SPECTRALIZATION: GENERALIZING SPECTRAL INVARIANTS FOR EXPRESSIVE GRAPH LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Spectral invariant graph neural networks (GNNs), which use spectral invariants to avoid eigenvector ambiguities, have demonstrated strong expressive power. Rooted graph invariants (RGIs), such as distance measures and substructure counts, are also commonly incorporated as structural encodings to improve the expressiveness of GNNs and graph Transformers. We connect these two threads by reformulating these spectral invariants in terms of spectral responses to the identity signal, which is itself an elementary RGI. The construction of these spectral invariants can thus be understood as spectrally enriching the identity signal by capturing its interactions with graph topology. We therefore propose *Spectralization*, a general framework that extends this construction from the identity signal to arbitrary RGIs, thereby substantially broadening the scope of spectral invariants. To demonstrate its benefits, we spectralize the atomic types underlying Weisfeiler–Leman (WL) refinement to obtain Spectralized k-FWL, which we prove to be strictly more expressive than k-FWL for every $k \geq 2$ and place within a hierarchy of classical and spectral WL variants. Beyond atomic types, we further show that spectralizing diverse RGIs can improve model expressiveness compared with using their raw forms directly. We then develop structural encoding schemes guided by Spectralization, with experiments corroborating these expressiveness gains on graph-isomorphism benchmarks and demonstrating improved predictive performance across multiple neural architectures and downstream tasks.

1 INTRODUCTION

Graph spectral information, including the eigenvalues and eigenvectors of the graph Laplacian, has been widely incorporated into graph neural networks (GNNs) and graph Transformers (GTs) to encode graph structure, with strong empirical performance (Dwivedi & Bresson, 2020; Kreuzer et al., 2021; Rampášek et al., 2022; Dwivedi et al., 2023). A central challenge, however, is the inherent ambiguity of eigenvectors: their signs are arbitrary, and repeated eigenvalues admit different orthonormal bases of the same eigenspace (Lim et al., 2023). This motivates the use of *spectral invariants*, which are spectral representations inherently independent of the choice of eigenbasis (Fürer, 2010; Rattan & Seppelt, 2023; Arvind et al., 2025). These spectral invariants then enabled the development of spectral invariant GNNs with strong expressive power (Lim et al., 2023; Zhang et al., 2024b) and stability under graph perturbations (Huang et al., 2024). A general pairwise representation of these spectral invariants is provided by the eigenspace projection invariant (Zhang et al., 2024b), which records, for each node pair, the projection entries across all eigenspaces together with their respective eigenvalues. The corresponding Eigenspace Projection GNNs provide an upper bound on the expressiveness of all prior spectral invariant GNNs.

A seemingly separate line of work enhances the expressiveness of GNNs and GTs by exploiting local or global graph structural information through positional and structural encodings. One common approach encodes graph connectivity using distance measures, such as shortest-path and resistance distances (Li et al., 2020; Ying et al., 2021; Zhang et al., 2023b), as well as diffusion kernels (Feng et al., 2022) and random-walk statistics (Rampášek et al., 2022; Ma et al., 2023). Another uses substructure counts to describe how nodes or node pairs[1] participate in prescribed graph patterns, such as paths, cycles, and other motifs (Barceló et al., 2021; Jin et al., 2024; Bao et al., 2025; Airale

[1]Hereafter, we use *node tuples* to include individual nodes (1-tuples) and node pairs (2-tuples).

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.

et al., 2025). All these structural quantities are instances of *rooted graph invariants* (RGIs), which assign the same value to a node tuple and its image under any graph isomorphism, thereby avoiding artificial distinctions between corresponding node tuples in isomorphic graphs (Black et al., 2024).

Our contributions. We uncover a deeper connection between these two lines through a key observation about spectral invariants, specifically on their unified representation, the eigenspace projection invariant. As established in Theorem 4.4, this invariant is equivalently characterized by the graph spectrum and the complete response profile of the identity signal under second-order spectral filtering, where the identity signal $\mathbf{1}[u = v]$ is an elementary RGI that assigns 1 when the two nodes coincide and 0 otherwise. This reformulation offers a new understanding of spectral invariants: they are spectral enrichments of the identity signal, capturing its interactions with graph topology through graph spectral filtering. Building on this understanding, we introduce *Spectralization*, a general framework that extends this spectral enrichment from the identity signal to arbitrary RGIs, thereby substantially broadening the scope of spectral invariants. Specifically, it treats RGIs as structural signals defined on node tuples and records their responses across spectral filters of the corresponding order, together with the graph spectrum. The resulting spectralized RGIs retain the information in their raw forms and can reveal additional structural distinctions, providing a broad family of spectral representations for improving model expressiveness.

To demonstrate the benefits of Spectralization, we first apply it to atomic types, which are RGIs encoding the basic structural information underlying Weisfeiler–Leman (WL) refinement (Dell et al., 2018). Using these spectralized atomic types, we generalize k-FWL refinement (Zhang et al., 2024a) to obtain *Spectralized k-FWL* (Spec k-FWL). Theorem 4.7 establishes that Spec k-FWL is strictly more expressive than k-FWL for every $k \geq 2$. We further compare its expressive power with that of Local k-WL (Morris et al., 2020) and higher-order Spectral Invariant WL (Gai et al., 2025), establishing an expressiveness hierarchy among these WL variants. Beyond atomic types, Theorem 4.10 shows that spectralizing diverse RGIs, including many substructure counts as well as several graph distance measures, can improve model expressiveness compared with using their raw forms directly.

Finally, we develop practical structural encoding schemes based on Spectralization, using finite filter banks to compute spectral responses of RGIs. These responses are then combined through learnable transformations into node-pair encodings that can be integrated into existing GNNs and GTs. Experiments on graph-isomorphism benchmark corroborate the expressiveness gains of Spec 2-FWL; evaluations on molecular and image datasets further demonstrate improved predictive performance across multiple neural architectures.

2 Related Works

We discuss the works most closely related to our framework; a broader review is provided in Appendix B. Zhang et al. (2024b) formulate spectral invariant GNNs in terms of the eigenspace projection invariant, providing a direct starting point for our work. Gai et al. (2025) extend spectral invariant GNNs to higher orders, initializing node tuples with the corresponding entries of eigenspace projection invariant in place of atomic types. We will compare the resulting WL algorithms and their expressive power in Section 4.2. Earlier, Feldman et al. (2023) showed that node pre-colorings derived from heat kernels can strictly enhance the expressive power of 1-WL. Viewed through our framework, their construction can be recovered by restricting the spectral filters to the heat-kernel family and applying them to the identity signal. From the filtering perspective, Wang et al. (2026) introduce second-order spectral filtering for node-pair signals and then extend to arbitrary orders, we utilize them to define a general operation for enriching RGI. Regarding the scope of this work, we focus on a mainstream class of deterministic positional and structural encodings that are equivariant under node relabeling and therefore admit a formulation as RGIs. Probabilistic symmetry-breaking encodings (Abboud et al., 2021; Eliasof et al., 2023; Lawrence et al., 2025), fall outside this scope.

3 Preliminaries

Notations. Let $G = (V, E)$ be a graph with n nodes, and fix a standard graph operator $T_G \in \{A_G, L_G, \widetilde{L}_G\}$, where A_G, L_G, and $\widetilde{L}_G$ denote the adjacency matrix, graph Laplacian, and symmetric normalized graph Laplacian, respectively. Let $\lambda_1^G < \cdots < \lambda_{m_G}^G$ be the distinct eigenvalues

of T, with respective multiplicities $d_1, \dots, d_{m_G}$. For each r, let U_r^G be a matrix whose d_r columns are orthonormal eigenvectors corresponding to λ_r^G, and let $P_r^G := U_r^G (U_r^G)^\top$ be the corresponding eigenspace projection matrix. With $U_G = [U_1^G, \dots, U_{m_G}^G]$ and $\Lambda_G = \mathrm{diag}(\lambda_1^G I_{d_1}, \dots, \lambda_{m_G}^G I_{d_{m_G}})$, the spectral decomposition of T_G is $T_G = U_G \Lambda_G U_G^\top = \sum_{r=1}^{m_G} \lambda_r^G P_r^G$. We write $\mathrm{spec}(T_G) := (\lambda_1^G, \dots, \lambda_{m_G}^G)$ for the ordered tuple of distinct eigenvalues. We further denote $[n] := \{1, \dots, n\}$.

Spectral Invariants. For a fixed standard graph operator T, spectral invariants are representations constructed from $\mathrm{spec}(T)$ and the eigenspace projection matrices $\{P_r^G\}_{r=1}^{m_G}$. These representations are independent of the choice of eigenbasis: $\mathrm{spec}(T)$ does not depend on this choice, and any orthonormal basis of the eigenspace corresponding to λ_r^G has the form $U_r^G Q$ for some orthogonal matrix Q, yielding $(U_r^G Q)(U_r^G Q)^\top = U_r^G Q Q^\top (U_r^G)^\top = P_r^G$. The eigenspace projection invariant (Zhang et al., 2024b) provides a unified node-pair representation of this spectral information:

$$\mathcal{P}^T(G, u, v) := \left\{ \left(\lambda_r^G, P_r^G(u, v)\right) : \; r = 1, \dots, m_G \right\}. \tag{1}$$

Rooted Graph Invariants. Let $\mathcal{G}_k$ denote the class of graphs equipped with an ordered k-tuple of distinguished vertices. More precisely, an element of $\mathcal{G}_k$ is of the form $G_{\mathbf{u}} = (G, \mathbf{u}) = (G, u_1, \dots, u_k)$, where the entries of $\mathbf{u}$ are not necessarily distinct. For $G_{\mathbf{u}}, H_{\boldsymbol{v}} \in \mathcal{G}_k$, we write $G_{\mathbf{u}} \cong H_{\boldsymbol{v}}$ if there exists a graph isomorphism $\phi : G \to H$ such that $\phi(u_t) = v_t$ for $t \in [k]$. Given an arbitrary set $\mathcal{S}$, a function $\Phi : \mathcal{G}_k \to \mathcal{S}$ is called a $\mathcal{S}$-valued k-rooted graph invariant if $G_{\mathbf{u}} \cong H_{\boldsymbol{v}} \implies \Phi(G_{\mathbf{u}}) = \Phi(H_{\boldsymbol{v}})$ for all $G_{\mathbf{u}}, H_{\boldsymbol{v}} \in \mathcal{G}_k$. See Figure 1 for an illustration. For example, the degree $\Phi_{\deg}(G, u) = \deg_G(u)$ is a real-valued 1-RGI, while the shortest-path distance $\Phi_{\mathrm{spd}}(G, u, v) = d_G(u, v)$ is a real-valued 2-RGI. Two k-RGIs Φ and Ψ are said to be equivalent, denoted by $\Phi \equiv \Psi$, if for any $(G, \mathbf{u}), (H, \boldsymbol{v}) \in \mathcal{G}_k$, $\Phi(G, \mathbf{u}) = \Phi(H, \boldsymbol{v}) \iff \Psi(G, \mathbf{u}) = \Psi(H, \boldsymbol{v})$.

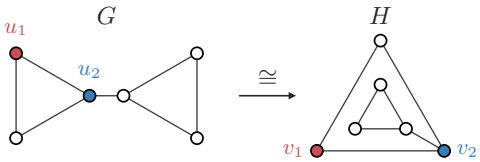

Figure 1: Illustration of an RGI. An isomorphism from G to H maps (u_1, u_2) to (v_1, v_2), preserving the value of Φ on corresponding node pairs.

GNN Expressiveness. We measure GNN expressiveness through the graph-distinguishing power of corresponding color-refinement algorithms (Xu et al., 2019; Zhang et al., 2024a). We write $\mathcal{A} \succeq \mathcal{B}$ if $\mathcal{A}$ distinguishes every graph pair distinguished by $\mathcal{B}$, with $\succ$ and $\equiv$ denoting strict and equal expressive power. In this work, we compare our refinements with three classical WL variants: folklore k-WL (k-FWL), oblivious k-WL (k-OWL), and local k-WL (k-LWL) (Morris et al., 2020). Their definitions and the relation k-FWL $\equiv (k+1)$-OWL for $k \geq 2$ are discussed in Appendix C.

4 Spectralization

We now formalize Spectralization. Since a real-valued k-RGI defines a structural signal on k-tuples of nodes, we begin by introducing higher-order spectral filtering for such signals. Throughout this section, we follow the notation introduced in Section 3.

A real-valued order-k graph signal is a function $\mathbf{X} : V(G)^k \to \mathbb{R}$, identified with its column-wise vectorization in $\boldsymbol{x} \in \mathbb{R}^{n^k}$. Define the order-$k$ Fourier basis as $U_G^{\otimes k} := U_G \otimes \cdots \otimes U_G$. The associated Fourier transform and its inverse are $\widehat{\boldsymbol{x}} = (U_G^\top)^{\otimes k} \boldsymbol{x}$ and $\boldsymbol{x} = U_G^{\otimes k} \widehat{\boldsymbol{x}}$. For a continuous spectral response $g : \mathbb{R}^k \to \mathbb{R}$, define the response tensor $\mathbf{G}_g^{(k)} := \left(g\big((\Lambda_G)_{i_k i_k}, \dots, (\Lambda_G)_{i_1 i_1}\big)\right)_{i_1, \dots, i_k \in [n]}$. The corresponding order-k spectral filter acts by pointwise multiplication in the Fourier domain:

$$\mathcal{F}_{G,g}^{(k)} := U_G^{\otimes k} \, \mathrm{diag}\left(\mathrm{vec}(\mathbf{G}_g^{(k)})\right) (U_G^\top)^{\otimes k}.$$

Proposition 4.1. *With P_r^G is the eigenspace projection matrix corresponding to λ_r,*

$$\mathcal{F}_{G,g}^{(k)} = g(T_G \otimes I_n^{\otimes(k-1)}, \cdots, I_n^{\otimes(k-1)} \otimes T_G) = \sum_{r_1, \dots, r_k = 1}^{m_G} g(\boldsymbol{\lambda}_{\boldsymbol{r}}^G) \, P_{\boldsymbol{r}}^G, \tag{2}$$

where $\boldsymbol{r} := (r_1, \dots, r_k)$, $\boldsymbol{\lambda}_{\boldsymbol{r}}^G := (\lambda_{r_1}^G, \dots, \lambda_{r_k}^G)$ and $P_{\boldsymbol{r}}^G := P_{r_1}^G \otimes \cdots \otimes P_{r_k}^G$.

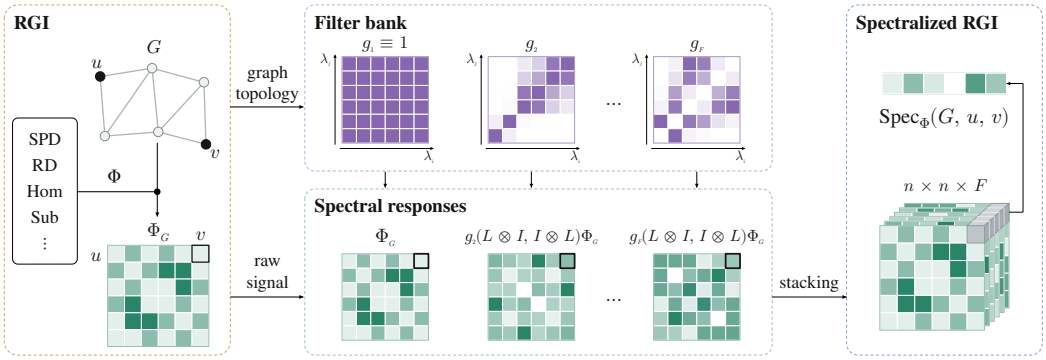

Figure 2: Spectralization of 2-RGIs, illustrated for a single chosen RGI Φ. The structural signal Φ_G on G is filtered using the graph Laplacian $T = L$ and a bank of bivariate spectral filters. The resulting spectral response matrices are stacked to collect the responses at each node pair (u, v). Together with the graph spectrum (omitted for clarity), these responses constitute the spectralized RGI. Moreover, note that Definition 4.3 uses all continuous bivariate filters, whereas the finite bank of F filters illustrated here corresponds to the restriction in Section 5.

Example 4.2. *Let T_G be the graph Laplacian L_G, the first-order case recover the conventional graph spectral filters $\mathcal{F}^{(1)}_{G,g} = g(L_G) = \sum_r g(\lambda^G_r) P^G_r$, while the second-order case introduced in Wang et al. (2026) takes the form $\mathcal{F}^{(2)}_{G,g} = g(L_G \otimes I_n,\ I_n \otimes L_G) = \sum_{r_1,r_2} g(\lambda^G_{r_1}, \lambda^G_{r_2}) P^G_{r_1} \otimes P^G_{r_2}$.*

4.1 Spectralization of Rooted Graph Invariants

Definition 4.3. Let Φ_k be a k-RGI, and let $C(\mathbb{R}^k)$ denote the space of continuous functions on $\mathbb{R}^k$. For graph G, write $\Phi_G := \Phi_k(G, \cdot)$. The *spectralization* of Φ_k is defined, for each $\boldsymbol{v} \in V(G)^k$, by

$$\mathrm{Spec}^T_{\Phi_k}(G, \boldsymbol{v}) := \Big(\mathrm{spec}(T),\ \big([\mathcal{F}^{(k)}_{G,g}\Phi_G](\boldsymbol{v})\big)_{g \in C(\mathbb{R}^k)}\Big).$$

Let $\mathfrak{C}_k = (\Phi_{k,\alpha})_{\alpha \in \mathcal{A}}$ be a possibly infinite family of real-valued k-RGIs. Its spectralization is defined componentwise by $\mathrm{Spec}^T_{\mathfrak{C}_k} := \big(\mathrm{Spec}^T_{\Phi_{k,\alpha}}\big)_{\alpha \in \mathcal{A}}$.

Figure 2 illustrates the spectralization of 2-RGIs. The following theorem gives an equivalent characterization of Spectralization in terms of eigenspace projections. In particular, it shows that spectralizing the identity RGI recovers the eigenspace projection invariant (1).

Theorem 4.4. *Let Φ_k be a k-RGI. With notations in Proposition 4.1 and Definition 4.3, define*

$$\mathcal{P}^T_{\Phi_k}(G, \boldsymbol{v}) := \left\{ \left(\boldsymbol{\lambda}^G_{\boldsymbol{r}}, \left[P^G_{\boldsymbol{r}}\Phi_G\right](\boldsymbol{v})\right) : \boldsymbol{r} \in [m_G]^k \right\}.$$

Then $\mathrm{Spec}^T_{\Phi_k} \equiv \mathcal{P}^T_{\Phi_k}$. In particular, for $k = 2$, let $\Phi_{\mathrm{id}}(G, u, v) := \mathbf{1}[u = v]$ be the identity RGI. Then its spectralization is equivalent to the eigenspace projection invariant (1), namely,

$$\mathrm{Spec}^T_{\Phi_{\mathrm{id}}} \equiv \mathcal{P}^T_{\Phi_{\mathrm{id}}} \equiv \mathcal{P}^T.$$

This provides a new perspective on the expressive power of spectral invariant GNNs. Initially, the identity RGI carries limited structural information and can be written as the sum of its eigenspace projection components: $\Phi_{\mathrm{id}} = \sum P^G_{\boldsymbol{r}}\Phi_{\mathrm{id}}$. By retaining these components separately, Spectralization refines the identity RGI and can distinguish node pairs with identical raw values but different projection components, thereby capturing richer structural information that can enhance GNN expressiveness. Since $\{P^G_{\boldsymbol{r}}\}_{\boldsymbol{r}}$ are determined by graph topology, $\{P^G_{\boldsymbol{r}}\Phi_{\mathrm{id}}\}_{\boldsymbol{r}}$ then capture the signal's interactions with graph topology, giving a concrete interpretation of spectral enrichment. More importantly, the same operation applies to arbitrary RGIs, allowing more informative structural signals to undergo this enrichment. This generalizes spectral invariants, yielding a broader class of *spectralized RGIs* that remain independent of the choice of eigenbasis and can capture richer structural information.

Examples of RGIs. Beyond the identity signal, shortest-path and resistance distances define 2-RGIs on connected graphs. Rooted homomorphism counts $\Phi^{\mathrm{hom}}_{F_{\boldsymbol{r}}}$ count maps from a template F to G that send its ordered roots $\boldsymbol{r}$ to the prescribed tuple; rooted subgraph counts $\Phi^{\mathrm{sub}}_{F_{\boldsymbol{r}}}$ count copies containing the prescribed rooted pattern. These provide RGIs of arbitrary order, with detailed definitions and examples in Appendix C.2.

4.2 Spectralized Weisfeiler-Leman Refinement

We next focus on atomic types, which are RGIs and represent tuple information in WL algorithms. Their spectralization connects the general construction above to the study of WL expressiveness.

Atomic types and FWL. An atomic type records the equality and adjacency relations among the entries of an ordered node tuple. Let $\mathbf{AT}_k$ be the finite set of possible atomic types of k-tuples and let $\Phi_\tau(G, \boldsymbol{v})$ indicate whether $\boldsymbol{v}$ has type $\tau \in \mathbf{AT}_k$. These indicators form the family $\mathcal{AT}_k = (\Phi_\tau)_{\tau\in\mathbf{AT}_k}$ of k-RGIs, with one-hot encoding $\mathbf{atp}_G(\boldsymbol{v}) = \mathcal{AT}_k(G, \boldsymbol{v})$. In the following discussion, we use the augmented form of k-FWL (12), which has the same graph-distinguishing power as the standard form (8) for $k \geq 2$. For $k = 1$, the corresponding all-node update is likewise equivalent to 1-WL (Zhang et al., 2024b). The full formulation and proofs are given in Appendix C.3.

We now apply Spectralization to the atomic-type family $\mathcal{AT}_k$ and define the spectralized k-FWL refinement by replacing every occurrence of $\mathbf{atp}$ with its spectralized counterpart.

Definition 4.5. Choose a standard graph operator $T \in \{A, L, \widetilde{L}\}$, the *spectralized atomic type* $\mathbf{Satp}$ of k-tuples is the spectralization $\mathrm{Spec}^T_{\mathcal{AT}_k}$, namely,

$$\mathbf{Satp}_G(\boldsymbol{v}) \coloneqq \mathrm{Spec}^T_{\mathcal{AT}_k}(G, \boldsymbol{v}) = \left(\mathrm{Spec}^T_{\Phi_\tau}(G, \boldsymbol{v})\right)_{\tau\in\mathbf{AT}_k}.$$

Definition 4.6. For $k \geq 2$, the *Spectralized k-FWL refinement*, abbreviated by Spec k-FWL, is

$$\widetilde{\chi}^{(0)}_G(\boldsymbol{v}) = \mathbf{Satp}_G(\boldsymbol{v}), \qquad \widetilde{\chi}^{(t+1)}_G(\boldsymbol{v}) = \mathrm{hash}\left(\widetilde{\chi}^{(t)}_G(\boldsymbol{v}), \widetilde{M}^{(t)}_G(\boldsymbol{v})\right),$$

where $\widetilde{M}^{(t)}_G(\boldsymbol{v}) = \left\{\!\!\left\{\left(\widetilde{\chi}^{(t)}_G(\boldsymbol{v}^{\,1\leftarrow w}), \ldots, \widetilde{\chi}^{(t)}_G(\boldsymbol{v}^{\,k\leftarrow w}), \mathbf{Satp}_G(\boldsymbol{v}, w)\right) : w \in V(G)\right\}\!\!\right\}$,

where hash is a perfect hashing; $\boldsymbol{v}^{\,i\leftarrow w}$ denote the tuple obtained by replacing its i-th entry with w. For $k = 1$, Spec 1-WL analogously update all nodes with $\mathbf{Satp}_G(v, w)$, see Appendix C.3.

We then compare Spec k-FWL with FWL, OWL, and LWL refinements and the higher-order spectral invariant WL refinement of Gai et al. (2025). The latter enriches coordinate-wise k-OWL (9) aggregation with eigenspace projection invariant; we call it SI k-OWL (11), with its complete definition in Appendix C.3. Its order-one case is EPWL (Zhang et al., 2024b), also denoted SI 1-WL. Moreover, all comparisons involving SI refinements use $T = A$, consistent with Gai et al. (2025).

For $k \geq 2$, the following theorem shows that Spec k-FWL is strictly more expressive than k-FWL and bounded above by $(k+1)$-LWL. It also places SI k-OWL strictly between k-LWL and k-FWL. Combining Theorem 4.7 and Proposition 4.9, we obtain the expressiveness hierarchy in Figure 3.

Theorem 4.7. *For every $k \geq 2$,*

(i) k-FWL $\prec$ Spec k-FWL $\preceq$ $(k+1)$-LWL

(ii) k-LWL $\prec$ SI k-OWL $\prec$ k-FWL

Example 4.8. *The Shrikhande graph and the 4×4 rook graph, a classical pair of non-isomorphic strongly regular graphs, are known to be indistinguishable by 2-FWL (Papp & Wattenhofer, 2022). In contrast, Spec 2-FWL distinguishes them, see Corollary D.4.*

Proposition 4.9. *Spec 1-WL $\equiv$ SI 1-WL $\prec$ 2-LWL.*

Beyond atomic types, we investigate whether Spectralization also improves the expressive power of WL refinement algorithms augmented with other 2-RGIs, which can be used as relative positional encodings or pairwise structural encodings. To compare each RGI with its spectralization, we use the all-node formulation of 1-WL (13) as a common backbone. Since Spectralization already includes the graph spectrum, we supply the same information to both variants through their initial node

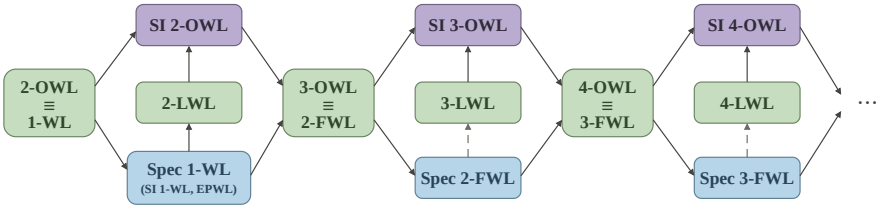

Figure 3: Expressiveness hierarchy of WL variants. Green, purple, and blue boxes denote classical, spectral invariant (SI), and spectralized (Spec) refinement algorithms, respectively. A solid arrow $\mathcal{A} \to \mathcal{B}$ indicates $\mathcal{A} \prec \mathcal{B}$, while a dashed arrow $\mathcal{A} \dashrightarrow \mathcal{B}$ indicates $\mathcal{A} \preceq \mathcal{B}$.

colors to ensure a fair comparison. We then augment the backbone with the pairwise RGI values. Specifically, for a 2-RGI Φ, define $\chi^{(0)}_{\Phi,G}(u) := \mathrm{spec}(A_G)$,

$$\chi^{(t+1)}_{\Phi,G}(u) = \mathrm{hash}\left(\chi^{(t)}_{\Phi,G}(u), \left\{\!\!\left\{ \left(\chi^{(t)}_{\Phi,G}(v), \mathbf{atp}_G(u,v), \Phi(G,u,v)\right) : v \in V(G) \right\}\!\!\right\}\right).$$

We denote this refinement by Aug_Φ 1-WL. Replacing Φ with Spec^A_Φ gives the corresponding refinement using spectralized RGI values. Since the constant filter $g \equiv 1$ recovers the raw RGI signal, this replacement cannot decrease expressiveness. The following theorem establishes strict gains for individual RGIs from three classes.

Theorem 4.10. *Let $F_{\boldsymbol{r}} = (F, \boldsymbol{r})$ range over all connected template graphs with $2 \le |V(F)| \le 5$ and an ordered pair of distinct roots $\boldsymbol{r} = (r_1, r_2) \in V(F)^2$. The strict relation*

$$\mathrm{Aug}_\Phi \text{ 1-}WL \prec \mathrm{Aug}_{\mathrm{Spec}^A_\Phi} \text{ 1-}WL$$

holds for each individual 2-RGI Φ in the following three classes:

(i) Rooted homomorphism counts $\Phi^{\mathrm{hom}}_{F_{\boldsymbol{r}}}$,

(ii) Rooted subgraph counts $\Phi^{\mathrm{sub}}_{F_{\boldsymbol{r}}}$,

(iii) Shortest-path distance, diffusion distance at time $t = 1$, and biharmonic distance.

The comparisons in (iii) do not involve the template $F_{\boldsymbol{r}}$ and are restricted to connected input graphs.

5 Structural Encodings via Spectralization

We use Spectralization to construct structural encodings. Full Spectralization records responses from all continuous filters, whereas practical implementations can only use finitely many filters. In the following, we formalize the restriction of Spectralization to a filter bank and then give sufficient conditions under which the restricted version retains all the information contained in Spectralization.

Definition 5.1. Let $\mathcal{B} = (g_b)_{b=1}^F \subseteq C(\mathbb{R}^k)$ be a finite filter bank and Φ_k be a k-RGI. For each graph G, write $\Phi_G := \Phi_k(G, \cdot)$. The *$\mathcal{B}$-restricted spectralization* of Φ_k is defined, for every $\boldsymbol{v} \in V(G)^k$, by

$$\mathrm{Spec}^T_{\Phi_k}\big|_{\mathcal{B}}(G, \boldsymbol{v}) := \left(\mathrm{spec}(T), \left([\mathcal{F}^{(k)}_{G,g_b}\Phi_G](\boldsymbol{v})\right)_{b=1}^F\right).$$

Proposition 5.2. *Using the notation of Proposition 4.1, define the evaluation matrix of $\mathcal{B}$ by $A^{\mathcal{B}}_G := \left(g_b(\boldsymbol{\lambda}^G_{\boldsymbol{r}})\right)_{b\in[F],\,\boldsymbol{r}\in[m_G]^k} \in \mathbb{R}^{F\times m_G^k}$. Fix a graph G. If $\mathrm{rank}\, A^{\mathcal{B}}_G = m^k_G$, then, for every real-valued k-RGI Φ_k and every $\boldsymbol{v} \in V(G)^k$, $\mathrm{Spec}^T_{\Phi_k}\big|_{\mathcal{B}}(G, \boldsymbol{v})$ determines $\mathrm{Spec}^T_{\Phi_k}(G, \boldsymbol{v})$.*

Consequently, for any graph family whose graph operators have at most M distinct eigenvalues and any fixed order k, there exists a finite, graph-independent filter bank $\mathcal{B}_k(M)$ such that the $\mathcal{B}_k(M)$-restricted spectralization exactly recovers full spectralization for every real-valued k-RGI Φ_k and every graph in the family. Nonetheless, Proposition 5.2 implies that exact recovery may require a prohibitively large number of filters. To balance expressiveness and scalability, we therefore combine a small set of fixed filters with clear structural interpretations, such as the constant filter $g \equiv 1$

that preserves the raw RGI signal, with learnable filters that adapt to the downstream task. We consider two approaches to parameterizing the learnable filters, depending on whether spectral decomposition of the graph operator is computationally feasible. The following discussion focuses on 2-RGIs; the general case is presented in Appendix E.

Implementation via Spectral Decomposition. When spectral decomposition is computationally feasible, we directly parameterize the spectral response functions of the learnable filters. Specifically, let $\mathcal{B}_{\text{learn}} = (g_{b,\theta})_{b=1}^{F_{\text{learn}}}$ be a bank of F_{learn} learnable filters, jointly modeled by an MLP $f_\theta : \mathbb{R}^2 \to \mathbb{R}^{F_{\text{learn}}}$ shared across graphs. The bth output of f_θ defines the filter $g_{b,\theta}$.

To apply this filter bank to a real-valued 2-RGI Φ_2, consider a graph G with n vertices and identify $\Phi_G \coloneqq \Phi_2(G, \cdot)$ with its $n \times n$ signal matrix. Using the spectral decomposition $T = U_G \Lambda_G U_G^\top$, we evaluate each filter on pairs of eigenvalues to form a spectral response matrix $S_{b,G}$. This matrix scales the Fourier coefficients of Φ_G, giving the filtered RGI signal $\mathcal{F}^{(2)}_{G,g_b}\Phi_G$ after the inverse transform:

$$Y_{b,G} \coloneqq \mathcal{F}^{(2)}_{G,g_b}\Phi_G = g_b(T \otimes I_n, I_n \otimes T)\Phi_G = U_G \left[S_{b,G} \odot \left(U_G^\top \Phi_G U_G\right)\right] U_G^\top,$$

where $(S_{b,G})_{ij} = g_{b,\theta}\big((\Lambda_G)_{jj}, (\Lambda_G)_{ii}\big)$ and $\odot$ denotes entry-wise multiplication. For each node pair (u, v), we then collect $Y_{b,G}(u, v)$ across the learnable filters together with the fixed-filter responses into a response vector, which serves as the input for constructing its structural encoding.

Implementation via Polynomial Filters. When spectral decomposition is computationally prohibitive, we parameterize the learnable filters as bivariate polynomials of bounded total degree. Specifically, let $\mathbb{P}_{\le K}(\mathbb{R}^2)$ denote the space of bivariate polynomials of total degree at most K. Choose fixed univariate polynomials $p_0, \dots, p_K$ with $\deg p_\ell = \ell$ and $p_0 \equiv 1$, such as the Chebyshev polynomials. Their products define the *polynomial basis filter bank* $\mathcal{B}_{2,K} \coloneqq \{p_i(x)p_j(y) : 0 \le i + j \le K\}$, where i and j range over nonnegative integers.

To apply this filter bank to a real-valued 2-RGI Φ_2, identify $\Phi_G \coloneqq \Phi_2(G, \cdot)$ with its signal matrix. Responses to all filters in $\mathbb{P}_{\le K}(\mathbb{R}^2)$ can be obtained by linearly combining the responses to this filter bank. For every $g \in \mathbb{P}_{\le K}(\mathbb{R}^2)$, its response has the form

$$\mathcal{F}^{(2)}_{G,g}\Phi_G = \sum_{0 \le i+j \le K} w_{i,j}\, p_j(T)\Phi_G p_i(T), \tag{3}$$

with unique polynomial-basis coefficients $w_{i,j}$; the full statement is in Proposition E.3. Therefore, when constructing structural encodings, we can precompute those spectral responses $p_j(T)\Phi_G p_i(T)$ and then mix them through a learnable linear layer with F_{learn} output channels. Each output channel can then realize the response of any polynomial filter in $\mathbb{P}_{\le K}(\mathbb{R}^2)$. During training, learning these filters therefore requires only optimizing the linear layer, without recomputing the basis responses.

Spectralized Structural Encodings. We now describe how to construct a *spectralized structural encoding*, abbreviated as SpecSE. First, we select a finite family of 2-RGIs $(\Phi_\alpha)_{\alpha \in \mathcal{A}}$, whose values at each node pair already form an edge-level structural encoding. We then enrich this encoding by applying a bank of learnable spectral filters $(g_b)_{b=1}^{F_{\text{learn}}}$ to each RGI, using either of the two implementations described above. Alongside these learnable filters, we retain the constant filter $g \equiv 1$ as the sole fixed filter to preserve the raw RGI values. Concatenating the raw values and learned spectral responses across all selected RGIs gives

$$\text{SpecSE}(G, u, v) \coloneqq \Big\|_{\alpha \in \mathcal{A}} \left[\Phi_\alpha(G, u, v) \,\Big\|\, \left(\left[\mathcal{F}^{(2)}_{G,g_b}\Phi_\alpha(G, \cdot)\right](u, v)\right)_{b=1}^{F_{\text{learn}}}\right], \tag{4}$$

where $\|$ denotes concatenation. Pseudocode for both implementations is provided in Appendix E.3.

6 Experiments

We benchmark Spectralization on both synthetic and real-world datasets. On the synthetic dataset, graph isomorphism (GI) testing assesses the expressiveness gains predicted by our theoretical analysis. On real-world datasets, we evaluate the effectiveness of spectralized structural encodings for molecular property regression and image classification. We further examine their performance across different neural architectures.

Table 1: Pair distinguishing results of different Weisfeiler–Leman variants and corresponding models on BREC. The Overall column reports the number of distinguished graph pairs, with accuracy in parentheses. Equal results on BREC do not imply theoretical equivalence.

Model	WL	Basic (60)	Regular (140)	Extension (100)	CFI (100)	Overall (400)
Non-GNN	1-WL	0	0	0	0	0 (0%)
	SPD-WL	16	14	41	12	83 (20.8%)
	SI 1-WL (EPWL)	60	50	100	26	236 (59.0%)
	Spec 1-WL	60	50	100	26	236 (59.0%)
	2-LWL	60	50	100	57	267 (66.8%)
	SI 2-OWL	60	50	100	60	270 (67.5%)
	2-FWL	60	50	100	60	270 (67.5%)
	Spec 2-FWL	60	137	100	70	367 (91.8%)
GIN	1-WL	0.0	0.0	0.0	0.0	0.0 (0%)
Graphormer	SPD-WL	16.0	12.0	41.0	10.0	79.0 (19.8%)
EPNN	EPWL	60.0	50.0	100.0	5.0	215.0 (53.8%)
PPGN	2-FWL	60.0	50.0	100.0	20.3	230.3 (57.6%)
SpecPPGN	Spec 2-FWL	60.0	93.7	100.0	38.0	291.7 (72.9%)

Datasets. For GI testing, we use BREC (Wang & Zhang, 2024), which contains 400 non-isomorphic graph pairs grouped into four categories: Basic, Regular, Extension, and CFI. For downstream prediction, we conduct molecular property regression on QM9 (Wu et al., 2018) and ZINC (Irwin et al., 2012), and image classification on CIFAR10 (Krizhevsky et al., 2009). Further dataset details are provided in Appendix F.

Baselines. For GI testing, we follow Wang & Zhang (2024); Zhang et al. (2024b); Gai et al. (2025), comparing against representative WL variants (1-WL, SPD-WL, EPWL, 2-LWL, SI 2-OWL, and 2-FWL) and neural models, including GIN(Xu et al., 2019), Graphormer(Ying et al., 2021), EPNN(Zhang et al., 2024b), and PPGN(Maron et al., 2019a). For downstream prediction, we follow the experimental setup of Bao et al. (2025). On QM9, we compare SpecSE with random-walk structural encoding (Dwivedi et al., 2022, RWSE) and motif structural encoding (Bao et al., 2025, MoSE), and with several established GNN models (Brockschmidt, 2020; Alon & Yahav, 2021; Abboud et al., 2022; Dimitrov et al., 2023). On ZINC-12K and CIFAR10, we further compare these structural encodings using backbone architectures ranging from the multilayer perceptron baseline (MLP-E), GIN-E and its virtual-node extension GIN-E+VN (Hu et al., 2020), and GPS (Rampášek et al., 2022).

6.1 Graph Isomorphism Testing

We evaluate both the spectralized WL refinement algorithms and their neural implementations on BREC. Table 1 reports the number of distinguished graph pairs in each category and overall. Our reproductions exactly match the published results for non-neural baselines, while yielding slightly lower but comparable results for neural baselines. We therefore use the baseline results reported in Wang & Zhang (2024). In our experiments, Spec 2-FWL distinguishes 367 of the 400 pairs, compared with 270 for 2-FWL, which is consistent with the strict expressiveness gain established in Theorem 4.7. The neural implementations show a similar trend, with results averaged over 10 runs. SpecPPGN, which incorporates spectralized atomic-type features into PPGN, distinguishes an average of 291.7 pairs, outperforming PPGN and indicating that the benefits of Spectralization also extend to learned models.

6.2 Downstream Prediction

We evaluate the effectiveness of SpecSE on molecular property regression and image classification, reporting mean absolute error (MAE) on ZINC-12K and QM9 and test accuracy on CIFAR10. Tables 2 and 3 summarize the results. We use the baseline results reported by Bao et al. (2025), as our reproductions closely match their reported performance.

Table 2: MAE on QM9. Best, second-best, and third-best results are marked in red, blue, and olive, respectively.

Property	R-GIN	R-GIN+FA	R-SPN	E-BasePlanE	GPS		
					RWSE	MoSE	SpecSE (ours)
μ	$2.64_{\pm 0.11}$	$2.54_{\pm 0.09}$	$2.21_{\pm 0.21}$	$1.97_{\pm 0.03}$	$1.47_{\pm 0.02}$	$1.43_{\pm 0.02}$	$1.45_{\pm 0.02}$
α	$4.67_{\pm 0.52}$	$2.28_{\pm 0.04}$	$1.66_{\pm 0.06}$	$1.63_{\pm 0.01}$	$1.52_{\pm 0.27}$	$1.48_{\pm 0.16}$	$1.28_{\pm 0.03}$
HOMO	$1.42_{\pm 0.01}$	$1.26_{\pm 0.02}$	$1.20_{\pm 0.08}$	$1.15_{\pm 0.01}$	$0.91_{\pm 0.01}$	$0.91_{\pm 0.01}$	$0.90_{\pm 0.01}$
LUMO	$1.50_{\pm 0.09}$	$1.34_{\pm 0.04}$	$1.20_{\pm 0.06}$	$1.06_{\pm 0.02}$	$0.90_{\pm 0.06}$	$0.86_{\pm 0.01}$	$0.85_{\pm 0.01}$
Gap	$2.27_{\pm 0.09}$	$1.96_{\pm 0.04}$	$1.77_{\pm 0.06}$	$1.73_{\pm 0.02}$	$1.47_{\pm 0.02}$	$1.45_{\pm 0.02}$	$1.43_{\pm 0.02}$
R^2	$15.63_{\pm 1.40}$	$12.61_{\pm 0.37}$	$10.63_{\pm 1.01}$	$10.53_{\pm 0.55}$	$6.11_{\pm 0.16}$	$6.22_{\pm 0.19}$	$6.09_{\pm 0.07}$
ZPVE	$12.93_{\pm 1.81}$	$5.03_{\pm 0.36}$	$2.58_{\pm 0.13}$	$2.81_{\pm 0.16}$	$2.63_{\pm 0.44}$	$2.43_{\pm 0.27}$	$2.58_{\pm 0.10}$
U_0	$5.88_{\pm 1.01}$	$2.21_{\pm 0.12}$	$0.89_{\pm 0.05}$	$0.95_{\pm 0.04}$	$0.83_{\pm 0.17}$	$0.85_{\pm 0.08}$	$0.80_{\pm 0.05}$
U	$18.71_{\pm 23.36}$	$2.32_{\pm 0.18}$	$0.93_{\pm 0.03}$	$0.94_{\pm 0.04}$	$0.83_{\pm 0.15}$	$0.75_{\pm 0.03}$	$0.78_{\pm 0.06}$
H	$5.62_{\pm 0.81}$	$2.26_{\pm 0.19}$	$0.92_{\pm 0.03}$	$0.92_{\pm 0.04}$	$0.86_{\pm 0.15}$	$0.83_{\pm 0.09}$	$0.75_{\pm 0.03}$
G	$5.38_{\pm 0.75}$	$2.04_{\pm 0.24}$	$0.83_{\pm 0.05}$	$0.88_{\pm 0.04}$	$0.83_{\pm 0.12}$	$0.80_{\pm 0.14}$	$0.73_{\pm 0.05}$
C_v	$3.53_{\pm 0.37}$	$1.86_{\pm 0.03}$	$1.23_{\pm 0.06}$	$1.20_{\pm 0.06}$	$1.25_{\pm 0.05}$	$1.02_{\pm 0.04}$	$0.94_{\pm 0.01}$
Ω	$1.05_{\pm 0.11}$	$0.80_{\pm 0.04}$	$0.52_{\pm 0.02}$	$0.45_{\pm 0.01}$	$0.39_{\pm 0.02}$	$0.38_{\pm 0.01}$	$0.36_{\pm 0.01}$

On QM9, we follow Bao et al. (2025) in choosing all connected graphs with at most five vertices, together with the six-cycle, as templates. We compute the two-rooted homomorphism counts based on these templates as 2-RGIs and spectralize them via the spectral decomposition implementation in Section 5 to construct SpecSE (4).

Table 3: Performance comparison of different structural encodings on ZINC-12K and CIFAR10.

Dataset	Model	Encoding Type		
		RWSE	MoSE	SpecSE (ours)
ZINC-12K	MLP-E	$0.361_{\pm 0.010}$	$0.347_{\pm 0.003}$	$\mathbf{0.329}_{\pm \mathbf{0.006}}$
	GIN-E	$0.122_{\pm 0.003}$	$0.118_{\pm 0.007}$	$\mathbf{0.107}_{\pm \mathbf{0.004}}$
	GINE+VN	$0.085_{\pm 0.003}$	$0.068_{\pm 0.004}$	$\mathbf{0.066}_{\pm \mathbf{0.002}}$
	GPS	$0.069_{\pm 0.001}$	$0.062_{\pm 0.002}$	$\mathbf{0.061}_{\pm \mathbf{0.005}}$
CIFAR10	GPS	$71.96_{\pm 0.40}$	$73.50_{\pm 0.44}$	$\mathbf{74.27}_{\pm \mathbf{0.40}}$

For ZINC-12K, we instead use the template family $\mathrm{Spasm}(C_7) \cup \mathrm{Spasm}(C_8)$ considered by Bao et al. (2025); Jin et al. (2024). This family consists of the loop-free quotient graphs obtained from the seven- and eight-cycles by identifying vertices, up to isomorphism. For CIFAR10, we again follow Bao et al. (2025) to use all connected graphs with at most five vertices as templates. We likewise spectralize the two-rooted homomorphism counts of these templates to construct SpecSE (4), but use the polynomial filter implementation for these datasets. For each model, we adopt the corresponding MoSE-based configuration (Bao et al., 2025), retaining its backbone layer width and depth, training protocol, data splits, and random seeds, while replacing the MoSE encoding with SpecSE.

On QM9, GPS with SpecSE achieves lower mean MAE than GPS with RWSE on all 13 targets and the lowest mean MAE among all compared methods on 10 targets. On ZINC-12K, SpecSE achieves the lowest mean MAE among the evaluated encodings for MLP-E, GIN-E, and GINE+VN. On CIFAR10, GPS with SpecSE also outperforms its counterparts using RWSE and MoSE. Together, these results demonstrate the effectiveness of SpecSE across molecular and image datasets and multiple neural architectures.

7 Conclusion

We introduced Spectralization, which generalizes spectral invariants by extending spectral enrichment from the identity signal to arbitrary RGIs. We proved that this operation strictly strengthens k-FWL for every $k \geq 2$ and provides expressiveness gains for refinements augmented with several other classes of RGIs. We further developed practical implementations that turn these enriched signals into structural encodings. Experiments on graph isomorphism testing corroborate the expressiveness gains, while downstream evaluations demonstrate the effectiveness of the resulting encodings across multiple tasks and neural architectures.

AI USE STATEMENT

In this work, we used generative AI tools to assist with polishing writing, searching literature on positional and structural encodings, parts of the mathematical proofs (particularly part (ii) of Theorem 4.7 and Theorem 4.10), and coding. We have not used generative AI tools to generate the research ideas, draft sections of the paper, or generate synthetic datasets. All AI-assisted mathematical proofs were independently reviewed by the authors, and all AI-assisted code was verified and tested for correctness. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

The assumptions underlying our theoretical results are stated in the corresponding definitions and results, and complete proofs are provided in Appendices D and E. Section 5 describes the construction and implementation of SpecSE. Appendix F provides experimental details, including the construction and preprocessing of input signals, model configurations, training settings, and evaluation protocols. Our source code and experiment configurations are available at this repository.

REFERENCES

Ralph Abboud, Ismail Ilkan Ceylan, Martin Grohe, and Thomas Lukasiewicz. The surprising power of graph neural networks with random node initialization, 2021. URL https://openreview.net/forum?id=L7Irrt5sMQa.

Ralph Abboud, Radoslav Dimitrov, and Ismail Ilkan Ceylan. Shortest path networks for graph property prediction. In *Learning on graphs conference*, pp. 5–1. PMLR, 2022.

Louis Airale, Antonio Longa, Mattia Rigon, Andrea Passerini, and Roberto Passerone. Simple path structural encoding for graph transformers. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 857–873. PMLR, 13–19 Jul 2025.

Uri Alon and Eran Yahav. On the bottleneck of graph neural networks and its practical implications. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=i80OPhOCVH2.

Juan P Amboage, Ernst Roell, Patrick Schnider, and Bastian Rieck. Leap: Local ect-based learnable positional encodings for graphs. In *International Conference on Learning Representations*, volume 2026, pp. 41977–41995, 2026.

Vikraman Arvind, Frank Fuhlbrück, Johannes Köbler, and Oleg Verbitsky. On a hierarchy of spectral isomorphism invariants. *computational complexity*, 34(2):19, 2025.

Alan Baker. *Transcendental number theory*. Cambridge university press, 2022.

Linus Bao, Emily Jin, Michael Bronstein, Ismail I Ceylan, and Matthias Lanzinger. Homomorphism counts as structural encodings for graph learning. In *International Conference on Learning Representations*, volume 2025, pp. 96873–96901, 2025.

Pablo Barceló, Floris Geerts, Juan Reutter, and Maksimilian Ryschkov. Graph neural networks with local graph parameters. *Advances in Neural Information Processing Systems*, 34:25280–25293, 2021.

Mitchell Black, Zhengchao Wan, Gal Mishne, Amir Nayyeri, and Yusu Wang. Comparing graph transformers via positional encodings. In *Forty-first International Conference on Machine Learning*, 2024. URL https://openreview.net/forum?id=va3r3hSA6n.

Giorgos Bouritsas, Fabrizio Frasca, Stefanos Zafeiriou, and Michael M Bronstein. Improving graph neural network expressivity via subgraph isomorphism counting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1):657–668, 2022.

Marc Brockschmidt. Gnn-film: Graph neural networks with feature-wise linear modulation. In *International conference on machine learning*, pp. 1144–1152. PMLR, 2020.

Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. Spectral networks and locally connected networks on graphs. *arXiv preprint arXiv:1312.6203*, 2013.

Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in neural information processing systems*, 29, 2016.

Holger Dell, Martin Grohe, and Gaurav Rattan. Lovász meets weisfeiler and leman. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, volume 107 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 40:1–40:14. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018.

Radoslav Dimitrov, Zeyang Zhao, Ralph Abboud, and Ismail Ceylan. Plane: Representation learning over planar graphs. *Advances in Neural Information Processing Systems*, 36:16028–16054, 2023.

Vijay Prakash Dwivedi and Xavier Bresson. A generalization of transformer networks to graphs. *arXiv preprint arXiv:2012.09699*, 2020.

Vijay Prakash Dwivedi, Anh Tuan Luu, Thomas Laurent, Yoshua Bengio, and Xavier Bresson. Graph neural networks with learnable structural and positional representations. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=wTTjnvGphYj.

Vijay Prakash Dwivedi, Chaitanya K Joshi, Anh Tuan Luu, Thomas Laurent, Yoshua Bengio, and Xavier Bresson. Benchmarking graph neural networks. *Journal of Machine Learning Research*, 24(43):1–48, 2023.

Moshe Eliasof, Fabrizio Frasca, Beatrice Bevilacqua, Eran Treister, Gal Chechik, and Haggai Maron. Graph positional encoding via random feature propagation. In *International conference on machine learning*, pp. 9202–9223. PMLR, 2023.

Or Feldman, Amit Boyarski, Shai Feldman, Dani Kogan, Avi Mendelson, and Chaim Baskin. Weisfeiler and leman go infinite: Spectral and combinatorial pre-colorings. *Transactions on Machine Learning Research*, 2023. ISSN 2835-8856. URL https://openreview.net/forum?id=YJDqQSAuB6.

Jiarui Feng, Yixin Chen, Fuhai Li, Anindya Sarkar, and Muhan Zhang. How powerful are k-hop message passing graph neural networks. *Advances in neural information processing systems*, 35: 4776–4790, 2022.

Martin Fürer. On the power of combinatorial and spectral invariants. *Linear algebra and its applications*, 432(9):2373–2380, 2010.

Jingchu Gai, Yiheng Du, Bohang Zhang, Haggai Maron, and Liwei Wang. Homomorphism expressivity of spectral invariant graph neural networks. In *The Thirteenth International Conference on Learning Representations*, 2025.

Martin Grohe. The logic of graph neural networks. In *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '21. IEEE Press, 2021. ISBN 9781665448956. doi: 10.1109/LICS52264.2021.9470677. URL https://doi.org/10.1109/LICS52264.2021.9470677.

Weihua Hu, Matthias Fey, Marinka Zitnik, Yuxiao Dong, Hongyu Ren, Bowen Liu, Michele Catasta, and Jure Leskovec. Open graph benchmark: Datasets for machine learning on graphs. *Advances in neural information processing systems*, 33:22118–22133, 2020.

Yinan Huang, William Lu, Joshua Robinson, Yu Yang, Muhan Zhang, Stefanie Jegelka, and Pan Li. On the stability of expressive positional encodings for graphs. In *International Conference on Learning Representations*, volume 2024, pp. 39745–39774, 2024.

John J Irwin, Teague Sterling, Michael M Mysinger, Erin S Bolstad, and Ryan G Coleman. Zinc: a free tool to discover chemistry for biology. *Journal of chemical information and modeling*, 52 (7):1757–1768, 2012.

Mattie Ji, Amauri Souza, and Vikas Garg. Graph persistence goes spectral. *Advances in Neural Information Processing Systems*, 38:160642–160680, 2026.

Emily Jin, Michael M. Bronstein, Ismail Ilkan Ceylan, and Matthias Lanzinger. Homomorphism counts for graph neural networks: All about that basis. 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. 22075–22098. PMLR, 21–27 Jul 2024.

Devin Kreuzer, Dominique Beaini, Will Hamilton, Vincent Létourneau, and Prudencio Tossou. Rethinking graph transformers with spectral attention. *Advances in neural information processing systems*, 34:21618–21629, 2021.

Alex Krizhevsky et al. Learning multiple layers of features from tiny images. 2009.

Hannah Lawrence, Vasco Portilheiro, Yan Zhang, and Sékou-Oumar Kaba. Improving equivariant networks with probabilistic symmetry breaking. In *International Conference on Learning Representations*, volume 2025, pp. 80179–80206, 2025.

Pan Li, Yanbang Wang, Hongwei Wang, and Jure Leskovec. Distance encoding: Design provably more powerful neural networks for graph representation learning. *Advances in Neural Information Processing Systems*, 33:4465–4478, 2020.

Derek Lim, Joshua David Robinson, Lingxiao Zhao, Tess Smidt, Suvrit Sra, Haggai Maron, and Stefanie Jegelka. Sign and basis invariant networks for spectral graph representation learning. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=Q-UHqMorzil.

Yuankai Luo, Hongkang Li, Lei Shi, and Xiao-Ming Wu. Enhancing graph transformers with hierarchical distance structural encoding. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, NIPS '24, Red Hook, NY, USA, 2024. Curran Associates Inc. ISBN 9798331314385.

Liheng Ma, Chen Lin, Derek Lim, Adriana Romero-Soriano, Puneet K Dokania, Mark Coates, Philip Torr, and Ser-Nam Lim. Graph inductive biases in transformers without message passing. In *International Conference on Machine Learning*, pp. 23321–23337. PMLR, 2023.

Haggai Maron, Heli Ben-Hamu, Hadar Serviansky, and Yaron Lipman. Provably powerful graph networks. *Advances in neural information processing systems*, 32, 2019a.

Haggai Maron, Heli Ben-Hamu, Nadav Shamir, and Yaron Lipman. Invariant and equivariant graph networks. In *International Conference on Learning Representations*, 2019b. URL https://openreview.net/forum?id=Syx72jC9tm.

Grégoire Mialon, Dexiong Chen, Margot Selosse, and Julien Mairal. Graphit: Encoding graph structure in transformers. *arXiv preprint arXiv:2106.05667*, 2021.

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

Christopher Morris, Gaurav Rattan, and Petra Mutzel. Weisfeiler and leman go sparse: Towards scalable higher-order graph embeddings. *Advances in neural information processing systems*, 33: 21824–21840, 2020.

Pál András Papp and Roger Wattenhofer. A theoretical comparison of graph neural network extensions. In *International Conference on Machine Learning*, pp. 17323–17345. PMLR, 2022.

Ladislav Rampášek, Michael Galkin, Vijay Prakash Dwivedi, Anh Tuan Luu, Guy Wolf, and Dominique Beaini. Recipe for a general, powerful, scalable graph transformer. *Advances in neural information processing systems*, 35:14501–14515, 2022.

Gaurav Rattan and Tim Seppelt. Weisfeiler-leman and graph spectra. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 2268–2285. SIAM, 2023.

David E Roberson. Oddomorphisms and homomorphism indistinguishability over graphs of bounded degree. *arXiv preprint arXiv:2206.10321*, 2022.

Ryoma Sato, Makoto Yamada, and Hisashi Kashima. Random features strengthen graph neural networks. In *Proceedings of the 2021 SIAM international conference on data mining (SDM)*, pp. 333–341. SIAM, 2021.

Ameya Velingker, Ali Sinop, Ira Ktena, Petar Veličković, and Sreenivas Gollapudi. Affinity-aware graph networks. *Advances in Neural Information Processing Systems*, 36:67847–67865, 2023.

Haorui Wang, Haoteng Yin, Muhan Zhang, and Pan Li. Equivariant and stable positional encoding for more powerful graph neural networks. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=e95i1IHcWj.

Xiaohan Wang, Deyu Bo, Longlong Li, and Kelin Xia. Full-spectrum graph neural networks: Expressive and scalable. In *Forty-third International Conference on Machine Learning*, 2026.

Xiyuan Wang and Muhan Zhang. How powerful are spectral graph neural networks. In *International conference on machine learning*, pp. 23341–23362. PMLR, 2022.

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, 21–27 Jul 2024.

Zhenqin Wu, Bharath Ramsundar, Evan N Feinberg, Joseph Gomes, Caleb Geniesse, Aneesh S Pappu, Karl Leswing, and Vijay Pande. Moleculenet: a benchmark for molecular machine learning. *Chemical science*, 9(2):513–530, 2018.

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.

Chengxuan Ying, Tianle Cai, Shengjie Luo, Shuxin Zheng, Guolin Ke, Di He, Yanming Shen, and Tie-Yan Liu. Do transformers really perform badly for graph representation? *Advances in neural information processing systems*, 34:28877–28888, 2021.

Bohang Zhang, Guhao Feng, Yiheng Du, Di He, and Liwei Wang. A complete expressiveness hierarchy for subgraph gnns via subgraph weisfeiler-lehman tests. In *International Conference on Machine Learning*, pp. 41019–41077. PMLR, 2023a.

Bohang Zhang, Shengjie Luo, Liwei Wang, and Di He. Rethinking the expressive power of GNNs via graph biconnectivity. In *The Eleventh International Conference on Learning Representations*, 2023b. URL https://openreview.net/forum?id=r9hNv76KoT3.

Bohang Zhang, Jingchu Gai, Yiheng Du, Qiwei Ye, Di He, and Liwei Wang. Beyond weisfeiler-lehman: A quantitative framework for gnn expressiveness. In *International Conference on Learning Representations*, volume 2024, pp. 53674–53746, 2024a.

Bohang Zhang, Lingxiao Zhao, and Haggai Maron. On the expressive power of spectral invariant graph neural networks. In *Proceedings of the 41st International Conference on Machine Learning*, 2024b.

CONTENTS OF THE APPENDIX

A Limitations and Discussions **15**

B More Related Work **16**

C More Background Knowledge **18**

C.1 Higher-Order Graph Spectral Filtering 18

C.2 Rooted Graph Invariants 19

C.3 GNN Expressiveness 21

D Proofs for Section 4 **24**

D.1 Proof of Theorem 4.4 25

D.2 Proof of Theorem 4.7 25

D.3 Proof of Proposition 4.9 44

D.4 Proof of Theorem 4.10 46

E Proofs for Section 5 **49**

E.1 Proof of Proposition 5.2 49

E.2 Proof of Equation (3) 50

E.3 Pseudocode for Spectralized Structural Encodings 52

F Experimental Details **52**

A Limitations and Discussions

Theoretical Analysis

It remains open whether Spec k-FWL is strictly less expressive than $(k+1)$-LWL. Two natural routes toward proving strictness remain unresolved. The first is to construct a family of non-isomorphic graph pairs, for example through CFI-type constructions, that cannot be distinguished by Spec k-FWL but can be distinguished by $(k+1)$-LWL. So far, however, we have not succeeded in obtaining such a separation. The second approach is to characterize the homomorphism expressivity of Spec k-FWL following Zhang et al. (2024a, Definition 3.1). In the corresponding part of Appendix D, we express spectralized atomic-type responses through rooted homomorphism counts over certain templates and use these representations to establish the upper bound by $(k+1)$-LWL. A complete homomorphism characterization of Spec k-FWL, including the required maximality condition, remains open.

In Theorem 4.10, we establish the strict expressiveness gain of Spectralization through a case-by-case separation argument. It is straightforward to show that refinement augmented with a raw RGI is upper-bounded by refinement augmented with its spectralized counterpart. To prove that the gain is strict, however, we construct a specific pair of graphs that cannot be distinguished using the raw RGI but can be distinguished after spectralization. While this approach suffices for Theorem 4.10, a more systematic understanding of when Spectralization yields a strict expressiveness gain remains open. In particular, it would be desirable to characterize the boundary between RGIs for which Spectralization provides additional distinguishing power and those for which it does not. We suspect that such a characterization may be closely related to homomorphism expressivity, although obtaining a complete description of this gain boundary appears substantially more challenging. Developing such a theory would be valuable both for understanding Spectralization as a graph-isomorphism refinement mechanism and for guiding the design of more expressive GNNs.

Since higher-order spectral invariant GNNs are formulated with the standard graph operator fixed as $T = A$ in Gai et al. (2025), we establish part (ii) of Theorem 4.7 and Proposition 4.9 under the same restriction. This choice is also natural from the perspective of our proof techniques: when $T = A$, spectral invariants admit a characterization in terms of walk counts, which in turn leads to a clean homomorphism-based characterization. Nevertheless, this restriction leaves open the corresponding results for other standard graph operators. In the first-order setting, the Eigenspace Projection GNN of Zhang et al. (2024b) considers all three choices $T \in \{A, L, \widetilde{L}\}$. Moreover, in our proof of Proposition 4.9, the assumption $T = A$ is essential to the current argument. Without it, we can establish only one direction, SI 1-WL $\preceq$ Spec 1-WL, while whether the converse relation continues to hold for $T = L$ or $T = \widetilde{L}$ remains open. More generally, extending the expressiveness hierarchy in part (ii) of Theorem 4.7 from the adjacency operator to Laplacian operators remains an open problem.

Practical realizations

In our implementations, we primarily focus on spectralized 2-RGIs and use the resulting representations as node-pair-level structural encodings. This choice is motivated by the fact that many widely used RGIs are naturally 2-rooted. In particular, distance-based and random-walk-based structural quantities are defined on node pairs and have been extensively adopted in graph Transformers, as discussed in Appendix B. We therefore take 2-RGIs as a practical starting point, so that the resulting structural encodings can be readily integrated into existing architectures. Nevertheless, spectralized RGIs can be incorporated into existing architectures in considerably more flexible ways. First, one may directly spectralize 1-RGIs to obtain node-level structural encodings. More generally, higher-order spectralized RGIs can be transformed into lower-order representations through equivariant graph networks (Maron et al., 2019b). This follows the same principle as Wang et al. (2022), and is formalized and discussed in the context of positional encodings by Black et al. (2024), where node-pair encodings can be converted into node-level encodings without loss of expressive power. The same idea can be extended to higher orders; for example, spectralized 3-RGIs may be mapped to node-pair structural encodings, potentially providing more expressive pairwise representations. Exploring such higher-order realizations of Spectralization is therefore a natural direction for future work.

Finally, as this work is primarily theoretical, a more systematic empirical study of spectralizing a broader range of RGIs across diverse neural architectures, particularly graph Transformers, would be valuable. Such experiments could help establish practical guidelines for selecting which RGIs to spectralize and how to incorporate their spectralized forms into different architectures. More broadly, they may shed light on when spectralization leads to meaningful performance gains, and whether such gains are associated with particular physical or chemical processes or properties underlying the prediction tasks.

B More Related Work

Graph Spectral Filtering

Graph spectral filtering provides the computational foundation for our construction. Early spectral networks parameterize graph convolutions in the Laplacian eigenbasis (Bruna et al., 2013), while ChebNet uses Chebyshev polynomials to obtain localized filters without explicit eigendecomposition (Defferrard et al., 2016). Wang & Zhang (2022) analyze the expressive power of spectral GNNs in terms of both graph spectra and the frequency components present in input features, and motivate polynomial filter parameterizations based on orthogonal bases. These results emphasize that the information accessible through filtering depends on the input signal as well as the filter family.

Full-Spectrum GNNs extend filtering from node signals to node-pair signals through bivariate spectral responses, and further to arbitrary orders (Wang et al., 2026). We use this as our starting point. Our focus is on the structural information retained in the complete response profile of an RGI, together with the graph spectrum, and on how this profile can strengthen graph encodings.

Spectral Invariants and Spectral Invariant GNNs

Spectral invariants have long been studied for distinguishing cospectral graphs, namely, graphs with the same spectra. Fürer (2010) investigate the distinguishing power of combinatorial and spectral invariants, including information derived from eigenspace projections. Rattan & Seppelt (2023) construct a family of graph matrices based on the adjacency of those graphs. For a fixed WL dimension and a prescribed number of refinement rounds, they study whether two graphs remain indistinguishable, and they show that this is equivalent to the corresponding matrices satisfying certain spectral conditions, thereby connecting finite-round WL refinement with the spectral information of these matrices. These results place spectral information within the broader study of graph isomorphism and motivate its use in expressive graph learning.

In graph learning, a central issue is how to use eigenvectors without making predictions depend on their arbitrary signs or on the choice of basis within repeated eigenspaces. SignNet handles sign ambiguity, while BasisNet addresses orthogonal changes of eigenspace bases through eigenspace-projection-based representations (Lim et al., 2023). PEG instead imposes orthogonal equivariance on positional features alongside permutation equivariance on node features (Wang et al., 2022). SPE further studies stability under graph perturbations, using soft spectral aggregation to avoid the instability associated with separating nearby eigenspaces (Huang et al., 2024). The Eigenspace Projection GNN framework of Zhang et al. (2024b) organizes the expressive power of a broad class of spectral invariant architectures through eigenspace projection information, and Gai et al. (2025) extend the analysis to higher-order spectral invariant GNNs using homomorphism expressivity (Zhang et al., 2024a).

Other approaches combine spectral information with additional structural summaries. Feldman et al. (2023) construct spectral pre-colorings from heat-kernel information, including diagonal entries and nodewise summaries of off-diagonal entries, before applying WL refinement. SpectRe enriches persistent-homology representations with spectra of connected components arising along graph filtrations (Ji et al., 2026). Its spectral information is therefore collected from a family of evolving subgraphs, whereas our construction applies higher-order filters of a chosen graph operator to structural signals on node tuples. These approaches illustrate distinct ways to combine spectral and combinatorial information.

Rooted Graph Invariants as Positional and Structural Encodings

RGIs provide a common formulation for deterministic positional and structural encodings that are equivariant under node relabeling. Node-level encodings assign structural information to individual vertices, while pairwise encodings describe relations between vertices; both fit the isomorphism-based formulation of absolute and relative positional encodings studied by Black et al. (2024). Such encodings can augment node or edge features, modify attention scores, or be updated jointly with learned representations (Ying et al., 2021; Dwivedi et al., 2022; Rampášek et al., 2022; Ma et al., 2023). Their use is particularly important in graph Transformers, where they provide structural information that is absent from unrestricted self-attention alone.

A major family of those encodings describes connectivity through distances and propagation statistics. Distance Encoding incorporates shortest-path and random-walk information into graph representations (Li et al., 2020), while Graphormer uses shortest-path distances to define pairwise attention biases (Ying et al., 2021). Zhang et al. (2023b) analyze the expressive roles of shortest-path and resistance distances through graph biconnectivity. Related constructions include hierarchical distances (Luo et al., 2024), diffusion kernels in GraphiT (Mialon et al., 2021), random-walk return probabilities in structural encodings (Dwivedi et al., 2022; Rampášek et al., 2022), and relative random-walk probabilities in GRIT (Ma et al., 2023). Affinity-aware graph networks further exploit relations such as effective resistance and hitting times (Velingker et al., 2023). These quantities capture different aspects of connectivity and need not have the same distinguishing power (Black et al., 2024).

Another family records how nodes or node pairs participate in prescribed substructures. Graph Substructure Networks augment message passing with structural features obtained by counting subgraph isomorphisms (Bouritsas et al., 2022), while Barceló et al. (2021) study GNNs augmented with rooted homomorphism counts as local graph parameters. Jin et al. (2024) show how separating the homomorphism counts underlying motif counts can yield more expressive augmentations. Building on this perspective, MoSE constructs structural encodings from homomorphism counts and integrates them into graph learning architectures (Bao et al., 2025). Simple Path Structural Encoding instead uses counts of simple paths between node pairs to capture connectivity beyond random-walk probabilities (Airale et al., 2025). Local topological encodings offer a related source of structural information: LEAP computes differentiable local Euler characteristic transforms using node features as coordinates (Amboage et al., 2026) (For such feature-dependent constructions, the RGI interpretation requires feature-preserving isomorphisms, or node features derived equivariantly from graph structure.). Our framework provides a common operation for enriching these structural signals through their spectral responses, with the chosen RGI determining which information is supplied to that operation.

GNN Expressiveness

The expressive power of GNNs is commonly studied through their ability to distinguish non-isomorphic graphs. Message-passing GNNs are bounded by 1-WL, and this bound can be attained with sufficiently expressive aggregation and update functions (Xu et al., 2019; Morris et al., 2019). Higher-order architectures strengthen these models by maintaining representations of node tuples (Morris et al., 2019; Maron et al., 2019a). To reduce the cost of updating tuple representations, Morris et al. (2020) introduce local higher-order refinement, which restricts coordinate replacements to graph neighbors. Another approach processes collections of rooted or otherwise selected subgraphs and combines their representations; corresponding subgraph WL tests provide an expressiveness hierarchy for representative architectures (Zhang et al., 2023a). These constructions provide reference points for comparing the refinement procedures studied in this work.

Homomorphism counting gives a structural characterization of these expressive capabilities. Dell et al. (2018) characterize WL indistinguishability through agreement of homomorphism counts from graphs of bounded treewidth. In our indexing convention, for $k \geq 2$, two graphs are indistinguishable by k-FWL if and only if their homomorphism counts agree for every graph of treewidth at most k. Zhang et al. (2024a) develop homomorphism expressivity as a quantitative framework for comparing GNN families and analyzing their ability to count substructures. This viewpoint is relevant to Spectralization in two ways: rooted homomorphism counts supply structural signals, and homomorphism characterizations can provide tools for comparing the resulting refinement algorithms.

Probabilistic Positional Encodings and Symmetry Breaking

Probabilistic encodings provide another route to improving GNN expressiveness. Sato et al. (2021) show that random node features strengthen standard GNNs, and Abboud et al. (2021) establish probabilistic universal approximation of invariant functions on graphs of bounded size using random node initialization. Random Feature Propagation connects this approach with spectral methods by constructing positional encodings from intermediate graph-propagation states initialized with random features (Eliasof et al., 2023). More recently, Lawrence et al. (2025) formulate symmetry breaking through equivariant conditional distributions and introduce SymPE, which uses randomized canonicalization to supply symmetry-breaking positional encodings to equivariant networks.

The distinction from deterministic RGIs concerns how these encodings behave under graph automorphisms. A node-level RGI assigns the same value to automorphic nodes; more generally, a k-RGI is constant on each automorphism orbit of node k-tuples. Probabilistic encodings can distinguish automorphic nodes in individual samples while preserving equivariance at the distributional level (Lawrence et al., 2025). A construction may also be equivariant when the graph and its sampled auxiliary features are relabeled jointly, but this does not make the sampled encoding an RGI of the graph alone. Regarding the scope of this work, Spectralization operates on deterministic RGIs and preserves their isomorphism invariance. Its expressiveness gains arise from exposing additional structural information while respecting automorphism orbits.

C More Background Knowledge

We provide additional background on higher-order spectral filtering, rooted graph invariants, and the color-refinement view of GNN expressiveness. Throughout, we follow the notation in Section 3 and work with finite, simple, undirected graphs unless otherwise stated.

C.1 Higher-Order Graph Spectral Filtering

A real-valued order-k graph signal is a function $\mathbf{X} : V(G)^k \to \mathbb{R}$, which we identify with a tensor $\mathbf{X} \in \mathbb{R}^{n\times\cdots\times n}$, or equivalently with its column vectorization $\boldsymbol{x} := \mathrm{vec}(\mathbf{X}) \in \mathbb{R}^{n^k}$. Let u_i^G denote the ith column of U_G. The order-k graph Fourier basis is defined by $U_G^{\otimes k} \in \mathbb{R}^{n^k\times n^k}$, whose columns $\{u_{i_1}^G \otimes \cdots \otimes u_{i_k}^G\}_{(i_1,\ldots,i_k)\in[n]^k}$ form an orthonormal basis of $\mathbb{R}^{n^k}$. Accordingly, the order-k graph Fourier transform and its inverse are

$$\widehat{\boldsymbol{x}} = (U_G^\top)^{\otimes k}\boldsymbol{x}, \qquad \boldsymbol{x} = U_G^{\otimes k}\widehat{\boldsymbol{x}}.$$

Writing $\widehat{\boldsymbol{x}} = \mathrm{vec}(\widehat{\mathbf{X}})$, these transformations can also be expressed in tensor-index notation as

$$\widehat{\mathbf{X}}_{i_1\cdots i_k} = \sum_{v_1,\ldots,v_k\in V(G)} \left(\prod_{\ell=1}^{k}(U_G)_{v_\ell i_\ell}\right)\mathbf{X}_{v_1\cdots v_k}, \quad \mathbf{X}_{v_1\cdots v_k} = \sum_{i_1,\ldots,i_k\in[n]} \left(\prod_{\ell=1}^{k}(U_G)_{v_\ell i_\ell}\right)\widehat{\mathbf{X}}_{i_1\cdots i_k}.$$

Given a continuous spectral response function $g : \mathbb{R}^k \to \mathbb{R}$, define the spectral response tensor by

$$\mathbf{G}_g^{(k)} := \big(g\big((\Lambda_G)_{i_k i_k},\ldots,(\Lambda_G)_{i_1 i_1}\big)\big)_{i_1,\ldots,i_k\in[n]}.$$

The reversed argument order matches the order of the Kronecker factors under column-wise vectorization. Here, $(\Lambda_G)_{ii}$, $i \in [n]$, lists eigenvalues with multiplicity, whereas λ_r^G, $\boldsymbol{r} \in [m_G]$, denotes the distinct eigenvalues. The corresponding order-k spectral filter is

$$\mathcal{F}_{G,g}^{(k)} := U_G^{\otimes k}\,\mathrm{diag}\Big(\mathrm{vec}(\mathbf{G}_g^{(k)})\Big)\,(U_G^\top)^{\otimes k}.$$

Thus, filtering multiplies the Fourier coefficients pointwise by $\mathbf{G}_g^{(k)}$ before applying the inverse transform.

Equivalently, define the commuting operators

$$\mathbf{T}_i := I_n^{\otimes(i-1)} \otimes T \otimes I_n^{\otimes(k-i)}, \qquad i \in [k],$$

where an empty tensor product is omitted. For $\boldsymbol{r} = (r_1, \ldots, r_k) \in [m_G]^k$, let $\boldsymbol{\lambda}^G_{\boldsymbol{r}} := (\lambda^G_{r_1}, \ldots, \lambda^G_{r_k})$ and $P^G_{\boldsymbol{r}} := P^G_{r_1} \otimes \cdots \otimes P^G_{r_k}$. The latter is the orthogonal projector onto the joint eigenspace $E^G_{r_1} \otimes \cdots \otimes E^G_{r_k}$, where $E^G_r = \ker(T - \lambda^G_r I_n)$. Proposition 4.1 then gives

$$\mathcal{F}^{(k)}_{G,g} = g(\mathbf{T}_1, \ldots, \mathbf{T}_k) = \sum_{\boldsymbol{r} \in [m_G]^k} g(\boldsymbol{\lambda}^G_{\boldsymbol{r}}) P^G_{\boldsymbol{r}}. \tag{5}$$

This expansion shows that the filter is independent of the choice of orthonormal basis within each eigenspace, even though the individual Fourier coefficients depend on that choice. It also makes clear that only the values of g on the finite set $\{\boldsymbol{\lambda}^G_{\boldsymbol{r}} : \boldsymbol{r} \in [m_G]^k\}$ affect filtering on a fixed graph.

The order-1 specialization recovers conventional graph spectral filtering, $\mathcal{F}^{(1)}_{G,g} = g(T)$. For order 2, identifying $\mathbf{X}$ with its signal matrix $X \in \mathbb{R}^{n \times n}$, column-wise vectorization gives

$$(P^G_r \otimes P^G_s)\,\mathrm{vec}(X) = \mathrm{vec}(P^G_s X P^G_r).$$

Consequently, the second-order filtering of Wang et al. (2026) takes the matrix form

$$\mathcal{F}^{(2)}_{G,g} X = \sum_{r,s=1}^{m_G} g(\lambda^G_r, \lambda^G_s) P^G_s X P^G_r,$$

where we identify the filtered vector with its matrix representation. In particular, a separable response $g(x, y) = f(x)h(y)$ yields $\mathcal{F}^{(2)}_{G,g} X = h(T) X f(T)$. At any order, the constant response $g \equiv 1$ gives the identity operator and therefore recovers the original signal.

C.2 Rooted Graph Invariants

As defined in Section 3, a k-RGI assigns the same value to isomorphic graphs with corresponding ordered root tuples. The roots need not be distinct, and their positions in the tuple matter. For a fixed graph, an RGI therefore defines a signal on ordered node tuples that transforms equivariantly under node relabeling. More precisely, let $\varphi : G \to H$ be a graph isomorphism and write $\varphi(\mathbf{u}) = (\varphi(u_1), \ldots, \varphi(u_k))$. Then

$$\Phi_k(H, \varphi(\mathbf{u})) = \Phi_k(G, \mathbf{u}).$$

For a real-valued RGI, let Π_φ be the permutation matrix satisfying $\Pi_\varphi e_u = e_{\varphi(u)}$, where e_u is the standard basis vector indexed by u. With $\Phi_G := \Phi_k(G, \cdot)$ identified with its vectorization, the same condition becomes $\Phi_H = \Pi_\varphi^{\otimes k} \Phi_G$. This connects the rooted-invariant formulation to deterministic equivariant encodings on nodes, node pairs, and higher-order tuples.

RGIs can be combined componentwise or followed by a fixed function without losing invariance. They can also be restricted to repeated roots; for example, a 2-RGI Φ induces the 1-RGI $(G, u) \mapsto \Phi(G, u, u)$. Within a graph, every RGI is constant on root tuples related by a graph automorphism, so it cannot distinguish such tuples. The following proposition shows that Spectralization preserves the same equivariance property.

Proposition C.1. *Let Φ_k be a real-valued k-RGI and let $\mathfrak{C}_k = (\Phi_{k,\alpha})_{\alpha \in \mathcal{A}}$ be a family of real-valued k-RGIs. For every standard graph operator T, $\mathrm{Spec}^T_{\Phi_k}$ and $\mathrm{Spec}^T_{\mathfrak{C}_k}$ are k-RGIs.*

Proof. Let $G_{\mathbf{u}} = (G, u_1, \ldots, u_k)$ and $H_{\mathbf{v}} = (H, v_1, \ldots, v_k)$ be isomorphic k-rooted graphs, and let $\varphi : G \to H$ be an isomorphism satisfying $\varphi(u_i) = v_i$ for every $i \in [k]$. Let Π_φ denote the corresponding permutation matrix. Writing T_G and T_H for the chosen graph operator on the two graphs, equivariance gives $T_H = \Pi_\varphi T_G \Pi_\varphi^\top$. Hence $\mathrm{spec}(T_G) = \mathrm{spec}(T_H)$, and their spectral projectors satisfy

$$P^H_r = \Pi_\varphi P^G_r \Pi_\varphi^\top, \qquad r \in [m_G].$$

Consequently, for every $\boldsymbol{r} \in [m_G]^k$,

$$P^H_{\boldsymbol{r}} = \Pi_\varphi^{\otimes k} P^G_{\boldsymbol{r}} (\Pi_\varphi^{\otimes k})^\top.$$

By Proposition 4.1, for every $g \in C(\mathbb{R}^k)$,

$$\mathcal{F}^{(k)}_{H,g} = \Pi_\varphi^{\otimes k} \mathcal{F}^{(k)}_{G,g} (\Pi_\varphi^{\otimes k})^\top.$$

Since Φ_k is a k-RGI, its associated signals satisfy $\Phi_H = \Pi_\varphi^{\otimes k}\Phi_G$. Therefore, $\mathcal{F}_{H,g}^{(k)}\Phi_H = \Pi_\varphi^{\otimes k}\mathcal{F}_{G,g}^{(k)}\Phi_G$. Evaluating at $\mathbf{v} = \varphi(\mathbf{u})$ gives

$$\left[\mathcal{F}_{H,g}^{(k)}\Phi_H\right](\mathbf{v}) = \left[\mathcal{F}_{G,g}^{(k)}\Phi_G\right](\mathbf{u}).$$

This holds for every $g \in C(\mathbb{R}^k)$. Together with equality of the spectra, it implies

$$\mathrm{Spec}_{\Phi_k}^T(H, \mathbf{v}) = \mathrm{Spec}_{\Phi_k}^T(G, \mathbf{u}).$$

Thus $\mathrm{Spec}_{\Phi_k}^T$ is a k-RGI. Finally, $\mathrm{Spec}_{\mathfrak{C}_k}^T$ is obtained by collecting $\mathrm{Spec}_{\Phi_{k,\alpha}}^T$ componentwise over $\alpha \in \mathcal{A}$, so it is also a k-RGI. □

Homomorphism and subgraph counts. Given graphs F and G, a homomorphism from F to G is a map $f : V(F) \to V(G)$ that preserves edges. We denote the set of such maps by $\mathrm{Hom}(F, G)$ and its cardinality by $\mathrm{hom}(F, G)$. An injective homomorphism is a homomorphism that is injective on vertices; we denote the corresponding set and count by $\mathrm{Hom}_{\mathrm{inj}}(F, G)$ and $\mathrm{hom}_{\mathrm{inj}}(F, G)$. An isomorphism is a bijection $f : V(F) \to V(G)$ such that both f and f^{-1} are homomorphisms. We denote by $\mathrm{Sub}(F, G)$ the set of subgraphs of G isomorphic to F, and define $\mathrm{sub}(F, G) := |\mathrm{Sub}(F, G)|$.

In the following, we refer to the domain graph F as a *template*. For rooted graphs, homomorphisms and isomorphisms are additionally required to preserve the distinguished roots. For a tuple $\boldsymbol{v} = (v_1, \ldots, v_k) \in V(G)^k$, recall that $\mathbf{atp}_G(\boldsymbol{v})$ denotes its atomic type. For a template root tuple $\boldsymbol{r} = (r_1, \ldots, r_k) \in V(F)^k$, $\mathbf{atp}_G(\boldsymbol{v}) = \mathbf{atp}_F(\boldsymbol{r})$ if and only if, for all $s, t \in [k]$,

$$v_s = v_t \iff r_s = r_t, \qquad \{v_s, v_t\} \in E(G) \iff \{r_s, r_t\} \in E(F).$$

A natural source of RGIs is homomorphism counting. Let $F_{\boldsymbol{r}} = (F, r_1, \ldots, r_k) \in \mathcal{G}_k$ be a fixed k-rooted template. For $G_{\boldsymbol{v}} = (G, v_1, \ldots, v_k) \in \mathcal{G}_k$, define

$$\mathrm{Hom}(F_{\boldsymbol{r}}, G_{\boldsymbol{v}}) \coloneqq \{f \in \mathrm{Hom}(F, G) : f(r_t) = v_t \text{ for every } t \in [k]\},$$

and

$$\Phi_{F_{\boldsymbol{r}}}^{\mathrm{hom}}(G, \boldsymbol{v}) \coloneqq \mathrm{hom}(F_{\boldsymbol{r}}, G_{\boldsymbol{v}}) = |\mathrm{Hom}(F_{\boldsymbol{r}}, G_{\boldsymbol{v}})|.$$

Requiring homomorphisms to be injective analogously defines

$$\Phi_{F_{\boldsymbol{r}}}^{\mathrm{inj}}(G, \boldsymbol{v}) \coloneqq \mathrm{hom}_{\mathrm{inj}}(F_{\boldsymbol{r}}, G_{\boldsymbol{v}}).$$

Injective homomorphisms preserve edges but need not preserve non-edges, so their image copies need not be induced. Ordinary homomorphisms may identify distinct template vertices, whereas injective homomorphisms cannot.

For a fixed rooting, let $\mathrm{Aut}(F_{\boldsymbol{r}})$ be the group of automorphisms of F fixing every root r_t. The rooted subgraph count is

$$\Phi_{F_{\boldsymbol{r}}}^{\mathrm{sub}}(G, \boldsymbol{v}) \coloneqq \frac{\mathrm{hom}_{\mathrm{inj}}(F_{\boldsymbol{r}}, G_{\boldsymbol{v}})}{|\mathrm{Aut}(F_{\boldsymbol{r}})|}.$$

It counts subgraphs $Q \subseteq G$ for which $(Q, \boldsymbol{v}) \cong (F, \boldsymbol{r})$, since each such rooted copy admits exactly $|\mathrm{Aut}(F_{\boldsymbol{r}})|$ root-preserving isomorphisms from $F_{\boldsymbol{r}}$.

Subgraph counts that do not prescribe a particular rooting of the template can instead be obtained by aggregating rooted counts. For a fixed template F and $\boldsymbol{v} \in V(G)^k$, define the admissible rootings

$$\mathcal{R}_F(\boldsymbol{v}) \coloneqq \left\{\boldsymbol{r} \in V(F)^k : \mathbf{atp}_F(\boldsymbol{r}) = \mathbf{atp}_G(\boldsymbol{v})\right\}.$$

The number of copies of F containing the subgraph induced by $\boldsymbol{v}$ is then

$$\Phi_F^{\mathrm{sub}}(G, \boldsymbol{v}) \coloneqq \mathrm{sub}_F(G; \boldsymbol{v}) = \frac{1}{|\mathrm{Aut}(F)|} \sum_{\boldsymbol{r} \in \mathcal{R}_F(\boldsymbol{v})} \mathrm{hom}_{\mathrm{inj}}(F_{\boldsymbol{r}}, G_{\boldsymbol{v}}), \tag{6}$$

where $\mathrm{Aut}(F)$ is the automorphism group of F. Here, a copy is a subgraph $Q \subseteq G$ isomorphic to F, and containment means $G[\{v_1, \ldots, v_k\}] \subseteq Q$. In particular, the copy must contain every edge of G between the specified vertices, although it need not be induced on all its vertices. Each such copy

contributes a total of $|\operatorname{Aut}(F)|$ to the sum, one for each isomorphism from F onto the copy, which explains the normalization.

For $k = 1$, $\Phi_F^{\mathrm{sub}}(G, u)$ counts copies of F containing u. For $k = 2$, when (u, v) spans an edge, it counts copies containing that undirected edge; for $k = 3$, when the tuple spans a triangle, it counts copies containing that triangle. The node and edge cases correspond to the substructure counts considered in Table 3 of Zhang et al. (2024a).

Proposition C.2. *For every $F_{\boldsymbol{r}} \in \mathcal{G}_k$, $\Phi_{F_{\boldsymbol{r}}}^{\mathrm{hom}}$, $\Phi_{F_{\boldsymbol{r}}}^{\mathrm{inj}}$, $\Phi_{F_{\boldsymbol{r}}}^{\mathrm{sub}}$, and Φ_F^{sub} are k-RGIs.*

Proof. Let $G_{\mathbf{v}}$ and $H_{\mathbf{w}}$ be isomorphic k-rooted graphs, and let $\varphi : G \to H$ be an isomorphism satisfying $\varphi(v_i) = w_i$ for every $i \in [k]$. For a fixed k-rooted template $F_{\boldsymbol{r}}$, composition with φ defines a map

$$\operatorname{Hom}(F_{\boldsymbol{r}}, G_{\mathbf{v}}) \longrightarrow \operatorname{Hom}(F_{\boldsymbol{r}}, H_{\mathbf{w}}), \qquad f \longmapsto \varphi \circ f.$$

This map is a bijection, with inverse $h \mapsto \varphi^{-1} \circ h$. Therefore,

$$\hom(F_{\boldsymbol{r}}, G_{\mathbf{v}}) = \hom(F_{\boldsymbol{r}}, H_{\mathbf{w}}),$$

so $\Phi_{F_{\boldsymbol{r}}}^{\mathrm{hom}}$ is a k-RGI. Since composition with an isomorphism preserves injectivity, the same bijection restricts to

$$\operatorname{Hom}_{\mathrm{inj}}(F_{\boldsymbol{r}}, G_{\mathbf{v}}) \longleftrightarrow \operatorname{Hom}_{\mathrm{inj}}(F_{\boldsymbol{r}}, H_{\mathbf{w}}).$$

This proves invariance of $\Phi_{F_{\boldsymbol{r}}}^{\mathrm{inj}}$, and division by the fixed number $|\operatorname{Aut}(F_{\boldsymbol{r}})|$ proves invariance of $\Phi_{F_{\boldsymbol{r}}}^{\mathrm{sub}}$. Finally, isomorphic rooted graphs have the same atomic type, so $\mathcal{R}_F(\mathbf{v}) = \mathcal{R}_F(\mathbf{w})$. Every summand in (6) is preserved by the preceding bijection, and hence $\Phi_F^{\mathrm{sub}}(G, \mathbf{v}) = \Phi_F^{\mathrm{sub}}(H, \mathbf{w})$. □

Atomic-type indicators form another useful family of real-valued RGIs. For $\tau \in \mathbf{AT}_k$,

$$\Phi_\tau(G, \boldsymbol{v}) = \mathbf{1}\big[(G[\{v_1, \ldots, v_k\}], \boldsymbol{v}) \cong \tau\big].$$

As in Section 4.2, this indicator records the full equality and adjacency pattern of the ordered tuple, including non-adjacency. Its invariance follows because an isomorphism preserves each of these relations. Unlike a count that sums over template rootings, an atomic-type indicator retains the correspondence between individual root positions.

C.3 GNN Expressiveness

We study expressiveness through the ability to distinguish non-isomorphic graphs. A GNN architecture distinguishes graphs G and H if some choice of its parameters and depth produces different graph representations. This is a statement about the capacity of an architecture, rather than the behavior of a particular trained model. For architectures based on message passing, color-refinement algorithms provide a combinatorial description of the structural information available to the network (Xu et al., 2019; Grohe, 2021). They replace learned transformations by injective hashes, in turn preserving all information supplied by the initialization and aggregation rules.

Color refinement and graph distinguishability. For a WL color refinement algorithm $a \in \{F, O, L\}$ acting on ordered k-tuples, define its round-t graph representation by

$$\mathcal{C}_{a,G}^{(t)} := \left\{\!\!\left\{ \chi_{a,G}^{(t)}(\boldsymbol{v}) : \boldsymbol{v} \in V(G)^k \right\}\!\!\right\}.$$

The algorithm distinguishes G and H if $\mathcal{C}_{a,G}^{(t)} \neq \mathcal{C}_{a,H}^{(t)}$ for some t. Colors are compared using the same injective hash across graphs. Since each update retains the previous color, the induced partitions can only become finer and stabilize on finite graphs. Accordingly, we may equivalently compare the multisets of stable tuple colors, with stabilization understood for the common coloring of the two tuple spaces. The partitions stabilize even though the numerical or symbolic names assigned to their colors may continue to change. Every refinement considered here preserves graph isomorphisms, so a difference between the color multisets certifies non-isomorphism; equality does not, in general, certify isomorphism. The relations $\succeq$, $\succ$, and $\equiv$ introduced in Section 3 compare this graph-distinguishing power.

Ordinary 1-WL and message passing. The ordinary 1-WL algorithm, also called color refinement algorithm, assigns colors to individual nodes and updates them using neighboring colors:

$$\chi_G^{(0)}(v) = \mathbf{atp}_G(v), \qquad \chi_G^{(t+1)}(v) = \mathrm{hash}\left(\chi_G^{(t)}(v), \left\{\!\left\{\chi_G^{(t)}(w) : w \in \mathcal{N}_G(v)\right\}\!\right\}\right). \tag{7}$$

For the unlabeled graphs considered in our expressiveness analysis, $\mathbf{atp}_G(v)$ is constant. Standard message-passing GNNs initialized with the same node information cannot distinguish graphs that 1-WL fails to distinguish. Conversely, injective neighborhood aggregation, injective node updates, and an injective graph readout recover the distinguishing power of 1-WL when sufficiently many layers are available (Xu et al., 2019). The same principle motivates higher-order architectures whose representations and aggregation rules operate on node tuples. An exact correspondence with a particular refinement requires that the neural architecture preserve the information used by that refinement.

Higher-order refinements and the order-one convention. For $\boldsymbol{v} \in V(G)^k$, $i \in [k]$, and $w \in V(G)$, write $\boldsymbol{v}^{\,i\leftarrow w} := (v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_k)$ for the tuple obtained by replacing the i-th entry of $\boldsymbol{v}$ with w. For $k \geq 2$, all three algorithms share the initialization and update rule

$$\chi_{a,G}^{(0)}(\boldsymbol{v}) = \mathbf{atp}_G(\boldsymbol{v}), \qquad \chi_{a,G}^{(t+1)}(\boldsymbol{v}) = \mathrm{hash}\left(\chi_{a,G}^{(t)}(\boldsymbol{v}), M_{a,G}^{(t)}(\boldsymbol{v})\right),$$

where $\mathbf{atp}_G(\boldsymbol{v})$ denotes the atomic type of $\boldsymbol{v}$ (Dell et al., 2018), $a \in \{F, O, L\}$ identifies the variant, and hash is an injective encoding used consistently across graphs. The three variants differ in their aggregation rules.

The folklore k-dimensional Weisfeiler–Leman (k-FWL) algorithm jointly aggregates the colors of the k replacements associated with each node w:

$$M_{F,G}^{(t)}(\boldsymbol{v}) = \left\{\!\left\{\left(\chi_{F,G}^{(t)}(\boldsymbol{v}^{\,1\leftarrow w}), \ldots, \chi_{F,G}^{(t)}(\boldsymbol{v}^{\,k\leftarrow w})\right) : w \in V(G)\right\}\!\right\}. \tag{8}$$

The oblivious k-dimensional Weisfeiler–Leman (k-OWL) algorithm instead aggregates replacement colors separately for each coordinate:

$$M_{O,G}^{(t)}(\boldsymbol{v}) = \left(\left\{\!\left\{\chi_{O,G}^{(t)}(\boldsymbol{v}^{\,1\leftarrow w}) : w \in V(G)\right\}\!\right\}, \ldots, \left\{\!\left\{\chi_{O,G}^{(t)}(\boldsymbol{v}^{\,k\leftarrow w}) : w \in V(G)\right\}\!\right\}\right). \tag{9}$$

The local k-dimensional Weisfeiler–Leman (k-LWL) algorithm, denoted by δ-k-LWL in Morris et al. (2020), restricts the aggregation for each coordinate to neighborhood $\mathcal{N}_G$ of its current entry:

$$M_{L,G}^{(t)}(\boldsymbol{v}) = \left(\left\{\!\left\{\chi_{L,G}^{(t)}(\boldsymbol{v}^{\,1\leftarrow w}) : w \in \mathcal{N}_G(v_1)\right\}\!\right\}, \ldots, \left\{\!\left\{\chi_{L,G}^{(t)}(\boldsymbol{v}^{\,k\leftarrow w}) : w \in \mathcal{N}_G(v_k)\right\}\!\right\}\right). \tag{10}$$

For $k \geq 2$, the three variants in Section 3 start from the same atomic types and differ in how they aggregate replacement colors. The joint aggregation of k-FWL in (8) retains the association between all replacements made with the same node w. The coordinate-wise aggregation of k-OWL in (9) discards this association. The k-LWL update in (10) is also coordinate-wise, but permits a replacement in coordinate i only when $w \in \mathcal{N}_G(v_i)$. Thus, locality changes the structural information supplied to the refinement, in addition to reducing the number of replacements. Our k-LWL is the δ-k-LWL of Morris et al. (2020); it does not include the additional counting information used in their δ-k-LWL$^+$ variant.

In graph-distinguishing power, the dimension shift between the joint and oblivious variants is

$$k\text{-FWL} \equiv (k+1)\text{-OWL} \quad (k \geq 2), \qquad 1\text{-WL} \equiv 2\text{-OWL},$$

as established by Grohe (2021, Theorem V.6 and Corollary V.7). The order-one case requires separate treatment: directly substituting $k = 1$ into either global replacement rule aggregates all node colors without recording adjacency, and hence does not yield ordinary 1-WL. Throughout this work, the order-one baseline is instead (7), or its equivalent all-node formulation in (13). The latter records the atomic type of (u, v) alongside the color of each v and therefore distinguishes the root, its neighbors, and its non-neighbors. This has the same graph-distinguishing power as ordinary 1-WL: the global color multiset determines the total number of nodes of each color, from which the root and neighbor multiplicities determine the non-neighbor multiplicities. This graph-level equivalence does not assert equality of the node colorings across different graphs.

Moreover, the spectral invariant k-dimensional Weisfeiler–Leman (SI k-OWL) algorithm (Gai et al., 2025), with its tuple colors $\chi^{(t)}_{\mathrm{SI},G}(\boldsymbol{v})$ initialized and updated by

$$\chi^{(0)}_{\mathrm{SI},G}(\boldsymbol{v}) = \big(\mathcal{P}^A(G, v_i, v_j)\big)_{1\le i<j\le k},$$

$$\chi^{(t+1)}_{\mathrm{SI},G}(\boldsymbol{v}) = \mathrm{hash}\Big(\chi^{(t)}_{\mathrm{SI},G}(\boldsymbol{v}), \Big(\Big\{\!\Big\{\big(\chi^{(t)}_{\mathrm{SI},G}(\boldsymbol{v}^{\,i\leftarrow w}), \mathcal{P}^A(G, v_i, w)\big) : w \in V(G)\Big\}\!\Big\}\Big)_{i\in[k]}\Big). \tag{11}$$

This refinement follows the coordinate-wise aggregation of k-OWL (9), enriching both its initialization and aggregation with the eigenspace projection invariant (1).

The connection to homomorphism counts. WL distinguishability also admits a characterization in terms of graph homomorphisms (Dell et al., 2018). Two graphs are indistinguishable by 1-WL if and only if they have the same number of homomorphisms from every finite tree. More generally, for $k \ge 2$, they are indistinguishable by k-FWL if and only if

$$\hom(F, G) = \hom(F, H) \qquad \text{for every finite graph } F \text{ of treewidth at most } k.$$

Here a tree decomposition of F consists of bags of vertices indexed by the nodes of a tree: the bags cover $V(F)$, the endpoints of every edge belong to a common bag, and the bags containing any fixed vertex form a connected subtree. The treewidth is the minimum, over all such decompositions, of the largest bag size minus one. This characterization connects the comparison of refinement algorithms to the structural patterns whose homomorphism counts they determine, and complements the rooted homomorphism counts introduced above.

Atomic Types and Augmented FWL An atomic type records the equality and adjacency relations among the entries of an ordered node tuple. To describe them, fix $k \ge 1$ and consider a simple graph F with an ordered root tuple $\boldsymbol{r} = (r_1, \dots, r_k) \in V(F)^k$ containing every vertex of F. The entries of $\boldsymbol{r}$ may repeat, so F has at most k vertices. As defined in Section 3, isomorphisms between rooted graphs must map roots in corresponding tuple positions to one another. There are finitely many such isomorphism classes; let $\mathbf{AT}_k$ be the set containing one representative from each class. For a tuple $\boldsymbol{v} = (v_1, \dots, v_k) \in V(G)^k$, let $G[\{v_1, \dots, v_k\}]$ denote the subgraph consisting of the distinct vertices appearing in $\boldsymbol{v}$ and all edges of G between them. Together with the ordered root tuple $\boldsymbol{v}$, this gives the rooted graph $(G[\{v_1, \dots, v_k\}], \boldsymbol{v})$. For each $\tau \in \mathbf{AT}_k$, define the atomic-type indicator $\Phi_\tau(G, \boldsymbol{v}) \coloneqq \mathbf{1}\big[(G[\{v_1, \dots, v_k\}], \boldsymbol{v}) \cong \tau\big]$. These indicators are real-valued k-RGIs and form the finite family $\mathcal{AT}_k \coloneqq (\Phi_\tau)_{\tau\in\mathbf{AT}_k}$. We define the atomic type of $\boldsymbol{v}$ by the one-hot vector

$$\mathbf{atp}_G(\boldsymbol{v}) \coloneqq \mathcal{AT}_k(G, \boldsymbol{v}) = \big(\Phi_\tau(G, \boldsymbol{v})\big)_{\tau\in\mathbf{AT}_k}.$$

Example C.3. *For $k = 2$, a node pair (u, v) has one of three atomic types: the nodes are identical, adjacent, or distinct and non-adjacent. Ordering the corresponding indicators in this way gives*

$$\mathbf{atp}_G(u, v) = \Big(\mathbf{1}[u = v], \mathbf{1}[\{u, v\} \in E(G)], \mathbf{1}[u \ne v \text{ and } \{u, v\} \notin E(G)]\Big).$$

We then study how spectralizing atomic types affects the expressive power of k-FWL (8). We first introduce an augmented formulation that makes the atomic-type information used in aggregation explicit. For $k \ge 2$, denote the colors of this refinement by $\widehat{\chi}^{(t)}_{F,G}$. We retain the atomic-type initialization and hash update of standard k-FWL, and replace the aggregation multiset by

$$\widehat{M}^{(t)}_{F,G}(\boldsymbol{v}) = \Big\{\!\Big\{\Big(\widehat{\chi}^{(t)}_{F,G}(\boldsymbol{v}^{\,1\leftarrow w}), \dots, \widehat{\chi}^{(t)}_{F,G}(\boldsymbol{v}^{\,k\leftarrow w}), \mathbf{atp}_G(\boldsymbol{v}, w)\Big) : w \in V(G)\Big\}\!\Big\}, \tag{12}$$

where $(\boldsymbol{v}, w) \coloneqq (v_1, \dots, v_k, w)$, recording the atomic type of a $(k+1)$-tuple.

Lemma C.4. *For $k \ge 2$, the augmented k-FWL refinement is as expressive as the standard k-FWL refinement.*

For $k = 1$, the same formula gives an all-node formulation of 1-WL:

$$\widehat{\chi}^{(l+1)}_G(u) = \mathrm{hash}\left(\widehat{\chi}^{(l)}_G(u), \Big\{\!\Big\{\big(\widehat{\chi}^{(l)}_G(v), \mathbf{atp}_G(u, v)\big) : v \in V_G\Big\}\!\Big\}\right), \tag{13}$$

which has the same graph-distinguishing power as standard 1-WL (Zhang et al., 2024b).

For $k = 1$, the spectralized 1-WL refinement, abbreviated as Spec 1-WL, starts from $\widetilde{\chi}_G^{(0)}(v) = \mathbf{Satp}_G(v)$ and is defined by

$$\widetilde{\chi}_G^{(t+1)}(v) = \mathrm{hash}\left(\widetilde{\chi}_G^{(t)}(v), \widetilde{M}_G^{(t)}(v)\right), \quad \widetilde{M}_G^{(t)}(v) = \left\{\!\!\left\{ \left(\widetilde{\chi}_G^{(t)}(w), \mathbf{Satp}_G(v,w)\right) : w \in V(G) \right\}\!\!\right\}.$$

Proof of Lemma C.4. Let $\chi_{F,G}^{(t)}$ and $\widehat{\chi}_{F,G}^{(t)}$ be the round-t colors of standard and augmented k-FWL, respectively. We show by induction on t that each determines the other by a graph-independent transformation. Both initial colors are $\mathbf{atp}_G(\boldsymbol{v})$, so the assertion holds at $t = 0$.

Suppose it holds at round t. Each augmented multiset element is

$$\left(\widehat{\chi}_{F,G}^{(t)}(\boldsymbol{v}^{1\leftarrow w}), \ldots, \widehat{\chi}_{F,G}^{(t)}(\boldsymbol{v}^{k\leftarrow w}), \mathbf{atp}_G(\boldsymbol{v}, w)\right).$$

Discarding its last coordinate and converting the child colors by the induction hypothesis gives the corresponding standard multiset element. Thus the augmented update determines the standard update.

For the converse, every refinement color retains its initial atomic type. Fix $\boldsymbol{v} = (v_1, \ldots, v_k)$ and $w \in V(G)$. The parent color $\chi_{F,G}^{(t)}(\boldsymbol{v})$ determines all relations among $v_1, \ldots, v_k$. For each $j \in [k]$, choose $i \neq j$, which is possible because $k \geq 2$. The tuple $\boldsymbol{v}^{i\leftarrow w}$ contains both w and v_j. Its atomic type, retained by its child color, determines whether $w = v_j$ and whether w is adjacent to v_j. Doing this for every j reconstructs $\mathbf{atp}_G(\boldsymbol{v}, w)$.

Each augmented multiset element can therefore be reconstructed from the corresponding standard joint-replacement element together with the parent color. The induction hypothesis also converts all child colors to their augmented versions. This recovers the entire augmented update, including multiplicities, and completes the induction. The two refinements consequently have the same graph-distinguishing power. □

D Proofs for Section 4

We use the notation of Section 4. In particular, $\lambda_1^G < \cdots < \lambda_{m_G}^G$ are the distinct eigenvalues of T_G, and P_r^G is the orthogonal projector onto $E_r^G := \ker(T_G - \lambda_r^G I_n)$, where $n = |V(G)|$. Thus $T_G = \sum_{r=1}^{m_G} \lambda_r^G P_r^G$ and $I_n = \sum_{r=1}^{m_G} P_r^G$.

Proof of Proposition 4.1. For each $i \in [k]$, let

$$\mathbf{T}_i := I_n^{\otimes(i-1)} \otimes T_G \otimes I_n^{\otimes(k-i)}.$$

These operators commute because they act on different tensor factors. Moreover,

$$\mathbf{T}_i = \sum_{\boldsymbol{r} \in [m_G]^k} \lambda_{r_i}^G P_{\boldsymbol{r}}^G, \qquad P_{\boldsymbol{r}}^G = P_{r_1}^G \otimes \cdots \otimes P_{r_k}^G.$$

Consequently, the tensor-product subspaces $E_{r_1}^G \otimes \cdots \otimes E_{r_k}^G$ are their joint eigenspaces, with joint eigenvalue tuple $\boldsymbol{\lambda}_{\boldsymbol{r}}^G = (\lambda_{r_1}^G, \ldots, \lambda_{r_k}^G)$. The joint functional calculus therefore gives

$$g(\mathbf{T}_1, \ldots, \mathbf{T}_k) = \sum_{\boldsymbol{r} \in [m_G]^k} g(\boldsymbol{\lambda}_{\boldsymbol{r}}^G) P_{\boldsymbol{r}}^G.$$

To identify this with the spectral filter, write $U_G = [U_1^G, \ldots, U_{m_G}^G]$, where the columns of U_r^G form an orthonormal basis of E_r^G. Then $U_G^{\otimes k}$ is an orthonormal joint eigenbasis, and the defining spectral multiplier of $\mathcal{F}_{G,g}^{(k)}$ on the joint eigenspace indexed by $\boldsymbol{r}$ is $g(\boldsymbol{\lambda}_{\boldsymbol{r}}^G)$. Hence $\mathcal{F}_{G,g}^{(k)} = g(\mathbf{T}_1, \ldots, \mathbf{T}_k)$, proving the claimed expansion. □

D.1 Proof of Theorem 4.4

Proof of Theorem 4.4. By Proposition 4.1, for every $g \in C(\mathbb{R}^k)$,

$$\big[\mathcal{F}^{(k)}_{G,g}\Phi_G\big](\boldsymbol{v}) = \sum_{\boldsymbol{r}\in[m_G]^k} g(\boldsymbol{\lambda}^G_{\boldsymbol{r}})\big[P^G_{\boldsymbol{r}}\Phi_G\big](\boldsymbol{v}).$$

Thus $\mathcal{P}^T_{\Phi_k}(G,\boldsymbol{v})$ determines every spectral response and the distinct spectrum, and therefore determines $\mathrm{Spec}^T_{\Phi_k}(G,\boldsymbol{v})$.

Conversely, the spectrum recorded by $\mathrm{Spec}^T_{\Phi_k}(G,\boldsymbol{v})$ determines the Lagrange polynomials

$$\ell^G_r(z) := \prod_{\substack{s\in[m_G]\\ s\neq r}} \frac{z-\lambda^G_s}{\lambda^G_r-\lambda^G_s}, \qquad r\in[m_G],$$

where an empty product equals 1. For $\boldsymbol{r} = (r_1,\dots,r_k)$, define the continuous filter $g^G_{\boldsymbol{r}}(z_1,\dots,z_k) := \prod_{j=1}^k \ell^G_{r_j}(z_j)$. Since $g^G_{\boldsymbol{r}}(\boldsymbol{\lambda}^G_{\mathbf{s}}) = \mathbf{1}[\boldsymbol{r}=\mathbf{s}]$, its recorded response is

$$\big[\mathcal{F}^{(k)}_{G,g^G_{\boldsymbol{r}}}\Phi_G\big](\boldsymbol{v}) = \big[P^G_{\boldsymbol{r}}\Phi_G\big](\boldsymbol{v}).$$

All projector-response coefficients are therefore recovered. Although these interpolation filters depend on the spectrum, that spectrum is part of the invariant being decoded. The construction thus gives a well-defined decoder from spectralization to $\mathcal{P}^T_{\Phi_k}$, proving the equivalence.

For the identity RGI $\Phi_{\mathrm{id}}(G,u,v) = \mathbf{1}[u=v]$, its signal matrix is I_n. Orthogonality and symmetry of the projectors give

$$\big[(P^G_i\otimes P^G_j)\Phi_{\mathrm{id},G}\big](u,v) = \big(P^G_jP^G_i\big)(u,v) = \mathbf{1}[i=j]P^G_i(u,v),$$

where the matrix identity uses column-wise vectorization, as in Section 5. It follows that

$$\mathcal{P}^T_{\Phi_{\mathrm{id}}}(G,u,v) = \big\{\big((\lambda^G_i,\lambda^G_j),\mathbf{1}[i=j]P^G_i(u,v)\big) : i,j\in[m_G]\big\}.$$

Selecting the entries with equal eigenvalue coordinates recovers $\mathcal{P}^T(G,u,v)$. Conversely, this pair invariant reconstructs the displayed set by assigning zero to every pair of unequal eigenvalues. Hence $\mathrm{Spec}^T_{\Phi_{\mathrm{id}}} \equiv \mathcal{P}^T_{\Phi_{\mathrm{id}}} \equiv \mathcal{P}^T$. □

D.2 Proof of Theorem 4.7

Throughout this subsection, graphs are finite, nonempty, simple, undirected, and unlabeled. Tuple coordinates are ordered and may repeat. All hashes are injective and shared across graphs, and graph readout is the multiset of stable tuple colors. Stability refers to the induced partition; when comparing graphs, we refine them synchronously with a common color dictionary. Equivalently, they are indistinguishable if their color multisets agree at every finite round. Spectral information is exact and uses the orthogonal projector onto the complete eigenspace of each distinct eigenvalue. We use the recursive Spec refinement of Section 4.2, so its child colors at round t are $\widetilde{\chi}^{(t)}_G$.

For a graph F, write $\mathrm{tw}(F)$ for its treewidth. A tree decomposition consists of a tree with vertex bags that cover all vertices and edges of F, such that the bags containing each fixed vertex induce a connected subtree. Its width is one less than the largest bag size, and treewidth is the minimum such width. We also write $\mathbb{F}_2 = \{0,1\}$ for the field with two elements, and use $\oplus$ for addition in this field.

We prove the two parts of the theorem in four steps. The first two establish the lower and upper bounds for Spec k-FWL. The last two establish the strict bounds for SI k-OWL, using the adjacency operator as specified in the main text.

D.2.1 The First Separation

We first prove the strict separation between k-FWL and its spectralized counterpart. We begin with the clique-counting response used to construct the witnesses. The notation $\mathrm{sub}(K_{r+1},G;u)$ counts copies of K_{r+1} in G containing u.

Proposition D.1. *Let $r \geq 2$, and let $\kappa_r \in \mathbf{AT}_r$ be the rooted isomorphism type represented by K_r with its r distinct vertices as ordered roots. Thus $\Phi_{\kappa_r}(G, \mathbf{x})$ is the indicator that the entries of $\mathbf{x}$ are distinct and span a clique. Write $\mathrm{diag}_r(u) := (u, \dots, u)$. If G is d-regular, then for every $T \in \{A, L, \widetilde{L}\}$ there is a response $g_{T,d} \in C(\mathbb{R}^r)$, common to all d-regular graphs, such that*

$$\left[\mathcal{F}^{(r)}_{G,g_{T,d}} \Phi_{\kappa_r}(G, \cdot)\right](\mathrm{diag}_r(u)) = r!\, \mathrm{sub}(K_{r+1}, G; u) \qquad (u \in V(G)).$$

For $T = A$, regularity is unnecessary and the response is $g_A(z_1, \dots, z_r) = \prod_{j=1}^r z_j$. The spectralized atomic type at $\mathrm{diag}_r(u)$ therefore determines this rooted clique count on the stated graph class.

Proof. If $d = 0$, then G is edgeless, the signal $\Phi_{\kappa_r}(G, \cdot)$ is zero, and the assertion follows by taking the zero response. Suppose henceforth that $d > 0$, and set

$$h_{T,d}(z) = \begin{cases} z, & T = A, \\ d - z, & T = L, \\ d(1 - z), & T = \widetilde{L}, \end{cases} \qquad g_{T,d}(z_1, \dots, z_r) = \prod_{j=1}^r h_{T,d}(z_j).$$

Since $L_G = dI - A_G$ and $\widetilde{L}_G = I - A_G/d$, we have $h_{T,d}(T) = A_G$. The tensor spectral calculus therefore gives

$$\mathcal{F}^{(r)}_{G,g_{T,d}} = A_G^{\otimes r}.$$

At a diagonal tuple, the resulting response is

$$\sum_{x_1,\dots,x_r \in V(G)} \left(\prod_{j=1}^r A_G(u, x_j) \right) \Phi_{\kappa_r}(G, x_1, \dots, x_r).$$

A summand equals one exactly when $x_1, \dots, x_r$ are distinct, span K_r, and are all adjacent to u. Since G has no loops, these conditions are equivalent to $\{u, x_1, \dots, x_r\}$ spanning a copy of K_{r+1}. Each copy containing u contributes $r!$ ordered tuples, giving the claimed identity. When $T = A$, the same argument applies to an arbitrary graph because the response $\prod_j z_j$ always yields $A_G^{\otimes r}$. Finally, the spectralized atomic type contains the response of every atomic-type channel to every continuous filter, so it determines the count just computed. □

Proposition D.2. *For every $k \geq 1$ and every standard graph operator $T \in \{A, L, \widetilde{L}\}$,*

$$k\text{-FWL} \prec \text{Spec } k\text{-FWL},$$

where 1-FWL and Spec 1-FWL refer to 1-WL and Spec 1-WL, respectively.

We first record the elementary recovery property of spectralized atomic types.

Lemma D.3. *For every $r \geq 1$ and $\mathbf{x} \in V(G)^r$, the spectralized atomic type $\mathbf{Satp}_G(\mathbf{x})$ determines $\mathbf{atp}_G(\mathbf{x})$.*

Proof. Write $n = |V(G)|$. Since the eigenspace projectors form a resolution of the identity, $\sum_{i=1}^{m_G} P_i^G = I_n$, and hence

$$\sum_{\mathbf{i} \in [m_G]^r} P_{\mathbf{i}}^G = \left(\sum_{i=1}^{m_G} P_i^G \right)^{\otimes r} = I_n^{\otimes r}.$$

Therefore, for every $\tau \in \mathbf{AT}_r$,

$$\sum_{\mathbf{i} \in [m_G]^r} \left[P_{\mathbf{i}}^G \Phi_\tau(G, \cdot) \right](\mathbf{x}) = \Phi_\tau(G, \mathbf{x}).$$

The recovered channel values form precisely the one-hot vector $\mathbf{atp}_G(\mathbf{x}) = (\Phi_\tau(G, \mathbf{x}))_{\tau \in \mathbf{AT}_r}$. Thus $\mathbf{Satp}_G(\mathbf{x})$ determines $\mathbf{atp}_G(\mathbf{x})$. □

Proof of Proposition D.2. We first prove the non-strict direction.

For $k \geq 2$, use the augmented formulation of k-FWL from Lemma C.4. Let $\widehat{\chi}^{(t)}_{F,G}$ denote its round-t color and $\widetilde{\chi}^{(t)}_{G}$ the corresponding Spec k-FWL color. At round zero, Lemma D.3 implies that $\widetilde{\chi}^{(0)}_{G}(\boldsymbol{v})$ determines $\widehat{\chi}^{(0)}_{F,G}(\boldsymbol{v})$.

Suppose inductively that every round-t spectralized color determines the corresponding augmented k-FWL color. Each element of the spectralized replacement multiset contains $\mathbf{Satp}_G(\boldsymbol{v}, w)$, which determines $\mathbf{atp}_G(\boldsymbol{v}, w)$ by Lemma D.3; the remaining coordinates determine the k replacement colors by the induction hypothesis. Hence $\widetilde{\chi}^{(t+1)}_{G}(\boldsymbol{v})$ determines $\widehat{\chi}^{(t+1)}_{F,G}(\boldsymbol{v})$. By Lemma C.4,

$$k\text{-FWL} \preceq \text{Spec } k\text{-FWL}.$$

For $k = 1$, the same argument applies directly. From each relation label $\mathbf{Satp}_G(u, w)$, it recovers $\mathbf{atp}_G(u, w)$, and hence identifies exactly those messages with $w \in \mathcal{N}_G(u)$. Inductively recovering the child colors therefore recovers the ordinary 1-WL neighborhood multiset. Thus 1-WL $\preceq$ Spec 1-WL. We then prove strictness.

Strictness for $k \geq 2$. Fix $k \geq 2$, and put $q := k + 1$ and $s := k + 2 = q + 1$.

Fürer witnesses. We use the Fürer-graph construction in Zhang et al. (2024a, Definitions C.44–C.45). This construction goes back to Fürer and is a variant of the classical CFI construction. For completeness, we state the specialization needed here.

For a twist function $\sigma : E(K_s) \to \mathbb{F}_2$, define Γ^{σ}_s with

$$V(\Gamma^{\sigma}_s) = \big\{(v, S) : v \in V(K_s),\ S \subseteq \mathcal{N}_{K_s}(v),\ |S| \equiv 0 \pmod 2\big\}.$$

For a base edge $vw \in E(K_s)$, two states (v, S) and (w, R) are adjacent precisely when

$$\mathbf{1}[w \in S] \oplus \mathbf{1}[v \in R] = \sigma(vw).$$

Let $\sigma_0 \equiv 0$, and let σ_1 be 1 on one fixed base edge and 0 elsewhere. We write $\Gamma^0_s := \Gamma^{\sigma_0}_s$ and $\Gamma^1_s := \Gamma^{\sigma_1}_s$. Thus Γ^0_s and Γ^1_s are respectively the untwisted and a one-edge-twisted Fürer graph of K_s.

Taking all base-vertex labels to be equal gives the unlabeled graphs used here. The model $F(k)$ in Zhang et al. (2024a, Appendix E.1) is exactly the k-FWL used here. Their Theorems E.6 and E.9 imply that whenever $\mathrm{tw}(F) > k$, the untwisted and one-edge-twisted Fürer graphs of F are indistinguishable by k-FWL. Since $\mathrm{tw}(K_s) = s - 1 = k + 1 > k$, we obtain

$$\Gamma^0_s \equiv_{k\text{-FWL}} \Gamma^1_s.$$

We therefore only need to show that these witnesses satisfy the regularity required by Proposition D.1, and that their K_s-counts differ.

Regularity and connectedness. For each $v \in V(K_s)$, let

$$F_v := \{(v, S) : S \subseteq \mathcal{N}_{K_s}(v),\ |S| \equiv 0 \pmod 2\}$$

be the fiber over v. Since $|\mathcal{N}_{K_s}(v)| = s - 1$, $|F_v| = 2^{s-2}$.

Fix $(v, S) \in F_v$ and a base neighbor $w \neq v$. The adjacency condition fixes the single bit $\mathbf{1}[v \in R]$ of a state $(w, R) \in F_w$. Among the remaining $s - 2$ bits, one parity constraint remains, so there are exactly 2^{s-3} valid choices for R. Hence every vertex has degree

$$d_s = (s - 1)2^{s-3}.$$

Thus both Γ^0_s and Γ^1_s are d_s-regular.

They are also connected. Let (v, S) and (w, R) be arbitrary states with $v \neq w$. Choose distinct $a, b \in V(K_s) \setminus \{v, w\}$, which is possible since $s \geq 4$. Choose an arbitrary bit $\varepsilon \in \mathbb{F}_2$. A state (a, S_a) can be chosen so that its bit on av satisfies the adjacency condition with (v, S), its bit on ab equals ε, and its remaining bits satisfy even parity. Such a choice exists because a has $s - 1 \geq 3$ incident bits and only two have been prescribed. Similarly, choose (b, S_b) so that its bit on ab

satisfies the adjacency condition with (a, S_a), its bit on bw satisfies the adjacency condition with (w, R), and its remaining bits satisfy even parity. Therefore

$$(v, S) \sim (a, S_a) \sim (b, S_b) \sim (w, R).$$

If $v = w$, the same construction applied to the base walk $v - a - b - v$ connects (v, S) to (v, R). Hence every Γ_s^σ, and in particular Γ_s^0, Γ_s^1, is connected.

Clique counts. There are no edges within a fiber, so any K_s in Γ_s^σ must choose exactly one state (v, S_v) from every fiber F_v. For every incidence of a base edge $e = vw$, write $x_{v,e} := \mathbf{1}[w \in S_v]$. Since every S_v has even cardinality, $\sum_{e \ni v} x_{v,e} = 0$ in $\mathbb{F}_2$ for every base vertex v. Pairwise adjacency of the selected states is equivalent to $x_{v,e} \oplus x_{w,e} = \sigma(e)$ for every $e = vw$.

For Γ_s^0, the two incidence bits of each edge agree, so they define a global edge set $X \subseteq E(K_s)$. The parity conditions say exactly that every vertex has even degree in X. Thus X belongs to the binary cycle space of K_s. Since K_s is connected, this cycle space has dimension $\binom{s}{2} - s + 1$. Consequently,

$$\mathrm{sub}(K_s, \Gamma_s^0) = 2^{\binom{s}{2} - s + 1}.$$

For Γ_s^1, summing all vertex-parity equations gives

$$0 = \sum_v \sum_{e \ni v} x_{v,e} = \sum_{e=vw} \big(x_{v,e} + x_{w,e}\big) = \sum_e \sigma_1(e) = 1 \qquad \text{in } \mathbb{F}_2,$$

a contradiction. Hence

$$\mathrm{sub}(K_s, \Gamma_s^1) = 0.$$

The general homomorphism-count separation of the untwisted and twisted Fürer graphs is already established in Zhang et al. (2024a, Theorem E.6); the computation above records the stronger exact clique counts needed for our explicit separator.

The witnesses are cospectral. Although cospectrality is not required for strictness, we record it to show that the separation is not caused merely by different eigenvalue multisets.

Let $h := 2^{s-2}$, $m := \binom{s}{2}$, and $d_s = (s-1)h/2$. For every fiber F_v, define $a_v := h^{-1/2}\mathbf{1}_{F_v}$. For every base edge $e = vw$ incident to v, let $b_{v,e}$ be supported on F_v and defined by

$$b_{v,e}(v, S) = h^{-1/2}(-1)^{\mathbf{1}[w \in S]}.$$

Since $s - 1 \geq 3$, the constant character and the singleton-coordinate characters are mutually orthogonal on the even-parity fiber. Thus $\{a_v\} \cup \{b_{v,e} : e \ni v\}$ is an orthonormal family.

Let A_σ be the adjacency matrix of Γ_s^σ. The adjacency condition gives the exact matrix identity

$$A_\sigma = \frac{h}{2} \sum_{e=vw \in E(K_s)} \Big(a_v a_w^\top + a_w a_v^\top + (-1)^{\sigma(e)}(b_{v,e} b_{w,e}^\top + b_{w,e} b_{v,e}^\top)\Big).$$

On $U_0 := \mathrm{span}\{a_v : v \in V(K_s)\}$, A_σ is $h/2$ times the adjacency matrix of K_s. It therefore contributes the eigenvalue d_s once and $-h/2$ with multiplicity $s - 1$.

For every base edge $e = vw$, the space $U_e := \mathrm{span}\{b_{v,e}, b_{w,e}\}$ is invariant, and the restriction of A_σ to U_e is

$$\begin{pmatrix} 0 & (-1)^{\sigma(e)} h/2 \\ (-1)^{\sigma(e)} h/2 & 0 \end{pmatrix},$$

whose eigenvalues are $h/2$ and $-h/2$, independently of $\sigma(e)$. The orthogonal complement of $U_0 \oplus \bigoplus_{e \in E(K_s)} U_e$ lies in the kernel of A_σ. The eigenvalues and their multiplicities are therefore

eigenvalue	d_s	$h/2$	$-h/2$	0
multiplicity	1	m	$m + s - 1$	$sh - s^2$

(with the zero eigenvalue omitted when its multiplicity is zero). These multiplicities are independent of the twist function σ. Thus Γ_s^0 and Γ_s^1 are adjacency-cospectral. Since both graphs are d_s-regular, $L = d_s I - A$ and $\widetilde{L} = I - A/d_s$; therefore they are also cospectral for both Laplacian choices.

Separation by one spectralized refinement. Define $\mathrm{diag}_r(u) := (u, \dots, u) \in V(G)^r$. Consider the first Spec k-FWL update at $\mathrm{diag}_k(u)$. Its relation multiset contains $\mathbf{Satp}_G(\mathrm{diag}_k(u), w)$ for every $w \in V(G)$. By Lemma D.3, each such label determines $\mathbf{atp}_G(\mathrm{diag}_k(u), w)$. Hence the unique message for which all q coordinates are equal is exactly the one with $w = u$. Therefore the round-one color of $\mathrm{diag}_k(u)$ determines $\mathbf{Satp}_G(\mathrm{diag}_q(u))$.

Both witnesses are d_s-regular. Applying Proposition D.1 at order $q = s - 1$, there exists a common spectral response g_{T,d_s} such that

$$\Theta_G^T(u) := \left[\mathcal{F}^{(q)}_{G,g_{T,d_s}} \Phi_{\kappa_q}(G, \cdot)\right](\mathrm{diag}_q(u)) = q!\,\mathrm{sub}(K_s, G; u).$$

Thus the round-one color of each diagonal k-tuple determines $\Theta_G^T(u)$. Since the initial spectralized color determines the atomic type and every refinement retains its previous color, the graph-level readout can identify exactly the diagonal k-tuples. Consequently it determines

$$\mathrm{sub}(K_s, G) = \frac{1}{s!} \sum_{u \in V(G)} \Theta_G^T(u).$$

Since

$$\mathrm{sub}(K_s, \Gamma_s^0) = 2^{\binom{s}{2}-s+1} \neq 0 = \mathrm{sub}(K_s, \Gamma_s^1),$$

the two graphs are distinguished after one Spec k-FWL refinement round, for every $T \in \{A, L, \widetilde{L}\}$. Together with $\Gamma_s^0 \equiv_{k\text{-FWL}} \Gamma_s^1$, this proves strictness for every $k \geq 2$.

Strictness for $k = 1$. Let $G = C_3 \square K_2$ and $H = K_{3,3}$. Both are 3-regular graphs on six vertices, so ordinary 1-WL with constant initial colors does not distinguish them. The prism contains two triangles, whereas $K_{3,3}$ contains none.

At the first Spec 1-WL update of a vertex u, the relation multiset contains $\mathbf{Satp}_G(u, w)$ for all w. Lemma D.3 identifies the unique element with $w = u$, and hence recovers $\mathbf{Satp}_G(u, u)$. Proposition D.1 with order 2 then determines the number of triangles containing u. The graph-level readout therefore determines the total triangle count, and distinguishes the triangular prism from $K_{3,3}$.

This proves

$$k\text{-FWL} \prec \text{Spec } k\text{-FWL}$$

for every $k \geq 1$ and every $T \in \{A, L, \widetilde{L}\}$. □

D.2.2 An Example at Order Two

We give a classical concrete witness for the strict separation between 2-FWL and Spec 2-FWL.

Let R denote the 4×4 rook graph,

$$R = K_4 \square K_4,$$

and let Sh denote the Shrikhande graph, which may be realized as the Cayley graph

$$\mathrm{Sh} = \mathrm{Cay}\left(\mathbb{Z}_4^2, \{\pm(1,0), \pm(0,1), \pm(1,1)\}\right).$$

Both R and Sh are connected strongly regular graphs with parameters

$$(n, d, \lambda, \mu) = (16, 6, 2, 2).$$

Graphs with the same strongly regular parameters are indistinguishable by joint 2-FWL. Indeed, the initial pair color records equality, adjacency, or distinct non-adjacency. For each of these three relations, the numbers of vertices w realizing each ordered pair of relations between w and the two coordinates depend only on (n, d, λ, μ). Consequently, every refinement round assigns the same color to all pairs having the same relation, in either graph. The three relation classes also have the same sizes in the two graphs. This proves $R \equiv_{2\text{-FWL}} \mathrm{Sh}$; see also Papp & Wattenhofer (2022, Sec. 4.1 and Appendix G).

The relevant structural difference can also be seen directly. In the rook graph, the four vertices in any row or column form a K_4, and every clique of size at least three lies in a single row or column. Thus each vertex lies in exactly two copies of K_4. In the Shrikhande graph, the neighborhood of

every vertex induces a six-cycle, as follows from the displayed Cayley representation. It therefore contains no triangle, so no K_4 can contain that vertex. Hence, for every $u \in V(R)$ and $v \in V(\mathrm{Sh})$,

$$\mathrm{sub}(K_4, R; u) = 2, \qquad \mathrm{sub}(K_4, \mathrm{Sh}; v) = 0.$$

We now show that this difference is detected by one spectralized refinement round.

Corollary D.4. *For every $T \in \{A, L, \widetilde{L}\}$,*

$$R \equiv_{\text{2-FWL}} \mathrm{Sh}, \qquad R \not\equiv_{\text{Spec 2-FWL}} \mathrm{Sh}.$$

Moreover, Spec 2-FWL distinguishes R and Sh after one refinement round.

Proof. The 2-FWL equivalence was established above.

It remains to prove the spectralized separation. Fix a graph $G \in \{R, \mathrm{Sh}\}$ and $u \in V(G)$. Consider the diagonal pair (u, u). Its first Spec 2-FWL refinement contains, for each $w \in V(G)$, the relation label

$$\mathbf{Satp}_G(u, u, w).$$

By Lemma D.3, $\mathbf{Satp}_G(u, u, w)$ determines $\mathbf{atp}_G(u, u, w)$. Hence the unique message for which all three coordinates are equal is exactly the one corresponding to $w = u$. Therefore the round-one color of (u, u) determines

$$\mathbf{Satp}_G(u, u, u).$$

Both R and Sh are 6-regular. Applying Proposition D.1 with order 3, there exists a spectral response $g_{T,6}$ such that

$$\left[\mathcal{F}^{(3)}_{G, g_{T,6}} \Phi_{\kappa_3}(G, \cdot)\right](u, u, u) = 3!\ \mathrm{sub}(K_4, G; u).$$

Thus $\mathbf{Satp}_G(u, u, u)$, and hence the round-one color of (u, u), determines the rooted K_4-count at u. The displayed response is 12 on every rook-graph diagonal and 0 on every Shrikhande diagonal.

Since every vertex of R lies in a K_4, while no vertex of Sh does, the round-one colors of diagonal pairs in the two graphs are different. Furthermore, Lemma D.3 implies that the initial spectralized color determines whether a pair is diagonal, and each refinement retains its previous color. Hence diagonal pairs can be identified from the graph-level color multiset, so the two graph-level representations differ after one round.

Therefore

$$R \not\equiv_{\text{Spec 2-FWL}} \mathrm{Sh}.$$

□

Remark D.5. The separation is not caused by different eigenvalue multisets. Since R and Sh are strongly regular with the same parameters, they are adjacency-cospectral. As both graphs are 6-regular, they are also cospectral with respect to the combinatorial and normalized Laplacians. Thus the separation comes from spectralizing the triangle atomic-type signal, rather than from the spectrum alone.

D.2.3 The Upper Bound by $(k+1)$-LWL

We prove that $(k+1)$-LWL determines every color produced by Spec k-FWL for each $T \in \{A, L, \widetilde{L}\}$. The argument has three steps. We express spectralized atomic types using rooted homomorphism counts, show that the operations in a spectralized refinement preserve an appropriate class of rooted templates, and then apply the homomorphism characterization of LWL.

Throughout this subsection, fix $T \in \{A, L, \widetilde{L}\}$, $k \geq 2$, and $q = k + 1$. We retain the notation $\widetilde{\chi}^{(t)}_G$ and $\mathbf{Satp}_G$ from Section 4.2. The auxiliary graph families and the imported characterization below are defined for every $q \geq 2$.

We first state precisely the graph family used in the proof.

Definition D.6. A strong canonical decomposition of width q of a finite nonempty graph F is a finite rooted tree $\mathcal{D}$ with a multiset bag $\beta(t)$ of vertices of F at each tree node t, satisfying:

(i) Every vertex occurs in a bag, every edge has both endpoints in a bag, and the nodes whose bags contain a fixed vertex induce a connected subtree.

(ii) An even-depth bag has cardinality q, and an odd-depth bag has cardinality $q+1$, with multiplicities counted.

(iii) If an even-depth node and an odd-depth node are adjacent, the even bag is a multiset-subbag of the odd bag.

(iv) Every odd-depth node has exactly one child.

Let $\mathfrak{S}_q$ denote the family of graphs admitting such a decomposition, and let $\mathcal{S}_q$ denote its connected subfamily. All unions, deletions, and inclusions of bags below are multiset operations; deleting a vertex from a bag removes one occurrence.

This is the family $S^{L(q)}$ of Zhang et al. (2024a, Definitions C.4 and E.4(c)), with the decomposition itself suppressed from the notation. Its leaves have even depth. Equivalently, one may allow an odd node to have at most one child: an odd leaf can be given a dummy even child consisting of any q entries of its bag.

Theorem D.7. *For every $q \geq 2$ and finite nonempty simple graphs G, H,*

$$G \equiv_{q\text{-LWL}} H \quad \Longleftrightarrow \quad \hom(F,G) = \hom(F,H) \quad \textit{for every } F \in \mathfrak{S}_q. \tag{14}$$

Moreover, $\mathfrak{S}_q$ is the family of graphs admitting a componentwise q-order strong nested-ear decomposition. If a connected graph F does not belong to $\mathcal{S}_q$, its untwisted and single-edge-twisted Fürer graphs $G(F), H(F)$ satisfy

$$G(F) \equiv_{q\text{-LWL}} H(F), \qquad \hom(F,G(F)) \neq \hom(F,H(F)). \tag{15}$$

These statements are, respectively, Zhang et al. (2024a, Theorems E.5, E.8, and E.6). The q-LWL update in (10) corresponds to Local q-GNN in that reference. Isolated vertices are included in $\mathfrak{S}_q$ through the decomposition definition.

Rooted templates and spectral moments. An r-rooted template is a finite simple graph F with an ordered root tuple $\mathbf{x} = (x_1, \ldots, x_r)$; entries may repeat. Write

$$h^G_{(F,\mathbf{x})}(\boldsymbol{v}) := \big|\{f : F \to G \text{ a homomorphism} : f(x_i) = v_i \text{ for every } i \in [r]\}\big|.$$

Thus a repeated root enforces equality of its corresponding host coordinates. No injectivity is imposed on the homomorphism. Template families are understood up to rooted isomorphism.

For $1 \leq r \leq q$, let $\mathcal{T}_{q,r}$ be the class of r-rooted templates admitting a strong canonical width-q decomposition whose root bag is

$$\rho_{q,r}(\mathbf{x}) := \{\{x_1, \ldots, x_r, \underbrace{x_r, \ldots, x_r}_{q-r \text{ extra copies}}\}\}. \tag{16}$$

We will use $r = k$ and $r = q$. The distinguished roots are vertices of the template; the extra bag entries in (16) introduce no additional vertices.

For an atomic type $\tau \in \mathbf{AT}_r$ and $\mathbf{a} = (a_1, \ldots, a_r) \in \mathbb{N}_0^r$, where $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$, write $\Phi_{\tau,G} := \Phi_\tau(G, \cdot)$ and define

$$M^G_{\tau,\mathbf{a}}(\boldsymbol{v}) := \sum_{\mathbf{z} \in V(G)^r} \Phi_\tau(G, \mathbf{z}) \prod_{j=1}^{r} (T^{a_j})_{v_j z_j}. \tag{17}$$

Under columnwise vectorization, this is the response to

$$g_{\mathbf{a}}(z_1, \ldots, z_r) = \prod_{j=1}^{r} z_j^{a_{r+1-j}}.$$

The coordinate reversal follows the Kronecker-factor ordering and does not change the family of responses as $\mathbf{a}$ varies.

Definition D.8. Let $\mathcal{B}_r$ consist of the following r-rooted templates. Choose a finite simple core graph K with ordered ports $z_1, \dots, z_r$ covering all its vertices; ports may repeat. For each i, attach a path of length $a_i \in \mathbb{N}_0$ from an outer root x_i to z_i. If $a_i = 0$, set $x_i = z_i$. If $a_i > 0$, every vertex of this path except z_i, including x_i, is fresh. These fresh vertices are distinct for different arms. In particular, different arms can meet only at their core endpoints. Let $\mathcal{B}_r^+$ be the family obtained by attaching any finite number of fresh pendant leaves to each vertex of a template in $\mathcal{B}_r$, retaining its ordered roots. Zero leaves are allowed.

To handle all three graph operators within the same template family, write $d_G(v)$ for the degree of v and $D_G = \mathrm{diag}(d_G(v))$. Each operator has the form

$$T_G = \alpha(D_G) + \varepsilon\beta(D_G)A_G\beta(D_G), \tag{18}$$

where the scalar functions and sign are

$$(\alpha(d), \beta(d), \varepsilon) = \begin{cases} (0, 1, 1), & T = A, \\ (d, 1, -1), & T = L, \\ (\mathbf{1}[d > 0], b(d), -1), & T = \widetilde{L}, \end{cases} \qquad b(d) = \begin{cases} d^{-1/2}, & d > 0, \\ 0, & d = 0. \end{cases}$$

Here the normalized Laplacian is $D_G^{-1/2} L_G D_G^{-1/2}$, with zero inverse degree at isolated vertices. The alternative convention $I - D_G^{-1/2} A_G D_G^{-1/2}$ is covered by taking $\alpha(d) = 1$ instead.

For a fixed pair of finite graphs G, H, both α and β agree with real polynomials on the finite set

$$\Delta_{G,H} := \{d_X(v) : X \in \{G, H\},\ v \in V(X)\}.$$

We may therefore replace them by these polynomials in any computation on G or H, using the same coefficients for both graphs. A factor $d_G(f(x))^j$ in a homomorphism sum is represented by attaching j fresh leaves to the template vertex x: after $f(x)$ is fixed, each leaf has independently $d_G(f(x))$ possible images. Thus polynomial degree weights become finite real linear combinations of ordinary pendant-decorated homomorphism counts. The interpolating polynomials may depend on G, H; this is sufficient because every comparison below concerns those same two graphs. No bound on the degrees of the graph class is assumed.

Lemma D.9. *Suppose T_G and T_H have the same distinct eigenvalues. Equality of all $\mathcal{B}_r^+$ rooted counts at $\boldsymbol{v} \in V(G)^r$ and $\boldsymbol{u} \in V(H)^r$ implies*

$$\mathbf{Satp}_G(\boldsymbol{v}) = \mathbf{Satp}_H(\boldsymbol{u}).$$

Furthermore, $\mathcal{B}_r^+ \subseteq \mathcal{T}_{q,r}$ for every $1 \le r \le q$.

Proof. Let $\lambda_1, \dots, \lambda_m$ be the common distinct eigenvalues, and let P_i^G, P_i^H be the corresponding eigenspace projectors. For $\mathbf{i} = (i_1, \dots, i_r) \in [m]^r$, define the coordinatewise projector response

$$C^G_{\tau,\mathbf{i}}(\boldsymbol{v}) := \sum_{\mathbf{z} \in V(G)^r} \Phi_\tau(G, \mathbf{z}) \prod_{j=1}^r P^G_{i_j}(v_j, z_j).$$

With columnwise vectorization, this corresponds to the Kronecker product $P^G_{i_r} \otimes \cdots \otimes P^G_{i_1}$. For each atomic-type channel, spectral resolution gives

$$M^G_{\tau,\mathbf{a}}(\boldsymbol{v}) = \sum_{i_1,\dots,i_r=1}^m \lambda_{i_1}^{a_1} \cdots \lambda_{i_r}^{a_r} C^G_{\tau,(i_1,\dots,i_r)}(\boldsymbol{v}). \tag{19}$$

Taking $0 \le a_j < m$ makes this an invertible tensor product of Vandermonde transformations. Consequently, equality of these moments recovers every projector-response coefficient and hence every continuous spectral response. Together with the common eigenvalues, these responses determine $\mathbf{Satp}$.

It remains to express the moments using the stated templates. Using the indicator definition of Φ_τ gives

$$M^G_{\tau,\mathbf{a}}(\boldsymbol{v}) = \sum_{\substack{\mathbf{z} \in V(G)^r \\ \Phi_\tau(G,\mathbf{z}) = 1}} \prod_{j=1}^r (T^{a_j})_{v_j z_j}. \tag{20}$$

Expand each factor $(T^{a_j})_{v_j z_j}$ using (18). At each matrix multiplication, choose either a diagonal factor, which keeps the current vertex fixed, or an adjacency factor, which takes an edge step. After collapsing the diagonal steps, the corresponding template is a path of at most a_j edges from x_j to z_j. Its vertices carry products of α and β evaluated at their image degrees, and the coefficient includes the sign ε of every edge step. The polynomial substitutions above turn these vertex weights into finite sums of degree monomials, represented by fresh pendant leaves. Zero edge steps give a zero-length arm and impose $x_j = z_j$.

Next identify ports that τ requires to be equal and insert the edges required by τ. Enforce its remaining port inequalities and forbidden adjacencies by expanding $\mathbf{1}[a \neq b] = 1 - \mathbf{1}[a = b]$ and $\mathbf{1}[\{a, b\} \notin E(G)] = 1 - A_G(a, b)$. Every term identifies additional core ports or inserts additional core edges, with pendant leaves retained. Parallel edges can be collapsed on simple targets, and loop-producing terms vanish. Hence every moment in (20) is a finite real linear combination of $\mathcal{B}_r^+$ rooted counts. The coefficients are the same for G and H. Equal rooted counts therefore give equal moments and prove the first assertion.

For the decomposition claim, start with the multiset bag of the r outer roots. Advance the arms one edge at a time. If the current entry on an arm is x and its next vertex is z, use the chain

$$B_{\text{even}} \longrightarrow (B \uplus \{\{z\}\})_{\text{odd}} \longrightarrow ((B \setminus \{\{x\}\}) \uplus \{\{z\}\})_{\text{even}}. \tag{21}$$

The odd bag covers the traversed arm edge. After all arms have advanced, the final even bag contains all core ports and therefore covers every core edge. Each arm vertex occurs along a contiguous part of this chain. A core vertex, once reached by any arm, remains in a bag entry until the end, which also handles repeated ports. This is a strong canonical width-r decomposition. To obtain width q, add $q - r$ copies of the fixed vertex x_r to *every* bag. The occurrence set of x_r is then the whole chain; other occurrence sets are unchanged. The root bag becomes $\rho_{q,r}(\mathbf{x})$. Every vertex of the undecorated template occurs in an even bag. To attach a fresh leaf y at such a vertex x, choose an even bag B containing x and append a new branch

$$B_{\text{even}} \longrightarrow (B \uplus \{\{y\}\})_{\text{odd}} \longrightarrow B_{\text{even}}. \tag{22}$$

The odd bag covers xy, its only child has the original bag B, and the occurrence set of every old vertex remains connected. Repeating this for all pendant leaves proves $\mathcal{B}_r^+ \subseteq \mathcal{T}_{q,r}$. □

The root product of equally rooted templates takes their disjoint union and identifies roots with the same coordinate label. Root equalities already present in one factor are propagated to all factors. Homomorphisms into a simple graph then give

$$h^G_{R_1 \cdots R_s}(\boldsymbol{v}) = \prod_{j=1}^{s} h^G_{R_j}(\boldsymbol{v}). \tag{23}$$

The empty product consists of r distinct isolated roots and has rooted count one. Any product with a loop has count identically zero and may be discarded. Forgetting a distinguished root retains the underlying vertex and sums over its image; in particular,

$$h^G_{\mathrm{For}_w U}(\boldsymbol{v}) = \sum_{w \in V(G)} h^G_U(\boldsymbol{v}, w). \tag{24}$$

This identity remains valid if the forgotten root equals another root: only the compatible value of w contributes.

Lemma D.10. *The following operations preserve the indicated rooted-template classes.*

(i) *A root product either has identically zero count on simple targets or belongs to $\mathcal{T}_{q,r}$.*

(ii) *If $R \in \mathcal{T}_{q,k}$ and $i \in [k]$, its child lift $\iota_i(R)$, characterized by $h^G_{\iota_i(R)}(\boldsymbol{v}, w) = h^G_R(\boldsymbol{v}^{i \leftarrow w})$, belongs to $\mathcal{T}_{q,q}$.*

(iii) *If $U \in \mathcal{T}_{q,q}$ has roots $(v_1, \ldots, v_k, w)$, then $\mathrm{For}_w U \in \mathcal{T}_{q,k}$.*

Forgetting all distinguished roots of any template in these classes leaves an unrooted graph in $\mathfrak{S}_q$.

Proof. We first address the identifications in a root product. Within each factor, all root vertices occur in its root bag. Identifying any collection of these vertices preserves the running-intersection condition: the new occurrence set is a union of connected subtrees that all contain the root node. Applying the quotient map to each multiset bag preserves its cardinality and all multiset inclusions. Edge coverage is also preserved, except that loop-producing terms are discarded as zero. After all root identifications, the factors have the same root multiset. Identify their root tree nodes. This joins the decomposition trees at a single even node, so each odd node still has one child. Shared vertices occur in the common root bag, and their occurrence subtrees remain connected. This proves (i), including products whose factors impose different root equalities.

For (ii), rename the roots of R to $(v_1, \dots, v_{i-1}, w, v_{i+1}, \dots, v_k)$ and add the missing v_i as a fresh isolated distinguished vertex. Let d be the last root of this child tuple, which is the duplicated padding entry in its existing root bag, and put $E = \{\{v_1, \dots, v_k, w\}\}$. Prepend the chain

$$E_{\text{even}} \longrightarrow (E \uplus \{\{d\}\})_{\text{odd}} \longrightarrow ((E \setminus \{\{v_i\}\}) \uplus \{\{d\}\})_{\text{even}}, \tag{25}$$

where the last node is the original root node of the decomposition. All previously occurring vertices in the added bags also occur in that original root bag; the new isolated vertex v_i occurs only in the first two bags. Hence running intersection is preserved. Both bag inclusions and the one-child condition are immediate.

For (iii), put $R = \{\{v_1, \dots, v_k, v_k\}\}$ and prepend

$$R_{\text{even}} \longrightarrow (R \uplus \{\{w\}\})_{\text{odd}} \longrightarrow E_{\text{even}} \tag{26}$$

to the decomposition of U. Its last node is the old root node. This gives the required root bag after forgetting w, while preserving edge coverage and running intersection. Multiset notation makes the construction valid even when some roots coincide. Finally, forgetting all root labels does not change the underlying decomposition at all. □

The following elementary fact explains why products suffice to reconstruct an entire refinement multiset.

Lemma D.11. *Let I be any index set. Two finite multisets of vectors in $\mathbb{R}^I$ are equal if and only if the sums of all monomials in finitely many coordinates are equal, including the constant monomial.*

Proof. The forward implication is immediate. For the converse, the union of the two finite supports can be separated by a finite set of coordinates. On this finite projection, for each support point a and every other support point b, choose a coordinate $j(a, b)$ on which they differ. The polynomial

$$p_a(z) = \prod_{b \neq a} \frac{z_{j(a,b)} - b_{j(a,b)}}{a_{j(a,b)} - b_{j(a,b)}}$$

equals one at a and zero at the other support points. Equality of the monomial sums gives equality of the sums of p_a, and hence the multiplicity of a. This holds for every support point. □

Lemma D.12. *Suppose T_G, T_H have a common spectrum. If*

$$h_R^G(\boldsymbol{v}) = h_R^H(\boldsymbol{u}) \qquad \text{for all } R \in \mathcal{T}_{q,k}, \tag{27}$$

then $\widetilde{\chi}_G^{(t)}(\boldsymbol{v}) = \widetilde{\chi}_H^{(t)}(\boldsymbol{u})$ for every $t \geq 0$.

Proof. At round zero, the claim follows from $\mathcal{B}_k^+ \subseteq \mathcal{T}_{q,k}$ and Lemma D.9. Assume it holds at round t. For each message vertex $w \in V(G)$, consider the numerical vector

$$Z_G(\boldsymbol{v}, w) := \left((h_B^G(\boldsymbol{v}, w))_{B \in \mathcal{B}_q^+},\ (h_R^G(\boldsymbol{v}^{i \leftarrow w}))_{i \in [k],\, R \in \mathcal{T}_{q,k}} \right).$$

A monomial in finitely many coordinates of this vector is the rooted count of a root product of $\mathcal{B}_q^+$ templates and child lifts $\iota_i(R)$. Lemmas D.9 and D.10 put that product in $\mathcal{T}_{q,q}$. Summing the monomial over w forgets its last root and therefore gives the count of a template in $\mathcal{T}_{q,k}$. Thus (27) gives equality of every mixed-moment sum of the two message-vector multisets. By Lemma D.11,

$$\{\{Z_G(\boldsymbol{v}, w) : w \in V(G)\}\} = \{\{Z_H(\boldsymbol{u}, z) : z \in V(H)\}\}.$$

For matching vectors, the first block determines equal relation labels $\mathbf{Satp}$ by Lemma D.9. Each child block determines equal round-t child colors by the induction hypothesis. We therefore obtain equality of the complete *joint* replacement multisets, with the relation label and all k child colors associated with the same message vertex. The old colors are equal by the induction hypothesis as well. Applying the shared refinement hash proves equality at round $t+1$. □

Proposition D.13. *For every $k \geq 2$, with $q = k+1$ and any graph operator $T \in \{A, L, \widetilde{L}\}$,*

$$\textit{Spec } k\textit{-FWL} \preceq q\textit{-LWL}. \tag{28}$$

Proof. Suppose $G \equiv_{q\text{-LWL}} H$. Theorem D.7 gives equal homomorphism counts from every graph in $\mathfrak{S}_q$.

First, these counts determine a common spectrum of T_G, T_H. Fix the degree-interpolating polynomials for the pair G, H as above. For each integer $a \geq 1$, expand

$$\operatorname{tr}(T_G^a) = \sum_{v_0,\dots,v_{a-1} \in V(G)} \prod_{j=0}^{a-1} (T_G)_{v_j v_{j+1}}, \qquad v_a = v_0,$$

using (18). After collapsing diagonal steps, a term with $b \geq 3$ edge steps is a homomorphism sum from C_b with polynomial degree weights at its vertices. For $b = 2$, the two parallel traversals give $A_G(u,v)^2 = A_G(u,v)$, so the template is K_2. A term with $b = 1$ vanishes because the graph is loopless, and $b = 0$ gives a degree-weighted K_1. Polynomial degree weights again correspond to pendant decorations, with common coefficients for G, H.

Each resulting decorated graph belongs to $\mathfrak{S}_q$. The graphs K_1, K_2 have immediate one-bag decompositions after multiset padding. A cycle has a width-two chain that keeps one cycle vertex fixed while advancing the other entry around the cycle; adding copies of the fixed vertex to every bag gives width q. Every vertex of these base graphs occurs in an even bag, so (22) adds all pendant decorations. Their homomorphism counts consequently agree in G, H, proving

$$\operatorname{tr}(T_G^a) = \operatorname{tr}(T_H^a) \qquad (a \geq 1).$$

Also, the count of K_1 determines the common number n of vertices. Newton identities applied to the first n traces recover the characteristic polynomial. Thus T_G, T_H have the same spectrum, including multiplicities.

Now associate with each $\boldsymbol{v} \in V(G)^k$ its rooted-count profile

$$Y_G(\boldsymbol{v}) := (h_R^G(\boldsymbol{v}))_{R \in \mathcal{T}_{q,k}}.$$

A monomial in finitely many coordinates is a root-product count. Summing it over all $\boldsymbol{v}$ forgets all root labels. By Lemma D.10, the resulting unrooted template belongs to $\mathfrak{S}_q$. Its homomorphism count agrees in G, H, so all mixed-moment sums of the two profile multisets agree. Lemma D.11 therefore yields

$$\{\{Y_G(\boldsymbol{v}) : \boldsymbol{v} \in V(G)^k\}\} = \{\{Y_H(\boldsymbol{u}) : \boldsymbol{u} \in V(H)^k\}\}.$$

Using the common spectrum, Lemma D.12 converts each profile into its Spec k-FWL color at every finite round. The color histograms consequently agree at every round and at stabilization. Thus $G \equiv_{\text{Spec } k\text{-FWL}} H$, as required. □

D.2.4 Strict separation between k-LWL and SI k-OWL

We next compare k-LWL with the pairwise spectral-invariant refinement SI k-OWL. Throughout this subsection, $k \geq 2$ and the spectral operator is the adjacency matrix. We use the strong canonical family $\mathfrak{S}_k$ and its connected subfamily $\mathcal{S}_k$ from Definition D.6.

Proposition D.14. *For every $k \geq 2$,*

$$k\textit{-LWL} \prec \textit{SI } k\textit{-OWL}.$$

The strict separation can be witnessed by connected graphs.

The non-strict comparison follows by recovering adjacency from the pairwise spectral labels. For strictness, we construct a graph whose homomorphism count is determined by one round of SI k-OWL, but is not determined by k-LWL. We first prepare these two ingredients.

Recovering walk counts from spectral labels. Write

$$\mathcal{P}^A(G,u,v) = \{(\lambda_a^G, P_a^G(u,v)) : a \in [m_G]\},$$

where $\lambda_1^G, \dots, \lambda_{m_G}^G$ are the distinct adjacency eigenvalues and P_a^G is the projector onto the corresponding full eigenspace. The spectral decomposition gives, for every integer $r \geq 0$,

$$(A_G^r)(u,v) = \sum_{a=1}^{m_G} (\lambda_a^G)^r P_a^G(u,v). \tag{29}$$

For $r = 0$, the right-hand side means $\sum_a P_a^G(u,v)$, so in particular

$$\sum_a P_a^G(u,v) = \mathbf{1}[u=v], \qquad \sum_a \lambda_a^G P_a^G(u,v) = A_G(u,v). \tag{30}$$

Thus a pair label determines equality, adjacency, and the number of walks of every length between its two vertices.

The separating template. Let $C = \{c_0, \dots, c_k\}$ be a set of $k+1$ vertices. Define F_k by making C a clique and, for each unordered pair $\{a,b\} \subseteq C$, adding one fresh vertex x_{ab} adjacent precisely to a and b. In particular, the central edge ab is retained. Hence each central edge belongs to a triangle abx_{ab}, and different added vertices are distinct and nonadjacent. The graph F_k is connected; every central vertex has degree $2k$, and every added vertex has degree two.

Lemma D.15. *The graph F_k does not belong to $\mathcal{S}_k$.*

Proof. Suppose that F_k has a strong canonical decomposition (T, β) of width k. For each graph vertex v, the set of tree nodes whose bags contain v is a connected subtree of T. Whenever two graph vertices are adjacent, their occurrence subtrees intersect, because some bag must cover their edge. Therefore, the occurrence subtrees of the vertices in any clique pairwise intersect. The Helly property for subtrees of a tree implies that every clique is contained in a single bag.

Apply this observation to the central clique C. There is a node t_C whose bag contains all $k+1$ distinct central vertices. Even-depth bags have cardinality k, whereas odd-depth bags have cardinality $k+1$. Consequently, t_C has odd depth and $\beta(t_C) = C$. Moreover, it is the only node whose bag contains all of C: if two such nodes existed, the path between them would contain an even-depth node, and running intersection would force its cardinality-k bag to contain all $k+1$ vertices of C.

Since the root has even depth, t_C has one parent. It has at most one child by the strong canonical condition, so it has either one or two neighbors. Each neighboring bag is a cardinality-k multiset-subbag of C. As C has no repetitions, every such bag is obtained by omitting exactly one central vertex.

We can therefore choose two distinct vertices $a, b \in C$ such that no neighbor of t_C contains both. If there are two neighbors omitting different vertices, choose those two omitted vertices. If the omitted vertex is the same for all neighbors, choose it and any other central vertex.

The three vertices a, b, x_{ab} form a clique, so some bag contains all three. This bag is different from $\beta(t_C)$, which contains only central vertices. Both endpoint bags of the path from t_C to this triangle bag contain a and b. Running intersection therefore forces every bag on that path, including the first neighbor of t_C, to contain both a and b. This contradicts the choice of a, b. □

Lemma D.16. *The graph-level color multiset after one round of SI k-OWL determines $\hom(F_k, G)$ for every finite simple graph G.*

Proof. Define the pairwise quantity

$$\phi_G(u,v) := A_G(u,v)(A_G^2)(u,v). \tag{31}$$

By (29), $\mathcal{P}^A(G,u,v)$ determines $\phi_G(u,v)$.

First fix the images $z_0, \dots, z_k$ of the central vertices of F_k. For every central edge $c_a c_b$, the factor $A_G(z_a, z_b)$ enforces preservation of that edge, while $(A_G^2)(z_a, z_b)$ counts the possible images of

$x_{c_a c_b}$. These choices are independent: added vertices have no other incident edges, and homomorphisms are allowed to identify nonadjacent vertices. It follows that

$$\hom(F_k, G) = \sum_{z_0,\dots,z_k \in V(G)} \prod_{0 \le a < b \le k} \phi_G(z_a, z_b). \tag{32}$$

This counts homomorphisms of the fixed template; no division by an automorphism factor is needed.

We now show how one SI round recovers this sum. Let $\chi^{(t)}_{\mathrm{SI},G}(\boldsymbol{u})$ be the round-t SI color of $\boldsymbol{u} = (u_1, \dots, u_k)$. Its initial color contains the pair labels $\mathcal{P}^A(G, u_a, u_b)$ for $a < b$, and hence determines

$$W_G(\boldsymbol{u}) := \prod_{1 \le a < b \le k} \phi_G(u_a, u_b).$$

Fix any coordinate $i \in [k]$. The ith first-round replacement multiset is

$$\left\{\!\!\left\{ \left(\chi^{(0)}_{\mathrm{SI},G}(\boldsymbol{u}^{i \leftarrow v}), \mathcal{P}^A(G, u_i, v) \right) : v \in V(G) \right\}\!\!\right\}.$$

Within one element indexed by v, the explicit pair label determines $\phi_G(u_i, v)$. The child color determines $\phi_G(v, u_j)$ for every $j \neq i$, because the child tuple contains v in coordinate i and u_j in coordinate j. Thus that same multiset element determines the product

$$\phi_G(u_i, v) \prod_{j \neq i} \phi_G(v, u_j).$$

Applying this decoder to every multiset element and adding the results recovers

$$S_{G,i}(\boldsymbol{u}) := \sum_{v \in V(G)} \phi_G(u_i, v) \prod_{j \neq i} \phi_G(v, u_j).$$

The round-one color also retains the initial color, so it determines $W_G(\boldsymbol{u}) S_{G,i}(\boldsymbol{u})$. Summing this quantity over the graph-level multiset of all ordered k-tuple colors gives

$$\sum_{\boldsymbol{u} \in V(G)^k} W_G(\boldsymbol{u}) S_{G,i}(\boldsymbol{u}) = \hom(F_k, G), \tag{33}$$

by (32), with v providing the remaining central image. All decoders used above depend only on the corresponding colors, so equality of round-one graph representations forces equality of this homomorphism count. □

Proof of Proposition D.14. We first show k-LWL $\preceq$ SI k-OWL by induction on the refinement round. Write $\chi^{(t)}_{L,G}$ for the k-LWL colors. At round zero, (30) shows that the initial SI color determines all equality and adjacency relations between the tuple coordinates, and therefore determines $\chi^{(0)}_{L,G}(\boldsymbol{u}) = \mathbf{atp}_G(\boldsymbol{u})$.

Suppose that $\chi^{(t)}_{\mathrm{SI},G}$ determines $\chi^{(t)}_{L,G}$ by a graph-independent decoder. Consider the ith SI replacement multiset at $\boldsymbol{u}$. In each element

$$\left(\chi^{(t)}_{\mathrm{SI},G}(\boldsymbol{u}^{i \leftarrow v}), \mathcal{P}^A(G, u_i, v) \right),$$

the pair label determines whether $v \in \mathcal{N}_G(u_i)$ by (30). Retain exactly those elements for which this holds, and apply the induction decoder to the child colors. The resulting multiset is precisely

$$\left\{\!\!\left\{ \chi^{(t)}_{L,G}(\boldsymbol{u}^{i \leftarrow v}) : v \in \mathcal{N}_G(u_i) \right\}\!\!\right\}.$$

Doing this separately for every coordinate, and decoding the retained old color, recovers the entire k-LWL update. Hence $\chi^{(t+1)}_{\mathrm{SI},G}$ determines $\chi^{(t+1)}_{L,G}$. Passing to graph-level color multisets proves the non-strict comparison.

For strictness, Lemma D.15 gives $F_k \notin \mathcal{S}_k$. The Fürer-witness statement in Theorem D.7, which is the specialization of Zhang et al. (2024a, Theorem E.6) to their model $L(k)$, therefore provides the untwisted and single-edge-twisted lifts

$$G_k := G(F_k), \qquad H_k := H(F_k)$$

satisfying

$$G_k \equiv_{k\text{-LWL}} H_k, \qquad \hom(F_k, G_k) \neq \hom(F_k, H_k). \tag{34}$$

The construction of Zhang et al. (2024a, Definitions C.44–C.45) copies each base vertex label to its fiber. Taking all base labels equal therefore gives unlabeled witnesses, as required here. By Lemma D.16, SI k-OWL distinguishes these witnesses after one refinement round. Combined with the non-strict comparison, this proves the claimed strict inequality.

The two witnesses are connected by Lemma D.23. □

D.2.5 SI k-OWL IS STRICTLY BOUNDED BY k-FWL

Fix $k \geq 2$ and $T = A$. We first prove that k-FWL simulates SI k-OWL. For strictness, we use path-expanded homomorphism patterns and a parity argument to construct connected Fürer witnesses. The strong families $\mathfrak{S}_k$ and $\mathcal{S}_k$ are those of Definition D.6.

Lemma D.17. *For every $k \geq 2$,*

$$\textit{SI } k\textit{-OWL} \preceq k\textit{-FWL}. \tag{35}$$

Proof. Fix two nonempty graphs G, H that are equivalent under k-FWL. Run the refinement with a common injective color encoding on their two tuple domains, and write e_∞ for the resulting stable colors. Stability means that a tuple's color determines the multiset of joint replacement-color vectors

$$J^G_\infty(\boldsymbol{u}) = \left\{\!\!\left\{ \left(e^G_\infty(\boldsymbol{u}^{1\leftarrow w}), \ldots, e^G_\infty(\boldsymbol{u}^{k\leftarrow w})\right) : w \in V(G) \right\}\!\!\right\}.$$

The corresponding decoder is common to G and H.

Step 1: recover pairwise walk counts. We show inductively that, for all $r \geq 0$ and $i, j \in [k]$, $e^G_\infty(\boldsymbol{u})$ determines

$$(A^r_G)_{u_i u_j}. \tag{36}$$

The cases $r = 0, 1$ are equality and adjacency, respectively, and are part of the initial atomic type retained by every stable color. For $i \neq j$, use

$$(A^{r+1}_G)_{u_i u_j} = \sum_{w \in V(G)} (A^r_G)_{u_i w} A_G(w, u_j). \tag{37}$$

Within a single element of $J^G_\infty(\boldsymbol{u})$, its jth component is the color of a tuple containing u_i, w and hence determines the first factor by induction. Its ith component contains w, u_j and determines the second factor. Both factors refer to the same replacement vertex w, so their product can be decoded elementwise and then summed over the joint multiset. For $i = j$, choose $h \neq i$. The hth child tuple contains both u_i and w, and its color determines both factors in the corresponding diagonal recurrence. This proves the induction, including repeated coordinates of $\boldsymbol{u}$.

Step 2: recover a common spectrum and spectral pair labels. The graph-level histogram determines the graph order n from its cardinality n^k. Moreover,

$$\sum_{\boldsymbol{u} \in V(G)^k} (A^r_G)_{u_1 u_1} = n^{k-1} \operatorname{tr}(A^r_G). \tag{38}$$

Thus G and H have equal adjacency power traces, and hence the same eigenvalues with multiplicities, for example by Newton's identities applied to the first n traces. Let Λ be their common set of distinct eigenvalues. The projector onto the λ-eigenspace is the polynomial

$$P_\lambda = \prod_{\mu \in \Lambda \setminus \{\lambda\}} \frac{A - \mu I}{\lambda - \mu}. \tag{39}$$

Its coefficients are common to the two graphs. Together with (36), this recovers $\mathcal{P}^A(G, u_i, u_j)$ for every coordinate pair and therefore the initial SI k-OWL color.

Step 3: simulate each SI update. Assume that stable k-FWL color, together with the common spectrum, determines the round-t SI k-OWL color. In an element of $J^G_\infty(\boldsymbol{u})$ indexed by w, its ith

component therefore determines $\chi^{(t)}_{\mathrm{SI},G}(\boldsymbol{u}^{i\leftarrow w})$. Choose $h \neq i$. Its hth component contains u_i, w and, by Steps 1–2, determines $\mathcal{P}^A(G, u_i, w)$. Elementwise decoding consequently recovers

$$M^{(t)}_{\mathrm{SI},G,i}(\boldsymbol{u}) = \left\{\!\!\left\{ \left(\chi^{(t)}_{\mathrm{SI},G}(\boldsymbol{u}^{i\leftarrow w}), \mathcal{P}^A(G, u_i, w)\right) : w \in V(G) \right\}\!\!\right\}$$

where $M^{(t)}_{\mathrm{SI},G,i}$ denotes the ith SI message multiset. Together with the old SI color, these are precisely the inputs to the next SI update. This proves the induction. Applying the same decoder to the equal stable k-FWL histograms gives equal SI histograms at every finite round, and in particular at stabilization. Empty graphs, if allowed, are already handled by the histogram cardinality. □

Definition D.18. Let $k \geq 2$. A skeleton is a finite multigraph S, possibly with loops, such that its underlying loopless simple graph admits a strong canonical width-k decomposition. A path expansion of S replaces each non-loop edge occurrence ab by a positive-length a–b path, and each loop occurrence at a by a closed path based at a of length at least two. Every internal vertex is fresh, and internal vertices of different paths are disjoint. A closed path of length two is interpreted as a pendant edge: its two traversals impose the same adjacency condition on a simple target graph. Repeated edges in the result are likewise suppressed. In particular, we retain at most one length-one path between any pair of distinct skeleton vertices, so path multiplicities below always refer to the resulting simple graph. Isolated skeleton vertices are retained. Denote by $\mathcal{Q}_k$ the resulting family of finite simple graphs, allowing disconnected graphs.

A k-rooted expanded template consists of such an expansion together with an ordered list of k skeleton vertices, with repetitions allowed, and a strong canonical decomposition of the skeleton whose even root bag is precisely the multiset of these distinguished vertices. Write $\mathcal{Q}^{\mathrm{rt}}_k$ for this rooted family.

The skeleton decomposition is as in Definition D.6; loops impose no additional edge-coverage requirement.

For a rooted template $T = (F; z_1, \ldots, z_k)$, let

$$h^G_T(\boldsymbol{u}) := \#\{f : F \to G \text{ a homomorphism} : f(z_j) = u_j \text{ for every } j \in [k]\}.$$

If two distinguished roots coincide but their prescribed images do not, this count is zero.

Lemma D.19. *For every $k \geq 2$ and all finite simple graphs G, H,*

$$\hom(F, G) = \hom(F, H) \quad (F \in \mathcal{Q}_k) \quad \Longrightarrow \quad G \equiv_{\textit{SI } k\textit{-OWL}} H.$$

Proof. We first record the operations on rooted templates needed by the refinement. The constant function 1 is represented by a template consisting of k distinct isolated roots, with a single even bag. This also supplies the empty root product used below. Given several members of $\mathcal{Q}^{\mathrm{rt}}_k$, take disjoint copies and identify their equally indexed roots. If a factor has repeated roots, these identifications may also identify different roots in another factor. The resulting root product either has identically zero rooted count or belongs to $\mathcal{Q}^{\mathrm{rt}}_k$. In either case, its rooted count is the product of the factor counts. To see that the decomposition is preserved, first apply the root identifications to every multiset bag. For every identified vertex, its new occurrence set is a union of subtrees containing the old root node, hence is connected. Bag cardinalities and multiset inclusions are unchanged. Then identify the decomposition roots of the factors. Only even nodes are joined, so the one-child restriction on odd nodes is preserved. The path interiors remain fresh and disjoint. An identified pair of path endpoints yields a closed path; a length-one closed path makes the entire rooted count identically zero on simple targets, so that zero template term is discarded. All other closed paths are allowed by Definition D.18.

The second operation corresponds to an ith-coordinate message. Start with a template T whose roots are named $(z_1, \ldots, z_{i-1}, w, z_{i+1}, \ldots, z_k)$, introduce the missing old root z_i as a fresh vertex, and add any finite number of paths between z_i and w, with fresh internal vertices. Forget the distinguished label of w, retaining the ordered old roots $(z_1, \ldots, z_k)$. If the added path lengths are $r_1, \ldots, r_d \geq 1$, the new rooted count is

$$\sum_{w \in V(G)} h^G_T(\boldsymbol{u}^{i\leftarrow w}) \prod_{j=1}^{d} (A^{r_j}_G)_{u_i w}. \tag{40}$$

The empty path product is permitted. A factor with length zero means identifying z_i with w, since $(A_G^0)_{u_i w} = \mathbf{1}[u_i = w]$.

This message construction also preserves $\mathcal{Q}_k^{\mathrm{rt}}$. Set $R = \{\{z_1, \dots, z_k\}\}$ and prepend the chain

$$R_{\mathrm{even}} \longrightarrow (R \uplus \{\{w\}\})_{\mathrm{odd}} \longrightarrow ((R \setminus \{\{z_i\}\}) \uplus \{\{w\}\})_{\mathrm{even}}$$

to the old decomposition, identifying the last node with its old root. The added skeleton edges $z_i w$ are covered by the middle bag. The new root occurs only in the first two bags, so running intersection holds. Repeated child roots or length-zero factors amount to identifying vertices all present in the middle bag; their occurrence subtrees therefore have a common node, and the same quotient argument applies. This proves the claim, including the case that w coincides with a retained root or with z_i.

Now assume equality of the unrooted counts in the lemma. The family $\mathcal{Q}_k$ contains a single isolated vertex, a single edge, and every cycle of length at least three. For example, cycles are the expansions of a loop on a one-vertex skeleton, whose bags can be padded by repetitions. Consequently G and H have the same order; if it is zero, the conclusion is immediate. Otherwise they have the same traces $\mathrm{tr}(A^r)$ for every $r \geq 1$, where $r = 1$ is zero and $r = 2$ is the homomorphism count of a single edge. Hence they have a common adjacency spectrum. If Λ is its set of distinct eigenvalues, then

$$P_\lambda = \prod_{\mu \in \Lambda \setminus \{\lambda\}} \frac{A - \mu I}{\lambda - \mu} \qquad (\lambda \in \Lambda).$$

Thus every pair-label coordinate is a common linear combination of the walk functions $(A^r)_{uv}$ for $0 \leq r < |\Lambda|$.

We claim that, on the finite domain $V(G)^k \sqcup V(H)^k$, the indicator of every round-t SI k-OWL color is a finite real linear combination of the functions h_T, with $T \in \mathcal{Q}_k^{\mathrm{rt}}$. All coefficients are chosen jointly for the two graphs. For $t = 0$, each positive-order walk function between two root positions is represented by a path template, with the other roots isolated. The zeroth-order function is represented by identifying the two roots. Their skeletons have a single even root bag. The common projector formula and closure under root products therefore represent every polynomial in the initial color coordinates. On a finite set of possible coordinate vectors, the indicator of each vector is a polynomial: for a target vector a, select for every other vector b a coordinate $j(a, b)$ where they differ and use

$$\prod_{b \neq a} \frac{x_{j(a,b)} - b_{j(a,b)}}{a_{j(a,b)} - b_{j(a,b)}}.$$

This proves the base case.

For the inductive step, an SI k-OWL update is determined by the old color and the multiplicities

$$N_{i,c,p}^G(\boldsymbol{u}) := \#\{w \in V(G) : \chi_{\mathrm{SI},G}^{(t)}(\boldsymbol{u}^{i \leftarrow w}) = c, \mathcal{P}^A(G, u_i, w) = p\},$$

for every coordinate i, old color c, and pair label p occurring in either graph. By the induction hypothesis, the first indicator in this count is a finite linear combination of child rooted counts. By the same finite-set interpolation argument, the second indicator is a polynomial in $(A^r)_{u_i w}$, with coefficients common to G, H. Expanding the product and summing over w gives exactly the message constructions in (40), including the permitted length-zero identifications. Thus every $N_{i,c,p}$ is a linear combination of admissible rooted counts. There are only finitely many attained vectors consisting of the old-color indicators and these multiplicities. Polynomial interpolation on those vectors, followed by root-product closure, represents every round-$(t+1)$ color indicator. This completes the induction.

Finally, summing a rooted template count over all ordered tuples forgets its distinguished roots:

$$\sum_{\boldsymbol{u} \in V(G)^k} h_T^G(\boldsymbol{u}) = \hom(F_T, G).$$

This identity remains valid with repeated roots because each homomorphism prescribes exactly one compatible tuple. The underlying graph F_T belongs to $\mathcal{Q}_k$. Our hypothesis therefore gives equal sums of every color indicator, and hence equal SI color histograms at every finite round. Applying this to a round at which the joint refinement on the two finite domains has stabilized proves $G \equiv_{\text{SI } k\text{-OWL}} H$. □

The separating base graphs. We construct a strong-decomposition obstruction B_k and a graph O_k of the same treewidth whose Fürer lifts will separate SI from k-FWL. Write $\delta(F) := \min_{v \in V(F)} \deg_F(v)$ for the minimum degree of a nonempty graph. The added leaves ensure that all vertices of the core have degree at least three; this is the condition needed in the parity argument below.

Lemma D.20. *For every $k \geq 2$, there are connected graphs $B_k \subseteq O_k$ such that*

$$B_k \notin \mathcal{S}_k, \qquad \mathrm{tw}(O_k) = k, \qquad \delta(B_k) = k. \tag{41}$$

The graph O_k is obtained from B_k by attaching at most one pendant leaf to each vertex. Every vertex of B_k has degree at least three in O_k, and all degrees of O_k are odd.

Proof. Let $C = \{c_0, \ldots, c_k\}$ induce K_{k+1}. Add independent vertices $x_0, \ldots, x_k$ with

$$N_{B_k}(x_i) = C \setminus \{c_i\}.$$

This defines B_k. Attach a fresh pendant leaf to each c_i; when k is even, also attach a fresh pendant leaf to each x_i. The resulting graph is O_k.

The treewidth is k. The central clique implies $\mathrm{tw}(O_k) \geq k$. For the reverse inequality, take the bag C and attach the $k+1$ bags

$$D_i = (C \setminus \{c_i\}) \cup \{x_i\}$$

as its children. These bags cover all edges of B_k and satisfy running intersection. Add a two-vertex bag for each pendant edge, attached to a bag containing its core endpoint. The maximum bag size is $k+1$, so $\mathrm{tw}(O_k) = k$.

The core has no strong canonical width-k decomposition. Suppose such a decomposition of B_k existed. The occurrence subtrees of vertices in a clique pairwise intersect, because every edge is covered. The Helly property for subtrees of a tree implies that the whole clique lies in one bag. Since C and each D_i are $(k+1)$-cliques, they occur in odd bags t_C, t_i whose supports are exactly C, D_i, respectively; all their multiplicities are one. The bag t_C is unique: two occurrences of all of C would force an intervening even bag of cardinality k to contain its $k+1$ distinct vertices, contradicting running intersection.

For every i, all bags on the path from t_C to t_i contain $C \setminus \{c_i\}$, by running intersection for those k vertices. Its first neighbor s_i after t_C is an even bag of cardinality k, so

$$\beta(s_i) = C \setminus \{c_i\}.$$

The $k+1$ different supports force $k+1$ different neighbors of t_C. But the root is even, so an odd node has one parent and at most one child; its degree is at most two. Since $k+1 \geq 3$, this is impossible. Therefore $B_k \notin \mathcal{S}_k$.

Finally, the degrees in B_k are $\deg(c_i) = 2k$ and $\deg(x_i) = k$. The added leaves make them $2k+1$ and, respectively, k when k is odd or $k+1$ when k is even. All these degrees are odd and at least three; each new leaf has degree one. Also $\delta(B_k) = k$, completing the proof. □

Parity certificates for Fürer lifts. We use the following consequences of Roberson's oddomorphism theorem (Roberson, 2022, Definition 3.9 and Theorem 3.13). For a connected graph F with at least one edge, let F^0 and F^1 denote its untwisted and single-edge-twisted Fürer lifts, respectively. A homomorphism $f : R \to F$ is an *oddomorphism* if there is a function $p : V(R) \to \mathbb{F}_2$ such that

$$\big|N_R(x) \cap f^{-1}(y)\big| \equiv p(x) \pmod 2 \qquad \big(x \in V(R),\ y \in N_F(f(x))\big), \tag{42}$$

$$\sum_{x \in f^{-1}(v)} p(x) = 1 \qquad \big(v \in V(F)\big). \tag{43}$$

Here $\mathbb{F}_2$ is the field with two elements; all sums involving p are taken in this field. A vertex with $p(x) = 1$ is called odd; otherwise it is called even. In particular, an odd vertex has a neighbor in every neighbor fiber of its image.

Theorem D.21. *Let F be connected with at least one edge and let T be any finite simple graph.*

(i) *If $\hom(T, F^0) \neq \hom(T, F^1)$, some subgraph $R \subseteq T$ admits an oddomorphism to F.*

(ii) *If T admits an oddomorphism to F, then* $\hom(T, F^0) \neq \hom(T, F^1)$.

Proof. Both assertions follow from the cited Theorem 3.13. That theorem uses a lift with odd parity at one base vertex in place of an edge twist. Toggling the bit of one incident edge at that vertex is an isomorphism to the even-parity, single-edge-twisted lift used here. □

Lemma D.22. *Let O be obtained from a nonempty simple graph B by attaching pendant leaves. Assume that every vertex of B has degree at least three in O, and $\delta(B) \geq 2$. Let $T \in \mathcal{Q}_k$ have underlying strong skeleton S. If a subgraph of T admits an oddomorphism to O, then some graph $S' \in \mathfrak{S}_k$ admits an oddomorphism to B.*

Proof. Write $W = V(S) \subseteq V(T)$ for the skeleton vertices. Let $R \subseteq T$ and $f : R \to O$ be an oddomorphism, with parity labels p as in (42)–(43). We will remove the path interiors while preserving the parity conditions on the core B.

Step 1: restrict the certificate to the core. Every non-skeleton vertex $x \in V(R) \setminus W$ has degree at most two in R. If $f(x) \in V(B)$ and $p(x) = 1$, then (42) requires a neighbor of x in each of at least three distinct neighbor fibers of $f(x)$ in O. This is impossible. Consequently,

$$p(x) = 0 \qquad \text{whenever } x \notin W \text{ and } f(x) \in V(B). \tag{44}$$

Let $R_B = R[f^{-1}(V(B))]$. The restricted map $f_B : R_B \to B$ is still an oddomorphism with the same labels p: for each retained vertex, the conditions indexed by neighbors in B are a subset of its original conditions, and the fibers over vertices of B have not changed. A retained non-skeleton vertex therefore has degree zero or two. In the latter case, its two neighbors have the same image, since every neighbor fiber has even cardinality. Notice that the parity labels are inherited from O before this restriction; this also covers the case in which a vertex of B has degree two.

Delete isolated non-skeleton vertices and components with no skeleton vertex. All their parity labels are zero, so this does not change (43). Every remaining non-skeleton vertex has degree two. Thus a surviving replacement path cannot stop at a path interior: it is an entire path between skeleton vertices, or a cycle based at one skeleton vertex. Along each such path, images alternate between two fixed adjacent vertices of B. An even-length path has same-image endpoints; an odd-length path has adjacent-image endpoints; and a based cycle necessarily has even length. The pendant edges representing length-two closed walks cannot survive, because their non-skeleton endpoint would have degree one in R_B.

Step 2: compress the surviving paths. In S, contract every edge that has a surviving even-length replacement path. Its endpoints have the same image under f_B, so f_B is constant on every contraction class. A skeleton edge cannot have surviving paths of both parities, since its endpoint images cannot be both equal and adjacent in a simple graph. For each pair of contraction classes, retain one edge if the total number of surviving odd-length paths between them is odd, and no edge otherwise. Delete all other skeleton edges. Based cycles are discarded. Call the resulting simple graph S'.

Skeleton vertices absent from R_B are retained in S' as isolated vertices. Give them parity zero and map them to an arbitrary fixed vertex of B. Thus no closure under vertex deletion is required. For a contraction class C of retained skeleton vertices, set

$$f'(C) = f_B(x) \quad (x \in C), \qquad p'(C) = \sum_{x \in C} p(x).$$

The first definition is well-defined, and $f' : S' \to B$ is a homomorphism because every retained edge comes from an odd-length path whose endpoint images are adjacent.

We verify the parity conditions. Fix C and $y \in N_B(f'(C))$, and sum (42) over all $x \in C$. Every even-length path within C contributes twice to the same neighbor fiber: its first and last internal vertices have the same image by alternation. The same cancellation applies to every based even cycle. The remaining contributions are precisely the odd-length paths leaving C, with multiplicity reduced modulo two in the construction of S'. Hence

$$\left|N_{S'}(C) \cap (f')^{-1}(y)\right| \equiv p'(C) \pmod 2.$$

Moreover, (44) implies that all odd vertices of R_B are skeleton vertices. Therefore, for every $v \in V(B)$,

$$\sum_{C \in (f')^{-1}(v)} p'(C) = \sum_{x \in f_B^{-1}(v)} p(x) = 1.$$

Thus f' is an oddomorphism.

Step 3: preserve the strong decomposition. The graph S' is obtained from S by edge contractions and edge deletions, retaining isolated vertices. Edge deletion does not affect any bag. To contract an edge ab, replace every occurrence of a or b in every multiset bag by the new quotient vertex, preserving all multiplicities. Its occurrence subtree is the union of the old a- and b-subtrees; these intersect in a bag covering ab. Running intersection, multiset cardinalities, bag inclusions, depths, and the one-child condition at odd nodes are all preserved. Consequently $S' \in \mathfrak{S}_k$, as required. □

Connected witnesses. The following lemma verifies connectedness for this separation and for the preceding k-LWL versus SI k-OWL separation.

Lemma D.23. *For every $k \geq 2$, every twisted Fürer lift of F_k or O_k is connected, including the untwisted lift.*

Proof. For a base graph F and a twist function $\sigma : E(F) \to \mathbb{F}_2$, write the lift states over v as (v, ξ), where $\xi \in \mathbb{F}_2^{E_F(v)}$ has even Hamming weight and $E_F(v)$ is the set of edges incident to v. States (u, ξ) and (v, η) above $e = uv$ are adjacent precisely when

$$\xi(e) + \eta(e) = \sigma(e). \tag{45}$$

First consider F_k. Any two even states over a central vertex c differ by a sequence of two-bit toggles. For a toggle on incident edges e, f, choose a third incident edge $g = cu$; this is possible because $\deg(c) = 2k \geq 4$. The states before and after the toggle agree on g. Since $\deg(u) \geq 2$, the bit required by (45) extends to an even state over u, giving a common lifted neighbor. Thus all states in a central fiber lie in one connected component. Every state over x_{ab} has a lifted neighbor over a, and every central base edge has a lifted edge. The central clique therefore joins all central fibers and all other fibers attach to them, proving connectedness.

For O_k, let $R = V(B_k)$ be the non-leaf core. Every $v \in R$ has at least two core neighbors and at most one leaf neighbor. If two even states over v agree on a core edge vu, they have a common lifted neighbor over u: prescribe the required bit at u and extend it to an even state, using $\deg(u) \geq 2$. Identify an even state over v with the set X of neighbors whose incidence bits are one. If a core-neighbor bit is zero, this observation connects X to the all-zero state. Otherwise X contains every core neighbor. If v has no leaf, their number is odd, contradicting the even parity of X. If v has a leaf ℓ, the number of core neighbors is even and parity forces $\ell \notin X$. Pick distinct core neighbors a, b; then X connects to $\{a, \ell\}$ through the a coordinate, and $\{a, \ell\}$ connects to the empty state through the b coordinate. Thus all states in each core fiber belong to one connected component. A leaf fiber is a singleton. Finally, every base edge has a lifted edge: orient it so that one endpoint is a core vertex, prescribe its required bit, and extend to an even core state. Since O_k is connected, every twisted lift is connected. □

Proposition D.24. *For every $k \geq 2$, SI k-OWL $\prec$ k-FWL.*

Proof. The non-strict containment is Lemma D.17. Let B_k and O_k be the graphs constructed in Lemma D.20. Recall that $B_k \notin \mathcal{S}_k$, $\mathrm{tw}(O_k) = k$, and all degrees of O_k are odd. Each core vertex has degree at least three in O_k, and $\delta(B_k) = k \geq 2$. Let O_k^0, O_k^1 be the untwisted and single-edge-twisted Fürer lifts of O_k.

We first show that

$$\hom(T, O_k^0) = \hom(T, O_k^1) \qquad \text{for every } T \in \mathcal{Q}_k. \tag{46}$$

If this failed, Theorem D.21(i) would provide an oddomorphism from a subgraph of T to O_k. Lemma D.22 would then give a strong template $S' \in \mathfrak{S}_k$ with an oddomorphism to B_k. By Theorem D.21(ii),

$$\hom(S', B_k^0) \neq \hom(S', B_k^1).$$

However, $B_k \notin \mathcal{S}_k$, so Theorem D.7 gives $B_k^0 \equiv_{k\text{-LWL}} B_k^1$ and equality of their homomorphism counts from every template in $\mathfrak{S}_k$. This is a contradiction, proving (46). Lemma D.19 now implies

$$O_k^0 \equiv_{\text{SI } k\text{-OWL}} O_k^1.$$

The identity map of O_k is an oddomorphism, with every vertex odd. Theorem D.21(ii) therefore gives

$$\hom(O_k, O_k^0) \neq \hom(O_k, O_k^1).$$

Since $\mathrm{tw}(O_k) = k$, the treewidth-k homomorphism characterization of k-FWL (Zhang et al., 2024a, Theorems 3.10 and E.9) implies that it distinguishes these two graphs. They are connected by Lemma D.23. This establishes the strict separation. □

Proof of Theorem 4.7. For every $k \geq 2$, Proposition D.2 gives k-FWL $\prec$ Spec k-FWL, and Proposition D.13 gives Spec k-FWL $\preceq$ $(k+1)$-LWL. Both comparisons hold for each $T \in \{A, L, \widetilde{L}\}$, proving part (i). With the adjacency operator used in the SI definition, Propositions D.14 and D.24 give k-LWL $\prec$ SI k-OWL $\prec$ k-FWL, which is part (ii). □

D.3 Proof of Proposition 4.9

Throughout this subsection, the spectral operator is $T = A$. We first prove the equivalence between Spec 1-WL and SI 1-WL by explicitly comparing their refinement colors. We then obtain the strict comparison with 2-LWL from the corresponding published result for spectral invariant GNNs.

For completeness, the order-one initializations used here are

$$\widetilde{\chi}_G^{(0)}(u) = \mathbf{Satp}_G(u), \qquad \chi_{\mathrm{SI},G}^{(0)}(u) = *,$$

where $*$ is a common constant. The second initialization is the empty tuple of pairwise labels in the SI definition when $k = 1$. The first is the unary instance of spectralized atomic-type initialization. In particular, the comparison concerns stable expressiveness and does not assert equality at every matching refinement round.

The spectralized unary and binary atomic types. Let $\mathcal{L}_G$ be the set of distinct eigenvalues of A_G, and write $A_G = \sum_{\lambda \in \mathcal{L}_G} \lambda P_\lambda^G$. For each $\lambda \in \mathcal{L}_G$, define the projected constant signal

$$a_\lambda^G(u) := (P_\lambda^G \mathbf{1})(u) = \sum_{v \in V(G)} P_\lambda^G(u, v).$$

There is only one unary atomic type, whose indicator is the constant signal $\mathbf{1}$. By Theorem 4.4, its spectralized label $\mathbf{Satp}_G(u)$ is therefore determined by the distinct eigenvalues and the values $(a_\lambda^G(u))_{\lambda \in \mathcal{L}_G}$, and determines them in turn.

For binary atomic types, let $J = \mathbf{1}\mathbf{1}^\top$. The identity, adjacency, and distinct-nonadjacency indicator matrices are

$$C_G^{\mathrm{id}} = I, \qquad C_G^{\mathrm{adj}} = A_G, \qquad C_G^{\mathrm{nonadj}} = J - I - A_G.$$

Under the column-vectorization convention, $(P_\lambda^G \otimes P_\mu^G)\,\mathrm{vec}(C) = \mathrm{vec}(P_\mu^G C P_\lambda^G)$. The projector characterization of spectralization therefore shows that $\mathbf{Satp}_G(u, v)$ is equivalently described by the distinct eigenvalues together with all entries

$$(P_\mu^G C_G^\tau P_\lambda^G)(u, v), \qquad \lambda, \mu \in \mathcal{L}_G, \quad \tau \in \{\mathrm{id}, \mathrm{adj}, \mathrm{nonadj}\}.$$

Orthogonality of the eigenspace projectors yields

$$P_\mu^G I P_\lambda^G = \mathbf{1}[\lambda = \mu] P_\lambda^G, \tag{47}$$

$$P_\mu^G A_G P_\lambda^G = \mathbf{1}[\lambda = \mu] \lambda P_\lambda^G. \tag{48}$$

Consequently, the identity channel determines $\mathcal{P}^A(G, u, v)$, and this pair label determines the adjacency channel as well. For the remaining channel, symmetry of the projectors gives

$$(P_\mu^G J P_\lambda^G)(u, v) = a_\mu^G(u) a_\lambda^G(v),$$

and therefore

$$\begin{aligned}&(P_\mu^G(J - I - A_G)P_\lambda^G)(u, v)\\ &\quad = a_\mu^G(u)a_\lambda^G(v) - \mathbf{1}[\lambda = \mu](1 + \lambda)P_\lambda^G(u, v).\end{aligned} \tag{49}$$

Thus the complete binary spectralized label is determined by

$$\mathcal{P}^A(G, u, v), \qquad (a_\lambda^G(u))_{\lambda\in\mathcal{L}_G}, \qquad (a_\lambda^G(v))_{\lambda\in\mathcal{L}_G}. \tag{50}$$

This is the key point: the additional binary atomic-type channels require only the pairwise spectral label and projected constant signals at the two endpoints.

Proof of Proposition 4.9. We first establish SI 1-WL $\preceq$ Spec 1-WL. At initialization, the SI color is constant and is therefore trivially determined by the spectralized color. Suppose inductively that $\widetilde{\chi}_G^{(t)}$ determines $\chi_{\mathrm{SI},G}^{(t)}$. Each element of the next Spec replacement multiset contains

$$\big(\widetilde{\chi}_G^{(t)}(v), \mathbf{Satp}_G(u, v)\big).$$

The first entry determines the round-t SI child color by the induction hypothesis; the second determines $\mathcal{P}^A(G, u, v)$ by (47). Applying these decoders to every multiset element recovers the entire SI update. Hence $\widetilde{\chi}_G^{(t)}$ determines $\chi_{\mathrm{SI},G}^{(t)}$ at every round.

For the converse, we prove the finite-round statement

$$\chi_{\mathrm{SI},G}^{(t+1)}(u) \quad \text{determines} \quad \widetilde{\chi}_G^{(t)}(u) \qquad (t \geq 0). \tag{51}$$

Because SI starts from a constant, its first-round color determines

$$\left\{\left\{\mathcal{P}^A(G, u, v) : v \in V(G)\right\}\right\}.$$

Every pair label contains the distinct adjacency eigenvalues, and summing its λ-projector entries over this multiset recovers $a_\lambda^G(u)$. Thus the first-round SI color determines all the data needed for $\mathbf{Satp}_G(u)$, proving (51) for $t = 0$. These data remain available at every subsequent round because the update retains its old color.

Suppose (51) holds at round t. The round-$(t + 2)$ SI color at u determines its old color and the multiset

$$\left\{\left\{\big(\chi_{\mathrm{SI},G}^{(t+1)}(v), \mathcal{P}^A(G, u, v)\big) : v \in V(G)\right\}\right\}.$$

For each element, the child color determines $\widetilde{\chi}_G^{(t)}(v)$ by induction and all values $a_\mu^G(v)$ by the first-round recovery above. The old color at u similarly determines $\widetilde{\chi}_G^{(t)}(u)$ and all $a_\lambda^G(u)$. Together with the pair label, (50) therefore reconstructs $\mathbf{Satp}_G(u, v)$ within each same-vertex message. Consequently, the complete round-$(t + 1)$ Spec update is determined. This proves (51) for every t.

All these decoders are shared across graphs. Applying them to graph-level color multisets in both directions proves

$$\text{Spec 1-WL} \equiv \text{SI 1-WL}.$$

It remains to establish the strict upper bound. The SI 1-WL update and constant initialization used here agree exactly with the spectral invariant refinement of Gai et al. (2025, Section 2.1). Their Corollary 3.8, with the number of iterations set to ∞, proves that this stable refinement is strictly less expressive than Local 2-GNN. The latter model is precisely 2-LWL as defined in (10): its two replacement multisets independently range over the neighbors of the corresponding coordinate (Zhang et al., 2024a, Appendix E.1, equation (27)). Hence

$$\text{SI 1-WL} \prec \text{2-LWL}.$$

Combining this published strict comparison with the equivalence proved above establishes Proposition 4.9. □

D.4 Proof of Theorem 4.10

We first establish the non-strict comparison for an arbitrary real-valued 2-RGI. We then prove strictness for the RGIs in the theorem by a verification on explicit graph pairs. Throughout this subsection, Spectralization uses $T = A$, and both augmented refinements are initialized with $\mathrm{spec}(A)$.

For a connected simple template F with $2 \le |V(F)| \le 5$, fix an ordered pair of distinct roots $\boldsymbol{r} = (r_1, r_2) \in V(F)^2$, and write $F_{\boldsymbol{r}} = (F, \boldsymbol{r})$ for the rooted template. We use the rooted homomorphism and subgraph counts from Appendix C.2, namely,

$$\Phi^{\hom}_{F_{\boldsymbol{r}}}(G,u,v) = \hom\big((F,r_1,r_2),(G,u,v)\big), \qquad \Phi^{\mathrm{sub}}_{F_{\boldsymbol{r}}}(G,u,v) = \frac{\hom_{\mathrm{inj}}\big((F,r_1,r_2),(G,u,v)\big)}{|\operatorname{Aut}(F_{\boldsymbol{r}})|}.$$

Here $\operatorname{Aut}(F_{\boldsymbol{r}})$ fixes each root individually. The root order is retained, so these signals need not be symmetric in (u, v). Ordinary homomorphisms may map the two distinct roots to the same vertex, whereas injective homomorphisms cannot. Subgraph copies are not required to be induced. For each fixed rooted template, using the injective count instead of the subgraph count multiplies the signal and every spectral response by the same positive constant. This does not change either refinement's distinguishing power.

For a connected input graph G, let $L_G = D_G - A_G$ be its combinatorial Laplacian, where D_G is the diagonal degree matrix, and let $L_G^{\dagger}$ be its Moore–Penrose pseudoinverse. For a symmetric matrix M indexed by $V(G)$, define its distance lift by

$$\Delta_G(M)(u,v) := M(u,u) + M(v,v) - 2M(u,v).$$

The three distance signals in the theorem are

$$\begin{aligned}
\Phi_{\mathrm{SPD}}(G,u,v) &:= \mathrm{dist}_G(u,v),\\
\Phi_{\mathrm{DD}}(G,u,v) &:= \sqrt{\Delta_G\big(e^{-2L_G}\big)(u,v)},\\
\Phi_{\mathrm{BD}}(G,u,v) &:= \sqrt{\Delta_G\big((L_G^{\dagger})^2\big)(u,v)}.
\end{aligned}$$

Here $\mathrm{dist}_G(u,v)$ is the length of a shortest path from u to v, and Φ_{DD} is diffusion distance at time $t = 1$. These definitions use L_G to specify the distances; the subsequent Spectralization of each distance signal still uses A_G.

Proof of Theorem 4.10. Fix a real-valued 2-RGI Φ, and write $\Phi_G = (\Phi(G,u,v))_{u,v\in V(G)}$ for its signal matrix. The constant filter $g \equiv 1$ acts as the identity, so $\mathrm{Spec}^A_\Phi(G,u,v)$ determines $\Phi(G,u,v)$. Let $\chi_G^{\mathrm{raw},(t)}$ and $\chi_G^{\mathrm{spec},(t)}$ denote the node colors after t rounds of Aug_Φ 1-WL and $\mathrm{Aug}_{\mathrm{Spec}^A_\Phi}$ 1-WL, respectively. We refer to these refinements as Raw and Spec. We claim that, for all graphs G, H, vertices $u \in V(G)$, $v \in V(H)$, and $t \ge 0$,

$$\chi_G^{\mathrm{spec},(t)}(u) = \chi_H^{\mathrm{spec},(t)}(v) \quad\Longrightarrow\quad \chi_G^{\mathrm{raw},(t)}(u) = \chi_H^{\mathrm{raw},(t)}(v).$$

The claim holds at initialization because both variants use the same graph spectrum. If it holds at round t, each message in the spectralized update determines the corresponding raw message: the previous node color can be decoded by the induction hypothesis, the atomic type is retained, and the raw RGI value is recovered by the constant filter. The same decoding therefore applies to the message multiset and the retained color of the updated vertex. Injectivity of the shared hash proves the claim at round $t + 1$. Applying this decoding to graph-level color multisets yields

$$\mathrm{Aug}_\Phi \text{ 1-WL} \preceq \mathrm{Aug}_{\mathrm{Spec}^A_\Phi} \text{ 1-WL}. \tag{52}$$

To construct the separating graph pairs, we use the following form of the Fürer construction (Gai et al., 2025, Appendix B.5). Let B be a connected simple graph, write $\mathcal{N}_B(x)$ for the neighborhood of $x \in V(B)$, and let $\sigma : E(B) \to \{0,1\}$ assign a twist to each edge. Define $\Gamma^\sigma(B)$ to have vertex set

$$V(\Gamma^\sigma(B)) := \{(x,S) : x \in V(B),\ S \subseteq \mathcal{N}_B(x),\ |S| \equiv 0 \pmod 2\}.$$

Two vertices (x, S) and (y, R) are adjacent exactly when

$$\{x,y\} \in E(B) \quad\text{and}\quad \mathbf{1}[y \in S] + \mathbf{1}[x \in R] \equiv \sigma(\{x,y\}) \pmod 2.$$

Table 4: Base graphs for the separating pairs. A base graph with q vertices has vertex set $\{0, \dots, q-1\}$, and ij abbreviates the undirected edge $\{i, j\}$.

Base	q	Edge set	Twisted edge
B_b	6	$\{01, 02, 04, 05, 12, 23, 25, 34\}$	01
B_c	7	$\{01, 12, 15, 16, 23, 34, 45, 46\}$	01
B_d	5	$\{01, 04, 12, 13, 14, 23, 34\}$	01
B_e	6	$\{01, 02, 03, 04, 14, 23, 35\}$	01
B_f	6	$\{02, 05, 13, 14, 23, 34, 35\}$	02

Thus $\sigma = 0$ gives the untwisted graph. For an edge $e \in E(B)$, let σ_e equal 1 on e and 0 elsewhere; this reverses adjacency between the two vertex fibers over the endpoints of e.

Table 4 specifies five base graphs B_b, B_c, B_d, B_e, B_f and their selected edges e_b, e_c, e_d, e_e, e_f. For $i \in \{b, c, d, e, f\}$, set

$$G_i := \Gamma^0(B_i), \qquad H_i := \Gamma^{\sigma_{e_i}}(B_i).$$

We use two further pairs obtained by replacing each vertex of G_i and H_i by a two-vertex clique, for $i \in \{b, c\}$. More precisely, for any simple graph Q, the graph $Q[K_2]$ has vertex set $V(Q) \times \{0, 1\}$, with distinct (u, a) and (v, b) adjacent if $u = v$ or $\{u, v\} \in E(Q)$.

All seven pairs consist of connected graphs. For each pair, exact comparison of the adjacency characteristic polynomials verifies cospectrality, including eigenvalue multiplicities. In particular, the two graphs have identical initial colors. Consequently, using a common constant as the initial color gives exactly the same equality pattern in every refinement round on these pairs.

We next explain how a finite calculation verifies the two refinements. Fix one of the cospectral pairs (G, H), and write $\lambda_1 < \cdots < \lambda_d$ for their common distinct adjacency eigenvalues. For $Q \in \{G, H\}$, let P_r^Q be the projector associated with λ_r. The projector characterization in Theorem 4.4 identifies the spectralized signal with the common spectrum and the entries of $P_r^Q \Phi_Q P_s^Q$. For $0 \le a, b < d$,

$$A_Q^a \Phi_Q A_Q^b = \sum_{r,s=1}^{d} \lambda_r^a \lambda_s^b P_r^Q \Phi_Q P_s^Q. \tag{53}$$

The matrix $V = (\lambda_r^a)_{0 \le a < d,\, 1 \le r \le d}$ is an invertible Vandermonde matrix. Applying its inverse in both indices shows that the moment entries

$$\big((A_Q^a \Phi_Q A_Q^b)(u, v)\big)_{0 \le a, b < d}$$

and the projector-response entries determine one another, with the same change of coordinates for both graphs. Hence they induce exactly the same pair-label equalities. If Φ_Q commutes with A_Q in both graphs, the responses with $r \neq s$ vanish, and the one-sided moments $\Phi_Q A_Q^b$, $0 \le b < d$, already suffice.

For rooted homomorphism counts, injective counts, and shortest-path distance, the signal matrices and all these moments have integer entries. Hom counts are computed from their defining sums over root-preserving maps. Injective counts can be obtained by Möbius inversion over partitions of the template vertices, discarding quotients with loops; when the two host roots are distinct, quotients merging the template roots contribute zero. The diagonal of an injective-count signal is zero because the template roots are distinct. All resulting labels are therefore computed with exact integer arithmetic. Dividing the injective counts by $|\operatorname{Aut}(F_{\boldsymbol{r}})|$ gives the subgraph counts; equivalently, the fixed scaling can be omitted when comparing colors.

Biharmonic distance is handled exactly as well. Its squared values are rational, since

$$L_Q^\dagger = \left(L_Q + \frac{1}{|V(Q)|}\mathbf{1}\mathbf{1}^\mathsf{T}\right)^{-1} - \frac{1}{|V(Q)|}\mathbf{1}\mathbf{1}^\mathsf{T},$$

where $\mathbf{1}$ is the all-ones vector of the appropriate size. Each distance value can thus be written as a rational multiple of $\sqrt{s}$ for a positive squarefree integer s. The distinct squarefree radicals occurring in the two graphs are linearly independent over $\mathbb{Q}$. Representing the signal and its moments by their

Table 5: Separating pairs and the number of assigned ordered-root template cases in the Hom and Sub columns, respectively. Both graphs in a row have the stated numbers of vertices and edges. Each assigned RGI is used individually. In every comparison, Raw stabilizes without distinguishing the graphs, while Spec distinguishes them after one round. DD is evaluated at time $t = 1$.

Pair	Vertices	Edges	Hom	Sub	Distances
(G_b, H_b)	24	82	0	0	SPD
(G_c, H_c)	21	42	59	56	—
(G_d, H_d)	20	64	30	20	BD, DD
(G_e, H_e)	19	48	26	9	—
(G_f, H_f)	18	38	0	6	—
$(G_b[K_2], H_b[K_2])$	48	352	69	86	—
$(G_c[K_2], H_c[K_2])$	42	189	6	13	—
Total			190	190	

rational coefficients in this common basis gives exact equality tests for the scalar distance and all its responses.

For diffusion distance, we use exact raw labels and a certified separating response. Let $\mathcal{L}$ be the union of the distinct Laplacian eigenvalues of G and H, and set $\ell = |\mathcal{L}|$. For $\nu \in \mathcal{L}$, let E_ν^Q denote the corresponding Laplacian projector, taking $E_\nu^Q = 0$ if ν is absent from Q. Then

$$\Phi_{\mathrm{DD}}(Q, u, v)^2 = \sum_{\nu \in \mathcal{L}} e^{-2\nu} \Delta_Q(E_\nu^Q)(u, v).$$

The eigenvalues and projector entries are algebraic. By the Lindemann–Weierstrass theorem (Baker, 2022), the exponentials $e^{-2\nu}$ for distinct $\nu \in \mathcal{L}$ are linearly independent over the algebraic numbers. Equality of squared distances is therefore equivalent to equality of all coefficients in this expression. A common Vandermonde change of coordinates replaces those coefficients by

$$\big(\Delta_Q(L_Q^j)(u, v)\big)_{0 \le j < \ell},$$

which are integers. Since the nonnegative square root is injective, these vectors also give exact equality labels for the raw distance.

To certify the spectralized separation, we evaluate $\Phi_{\mathrm{DD},Q} A_Q$ by outward-rounded interval arithmetic, where $\Phi_{\mathrm{DD},Q}$ is the matrix of diffusion distances. This is a spectral response of the distance itself, including its square root. We pool the intervals for all entries in both graphs, join two intervals when they overlap, and assign the same label to intervals in the same connected component of this overlap relation. Equal exact values necessarily receive the same label. Thus these labels, combined with the exact raw labels, form a coarsening of the labels supplied by full Spectralization. A distinction obtained with this coarsening certifies a distinction for the full spectralized refinement. For the pair (G_d, H_d), 256-bit interval arithmetic already certifies such a distinction at the first round.

We enumerate connected simple templates on two to five vertices and their ordered pairs of distinct roots, up to root-preserving isomorphism. There are respectively $1, 4, 23, 162$ rooted types at these four sizes, giving 190 types in total. The stored catalogue has 118 representatives modulo reversal of the root pair: 46 admit an automorphism exchanging the roots, and both root orders are checked separately for the remaining 72. Homomorphism and subgraph counts are then verified individually, giving 380 comparisons. Table 5 summarizes their assigned graph pairs and the distance examples.

For each comparison, the raw relation label is $(\mathbf{atp}_Q(u, v), \Phi(Q, u, v))$, and the spectralized relation label replaces $\Phi(Q, u, v)$ by its complete moment profile, or by the certified coarsening described above for DD. The atomic type itself is retained without Spectralization. At each round, the node update uses all vertices of the same graph, exactly as in the definition of Aug_Φ 1-WL, with a common injective color encoding across the two graphs. We compare the full node-color multisets after each update. Because the previous color is retained, the joint partition of the two vertex sets can only refine. If its number of classes no longer increases, the partition is stable; equality of the two color multisets at that stage therefore proves indistinguishability at all subsequent rounds.

The exact verification gives the following result for every assigned template or distance signal: Raw has identical graph-level color multisets, with stability certified at round 2 or 3, whereas Spec has

different multisets at round 1. For example, take a triangle with any two distinct ordered roots. Its homomorphism-count signal is $\Phi_Q(u,v) = A_Q(u,v)(A_Q^2)(u,v)$. On (G_d, H_d), Raw has common color multiplicities $(12, 8)$ at round 1 and $(4, 8, 8)$ at rounds 2 and 3, with the last update confirming stability. Using only Φ_Q and the response $\Phi_Q A_Q$ already yields different node-color multisets at round 1: under a common color encoding, one color occurs eight times in G_d and zero times in H_d. Thus even this restricted response profile certifies the strict separation for the triangle count.

The graph data, rooted-template catalogue, assignments, and verification code are provided in the `Theorem412` directory of the accompanying repository. The assignments specify a witness for each individual RGI; the graph pair may depend on the chosen RGI. For every Φ covered by the theorem, the verified pair (G_Φ, H_Φ) satisfies

$$G_\Phi \equiv_{\mathrm{Aug}_\Phi \text{ 1-WL}} H_\Phi, \qquad G_\Phi \not\equiv_{\mathrm{Aug}_{\mathrm{Spec}_\Phi^A} \text{ 1-WL}} H_\Phi.$$

Together with (52), this proves the strict comparison claimed in Theorem 4.10. □

E Proofs for Section 5

We first prove the finite-bank recovery criterion and give an explicit filter bank that works uniformly when the number of distinct eigenvalues is bounded. We then establish the polynomial realization for arbitrary order k, including the bivariate matrix form used in the main text. Throughout, $k \geq 1$, Φ_k is a real-valued k-RGI, and $\Phi_G = \Phi_k(G, \cdot)$. We use the graph operator T and the joint spectral data $\boldsymbol{\lambda}_{\boldsymbol{r}}^G, P_{\boldsymbol{r}}^G$ from Proposition 4.1. The arguments require neither connected input graphs nor distinct entries of the rooted tuple.

E.1 Proof of Proposition 5.2

Proof of Proposition 5.2. Fix a graph G, a tuple $\boldsymbol{v} \in V(G)^k$, and a filter bank $\mathcal{B} = (g_b)_{b=1}^F$. Define the joint eigenspace-response vector and the observed response vector by

$$\mathbf{c}_{G,\Phi_k}(\boldsymbol{v}) := \left([P_{\boldsymbol{r}}^G \Phi_G](\boldsymbol{v})\right)_{\boldsymbol{r} \in [m_G]^k}, \tag{54}$$

$$\mathbf{y}_{G,\Phi_k}^{\mathcal{B}}(\boldsymbol{v}) := \left([\mathcal{F}_{G,g_b}^{(k)} \Phi_G](\boldsymbol{v})\right)_{b=1}^F. \tag{55}$$

The projector expansion in Proposition 4.1 gives

$$\mathbf{y}_{G,\Phi_k}^{\mathcal{B}}(\boldsymbol{v}) = A_G^{\mathcal{B}} \mathbf{c}_{G,\Phi_k}(\boldsymbol{v}), \qquad A_G^{\mathcal{B}} = \left(g_b(\boldsymbol{\lambda}_{\boldsymbol{r}}^G)\right)_{b \in [F],\, \boldsymbol{r} \in [m_G]^k}. \tag{56}$$

If $A_G^{\mathcal{B}}$ has full column rank, it has the left inverse

$$L_G^{\mathcal{B}} := \left((A_G^{\mathcal{B}})^\top A_G^{\mathcal{B}}\right)^{-1} (A_G^{\mathcal{B}})^\top.$$

Consequently,

$$\mathbf{c}_{G,\Phi_k}(\boldsymbol{v}) = L_G^{\mathcal{B}} \mathbf{y}_{G,\Phi_k}^{\mathcal{B}}(\boldsymbol{v}). \tag{57}$$

For any $g \in C(\mathbb{R}^k)$, let $\mathbf{a}_G(g) = (g(\boldsymbol{\lambda}_{\boldsymbol{r}}^G))_{\boldsymbol{r} \in [m_G]^k}$. Then

$$[\mathcal{F}_{G,g}^{(k)} \Phi_G](\boldsymbol{v}) = \mathbf{a}_G(g)^\top L_G^{\mathcal{B}} \mathbf{y}_{G,\Phi_k}^{\mathcal{B}}(\boldsymbol{v}). \tag{58}$$

This recovers every component of full spectralization.

The decoder does not need additional information about the eigenvectors or the graph: both $A_G^{\mathcal{B}}$ and $\mathbf{a}_G(g)$ are determined by the fixed bank and the ordered distinct eigenvalues in $\mathrm{spec}(T)$, which are retained in the restricted spectralization. Thus the same decoding rule applies to all graphs on which the rank condition holds. The construction is independent of Φ_k and $\boldsymbol{v}$, proving the proposition. □

The following construction proves the uniform recovery claim following Proposition 5.2.

Proposition E.1. *Fix positive integers k, M, and define*

$$\mathcal{B}_k(M) := \left((z_1, \ldots, z_k) \longmapsto \prod_{j=1}^k z_j^{\alpha_j} \right)_{\boldsymbol{\alpha} \in \{0,\ldots,M-1\}^k}. \tag{59}$$

This bank consists of M^k polynomial filters. For every graph whose operator T has at most M distinct eigenvalues, its evaluation matrix has full column rank. Hence its restricted spectralization determines full spectralization for every real-valued k-RGI.

Proof. Write $m = m_G \le M$ and form the matrix

$$V_G := \left((\lambda_r^G)^a\right)_{a\in\{0,\dots,M-1\},\, r\in[m]} \in \mathbb{R}^{M\times m}.$$

Its first m rows form a square Vandermonde matrix. Its determinant is $\prod_{1\le r<s\le m}(\lambda_s^G - \lambda_r^G)$, which is nonzero because the eigenvalues are distinct. Thus $\operatorname{rank} V_G = m$. With consistent product orderings of the rows and columns,

$$A_G^{\mathcal{B}_k(M)} = V_G^{\otimes k}, \qquad \operatorname{rank} A_G^{\mathcal{B}_k(M)} = (\operatorname{rank} V_G)^k = m^k.$$

Proposition 5.2 now applies. The monomials in (59) depend only on k and M; neither their coefficients nor their exponents depend on the graph or on its eigenvalues. □

Full column rank is a sufficient condition for recovery of all joint eigenspace-response coefficients. It is not asserted to be necessary for a particular invariant, whose response coefficients may satisfy additional constraints. In particular, the requirement of at least m_G^k filters belongs to this full-rank criterion.

E.2 Proof of Equation (3)

Fix a nonnegative integer K and univariate polynomials $p_0, \dots, p_K$ with $p_0 \equiv 1$ and $\deg p_\ell = \ell$. The cutoff K and these polynomials are fixed across graphs. Let $\mathbb{P}_{\le K}(\mathbb{R}^k)$ denote the space of k-variate polynomials of total degree at most K, write $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ and $|\boldsymbol{\alpha}|_1 := \sum_{j=1}^k \alpha_j$, and set

$$\mathcal{I}_{k,K} := \{\boldsymbol{\alpha} \in \mathbb{N}_0^k : |\boldsymbol{\alpha}|_1 \le K\}, \tag{60}$$

$$g_{\boldsymbol{\alpha}}(z_1, \dots, z_k) := \prod_{j=1}^k p_{\alpha_j}(z_j), \qquad \mathcal{B}_{k,K} := (g_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathcal{I}_{k,K}}. \tag{61}$$

Choose one ordering of $\mathcal{I}_{k,K}$ and use it for the bank, the response channels, and the columns of each mixing matrix.

Lemma E.2. *The bank $\mathcal{B}_{k,K}$ is a basis of $\mathbb{P}_{\le K}(\mathbb{R}^k)$, and*

$$|\mathcal{B}_{k,K}| = |\mathcal{I}_{k,K}| = \binom{K+k}{k}.$$

Proof. Write $p_\ell(z) = a_\ell z^\ell + q_\ell(z)$, where $a_\ell \ne 0$ and $\deg q_\ell < \ell$. For $\ell = 0$, the remainder is zero. For each $\boldsymbol{\alpha} \in \mathcal{I}_{k,K}$,

$$g_{\boldsymbol{\alpha}}(\mathbf{z}) = \left(\prod_{j=1}^k a_{\alpha_j}\right) z_1^{\alpha_1}\cdots z_k^{\alpha_k} + q_{\boldsymbol{\alpha}}(\mathbf{z}), \qquad \deg q_{\boldsymbol{\alpha}} < |\boldsymbol{\alpha}|_1,$$

where the remainder may again be zero. Suppose a nonzero linear combination of the $g_{\boldsymbol{\alpha}}$ vanishes, and let d be the largest total degree of an index with nonzero coefficient. Its homogeneous component of degree d is a linear combination of distinct monomials $z_1^{\alpha_1}\cdots z_k^{\alpha_k}$, each multiplied by the nonzero scalar $\prod_j a_{\alpha_j}$. Their linear independence forces all coefficients of total degree d to vanish, a contradiction. Thus $\mathcal{B}_{k,K}$ is linearly independent.

The monomials indexed by $\mathcal{I}_{k,K}$ form the usual basis of $\mathbb{P}_{\le K}(\mathbb{R}^k)$, so this independent family has the dimension of the polynomial space and is therefore a basis. Finally, adjoining the slack variable $\alpha_{k+1} = K - |\boldsymbol{\alpha}|_1$ identifies $\mathcal{I}_{k,K}$ with the nonnegative integer solutions of $\alpha_1 + \cdots + \alpha_{k+1} = K$. Stars and bars gives the stated cardinality. □

Proposition E.3. *For every $g \in \mathbb{P}_{\le K}(\mathbb{R}^k)$, there are unique coefficients $(w_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathcal{I}_{k,K}}$ satisfying*

$$g(\mathbf{z}) = \sum_{\boldsymbol{\alpha}\in\mathcal{I}_{k,K}} w_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}}(\mathbf{z}). \tag{62}$$

For every graph G, the corresponding filter operator is

$$\mathcal{F}^{(k)}_{G,g} = \sum_{\boldsymbol{\alpha}\in\mathcal{I}_{k,K}} w_{\boldsymbol{\alpha}} \bigotimes_{j=1}^{k} p_{\alpha_j}(T). \tag{63}$$

Consequently, for any fixed collection $g_1, \dots, g_F \in \mathbb{P}_{\le K}(\mathbb{R}^k)$, there is a matrix $W \in \mathbb{R}^{F\times\binom{K+k}{k}}$ such that

$$\left([\mathcal{F}^{(k)}_{G,g_b}\Phi_G](\boldsymbol{v})\right)_{b=1}^{F} = W\left(\left[\left(\bigotimes_{j=1}^{k} p_{\alpha_j}(T)\right)\Phi_G\right](\boldsymbol{v})\right)_{\boldsymbol{\alpha}\in\mathcal{I}_{k,K}}. \tag{64}$$

The matrix W depends only on the chosen filters $g_1, \dots, g_F$, the fixed polynomial basis, and its ordering; it is independent of G, Φ_k, and $\boldsymbol{v}$.

Proof. The unique expansion (62) follows from Lemma E.2. To identify the operator corresponding to a basis element, use the spectral decomposition of T in each tensor factor:

$$\bigotimes_{j=1}^{k} p_{\alpha_j}(T) = \bigotimes_{j=1}^{k}\left(\sum_{r_j=1}^{m_G} p_{\alpha_j}(\lambda^G_{r_j})P^G_{r_j}\right) = \sum_{\boldsymbol{r}\in[m_G]^k}\left(\prod_{j=1}^{k} p_{\alpha_j}(\lambda^G_{r_j})\right)P^G_{\boldsymbol{r}} = \mathcal{F}^{(k)}_{G,g_{\boldsymbol{\alpha}}}. \tag{65}$$

The last equality is Proposition 4.1. The map $g \mapsto \mathcal{F}^{(k)}_{G,g}$ is linear, so substituting (62) yields (63).

For each $b \in [F]$, let $(w_{b,\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathcal{I}_{k,K}}$ be the unique basis coefficients of g_b, and take $W = (w_{b,\boldsymbol{\alpha}})_{b,\boldsymbol{\alpha}}$. Apply (63) to Φ_G, evaluate at $\boldsymbol{v}$, and collect the identities over b to obtain (64). These coefficients arise from identities of polynomials on $\mathbb{R}^k$, before any graph or spectrum is chosen. This proves the asserted independence of W from the graph and the signal. □

Uniqueness here concerns the polynomial coefficients. It does not require the evaluated basis responses on a particular graph to be linearly independent. The constant basis element $g_{(0,\dots,0)} \equiv 1$ has operator $I_n^{\otimes k}$ and therefore preserves the raw signal.

Bivariate matrix form. For $k = 2$, identify Φ_G with its $n \times n$ signal matrix and use column-wise vectorization. The identity

$$\operatorname{vec}(BXA^\top) = (A \otimes B)\operatorname{vec}(X)$$

and the symmetry of $p_i(T)$ give

$$(p_i(T) \otimes p_j(T))\operatorname{vec}(\Phi_G) = \operatorname{vec}\big(p_j(T)\Phi_G p_i(T)\big). \tag{66}$$

Applying (63) with $k = 2$ therefore proves Proposition 3 in the main text, with exactly the stated order of the left and right matrix factors.

The same convention gives the spectral-decomposition implementation. Let $\widehat{\Phi}_G = U_G^\top \Phi_G U_G$ and set $(S_{g,G})_{ij} = g((\Lambda_G)_{jj}, (\Lambda_G)_{ii})$. Then the Fourier coefficient at row i and column j corresponds to the Kronecker basis vector with the first spectral index j and the second spectral index i. Hence

$$\mathcal{F}^{(2)}_{G,g}\Phi_G = U_G\big(S_{g,G} \odot \widehat{\Phi}_G\big)U_G^\top, \tag{67}$$

where the operator on the left is identified with its matrix-valued output. Its projector expansion shows that this expression is independent of the eigenbasis chosen within repeated eigenspaces.

A uniform total-degree cutoff. If the graph family satisfies $m_G \le M$, every monomial in $\mathcal{B}_k(M)$ has total degree at most $k(M-1)$. Therefore, for any fixed $K \ge k(M-1)$, each member of $\mathcal{B}_k(M)$ is a graph-independent linear combination of the basis bank $\mathcal{B}_{k,K}$. The basis responses thus determine the monomial-bank responses, and Proposition E.1 gives exact recovery of full spectralization on that family. For an arbitrary chosen cutoff K, Proposition E.3 instead guarantees exact realization of every filter in $\mathbb{P}_{\le K}(\mathbb{R}^k)$; it makes no claim that this truncated space recovers all continuous-filter responses.

With K and the basis fixed, the response vector on the right of (64) can be precomputed for each graph and input RGI. A learnable linear layer with F_{learn} rows then parameterizes an arbitrary collection of F_{learn} filters in the chosen polynomial space, using the same coefficients across graphs.

E.3 Pseudocode for Spectralized Structural Encodings

Algorithms 1 and 2 summarize the two implementations. Here, $n = |V(G)|$, and $W \in \mathbb{R}^{F_{\text{learn}} \times \frac{1}{2}(K+2)(K+1)}$ is the learnable linear layer, whose columns follow a fixed ordering of the nonnegative integer pairs (i, j) with $i + j \le K$. Both algorithms include the response to the fixed filter $g \equiv 1$. The returned matrices are evaluated at each node pair to form its response vector, which is consistent with (4).

Algorithm 1: SpecSE via Spectral decomposition

Input: T, Φ_G, MLP f_θ

[Precompute once]
Decompose $T = U_G \Lambda_G U_G^\top$
$\widehat{\Phi}_G \leftarrow U_G^\top \Phi_G U_G$

[Forward pass]
for $i, j = 1, \dots, n$ **do**
  $\boldsymbol{s} \leftarrow f_\theta((\Lambda_G)_{jj}, (\Lambda_G)_{ii})$
  $(S_{b,G})_{ij} \leftarrow s[b], b \in \{1, \dots, F_{\text{learn}}\}$
end
for $b = 1, \dots, F_{\text{learn}}$ **do**
  $Y_{b,G} \leftarrow U_G (S_{b,G} \odot \widehat{\Phi}_G) U_G^\top$
end
return $\left[\Phi_G \,\|\, (Y_{b,G})_{b=1}^{F_{\text{learn}}}\right]$

Algorithm 2: SpecSE via Polynomial filters

Input: T, Φ_G, $(p_\ell)_{\ell=0}^K$, W

[Precompute once]
for $i, j \ge 0$ *with* $i + j \le K$ **do**
  $H_i \leftarrow \Phi_G p_i(T)$
  $Z_{ij} \leftarrow p_j(T) H_i$
end

[Forward pass]
for $u, v \in V(G)$ **do**
  $\boldsymbol{z} \leftarrow (Z_{ij}(u, v))_{0 \le i+j \le K}$
  $\boldsymbol{y} \leftarrow W \boldsymbol{z}$
  $Y_{b,G}(u, v) \leftarrow y[b], b \in \{1, \dots, F_{\text{learn}}\}$
end

return $\left[\Phi_G \,\|\, (Y_{b,G})_{b=1}^{F_{\text{learn}}}\right]$

F Experimental Details

Datasets. QM9 is a real-world molecular dataset containing over 130,000 molecular graphs, annotated with 13 quantum chemical properties for regression. Our ZINC experiments use the ZINC-12K subset (Dwivedi et al., 2023), which contains 12,000 molecules with a single regression target. For CIFAR10, we adopt the graph representation of Dwivedi et al. (2023), in which each of the 60,000 images is represented as an 8-nearest-neighbor graph over its SLIC superpixels. The resulting task is to classify each graph into one of the 10 image categories. We use their benchmark split of 45,000 training, 5,000 validation, and 10,000 test graphs.

Input RGIs and preprocessing. Let $\mathcal{F}_5$ contain one fixed representative of each connected simple undirected graph with two to five vertices. We use $\mathcal{F}_5 \cup \{C_6\}$ for QM9 and $\mathcal{F}_5$ for CIFAR10. Each template contributes one two-rooted homomorphism-count channel, using the fixed ordered roots $r_1 = 0$ and $r_2 = \min N_F(0)$ in its supplied vertex labeling. For ZINC-12K, we use 36 direct two-rooted homomorphism-count channels selected from the published C_7/C_8 template bases, together with six prescribed rational linear combinations of counts from the C_3–C_8 bases following Bao et al. (2025). These latter channels are count-derived RGIs and may take signed or fractional values.

Table 6: Training settings for the GPS configurations.

Setting	QM9	ZINC-12K	CIFAR10
Training graphs	110,462	10,000	45,000
Validation graphs	10,000	1,000	5,000
Test graphs	10,000	1,000	10,000
Batch size	128	32	16
Epochs	1,200	2,000	100
Warmup epochs	20	50	5
Weight decay	10^{-5}	10^{-5}	10^{-4}

Spectral implementations. We use $T = -D^{-1/2}AD^{-1/2}$, where A is the unweighted, loopless, undirected adjacency matrix and inverse degrees are set to zero at isolated nodes. For CIFAR10, this topology is obtained by symmetrizing the original edge set. On QM9, we use full eigendecomposition and a compressed variant of the construction in Section 5. A bias-free linear map first mixes the input RGIs into 16 latent signals. An MLP of dimensions $2 \to 32 \to 32 \to 64$, with GELU after each hidden layer, maps each ordered eigenvalue pair to four filter values per latent signal. The resulting 64 response channels are mapped to the model width by an MLP with hidden width 64, batch normalization, and LeakyReLU.

On ZINC-12K and CIFAR10, we use a bivariate Chebyshev basis of total degree at most five. This gives 21 responses $p_i(T)\Phi_\alpha p_j(T)$ with $i+j \leq 5$, where p_i and p_j are Chebyshev polynomials of the first kind. A shared bias-free linear map $\mathbb{R}^{21} \to \mathbb{R}^{10}$ produces ten learnable responses for each RGI, alongside a separate raw-response channel. Combining the identity signal, a shared $11 \to 32 \to 1$ MLP reduces the responses of each RGI, followed by an MLP that mixes information across RGIs.

Integration with GPS. SpecSE enters the local message-passing branch as an edge encoding, and the original node features are projected to the full node-encoder width. All QM9 configurations use an R-GCN local branch with the depths and widths in Table 7. This changes MoSE's GIN-E local branch to R-GCN for U0, U, H, G, and ZPVE. Bond types remain discrete relations, while SpecSE vectors are summed over incoming edges and added to the R-GCN output. For G and ZPVE, diagonal spectral responses also provide the encodings of the added self-loops; the other released QM9 configurations assign zero encodings to these self-loops. The added self-loops do not enter the spectral operator. ZINC-12K uses eight GIN-E+Transformer layers, with widths 92 for the first two layers and 64 thereafter; CIFAR10 uses CustomGatedGCN as its local branch.

Training settings. We use AdamW with learning rate 10^{-3} and cosine decay following a warmup period. Table 6 lists the dataset-specific settings. The molecular tasks use L1 loss and sum pooling, while CIFAR10 uses cross-entropy loss and mean pooling. Attention dropout is 0.5 throughout; GPS dropout is zero for QM9 and ZINC-12K and 0.2 for CIFAR10. Checkpoints are selected using validation performance for each run. Following Bao et al. (2025), ZINC-12K and CIFAR10 uses four seeds, $0, 14, 48, 96$. QM9 uses five seeds, $0, 14, 48, 96, 100$.

Table 7: GPS configurations for QM9. Each tuple specifies (MPNN type, depth, width). All configurations use four attention heads.

	None	RWSE	MoSE	SpecSE
mu	(R-GCN, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)
alpha	(R-GCN, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)
HOMO	(GIN-E, 8, 128)	(R-GCN, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)
LUMO	(R-GCN, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)
gap	(R-GCN, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)
R2	(R-GCN, 10, 64)	(R-GCN, 8, 128)	(R-GCN, 10, 64)	(R-GCN, 10, 64)
ZPVE	(GIN-E, 10, 64)	(GIN-E, 8, 128)	(GIN-E, 8, 128)	(R-GCN, 8, 128)
U0	(GIN-E, 10, 64)	(GIN-E, 8, 128)	(GIN-E, 8, 128)	(R-GCN, 8, 128)
U	(R-GCN, 10, 64)	(GIN-E, 10, 64)	(GIN-E, 8, 128)	(R-GCN, 8, 128)
H	(GIN-E, 10, 64)	(GIN-E, 10, 64)	(GIN-E, 8, 128)	(R-GCN, 8, 128)
G	(R-GCN, 10, 64)	(GIN-E, 10, 64)	(GIN-E, 8, 128)	(R-GCN, 8, 128)
Cv	(R-GCN, 10, 64)	(GIN-E, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)
Omega	(GIN-E, 10, 64)	(GIN-E, 10, 64)	(R-GCN, 10, 64)	(R-GCN, 10, 64)