3-WL IN DEPTH: UNIVERSAL GRAPH IDENTIFICATION VIA ITERATED LINE GRAPH TRANSFORMATION

Anonymous authors
Paper under double-blind review

ABSTRACT

The expressive power of graph neural networks is commonly studied through the Weisfeiler–Leman (WL) hierarchy, where higher dimensions distinguish more graph pairs. Through an injective graph transformation, the same model may also distinguish additional pairs while the original graph remains recoverable up to isomorphism. We study iterated line graph transformation and prove that every pair of non-isomorphic finite simple graphs is distinguished by 3-WL at some finite transformation depth r. Each additional transformation can distinguish pairs missed at all earlier depths, although no fixed depth suffices for all graphs. Based on these results, we introduce a universal line graph adapter for applying 3-WL-equivalent model families such as PPGN++ and G^2N^2 to iterated line graphs, and extend the implicit line graph neural network (ILG-3-GNN) to arbitrary line depth. On BREC, line graph transformation improves separation for all tested 3-WL-equivalent GNNs, with the best results matching 3-WL on $L^r(G)$ for $r \leq 3$, which separates 399 of the 400 pairs. Compared with 4-WL on the original graphs, shallow transformations can reduce time and memory per refinement round at lower edge densities, while density and degree variation raise the cost of further steps.

1 INTRODUCTION

The expressive power of graph neural networks (GNNs) describes their ability to distinguish non-isomorphic graphs. The Weisfeiler–Leman (WL) hierarchy is commonly used to study this ability. If a GNN is no more expressive than a WL test, it cannot distinguish graph pairs that the test fails to separate, regardless of training (Xu et al., 2019; Morris et al., 2019). Higher-order architectures improve expressivity by using larger vertex tuples or aggregating over more vertices jointly.

Graph transformation offers another way to learn from graphs without designing a new GNN architecture. If a transformation preserves isomorphism and is injective up to isomorphism, two original graphs are isomorphic exactly when their transformed graphs are. We can therefore apply a GNN to the transformed graphs and project any separation back to the root graphs (Yang & Huang, 2026). The change in graph structure may make pairs easier or harder to distinguish, depending on the transformation and the expressive power of the GNN.

The line graph $L(G)$ represents each edge of G by a vertex and joins two vertices when their edges share an endpoint. For connected graphs, Whitney's theorem states that the line graph determines the original graph up to isomorphism, except for the triangle–claw pair K_3 and $K_{1,3}$ (Whitney, 1932). On this class, Yang (2026) proves that 3-WL on line graphs is strictly more expressive than 3-WL on the original graphs. This containment ensures that every pair distinguished by 3-WL remains distinguishable after transformation. The additional separations suggest that repeating line graph transformation may make 3-WL increasingly expressive.

The effect of line transformation on regular graphs motivates the choice of 3-WL. With identical initial vertex colors, 1-WL first distinguishes vertices by degree. Since every vertex of a regular graph has the same degree, all vertices receive the same color in every subsequent refinement round. Line graph transformation preserves regularity (Ramane et al., 2005). Thus, non-isomorphic regular graphs with the same order and degree remain indistinguishable to 1-WL at every line depth. Under our convention, 3-WL updates the color of each vertex pair (u, v) using every vertex w and can therefore count triangles. It can distinguish regular graphs such as the triangular prism and $K_{3,3}$, but cannot distinguish strongly regular pairs with equal parameters. Yang & Huang (2026) show that

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.

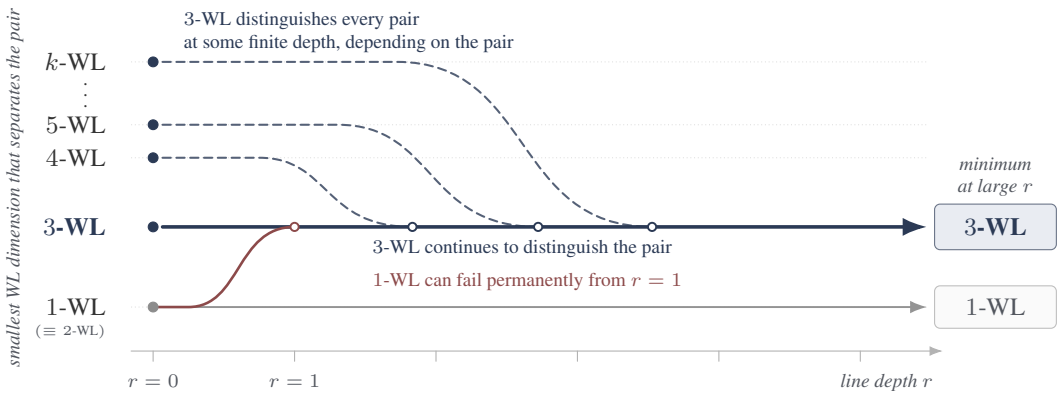


Figure 1: Line transformation can change the WL dimension needed to distinguish two connected graphs that are neither paths nor claws. Once 3-WL distinguishes such a pair, it continues to do so at every later depth. Every pair that initially requires a higher dimension becomes distinguishable by 3-WL at some finite depth. A 1-WL separation can be lost at the first transformation, and further transformations cannot revert this loss.

connected strongly regular graphs lose strong regularity after at most two transformations, apart from C_3, C_4, C_5. Although losing strong regularity does not guarantee separation, it suggests that 3-WL may distinguish the transformed graphs even though 1-WL still fails on them.

We therefore study whether iterated line graph transformation enables 3-WL to distinguish every pair of non-isomorphic graphs. We prove that for every $G \ncong H$, there is a finite $r \geq 0$ such that 3-WL distinguishes $L^r(G)$ and $L^r(H)$. By comparison, if 1-WL cannot distinguish G and H, it cannot distinguish their iterated line graphs at any depth (Yang, 2026). Figure 1 summarizes how transformation can change the WL dimension needed to distinguish a pair.

These results also show how to improve the expressive power of existing 3-WL-equivalent GNNs. In its general form, our *line graph adapter* applies a model to the original graph and its final iterated line graph, then concatenates the two outputs. We prove that the resulting model family has the same distinguishing power as applying the model to each of $G, L(G), \ldots, L^r(G)$ and concatenating all outputs. For connected graphs, the model output on the root graph can be replaced by its vertex count without changing which pairs the resulting model family can distinguish. The vertex count tells the triangle from the claw, which have the same line graph, and tells apart short paths whose iterates have all become empty.

The required number of transformations depends on the pair. For any fixed depth, there are connected non-isomorphic 3-regular graphs that 3-WL cannot distinguish at that depth or any earlier one. There is therefore no fixed depth at which 3-WL distinguishes every non-isomorphic graph pair.

We test whether GNNs of fixed width and number of layers can distinguish the additional pairs predicted by the theory. On BREC, we compare PPGN++, G^2N^2 and an extension of the native implicit line graph (ILG) neural network of Yang (2026) with exact WL tests. Each additional pair is evaluated at the first depth where exact 3-WL separates it, using only the model output on $L^r(G)$. To evaluate large iterated line graphs, we store one feature vector for each class of vertex pairs whose features remain equal throughout the network. This reduces memory use without changing the model's formulation.

Contributions.

1. We prove that 3-WL distinguishes every pair of non-isomorphic finite simple graphs after finitely many line graph transformations. Each additional depth can distinguish pairs missed at all earlier depths, but no fixed depth suffices for all pairs (Section 3).
2. We construct an adapter that applies a GNN to the original and final graphs. For a 3-WL-equivalent model family, we prove that also encoding the intermediate graphs adds no further distinguishing power (Section 4).

3. We evaluate PPGN++ and G^{2}N^2 on iterated line graphs, and extend ILG-3-GNN to repeated transformations (Section 4). On BREC, we measure which additional graph pairs these models distinguish (Section 5).
4. We compare the time and memory costs of increasing line depth with those of increasing WL dimension (Section 6).

2 Background and Notation

All graphs are finite, undirected and simple. Write $V(G), E(G)$ for their vertex and edge sets, $n = |V(G)|$, $m = |E(G)|$, $d_G(v)$ for degree and A_G for adjacency. I, J are the identity and all-ones matrices on the indicated set. Edges are unordered pairs. P_n, C_n, K_n denote the path, cycle and complete graph on n vertices. $K_{p,q}$ denotes the complete bipartite graph with parts of sizes p and q; $K_{1,3}$ is the claw.

Line graph transformation. The *line graph* $L(G)$ has one vertex for each edge of G; two vertices are adjacent exactly when their edges share an endpoint. Formally,

$$V(L(G)) = E(G), \qquad E(L(G)) = \big\{\{e, f\} : e, f \in E(G),\ e \neq f,\ e \cap f \neq \varnothing\big\}.$$

Define the iterates by $L^0(G) = G$ and $L^{r+1}(G) = L(L^r(G))$. We call r the *line depth*, distinct from WL refinement rounds or neural layers. Figure 2 illustrates two steps.

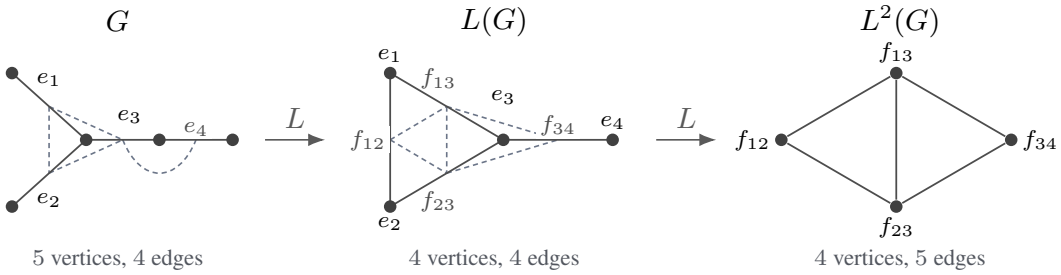


Figure 2: Edges become vertices, and incident edge pairs become adjacent vertices. Dashed lines connect the midpoints of edges that share a vertex in G and $L(G)$. Each dashed line becomes a solid edge in the next panel. Labels show how edges become vertices.

Weisfeiler–Leman refinement. The k-dimensional Weisfeiler–Leman (k-WL) test iteratively refines colors on ordered k-tuples of vertices. Here, 1-WL is ordinary vertex color refinement. For $k \geq 2$, we use the oblivious WL convention, which aggregates replacement colors separately for each coordinate. At stabilization, k-WL and $(k-1)$-dimensional folklore WL (FWL) distinguish the same graph pairs for $k \geq 3$ (Grohe, 2021). We therefore implement 3-WL using 2-FWL. The color $c_t(u, v)$ of an ordered pair (u, v) at round t is updated by

$$c_{t+1}(u, v) = \mathrm{HASH}(c_t(u, v), \{\{(c_t(u, w), c_t(w, v)) : w \in V(G)\}\}). \tag{1}$$

Initial colors record equality, adjacency and non-adjacency. HASH is injective, and the same encoding is used for both graphs; each replacement vertex belongs to the graph being refined. Write $G \equiv_{3\text{-WL}} H$ when the stable color histograms agree.

Implicit line-graph WL. Implicit line-graph WL (ILG-k-WL) evaluates k-WL on $L(G)$ by representing its vertices as edges of G and determining adjacency from shared endpoints (Yang, 2026). For $r \geq 1$, ILG^r-k-WL applies ILG-k-WL to $L^{r-1}(G)$, and ILG^0-k-WL is k-WL on G. It has the same distinguishing power as k-WL on $L^r(G)$. For $k = 3$, let $S_r(G)$ be the stable signature of ILG^r-3-WL on G, and write $S_{\leq r}(G) = (S_0(G), \ldots, S_r(G))$. For graph signatures, $S \preceq T$ means that T distinguishes every pair in the stated domain that S distinguishes: $T(G) = T(H)$ implies $S(G) = S(H)$. $S \prec T$ means that, in addition, T distinguishes some pair that S does not.

For an invariant encoder family $\mathcal{F}$, write $G \equiv_{\mathcal{F}} H$ if $f(G) = f(H)$ for every $f \in \mathcal{F}$. A 3-WL-equivalent family has exactly the relation $\equiv_{3\text{-WL}}$. This means that every pair separated by 3-WL is separated by some model in the family, with a suitable number of layers, width and parameters. It need not be separated by every such model.

3 LINE GRAPH ITERATION AND UNIVERSAL IDENTIFICATION

One line transformation can distinguish graph pairs that 3-WL misses on the original graphs (Yang, 2026). We first show that further transformations continue to increase expressivity. We then prove that every non-isomorphic pair is eventually distinguished, explain why the original graph must sometimes be retained, and show that no fixed number of transformations suffices for all pairs.

3.1 FROM ONE LINE STEP TO A STRICT HIERARCHY

On connected graphs outside $\{K_3, K_{1,3}\}$, the backward-transfer theorem of Yang (2026) gives

$$L(G) \equiv_{\text{3-WL}} L(H) \implies G \equiv_{\text{3-WL}} H.$$

Thus, if 3-WL distinguishes G and H, it also distinguishes $L(G)$ and $L(H)$. The converse fails. For example, 3-WL cannot distinguish the Shrikhande and 4×4 rook graphs, but distinguishes their line graphs (Yang, 2026). Therefore ILG-3-WL is strictly more expressive than 3-WL on connected graphs outside the Whitney exception.

We can apply the same theorem to successive line graphs. For a connected graph that is not a path, every iterate is nonempty and connected (van Rooij & Wilf, 1965). Every positive iterate is also claw-free (Beineke, 1970). Thus no later step can produce the triangle–claw exception. Consequently, for two connected graphs that are not paths, once 3-WL distinguishes their iterates at some positive depth, it distinguishes them at every later depth. The next result shows that every additional transformation also separates some previously indistinguishable pairs.

Theorem 1 (Strict depth hierarchy)**.** *For every $r \geq 1$, $S_r \prec S_{r+1}$ on connected graphs that are not paths. If $K_{1,3}$ is also excluded, this holds for $r = 0$. The hierarchy remains strict even when restricted to connected regular graphs of degree at least three. On all finite simple graphs, $S_{\leq r} \prec S_{\leq r+1}$ for every $r \geq 0$.*

Proof idea. The containment follows by applying one-step backward transfer (Theorem 9) to consecutive iterates. To prove strictness, Theorem 4 gives a connected 3-regular pair that remains indistinguishable through depth r. By universal identification (Theorem 2), it is first distinguished at some depth $t > r$. Take its $(t - r - 1)$st iterates as a new pair. This pair is indistinguishable through depth r and distinguishable at depth $r + 1$. Both graphs remain connected and regular, with degree at least three. The same examples prove $S_{\leq r} \prec S_{\leq r+1}$ on all graphs (Appendix B.10). □

Paths require separate treatment because they eventually become empty. We return to paths after proving universal identification.

Why 3-WL? Line transformation cannot make 1-WL distinguish every graph pair. For example, the triangular prism and $K_{3,3}$ have the same number of vertices and the same degree, as do their corresponding iterates. Hence 1-WL cannot distinguish them at any depth. More generally, any pair that is 1-WL-indistinguishable remains so after transformation (Yang, 2026). Under our convention, 2-WL has the same graph-distinguishing power as 1-WL. The next theorem shows that 3-WL suffices, which makes it the smallest WL dimension with which line transformation identifies all graphs.

3.2 DOES INCREASING LINE DEPTH IDENTIFY EVERY GRAPH?

Strictly increasing expressivity does not by itself imply that every non-isomorphic pair will be distinguished. We now prove this stronger statement for iterated line graphs.

Theorem 2 (Universal graph identification)**.** *For finite simple graphs G, H,*

$$G \cong H \quad \iff \quad L^r(G) \equiv_{\text{3-WL}} L^r(H) \text{ for every } r \geq 0.$$

Thus every non-isomorphic pair is distinguished at a finite line depth.

Proof idea. Suppose 3-WL cannot distinguish G, H at any depth. We show that they are isomorphic. First consider connected graphs with minimum degree at least five. A vertex of $L^2(G)$ represents two edges of G with a common endpoint. Stable 3-WL identifies the classes of such vertices sharing the

same endpoint. At greater depths, we combine these classes to represent a tuple listing all vertices of G, together with the equality and adjacency relations between its entries. Matching stable colors in the corresponding iterate of H then gives a tuple of distinct vertices with the same edges and nonedges. The two tuples define an isomorphism.

For a general connected graph, paths, cycles and the claw are identified by their order and degrees. Every other connected graph reaches minimum degree at least five after finitely many transformations (Hartke & Higgins, 2003). The preceding argument identifies those iterates, and Whitney's theorem recovers the original graphs. For disconnected graphs, we match connected components and apply the connected case. Appendix A.1 explains the construction; the complete proof is in Appendix B. □

For graph learning, the theorem means that a 3-WL-equivalent GNN family can distinguish every non-isomorphic pair by choosing a suitable transformation depth and suitable model parameters. No higher WL dimension is required. A particular finite network may still fail to distinguish a pair, which we examine experimentally.

3.3 Retaining the root makes the hierarchy general

The theorem guarantees separation at *some* depth, including zero, but not at every later depth. For example, $L(K_3) \cong L(K_{1,3}) \cong K_3$, and a path component P_j becomes empty after j steps. Removing isolated vertices from the original graph does not avoid this: P_2 and P_3 are separated at depth two, where $L^2(P_2) = \varnothing$ and $L^2(P_3) \cong P_1$, but not at depth three, where both iterates are empty. The same losses can occur inside disconnected graphs.

We can avoid this loss by retaining the signature of the original graph. In fact, its signature and that of $L^r(G)$ determine the signatures at all intermediate depths. To state the result, let $\kappa(G)$ count claw components and $p_j(G)$ count P_j components. Set $D_r(G) = (\kappa(G), p_1(G), \ldots, p_r(G))$ for $r \geq 1$.

Theorem 3 (First and last depths suffice)**.** *For finite simple graphs G, H and $r \geq 1$,*

$$\begin{aligned} S_{\leq r}(G) = S_{\leq r}(H) &\iff (S_0(G), S_r(G)) = (S_0(H), S_r(H)) \\ &\iff (D_r(G), S_r(G)) = (D_r(H), S_r(H)). \end{aligned}$$

Proof idea. The full list $S_{\leq r}$ contains S_0 and S_r, and S_0 determines D_r, because 3-WL equivalence preserves the numbers of claw and path components. For the converse, we recover the signatures from depth r down to zero. The signature of $L(G)$ determines that of G once the numbers of isolated vertices and claw components of G are known. These counts come from D_r: $L^q(G)$ has $p_{q+1}(G)$ isolated vertices, $L^q(G)$ is claw-free for $q \geq 1$, and $\kappa(G)$ is part of D_r (Appendix B.9). □

Therefore, comparing only the original and final signatures gives the same strict hierarchy and universal identification on all finite simple graphs. Section 4 uses this result to construct adapters requiring at most two model evaluations.

3.4 No fixed depth is universal

The depth required for identification can be arbitrarily large, even for connected 3-regular graphs.

Theorem 4 (No fixed depth suffices)**.** *For every $R \geq 0$, there are connected non-isomorphic 3-regular graphs G, H with $L^r(G) \equiv_{3\text{-}WL} L^r(H)$ for all $0 \leq r \leq R$.*

Proof idea. 3-WL equivalence is equivalence in first-order logic with counting quantifiers and three variables (Cai et al., 1992). A vertex of $L^r(G)$ can be described using at most $r + 1$ vertices of G. These descriptions translate each such formula on $L^r(G)$ into one on G, using $r + 1$ root variables for each line-graph variable. Graphs equivalent in counting logic with $3(r + 1)$ variables therefore remain 3-WL equivalent after r transformations. For any fixed number of variables, the uncolored 3-regular Cai–Fürer–Immerman (CFI) construction provides a non-isomorphic pair that this logic cannot distinguish (Cai et al., 1992; Chen et al., 2025). Taking $3(R + 1)$ variables therefore gives a pair indistinguishable at every depth up to R (Appendix B.10). These pairs are also used in the proof of the strict hierarchy. □

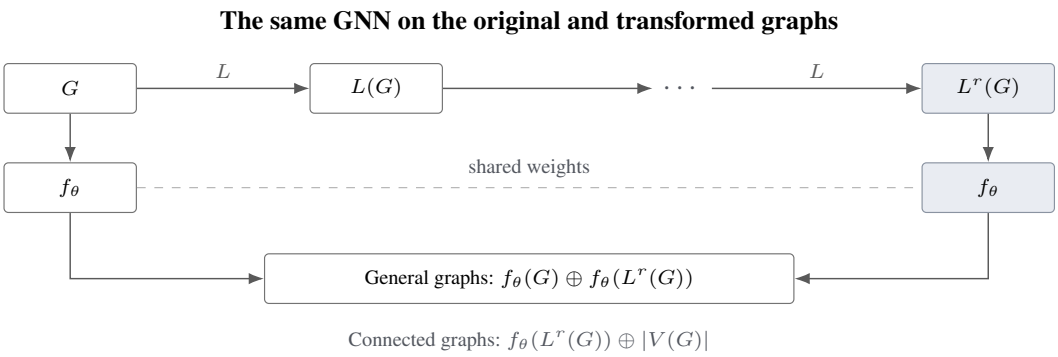


Figure 3: The general line graph adapter applies the same model to G and $L^r(G)$ and concatenates their outputs. Intermediate line graphs are used to construct $L^r(G)$ but are not passed through the model. For connected graphs, $|V(G)|$ can replace the output on G.

4 A Universal Adapter for Graph Encoders

The adapter uses only the model's graph input and graph-level output. It is independent of the model's internal layers, so the same construction applies to PPGN (Maron et al., 2019), PPGN++, and G^{2}N^2.

For the expressivity guarantee below, let f_θ belong to a 3-WL-equivalent GNN family, and let $\oplus$ denote concatenation. For arbitrary simple graphs, define the *general adapter*

$$z^{\mathrm{gen}}_{\theta,r}(G) = f_\theta(G) \oplus f_\theta(L^r(G)). \tag{2}$$

Figure 3 shows the two model evaluations. The weights may be shared, as in our default implementation. Allowing separate weights does not change which graph pairs the model family can distinguish. At $r = 0$, one evaluation gives both outputs.

For connected graphs, the *connected adapter* uses one neural branch:

$$z^{\mathrm{con}}_{\theta,r}(G) = f_\theta(L^r(G)) \oplus |V(G)|. \tag{3}$$

The vertex count of G separates the triangle from the claw, and separates short paths whose iterates have become empty. This definition therefore applies to every connected finite simple graph. If paths and the claw are excluded, the scalar can be omitted and only $L^r(G)$ is encoded.

Corollary 5 (Universal expressivity transfer)**.** *The general adapter family distinguishes exactly the same graph pairs as $S_{\leq r}$ on all finite simple graphs. The connected adapter has the same property on connected graphs. Each family therefore distinguishes every non-isomorphic pair in its domain at all sufficiently large depths, and becomes strictly more expressive with each additional depth.*

The general statement follows from Theorem 3. For connected graphs, the vertex count identifies a path whose iterate is empty and distinguishes the Whitney pair; in all other cases, backward transfer recovers the earlier signatures. Appendix C gives the proof. The argument applies to any 3-WL-equivalent family, using its existing expressivity result.

GNN instantiations. We instantiate the adapter with PPGN++ (Puny et al., 2023) and G^{2}N^2 (Piquenot et al., 2024). Both maintain pair features; G^{2}N^2 also maintains vertex features. The adapter supplies the equality and adjacency relations of $L^r(G)$ to their existing update rules. We also extend ILG-k-GNN (Yang, 2026) to arbitrary line depth. Let J_k denote its underlying graph model. Replacing $J_k(L(G))$ with $J_k(L^r(G))$ defines ILG^r-k-GNN. At positive depths, it represents vertices of $L^r(G)$ by edges of $L^{r-1}(G)$ and determines adjacency from shared endpoints. We test $k = 3$, matching the other model families and our expressivity guarantee. Appendix D.1 gives the model equations and implementation details.

5 Expressivity Evaluation on BREC

We test whether the adapted GNNs distinguish graph pairs that 3-WL cannot distinguish before transformation. BREC contains 400 non-isomorphic pairs and relabeled isomorphic pairs for reliability

Table 1: **BREC coverage grows with line depth.** Cumulative separations out of 400 through depth R. After the full depth-zero run, we evaluate each pair at the first depth where exact 3-WL separates it, avoiding the cost of larger line graphs for pairs already separated. *Pairs run* gives the number evaluated at that depth; native ILG differs at depth one (†). Indented rows split each depth's new separations by BREC category. Exact counts are certified 3-WL results; every neural separation passes its own isomorphic control.

				Neural separations, untrained models		
R	Pairs first separated at this depth	Pairs run	Exact	PPGN++	G^2N^2	Native ILG†
0	Root-separable groups (270)	400	270	270	270	—
1	Harder regular graphs (89)	89	359	359	359	359
	Strongly regular	50	+50	+50	+50	+50
	Four-vertex condition	20	+20	+20	+20	+20
	Distance regular	19	+19	+19	+19	+19
2	Remaining distance-regular pair	1	360	360	360	360
3	**Harder CFI pairs** (39)	**39**	**399**	**399**	**384**	**399**
	CFI, 3-WL-hard	20	+20	+20	+15	+20
	CFI, 4-WL-hard	19	+19	+19	+9	+19
6	Final CFI pair 354 (1)	—	400	—	—	—

All models use untrained FP64/$\sin(5x)$ evaluation on the stated graph transformation. The root row is a full 400-pair control; later rows evaluate only the pairs assigned to that depth. Counts through $R \leq 2$ agree across ten seeds; the $R = 3$ row and its sub-rows use seed 100 only. †Native ILG starts with all 400 pairs at $r = 1$. —: not evaluated. Depth six is certified only, with neural evaluation omitted due to computational limits.

checks (Wang & Zhang, 2024). We use the official dataset and keep its pair numbering at every depth. 3-WL distinguishes $270/400$ original pairs. Separating any of the other 130 pairs therefore demonstrates expressivity beyond the original 3-WL-equivalent model.

Exact WL results. For each pair, we determine the first depth at which 3-WL distinguishes its iterated line graphs, using exact refinement or a mathematical certificate. The cumulative numbers of separated pairs through depths 0–6 are $270, 359, 360, 399, 399, 399, 400$. One transformation adds 50 strongly regular, 20 four-vertex-condition and 19 distance-regular pairs. A second transformation separates the remaining distance-regular pair, and a third separates another 39 CFI pairs. Pair 354 is indistinguishable through depth five and distinguished at depth six. Appendix D.6 proves this without constructing the full sixth line graph.

Neural protocol. We first evaluate PPGN++ and G^2N^2 on all 400 original pairs and native ILG on all 400 pairs at depth one. We then evaluate each remaining pair at the first depth where exact 3-WL separates it. We test 89 pairs at depth one for PPGN++ and G^2N^2, one distance-regular pair at depth two, and 39 CFI pairs at depth three. All models use width 16, five neural layers, FP64 arithmetic, $\sin(5x)$ activation and a 16-dimensional output on $L^r(G)$ alone. No original-graph embedding or vertex count is appended in these experiments; the general adapter in Equation (2) is not evaluated.

All results in Table 1 use randomly initialized models without training. Counts through depth two agree across ten seeds $(100, 200, \ldots, 1000)$; depth three uses seed 100 for each backbone. Appendix D.2 reports activation and precision ablations that support our numerical settings.

Following the BREC evaluation protocol (Wang & Zhang, 2024), we evaluate 32 relabeled non-isomorphic comparisons and 32 isomorphic control comparisons for each graph pair and seed. All 128 graph inputs are evaluated independently; neural outputs are not reused across relabelings. We report separation only when both the main test and its own reliability control pass.

Compressed execution reduces memory without changing the model's formulation (Appendix C.3); Appendix D gives the remaining details.

Observed gains. PPGN++ rises from 270 root separations to 399 through depth three; native ILG rises from 359 at depth one to 399. G^2N^2 matches the exact WL counts through depth two and separates another 24 of the 39 depth-three CFI pairs, reaching 384. All 117 depth-three evaluations pass their own isomorphic control. Thus every model distinguishes more pairs after transformation, and PPGN++ and native ILG match the exact WL counts through depth three. At depth three, however, G^2N^2 outputs for relabeled copies of one graph lie up to 3.0 from their mean (output norm about 1.7), compared with below 10^{-4} for the other models. In each of its 15 misses, this variation exceeds the distance between the mean outputs of the two graphs. A finite network evaluated in floating point can therefore fall short of its family's expressivity.

Interpretation. For PPGN++, the gain from 270 to 399 pairs shows that an unchanged 3-WL-equivalent GNN applied to iterated line graphs can separate pairs beyond the 3-WL limit on the original graphs. The 39 depth-three pairs show that a third transformation adds further separations.

These results measure separation on BREC with the stated model parameters. They do not measure generalization to unseen graph families or performance on a prediction task. The table also does not establish that neural separation persists at greater depths, since each additional pair is evaluated only at its assigned depth.

6 Expressivity and Computational Cost

Increasing WL dimension uses larger vertex tuples, whereas increasing line depth uses the same pair update on a transformed graph. One-step ILG-3-WL and root 4-WL are incomparable (Yang, 2026). We also obtain $S_{\le r} \preceq (3r+2)$-FWL for every $r \ge 0$ (Appendix E.1). In particular, 6-WL on the original graphs distinguishes every pair distinguished by ILG-3-WL. Since one line step and one extra WL dimension do not subsume each other, the choice between them also depends on cost. We examine how graph structure affects the cost of these two ways to increase expressivity.

The cost of line refinement is set by the order of the transformed graph, which depends on the degrees of G. Write $n_r = |V(L^r(G))|$, and let $\bar{d}$ and σ_d^2 be the mean and variance of the degrees in G. Then

$$n_1 = m, \qquad n_2 = \sum_{v \in V(G)} \binom{d_G(v)}{2} = \frac{n}{2}\big(\bar{d}^2 - \bar{d} + \sigma_d^2\big). \tag{4}$$

Direct 4-WL, implemented as 3-FWL, stores n^3 tuple colors; ILG^r-3-WL stores n_r^2 pair colors. In each refinement round, the update for every pair on $L^r(G)$ uses all n_r replacement vertices, giving n_r^3 combinations. Root 4-WL instead considers n^4 combinations of a triple and a replacement vertex. The line test can therefore store fewer colors than 4-WL ($n_r^2 < n^3$) yet perform more work per round ($n_r^3 > n^4$). These counts describe direct enumeration; actual runtime also depends on sorting and assigning colors.

Figure 4 measures time and peak process memory for one refinement round, including construction and initialization, on an Apple M2 CPU. Both FWL dimensions use the same batched refinement routine. We use 24 settings, five graph seeds and five timing repetitions per graph and method. Panels A–C vary graph size, mean degree and degree variance in turn.

Sparsity can favor shallow transformation. At fixed mean degree, $n_1 = n\bar{d}/2$ grows linearly with n, so one transformation retains a quadratic pair-state count. In panel A, uniform graphs with 64 vertices and mean degree four take 0.061 seconds with ILG-3-WL versus 0.771 seconds with root 4-WL, a 12.7× speedup, with 58 versus 81 MiB peak memory. This advantage concerns the measured refinement step, not time to identification.

Density can reverse the cost ordering. Panel B holds $n = 24$ fixed. Increasing mean degree from two to six raises depth-two time on uniform graphs from 0.0057 to 1.588 seconds (279×), compared with 0.0256 to 0.0305 seconds (1.19×) for root 4-WL. ILG^2-3-WL changes from faster to slower than root 4-WL. Added edges enlarge the line graphs but leave the root tuple count unchanged.

Degree variation adds a cost at the second step. Panel C fixes both n and m, so n_1 stays constant. Increasing degree variance from zero to nine enlarges n_2 from 192 to 336 and raises depth-two

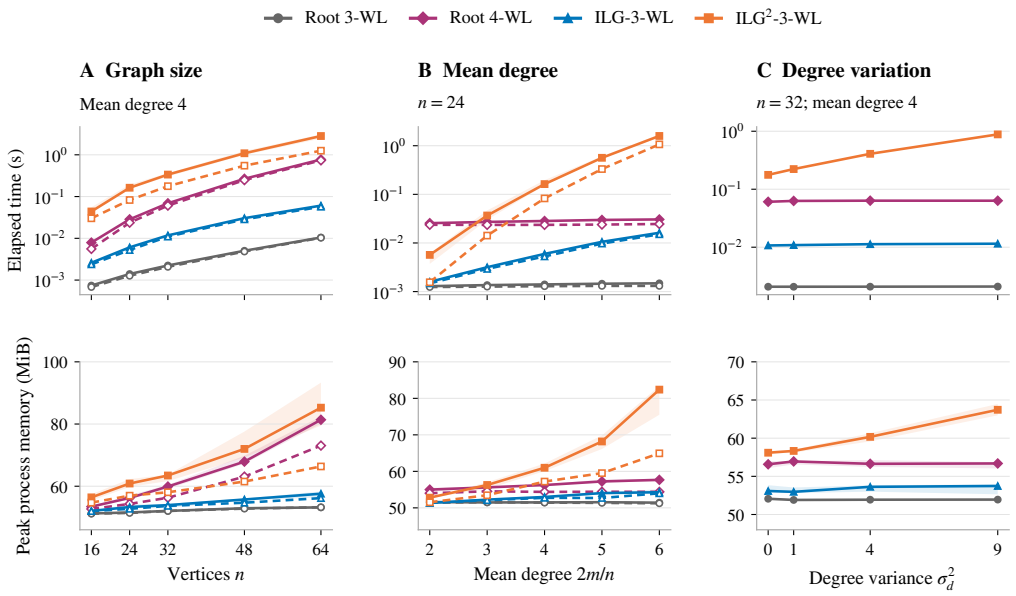


Figure 4: **Time and memory costs across graph sizes and degree distributions.** Solid curves represent uniform fixed-edge graphs in A–B and two-degree graphs in C; dashed curves represent regular graphs. Appendix E.4 gives the protocol.

time $4.99\times$, while depth-one and root costs change little. Thus with the same numbers of vertices and edges, more diverse node degrees make the second transformation much more expensive. Peak memory rises from 58 to 64 MiB. Its relative increase is smaller because these totals include roughly 51 MiB of baseline process memory.

Overall, the direct work of a line transformation grows with the degrees of the graph being transformed, whereas that of a higher WL dimension depends only on the number of vertices. Shallow transformation therefore gains most over root 4-WL on sparse graphs with similar degrees.

7 Discussion and Conclusion

We prove that 3-WL distinguishes every non-isomorphic graph pair at some finite line transformation depth. Each additional depth can distinguish new pairs, and retaining the original and final graph signatures is sufficient to preserve all earlier separations. The resulting adapter extends 3-WL-equivalent GNN families without changing their update rules. On BREC, all three tested models improve with line transformation, and PPGN++ and native ILG match the exact separation counts through depth three.

The required depth and graph size can still be large. For some families of 3-regular CFI graph pairs, the first separating depth grows at least linearly with the original vertex count. By that depth, the number of vertices in the iterated graph grows at least exponentially with the square of the original vertex count (Appendix E.3). These bounds concern the size of explicit line graphs; compressed execution may avoid storing them in full.

Our compressed evaluation uses graph symmetries to make the tested cases feasible. Efficient execution on less symmetric graphs remains an open problem. Another question is the width, number of layers and numerical conditioning needed for a finite network to match 3-WL on large line graphs. The depth-three G^2N^2 misses in Section 5 show that this question matters in practice.

We study uncolored graphs and graph-level representations. Extending the adapter to attributed graphs or node and edge prediction requires specifying how attributes and predictions map between the original and transformed graphs.

Reproducibility

The supplementary code includes both adapters, all three models, tests comparing dense and compressed evaluation, and the runners and configurations that regenerate every reported result: the exact WL depths and their per-pair certificates, the pair-354 certificates, the neural BREC evaluations, the numerical ablation and the cost measurements. It also contains the Lean formalization, its theorem index and its explicitly stated published premises. Appendices D and E specify the evaluation and resource measurements.

Ethics Statement

This work studies finite graph invariants and synthetic expressivity benchmarks. It uses no personal or sensitive data.

Use of AI Tools

AI tools assisted with proofs, implementation, synthetic graph generation, and writing TikZ code for diagrams. AI tools also assisted in writing and revising the manuscript. The authors formulated the research questions and mathematical claims, designed the experiments, and are responsible for all claims, citations and experimental results.

References

Lowell W. Beineke. Characterizations of derived graphs. *Journal of Combinatorial Theory*, 9(2): 129–135, 1970. doi: 10.1016/S0021-9800(70)80019-9.

Jin-Yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12(4):389–410, 1992. doi: 10.1007/BF01305232.

Gang Chen and Ilia Ponomarenko. *Coherent Configurations*. Central China Normal University Press, Wuhan, 2019.

Yijia Chen, Jörg Flum, and Mingjun Liu. Some remarks on the uncolored versions of the original CFI-graphs, 2025. URL https://arxiv.org/abs/2507.01459.

Frank Fuhlbrück, Johannes Köbler, and Oleg Verbitsky. Identifiability of graphs with small color classes by the Weisfeiler–Leman algorithm. *SIAM Journal on Discrete Mathematics*, 35(3): 1792–1853, 2021. doi: 10.1137/20M1327550.

Martin Grohe. The logic of graph neural networks. In *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, pp. 1–17, 2021. doi: 10.1109/LICS52264.2021.9470677.

Stephen G. Hartke and Aparna W. Higgins. Minimum degree growth of the iterated line graph. *Ars Combinatoria*, 69:275–283, 2003.

Lauri Hella. Logical hierarchies in PTIME. *Information and Computation*, 129(1):1–19, 1996. doi: 10.1006/inco.1996.0070.

Neil Immerman and Rik Sengupta. The k-dimensional Weisfeiler-Leman algorithm, 2019. URL https://arxiv.org/abs/1907.09582.

Martin Knor and Ľudovít Niepel. Histories in iterated line graphs. In *Zborník z výročnej konferencie Stavebnej fakulty STU*, pp. 201–210, Bratislava, 1998. Slovak Technical University.

Haggai Maron, Heli Ben-Hamu, Hadar Serviansky, and Yaron Lipman. Provably powerful graph networks. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/file/bb04af0f7ecaee4aae62035497da1387-Paper.pdf.

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. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):4602–4609, 2019. doi: 10.1609/aaai.v33i01.33014602.

Martin Otto. *Bounded Variable Logics and Counting: A Study in Finite Models*, volume 9 of *Lecture Notes in Logic*. Springer, Berlin, 1997.

Jason Piquenot, Aldo Moscatelli, Maxime Bérar, Pierre Héroux, Romain Raveaux, Jean-Yves Ramel, and Sébastien Adam. G^2N^2: Weisfeiler and Lehman go grammatical. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=eZneJ55mRO.

Ilia Ponomarenko. On the WL-dimension of circulant graphs of prime power order. *Algebraic Combinatorics*, 6(6):1469–1490, 2023. doi: 10.5802/alco.315.

Omri Puny, Derek Lim, Bobak Kiani, Haggai Maron, and Yaron Lipman. Equivariant polynomials for graph neural networks. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 28191–28222. PMLR, 2023. URL https://proceedings.mlr.press/v202/puny23a.html.

H. S. Ramane, H. B. Walikar, S. B. Rao, B. D. Acharya, P. R. Hampiholi, S. R. Jog, and I. Gutman. Spectra and energies of iterated line graphs of regular graphs. *Applied Mathematics Letters*, 18(6): 679–682, 2005. doi: 10.1016/j.aml.2004.04.012.

A. C. M. van Rooij and H. S. Wilf. The interchange graph of a finite graph. *Acta Mathematica Academiae Scientiarum Hungaricae*, 16(3–4):263–269, 1965. doi: 10.1007/BF01904834.

Yanbo Wang and Muhan Zhang. An empirical study of realized GNN expressiveness. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 52134–52155. PMLR, 2024. URL https://proceedings.mlr.press/v235/wang24cl.html.

Hassler Whitney. Congruent graphs and the connectivity of graphs. *American Journal of Mathematics*, 54(1):150–168, 1932. doi: 10.2307/2371086.

Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=ryGs6iA5Km.

Fan Yang. On the expressive power of implicit line-graph higher-order Weisfeiler–Leman, 2026. URL https://arxiv.org/abs/2609.16412.

Fan Yang and Xingyue Huang. Theoretical insights into line graph transformation on graph learning. In Francisco Acosta, Simone Azeglio, Bahareh Tolooshams, Chase van de Geijn, Christian Shewmake, Sophia Sanborn, and Nina Miolane (eds.), *Proceedings of the 4th (2025) and 3rd (2024) NeurIPS Workshops on Symmetry and Geometry in Neural Representations*, volume 282 of *Proceedings of Machine Learning Research*, pp. 996–1016. PMLR, 2026. URL https://proceedings.mlr.press/v282/yang26a.html.

A Mathematical Background and Proof Roadmap

A.1 Proof roadmap

To prove Theorem 2, suppose two graphs G, H agree under 3-WL at every line depth. We show that they are isomorphic. In a sufficiently deep iterate of G, we find vertices that encode a list of all vertices of G, and whose stable pair colors reveal which entries of the list are equal or adjacent. It suffices to encode one such list; we do not need to encode every tuple of vertices.

Building the list. First assume the graphs are connected and have minimum degree at least five. Each vertex of $L^2(G)$ is a pair of edges of G with a common endpoint. The Kneser relations of Appendix B.4 group these vertices by that endpoint, giving one group for each vertex of G. Appendix B.5 relates groups whose vertices lie at a given distance in G, and Appendix B.6 creates disjoint copies of the groups. With these constructions, we add the vertices of G to the list one at a time, keeping every entry tied to the same copy of G (Appendix B.7). Because groups can have different sizes, each construction divides by the group sizes, and every relation therefore counts vertices of G rather than their representatives.

Recovering an isomorphism. Matching the stable colors of G and H at a given depth induces matchings at every lower depth, and the backward-transfer identity of Appendix B.2 makes these matchings compatible (Appendix B.3). Every relation constructed for G therefore has a counterpart for H. We combine the relations into one matrix expression that is nonzero exactly on encoded lists whose entries are distinct and have the adjacencies of the list for G. The expression is nonzero for G; hence its counterpart is nonzero for H, and any list on which the counterpart is nonzero gives an isomorphism. Appendix B.7 gives the construction and the extraction of the isomorphism.

Returning to general graphs. Every connected graph other than a path, a cycle or the claw eventually has minimum degree at least five (Hartke & Higgins, 2003). The preceding argument identifies its sufficiently high iterates; Whitney's theorem then recovers the original graph. Paths, cycles and the claw are handled directly (Appendix B.8). For disconnected graphs, Appendix B.9 matches components with multiplicity and shows that one matching works at all depths. The connected result applies to each matched pair.

A.2 Notation and cited results

We define the matrix notation used in the proofs. Stable pair refinement is represented by a coherent matrix algebra; its operations let us construct relations that are preserved between equivalent graphs.

A.2.1 Stable pair algebras and transport

For a finite graph X, let

$$\mathcal{A}(X) = \mathcal{A}_X = \mathrm{CC}\langle I_{V(X)}, J_{V(X)}, A_X \rangle$$

be the coherent matrix algebra generated by equality and adjacency. Here $\mathrm{CC}\langle\cdot\rangle$ denotes closure over $\mathbb{R}$ under linear combinations, ordinary multiplication, transpose, and Hadamard product. Stable joint refinement of ordered pairs partitions $V(X)^2$ into basis relations. We write R_ρ^X for both such a relation and its 0–1 matrix. The matrices R_ρ^X have disjoint supports, sum to $J_{V(X)}$, and form the distinguished basis of $\mathcal{A}(X)$. Their intersection numbers are defined by

$$p_{\rho\sigma}^\tau = \big|\{z : (x,z) \in R_\rho^X,\ (z,y) \in R_\sigma^X\}\big| \quad \big((x,y) \in R_\tau^X\big),$$

and satisfy

$$R_\rho^X R_\sigma^X = \sum_\tau p_{\rho\sigma}^\tau R_\tau^X.$$

The intersection numbers do not depend on the chosen pair (x, y). This fact and the characterization by coherent closure follow from stable pair refinement (Fuhlbrück et al., 2021); see also Ponomarenko (2023).

An *algebraic isomorphism* $\Phi : \mathcal{A}(X) \to \mathcal{A}(Y)$ is a linear bijection that maps the distinguished basis bijectively and preserves I, J, transpose, ordinary multiplication, and Hadamard multiplication.

We use the term *weak isomorphism* for the same map. Trace preservation, diagonal-class sizes, and relation cardinalities follow from the coherent operations (Lemma 6); they are not additional hypotheses. Matching stable signatures gives an adjacency-preserving weak isomorphism: the shared signature encoding supplies a canonical stable-color bijection. A general weak isomorphism need not preserve adjacency. After choosing matching indices so that $\Phi(R_\rho^X) = R_\rho^Y$, we say that Φ *transports* M_X to M_Y when $M_Y = \Phi(M_X)$. A construction is called *canonical* below when it is defined from the distinguished coherent basis, independently of vertex labels.

Within the tuple construction we also write $\mathcal{W}(X) = \mathcal{A}(X)$. Each diagonal basis relation has the form I_U for a nonempty vertex set U, called a *fiber*. These fibers partition the vertex set.

Lemma 6 (Trace and cardinalities are intrinsic)**.** *Every weak isomorphism of coherent matrix algebras preserves trace. If $\Phi(R) = R'$ for distinguished basis relations, then $|R| = |R'|$; in particular, corresponding diagonal fibers have the same cardinality.*

Proof. Weak isomorphisms are exactly the algebraic isomorphisms of the corresponding coherent configurations (Chen & Ponomarenko, 2019). These preserve fiber sizes and relation sizes (Chen & Ponomarenko, 2019). Trace preservation follows, since a diagonal basis relation I_U has trace $|U|$ and every other basis relation has trace zero. □

Lemma 7 (Restriction to generated coherent subalgebras)**.** *Let $\mathcal{B}_X, \mathcal{B}_Y$ be coherent matrix algebras and let $\Phi : \mathcal{B}_X \to \mathcal{B}_Y$ be a weak isomorphism. If $\mathcal{S}_X \subseteq \mathcal{B}_X$ and $\mathcal{S}_Y = \Phi(\mathcal{S}_X)$, then Φ restricts to a weak isomorphism*

$$\mathrm{CC}\langle I, J, \mathcal{S}_X\rangle \longrightarrow \mathrm{CC}\langle I, J, \mathcal{S}_Y\rangle.$$

Proof. By Chen & Ponomarenko (2019), Φ maps the left generated algebra onto a coherent algebra and restricts to a weak isomorphism onto it. The image contains $I, J, \mathcal{S}_Y$, so it contains the right generated algebra; applying the same argument to Φ^{-1} gives equality. □

For a coherent matrix M and a scalar a, write $[M = a]$ for the entrywise 0–1 indicator of the level set $M_{xy} = a$; similarly define $[M \neq 0]$ and $[M > 0]$. We also write $\mathrm{supp}_+(M) = [M > 0]$. These selectors belong to the same coherent algebra by finite Lagrange interpolation with Hadamard products. The same interpolation polynomial on corresponding matrices shows that every selector used below is transported.

Lemma 8 (Finite coherent expressions)**.** *Let S be a finite set of real $n \times n$ matrices whose real span contains I, J. Every member of $\mathrm{CC}_{\mathbb{R}}\langle S\rangle$, and in particular every distinguished basis relation, is the value of a finite expression over S using real linear combinations, transpose, ordinary multiplication and Hadamard multiplication. Taking $S = \{I, J, A_X\}$ gives $\mathcal{A}(X)$.*

Proof. The values of such expressions form a coherent algebra containing S, and every coherent algebra containing S contains them. They are therefore the coherent closure of S (Fuhlbrück et al., 2021). □

A.2.2 Parabolics and normalized quotients

A *parabolic* is an equivalence relation whose matrix belongs to the coherent algebra. We also allow a *restricted parabolic*, defined on a subset of the vertex set. Thus

$$E = \bigcup_{a \in A} C_a \times C_a$$

for pairwise disjoint nonempty classes C_a. Its diagonal domain projector is $Z_E = I \circ E$. If N is nonnegative diagonal, let $N^\dagger$ replace each positive diagonal entry by its reciprocal and leave zero fixed. The intrinsic class-size normalizer is

$$D_E = (I \circ E^2)^\dagger, \qquad D_E(x, x) = |C_a|^{-1} \quad (x \in C_a).$$

For a matrix $M = (M_{ab})_{a,b \in A}$, its *block inflation* copies each entry M_{ab} to every position in $C_a \times C_b$:

$$\iota_E(M)[x, y] = \begin{cases} M_{ab}, & x \in C_a,\ y \in C_b, \\ 0, & x \notin \mathrm{dom}(E) \text{ or } y \notin \mathrm{dom}(E). \end{cases}$$

Such a matrix is called *block-complete*. The quotient map q_E recovers its value on each class pair, so it is the inverse of ι_E. Dividing by the size of each intermediate class gives

$$\iota_E(MN) = \iota_E(M)D_E\iota_E(N). \tag{A.1}$$

Thus a weak isomorphism that transports E also transports D_E and the corresponding quotient operations. No ordering of the classes is needed.

A.2.3 LINE-GRAPH CONVENTIONS

For a finite simple graph X, let $B_X \in \{0,1\}^{V(X)\times E(X)}$ be its unsigned vertex–edge incidence matrix, so $B_X(v,e) = \mathbf{1}[v \in e]$. With D_X the diagonal degree matrix,

$$B_X^{\mathsf{T}}B_X = 2I + A_{L(X)}, \qquad B_X B_X^{\mathsf{T}} = D_X + A_X. \tag{A.2}$$

We call a connected graph *Whitney-general* when it is neither K_3 nor $K_{1,3}$. On this domain the line graph determines its root up to isomorphism by Whitney's theorem (Whitney, 1932). This term refers to graph reconstruction up to isomorphism. Yang (2026) also proves backward transfer for stable pair refinement. Appendix B.3 applies their incidence-matrix identity repeatedly to the stable algebras at successive line depths.

A.2.4 CITED RESULTS

The proofs use coherent-algebra theory (Ponomarenko, 2023); the ILG equivalence, backward transfer with its incidence form, and component matching of Yang (2026); Whitney reconstruction (Whitney, 1932); and eventual minimum-degree growth (Hartke & Higgins, 2003). The proof of Theorem 4 additionally uses line-history support bounds (Knor & Niepel, 1998), counting-logic interpretations (Otto, 1997), and uncolored CFI lower bounds (Cai et al., 1992; Chen et al., 2025). Each result is stated with its hypotheses where it is applied.

B PROOFS OF THE MAIN GRAPH-THEORETIC RESULTS

B.1 EXACT EXECUTION FROM PARENT PAIRS

Let $\Omega_0 = V(G)$. Recursively define

$$\Omega_{s+1} = \{\{x,y\} : x,y \in \Omega_s,\ x \neq y,\ x \sim_{L^s(G)} y\}.$$

Fix an arbitrary internal ID order and list each pair in that order. The map $\Omega_{s+1} \to E(L^s(G))$ is a bijection, and hence $\Omega_s \simeq V(L^s(G))$ for every s. Two elements of Ω_{s+1} are adjacent exactly when their two-element parent sets intersect. We call an element of Ω_s a *history*: its nested parent pairs record how it was obtained from G. For $s \geq 1$, its *root footprint* is the subgraph of G formed by the root edges appearing in those pairs.

This parent-pair representation of $L^r(G)$ is the implicit line-graph representation of Yang (2026) applied to $L^{r-1}(G)$. Their proof that ILG-k-WL and k-WL on the line graph produce the same partitions in every round therefore shows that refinement on parent pairs computes $S_r(G)$ exactly. The internal pair identifiers are used only for storage and do not enter the initial colors.

B.2 ONE-STEP BACKWARD TRANSFER

We use the one-step backward-transfer theorem of Yang (2026) and record the incidence-matrix identity from its proof, which Appendix B.3 needs. The cited result uses joint ordered-pair refinement, called 2-FWL in the present paper, and then applies the standard arity shift. Its graph-level equivalence relation is therefore exactly the invariant denoted 3-WL here.

Theorem 9 (One-step backward transfer (Yang, 2026))**.** *Let X, Y be nonempty connected finite simple graphs. If $L(X) \equiv_{\text{3-WL}} L(Y)$, then either $X \equiv_{\text{3-WL}} Y$ or $\{X,Y\} = \{K_3, K_{1,3}\}$ up to isomorphism. In particular, the implication holds for connected Whitney-general graphs.*

B.2.1 INCIDENCE-MATRIX IDENTITY FROM BACKWARD TRANSFER

The coordinate construction will use relations from several line depths in one matrix algebra. For this purpose, we need more than the implication in Theorem 9: we also need the reconstructed maps to preserve incidence-matrix products. We record this property from the proof of Yang (2026) in their notation.

For a connected Whitney-general graph X, outside the four small roots K_1, K_2, the paw, and the diamond handled directly in the cited proof, the stable basis of $\mathcal{A}(L(X))$ canonically determines endpoint degrees, wedge types, root star profiles, and endpoint dart profiles. Let $\mathfrak{C}_X$ be the resulting partition of $V(X)$ into star-profile classes, with diagonal projectors P_C. Let

$$\mathcal{I}_X = \{B_X\} \cup \{\widetilde{B}_{X,\alpha} : \alpha \text{ has two unequal endpoint dart profiles}\}, \tag{B.1}$$

where B_X is unsigned vertex–edge incidence and each $\widetilde{B}_{X,\alpha}$ is the signed incidence obtained by ordering the two transported dart-profile labels. Equal-profile edge fibers are left unoriented.

The reconstruction forms

$$\mathcal{W}_X = \operatorname{span}\big(\{P_C : C \in \mathfrak{C}_X\} \cup \{P_C S M T^{\mathsf{T}} P_D : C, D \in \mathfrak{C}_X,\ S, T \in \mathcal{I}_X,\ M \in \mathcal{A}(L(X))\}\big). \tag{B.2}$$

Given an adjacency-preserving weak isomorphism $\varphi : \mathcal{A}(L(X)) \to \mathcal{A}(L(Y))$, write primes for the transported profile and incidence objects and $\Omega : \mathcal{W}_X \to \mathcal{W}_Y$ for the atomwise root map constructed in the cited reconstruction proof. That proof establishes the following three facts.

(P1) *Mixed Grams.* For every $C \in \mathfrak{C}_X$ and $S, T \in \mathcal{I}_X$,

$$S^{\mathsf{T}} P_C T \in \mathcal{A}(L(X)), \tag{B.3}$$

and the corresponding matrix is transported to its counterpart for Y.

(P2) *Ordinary algebra and compression.* The space $\mathcal{W}_X$ is a unital transpose-closed real algebra containing $I_{V(X)}, J_{V(X)}, A_X$. The map Ω is a well-defined trace-preserving algebra isomorphism. For every $Z \in \mathcal{W}_X$ and $U, V \in \mathcal{I}_X$,

$$U^{\mathsf{T}} Z V \in \mathcal{A}(L(X)), \qquad \varphi(U^{\mathsf{T}} Z V) = U'^{\mathsf{T}} \Omega(Z) V', \tag{B.4}$$

where U', V' are the transported incidence counterparts. The membership of $I_{V(X)}, J_{V(X)}, A_X$ follows from $B_X B_X^{\mathsf{T}} = D_X + A_X$ and $B_X J_{E(X)} B_X^{\mathsf{T}} = D_X J_{V(X)} D_X$, together with the transported star-profile degree diagonal D_X. On this domain D_X has strictly positive diagonal, so finite interpolation also transports D_X^{-1}; explicitly,

$$A_X = B_X B_X^{\mathsf{T}} - D_X, \qquad J_{V(X)} = D_X^{-1}(B_X J_{E(X)} B_X^{\mathsf{T}}) D_X^{-1}.$$

(P3) *Coherent closure.* Canonical profile cells are constructed first on the diagonal and root edges and then by induction on root distance. They are disjoint 0–1 matrices summing to $J_{V(X)}$, form a transported basis of $\mathcal{W}_X$, and make it Hadamard closed. Hence $\mathcal{W}_X$ is coherent and the transported map restricts to the stable root algebra $\mathcal{A}(X)$.

Items (P1)–(P3) summarize the backward-transfer proof in Yang (2026). Below we use its restriction to $\mathcal{A}(X)$ and equation (B.4) with unsigned incidence on both sides.

B.3 COMPATIBLE STABLE MAPS AND MATRIX CONSTRUCTIONS

Suppose 3-WL does not distinguish $L^r(G)$ and $L^r(H)$. Backward transfer gives a weak isomorphism between the stable algebras at each lower depth. We show that these maps preserve incidence-matrix products between successive depths. This allows a matrix relation constructed at one depth to be used at a later depth while preserving the correspondence between G and H.

For graphs G, H, let

$$B_i = B_{L^i(G)} \in \{0,1\}^{V(L^i(G)) \times V(L^{i+1}(G))},$$
$$B'_i = B_{L^i(H)} \in \{0,1\}^{V(L^i(H)) \times V(L^{i+1}(H))}.$$

For $0 \le a \le b$, define the unsigned ancestry matrix

$$C_{a,b} = B_a B_{a+1} \cdots B_{b-1}, \qquad C_{a,a} = I_{V(L^a(G))}.$$

Write $C'_{a,b}$ for the corresponding ancestry matrix of H.

Lemma 10 (Uniqueness on a graph-generated stable algebra)**.** *If $\Phi, \Psi : \mathcal{A}(X) \to \mathcal{A}(Y)$ are weak isomorphisms and $\Phi(A_X) = \Psi(A_X) = A_Y$, then $\Phi = \Psi$.*

Proof. Both maps preserve I, J, transpose, ordinary multiplication, Hadamard multiplication, and real linear combinations. They therefore agree on every finite expression over I, J, A_X. By Lemma 8, those expressions span $\mathcal{A}(X) = \mathrm{CC}\langle I, J, A_X\rangle$. □

Definition 11 (Reconstruction-admissible prefix)**.** *A finite prefix $G, L(G), \ldots, L^r(G)$ is* reconstruction-admissible *when every $L^i(G)$, $i < r$, is connected and is none of*

$$K_1,\ K_2,\ K_3,\ K_{1,3},\ \textit{the paw},\ \textit{or the diamond}.$$

These are precisely the connected graphs for which we use properties (P1)–(P3) and equation (B.4) *at every step.*

Theorem 12 (Compatible stable maps at successive depths)**.** *Let G, H have reconstruction-admissible depth-r prefixes, and let*

$$\phi_r : \mathcal{A}(L^r(G)) \longrightarrow \mathcal{A}(L^r(H))$$

be a weak isomorphism satisfying $\phi_r(A_{L^r(G)}) = A_{L^r(H)}$. Then there are unique adjacency-preserving weak isomorphisms

$$\phi_i : \mathcal{A}(L^i(G)) \longrightarrow \mathcal{A}(L^i(H)), \qquad 0 \le i < r,$$

obtained by applying the reconstruction (P1)–(P3) at each depth, and for every $i < r$ and $M \in \mathcal{A}(L^i(G))$,

$$\begin{aligned} B_i^{\mathsf{T}} M B_i &\in \mathcal{A}(L^{i+1}(G)), \\ \phi_{i+1}(B_i^{\mathsf{T}} M B_i) &= B_i'^{\mathsf{T}} \phi_i(M) B_i'. \end{aligned} \tag{B.5}$$

Consequently, for $0 \le a \le b \le r$ and every $M \in \mathcal{A}(L^a(G))$,

$$\begin{aligned} C_{a,b}^{\mathsf{T}} M C_{a,b} &\in \mathcal{A}(L^b(G)), \\ \phi_b(C_{a,b}^{\mathsf{T}} M C_{a,b}) &= {C'_{a,b}}^{\mathsf{T}} \phi_a(M) C'_{a,b}. \end{aligned} \tag{B.6}$$

Proof. Descend from $i = r - 1$ to 0. Suppose that ϕ_{i+1} has already been constructed. Apply the reconstruction of Yang (2026) to $L^i(G), L^i(H)$ and the stable line-algebra map ϕ_{i+1}. Items (P1)–(P3) give a transported coherent root space and a root-side weak isomorphism. Restricting that map to $\mathcal{A}(L^i(G)) = \mathrm{CC}\langle I, J, A_{L^i(G)}\rangle$ gives an adjacency-preserving weak isomorphism ϕ_i. Its restriction is unique by Lemma 10.

Now take $M \in \mathcal{A}(L^i(G))$. The stable root algebra is contained in the reconstructed root space $\mathcal{W}_{L^i(G)}$, and B_i is the unsigned member of the incidence family (B.1). Substituting $Z = M$ and $U = V = B_i$ into (B.4) gives both the membership assertion and (B.5), including compatibility with the already fixed line-side map ϕ_{i+1}. This completes the downward induction.

For (B.6), start with M at level a and apply (B.5) successively to

$$M, \quad B_a^{\mathsf{T}} M B_a, \quad B_{a+1}^{\mathsf{T}} B_a^{\mathsf{T}} M B_a B_{a+1}, \quad \ldots.$$

Associativity identifies the final matrix with $C_{a,b}^{\mathsf{T}} M C_{a,b}$, and the transported side with its primed counterpart. □

Corollary 13 (Backward transfer at every earlier depth)**.** *Under the admissibility hypotheses of Theorem 12,*

$$L^r(G) \equiv_{\text{3-WL}} L^r(H) \quad \Longleftrightarrow \quad L^i(G) \equiv_{\text{3-WL}} L^i(H) \textit{ for every } 0 \le i \le r.$$

Moreover, the canonical top stable-color map gives the unique family $(\phi_i)_{i=0}^r$ satisfying (B.5)*.*

Proof. The reverse implication is immediate. For the forward implication, equality of the top stable signatures supplies the adjacency-preserving weak isomorphism ϕ_r described in Appendix A.2. Theorem 12 constructs the remaining maps, so the stable signatures agree at every lower level. Uniqueness follows from Lemma 10. □

Corollary 14 (Restriction to selected vertex sets)**.** *In the setting of Theorem 12, every matrix obtained from a stable input by finitely many coherent operations at one level and products of the form $C_{a,b}^{\mathsf{T}} M C_{a,b}$ lies in the stable algebra of its output level and is transported by the corresponding ϕ_i. In particular, diagonal extraction followed by a level-set selector gives a diagonal projector onto a corresponding vertex subset. If Z is such a projector and $R = ZRZ$ is a transported symmetric zero–one relation, then ϕ_i restricts to a weak isomorphism between the supported coherent algebras generated by*

$$Z, \quad ZJZ, \quad R$$

and their counterparts for H.

Proof. Same-level closure and finite interpolation hold in every coherent algebra; level changes use (B.5) or (B.6). The matrices supported on this subset are closed under the same operations, with identity Z and all-ones matrix ZJZ. The proof of Lemma 7, applied inside the corresponding subsets, gives the final restriction. □

Corollary 15 (Minimum-degree sufficient condition)**.** *If G, H are connected and have minimum degree at least five, then for every r, both depth-r prefixes are reconstruction-admissible. Hence an adjacency-preserving weak isomorphism at depth r determines compatible maps at every earlier depth, satisfying* (B.5)–(B.6).

Proof. For every graph X of minimum degree at least two,

$$\delta(L(X)) \geq 2\delta(X) - 2.$$

Thus $\delta(L^i(G)) \geq 2^i(\delta(G) - 2) + 2 \geq 5$, and likewise for H. Line graphs of connected graphs with at least two edges are connected. Minimum degree five excludes every exceptional graph in Definition 11. □

B.3.1 Finite stable matrix constructions

The proofs below construct matrices using coherent operations and incidence products. We record the permitted operations once so that each later proof can focus on what its matrices count.

Definition 16 (Finite stable matrix construction)**.** *Fix a reconstruction-admissible finite prefix $X, L(X), \ldots, L^r(X)$. For each step t, let $\ell(t)$ be its line depth and let $M_t = \mathrm{ev}(t)$ be the resulting matrix. Its row and column sets are specified by coherent diagonal projectors $Z_L(t), Z_R(t)$, so*

$$M_t = Z_L(t) M_t Z_R(t) \in \mathcal{A}(L^{\ell(t)}(X)).$$

Thus a rectangular block is represented as a square matrix that vanishes outside its specified row and column sets. A construction is a finite sequence of such matrices, each obtained from the inputs or earlier matrices.

The inputs are finitely many specified matrices in $\mathcal{A}(L^\ell(X))$, including I, J, A at each required depth. Their row and column projectors must also be inputs or previously constructed matrices; an unrestricted matrix uses I. At one level we allow real linear combinations, transpose, ordinary and Hadamard products, diagonal extraction, and entrywise level-set selectors.

To pass from level ℓ to $\ell + 1$, we use the incidence product

$$\mathsf{Gram}_\ell(t) = B_\ell^{\mathsf{T}} \mathrm{ev}(t) B_\ell \quad \textit{at level } \ell + 1. \tag{B.7}$$

Equation (B.5) *shows that the result is coherent and is preserved by the corresponding weak isomorphism.*

We also use block inflation when the required class relation and generator matrices have already been constructed at the target level. A normalized inflation certificate *from level a to level $b > a$ consists of a finite generator set $S = \{s_\nu\} \subseteq \mathcal{A}(L^a(X))$, a coherent parabolic $E \in \mathcal{A}(L^b(X))$ whose nonempty classes are indexed by $V(L^a(X))$, its normalizer $D_E \in \mathcal{A}(L^b(X))$, and block-complete lifts $\widetilde{s}_\nu = \iota_E(s_\nu) \in \mathcal{A}(L^b(X))$. Matched certificates satisfy*

$$\phi_b(E) = E', \qquad \phi_b(D_E) = D_{E'}, \qquad \phi_b(\widetilde{s}_\nu) = \widetilde{s}'_\nu = \iota_{E'}(\phi_a(s_\nu)).$$

For a finite expression u over S, define

$$\begin{aligned} \mathsf{Inf}_E(s_\nu) &= \widetilde{s}_\nu, \\ \mathsf{Inf}_E(\alpha u + \beta v) &= \alpha \mathsf{Inf}_E(u) + \beta \mathsf{Inf}_E(v), \\ \mathsf{Inf}_E(u^{\mathsf{T}}) &= \mathsf{Inf}_E(u)^{\mathsf{T}}, \\ \mathsf{Inf}_E(u \circ v) &= \mathsf{Inf}_E(u) \circ \mathsf{Inf}_E(v), \\ \mathsf{Inf}_E(uv) &= \mathsf{Inf}_E(u) D_E \mathsf{Inf}_E(v). \end{aligned} \tag{B.8}$$

By (A.1)*, its value is exactly $\iota_E(\mathrm{ev}(u))$. All matrices on the right have already been constructed at level b, so the result belongs to $\mathcal{A}(L^b(X))$. The two-line Kneser certificate uses $S = \{I, J, A_{L^\ell(X)}\}$ and the transported lifts $E, J, R_{L^\ell(X)}$ at level $\ell + 2$.*

Reciprocals of nonzero diagonal entries are obtained by interpolation at the same level:

$$N^\dagger = \sum_{a \in \mathrm{im}(N) \setminus \{0\}} a^{-1}[N = a]. \tag{B.9}$$

For every selected value $a \neq 0$, interpolate on $\{0\} \cup \mathrm{im}(N)$. Its indicator polynomial vanishes at zero, so (B.9) *uses only positive Hadamard powers and needs neither J nor matrix inversion. A construction starting at $L^a(X)$ adds a to all displayed line depths.*

Lemma 17 (Transport of finite matrix constructions)**.** *Suppose matching finite prefixes carry the stable maps (ϕ_ℓ) of Theorem 12. Let P, P' be finite supported stable matrix constructions with the same sequence of operations, supports, and inflation data. Suppose their inputs and input supports correspond under those maps. Then every pair of corresponding expressions t, t' satisfies*

$$\phi_{\ell(t)}(\mathrm{ev}_X(t)) = \mathrm{ev}_Y(t'). \tag{B.10}$$

Proof. Induct through the common sequence of operations. The claim holds for matched inputs. A weak isomorphism preserves every coherent operation, and a finite selector uses the same interpolation polynomial on both sides. The Gram case is (B.5).

For a normalized inflation, the parabolic E, normalizer D_E, and generator lifts have already been constructed, so they correspond by induction. Induction through the five rules in (B.8), using (A.1) for multiplication, gives

$$\mathrm{ev}(\mathsf{Inf}_E(u)) = \iota_E(\mathrm{ev}(u)), \qquad \phi_b(\mathrm{ev}(\mathsf{Inf}_E(u))) = \mathrm{ev}(\mathsf{Inf}_{E'}(u')). \tag{B.11}$$

Here u' is the same expression with each s_ν replaced by $\phi_a(s_\nu)$. Lemma 8 supplies a finite expression u for every stable-basis generator requested by the construction. Formula (B.9) handles diagonal reciprocals. For the two-line Kneser certificate, (B.11) also gives

$$q_{E'}(\phi_{\ell+2}(\iota_E(M))) = \phi_\ell(M). \tag{B.12}$$

Thus every constructed matrix belongs to the stable algebra at its stated depth and is preserved by the corresponding map ϕ_ℓ. □

B.4 Recovering root vertices by Kneser relations

If $\delta(G) \geq 5$, stable pair refinement on $L^2(G)$ recovers a class for each vertex of G and the adjacency relation between these classes. The key observation is that four common neighbors identify disjoint edge pairs belonging to the same root star.

All matrices claimed to be coherent below are constructed using the operations in Definition 16. Rectangular incidence matrices may be used to describe their entries; the coherent matrices themselves are square matrices at the stated line depth.

B.4.1 The four-common-neighbor relation

Let G be a finite simple graph, put

$$X = L(G), \qquad Y = L(X) = L^2(G),$$

and write $A = A_Y$. A vertex of Y is an edge of X. Define the binary relation

$$D = (J - I - A) \circ [A^2 = 4]. \tag{B.13}$$

Thus $D(e, f)$ says that the two represented X-edges are disjoint and have exactly four common neighbors in $Y = L(X)$. This relation belongs to $\mathcal{A}(Y)$ by closure under multiplication and entrywise selection.

For a root vertex u, its incident edges form a clique in $X = L(G)$. The edge sets of these cliques partition $E(X)$; they are the *Krausz classes* used below.

Lemma 18 (Krausz meaning of D)**.** *For disjoint edges $e = ab$ and $f = cd$ of the line graph X,*

$$\begin{aligned} D(e, f) = 1 \quad &\iff \quad \{a, b, c, d\} \textit{ induces } K_4 \textit{ in } X \\ &\iff \quad e, f \textit{ belong to the same Krausz class of } X. \end{aligned} \tag{B.14}$$

Proof. A common neighbor of e and f in $L(X)$ is an X-edge joining one endpoint of e to one endpoint of f. There are at most four:

$$ac,\ ad,\ bc,\ bd.$$

There are exactly four precisely when all four cross edges exist. Together with ab, cd, this says that the four endpoints induce K_4.

Every clique of size at least four in a line graph consists of root edges through one common root vertex; a root triangle has only three members. Hence the four X-vertices lie in one root-star clique. The two X-edges e, f therefore belong to the same edge class of the Krausz decomposition. Conversely, two disjoint edges inside one root-star clique have all four cross edges, proving (B.14). □

B.4.2 Kneser connectivity recovers each center block

For $u \in V(G)$, let

$$W_u = \{\{e, f\} : e, f \in E_G(u),\ e \neq f\} \subseteq V(Y). \tag{B.15}$$

Inside W_u, relation D joins two two-element subsets of $E_G(u)$ exactly when they are disjoint. Thus

$$D[W_u] \cong KG(d_G(u), 2). \tag{B.16}$$

For every $s \geq 5$, $KG(s, 2)$ is connected with diameter at most two: disjoint two-sets are adjacent, while two intersecting two-sets have a common disjoint two-set because their union has size three and at least two points remain.

Theorem 19 (Recovering the root-center classes)**.** *If $\delta(G) \geq 5$, then*

$$E_W = \operatorname{supp}_+(I + D + D^2) = \bigcup_{u \in V(G)} W_u \times W_u \tag{B.17}$$

belongs to $\mathcal{A}(L^2(G))$.

Proof. Lemma 18 shows that D has no pair joining distinct Krausz classes. Equation (B.16) and diameter two show that every two members of one W_u are joined by a D-walk of length at most two. The W_u's partition $V(L^2(G))$, so the positive support in (B.17) is exactly their equivalence relation. Every operation in (B.17) is coherent. □

The degree-five threshold is necessary for this formula. For $s = 4$, $KG(4, 2)$ is three disjoint edges, and for $s = 3$ it has no edges. In either case, (B.17) does not connect every pair in W_u.

B.4.3 RECOVERING ROOT ADJACENCY

For $uv \in E(G)$, write

$$W_{u|v} = \big\{\{uv, e\} : e \in E_G(u) \setminus \{uv\}\big\} \subseteq W_u.$$

The exact cross-block formula is

$$E_Y(W_u, W_v) = \begin{cases} W_{u|v} \times W_{v|u}, & uv \in E(G), \\ \varnothing, & uv \notin E(G). \end{cases} \tag{B.18}$$

Consequently, if

$$T = E_W A E_W,$$

then

$$R_G = \mathrm{supp}_+(T) \circ (J - E_W) \tag{B.19}$$

is constant on ordered pairs of distinct E_W-classes, and its quotient is exactly the adjacency relation of G.

Equations (B.17) and (B.19), together with ambient J, are the generator data of the normalized two-line certificate (B.8) from G to $Y = L^2(G)$:

$$E_W = \mathsf{Inf}_{E_W}(I), \qquad J = \mathsf{Inf}_{E_W}(J), \qquad R_G = \mathsf{Inf}_{E_W}(A_G).$$

Corollary 20 (Root quotient exposed at depth two)**.** *For connected simple roots of minimum degree at least five, stable pair refinement on $L^2(G)$ determines:*

1. *one quotient class W_u for every root vertex u; and*
2. *root adjacency between those classes.*

An adjacency-preserving algebraic isomorphism of the two second-line stable configurations therefore gives an algebraic isomorphism between coherent root-side algebras containing the stable root algebras.

Proof. The relation E_W is the coherent expression (B.17), so an adjacency-preserving algebraic isomorphism transports it to its counterpart. The same is then true of $T = E_W A E_W$, its positive support, and the relation R_G in (B.19). Standard parabolic quotient calculus therefore turns the map into an algebraic isomorphism of the two E_W-quotient configurations that preserves the quotient of R_G. Formula (B.18) identifies that quotient with root adjacency, which proves both assertions. □

The following appendices use these recovered root classes to represent several root coordinates jointly in one deeper stable algebra.

B.4.4 EVOLUTION OF THE MINIMUM DEGREE

For every graph X and every edge uv of X,

$$d_{L(X)}(uv) = d_X(u) + d_X(v) - 2, \qquad \delta(L^r(G)) \geq 2^r(\delta(G) - 2) + 2, \tag{B.20}$$

where the second inequality follows by iterating the first.

Thus minimum degree at least three becomes minimum degree at least six after two line transformations. At every later depth, Theorem 19 recovers the vertices from two levels earlier by the same formula.

B.5 Endpoint relations with unequal class sizes

We next group shortest-path histories by their two root endpoints. The classes recovered above may have different sizes, so ordinary matrix multiplication would count some root vertices more often than others. Dividing by the class sizes removes these multiplicities. This gives the endpoint relations needed for the tuple construction for every connected graph of minimum degree at least five.

B.5.1 Normalization by class size

Let G be a finite connected graph with minimum degree at least five, put $Y = L^2(G)$, and let

$$E = \bigcup_{s \in V(G)} S_s \times S_s \tag{B.21}$$

be the Kneser center parabolic on Y, with $S_s = W_s$ in the notation of Appendix B.4. Here

$$|S_s| = q_s = \binom{d_G(s)}{2}, \tag{B.22}$$

which need not be independent of s. With the notation of Appendix A.2, its intrinsic normalizer is

$$H = D_E = (I \circ E^2)^\dagger, \qquad H(h,h) = q_s^{-1} \quad (h \in S_s). \tag{B.23}$$

Let $W \in \{0,1\}^{V(G) \times V(Y)}$ be the class-incidence matrix, with $W(s,h) = 1$ exactly when $h \in S_s$, and write

$$\iota(M) = W^\mathsf{T} M W. \tag{B.24}$$

Equation (B.24) describes the entries of the square matrix $\iota(M)$; coherence will follow from the normalized products below. The Kneser quotient gives

$$\iota(I) = E, \qquad \iota(J) = J, \qquad \iota(A_G) = R_G, \tag{B.25}$$

where R_G is the full cross-block relation whose quotient is root adjacency. Moreover,

$$\begin{aligned} \iota(M \circ N) &= \iota(M) \circ \iota(N), \\ \iota(M^\mathsf{T}) &= \iota(M)^\mathsf{T}, \\ \iota(MN) &= \iota(M) H \iota(N). \end{aligned} \tag{B.26}$$

The product identity follows from $WHW^\mathsf{T} = I$. Every $M \in \mathcal{A}(G)$ is a finite expression over I, J, A_G by Lemma 8. Applying (B.26) to this expression constructs $\iota(M)$ from E, J, R_G, H, so $\iota(M) \in \mathcal{A}(Y)$.

B.5.2 Recovering the two outer vertices of a wedge

At the two-level window

$$L^{2m}(G) = L^2\big(L^{2m-2}(G)\big), \qquad m \geq 1,$$

write $S_s^{(m)}$ for its Kneser center class, E_m for the center parabolic, and

$$H_m = \big(I \circ E_m^2\big)^\dagger. \tag{B.27}$$

Here and below, ι_m is the normalized inflation $\mathsf{Inf}_{E_m} : L^{2m-2}(G) \to L^{2m}(G)$.

The minimum-degree hypothesis persists under line transformation, so all these Kneser parabolics exist. With $A_m = A_{L^{2m}(G)}$, define the coherent zero–one arm relation

$$P_m = \mathrm{supp}_+(A_m E_m) \circ (J - E_m). \tag{B.28}$$

If

$$h = (s; a, b) \in L^{2m}(G),$$

write s for the common endpoint of its two edges and a, b for their other endpoints in $L^{2m-2}(G)$. Then the P_m-row is the indicator of $S_a^{(m)} \cup S_b^{(m)}$. Indeed, multiplication by E_m expands each adjacent vertex to its center block, while $J - E_m$ removes the center block $S_s^{(m)}$. Define

$$\mathcal{L}_m(M) = \big(2I + P_m H_m\big)\iota_m(M). \tag{B.29}$$

Lemma 21 (Irregular exact arm transfer)**.** *For every square matrix M on $L^{2m-2}(G)$,*

$$\mathcal{L}_m(M)(h, \cdot) = \iota_m\big(M(a, \cdot) + 2M(s, \cdot) + M(b, \cdot)\big). \tag{B.30}$$

Proof. Every row of $\iota_m(M)$ indexed by $S_a^{(m)}$ is the block inflation of $M(a, \cdot)$. Without normalization, multiplying by the arm relation P_m would add $q_a^{(m)}$ identical copies of this row. The diagonal H_m gives each such row weight $1/q_a^{(m)}$. Hence $P_m H_m \iota_m(M)$ supplies one copy of the a-row and one copy of the b-row. The term $2I\iota_m(M)$ supplies two copies of the center s-row. This is (B.30). □

B.5.3 A RECURRENCE FOR PATH ENDPOINTS

Define nested blocks by

$$F_x^{(0)} = \{x\}, \qquad F_x^{(m)} = \bigcup_{s \in F_x^{(m-1)}} S_s^{(m)} \quad (m \geq 1). \tag{B.31}$$

Their sizes $f_x^{(m)} = |F_x^{(m)}|$ may vary with x, but their equivalence relation is coherent:

$$E_F^{(1)} = E_1, \qquad E_F^{(m)} = \iota_m(E_F^{(m-1)}). \tag{B.32}$$

On $L^2(G)$, set

$$U_2 = P_1, \qquad V_2 = P_1 - E_1. \tag{B.33}$$

For $m \geq 2$, set

$$\begin{aligned} U_{2m} &= \mathcal{L}_m(V_{2m-2}) - \iota_m(U_{2m-2}), \\ V_{2m} &= U_{2m} - \iota_m(V_{2m-2}). \end{aligned} \tag{B.34}$$

Every matrix in (B.34) is coherent. The two recurrences cancel the contributions of internal path vertices, leaving only the endpoints. Specifically, if $h \in L^{2m}(G)$ is the history of a shortest path

$$x_0 - x_1 - \cdots - x_{2m},$$

then

$$\begin{aligned} U_{2m}(h, \cdot) &= \mathbf{1}_{F_{x_0}^{(m)}} + \mathbf{1}_{F_{x_{2m}}^{(m)}}, \\ V_{2m}(h, \cdot) &= \sum_{i=0}^{2m} (-1)^i \mathbf{1}_{F_{x_i}^{(m)}}. \end{aligned} \tag{B.35}$$

For $m = 1$, the P_1-row of the history centered at x_1 is $\mathbf{1}_{F^{(1)}_{x_0}} + \mathbf{1}_{F^{(1)}_{x_2}}$; subtracting its E_1-row gives the alternating identity. Assume (B.35) for $m - 1$. The two arms and center in the top two-level window are respectively the canonical $(2m - 2)$-histories at shifts $0, 2$, and 1 along the same geodesic. They are again shortest histories, so the induction hypothesis applies to all three. Record their coefficients as polynomials:

$$u_{2m-2}(z) = 1 + z^{2m-2}, \qquad v_{2m-2}(z) = \sum_{i=0}^{2m-2} (-z)^i.$$

The transfer (B.30) adds the two arms and twice the center, hence multiplies v_{2m-2} by $(1 + z)^2$; the inflated center correction shifts u_{2m-2} by z. Therefore

$$u_{2m} = (1 + z)^2 v_{2m-2} - z u_{2m-2} = 1 + z^{2m},$$

and

$$v_{2m} = u_{2m} - z v_{2m-2} = \sum_{i=0}^{2m} (-z)^i.$$

This proves (B.35) by induction from (B.33).

B.5.4 A NORMALIZED ENDPOINT GRAM

We now divide by the nested class sizes so that each common endpoint is counted once. Define

$$N_m = \big(I \circ (E_F^{(m)})^2\big)^\dagger. \tag{B.36}$$

On a vertex of $F_x^{(m)}$, the diagonal value of N_m is $1/f_x^{(m)}$. Whenever a line history h has a root footprint that is a path, write ∂h for the unordered set of the two endpoints of that path. If $\delta = 2m$, put

$$K_\delta = U_{2m} N_m U_{2m}^{\mathsf{T}}. \tag{B.37}$$

For selected shortest-path histories h, h', equation (B.35) gives

$$K_\delta(h, h') = |\partial h \cap \partial h'|. \tag{B.38}$$

Indeed, a common endpoint x contributes $f_x^{(m)} \cdot (1/f_x^{(m)}) = 1$, and distinct nested blocks are disjoint. If $\delta = 2m + 1 \geq 3$, let B_δ be the unsigned incidence matrix from $L^{2m}(G)$ to $L^{2m+1}(G)$, and put

$$K_\delta = B_\delta^{\mathsf{T}} V_{2m} N_m V_{2m}^{\mathsf{T}} B_\delta. \tag{B.39}$$

The two length-$2m$ parent paths of a selected odd-length history start one position apart. Their alternating coefficients cancel at every internal vertex:

$$(1 + z) \sum_{i=0}^{2m} (-z)^i = 1 + z^{2m+1}. \tag{B.40}$$

Thus (B.38) holds for odd lengths as well. Coherence follows by applying the incidence product to $V_{2m} N_m V_{2m}^{\mathsf{T}}$.

B.5.5 SELECTING SHORTEST-PATH HISTORIES

For a square matrix M, write $\Delta(M) = I \circ M$. Let B_j be unsigned incidence from $L^{j-1}(G)$ to $L^j(G)$, and set

$$C_\delta = B_1 B_2 \cdots B_\delta. \tag{B.41}$$

Lemma 22 (Distance matrices and history realization)**.** *For every positive integer j, the matrix*

$$D_j = [(I + A_G)^j > 0] - [(I + A_G)^{j-1} > 0] \tag{B.42}$$

has value one exactly on vertex pairs at distance j. It is coherent and is transported by every adjacency-preserving weak isomorphism. Every shortest path of length j supports a genuine vertex of $L^j(G)$ whose root footprint is exactly that path.

Proof. The positive support of $(I + A_G)^j$ records distance at most j, because identity steps pad any shorter walk. Subtraction therefore gives exact distance. Matrix powers and their value selectors are coherent and transported. A shortest path is P_{j+1}, and $L^j(P_{j+1}) \cong P_1$. Its unique final vertex gives the claimed history. □

Let R be any coherent zero–one relation supported at a single positive distance δ, and let $\overline{R} = [R + R^\mathsf{T} > 0]$ be its symmetric union. Define

$$Z_R^{\mathrm{path}} = \frac{1}{2}\Delta(C_\delta^\mathsf{T} \overline{R} C_\delta). \tag{B.43}$$

Here C_δ abbreviates the incidence chain, so the middle matrix in (B.43) is the iterated Gram construction from G to $L^\delta(G)$.

Lemma 23 (Shortest-path selection)**.** *The matrix in* (B.43) *is the zero–one diagonal selector of those $h \in L^\delta(G)$ whose root footprint is a shortest path with unordered endpoint pair in $\overline{R}$. It is coherent and is transported by the stable-prefix map ϕ_δ.*

Proof. Let $\sigma(h) \subseteq E(G)$ be the set of root edges in the recursive line history h. Adjacent level-s histories a, b satisfy

$$|\sigma(a) \setminus \sigma(b)| \le 1, \qquad |\sigma(b) \setminus \sigma(a)| \le 1. \tag{B.44}$$

For $s = 1$ this is immediate. If at the next level $a = \{c, d\}$ and $b = \{c, e\}$, then $\sigma(a) \setminus \sigma(b) \subseteq \sigma(d) \setminus \sigma(c)$; the other inclusion is symmetric. This proves (B.44) by induction. Applying it once more to the two parents of h shows that $\sigma(h)$ is connected and $|\sigma(h)| \le \delta$.

If the ancestry of h contains x, y with $(x, y) \in \overline{R}$, then $d_G(x, y) = \delta$. A connected footprint with at most δ edges containing such a pair must be exactly a simple shortest x–y path. Conversely, every shortest path has one level-δ history: recursively join its prefix and suffix histories. It is unique because histories supported on that path form $L^\delta(P_{\delta+1}) \cong P_1$.

For this unique history, the column of C_δ has endpoint weights one (the interior weights are binomial coefficients). Its only support pair at distance δ is the two ordered endpoint pairs. Hence

$$(C_\delta^\mathsf{T} \overline{R} C_\delta)_{hh} = \begin{cases} 2, & h \text{ is selected,} \\ 0, & \text{otherwise.} \end{cases} \tag{B.45}$$

This proves the selector formula. Equation (B.6) shows that $C_\delta^\mathsf{T} \overline{R} C_\delta$ belongs to $\mathcal{A}(L^\delta(G))$ and is preserved by ϕ_δ. The same is true after diagonal extraction and multiplication by $1/2$. □

B.5.6 GROUPING HISTORIES BY THEIR ENDPOINTS

Theorem 24 (Endpoint classes at every distance)**.** *For every such relation R at distance $\delta \geq 1$,*

$$E_R^{\mathrm{end}} = \begin{cases} Z_R^{\mathrm{path}} I Z_R^{\mathrm{path}}, & \delta = 1, \\ Z_R^{\mathrm{path}} [K_\delta = 2] Z_R^{\mathrm{path}}, & \delta \geq 2, \end{cases} \tag{B.46}$$

is a restricted coherent equivalence relation whose classes are exactly the shortest δ-edge histories with one fixed unordered endpoint pair in $\overline{R}$. Every unordered pair in $\overline{R}$ indexes a nonempty class.

Proof. By Lemma 23, every selected history has exactly two endpoints. For $\delta \geq 2$, equation (B.38) says that $K_\delta(h, h')$ is their intersection cardinality; hence the selected value two is equivalent to equality of the unordered endpoint sets. For $\delta = 1$, a selected line vertex is the unique history of its root edge, so the restricted identity gives the same conclusion. Lemma 22 supplies a shortest history for every pair in R, proving nonemptiness as well as the asserted classes and transitivity. The matrices P_m, H_m, N_m, and Z_R^{path} are obtained by coherent operations. Equations (B.26), (B.34), and (B.39), together with Theorem 12, show that the compatible weak isomorphisms preserve every step of the construction. □

B.6 TWO COHERENT COPIES IN FOUR LINE STEPS

The tuple construction needs several distinguishable copies of each vertex class. We show how four line transformations produce two nonempty, disjoint copies, labeled by a bit 0 or 1. These labels are called *tags*. Relations between the original classes are copied at the same time. Repeating the construction k times gives 2^k tags after $4k$ line transformations.

B.6.1 INTERNAL AND BOUNDARY EDGES

Let Q be a connected simple graph with $\delta(Q) \geq 5$, and let

$$E = \bigcup_{C \in \mathcal{C}} C \times C \tag{B.47}$$

be a coherent restricted parabolic on $D = \bigcup \mathcal{C}$ with $E \in \mathcal{A}(Q)$, and let B be the unsigned incidence matrix from Q to $L(Q)$. Put

$$\Delta(M) = I \circ M.$$

$$K = B^{\mathsf{T}} E B. \tag{B.48}$$

By (B.7), K belongs to $\mathcal{A}(L(Q))$.

For a line vertex $e = xy$, its diagonal value is

$$K(e, e) = E(x, x) + E(y, y) + 2E(x, y). \tag{B.49}$$

Thus the four possibilities are

endpoints of e	diagonal value
both in one E-class	4
both in D, in different classes	2
exactly one in D	1
neither in D	0

Let $Z_a = [\Delta(K) = a]$. Direct endpoint counting gives

$$\frac{1}{4} Z_4 K Z_4 = \bigcup_C E(Q[C])^2 \tag{B.50}$$

and

$$Z_1 K Z_1 = \bigcup_C \partial_D(C)^2, \tag{B.51}$$

where $\partial_D(C)$ is the set of Q-edges with exactly one endpoint in D, that endpoint belonging to C.

These formulas suggest using internal edges for one copy of C and edges leaving D for the other. Either set, however, can be empty. If $E = J$, then $Z_1 = 0$; if $E = I$, then $Z_4 = 0$. We therefore first replace each class by a set in $L^3(Q)$ that has both kinds of edge.

B.6.2 Ensuring both edge types after three transformations

Keep the graph Q with $\delta(Q) \geq 5$. On $L^2(Q)$, let W_v be the Kneser center block indexed by $v \in V(Q)$. Let

$$\mathbf{W} \in \{0,1\}^{V(Q) \times V(L^2(Q))}, \qquad \mathbf{W}(v,w) = \mathbf{1}[w \in W_v], \tag{B.52}$$

and put

$$E_2 = \mathbf{W}^{\mathsf{T}} \mathbf{W}, \qquad H_2 = (I \circ E_2^2)^{\dagger}. \tag{B.53}$$

Thus $H_2(w,w) = |W_v|^{-1}$ for $w \in W_v$.

Lemma 25 (Normalized Kneser block inflation)**.** *For $M \in \mathcal{A}(Q)$, the matrix*

$$\iota_2(M) = \mathbf{W}^{\mathsf{T}} M \mathbf{W} \in \mathrm{Mat}_{V(L^2(Q))}(\mathbb{R}) \tag{B.54}$$

belongs to $\mathcal{A}(L^2(Q))$ and is constant with value $M(u,v)$ on $W_u \times W_v$. For every $M, N \in \mathcal{A}(Q)$, it satisfies

$$\begin{aligned} \iota_2(M \circ N) &= \iota_2(M) \circ \iota_2(N), \\ \iota_2(M^{\mathsf{T}}) &= \iota_2(M)^{\mathsf{T}}, \\ \iota_2(MN) &= \iota_2(M) H_2 \iota_2(N). \end{aligned} \tag{B.55}$$

The matrix is constructed by applying the normalized inflation rules (B.8) *to a finite expression for M.*

If $B_3 \in \{0,1\}^{V(L^2(Q)) \times V(L^3(Q))}$ is unsigned incidence, put

$$K_H = B_3^{\mathsf{T}} E_2 B_3, \qquad Z_H = [\Delta(K_H) = 4].$$

The support of Z_H is the disjoint union of $H_v = E(L^2(Q)[W_v])$: a line vertex has diagonal value four exactly when both its parents lie in the same W_v. Then

$$\widehat{M}_H = \frac{1}{4} Z_H B_3^{\mathsf{T}} \iota_2(M) B_3 Z_H \in \mathrm{Mat}_{V(L^3(Q))}(\mathbb{R}) \tag{B.56}$$

is coherent and has value $M(u,v)$ on $H_u \times H_v$. In particular, $\widehat{I}_H = \bigcup_v H_v^2 =: E_H$. Both (B.54) *and* (B.56) *are preserved by the compatible weak isomorphisms of Theorem 12.*

Proof. The Kneser quotient of Appendix B.4 gives $\iota_2(I) = E_2$, $\iota_2(J) = J$, and $\iota_2(A_Q) = R_Q$, which records root adjacency between classes. The first two identities in (B.55) are entrywise. For the third, $\mathbf{W}H_2\mathbf{W}^{\mathsf{T}} = I_{V(Q)}$; hence it follows by direct multiplication. Since $\mathcal{A}(Q)$ is the coherent closure of I, J, A_Q, structural induction using (B.55) proves coherent membership of (B.54). Every vertex of H_u has exactly two parents in W_u, so the middle Gram in (B.56) has value $4M(u, v)$ on $H_u \times H_v$. This proves the value formula and coherent membership. The root-adjacency relation, E_2, H_2, Z_H, and both incidence products are preserved by the compatible weak isomorphisms, by (B.5). □

The sets H_v are pairwise disjoint and their union H is a proper restricted domain in $L^3(Q)$. Let E be as in (B.47). By Lemma 25, its three-line block lift is

$$\widehat{E}_H = \bigcup_{C \in \mathcal{C}} K_C^2, \qquad K_C = \bigcup_{v \in C} H_v. \tag{B.57}$$

Explicitly,

$$\widehat{E}_H = \frac{1}{4} Z_H B_3^{\mathsf{T}} \iota_2(E) B_3 Z_H. \tag{B.58}$$

This is (B.56) with $M = E$.

For every $M \in \mathcal{A}(Q)$ supported on the domain of E and constant on its class rectangles, the same lift satisfies

$$\widehat{M}_H[K_C, K_{C'}] = m_{C,C'} J. \tag{B.59}$$

Lemma 26 (The padding has both kinds of edge)**.** *For every nonempty $C \in \mathcal{C}$:*

1. *$L^3(Q)[K_C]$ has an edge; and*
2. *an edge of $L^3(Q)$ has exactly one endpoint in the full padded domain $\widehat{D} = \bigcup_C K_C$, with that endpoint in K_C.*

Proof. Choose $v \in C$. Take four distinct root edges a, b, c, d incident with v. In $L^2(Q)$, put

$$p = \{a, b\}, \qquad q = \{a, c\}, \qquad r = \{a, d\}.$$

Then $\eta_0 = \{p, q\}$ and $\eta_1 = \{p, r\}$ lie in H_v, and they are adjacent in $L^3(Q)$ because they share the $L^2(Q)$-vertex p. This proves the first assertion.

For the second, write $a = vv_1$, take distinct $b, c \neq a$ at v, and choose a root edge $d_1 \neq a$ at v_1. Put

$$p = \{a, b\}, \qquad q = \{a, c\}, \qquad r_1 = \{a, d_1\}.$$

The history $\eta = \{p, q\}$ lies in H_v. The history $\zeta = \{p, r_1\}$ is adjacent to η, but its two wedge parents have different root centers v and v_1. Hence ζ lies outside the universal domain H, and therefore outside $\widehat{D}$, regardless of which vertices belong to the original domain D. The edge $\eta\zeta$ is the required boundary edge. □

The first construction needs four edges at v; the second needs three at v and two at v_1. Minimum degree five guarantees these choices and the Kneser classes used to define the sets H_v.

B.6.3 THE BINARY COPYING CONSTRUCTION

Let B_4 be the unsigned incidence matrix from $L^3(Q)$ to $L^4(Q)$, and put

$$K_E = B_4^{\mathsf{T}} \widehat{E}_H B_4, \qquad Z_4 = [\Delta(K_E) = 4], \qquad Z_1 = [\Delta(K_E) = 1]. \tag{B.60}$$

The incidence identity (B.7) and diagonal selection show that all three matrices belong to $\mathcal{A}(L^4(Q))$.

For each old class C, define two sets of $L^4(Q)$-vertices:

$$\begin{aligned} T_C^0 &= E(L^3(Q)[K_C]), \\ T_C^1 &= \{xy \in E(L^3(Q)) : |\{x, y\} \cap \widehat{D}| = 1, \text{ and the inside endpoint lies in } K_C\}. \end{aligned} \tag{B.61}$$

By (B.49), Z_4 and Z_1 are exactly the diagonal selectors of the two respective unions. Lemma 26 says that every T_C^i is nonempty. They are pairwise disjoint because a line vertex has one unambiguous endpoint status and, in the boundary case, one unique inside endpoint.

Theorem 27 (Four-step tag doubling)**.** *The matrix*

$$E_{\mathrm{tag}} = \frac{1}{4} Z_4 K_E Z_4 + Z_1 K_E Z_1 \tag{B.62}$$

is the coherent restricted parabolic

$$E_{\mathrm{tag}} = \bigcup_{C \in \mathcal{C}} \big((T_C^0)^2 \cup (T_C^1)^2 \big). \tag{B.63}$$

Thus its classes are indexed by

$$(C, i) \in \mathcal{C} \times \{0, 1\}. \tag{B.64}$$

Proof. An internal tag vertex has two endpoints in one K_C; a boundary tag vertex has one endpoint in one K_C and one endpoint outside $\widehat{D}$. For two internal vertices, the Gram K_E is four if their classes agree and zero otherwise. For two boundary vertices it is one if their inside-endpoint classes agree and zero otherwise. This gives (B.62)–(B.63). Coherent membership follows from forward Gram, diagonal value selection, restriction, scaling, and addition on disjoint supports. □

Theorem 28 (Simultaneous lift of block matrices)**.** *For M as in* (B.59)*, put*

$$K_M = B_4^{\mathsf{T}} \widehat{M}_H B_4 \tag{B.65}$$

and define

$$\begin{aligned} \mathcal{T}(M) = \frac{1}{4} Z_4 K_M Z_4 + \frac{1}{2} Z_4 K_M Z_1 \\ + \frac{1}{2} Z_1 K_M Z_4 + Z_1 K_M Z_1. \end{aligned} \tag{B.66}$$

All matrices in (B.66) *belong to $\mathcal{A}(L^4(Q))$.*

Then, for all tags $i, j \in \{0, 1\}$,

$$\mathcal{T}(M)[T_C^i, T_{C'}^j] = m_{C,C'} J. \tag{B.67}$$

Proof. A tag-i line vertex has $2 - i$ endpoints in the padded domain, all in its one class K_C. Since $\widehat{M}_H$ vanishes whenever an endpoint lies outside that domain, (B.65) has the constant value

$$(2-i)(2-j)m_{C,C'} \tag{B.68}$$

on $T_C^i \times T_{C'}^j$. The four coefficients in (B.66) are the reciprocals of these factors. □

The diagonal selectors Z_4 and Z_1 distinguish the two tags. Equation (B.66) preserves each original matrix entry between copies with either equal or different tags.

Line transformation preserves the minimum-degree-five hypothesis, so this binary construction can be repeated as coordinates are added. Each use lifts all required matrices together. Appendix B.7 applies it to the identity partition of the whole current graph, which keeps previously aligned relations consistent even when their domains overlap.

B.7 One enumerating tuple suffices

We construct a domain representing a subset of vertex tuples, including one prescribed tuple. The endpoint and copying constructions allow us to add one coordinate at a time. All coordinates refer to the same copy of the root graph, and one weak isomorphism transports the resulting matrices. When the prescribed tuple lists every root vertex, these matrices determine a graph isomorphism. We therefore need not represent every possible tuple.

Throughout, $D_E = (I \circ E^2)^\dagger$ and ι_E have the meanings fixed in Appendix A.2; every quotient multiplication uses the normalized identity (A.1). Each reciprocal D_E, including D_F, D_A, D_B below, is constructed by (B.9), and each inflation ι_E by (B.8).

B.7.1 Representing odd histories at an even depth

Let A be a finite connected graph with minimum degree at least five, and put $T = L^3(A)$. Let

$$E_S = \bigcup_{h \in E(A)} S_h^2, \qquad E_H = \bigcup_{t \in V(A)} H_t^2 \tag{B.69}$$

be the edge-indexed Kneser and vertex-indexed internal-history parabolics. In the matrix notation above,

$$E_S = \mathrm{ev}\big(\mathsf{Inf}_{E_S}(I_{L(A)})\big) : L(A) \longrightarrow L^3(A)$$

is the identity output of the two-line Kneser certificate over $L(A)$. For the other parabolic, let E_2 be the two-line Kneser parabolic over A, and put

$$K_H = \mathrm{ev}\big(\mathsf{Gram}_2\big(\mathsf{Inf}_{E_2}(I_A)\big)\big), \qquad Z_H = [\Delta(K_H) = 4], \qquad E_H = \frac{1}{4} Z_H K_H Z_H.$$

This is the supported three-line identity lift of (B.56); its inflation goes from A to $L^2(A)$ and its Gram goes from $L^2(A)$ to $L^3(A)$. Thus both parabolics in (B.69) are obtained by the previously established incidence and inflation operations. Their classes intersect in

$$|S_h \cap H_t| = \mathbf{1}[t \in h] \binom{d_A(t) - 1}{2}. \tag{B.70}$$

Put

$$C_{SH} = I \circ (E_S E_H), \qquad I_{SH} = E_S C_{SH}^\dagger E_H. \tag{B.71}$$

Equation (B.9) constructs the diagonal $C_{SH}^\dagger$ within the same stable algebra. Multiplication by E_S and E_H makes the result constant on each $S_h \times H_t$.

The middle reciprocal cancels the varying value in (B.70), so

$$I_{SH}[S_h, H_t] = \mathbf{1}[t \in h] J. \tag{B.72}$$

The selector above satisfies $Z_H = I \circ E_H$. Put $Z_\circ = I - Z_H$, and

$$S_h^\circ = S_h \setminus \mathrm{dom}(E_H). \tag{B.73}$$

If $h = uv$, then

$$|S_h^\circ| = (d_A(u) - 1)(d_A(v) - 1) > 0, \tag{B.74}$$

and $T[S_h^\circ]$ is the rectangular rook graph

$$K_{d_A(u)-1} \,\square\, K_{d_A(v)-1}, \tag{B.75}$$

of degree $d_A(u) + d_A(v) - 4$. Thus its internal-edge set is nonempty. The restricted cross matrix

$$I_{\circ H} = Z_\circ I_{SH} \tag{B.76}$$

still has the block values in (B.72), and its row and column domains are now disjoint.

For $V \in \mathcal{A}(A)$, let $\widehat{V}_H$ be its normalized three-line block lift to the H_t-classes. Appendix B.6 constructs this lift by applying Inf_{E_2}, then Gram_2, and restricting to the required support. Put

$$\mathcal{R}_\circ(V) = I_{\circ H} D_{E_H} \widehat{V}_H. \tag{B.77}$$

For $h = uv$,

$$\mathcal{R}_\circ(V)[S_h^\circ, H_t] = (V(u,t) + V(v,t))J = (B_A^{\mathsf{T}} V)(h,t)J. \tag{B.78}$$

The factor D_{E_H} cancels $|H_u|$ and $|H_v|$ separately; no equality of those sizes is used.

B.7.2 One internal lift and canonical containment

Apply one further line transformation and define

$$I_S(h) = E(T[S_h^\circ]), \qquad I_H(t) = E(T[H_t]). \tag{B.79}$$

We can construct the corresponding parabolics by the same matrix operations. Put $P_S^\circ = Z_\circ E_S Z_\circ$, $P_H = E_H$, and, for $P \in \{P_S^\circ, P_H\}$, define

$$K_P = B_T^{\mathsf{T}} P B_T = \mathrm{ev}(\mathsf{Gram}_3(P)), \qquad Z_P = [\Delta(K_P) = 4], \qquad E_{I(P)} = \frac{1}{4} Z_P K_P Z_P. \tag{B.80}$$

The internal-edge calculation (B.50) of Appendix B.6 gives $E_{I(P_S^\circ)} = E_{I_S}$ and $E_{I(P_H)} = E_{I_H}$, at $L^4(A)$. Set $Z_{I_S} = I \circ E_{I_S}$ and $Z_{I_H} = I \circ E_{I_H}$. Both parabolics therefore belong to the stable algebra of $L^4(A)$. The incidence Gram takes matrices from $L^3(A)$ to $L^4(A)$; selection, restriction, and scaling take place within the latter algebra. Applying the same operation to $\mathcal{R}_\circ(V)$ gives

$$\mathcal{R}^+(V) = \frac{1}{4} Z_{I_S} B_T^{\mathsf{T}} \mathcal{R}_\circ(V) B_T Z_{I_H} = \frac{1}{4} Z_{I_S} \,\mathrm{ev}(\mathsf{Gram}_3(\mathcal{R}_\circ(V))) Z_{I_H}. \tag{B.81}$$

This matrix retains the block values in (B.78). The factor four comes from the two parents of each selected line vertex, so it does not depend on the class sizes.

Let $W_t^{[1]} \subseteq V(L^2(A))$ denote the Kneser center class over $t \in V(A)$, and let $W_s^{[2]} \subseteq V(L^4(A))$ denote the center class in the next two-line Kneser window over $s \in V(L^2(A))$. Let $E_F^{[2]}$ be the canonical two-successive-Kneser parabolic on $L^4(A)$, with classes

$$F_t^{[2]} = \bigcup_{s \in W_t^{[1]}} W_s^{[2]}. \tag{B.82}$$

The endpoint recurrence constructs this parabolic by two successive inflations. Choose a finite expression $u_{E_F^{[1]}}$ with value $E_F^{[1]}$. Then $E_F^{[1]} = \mathrm{ev}(\mathsf{Inf}_{E_1}(I_A))$ from A to $L^2(A)$, followed by $E_F^{[2]} = \mathrm{ev}(\mathsf{Inf}_{E_2}(u_{E_F^{[1]}}))$ from $L^2(A)$ to $L^4(A)$.

The required containment is

$$I_H(t) \subseteq F_t^{[2]} \tag{B.83}$$

for every t: two adjacent internal $W_t^{[1]}$-edges share a center $s \in W_t^{[1]}$, so their four-line history lies in $W_s^{[2]}$. Different indices remain disjoint. Hence

$$J_{I_H,F} = E_{I_H} D_{E_{I_H}} E_F^{[2]} \tag{B.84}$$

is complete from $I_H(t)$ to $F_t^{[2]}$, even though $|I_H(t)|$ varies.

For a coherent relation supported at odd distance $\delta = 2r+1$, take $A = L^{2r}(G)$. Let Z_R be its diagonal shortest-history selector on $L(A)$ and put

$$Z_{R,\circ} = Z_\circ\big(I \circ \iota_{E_S}(Z_R)\big)Z_\circ, \qquad Q_R^{\mathrm{odd}} = Z_{I_S}\big[I \circ (B_T^{\mathsf{T}} Z_{R,\circ} B_T) = 2\big]Z_{I_S}. \tag{B.85}$$

An internal edge of S_h° has both parents selected exactly when h is selected. Hence this is precisely the required diagonal domain. Each selected S_h° contains an internal edge by the padding lemma, so no represented endpoint pair is lost.

The resulting endpoint row is

$$R_\delta^{\mathrm{can}} = Q_R^{\mathrm{odd}} \mathcal{R}^+(V) D_{E_{I_H}} E_F^{[2]}. \tag{B.86}$$

Here $\delta = 2r+1 \geq 3$ and $V = V_{2r}$ is the alternating-boundary matrix of (B.34) in Appendix B.5. For $\delta = 1$, set $F_x^{(0)} = \{x\}$, use $V = I$, and omit the empty recurrence. In either case a selected odd history has value one on its two endpoint blocks and zero on all other root blocks.

The nested block $F_x^{(r)}$ inflates under (B.82) to $F_x^{(r+2)}$ by definition. Thus odd- and even-distance endpoint rows refer to the same root classes at an even line depth. Further pairs of line transformations preserve the selected domain and its endpoint row under normalized inflation.

B.7.3 A PARTIAL PRODUCT AT ONE DISTANCE

Let W be the current connected graph, with $\delta(W) \geq 5$. Suppose its stable algebra contains two parabolics with disjoint domains,

$$E_A = \bigcup_{a \in U} A_a^2, \qquad E_B = \bigcup_{b \in V} B_b^2.$$

Fix one pair $(a_*, b_*) \in U \times V$ that we want to represent. Choose $p \in A_{a_*}$, $q \in B_{b_*}$, and let $\delta = d_W(p,q) > 0$. Write $Z_A = I \circ E_A$ and $Z_B = I \circ E_B$. The coherent relation

$$S = Z_A D_\delta Z_B \tag{B.87}$$

is supported at distance δ, lies in the ordered cross domain, and contains (p,q). It need not be a single stable color relation.

Apply Theorem 24 to $S \cup S^{\mathsf{T}}$. For odd distance, use the preceding odd-to-even construction. At one finite even depth we obtain a coherent diagonal domain Q, a root-indexed parabolic $E_F = \bigcup_{x \in V(W)} F_x^2$, and a coherent matrix R with $R = QR$ whose selected rows satisfy

$$R(h,\cdot) = \mathbf{1}_{F_x} + \mathbf{1}_{F_y}, \qquad (x,y) \in S. \tag{B.88}$$

Every pair in S has at least one selected history. The disjoint domains distinguish the coordinate in A from the coordinate in B, even though an endpoint history is unordered. Carry all matrices

retained from W through the same normalized inflation ι_F; write hats for their images and put $D_F = (I \circ E_F^2)^\dagger$.

Define

$$E_P = Q\big[RD_F(\widehat{E}_A + \widehat{E}_B)D_F R^{\mathsf{T}} = 2\big]Q, \tag{B.89}$$

$$J_A = RD_F\widehat{E}_A, \qquad J_B = RD_F\widehat{E}_B. \tag{B.90}$$

For selected histories h, h', normalization gives

$$\big(RD_F(\widehat{E}_A + \widehat{E}_B)D_F R^{\mathsf{T}}\big)(h, h') = \mathbf{1}[a(h) = a(h')] + \mathbf{1}[b(h) = b(h')].$$

Thus E_P is a parabolic with a nonempty class P_{ab} for precisely those class pairs (a, b) met by S. In particular, $P_{a_* b_*} \neq \varnothing$. The two matrices in (B.89) are complete incidences to the inflated classes $\widehat{A}_a = \bigcup_{x \in A_a} F_x$ and $\widehat{B}_b = \bigcup_{y \in B_b} F_y$, respectively. The factor D_F assigns unit total mass to each endpoint block before spreading it over its entire class. No uniform class-size assumption is used.

The construction represents the class pairs reached at the chosen distance, including (a_*, b_*). No endpoint construction at another distance is needed.

B.7.4 One root copy for all coordinates

Fix a connected root graph X with $\delta(X) \geq 5$, and fix a finite source tuple $\mathbf{x} = (x_1, \ldots, x_a)$. At stage $k \leq a$ we retain the following matrices in the stable algebra of one iterated line graph:

1. a tuple parabolic $E_A = \bigcup_{\mathbf{u} \in U_k} A^2_{\mathbf{u}}$, where $U_k \subseteq V(X)^k$ contains $(x_1, \ldots, x_k)$;
2. a root parabolic $E_B = \bigcup_{v \in V(X)} B^2_v$, with domain disjoint from the tuple domain;
3. complete coordinate incidences $J_i[A_{\mathbf{u}}, B_v] = \mathbf{1}[u_i = v]J$, for $1 \leq i \leq k$;
4. the root matrix lifts $\widehat{M}[B_u, B_v] = M(u, v)J$, for $M = I, A_X$.

Every coordinate is related to the same root copy B.

For the base case, apply the binary copying construction of Appendix B.6 directly to X with its identity parabolic. At depth four, use one copy for A_v, the other for B_v, the cross-copy lift of I for J_1, and the within-root-copy lifts of I, A_X. This represents every one-coordinate tuple, including x_1.

For the induction step, apply (B.89) to the current tuple domain and the common root domain. Choose the distance between one representative of $(x_1, \ldots, x_k)$ and one representative of x_{k+1}, and use S from (B.87). The partial product contains the desired extended tuple. At its endpoint depth, let $\widetilde{J}_i$ denote the common inflation of the old coordinate incidence and let $D_A = (I \circ \widehat{E}_A^2)^\dagger$. Set

$$K_i = J_A D_A \widetilde{J}_i \quad (1 \leq i \leq k), \qquad K_{k+1} = J_B. \tag{B.91}$$

On $P_{\mathbf{u},v} \times \widehat{B}_w$, the first product has value $\mathbf{1}[u_i = w]$: summing over $\widehat{A}_{\mathbf{u}}$ is canceled by its reciprocal size. The last incidence has value $\mathbf{1}[v = w]$.

The new tuple domain need not yet be disjoint from the inflated root domain. Apply the binary copying operation to the *whole current graph*, using its identity parabolic. If Λ_{ij}, $i, j \in \{0, 1\}$, denotes the jointly normalized lift to these two copies, retain

$$\begin{aligned} E_A^{\mathrm{new}} &= \Lambda_{00}(E_P), \quad & E_B^{\mathrm{new}} &= \Lambda_{11}(\widehat{E}_B), \\ J_i^{\mathrm{new}} &= \Lambda_{01}(K_i), \quad & \widehat{M}^{\mathrm{new}} &= \Lambda_{11}(\widetilde{M}) \quad (M = I, A_X). \end{aligned} \tag{B.92}$$

Here $\widetilde{M}$ is the common inflation of the retained root matrix lift. All these matrices pass through one copying operation, so their class indices stay aligned. The two domains are disjoint and the original matrix entries are unchanged. This proves the induction step. Repeated root coordinates are allowed: the tuple and root domains remain disjoint even when their root labels coincide.

Theorem 29 (A finite coordinate construction for one source tuple)**.** *For every connected X with $\delta(X) \geq 5$ and every nonempty finite tuple $\mathbf{x}$ of its vertices, the preceding finite construction produces the stated partial tuple parabolic, common root copy, coordinate incidences, and equality and adjacency lifts. Its tuple domain contains a class representing $\mathbf{x}$.*

Proof. The base case and the induction step preserve all four conditions. There are exactly $a-1$ induction steps. Each chooses a positive distance d in the current connected graph. The endpoint construction followed by binary copying uses $d+4$ line steps for even d, and $d+7$ for odd d. Thus the construction ends at finite depth. □

B.7.5 JOINT TRANSPORT OF THE FINITE CONSTRUCTION

Lemma 30 (Joint transport of the partial coordinate construction)**.** *Let X, Y be connected with minimum degree at least five and agree under 3-WL at every line depth. The construction for a prescribed source tuple can be performed jointly at one finite common depth r, with every matrix transported to the corresponding matrix on Y.*

Proof. Equivalence at every line depth gives an adjacency-preserving weak isomorphism between the stable pair algebras at each depth. Theorem 12 makes their restrictions compatible through each finite prefix. Apply these maps in the order of the construction. Coherent operations are transported within a level; Equation (B.5) transports incidence products between levels; and normalized Kneser inflation transports earlier matrices to the common even depth of the endpoint relations. The odd-to-even construction and the binary copying operation use these same operations, so all their outputs are transported jointly.

At each induction step, use on Y the same distance δ and the transported tuple and root domain selectors in (B.87). The resulting relation is coherent and supported at distance δ between the corresponding domains. The endpoint calculation and (B.89) therefore produce a partial product on Y, with the same projection and coordinate identities. Its tuple labels need not match those of X, nor need every possible tuple be represented. Only the nonempty classes in the resulting domain are needed.

The selected representatives determine only an integer distance. The construction uses the entire typed relation at that distance and adds no predicate selecting a named vertex. Equation (B.42) gives a shared finite matrix expression for it. All remaining normalizers and selectors are explicit coherent expressions, so every level change in the finite construction commutes with the compatible incidence and inflation maps. □

B.7.6 COORDINATE RELATIONS AND DIRECT EXTRACTION

Lemma 31 (Joint coordinate relations)**.** *Let X, Y be connected with minimum degree at least five and agree under 3-WL at every line depth. For a prescribed tuple $\mathbf{x} \in V(X)^a$, $a \geq 1$, there are a finite common depth r, a set $U \subseteq V(X)^a$ containing $\mathbf{x}$, and pairwise disjoint nonempty sets $P_{\mathbf{u}} \subseteq V(L^r(X))$, one for each $\mathbf{u} \in U$. Let Z be the diagonal indicator of their union. The stable algebra contains Z and matrices $C_{ij}(I), C_{ij}(A_X)$, for $1 \leq i, j \leq a$, satisfying*

$$C_{ij}(M)(p,q) = M(u_i, v_j) \quad \text{if } p \in P_{\mathbf{u}},\ q \in P_{\mathbf{v}}. \tag{B.93}$$

Under the adjacency-preserving weak isomorphism at depth r, the images satisfy the same coordinate identities for a partial tuple construction on Y.

Proof. Use the partial tuple construction of Theorem 29 and its joint transport from Lemma 30. At the final depth, write E_P for the partial tuple parabolic, $Z = I \circ E_P$, and $D_B = (I \circ E_B^2)^\dagger$. Define

$$C_{ij}(M) = J_i D_B \widehat{M} D_B J_j^{\mathsf{T}}, \qquad M = I, A_X. \tag{B.94}$$

On $P_{\mathbf{u}} \times P_{\mathbf{v}}$, this has value $M(u_i, v_j)$. Each of the two root-class sums is canceled by its normalizer. Lemma 30 transports all these matrices jointly, giving the same identities on Y. □

For an enumeration $\mathbf{x} = (x_1, \ldots, x_n)$ of $V(X)$, apply Lemma 31 and define

$$F_X = Z \circ \bigcirc_{i<j} \big(J - C_{ij}(I)\big) \circ \bigcirc_{\substack{i<j \\ x_i \sim_X x_j}} C_{ij}(A_X) \circ \bigcirc_{\substack{i<j \\ x_i \not\sim_X x_j}} \big(J - C_{ij}(A_X)\big). \tag{B.95}$$

Here $\bigcirc$ denotes entrywise product. Within the tuple domain, this diagonal matrix selects tuples with distinct entries and exactly the adjacencies of the fixed source enumeration.

Lemma 32 (Direct extraction of an isomorphism)**.** *Let X, Y have the same positive order n. Suppose one weak isomorphism transports a partial n-coordinate construction whose source domain contains an enumeration of $V(X)$. Then $X \cong Y$.*

Proof. Let $\mathbf{x}$ be that enumeration, and form F_X by (B.95). All factors belong to the final stable algebra. For every $p \in P_{\mathbf{x}}$, the coordinate identities give $F_X(p,p) = 1$, so $F_X \neq 0$. Apply the weak isomorphism to this single expression, keeping the source edge and nonedge index sets fixed. Its image F_Y is nonzero, diagonal, and supported on the target tuple domain. Choose a nonzero diagonal entry, and let $\mathbf{y}$ label its class. Every factor is zero or one, so

$$y_i \neq y_j \quad (i < j), \qquad x_i \sim_X x_j \iff y_i \sim_Y y_j \quad (i < j).$$

The n distinct target coordinates exhaust $V(Y)$. Therefore $x_i \mapsto y_i$ is a graph isomorphism. □

Theorem 33 (High-degree graph identification)**.** *Let X, Y be connected graphs of minimum degree at least five. If $L^r(X) \equiv_{3\text{-}WL} L^r(Y)$ for every $r \geq 0$, then $X \cong Y$. Consequently, every non-isomorphic such pair separates at a finite line depth.*

Proof. Root equivalence gives equal orders. Apply the finite coordinate construction to an enumeration of $V(X)$, with joint transport supplied by Lemma 30. Lemma 32 gives the required vertex bijection. The finite-depth conclusion is the contrapositive. □

B.8 Connected completeness of iterated line refinement

We extend the high-degree identification theorem from Appendix B.7 to all finite connected simple graphs. Paths, cycles, and the claw are identified at depth zero. Every other connected graph eventually has minimum degree at least five, where the preceding theorem applies.

B.8.1 Statement

Theorem 34 (Connected all-depth identification)**.** *Let G, H be finite connected simple graphs. Suppose*

$$L^r(G) \equiv_{3\text{-}WL} L^r(H) \qquad \text{for every } r \geq 0. \tag{B.96}$$

Then

$$G \cong H. \tag{B.97}$$

Equivalently, every pair of non-isomorphic connected graphs is separated by stable pair refinement at some finite line depth. The proof gives no useful uniform bound on that depth.

B.8.2 The elementary nonprolific cases

A connected graph is called prolific when it is not a path, a cycle, or the claw $K_{1,3}$. First suppose one of G, H, say G, is nonprolific.

Already the $r = 0$ instance of (B.96) preserves the vertex number and degree multiset. This follows at the first diagonal refinement: the diagonal color of a vertex records its adjacency row sum, and the stable graph signature records the multiplicity of every such color.

- If $G = P_n$, its degree multiset is (0) for $n = 1$, $(1, 1)$ for $n = 2$, and two ones with $n - 2$ twos for $n \geq 3$. A connected graph with that degree multiset is the same path.
- If $G = C_n$, every vertex has degree two. A finite connected 2-regular graph is the same cycle.
- If $G = K_{1,3}$, its degree multiset is $(3, 1, 1, 1)$, which identifies the claw among connected graphs.

Thus $H \cong G$ in every nonprolific case. In particular, the Whitney pair $K_3, K_{1,3}$ is not a counterexample: it collides after one line transformation, but it is already separated at level zero by order and degree.

It remains to consider two prolific graphs.

B.8.3 Reaching minimum degree five in both graphs

Hartke and Higgins prove that for every connected prolific graph X there is an index $c(X)$ such that

$$\delta(L^{k+1}(X)) = 2\delta(L^k(X)) - 2 \qquad (k \geq c(X)), \tag{B.98}$$

with $\delta(L^{c(X)}(X)) \geq 3$ (Hartke & Higgins, 2003). We use this published growth theorem without reproving it.

Writing $a_k = \delta(L^k(X)) - 2$, equation (B.98) gives

$$a_{k+1} = 2a_k, \tag{B.99}$$

so the minimum degree tends to infinity. Choose one common s large enough that

$$X := L^s(G), \qquad Y := L^s(H), \qquad \delta(X), \delta(Y) \geq 5. \tag{B.100}$$

Both X and Y remain connected and simple. Equation (B.96) shifts to

$$L^t(X) \equiv_{3\text{-WL}} L^t(Y) \qquad (t \geq 0). \tag{B.101}$$

B.8.4 One finite coordinate construction identifies the shifted graphs

The depth-zero instance of (B.101) gives $N = |V(X)| = |V(Y)|$. Apply Theorem 33 to the all-depth equivalence (B.101). It follows that

$$X \cong Y. \tag{B.102}$$

Lemma 32 obtains the isomorphism from one enumeration of $V(X)$.

B.8.5 Whitney descent through the finite prefix

Whitney's line-graph isomorphism theorem says that connected simple graphs with isomorphic line graphs are isomorphic, with the unique nonisomorphic pair $K_3, K_{1,3}$ (Whitney, 1932).

Starting from (B.102), descend one line level at a time. Suppose at some $1 \leq j \leq s$ that

$$L^j(G) \cong L^j(H). \tag{B.103}$$

Whitney gives

$$L^{j-1}(G) \cong L^{j-1}(H) \tag{B.104}$$

unless the two roots in (B.104) are K_3 and $K_{1,3}$ in some order. But in that exceptional case their common line graph is K_3, and every subsequent line graph remains K_3. It would follow that $X = L^s(G) \cong K_3$, contradicting $\delta(X) \geq 5$.

The exception is therefore impossible at every descent step. Induction from $j = s$ to $j = 0$ proves (B.97).

The coordinate construction uses compatible weak isomorphisms obtained from one final depth, as established in Lemma 30. We next treat disconnected graphs.

B.9 Arbitrary graphs and sufficiency of the first and last depths

For disconnected graphs, line transformation deletes isolated vertices and maps both triangle and claw components to triangles. We first identify the component counts needed for one-step backward transfer. We then prove Theorem 3 and extend graph identification to arbitrary finite simple graphs.

All graphs below are finite, undirected, and simple. The empty graph on zero vertices is allowed. The notation $X \equiv_{3\text{-WL}} Y$ means equivalence under the paper's stable joint ordered-pair refinement defined in Section 2.

B.9.1 Main statements

For a graph G, write

$$i(G) = \#\{\text{components of } G \text{ isomorphic to } K_1\},$$

$$\kappa(G) = \#\{\text{components of } G \text{ isomorphic to } K_{1,3}\},$$
$$t(G) = \#\{\text{components of } G \text{ isomorphic to } K_3\}.$$

For $m \geq 1$, also write

$$p_m(G) = \#\{\text{components of } G \text{ isomorphic to } P_m\}.$$

Theorem 35 (Exact all-graph one-step criterion)**.** *Let G, H be arbitrary finite simple graphs and suppose*

$$L(G) \equiv_{3\text{-WL}} L(H). \tag{B.105}$$

Then the following are equivalent:

1. $G \equiv_{3\text{-WL}} H$*;*
2. $i(G) = i(H)$ *and* $\kappa(G) = \kappa(H)$*;*
3. $i(G) = i(H)$ *and* $t(G) = t(H)$*;*
4. *the root degree multisets agree:* $\mathrm{Deg}(G) = \mathrm{Deg}(H)$.

Thus one line step loses exactly two pieces of root information at the 3-WL level:

- *isolated root vertices, which disappear; and*
- *for each line component K_3, whether its connected root was K_3 or $K_{1,3}$.*

The equivalence classes of all other connected root components under 3-WL are determined by their line graphs.

Theorem 36 (All-simple-graph completeness)**.** *Let G, H be arbitrary finite simple graphs. If*

$$L^r(G) \equiv_{3\text{-WL}} L^r(H) \qquad \text{for every } r \geq 0, \tag{B.106}$$

then

$$G \cong H. \tag{B.107}$$

Isolated vertices, triangle components, and claw components are all allowed. At depth zero, (B.106) *determines the multiplicity of each nonprolific component. We match the remaining components using their 3-WL signatures at successive line depths.*

B.9.2 MATCHING CONNECTED COMPONENTS BY STABLE PAIR REFINEMENT

Equivalent graphs have equally many connected components, which can be paired so that corresponding components are themselves equivalent.

Lemma 37 (Component matching (Yang, 2026)). *For finite graphs X, Y, $X \equiv_{\text{3-WL}} Y$ if and only if there is a bijection $\beta : \mathrm{Comp}(X) \to \mathrm{Comp}(Y)$ with $X[C] \equiv_{\text{3-WL}} Y[\beta(C)]$ for every $C \in \mathrm{Comp}(X)$. Components include isolated vertices, and two empty graphs have the empty bijection.*

B.9.3 PROOF OF THE EXACT ONE-STEP THEOREM

Every connected component of $L(G)$ comes from exactly one connected component of G that contains an edge; isolated root vertices contribute nothing. Apply Lemma 37 to (B.105), obtaining a componentwise pairing of $L(G)$ and $L(H)$.

Consider one paired line-component pair

$$L(C) \equiv_{\text{3-WL}} L(D), \tag{B.108}$$

where C, D are connected root components with at least one edge.

If one line component is K_3, the other is isomorphic to K_3: 3-WL equivalence preserves its order three and degree multiset $(2, 2, 2)$. Whitney's theorem then gives

$$C, D \in \{K_3, K_{1,3}\}. \tag{B.109}$$

If the paired line components are not K_3, neither connected root is a member of the Whitney pair. They are therefore Whitney-general, and Theorem 9 applies directly to (B.108):

$$C \equiv_{\text{3-WL}} D. \tag{B.110}$$

Appendix B.2 states the cited backward-transfer theorem and treats the Whitney exception.

Consequently, all nonexceptional nonisolated root components pair 3-WL equivalently, while

$$t(G) + \kappa(G) = t(H) + \kappa(H). \tag{B.111}$$

If $i(G) = i(H)$ and $\kappa(G) = \kappa(H)$, equation (B.111) also gives $t(G) = t(H)$. Re-pair the line-K_3 components so that triangle roots meet triangle roots and claw roots meet claw roots, and pair the isolated vertices. Lemma 37 now proves $G \equiv_{\text{3-WL}} H$. The same argument works starting from equality of i and t.

To obtain the degree-multiset formulation, subtract the common degree histograms of the paired nonexceptional components. The remaining counts are

	$d=0$	$d=1$	$d=2$	$d=3$
G	$i(G)$	$3\kappa(G)$	$3t(G)$	$\kappa(G),$

(B.112)

with zero entries at all other degrees. In view of (B.111), equality of these residual histograms is equivalent to equality of i and κ, and also to equality of i and t. This proves all reverse implications. The forward implication from root 3-WL equivalence to equal degree multisets is immediate from the first diagonal refinement. Theorem 35 follows. □

B.9.4 RECOVERING INTERMEDIATE SIGNATURES FROM THE FIRST AND LAST

Stable pair refinement determines the path and claw component counts needed to recover the intermediate signatures.

Lemma 38 (Path and claw multiplicities). *If $X \equiv_{\text{3-WL}} Y$, then*

$$p_m(X) = p_m(Y) \quad (m \geq 1), \qquad \kappa(X) = \kappa(Y).$$

Proof. Lemma 37 bijects the connected components of X, Y so that every matched pair is 3-WL equivalent. Such a pair has the same order and degree multiset. Among connected graphs, these data identify every path: P_1 is the one-vertex degree-zero graph, P_2 has two vertices of degree one, and for $m \geq 3$ the multiset with two degrees one and $m-2$ degrees two forces a connected graph to be P_m. They likewise identify the claw by its degree multiset $(3, 1, 1, 1)$. The component bijection therefore restricts to bijections of the P_m components for every m and of the claw components, including all repeated copies. □

Lemma 39 (Path extinction identity)**.** *For every finite simple graph X and $q \geq 0$,*

$$i(L^q(X)) = p_{q+1}(X).$$

Proof. For a connected graph C and $k \geq 1$, $L(C) \cong P_k$ if and only if $C \cong P_{k+1}$. The forward implication is elementary: a vertex of degree at least three would create a triangle in the line graph, while a cycle in C would create a cycle in $L(C)$. Thus C is connected with maximum degree at most two and is not a cycle, so it is a path; its number of edges fixes its order. The reverse implication is $L(P_{k+1}) \cong P_k$. Iterating gives $L^q(C) \cong K_1$ if and only if $C \cong P_{q+1}$.

Line iteration acts componentwise on components containing an edge, and an iterate of such a connected component remains connected until it becomes empty. Thus each isolated vertex of $L^q(X)$ comes from exactly one root P_{q+1} component, and every such component contributes one isolate. □

Proof of Theorem 3. Equality of $S_{\leq r}$ implies equality of the root and final signatures. Root equivalence, together with Lemma 38, implies equality of D_r. It remains to recover $S_{\leq r}$ from equality of D_r and of S_r.

Descend from $n = r - 1$ to $n = 0$. Suppose $S_{n+1}(G) = S_{n+1}(H)$. For $n \geq 1$, both $L^n(G)$ and $L^n(H)$ are claw-free, and Lemma 39 gives $i(L^n(G)) = p_{n+1}(G) = p_{n+1}(H) = i(L^n(H))$. At $n = 0$, D_r gives equality of both the isolate and claw counts. Theorem 35 therefore yields $S_n(G) = S_n(H)$ at every step. This proves all three equivalences. □

B.9.5 Proof of the all-simple-graph completeness theorem

We prove Theorem 36 by matching components consistently across all line depths.

B.9.6 Remove the nonprolific components at depth zero

A connected graph is prolific when it is not a path, a cycle, or the claw $K_{1,3}$. Apply Lemma 37 to the $r = 0$ instance of (B.106). A connected graph 3-WL equivalent to a nonprolific graph is isomorphic to it:

- order and degree multiset identify every path among connected graphs;
- a connected 2-regular graph of a fixed order is that cycle; and
- the degree multiset $(3, 1, 1, 1)$ identifies the claw.

This includes K_1, K_2 as paths and K_3 as a cycle. The same argument shows symmetrically that a prolific component cannot be paired with a nonprolific one. Hence the nonprolific components of G, H already have the same isomorphism multiset. Write the remaining prolific components as

$$C_1, \ldots, C_p \qquad \text{and} \qquad D_1, \ldots, D_p. \tag{B.113}$$

If $p = 0$, all matching statements below use the standard empty perfect matching and are vacuous.

Lemma 40 (Prolificity is line-closed)**.** *If X is connected and prolific, then $L(X)$ is connected and prolific.*

Proof. Since $X \neq P_1$, it has an edge, so its line graph is connected. It is claw-free: among three neighbors of the line vertex representing an edge uv, two meet the same endpoint of uv and hence are adjacent. If $L(X) \cong P_n$, then X has neither a vertex of degree at least three (which would create a line-graph triangle) nor a cycle (which would create a line-graph cycle). Thus X is a path.

If $L(X) \cong C_n$, then for $n \geq 4$, Whitney's theorem applied against $L(C_n) \cong C_n$ gives $X \cong C_n$; for $n = 3$, the connected simple roots are exactly K_3 and $K_{1,3}$. Every alternative contradicts prolificity. □

B.9.7 Nested perfect matchings at the original depth

For $t \geq 0$, let $B_t \subseteq [p] \times [p]$ be the bipartite graph defined by

$$i \, B_t \, j \quad \Longleftrightarrow \quad L^t(C_i) \equiv_{3\text{-WL}} L^t(D_j). \tag{B.114}$$

Lemma 40 makes every graph in (B.114) connected and prolific. Conversely, nonprolific components remain nonprolific or vanish: $L(P_n) \cong P_{n-1}$ for $n \geq 2$ (while $L(P_1)$ is empty), $L(C_n) \cong C_n$, and $L(K_{1,3}) \cong K_3$. Line iteration is componentwise, so the p descendants $L^t(C_i)$ and $L^t(D_j)$ are precisely the prolific components at that depth. Apply Lemma 37 to $L^t(G), L^t(H)$. As above, order and degree multiset prevent an equivalent connected prolific component from pairing with a nonprolific one. Hence the component matching restricts to a perfect matching of B_t.

The graphs are nested:

$$B_{t+1} \subseteq B_t. \tag{B.115}$$

Indeed, an edge of B_{t+1} gives equivalent line graphs of the connected roots $L^t(C_i), L^t(D_j)$. Theorem 9 reflects this equivalence unless the roots form the pair $K_3, K_{1,3}$, which prolificity excludes. Since the decreasing sequence lies in the finite set $[p] \times [p]$, it stabilizes; therefore $B_\infty = \bigcap_{t \geq 0} B_t$ is one of the B_t and has a perfect matching M (the empty matching when $p = 0$).

For every $(i, j) \in M$, the connected graphs C_i, D_j are 3-WL equivalent at every line depth. The connected completeness theorem (Appendix B.8) gives $C_i \cong D_j$. Together with the nonprolific components matched at depth zero, these isomorphisms prove $G \cong H$. □

Thus all-depth equivalence determines the graph up to isomorphism, even though an individual line transformation can lose component information. Theorem 3 retains precisely the missing counts: either (S_0, S_r) or (D_r, S_r) determines $S_{\leq r}$ for $r \geq 1$.

B.10 Pairs indistinguishable up to any depth, and strictness

To prove Theorem 4, we describe each vertex of $L^r(G)$ using at most $r + 1$ root vertices, following the history description of iterated line graphs (Knor & Niepel, 1998). The proof of Lemma 23 shows that the root support of a depth-r history is connected, with at most r edges and hence at most $r + 1$ vertices. The support does not uniquely determine a line-graph vertex, so we also record the ordered ancestry.

An *oriented history* records the ordered binary ancestry of a line vertex down to vertices of G. At each internal node, forgetting the order of the two children recovers the corresponding line vertex. Repeated occurrences are retained. Replace the root-vertex labels by indices in $[r + 1]$. The resulting binary trees form a finite set of templates $\mathcal{T}_r$ that depends only on r.

Lemma 41 (Compact history interpretation)**.** *Vertices of $L^r(G)$ admit codes $(\tau, \mathbf{x})$ with $\tau \in \mathcal{T}_r$ and $\mathbf{x} \in V(G)^{r+1}$. Validity of a code, equality of the represented vertices, and their adjacency are given by quantifier-free formulas on one or two coordinate blocks. The valid codes map onto $V(L^r(G))$.*

Proof. Label an oriented history using its at most $r + 1$ root vertices and fill unused coordinates by repeating a label. This gives a code for every line vertex. The validity formula checks root adjacency at depth one and, recursively, validity and adjacency of the two children. Equality allows the two children to be swapped at each internal node. At positive depth, two histories are adjacent exactly when they are distinct and share a child. All these tests are finite Boolean combinations of root equalities and edge predicates. Quotienting valid codes by represented equality therefore gives exactly $L^r(G)$. □

Write $G \equiv_{\mathrm{C}^q} H$ for equivalence in q-variable counting logic. At $q = 3$ this is the 3-WL relation used in this paper. We use the generalized interpretation theorem for counting logic (Otto, 1997). An interpretation using m coordinates for each output vertex pulls q-variable equivalence back to mq variables, provided its defining formulas fit this budget. The theorem allows a quotient of the code domain, so the number of codes for each vertex need not be constant.

Lemma 42 (Finite template tags)**.** *Let T be fixed and finite. Replace a structure A by $T \times A$, with unary tag predicates, equality of underlying elements, and each base relation evaluated after dropping the tags. In particular, lifted graph adjacency satisfies $E^*((t,x),(u,y)) \iff E(x,y)$ even when $t \neq u$. This construction preserves k-variable counting equivalence.*

Proof. Use the bijective k-pebble game, which characterizes k-variable counting equivalence (Hella, 1996). Whenever Duplicator's strategy for the untagged structures answers with a bijection f, she answers with $\mathrm{id}_T \times f$. Equality, tags and lifted relations of pebbled elements depend only on their tags and underlying elements, so every position remains a partial isomorphism. □

Theorem 43 (Counting-logic pullback)**.** *For finite simple graphs G, H, $r \geq 0$, and $q \geq 2$,*

$$G \equiv_{\mathrm{C}^{q(r+1)}} H \quad \Longrightarrow \quad L^r(G) \equiv_{\mathrm{C}^q} L^r(H). \tag{B.116}$$

Proof. Apply Lemma 42 with tags $\mathcal{T}_r$. The code in Lemma 41 becomes an $(r+1)$-tuple of tagged vertices whose coordinates have the same tag. Its domain formula uses $r+1$ variables; equality and adjacency use at most $2(r+1)$. The generalized interpretation theorem therefore applies within the $q(r+1)$-variable budget and gives the conclusion whenever the interpreted domains are nonempty. Existence of a valid code is itself expressible with $r+1$ variables, so the two interpreted graphs are either both nonempty or both empty. In the latter case they are isomorphic. This also covers empty roots because $r+1 \geq 1$. □

Proof of Theorem 4. Fix R. Choose a connected 3-regular base of sufficiently large treewidth and take the uncolored link/middle-vertex representatives $Y(B), \widetilde{Y}(B)$ of the cited CFI construction. The classical lower bound supplies arbitrarily large counting widths (Cai et al., 1992), while the uncolored representatives remain nonisomorphic and have the same exact treewidth threshold (Chen et al., 2025). Every middle vertex has three link neighbors, and every link vertex has two middle neighbors and one cross-gadget neighbor, so these representatives are 3-regular; connectedness follows from connectedness of the base and minimum base degree three. This is the link/middle convention of the cited uncolored construction, not the parity/wire convention used for the separate finite certificate in Appendix D.7. Choose B so the resulting pair is equivalent in C^s for some $s \geq 3(R+1)$, and set $G = Y(B)$ and $H = \widetilde{Y}(B)$. For every $0 \leq r \leq R$, Theorem 43 with $q = 3$ gives

$$L^r(G) \equiv_{\text{3-WL}} L^r(H),$$

as required.

Their roots are connected and 3-regular. The line graph of a connected graph with at least two edges is connected, and the line graph of a d-regular graph is $(2d-2)$-regular. Thus every iterate remains connected and regular of degree at least three. □

Proof of Theorem 1. First consider preservation. If G, H are connected and are not paths, their iterates remain nonempty and connected (Lemma 45). At every positive depth they are also claw-free. For $r \geq 1$, Theorem 9 therefore gives

$$S_{r+1}(G) = S_{r+1}(H) \quad \Longrightarrow \quad S_r(G) = S_r(H).$$

The same implication holds at $r = 0$ when the claw is excluded from the root domain, removing the only exceptional pair.

To prove strictness at any prescribed $r \geq 0$, choose nonisomorphic connected 3-regular graphs G, H that agree through depth r (Theorem 4). By Theorem 2, they have a first separating depth $t > r$. Set $X = L^{t-r-1}(G)$ and $Y = L^{t-r-1}(H)$. Then

$$S_r(X) = S_r(Y), \qquad S_{r+1}(X) \neq S_{r+1}(Y),$$

because these compare the original pair at depths $t-1$ and t. Both shifted roots remain connected and regular of degree at least three, so they lie in each of the asserted domains. Moreover, $S_j(X) = S_j(Y)$ for every $0 \leq j \leq r$, by the minimality of t. They therefore also witness strictness of $S_{\leq r+1}$ over $S_{\leq r}$ on all finite simple graphs, where refinement itself follows directly by retaining the prefix. □

C GNN ADAPTER AND EXACT EXECUTION

We prove that applying a 3-WL-equivalent family to the original graph and its final line graph gives exactly the cumulative distinguishing power of iterated line-graph refinement. We then prove that compressed execution gives the same output as dense execution for every choice of model parameters. The first result concerns the whole family; an individual network may distinguish fewer pairs.

C.1 ROOT AND FINAL ENCODERS ON ARBITRARY GRAPHS

Let $\mathcal{F} = \{f_\theta : \theta \in \Theta\}$ be a nonempty family of graph encoders, with a common output space. Its indistinguishability relation is

$$G \equiv_{\mathcal{F}} H \quad \Longleftrightarrow \quad f_\theta(G) = f_\theta(H) \text{ for every } \theta \in \Theta.$$

Assume that $\equiv_{\mathcal{F}}$ is exactly stable 3-WL equivalence. The family may contain networks of different widths and numbers of layers. An individual network need not distinguish every pair separated by 3-WL. The empty graph is assigned a fixed embedding.

Proposition 44 (Root and final augmentation)**.** *Fix $r \geq 0$. On all finite simple graphs, the family*

$$G \longmapsto \big(f_\theta(G), f_\theta(L^r(G))\big), \qquad \theta \in \Theta,$$

has exactly the indistinguishability relation given by $S_{\leq r}$. The same statement holds with independent parameters in the two branches.

Proof. Equality of the displayed pair for every θ holds exactly when $G \equiv_{\mathcal{F}} H$ and $L^r(G) \equiv_{\mathcal{F}} L^r(H)$. By the encoder assumption these are $S_0(G) = S_0(H)$ and $S_r(G) = S_r(H)$. Theorem 3 identifies their conjunction with equality of $S_{\leq r}$. The argument for independent parameters is the same, since the parameter family is nonempty. At $r = 0$ the two inputs coincide, so a shared encoder need only be evaluated once. □

The original graph encoding preserves the component counts required by Theorem 3. Concatenating encodings of the intermediate line graphs does not increase the family's distinguishing power, although it may affect learning at a fixed width and number of layers. The proof uses concatenation; averaging or gating the encodings can discard information.

C.2 ONE SCALAR SUFFICES FOR CONNECTED ROOTS

For a connected graph, its vertex count can replace the original graph encoding. This also covers paths and the triangle–claw pair.

Lemma 45 (Connected graphs that disappear)**.** *For a nonempty connected finite simple graph G and $r \geq 1$, $L^r(G)$ is empty if and only if $G \cong P_j$ for some $1 \leq j \leq r$.*

Proof. Each line step deletes one vertex from a nontrivial path, and the line graph of P_1 is empty. Conversely, a connected graph which is not a path either contains a cycle or has a vertex of degree at least three. A cycle persists as a subgraph after every line step. In the second case the three incident edges form a triangle in the first line graph, which then persists. Here an inclusion of root graphs induces an inclusion of their line graphs. Thus a connected non-path graph never disappears. □

Proposition 46 (Connected order and final signature)**.** *For nonempty connected finite simple graphs G, H and $r \geq 0$,*

$$S_{\leq r}(G) = S_{\leq r}(H) \quad \Longleftrightarrow \quad \big(|V(G)|, S_r(G)\big) = \big(|V(H)|, S_r(H)\big).$$

Proof. The forward direction follows because S_0 preserves graph order. The case $r = 0$ is immediate. For the converse with $r \geq 1$, suppose the vertex counts and final signatures agree. Stable 3-WL preserves the number of vertices, so $L^r(G)$ and $L^r(H)$ are either both empty or both nonempty. In the empty case, Lemma 45 makes both roots paths, and their common order makes them isomorphic.

Otherwise $L^j(G)$ and $L^j(H)$ are nonempty and connected for every $0 \leq j \leq r$. At every positive depth they are line graphs and hence claw-free. Starting from $L^r(G) \equiv_{\text{3-WL}} L^r(H)$, apply Theorem 9

successively down to the roots. No positive-depth step can encounter the exceptional pair $\{K_3, K_{1,3}\}$ because both graphs there are claw-free. The root step either gives $G \equiv_{3\text{-WL}} H$, or makes G, H the Whitney pair. The latter is excluded by the equal root orders, since those orders would be three and four. Therefore $S_0(G) = S_0(H)$, and Theorem 3 gives the result. □

Corollary 47 (Connected encoder augmentation)**.** *Under the encoder-family assumption above, the family*

$$G \longmapsto \big(f_\theta(L^r(G)), |V(G)|\big)$$

has exactly the indistinguishability relation $S_{\le r}$ on nonempty connected finite simple graphs.

Proof. Equality for all parameters is equality of root order and of the final stable signature. Apply Proposition 46. □

Corollary 48 (When the final encoder alone suffices)**.** *On connected graphs that are neither paths nor claws, the family $G \mapsto f_\theta(L^r(G))$ alone has exactly $S_{\le r}$. For a prescribed r, the same conclusion holds on the larger domain of connected non-claw graphs whose rth line graph is nonempty.*

Proof. Both final line graphs are nonempty. Apply backward transfer as above; the excluded claw removes the only possible root exception. Theorem 3 completes the argument. □

The triangle may remain in these domains because its competing root, the claw, is excluded. Excluding only K_3 and $K_{1,3}$ on all connected graphs is insufficient: already at $r = 2$, P_1 and P_2 have identical empty final line graphs. Appending the original vertex count, as in Corollary 47, handles these cases. The count must be represented exactly; a floating-point implementation must respect its exact-integer range.

C.3 Compressed execution of pair-based networks

Let $X = L^r(G)$ and $N = |V(X)|$. Dense pair features of width d require $N^2 d$ scalars. When pairs share the same features at every layer, compressed execution stores one feature vector per class. Suppose the ordered pairs are partitioned into relations $R_1, \dots, R_q$ forming a coherent configuration whose indicator algebra contains I, A_X, J. Stable pair colors and the ordered-pair orbitals of an automorphism group both provide such partitions. Write B_c for the indicator matrix of R_c. A compressed feature state stores $h_c \in \mathbb{R}^d$ and expands to the dense tensor

$$E(h)_{uv} = h_c \qquad \text{when } (u, v) \in R_c.$$

For $(u, v) \in R_c$, the intersection numbers

$$p^c_{ab} = \#\{w : (u, w) \in R_a,\ (w, v) \in R_b\}$$

do not depend on the representative (u, v). Consequently

$$B_a B_b = \sum_c p^c_{ab} B_c, \qquad B_a \circ B_b = \mathbf{1}[a = b] B_a, \qquad B_c^{\mathsf{T}} = B_{c^*}.$$

Here c^* indexes the reversed relation $R_{c^*} = \{(v, u) : (u, v) \in R_c\}$. Shared pointwise maps, channel mixing, transposition and Hadamard products can therefore be computed from the class values. Channelwise matrix multiplication and sum pooling are

$$(h \star k)_c = \sum_{a,b} p^c_{ab}\, h_a \circ k_b, \qquad \mathrm{Pool}(h) = \sum_c |R_c| h_c. \tag{C.1}$$

Diagonal relations specify vertex classes. Each pair relation has a fixed source class and target class, with a constant number of related vertices in each row and column. These counts give matrix–vector products and row and column sums; the diagonal classes give diagonal insertion and extraction. The trace readout uses only diagonal classes,

$$\mathrm{Tr}(h) = \sum_{c:\, R_c \subseteq \{(u,u) : u \in V(X)\}} |R_c| h_c.$$

These operations include every matrix operation used by PPGN++ and $\mathrm{G^2N^2}$. Class indices specify where values are stored and summed; the network does not receive them as input features.

Proposition 49 (Exact execution)**.** *Work in real arithmetic. Let a pair encoder be composed of the operations above, with initialization constant on each pair and vertex class. Replacing each operation by its weighted compressed implementation gives the same expanded states and graph output for every setting of the neural parameters. In particular, a compressed layer $\widehat{F}_\theta$ and its dense version F_θ satisfy*

$$E\big(\widehat{F}_\theta(h)\big) = F_\theta\big(E(h)\big) \quad \textit{for every parameter choice } \theta. \tag{C.2}$$

Proof. Initialization agrees by assumption. Each displayed identity shows that performing an operation on class values and then expanding gives the same result as dense evaluation. Shared pointwise functions commute because the expanded inputs are constant on each class. Induction over the computation graph therefore proves equality of every intermediate state. The weighted readout identities give the same graph output. □

A finer coherent configuration compatible with the initialization is also valid. The same equations hold with more classes. Residual additions and concatenations preserve the identity. The real functions have equal derivatives wherever those derivatives exist, although floating-point summation order can produce numerical differences.

For symmetry groups generated by commuting involutions, the implementation computes the matrix products in (C.1) blockwise after a real Fourier transform over the group, and returns to class values before each nonlinear map. This changes how the same products are computed, not Eq. (C.2). The supplementary code provides the implementations and tests comparing outputs and gradients with dense evaluation. Appendix D.1 gives the native ILG sine factorization.

D BREC Experiments and Exact Certificates

This appendix describes the models and neural evaluation. The reported untrained results use ten seeds through depth two and seed 100 on the 39 depth-three CFI pairs. PPGN++ and native ILG each separate $39/39$; $\mathrm{G^2N^2}$ separates $24/39$, and every pair passes its isomorphic control.

D.1 Backbone implementations

Our reference implementations follow the PPGN++ matrix layer of Puny et al. (2023) and the reduced $\mathrm{G^2N^2}$ grammar of Piquenot et al. (2024). The supplementary code checks the BREC dataset files by hash.

Let $\star$ denote channelwise matrix multiplication and $\odot$ entrywise product. Square brackets denote feature concatenation. Neural maps share parameters across the corresponding vertices or ordered pairs.

PPGN++. Following Puny et al. (2023), the pair-state update is

$$Z' = \bar{m}_3\big[\bar{m}_1([Z, Z^{\mathsf{T}}]) \star \bar{m}_2(Z),\ Z\big]\,. \tag{D.1}$$

Each $\bar{m}_i$ has separate diagonal and off-diagonal MLPs. Initialization uses $I, A, J - I - A$, with the redundant non-edge channel included explicitly. The readout preserves pair multiplicities by separately summing diagonal and off-diagonal features. We exclude the optional higher-degree polynomial input features, which would change the expressivity comparison.

$\mathrm{G^2N^2}$. The reduced matrix grammar of Piquenot et al. (2024) maintains pair memory C and vertex memory H, with updates

$$\begin{aligned} C' &= m_M[C,\ L_1(C) \star L_2(C),\ L_3(C) \odot L_4(C),\ \mathrm{Diag}(L_6(H))]\,, \\ H' &= m_V[H,\ L_5(C) \star L_7(H)]\,. \end{aligned} \tag{D.2}$$

Here $L_1, \ldots, L_7$ are learned linear maps on feature channels, distinct from the line graph transformation L. The product in the vertex update is channelwise matrix–vector multiplication. Both updates use the previous C, H, including the vertex-state skip connection. We initialize $C = [I, A]$, $H = \mathbf{1}$, and preserve node, diagonal-pair and off-diagonal-pair sums in the readout. Independent branch widths control the number of matrix products.

Evaluation configuration. The reported models use five layers, width 16, FP64 and $\sin(5x)$ activations, with each backbone's 16-output head and a readout of all layers. PPGN++ and G^2N^2 use hidden LayerNorm, which normalizes each pair and vertex feature vector over its channels. Native ILG has no hidden normalization and uses parameter-free output LayerNorm with $\epsilon = 10^{-5}$. Pair dropout is absent. These models replace the published activations with sine. As in BREC, each pair is evaluated with its own network. For a given seed, the networks for successive pairs are drawn from one seeded random stream with the default PyTorch parameter initialization. Since layer normalization acts separately on each pair, vertex or graph-output vector, it cannot separate inputs identified by stable 3-WL, and the exact counts in Table 1 remain upper bounds for these models. We do not rely on the normalized variants attaining 3-WL: the adapter results assume a 3-WL-equivalent family, and the experiments measure separation directly. We exclude raw identifiers and positional encodings.

Native ILG^r-3-GNN. Let J_k denote the graph encoder underlying ILG-k-GNN (Yang, 2026) on ordered $(k-1)$-tuples. ILG-k-GNN uses $J_k(L(G))$; at line depth r we use $J_k(L^r(G))$. Its input relations on $L^r(G)$ are computed from shared endpoints of edges of $L^{r-1}(G)$. This defines a model for each k; the adapter guarantee and the experiments here concern $k = 3$.

Write a_{uv} for the atomic equality/adjacency/nonedge relation on the graph being evaluated. With constant object inputs absorbed into the learned maps, the pair recurrence has the form

$$z_{uv}^{t+1} = \phi_t\left(z_{uv}^t, \sum_w \psi_t(z_{wv}^t, z_{uw}^t, a_{wu}, a_{wv})\right).$$

All replacement vertices participate jointly. On a line graph, these atomic relations coincide with the equality, incidence and disjointness codes for edges of ILG-k-GNN. They contain no endpoint identifiers or edge orientations. For efficiency, the implementation splits the message map so that the sum over replacement vertices becomes matrix products. With the sine activation, the identity $\sin(a+b) = \sin a \cos b + \cos a \sin b$ turns each message sum into two such products. These rewrites compute the same function in real arithmetic; the supplementary code gives the details.

D.2 Activation, precision and the root baseline

The original BREC study reports 233 separations for PPGN, compared with 270 for exact 3-WL (Wang & Zhang, 2024). A finite network can fall short of its family's expressivity because its parameters and numerical implementation affect which distinctions are observed. We examined these effects in a separate trained experiment on the original graphs.

We compared ReLU/FP32 with $\sin(5x)$/FP64 on all 400 pairs using seeds $100, 200, \ldots, 1000$. PPGN used the BREC implementation with five width-32 blocks. PPGN++ and G^2N^2 used five layers, width 16, hidden LayerNorm and an all-layer readout. Within each model, the two conditions used the same initialization seeds and training procedure. We trained with Adam and the BREC cosine loss for at most 20 epochs; the supplementary code provides the full configurations. The selected precision applied to the network, loss and evaluation statistic. These trained controls are separate from the untrained results in Table 1. Following BREC (Wang & Zhang, 2024), we report the best result over ten seeds. Table 2 also shows the individual scores.

The PPGN reproduction with ReLU/FP32 reaches 234, close to the published result of 233. With $\sin(5x)$/FP64, its best single-seed score rises to 267, and the ten seeds collectively separate all 270 theoretically separable pairs. PPGN++ and G^2N^2 reach 270 in every seed. The untrained root controls in Table 1 also reach 270 in every seed with these numerical settings. Thus the line graph experiments start from models that already attain the root 3-WL count; the additional separations arise after transformation.

Our PPGN runs use independent initialization seeds, do not reproduce the official data-loading order, and retain the upstream max-pooling routine. Its FP64 diagonal mask can admit diagonal entries into the off-diagonal maximum; the effect of correcting this mask on the reported decisions has not been tested. PPGN++ and G^2N^2 use separate sum readouts. The supplementary code includes the ablation harnesses.

Table 2: Root-graph separations under two numerical configurations. Each seed evaluates all 400 pairs; **Best** is the highest single-seed count. All isomorphic controls pass, and no pair outside the exact 3-WL separation set is counted.

Model	Setting	Seed										**Best**
		100	200	300	400	500	600	700	800	900	1000	
PPGN	ReLU / FP32	230	234	233	229	232	231	232	233	229	229	**234**
	$\sin(5x)$ / FP64	265	265	267	266	266	265	266	266	266	265	**267**
PPGN++	ReLU / FP32	202	206	206	202	205	204	204	204	203	205	**206**
	$\sin(5x)$ / FP64	270	270	270	270	270	270	270	270	270	270	**270**
G^2N^2	ReLU / FP32	207	207	204	205	205	206	204	208	207	209	**209**
	$\sin(5x)$ / FP64	270	270	270	270	270	270	270	270	270	270	**270**

D.3 BREC GRAPHS AND EVALUATION PAIRS

We preserve the official 400 pair identities and their isomorphic controls. Relabelings are copies of the same graph, not additional test pairs. The main study assigns each pair its first separating line depth. Pair 380 is the distance-regular pair evaluated at depth two. Depth-three evaluation covers 320–353 and 355–359; pair 354 is assigned depth six and is excluded from these 39 pairs. These evaluations apply each model's 16-output encoder only to $L^3(G)$. The original graphs are connected and are neither paths nor claws, so no original-graph encoding or vertex count is needed.

D.4 BREC SEPARATION STATISTIC

We use the reliable paired comparison of Wang & Zhang (2024) as implemented in their evaluation code. For the 32 relabeled pairs with output differences D_i, the statistic is $\bar{D}^{\mathsf{T}}\widehat{\Sigma}^{+}\bar{D}$, where $\bar{D}$ and $\widehat{\Sigma}$ are the sample mean and covariance of the D_i and $+$ denotes the pseudoinverse. The T^2 statistic stated by Wang & Zhang (2024) has an additional factor 32; we follow their evaluation code. A pair counts as separated when its statistic exceeds the threshold 72.34 and differs from that of its own isomorphic control (absolute tolerance 10^{-6}, relative tolerance 10^{-5}), and the control statistic stays below 72.34.

D.5 CHECKS ON NEURAL EVALUATION

In exact arithmetic, relabeled copies of one graph have identical outputs. At depth three, the largest distance of a copy's FP64 output from the mean over copies of the same graph is 7.4×10^{-6} for PPGN++ and 3.2×10^{-5} for native ILG, but 3.0 for G^2N^2, whose outputs have norm about 1.7. Every G^2N^2 nonseparation has spread at least 1.88, larger than the distance between the mean outputs of its two graphs. The same spread inflates both the main and control covariances, so these controls pass while the main test fails. These misses therefore reflect the numerical sensitivity of this untrained configuration on large line graphs; the experiment does not determine whether the same weights would separate these pairs in exact arithmetic.

Each model encodes the final line graph alone. The general adapter using both original and final graph representations is not evaluated here. Pairs 6, 17, 24, 52, 186, 187, 203, 265, 278 and 297 contain disconnected graphs, so Corollary 48 does not cover final-only encoding for them at depth one; native ILG nevertheless separates all ten in every seed. For each pair and seed, 32 main comparisons and 32 isomorphic-control comparisons use 128 independently computed graph representations. Separation requires finite statistics and both the main test and that pair's own reliability control to pass, using the statistic specified in Appendix D.4.

Cumulative counts take the union of separations at the assigned depths. The experiment does not retest earlier pairs at greater depths, so it does not measure whether their neural separations persist. All planned evaluations completed.

Configuration choices were developed using BREC. The depth-three configuration is fixed across all 39 depth-three pairs. Appendix D.1 specifies normalization and initialization, and Appendix C.3 specifies compressed execution.

D.6 THE DEPTH-SIX CERTIFICATE

Proposition 50 (Exact depth-six certificate)**.** *Let X_0, X_1 be BREC pair 354 in dataset order. Then*

$$S_j(X_0) = S_j(X_1) \quad (0 \le j \le 5), \qquad S_6(X_0) \ne S_6(X_1).$$

Thus their first separating line depth is exactly six.

The CFI inputs. Let $U = \{1, 2, 3\}$ and $W = \{4, 5, 6\}$. The base graph B has vertices $\{0\} \cup U \cup W$ and edges $0U$, all of $U \times W$, and the three edges within W. Write $E_B(v)$ for the edges incident to v. For a twist vector $\epsilon \in \mathbb{F}_2^{E(B)}$, the standard CFI construction has ports $p(v, e, b)$ for $e \in E_B(v)$ and $b \in \mathbb{F}_2$, and middle vertices $m(v, S)$ for even-cardinality $S \subseteq E_B(v)$. Its edges are

$$m(v, S) \sim p(v, e, \mathbf{1}[e \in S]), \qquad p(u, e, b) \sim p(v, e, b + \epsilon_e) \quad (e = uv).$$

The two inputs use $\epsilon = 0$ and a single twist on $\{5, 6\}$. The supplementary code constructs explicit bijections to BREC pair 354 and verifies every adjacency. The following proof applies to these two constructed graphs, and hence to the benchmark inputs.

Nonseparation through depth five. For a line vertex x, its *root support* is the set of root edges occurring in its recursive history. Its *middle support* $M(x)$ is the set of base vertices whose middle vertices occur in that history. Let $\mathcal{S}$ consist of the seven singletons and the fifteen endpoint sets of base edges.

Lemma 51 (Fifth-level supports and the residual base)**.** *Every vertex x of either fifth line graph has $M(x) \in \mathcal{S}$. Every root port occurring in its history lies over a base edge incident to $M(x)$. For $M_1, M_2 \in \mathcal{S} \cup \{\varnothing\}$, the graph $B - (M_1 \cup M_2)$ is connected and has at least three vertices.*

Proof. Root supports of adjacent histories differ by at most one edge in either direction. At depth one this is immediate. At the next depth, write the adjacent histories as $\{a, b\}, \{a, c\}$; their support difference is contained in the support difference of b and a, so induction applies. It follows that a depth-j history has connected root support with at most j edges. A fifth-level history contains at least two root edges. Since cross edges form a matching, its connected support includes a middle vertex.

A connected support meeting $k \ge 2$ middle gadgets needs at least $k - 1$ cross edges. In a minimal connected subgraph joining those middles, each endpoint of a cross edge also meets an internal middle–port edge. These $2(k - 1)$ internal edges are distinct, because cross edges form a matching. Thus the support has at least $3(k - 1)$ edges. Five root edges can therefore meet at most two middle gadgets, and the two must be adjacent in B. An occurring port either meets an internal support edge itself or connects by its unique cross edge to a port that does. Its base edge is consequently incident to $M(x)$.

Deleting two admissible supports removes at most four base vertices and at most two vertices of U. If some vertex of W remains, the complete U–W adjacency and the clique on W connect the remainder; 0, if present, attaches to the remaining U. If all of W is deleted, one support contains an edge within W and the other contains the third vertex of W. Neither can contain 0, so the remainder is a star centered at 0 with at least two vertices of U. □

Proof of the lower bound. All bit operations below are in $\mathbb{F}_2$. An incidence assignment $a = (a_{v,e} : v \in e)$ satisfying

$$a_{u,e} + a_{v,e} = \epsilon_e + \epsilon'_e, \qquad \sum_{e \in E_B(v)} a_{v,e} = \mathbf{1}[v = w]$$

has a single *defect* w. It defines a port map $p(v, e, b) \mapsto p(v, e, b + a_{v,e})$ and, for $v \ne w$, a middle map

$$m(v, S) \longmapsto m\big(v, S \triangle \{e \in E_B(v) : a_{v,e} = 1\}\big).$$

The parity equation makes the translated subsets even outside w; the edge equation preserves cross adjacency. These maps and their inverses therefore give an induced isomorphism after deleting the

middle vertices at w. Five line transformations give an induced isomorphism on precisely the histories avoiding those middles. It bijects each class with middle support M onto the corresponding class in the other graph whenever $w \notin M$.

Such an assignment exists by placing the twist at one endpoint of the twisted base edge. Flipping both incidence bits along a path from w to w' moves the sole defect to w'. A path avoiding a set V_0 changes no incidence over an edge incident to V_0. By Lemma 51, the change therefore leaves the image of every history with middle support contained in V_0 unchanged.

Duplicator maintains one assignment whose defect avoids all pebbled supports and whose fifth lift induces their correspondences. After a pebble is removed, let $V_0 = M_1 \cup M_2$ be the union of the two surviving supports, allowing empty supports for unoccupied pebbles. For each possible incoming support $M \in \mathcal{S}$, choose $w_M \notin V_0 \cup M$. This is possible because $B - V_0$ has at least three vertices and $|M| \leq 2$. Move the current defect to w_M along a path in the connected graph $B - V_0$. The resulting assignment fixes both surviving correspondences and gives a bijection on the entire class with support M.

Duplicator announces the union of these class bijections *before* Spoiler chooses the incoming vertex. The classes partition both fifth line graphs, and each maps bijectively onto the class with the same support, so the union is a global bijection. Its inverse handles a choice from the other graph. Once Spoiler chooses a class, its single assignment induces all three pebble correspondences. The lifted induced isomorphism preserves and reflects equality and adjacency, restoring the invariant. Duplicator can therefore play indefinitely, proving C^3 equivalence at depth five (Hella, 1996), or 3-WL equivalence under our convention. Backward preservation gives the same conclusion at every earlier depth. □

Separation at depth six. For either graph X, let $C = B_X B_{L(X)} \cdots B_{L^5(X)}$ be the product of its six unsigned incidence matrices, and let Z_M be the diagonal selector of its CFI middle vertices, determined by stable vertex colors. The ancestry Grams are $K_{\mathrm{all}} = C^{\mathsf{T}} C$ and $K_M = C^{\mathsf{T}} Z_M C$.

Proof of the upper bound. The roots are connected, have 136 vertices, and have minimum degree three. Their successive line graphs remain connected, grow in order, and satisfy $\delta(L(X)) \geq 2\delta(X) - 2$. They therefore avoid all exceptional small roots in Theorem 12. That theorem transports both ancestry Grams under any adjacency-preserving weak isomorphism at depth six. The middle selector is transported because its vertices form a union of three shared stable vertex-color classes, of sizes $4, 24, 48$.

The root-support argument above also bounds a sixth-level history by six root edges. Enumerating the 52 connected simple support graphs with at most six edges covers all such histories. Exact integer ancestry calculations on their $405{,}693$ sixth-level histories give minimum diagonal value $K_{\mathrm{all}}(x, x) = 924$. Equality occurs exactly on a six-edge path, with ancestry weights $(1, 6, 15, 20, 15, 6, 1)$. Among their squared weights, 402 can only be the sum $1 + 400 + 1$. Thus additionally requiring $K_M(x, x) = 402$ selects exactly the paths whose first, fourth and seventh vertices are middles.

On each CFI input, enumerating these paths gives $20{,}544$ selected histories. The certificate forms the graph on them whose adjacency is $K_M(x, y) \in \{40, 41, 401\}$; each graph has $819{,}264$ edges. Both its vertex selector and its adjacency are finite level sets of transported coherent matrices. By Corollary 14, equivalence of the sixth line graphs would therefore imply 3-WL equivalence of these selected graphs.

A refinement of the selected graphs that may merge exact 3-WL colors, but never splits them, already gives different color histograms at round three. Merging can only hide differences, so this mismatch shows that 3-WL distinguishes the selected graphs, contradicting their equivalence. □

Figure 5 summarizes the two certificates. The construction avoids materializing the sixth line graphs. It certifies 3-WL separation for this pair; it does not evaluate a neural model or give an efficient procedure for arbitrary deep line graphs.

The supplementary code regenerates both certificates and the maps from the BREC inputs to the CFI construction.

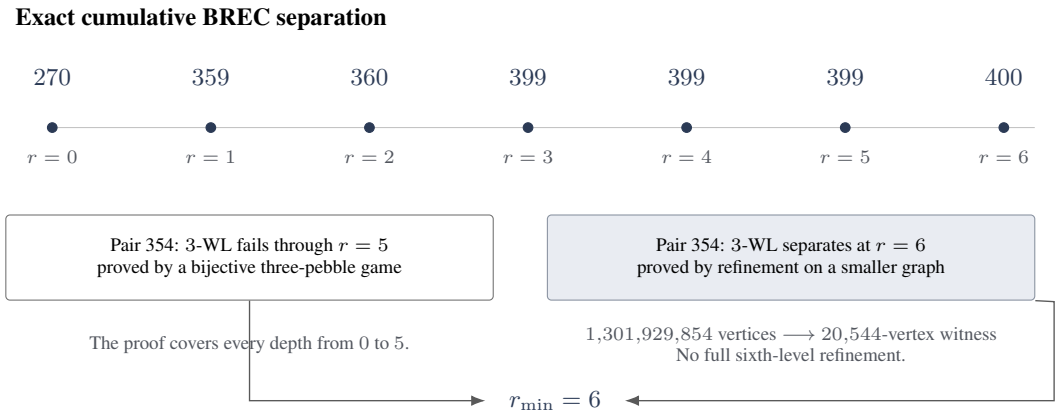


Figure 5: The last BREC pair requires depth six. The lower argument certifies equality through depth five. The upper argument interprets a smaller graph inside the sixth line level and certifies a refinement mismatch there. Together they establish six as the minimum separating depth.

D.7 An exact depth-two example

We give a concrete pair whose first separating line depth is two.

D.7.1 Construction

For a base graph X, the Cai–Fürer–Immerman graph $\mathrm{CFI}(X)$ has a *parity* vertex (v, S) for every $v \in V(X)$ and every even-cardinality $S \subseteq E(v)$, and two *wire* vertices e^0, e^1 for every $e \in E(X)$; the parity vertex (v, S) is joined to $e^{[e \in S]}$ for each $e \in E(v)$. The twisted graph $\widetilde{\mathrm{CFI}}(X)$ is obtained by exchanging e^0 and e^1 on one side of a single edge (Cai et al., 1992). Put

$$G_\star = \mathrm{CFI}(K_4), \qquad H_\star = \widetilde{\mathrm{CFI}}(K_4).$$

Both graphs are connected, bipartite, of girth six, and non-isomorphic, with 28 vertices and 48 edges: 16 parity vertices of degree three and 12 wire vertices of degree four. Because every edge joins a degree-three parity vertex to a degree-four wire vertex, every line vertex has degree $(3-1)+(4-1) = 5$, so $L(G_\star)$ is 5-regular on 48 vertices and $L^2(G_\star)$ is 8-regular on 120 vertices.

D.7.2 Certificate

Table 3: Exact joint 3-WL on the first three line depths of $G_\star, H_\star$. Both graphs use one color namespace; equality is checked to stability.

depth r	graph L^r	first split round	verdict
0	28 vertices, 48 edges	—	S_0 equal
1	48 vertices, 120 edges, 5-regular	—	S_1 equal
2	120 vertices, 480 edges, 8-regular	2	S_2 separates

Table 3 reports joint ordered-pair refinement with integer replacement counts. The runner reaches stable equality at depths zero and one, and separates the pair in round two at depth two. The supplementary code constructs the graphs and regenerates this certificate.

Because the exact-depth signatures agree at depths zero and one, the pair also witnesses $S_1 \prec S_2$ (and hence $S_{\le 1} \prec S_{\le 2}$), one level above the Shrikhande/rook witness of Section 3.

D.7.3 Comparing line depth and WL dimension

4-WL—the classical folklore update on triples—separates $G_\star$ from $H_\star$ at the root, in its second round. One additional tuple coordinate therefore distinguishes at depth zero a pair whose first line

graphs remain 3-WL equivalent. Increasing the line depth and increasing the WL dimension can require different numbers of steps to distinguish the same pair.

E Complexity Analysis and Measured Costs

E.1 What is known about relative expressive power

We compare exact tests on uncolored finite simple graphs. Under our convention, k-FWL has the distinguishing power of $(k+1)$-WL for $k \geq 2$. For fixed r, Theorem 43 implies

$$S_{\leq r} \preceq (3r+2)\text{-FWL}. \tag{E.1}$$

Indeed, root equivalence in $3(r+1)$-variable counting logic implies 3-WL equivalence at every depth $j \leq r$. This gives an upper bound; it does not identify the smallest sufficient FWL dimension. At fixed depth, the tests need not dominate one another. The incomparability of one-step ILG-3-WL and root 4-WL, equivalently root 3-FWL, on connected Whitney-general graphs was established by Yang (2026).

E.2 Total refinement time

Direct round costs are not total stable-refinement costs. For fixed k, optimized partition refinement gives $O(n^{k+1}\log n)$ total time for k-FWL (Immerman & Sengupta, 2019); in particular, stable pair refinement has an $O(n_r^3 \log n_r)$ bound. These bounds do not describe the measured running time of our reference implementations. With adjacency lists, reading the input and explicitly generating successive line graphs additionally costs $O(n+m+\sum_{j=1}^{r}(n_j+m_j))$, where $m_j = |E(L^j(G))|$. For sequential testing until separation, all earlier refinement runs also count. For testing a prescribed final depth, Theorem 3 allows their omission.

E.3 How graph structure changes the cost

The following statements concern explicit pair states and direct refinement work. They do not give lower bounds for compressed execution or include the number of refinement rounds needed for a decision.

Degree averages, variation and adjacency. Let $n > 0$, let $d(v)$ be the root degrees, and write $\bar{d} = 2m/n$ and $\sigma_d^2 = n^{-1}\sum_v (d(v)-\bar{d})^2$ for their mean and variance. Counting incident edge pairs gives

$$n_1 = m, \qquad n_2 = \sum_v \binom{d(v)}{2} = \frac{n}{2}(\bar{d}^2 - \bar{d} + \sigma_d^2), \tag{E.2}$$

$$n_3 = \sum_{uv \in E(G)} \binom{d(u)+d(v)-2}{2}. \tag{E.3}$$

Each pair of edges incident to a root vertex gives one edge of $L(G)$; since the graph is simple, it is counted once. Applying the same argument to $L(G)$ and using $d_{L(G)}(uv) = d(u)+d(v)-2$ proves the next identity. Thus, at fixed n, m, increasing degree variance increases n_2. For $m > 0$, the expansion from depth one to two is exactly

$$\frac{n_2}{n_1} = \bar{d} - 1 + \frac{\sigma_d^2}{\bar{d}}.$$

Degree averages alone therefore miss the effect of vertices of high degree. At fixed degree sequence, how these vertices are connected also matters, since expanding (E.3) gives

$$n_3 = \frac{1}{2}\sum_v d(v)^3 + \sum_{uv \in E(G)} d(u)d(v) - \frac{5}{2}\sum_v d(v)^2 + 3m.$$

Only the edge-product sum varies with the wiring. A degree-preserving switch from edges ax, by to ab, xy, when valid, changes n_3 by $(d(a)-d(y))(d(b)-d(x))$. In particular, n_3 increases when $d(a) > d(y)$ and $d(b) > d(x)$, as when the degrees of a and b both exceed those of x and y.

What one more parameter step costs. If a nonempty $L^j(G)$ has mean degree $\bar{d}_j$, then $n_{j+1}/n_j = \bar{d}_j/2$. Its explicit pair table grows by $(\bar{d}_j/2)^2$, and its direct pair-refinement work per round by $(\bar{d}_j/2)^3$. In contrast, raising FWL tuple arity from k to $k+1$ multiplies storage by n and literal color reads by $n(k+1)/k$. These root enumeration volumes depend on order and tuple arity, rather than edge density. Refinement rounds, color palettes and the smallest sufficient parameter can still depend on the graph.

For a d-regular root with $d \geq 2$, the final pair table after one line step uses $d^2/(4n)$ times the states of root 3-FWL, while its direct color reads per round are $d^3/(12n)$ times as many. Thus fixed degree gives cubic rather than quartic direct work in n; dense roots can reverse this advantage. More generally, with $N = n_r$, the final line test uses fewer direct color reads than root k-FWL precisely when

$$2N^3 < kn^{k+1}.$$

Its final pair table is smaller than the root k-FWL table precisely when $N < n^{k/2}$. These inequalities compare a single round and persistent entries; total running time and peak memory also include construction, temporary keys, earlier attempts and the actual number of rounds.

Examples of size growth. For a d-regular graph with $d \geq 2$, the exact recurrence is

$$d_r = 2^r(d-2) + 2, \qquad n_0 = n, \qquad n_{r+1} = n_r d_r/2. \tag{E.4}$$

This gives the following simple families. All asymptotics hold as the root order n grows, at fixed $r \geq 1$.

Root graph	n_1	n_2	n_r at fixed depth
Path P_n	$n-1$	$\max\{n-2, 0\}$	$\max\{n-r, 0\}$
Cycle C_n	n	n	n
3-regular graph	$3n/2$	$3n$	$\Theta(n)$
Star $K_{1,n-1}$	$n-1$	$\binom{n-1}{2}$	$\Theta(n^r)$
Complete graph K_n	$\binom{n}{2}$	$n(n-1)(n-2)/2$	$\Theta(n^{r+1})$

The star becomes a clique after one step. Its pair storage is still quadratic at depth one, but becomes quartic at depth two; for K_n it is already quartic at depth one. Paths shrink and cycles stay unchanged. Stars and complete graphs are identified by root degree information, so these are examples of representation growth, not of necessary costly separation. An adaptive test should stop before those transformations.

High symmetry or repeated pair states can reduce actual execution cost for any of the methods. The formulas above are not lower bounds on such implementations; even a large line graph may admit a small exact quotient. Conversely, regularity alone guarantees neither a small quotient nor a small sufficient depth.

Growth at larger depths. The degree identity $d_{L(G)}(uv) = d_G(u) + d_G(v) - 2$ implies

$$\Delta(L^r(G)) \leq 2^r(\Delta(G) - 2) + 2 \qquad (\Delta(G) \geq 2).$$

Since $n_{r+1} = m_r \leq n_r \Delta(L^r(G))/2$, induction gives the maximum-degree bound

$$n_r \leq n \prod_{j=0}^{r-1} \big(1 + 2^{j-1}(\Delta(G) - 2)\big).$$

For regular roots every inequality is an equality. For fixed degree $d \geq 3$,

$$\log_2 \frac{n_r}{n} = \sum_{j=0}^{r-1} \log_2\big(1 + 2^{j-1}(d-2)\big) = \frac{r^2}{2} + O_d(r).$$

In particular, fixed depth and bounded root degree give $n_r = O(n)$, with a constant that grows rapidly with depth. Removing intermediate WL runs does not change this representation size. On regular roots of degree at least three, $n_{j+1} \geq \frac{3}{2} n_j$, so

$$\sum_{j=0}^{R} n_j^3 \leq \frac{27}{19} n_R^3.$$

Thus omitting intermediate runs saves work but does not improve the leading exponent of the direct cubic per-round count on this family. This comparison holds for the same number of rounds at each level; it is not a bound on ratios of total observed running times.

Proposition 52 (Worst-case size at the first separating depth)**.** *There are connected* 3*-regular graph pairs of order* n *whose first separating line depth is* $\Omega(n)$*. At that depth the literal line graph has* $2^{\Omega(n^2)}$ *vertices.*

Proof. Use 3-regular CFI families with counting-variable lower bound linear in the order of the resulting graphs (Cai et al., 1992; Chen et al., 2025). Thus, along an unbounded sequence of orders, some constant $c > 0$ gives non-isomorphic pairs equivalent in $\mathrm{C}^{\lfloor cn \rfloor}$. Theorem 43 makes their rth line graphs 3-WL equivalent whenever $3(r+1) \leq \lfloor cn \rfloor$. The first separating depth therefore grows at least linearly in n. For 3-regular roots, the exact regular recurrence gives

$$n_r = n \prod_{j=0}^{r-1} (1 + 2^{j-1}) \geq n\, 2^{(r-1)(r-2)/2} \qquad (r \geq 2).$$

Substitution proves the claim. □

Full-width root tuple enumeration is complete with $n^{O(n)}$ states. This compares particular implementations and parameter choices, rather than giving an inherent lower bound for all implementations of iterated line refinement or for graph isomorphism.

E.4 Measured time and memory on controlled graph families

Figure 4 measures the exact WL implementations on an Apple M2 CPU with 8 GiB RAM. It uses 24 graph settings and five seeds per setting, each evaluated by root 3-WL, root 4-WL, and ILG at depths one and two. Five independent repetitions per graph and method give 2,400 executions. To vary graph order, we use $n \in \{16, 24, 32, 48, 64\}$ and mean degree four. To vary density, we fix $n = 24$ and use mean degrees $2, 3, 4, 5, 6$. Both include random regular graphs and uniform $G(n, m)$ graphs at the same vertex and edge counts. To vary the degree distribution, we fix $n = 32$ and mean degree four, with half the vertices of degree $4 - a$ and half of degree $4 + a$, for $a = 0, \ldots, 3$. We construct a simple realization with the Havel–Hakimi algorithm, then apply $10m$ successful random degree-preserving edge swaps, allowing at most $100m$ attempts. These are randomized realizations of the specified degree sequence, without a uniform-sampling claim. Seeds are $0, \ldots, 4$; disconnected graphs are retained. The grid was fixed before the full measurement run, and every combination of setting, seed and method is included.

Timed work. Each method receives the graph and a relabeled copy, so a run cannot stop early by separating its inputs. We time root indexing, line-graph construction where applicable, color initialization and one exact refinement round; graph generation and output writing are excluded. Each execution runs in a fresh process with one numerical-library thread, and executions run sequentially, with graph and method order varied across repetitions. Root 3-WL and ILG use 2-FWL and root 4-WL uses 3-FWL, all with the same refinement implementation.

Memory and aggregation. Memory is whole-process peak resident memory, including the interpreter. For each graph and method we take the median of the five repetitions; curves show medians across the five graph seeds, with 10th–90th percentile bands for solid curves. All 2,400 executions completed. These measurements concern the exact WL implementations, not the neural models.

The supplementary code contains the runner and instructions for reproducing the measurements and figure.

The neural evaluations use symmetry compression for large line graphs (Appendix C.3). Their runners record the time and memory of each evaluation.