000	100.0000	$\mathrm{SRG}(36, 15, 6, 6)$	0.0000	99.9996	100.0000	100.0000
$\mathrm{SRG}(26, 10, 3, 4)$	0.0000	95.5556	100.0000	100.0000	$\mathrm{SRG}(37, 18, 8, 9)$	0.0000	100.0000	100.0000	100.0000
$\mathrm{SRG}(28, 12, 6, 4)$	0.0000	100.0000	100.0000	100.0000	$\mathrm{SRG}(45, 12, 3, 3)$	0.0000	100.0000	100.0000	100.0000
$\mathrm{SRG}(29, 14, 6, 7)$	0.0000	99.2683	100.0000	100.0000	$\mathrm{SRG}(50, 21, 8, 9)$	0.0000	100.0000	100.0000	100.0000
$\mathrm{SRG}(40, 12, 2, 4)$	0.0000	98.9418	100.0000	100.0000	$\mathrm{SRG}(64, 18, 2, 6)$	0.0000	99.8269	99.9639	99.9856
$\mathrm{SRG}(35, 16, 6, 8)$	0.0000	99.9987	99.9999	100.0000	$\mathrm{SRG}(65, 32, 15, 16)$	0.0000	99.5968	100.0000	100.0000
$\mathrm{SRG}(35, 18, 9, 9)$	0.0000	99.9987	100.0000	100.0000					

Table 2: Ablation of global and local KnotState (KS) components on challenging strongly regular graph families. We report the number of remaining unresolved graph pairs; lower is better.

Family	Order	Global KS	Local KS	Joint KS	KSGL (ours)
$\mathrm{SRG}(35, 16, 6, 8)$	$K = 4$	281,272	729	729	**98**
	$K = 6$	146	34	34	**2**
$\mathrm{SRG}(35, 18, 9, 9)$	$K = 4$	281,272	729	729	**98**
	$K = 6$	209	25	25	**2**
$\mathrm{SRG}(36, 14, 4, 6)$	$K = 4$	1,314	377	377	**309**
	$K = 6$	16	9	8	**1**
$\mathrm{SRG}(36, 15, 6, 6)$	$K = 4$	21,599,051	36,412	36,412	**2,172**
	$K = 6$	687	110	84	**1**

6.1 Order-Dependent Structural Resolution

Experimental Settings To answer Q1, we evaluate KSGL on a collection of strongly regular graph (SRG) families. An $\mathrm{SRG}(n, k, \lambda, \mu)$ is a k-regular graph in which adjacent and non-adjacent vertex pairs have exactly λ and μ common neighbors, respectively (Brouwer & Van Maldeghem, 2022). Graphs with the same parameter tuple therefore share identical low-order statistics, making SRGs a challenging benchmark for structural discrimination (Spence, 2000; McKay & Spence, 2001; Haemers & Spence, 2001; Coolsaet et al., 2006; Gritsenko, 2021). In this experiment, we vary the maximum KnotState order K. For each K, KSGL uses all KnotState components up to order K, followed by deterministic 1-WL/GIN-style graph refinement. No additional supervised classifier is used for pair discrimination; two non-isomorphic graphs are considered separated whenever their resulting KSGL structural representations differ.

Findings Table 1 shows that pair-separation accuracy is preserved or improved as the maximum KnotState order K increases, consistent with Proposition 5.2. Once a graph pair is separated, the distinction is preserved at all higher orders, as predicted by Theorem 5.1. Consequently, the first order at which a pair becomes distinguishable corresponds to its KnotState separation order $\kappa_R(G, H)$ in Definition 5.1. Empirically, most families are resolved by $K = 4$ or 6, with the harder cases requiring higher orders.

Ablation study. Table 2 confirms that both global and local KnotState components contribute to structural discrimination. Their joint representation never performs worse than either component alone, consistent with Appendix B.3, while the KSGL structural refinement further reduces the number of unresolved graph pairs.

6.2 Discriminating Non-Isomorphic Graphs

Experimental Settings To answer Q2, we now evaluate KSGL on a broad collection of SRG families (Spence, 2000; Haemers & Spence, 2001; McKay & Spence, 2001; Coolsaet et al., 2006; Gritsenko, 2021). To evaluate the generalization capacity of our model for graph discrimination, we further use BREC (Wang & Zhang, 2024). BREC contains 400 non-isomorphic graph pairs

Table 3: Pair-separation accuracy (%) on challenging non-isomorphic graph benchmarks. Exact numbers of separated graph pairs are reported in Appendix C. T/O indicates that the experiment did not finish within the 48-hour time limit.

Benchmark	3-WL	4-ℓGIN	EPNN	CKGConv	PPGT	GDT+ RWSE	XP	Curv Pool	Resistance PE	Biharmonic PE	KSGL (ours)
SRG$(16, 6, 2, 2)$	0	100	100	0	0	0	100	0	0	0	**100**
SRG$(25, 12, 5, 6)$	0	100	0	0	0	0	100	90.48	0	0	**100**
SRG$(26, 10, 3, 4)$	0	97.78	51.11	0	0	0	100	93.33	0	0	**100**
SRG$(28, 12, 6, 4)$	0	100	0	0	0	0	100	83.33	0	0	**100**
SRG$(29, 14, 6, 7)$	0	99.88	13.90	0	0	0	100	0	0	0	**100**
SRG$(40, 12, 2, 4)$	0	99.21	0	0	0	0	100	0	0	0	**100**
SRG$(35, 16, 6, 8)$	0	T/O	100	0	0	0	99.85	94.33	50.03	1.24	**100**
BREC	67.50	64.25	53.75	54.50	58.50	50.75	31.00	39.25	54.25	55.00	**75**

spanning the Basic, Regular, Extension, and CFI categories, and provides a broader benchmark with structural patterns not restricted to strongly regular graphs. We compare KSGL with 10 state-of-the-art graph baselines reported on these benchmarks. The WL baseline is 3-WL (Morris et al., 2019), while 4-ℓGIN incorporates loopy structural information (Paolino et al., 2024). EPNN uses spectral information (Zhang et al., 2024), and CKGConv introduces continuous graph-kernel convolutions (Ma et al., 2024). We also include PPGT (Ma et al., 2026), GDT+RWSE (Stoll et al., 2025), the expressive pooling methods XP (Lachi et al., 2025) and CurvPool (Sanders et al., 2025), and the Resistance and Biharmonic positional encodings (Flora et al., 2026).

Findings KSGL distinguishes all graph pairs in the SRG families reported in Table 3 and achieves a pair-separation accuracy of 75% on BREC, correctly distinguishing 300 of 400 graph pairs. These results demonstrate KSGL's ability to distinguish structurally challenging graphs across the two evaluations. Exact pair counts are reported in Appendix C. Due to the computational cost of higher-order induced-subgraph enumeration, we evaluate BREC up to $K = 20$. Detailed results on BREC are provided in Appendix C (Table 8). Higher-order KnotState components may provide additional separation when their separation order exceeds 20, as predicted in Proposition 5.2.

6.3 Materials Property Prediction

Experimental Settings. To answer Q3, we use the zeolite adsorption benchmark of Petković et al. (2025), containing 27,648 structures from 108 framework topologies. We predict heat of adsorption and adsorption isotherms using MAE and MSE. For *generalization*, ITW and CHA are held out as unseen test topologies. For *interpolation*, each topology is split into training, validation, and test sets with an 80:10:10 ratio. We compare KSGL with GNN (Gilmer et al., 2017), and SymGNN (Petković et al., 2025). KSGL augments the same zeolite graph input with KnotState structural information and the same model comparison is used for both prediction tasks. All KSGL results are averaged over three different random seeds.

Findings Table 4 shows a clear performance gap between interpolation and unseen-topology generalization, indicating that transferring across different zeolite frameworks is substantially more challenging than predicting within known topologies. KSGL improves heat-of-adsorption prediction over GNN and SymGNN under both settings. These results suggest that the additional multi-order local–global KnotState information is useful when structural variation becomes important, consistent with the structural augmentation established in Section 5. More details are in Appendix D, where we further discuss the additional pressure-dependent factors involved in isotherm prediction.

6.4 Computational Efficiency

We evaluate computational efficiency on SRG$(35, 16, 6, 8)$ using a unified five-epoch protocol with a batch size of 16. Each method processes all 3,854 graphs, followed by an all-pairs comparison over $7,424,731$ graph pairs. We report the parameter count, peak allocated GPU memory measured by `torch.cuda.max_memory_allocated`, and synchronized no-gradient inference latency per graph. Offline feature preprocessing is excluded. For readability, the main text compares KSGL

Table 4: Zeolite prediction performance for heat of adsorption and adsorption isotherm under generalization and interpolation settings. Results are reported as mean ± standard deviation over three random seeds. Lower values are better (↓).

Model	Heat of Adsorption		Isotherm	
	MAE ↓	MSE ↓	MAE ↓	MSE ↓
Generalization				
GNN	1.6356 ± 0.3156	4.5152 ± 1.5961	0.9312 ± 0.3084	1.1379 ± 0.6977
SymGNN	1.5310 ± 0.1074	4.1802 ± 1.0658	$\mathbf{0.2951 \pm 0.0912}$	$\mathbf{0.2107 \pm 0.1473}$
KSGL (Ours)	$\mathbf{1.3666 \pm 0.0630}$	$\mathbf{3.1687 \pm 0.0690}$	0.3516 ± 0.0470	0.2351 ± 0.0595
Interpolation				
GNN	1.4490 ± 0.0044	3.8593 ± 0.0464	0.0729 ± 0.0056	0.0104 ± 0.0023
SymGNN	1.4671 ± 0.1093	3.6888 ± 0.1022	0.0767 ± 0.0043	0.0117 ± 0.0013
KSGL (Ours)	$\mathbf{1.3777 \pm 0.0191}$	$\mathbf{3.6189 \pm 0.0644}$	$\mathbf{0.0719 \pm 0.0083}$	$\mathbf{0.0103 \pm 0.0012}$

Table 5: Computational efficiency on SRG$(35, 16, 6, 8)$ after five training epochs. We report model size, peak allocated GPU memory, and synchronized inference latency per graph. Results are presented as mean ± standard deviation over three random seeds. Offline feature preprocessing is excluded.

Metric	KSGL (ours)	I^2-GNN	PPGT
Param. (M)	0.046 ± 0.000	0.303 ± 0.000	1.111 ± 0.000
Mem. (MiB)	43.788 ± 0.000	443.799 ± 0.000	508.343 ± 0.000
Latency (ms)	0.287 ± 0.000	15.505 ± 0.681	5.277 ± 0.423

with I^2-GNN and PPGT, while the complete comparison with all evaluated methods is provided in Appendix E (see Table 7).

KSGL has the smallest model, with only 0.046 million parameters, which is 38.2% fewer than the next-smallest baseline. It uses 43.788 MiB of peak GPU memory and requires 0.287 ms per graph, remaining substantially faster and more memory-efficient than several higher-order methods, including EPNN, I^2-GNN, and PPGT. Under this five-epoch efficiency protocol, order-6 KSGL separates $7,424,610$ of the $7,424,731$ SRG pairs (99.998%), demonstrating a favorable trade-off between discriminative power and computational cost.

7 Discussion, Limitations, and Future Directions

This paper introduces **KnotState**, an algebraic characterization of induced subgraphs through binary adjacency rank–nullity, and develops **KSGL**, which integrates multi-order local and global KnotState information with message passing. We have also established an order-dependent structural hierarchy in which distinctions obtained at lower orders persist at higher orders, while KSGL retains its underlying MPNN as a special case and can separate graph pairs left unresolved by standard message passing and, in certain cases, 3-WL. Experiments on multiple benchmarks have demonstrated the effectiveness of the same structural representation for both graph discrimination and practical prediction. The main current limitation of KSGL is the combinatorial cost of exact high-order KnotState computation, particularly for large graphs. Furthermore, Theorem 5.1 guarantees that a distinction, once obtained, persists at higher orders, but does not guarantee a finite separation order for every non-isomorphic graph pair. Future work will focus on scalable high-order computation, theoretical characterization of finite KnotState separation order, adaptive choices of K and R, and extending the KnotState principle to hypergraphs and other higher-order relational structures.

AI USE STATEMENT

We restricted the use of large language models (LLMs) to editing sentence structure and grammar for writing polishing, retrieval and discovery of related work, as well as code writing and debugging assistance. Experimental design, result generation, and the development of scientific ideas and analyses were conducted without LLM assistance.

REPRODUCIBILITY STATEMENT

We provide the source code, dataset access instructions, preprocessing procedures, model configurations, hyperparameters, random seeds, and evaluation scripts required to reproduce our experiments. The accompanying materials also document the training and evaluation protocols, software dependencies, and procedures used to aggregate results across independent runs.

ETHICS STATEMENT

This work focuses on methodological research in graph representation learning and does not involve human participants, animal subjects, or the collection of personal information. All experiments use publicly available or synthetic graph datasets. We follow the applicable dataset licenses and use the data solely for research purposes. We do not anticipate direct ethical risks beyond the general limitations and potential misuse associated with graph-based machine learning methods.

REFERENCES

Colin C. Adams. *The Knot Book: An Elementary Introduction to the Mathematical Theory of Knots*. American Mathematical Society, 2004.

Nurudin Alvarez-Gonzalez, Andreas Kaltenbrunner, and Vicenç Gómez. Improving subgraph-gnns via edge-level ego-network encodings. *Transactions on Machine Learning Research*, 2024.

Richard Arratia, Béla Bollobás, and Gregory B. Sorkin. The interlace polynomial of a graph. *Journal of Combinatorial Theory, Series B*, 92(2):199–233, 2004a.

Richard Arratia, Béla Bollobás, and Gregory B. Sorkin. A two-variable interlace polynomial. *Combinatorica*, 24(4):567–584, 2004b.

Federico Barbero, Ameya Velingker, Amin Saberi, Michael M. Bronstein, and Francesco Di Giovanni. Locality-aware graph rewiring in gnns. In *International Conference on Learning Representations*, 2024.

Maya Bechler-Speicher, Moshe Eliasof, Carola-Bibiane Schönlieb, Ran Gilad-Bachrach, and Amir Globerson. On the utilization of unique node identifiers in graph neural networks. *arXiv preprint arXiv:2411.02271*, 2024.

Maciej Besta, Florian Scheidl, Lukas Gianinazzi, Grzegorz Kwasniewski, Shachar Klaiman, Jurgen Muller, and Torsten Hoefler. Demystifying higher-order graph neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 48(3):2544–2565, 2026.

Beatrice Bevilacqua, Fabrizio Frasca, Derek Lim, Balasubramaniam Srinivasan, Chen Cai, Gopinath Balamurugan, Michael M. Bronstein, and Haggai Maron. Equivariant subgraph aggregation networks. In *International Conference on Learning Representations*, 2022.

Cristian Bodnar, Fabrizio Frasca, Yuguang Wang, Nina Otter, Guido F. Montufar, Pietro Lió, and Michael Bronstein. Weisfeiler and lehman go topological: Message passing simplicial networks. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 1026–1037, 2021.

Andries E. Brouwer and Hendrik Van Maldeghem. *Strongly Regular Graphs*, volume 182 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2022.

Sergei Chmutov, Sergei Duzhin, and Jacob Mostovoy. *Introduction to Vassiliev Knot Invariants*. Cambridge University Press, 2012.

Kris Coolsaet, Jan Degraer, and Edward Spence. The strongly regular $(45, 12, 3, 3)$ graphs. *The Electronic Journal of Combinatorics*, 13(1):R32, 2006.

Jiarui Feng, Yixin Chen, Fuhai Li, Anindya Sarkar, and Muhan Zhang. How powerful are k-hop message passing graph neural networks. In *Advances in Neural Information Processing Systems*, volume 35, pp. 4776–4790, 2022.

James Flora, Mitchell Black, Weng-Keen Wong, and Amir Nayyeri. Understanding truncated positional encodings for graph neural networks. In *International Conference on Machine Learning*, 2026.

Fabrizio Frasca, Beatrice Bevilacqua, Michael M. Bronstein, and Haggai Maron. Understanding and extending subgraph gnns by rethinking their symmetries. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning*, pp. 1263–1272, 2017.

Oleg Gritsenko. On strongly regular graph with parameters $(65, 32, 15, 16)$, 2021.

Willem H. Haemers and Edward Spence. The pseudo-geometric graphs for generalized quadrangles of order $(3, t)$. *European Journal of Combinatorics*, 22(6):839–845, 2001.

Max Horn, Edward De Brouwer, Michael Moor, Yves Moreau, Bastian Rieck, and Karsten Borgwardt. Topological graph neural networks. In *International Conference on Learning Representations*, 2022.

Yinan Huang, Xingang Peng, Jianzhu Ma, and Muhan Zhang. Boosting the cycle counting power of graph neural networks with I^2-gnns. In *International Conference on Learning Representations*, 2023.

Dongkwan Kim and Alice Oh. Generalizing weisfeiler-lehman kernels to subgraphs. In *International Conference on Learning Representations*, 2025.

Veronica Lachi, Alice Moallemy-Oureh, Andreas Roth, and Pascal Welke. Expressive pooling for graph neural networks. *Transactions on Machine Learning Research*, 2025.

W. B. Raymond Lickorish. *An Introduction to Knot Theory*, volume 175 of *Graduate Texts in Mathematics*. Springer, 1997.

Liheng Ma, Soumyasundar Pal, Yitian Zhang, Jiaming Zhou, Yingxue Zhang, and Mark Coates. Ckgconv: General graph convolution with continuous kernels. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, pp. 33902–33924. PMLR, 2024.

Liheng Ma, Soumyasundar Pal, Yingxue Zhang, Philip H. S. Torr, and Mark Coates. Plain transformers can be powerful graph learners. *Transactions on Machine Learning Research*, 2026.

Haggai Maron, Heli Ben-Hamu, Hadar Serviansky, and Yaron Lipman. Provably powerful graph networks. In *Advances in Neural Information Processing Systems*, volume 32, pp. 2156–2167, 2019.

Brendan D. McKay and Edward Spence. Classification of regular two-graphs on 36 and 38 vertices. *Australasian Journal of Combinatorics*, 24:293–300, 2001.

Christopher Morris, Martin Ritzert, Matthias Fey, William L. Hamilton, Jan Eric Lenssen, Gaurav Rattan, and Martin Grohe. Weisfeiler and leman go neural: Higher-order graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 4602–4609, 2019.

Ryan Murphy, Balasubramaniam Srinivasan, Vinayak Rao, and Bruno Ribeiro. Relational pooling for graph representations. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 4663–4673. PMLR, 2019.

Raffaele Paolino, Sohir Maskey, Pascal Welke, and Gitta Kutyniok. Weisfeiler and leman go loopy: A new hierarchy for graph representational learning. In *Advances in Neural Information Processing Systems*, volume 37, pp. 120780–120831, 2024.

Marko Petković, José-Manuel Vicent Luna, Elīza Beate Dinne, Vlado Menkovski, and Sofía Calero. Symmetry-informed graph neural networks for carbon dioxide isotherm and adsorption prediction in aluminum-substituted zeolites. *Journal of Materials Chemistry A*, 13:37018–37030, 2025.

Dale Rolfsen. *Knots and Links*. Publish or Perish, 1976.

Cedric Sanders, Andreas Roth, and Thomas Liebig. Curvature-based pooling within graph neural networks. In *Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, pp. 471–485, 2025.

Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1):61–80, 2009.

Li Shen, Hongsong Feng, Fengling Li, Fengchun Lei, Jie Wu, and Guo-Wei Wei. Knot data analysis using multiscale gauss link integral. *Proceedings of the National Academy of Sciences*, 121(42): e2408431121, 2024.

Ruzhi Song, Fengling Li, Jie Wu, Fengchun Lei, and Guo-Wei Wei. Multi-scale jones polynomial and persistent jones polynomial for knot data analysis. *AIMS Mathematics*, 10(1):1463–1487, 2025.

Joshua Southern, Yam Eitan, Guy Bar-Shalom, Michael M. Bronstein, Haggai Maron, and Fabrizio Frasca. Balancing efficiency and expressiveness: Subgraph GNNs with walk-based centrality. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 56615–56649. PMLR, 2025.

Edward Spence. The strongly regular $(40, 12, 2, 4)$ graphs. *The Electronic Journal of Combinatorics*, 7(1):R22, 2000.

Timo Stoll, Luis Müller, and Christopher Morris. Generalizable insights for graph transformers in theory and practice. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations*, 2018.

Yanbo Wang and Muhan Zhang. An empirical study of realized gnn expressiveness. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, pp. 52134–52155. PMLR, 2024.

Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *International Conference on Learning Representations*, 2019.

Tuo Xu. Towards a complete logical framework for gnn expressiveness. In *International Conference on Learning Representations*, 2025.

Jiaxuan You, Jonathan M. Gomes-Selman, Rex Ying, and Jure Leskovec. Identity-aware graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 10737–10745, 2021.

Bingxu Zhang, Changjun Fan, Shixuan Liu, Kuihua Huang, Xiang Zhao, Jincai Huang, and Zhong Liu. The expressive power of graph neural networks: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 37(3):1455–1474, 2025.

Bohang Zhang, Lingxiao Zhao, and Haggai Maron. On the expressive power of spectral invariant graph neural networks. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, pp. 60496–60526. PMLR, 2024.

Muhan Zhang and Pan Li. Nested graph neural networks. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. *AI Open*, 1:57–81, 2020.