# WHEN IRRELEVANT GRAPH GROWTH CHANGES OLD-NODE PREDICTIONS: GRAPH EXTENSION CONSISTENCY FOR SPECTRAL POSITIONAL ENCODINGS

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

Graph positional and structural encodings are often recomputed when an input graph grows. We ask whether this recomputation can change a frozen predictor's output on persistent nodes, even when their features, original edges, and local computational neighborhoods remain unchanged. We formalize this requirement as *graph extension consistency*. For global spectral encodings that retain the $k$ smallest strictly positive Laplacian modes, we identify spectral budget competition: added regions can displace modes representing the original graph. Under the stated graph and spectral-gap assumptions, we derive exact old-node spectral-projector discrepancies under disjoint extension and a strictly positive limit as a connecting bridge weakens to zero. We further introduce a paired extension audit that isolates the predictive consequences of encoding recomputation while holding learned parameters, inference state, raw features, and protected non-encoding computations fixed. Across WebKB and Airports datasets, we observe encoding-dependent probability drift and, in some configurations, prediction flips. Matched controls that preserve the original structural inputs reduce this drift to zero or the numerical-control scale, while targeted interventions distinguish spectral-input sensitivity from other encoder-domain effects. These results establish graph extension consistency as a distinct theoretical and empirical criterion for evaluating graph-dependent structural encodings.

## 1 INTRODUCTION

Modern graph models, including message-passing GNNs and graph Transformers, often augment node representations or attention mechanisms with positional or structural encodings that make graph structure explicit (Rampášek et al., 2022; Kanatsoulis et al., 2025; Ito et al., 2025; Amboage et al., 2026). Unlike raw node attributes, however, many such encodings are themselves functions of the current input graph: when the graph changes, the structural descriptions assigned to persistent nodes may be recomputed. Predictive performance on a fixed graph does not, by itself, reveal whether predictions on the same nodes remain consistent after graph extension. This distinction matters when a trained predictor is reused after its input graph has grown: can growth elsewhere rewrite the model's decision on an old node?

Consider a predictor trained on a graph $G$ and then frozen. At test time, we preserve every node, edge, and raw feature in $G$ and append a completely disconnected graph $H$. We recompute the structural encodings jointly on $G \sqcup H$, while keeping model parameters, inference state, and the audited nodes' non-encoding computational dependencies fixed. An old node $v$ gains no neighbor and receives no new message along graph edges from $H$. Nevertheless, our audits show that its predicted probability can move, and its predicted class can sometimes flip. An unchanged node can thus receive a different prediction merely because an unrelated region enters the graph used to compute its structural input.

We trace this apparently non-local effect to structural encodings recomputed on the expanded graph. Within this pathway, we identify a precise mechanism in fixed-budget global spectral encodings: *spectral budget competition*. Such encodings retain only $k$ low-frequency Laplacian modes selected across the entire input graph (Reid et al., 2026; Flora et al., 2026). Under extension, $G$ and $H$



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.

share this finite budget: modes contributed by $H$ can enter the selected set and displace modes that previously described $G$. Even under a disjoint extension, this can change the spectral subspace retained on old nodes while leaving their original graph unchanged. The change concerns which structural directions survive selection, not merely the signs or bases used to represent an unchanged subspace.

The weakly connected case reveals a further distinction between vanishing coupling and recovery. When two connected graphs are joined by a bridge of weight $\varepsilon$, a nonzero Laplacian eigenvalue approaches zero with $\varepsilon$, placing its mode among the lowest-frequency candidates. For encodings that select the $k$ lowest strictly positive modes, we prove that, under explicit eigengap and ordering conditions, the discrepancy between the original spectral projector and the expanded projector restricted to $G$ can converge to a strictly positive limit as $\varepsilon \to 0^+$. The expanded Laplacian approaches that of $G \sqcup H$, whereas recovery of the old representation requires agreement with $G$. Thus, vanishing coupling need not restore the original spectral representation.

We formalize this task-conditioned requirement as *Graph Extension Consistency* (GEC). For extensions that preserve old-node targets and supply no new task-relevant evidence, GEC asks whether representations and frozen predictions, restricted to corresponding old nodes, remain unchanged under disconnected additions or recover as weak coupling vanishes. Prior work studies disjoint-union invariance for node classifiers (Pflueger et al., 2024), sign and basis invariance (Lim et al., 2023), and perturbation-stability guarantees (Wang et al., 2022; Huang et al., 2024). Related work also examines representation drift in streaming networks (Zhang et al., 2023) and spatio-temporal spectral encodings for dynamic graphs (Karmim et al., 2024). Within this framework, we quantify how spectral budget competition changes the subspace retained on old nodes when structural encodings are jointly recomputed, and characterize conditions under which this mismatch persists as a weak bridge vanishes. Controlled audits of frozen predictors then examine the predictive consequences of jointly recomputing structural encodings, with targeted interventions distinguishing spectral-input changes from other encoding-dependent effects.

To test predictive consequences, we introduce the *Extension Counterfactual Audit* (ECA). ECA trains on $G$, freezes the predictor, any learned encoder, and inference state, and recomputes structural inputs while holding old raw features and the relevant non-encoding computational dependencies fixed. Across WebKB and Airports, we observe encoding-dependent probability drift and class flips in some configurations; under matched controlled conditions, retaining old-node encodings preserves the original predictions without retraining. Further interventions distinguish the effects of spectral-input changes, encoder-domain enlargement, and graph-dependent preprocessing. A controlled WIRE spectral-attention configuration (Reid et al., 2026) also exhibits small attention and probability changes. Meanwhile, no-PE, RWSE, and configuration-matched R-PEARL controls remain stable, showing that the phenomenon depends on the structural encoding and its computational pipeline rather than being an inevitable consequence of graph extension.

Our central contribution is an exact characterization of old-node spectral-projector mismatch for fixed-budget global spectral encodings that select the lowest strictly positive modes. We derive the mismatch under disjoint extension and its strictly positive weak-bridge limit under the stated assumptions, quantifying why vanishing coupling need not recover the representation computed on the original graph. Building on this analysis, we use ECA to examine the predictive consequences in trained, frozen models and distinguish different pathways of change through controlled interventions. Finally, we establish corresponding design conditions and boundaries: fixed full-spectrum matrix functions satisfy disjoint consistency, while heat-kernel and finite-step random-walk encodings recover under vanishing weak coupling under the stated assumptions. Together with stable reference methods and pipeline-level controls, these results distinguish consistency of the spectral construction itself from consistency of the complete encoding pipeline.

## 2 Graph Extension Consistency

### 2.1 Persistent-node inference

Let $G = (V, E)$ be a finite undirected, possibly weighted graph with node features $X$. An extension $G' = (V', E')$, with $V \subseteq V'$, preserves the old induced graph, including edge weights, and raw features: $G'[V] = G$ and $X'|_V = X$. We train on $G$ and compare predictions before and

after extension, recomputing structural inputs while freezing the predictor, any learned encoder, and inference state. Writing $\theta$ for all learned parameters, the pipeline is

$$Z_G = \Phi_\theta(G, X), \qquad P_\theta(G) = F_\theta(G, X, Z_G) \in [0,1]^{|V|\times C}, \tag{1}$$

where $Z_G$ contains the structural inputs supplied to the predictor and $C$ is the number of classes. Ordering old nodes first gives the restriction operator $R_G = [I_{|V|}\ 0]$.

We audit a nonempty set $\mathcal{P} \subseteq V$ of *protected* nodes whose raw inputs and non-encoding computational dependencies remain unchanged. Protection includes propagation weights and normalization factors, as specified in Section 4. We consider extensions that preserve $y_{G'}(v) = y_G(v)$ for $v \in \mathcal{P}$ and supply no new task-relevant evidence about these targets. This task condition determines when preserving old predictions is appropriate.

### 2.2 Encoding and prediction comparisons

We compare encodings through a declared node-pair relation $K_\Phi(G) \in \mathbb{R}^{|V|\times|V|}$ that respects their admissible coordinate symmetries, suppressing fixed inputs and parameters. For an orthonormal spectral basis $U_G$, the projector $K_\Phi(G) = U_G U_G^\top$ removes sign and basis ambiguities (Lim et al., 2023; Huang et al., 2024). The old-node encoding discrepancy is

$$D_{\mathrm{enc}}(G, G') = \left\| K_\Phi(G) - R_G K_\Phi(G') R_G^\top \right\|_F . \tag{2}$$

At the prediction level, we measure the mean probability discrepancy on protected nodes:

$$D_{\mathrm{pred}}(G, G'; \mathcal{P}) = \frac{1}{|\mathcal{P}|} \sum_{v\in\mathcal{P}} \left\| \left[R_G P_\theta(G')\right]_v - \left[P_\theta(G)\right]_v \right\|_1 . \tag{3}$$

Equality of a comparison relation need not imply equality of the full structural inputs consumed by a predictor. Conversely, a predictor may be insensitive to a nonzero encoding discrepancy. We therefore evaluate the two levels separately. Randomized encoders are compared under a declared coupling that preserves shared and old-node random draws; Appendix B specifies the comparison and protection conventions.

### 2.3 Disconnected and weakly connected extensions

We use *Graph Extension Consistency* (GEC) to organize two requirements, each defined separately at the encoding or prediction level. Write $q \in \{\mathrm{enc}, \mathrm{pred}\}$ for the selected level, suppressing the fixed protected set in $D_{\mathrm{pred}}$.

For $H = (V_H, E_H)$ with $V_H \cap V = \emptyset$, let $J_0 = G \sqcup H$. *Disjoint Extension Consistency* (DEC) requires

$$D_q(G, J_0) = 0. \tag{4}$$

For a weakly connected extension, fix endpoints $a \in V$ and $b \in V_H$, and let $J_\varepsilon$ join them by an undirected edge of weight $\varepsilon > 0$. Keep $G$, $H$, and the endpoints fixed; prediction comparisons also fix $\mathcal{P}$, which must remain protected for all sufficiently small positive $\varepsilon$. *Weak-Bridge Extension Stability* (WBES) requires

$$\lim_{\varepsilon\to 0^+} D_q(G, J_\varepsilon) = 0. \tag{5}$$

The recovery target is $G$, not $J_0$: continuity toward the disconnected expanded graph does not establish recovery of the original representation or prediction.

Encoding-family guarantees quantify over the graph pairs or extension paths in their stated domain; prediction guarantees also specify the frozen pipeline and protected nodes. Section 3 analyzes spectral encodings under these requirements, and Section 4 examines their predictive consequences.

## 3 Spectral Budget Competition

A fixed-budget spectral encoder retains a prescribed number $k$ of Laplacian modes. Graph extension changes the candidates for this budget. We analyze the resulting competition through spectral projectors, separating changes in the selected subspace from changes in its coordinates.

### 3.1 A SHARED BUDGET ACROSS DISCONNECTED COMPONENTS

Disconnected components occupy separate Laplacian blocks but share a budget when modes are selected globally. Throughout this section, $G$ and $H$ are finite, connected, undirected graphs with nonnegative edge weights, no self-loops, at least two nodes each, and positive degrees. Let $n_G = |V|$, with adjacency and degree matrices $A_G, D_G$, and use the symmetric normalized Laplacian

$$L_G = I - D_G^{-1/2} A_G D_G^{-1/2}, \tag{6}$$

with analogous notation for $H$. Write $\mu_1(G) \le \cdots \le \mu_{n_G-1}(G)$ for its strictly positive eigenvalues, counted with multiplicity, and $u_i^G$ for corresponding orthonormal eigenvectors. For $1 \le k \le n_G - 1$, define

$$\Pi_k^+(G) = \sum_{i=1}^{k} u_i^G (u_i^G)^\top, \qquad \Delta_k(G, J) = \frac{\|\Pi_k^+(G) - R_G \Pi_k^+(J) R_G^\top\|_F}{\sqrt{k}}. \tag{7}$$

Thus, $\Delta_k$ is the encoding discrepancy of Section 2, normalized by $\|\Pi_k^+(G)\|_F = \sqrt{k}$. We require $\mu_k(G) < \mu_{k+1}(G)$ when $k < n_G - 1$ and set $\Pi_0^+(G) = 0$. On every expanded graph, the selection rule excludes all exact zero modes and retains the $k$ smallest strictly positive modes.

Since $L_{G \sqcup H} = L_G \oplus L_H$, its positive spectrum is the multiset union of the component spectra, ordered as $\beta_1 \le \beta_2 \le \cdots$.

**Proposition 1 (Disjoint mode replacement)** *Suppose $\beta_k < \beta_{k+1}$, and let $r_0$ be the dimension contributed by $G$ to the invariant subspace associated with $\beta_1, \ldots, \beta_k$. Then*

$$R_G \Pi_k^+(G \sqcup H) R_G^\top = \Pi_{r_0}^+(G), \qquad \Delta_k(G, G \sqcup H) = \sqrt{\frac{k - r_0}{k}}. \tag{8}$$

Restriction to $G$ removes the selected modes supported on $H$ but cannot restore the $k - r_0$ displaced directions of $G$. Consequently, $r_0 < k$ violates DEC despite leaving $G$ unchanged. Because the discrepancy compares projectors, sign alignment or basis changes cannot remove it. Appendix A.1 provides the proof.

### 3.2 VANISHING COUPLING DOES NOT RESTORE THE ORIGINAL SUBSPACE

Let $J_\varepsilon$ join fixed endpoints of $G$ and $H$ by an edge of weight $\varepsilon > 0$, and let $L_\varepsilon$ include the resulting endpoint-degree changes. Joining the components creates a positive bridge mode whose eigenvalue approaches zero. For sufficiently small $\varepsilon$, this mode occupies one of the $k$ positions, leaving only $k - 1$ for modes approaching the original positive spectra. Define the pre-bridge quantities

$$v_G = \mathbf{1}^\top D_G \mathbf{1}, \qquad q_G = \frac{D_G^{1/2}\mathbf{1}}{\sqrt{v_G}}, \qquad \alpha = \frac{v_H}{v_G + v_H}, \tag{9}$$

with $v_H$ and $q_H$ defined analogously.

**Theorem 2 (Non-vanishing weak-bridge discrepancy)** *Under the standing graph and reference-gap assumptions, suppose $\beta_{k-1} < \beta_k$ when $k > 1$. Let $r_+$ be the dimension contributed by $G$ to the complete invariant subspace associated with $\beta_1, \ldots, \beta_{k-1}$, with $r_+ = 0$ when $k = 1$. As $\varepsilon \to 0^+$,*

$$R_G \Pi_k^+(J_\varepsilon) R_G^\top \longrightarrow \Pi_{r_+}^+(G) + \alpha q_G q_G^\top, \tag{10}$$

$$\lim_{\varepsilon \to 0^+} \Delta_k(G, J_\varepsilon) = \sqrt{\frac{k - r_+ + \alpha^2}{k}} > 0. \tag{11}$$

**Proof sketch.** The bridge-mode projector restricted to $G$ converges to $\alpha q_G q_G^\top$. The spectral gap makes the remaining selected subspace converge to the lowest $k-1$ merged positive modes, retaining $r_+$ directions on $G$. Orthogonality to $q_G$ gives squared discrepancy $k - r_+ + \alpha^2$. Appendix A.2 derives these limits using finite-dimensional perturbation theory (Kato, 1995).

Writing $s = k - 1 - r_+$, the squared limit is $(1 + s + \alpha^2)/k$: the bridge mode displaces one original direction, the selected positive modes from $H$ displace another $s$, and the restricted bridge

mode contributes $\alpha^2$. The discrepancy therefore persists even when $s = 0$. Crucially, its limit is computable from the component spectra and pre-bridge volumes, without training a predictor or fitting the observed curve.

Matrix continuity takes $L_\varepsilon$ toward $L_G \oplus L_H$, whereas WBES requires recovery of the representation on $G$. The positive-mode rule also excludes the bridge mode at exactly zero but includes it for sufficiently small positive weights. Appendix A.3 illustrates this distinction with a four-node example.

### 3.3 Consistency-preserving design principles

The shared selection budget can be avoided by using a fixed spectral matrix function. For $f : [0, 2] \to \mathbb{R}$, the relation $K_f(G) = f(L_G)$ satisfies

$$f(L_G \oplus L_H) = f(L_G) \oplus f(L_H), \tag{12}$$

and hence DEC (Higham, 2008). For fixed $t > 0$, the full heat kernel additionally satisfies

$$\left\| R_G e^{-tL_\varepsilon} R_G^\top - e^{-tL_G} \right\|_2 \le t \left\| L_\varepsilon - (L_G \oplus L_H) \right\|_2 = O(\varepsilon), \tag{13}$$

which gives WBES, also in Frobenius norm for a fixed graph pair. Likewise, finite-step return features $((M_G^j)_{vv})_{j=1}^T$, where $M_G = D_G^{-1} A_G$, are unchanged under disjoint extension and have maximum old-node coordinate change $O(T\varepsilon)$ along the weak-bridge path for fixed $T$. Proofs appear in Appendix A.4. These guarantees concern the specified relations or features; additional graph-dependent encoding operations require pipeline-level evaluation.

For an existing predictor, independent encoding of unchanged connected components removes cross-component budget competition under disjoint extension. A positive bridge merges those components, so this operation alone does not establish weak-bridge recovery. Alternatively, retaining all original structural inputs required by protected computations preserves a deterministic frozen predictor's outputs when its remaining dependencies are fixed. These controls concern old-node consistency, without guarantees for new-node accuracy or replacement-encoder performance. Section 4 tests the predictive consequences of encoding recomputation.

## 4 Extension Counterfactual Audit

We use the Extension Counterfactual Audit (ECA) to compare old-node predictions before and after extension using the same frozen checkpoint. Paired controls isolate the effect of encoding recomputation, and internal interventions examine its specific sources.

### 4.1 Paired frozen inference

We train each pipeline on the original graph $G$, select its checkpoint using the original validation split, and freeze the predictor, learned encoder, and inference state. Predictions on $G$ provide the reference. On the extension, we compare joint encoding recomputation with retention of the original structural inputs required by protected computations. Both conditions use the same expanded downstream setup, differing only in their structural inputs. Joint recomputation follows the encoder's preprocessing and selection rules, retaining its budget when applicable.

Under the task and protection conditions of Section 2, a change that disappears when the original structural inputs are retained can be attributed to the recomputation pathway. Randomness is coupled as specified in Appendix B.2.

### 4.2 Protecting non-encoding computation

For the two-layer GCN audit, we evaluate protected old test nodes whose relevant normalized propagation computations remain unchanged. For WIRE, a source-block mask excludes cross-source attention while allowing recomputed spectral coordinates to change attention within $G$. The mask fixes the permitted interaction domain, not the encoding-dependent attention weights. Appendix B.3 details these safeguards. Changes inside the structural encoder remain part of the recomputation pathway under audit.

### 4.3 CONTROLS AND OUTCOMES

Source-preserving controls evaluate the frozen encoder separately on the original $G/H$ groups; cached controls directly reuse original structural inputs. With a positive bridge, source preservation still uses the original groups rather than current connected components. These controls reuse the audited model, whereas alternative encoders and no-encoding references are trained separately (Appendix B.5).

Further interventions distinguish spectral-input changes from encoder-domain and preprocessing effects (Appendix B.6). We report probability discrepancy $D_{\mathrm{pred}}$, class-flip rates with fixed tie-breaking, and original and extended accuracy on the same evaluated nodes. Protected-node counts and statistical aggregation are reported with the experimental results.

## 5 EXPERIMENTS

We examine frozen prediction changes, test the spectral weak-bridge limits, and distinguish the sources of change within encoding pipelines.

### 5.1 EXPERIMENTAL SETUP

We use the Brazil and Europe Airports graphs (Ribeiro et al., 2017) and the Cornell, Texas, and Wisconsin WebKB graphs from the release of Pei et al. (2020). Airports labels retain their original regional definitions, and old raw features remain fixed throughout extension. Our main comparison uses two-layer GCNs with BasisNet, SPE, RWSE, and no PE, together with a configuration-matched local R-PEARL implementation (Lim et al., 2023; Huang et al., 2024; Kanatsoulis et al., 2025). BasisNet and SPE use $k = 16$. Structural curves select the lowest strictly positive modes; the separate WIRE audit includes zero modes, and LLPE uses the full spectrum. Implementation and training details appear in Appendix C.

The main weak-bridge comparison uses ten training splits, three bridge locations, and $\varepsilon \in \{1, 10^{-3}, 10^{-5}\}$. Table 1 reports medians across splits after taking within-split medians over bridge locations and, for R-PEARL, Monte Carlo banks. Descriptive 95% split-bootstrap intervals, corresponding accuracies, and evaluated-node counts appear in Appendix C.3. Repeated interventions on one model are not independent training replicates.

### 5.2 FROZEN PREDICTIONS UNDER GRAPH EXTENSION

A nonzero spectral-projector discrepancy does not by itself imply a change in a trained model's output; we first test whether encoding recomputation changes frozen predictions on old nodes. Joint recomputation can change predictions even under disconnected extension. For SPE on Brazil ← Europe, median probability drift is 0.5315 and median flip rate is 10.71%; cached original encodings make both zero. Encoder-domain enlargement is the dominant source in this case (Section 5.5). This disjoint comparison evaluates all old test nodes; the weak-bridge comparison evaluates the protected subset.

Table 1 shows that joint recomputation can also change protected-node predictions at $\varepsilon = 10^{-5}$, while source-preserving controls yield zero drift and flips for the same models. Separately trained RWSE, R-PEARL, and no-PE references remain stable to zero or numerical-scale drift.

Probability drift and class flips measure different effects: SPE on Brazil ← Europe has substantial median probability drift but zero median flips. Accuracy changes include both improvements and declines (Appendix C.3). Thus, encoding-induced prediction changes do not imply inevitable performance degradation.

### 5.3 DOES WEAKENING THE CONNECTION RESTORE THE REPRESENTATION?

We test the representation-level limit using eleven positive bridge weights from 1 to $10^{-5}$ and budgets $k \in \{8, 16, 32\}$. Figure 1 shows that the bridge eigenvalue decreases toward zero while the old-node projector discrepancy approaches a nonzero plateau.

Table 1: Frozen prediction changes under weak-bridge extension ($\varepsilon = 10^{-5}$). Entries are medians over ten training splits on protected old test nodes. Source-preserving rows reuse the preceding joint model. Intervals and corresponding accuracies are reported in Appendix C.3.

| | Brazil ← Europe | | Europe ← Brazil | |
|---|---|---|---|---|
| Encoding / handling | Prob. $L_1$ | Flips (%) | Prob. $L_1$ | Flips (%) |
| *Same-model paired comparisons* | | | | |
| BasisNet (joint) | 0.0588 | 17.16 | 0.0109 | 3.59 |
| Source-preserving | 0 | 0 | 0 | 0 |
| SPE (joint) | 0.4874 | 0.00 | 0.0908 | 9.92 |
| Source-preserving | 0 | 0 | 0 | 0 |
| *Separately trained references* | | | | |
| R-PEARL | $4.58 \times 10^{-7}$ | 0 | $1.73 \times 10^{-8}$ | 0 |
| RWSE | $8.62 \times 10^{-9}$ | 0 | 0 | 0 |
| No PE | 0 | 0 | 0 | 0 |

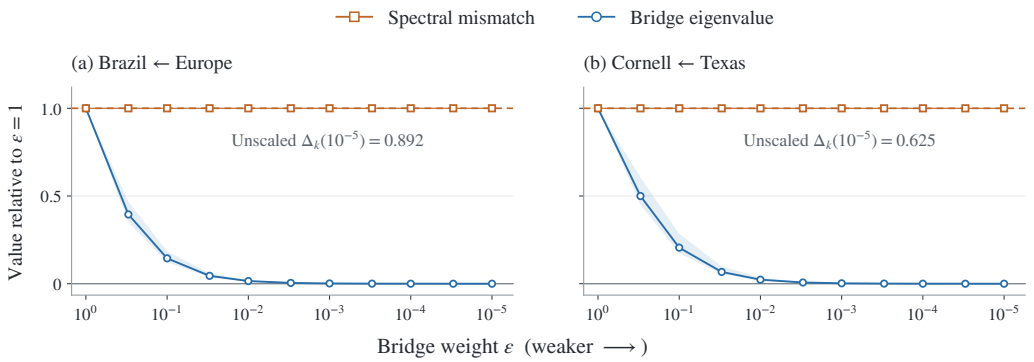



Figure 1: **Weakening the bridge need not restore the original spectral representation.** At $k = 16$, blue and orange curves show the bridge eigenvalue and projector discrepancy, each divided by its own value at $\varepsilon = 1$ along each fixed-endpoint path. Markers and bands denote the median and minimum–maximum range across three bridge locations; the bands are not confidence intervals. Dashed lines show the correspondingly rescaled theoretical limits, and annotations report $\Delta_k$ at $10^{-5}$ before display rescaling. All sampled weights are positive.

At the smallest weight, the minimum $\Delta_k$ across the full sweep is 0.6251 on WebKB and 0.5026 on Airports, before the figure's display rescaling. The maximum absolute difference from the theoretical limits is below $4.1 \times 10^{-9}$. The theoretical limits are computed from the component spectra and pre-bridge volumes, without fitting the observed curves. Agreement with these limits therefore checks the predicted magnitude of the residual mismatch across the evaluated graph pairs, bridge locations, and spectral budgets, rather than merely confirming the presence of a nonzero plateau. These results support spectral non-recovery under vanishing coupling. Predictive responses remain model-dependent; supplementary bridge-weight comparisons appear in Appendix C.5.

### 5.4 A CONTEMPORARY SPECTRAL-ATTENTION INSTANCE

We next ask whether extension sensitivity also appears when finite-spectrum coordinates modulate attention, even after cross-source interactions are blocked. Table 2 reports a controlled WIRE audit wrapper under this restriction (Reid et al., 2026). Its attention layer passes numerical parity checks against the official implementation (Appendix C.4). Entries are descriptive medians over intervention records within each dataset family, rather than medians over independent models.

Table 2: WIRE with cross-source attention blocked. Attention drift is the relative Frobenius change in first-layer old-to-old pre-softmax scores. The last column counts records with at least one class flip. Source-preserving controls reuse the WIRE model; no-WIRE references are trained separately.

| Family | Encoding / handling | Attention drift | Prob. $L_1$ | Records with flips |
|---|---|---|---|---|
| Airports | WIRE (joint) | $1.83 \times 10^{-3}$ | $3.73 \times 10^{-4}$ | $8/720$ |
| | Source-preserving | 0 | $3.28 \times 10^{-8}$ | $0/720$ |
| | No-WIRE | 0 | $3.07 \times 10^{-8}$ | $0/720$ |
| WebKB | WIRE (joint) | $1.80 \times 10^{-3}$ | $6.38 \times 10^{-7}$ | $8/1080$ |
| | Source-preserving | 0 | $3.90 \times 10^{-8}$ | $0/1080$ |
| | No-WIRE | 0 | $3.87 \times 10^{-8}$ | $0/1080$ |

Table 3: Encoding-pathway interventions under disjoint extension. SPE and BasisNet entries are medians over ten splits for Brazil ← Europe. LLPE entries are pooled medians over 150 intervention records across five directions. Comparisons are within each encoder block.

| Encoder | Intervention / pathway | Prob. $L_1$ |
|---|---|---|
| SPE | Joint recomputation | 0.5315 |
| | Spectral-input replacement only | $4.25 \times 10^{-5}$ |
| | Encoder-domain enlargement only | 0.5315 |
| BasisNet | Joint recomputation | 0.0896 |
| | Spectral-input replacement only | 0.0887 |
| LLPE | Author-preprocessing path | 0.4532 |
| | Combined diagnostic control | $1.41 \times 10^{-8}$ |

Joint recomputation changes old-to-old scores in all 1,800 WIRE records despite zero cross-source attention. Source-preserving coordinates eliminate this score drift, with remaining probability differences at the numerical control scale. The changed scores concern interactions among old nodes that were already permitted before extension. The paired comparison thus identifies a dependence on the expanded graph through the coordinates used to score those interactions, even though the attention mask prevents old nodes from attending to the added region. Only 16 records contain a class flip, and median flip rates are zero in every graph-pair, budget, and bridge-weight condition. The audited interface is therefore extension-sensitive, with heterogeneous and often small predictive effects.

### 5.5 Encoding-pathway attribution and consistency boundaries

Joint recomputation can change both spectral inputs and other graph-dependent encoder computations; we next test which changes account for the observed prediction drift. Table 3 summarizes disjoint-extension interventions that distinguish these sources. Implementation details appear in Appendix B.6.

For SPE, domain enlargement nearly reproduces the full drift and all class flips through accumulated responses to newly added zero-input columns. Replacing only the spectral kernel produces much smaller drift. BasisNet also responds to spectral-input replacement with its computational domain fixed: its median probability drift is $0.0887$, close to $0.0896$ under joint recomputation. These comparisons evaluate all old test nodes.

For full-spectrum LLPE (Ito et al., 2025), the full-basis projector discrepancy is below $1.6 \times 10^{-7}$, yet the author-preprocessing path changes protected-node predictions. The combined diagnostic controls bases and the spectral mapping domain, removes graph-level min–max preprocessing, and compares against its own original-graph output. Its near-zero drift demonstrates a consistent diagnostic path, not a performance-preserving replacement.

These results distinguish consistency of the spectral construction from consistency of the complete encoding pipeline. Additional external-graph interventions are reported in Appendix C.6.

## 6 Related Work

**Graph extension and prediction consistency.** Disjoint-union invariance has been formalized for GNN node classifiers (Pflueger et al., 2024), while metamorphic testing uses receptive-field invariances to detect GCN implementation bugs (Wang et al., 2020). Dynamic embedding methods study temporal stability (Gallagher et al., 2021); SIP analyzes streaming embedding-space drift and also evaluates recomputed embeddings with unchanged downstream classifiers (Zhang et al., 2023). Our paired audit focuses on extensions that preserve the targets of old nodes and supply no new task-relevant evidence. By holding model parameters and non-encoding computational dependencies fixed, it isolates the effect of joint structural-encoding recomputation on those nodes' predictions.

**Spectral selection and stability.** Classical spectral clustering identifies uneven coverage of disconnected blocks when selecting leading similarity-matrix eigenvectors (von Luxburg, 2007). Related work discusses localization of leading normalized adjacency eigenvectors on small disconnected or weakly connected subgraphs (Sarkar & Bickel, 2015). SignNet and BasisNet address eigenvector sign and basis symmetries (Lim et al., 2023), PEG and SPE study perturbation stability (Wang et al., 2022; Huang et al., 2024), and recent work characterizes truncation-dependent expressivity (Flora et al., 2026). We quantify changes in the spectral directions retained on old nodes under graph extension. For fixed-budget encodings selecting the smallest strictly positive Laplacian modes, we derive exact projector discrepancies under disjoint extension and an exact, strictly positive weak-bridge limit under the stated graph and spectral-gap assumptions. This distinguishes recovery of the original representation from stability around the expanded disconnected graph.

**Encoding designs and pipeline effects.** PEARL learns positional encodings through GNNs, LLPE uses the full Laplacian spectrum, and LEAP constructs learnable local topological features (Kanatsoulis et al., 2025; Ito et al., 2025; Amboage et al., 2026). These constructions offer alternatives to hard global spectral truncation, but avoiding mode eviction alone does not establish consistency of the complete pipeline. Graph-wide normalization can also couple disconnected components through shared statistics, as illustrated by the InstanceNorm example discussed in Eliasof et al. (2024, Example C.4). Our analysis establishes disjoint consistency for fixed spectral matrix functions and weak-coupling recovery for heat-kernel and finite-step random-walk features under the stated assumptions. Our experiments complement these results by separating spectral-input changes from encoder-domain and preprocessing effects.

## 7 Conclusion

We show that joint recomputation of structural encodings can change frozen predictions on old nodes even when graph extension supplies no new task-relevant evidence. A fixed global spectral budget allows newly added regions to displace spectral directions of the unchanged original graph, even under disconnected extension. For encodings selecting the lowest strictly positive Laplacian modes, we derive exact old-node projector discrepancies and strictly positive weak-bridge limits under explicit assumptions. Thus, vanishing coupling need not recover the original spectral representation. Our paired ECA experiments isolate the predictive effects of encoding recomputation and distinguish spectral-input sensitivity from encoder-domain and preprocessing effects. With the remaining computation fixed, retaining old structural inputs reduces measured drift to zero or the numerical-control scale in our matched controls. These findings complement the consistency conditions established for fixed spectral matrix functions and finite-step random-walk features. Together, these results show that avoiding spectral truncation alone does not guarantee consistency of the complete encoding pipeline. The graph dependence of a predictor must therefore be examined through both its structural-input construction and its downstream computation. Our analysis and paired audits provide concrete criteria for evaluating old-node consistency under task-irrelevant graph growth.

## AI USE STATEMENT

Generative AI tools were used solely for grammar and language polishing; all AI-assisted edits were reviewed by the authors, who take full responsibility for the final content.

## REPRODUCIBILITY STATEMENT

The graph-extension setting, consistency criteria, and theoretical assumptions are specified in Sections 2 and 3, with proofs and an illustrative example provided in Appendix A. Section 4 and Appendix B describe the frozen-inference audit, protected computations, randomness coupling, paired controls, and encoding-pathway interventions. Dataset preparation, data splits, model and training configurations, evaluation metrics, and statistical aggregation are documented in Section 5.1 and Appendices C.1–C.2. Appendices C.3–C.6 provide descriptive uncertainty intervals, corresponding accuracies, implementation checks, and additional graph-extension results.

## REFERENCES

Juan Amboage, Ernst Röell, Patrick Schnider, and Bastian Rieck. LEAP: Local ECT-based learnable positional encodings for graphs. In *International Conference on Learning Representations*, 2026.

Moshe Eliasof, Beatrice Bevilacqua, Carola-Bibiane Schönlieb, and Haggai Maron. GRANOLA: Adaptive normalization for graph neural networks. In *Advances in Neural Information Processing Systems*, volume 37, pp. 90514–90551, 2024.

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

Ian Gallagher, Andrew Jones, and Patrick Rubin-Delanchy. Spectral embedding for dynamic networks with stability guarantees. In *Advances in Neural Information Processing Systems*, volume 34, pp. 10158–10170, 2021.

Nicholas J. Higham. *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2008.

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*, 2024.

Michael Ito, Jiong Zhu, Dexiong Chen, Danai Koutra, and Jenna Wiens. Learning Laplacian positional encodings for heterophilous graphs. In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 2755–2763. PMLR, 2025.

Charilaos Kanatsoulis, Evelyn Choi, Stefanie Jegelka, Jure Leskovec, and Alejandro Ribeiro. Learning efficient positional encodings with graph neural networks. In *International Conference on Learning Representations*, 2025.

Yannis Karmim, Marc Lafon, Raphaël Fournier-S'niehotta, and Nicolas Thome. Supra-laplacian encoding for transformer on dynamic graphs. In *Advances in Neural Information Processing Systems*, volume 37, pp. 17215–17246, 2024.

Tosio Kato. *Perturbation Theory for Linear Operators*. Springer, Berlin, Heidelberg, second edition, 1995.

Derek Lim, Joshua Robinson, Lingxiao Zhao, Tess Smidt, Suvrit Sra, Haggai Maron, and Stefanie Jegelka. Sign and basis invariant networks for spectral graph representation learning. In *International Conference on Learning Representations*, 2023.

Hongbin Pei, Bingzhe Wei, Kevin Chen-Chuan Chang, Yu Lei, and Bo Yang. Geom-GCN: Geometric graph convolutional networks. In *International Conference on Learning Representations*, 2020.

Maximilian Pflueger, David Tena Cucala, and Egor V. Kostylev. Recurrent graph neural networks and their connections to bisimulation and logic. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 14608–14616, 2024.

Ladislav Rampášek, Michael Galkin, Vijay Prakash Dwivedi, Anh Tuan Luu, Guy Wolf, and Dominique Beaini. Recipe for a general, powerful, scalable graph transformer. In *Advances in Neural Information Processing Systems*, volume 35, pp. 14501–14515, 2022.

Isaac Reid, Arijit Sehanobish, Cederik Höfs, Bruno Mlodozeniec, Leonhard Vulpius, Federico Barbero, Adrian Weller, Krzysztof Choromanski, Richard E. Turner, and Petar Veličković. Rotary position encodings for graphs. In *International Conference on Machine Learning*, 2026.

Leonardo F. R. Ribeiro, Pedro H. P. Saverese, and Daniel R. Figueiredo. struc2vec: Learning node representations from structural identity. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 385–394. ACM, 2017.

Purnamrita Sarkar and Peter J. Bickel. Role of normalization in spectral clustering for stochastic blockmodels. *The Annals of Statistics*, 43(3):962–990, 2015.

Ulrike von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007.

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.

Yiwei Wang, Wei Wang, Yujun Cai, Bryan Hooi, and Beng Chin Ooi. Detecting implementation bugs in graph convolutional network based node classifiers. In *2020 IEEE 31st International Symposium on Software Reliability Engineering (ISSRE)*, pp. 313–324. IEEE, 2020.

Yanwen Zhang, Huiwen Wang, and Jichang Zhao. Space-invariant projection in streaming network embedding. *Information Sciences*, 649:119637, 2023.

# A Proofs and Examples for Spectral Budget Competition

This appendix proves the results in Section 3. We use its standing graph assumptions, fixed budget $k$, reference spectral gap, and strictly positive mode-selection rule. All limits keep the component graphs and bridge endpoints fixed. The volumes $v_G, v_H$ are measured before adding the bridge.

## A.1 Proof of disjoint mode replacement

For a disjoint extension,

$$L_{G \sqcup H} = L_G \oplus L_H.$$

Its positive spectrum is therefore the multiset union of the positive spectra of $G$ and $H$. The gap $\beta_k < \beta_{k+1}$ ensures that selecting the first $k$ positive modes includes complete eigenspaces at the selection boundary.

The selected invariant subspace decomposes into contributions supported on the two components. Its $r_0$-dimensional contribution on $G$ consists of the lowest $r_0$ positive modes of $G$. Consequently,

$$R_G \Pi_k^+(G \sqcup H) R_G^\top = \Pi_{r_0}^+(G).$$

This statement concerns spectral projectors and remains valid when eigenvalues coincide across the two components: complete eigenspaces are selected, so arbitrary choices of bases within them do not affect the projector.

Since $r_0 \le k$, the retained subspace on $G$ is contained in the reference subspace. The difference $\Pi_k^+(G) - \Pi_{r_0}^+(G)$ is thus an orthogonal projector of rank $k - r_0$, giving

$$\left\| \Pi_k^+(G) - \Pi_{r_0}^+(G) \right\|_F^2 = k - r_0.$$

Dividing by $\|\Pi_k^+(G)\|_F^2 = k$ and taking square roots proves Proposition 1.

### A.2 PROOF OF THE WEAK-BRIDGE LIMIT

Let $a \in V_G$ and $b \in V_H$ be the fixed bridge endpoints. All vectors below use the joint node ordering $(V_G, V_H)$. Write

$$A_\varepsilon = A_G \oplus A_H + \varepsilon(e_a e_b^\top + e_b e_a^\top), \qquad D_\varepsilon = D_G \oplus D_H + \varepsilon(e_a e_a^\top + e_b e_b^\top).$$

Set $L_0 = L_G \oplus L_H$.

**Regularity near zero.** All pre-bridge degrees are positive. Therefore $D_\varepsilon^{-1/2}$ is analytic in a neighborhood of zero, and

$$L_\varepsilon = I - D_\varepsilon^{-1/2} A_\varepsilon D_\varepsilon^{-1/2} = L_0 + O(\varepsilon)$$

is a real-symmetric analytic matrix family. For $\varepsilon > 0$, the joined graph is connected and has exactly one zero eigenvalue.

**Splitting the two zero modes.** Consider the equivalent generalized eigenproblem

$$\mathcal{L}_\varepsilon x_\varepsilon = \lambda_\varepsilon D_\varepsilon x_\varepsilon, \qquad \mathcal{L}_\varepsilon = D_\varepsilon - A_\varepsilon.$$

At zero, the nullspace of $\mathcal{L}_0$ is spanned by the component indicators, collected in

$$B = \begin{bmatrix} \mathbf{1}_G & 0 \\ 0 & \mathbf{1}_H \end{bmatrix}.$$

The bridge contributes

$$\mathcal{L}_0' = (e_a - e_b)(e_a - e_b)^\top.$$

For a branch emerging from zero, write $x_0 = Bc$ and $\lambda_\varepsilon = \varepsilon\nu + O(\varepsilon^2)$. Differentiating the generalized eigenproblem and multiplying by $B^\top$ gives

$$B^\top \mathcal{L}_0' B c = \nu B^\top D_0 B c.$$

The term involving $D_0'$ is multiplied by the zero-order eigenvalue and therefore vanishes. The degree changes have not been omitted. Explicitly, the reduced problem is

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} c = \nu \begin{bmatrix} v_G & 0 \\ 0 & v_H \end{bmatrix} c. \tag{14}$$

Its eigenvalues are $0$ and $1/v_G + 1/v_H$. One branch remains exactly zero. The other has positive first-order coefficient, while every remaining eigenvalue converges to a positive eigenvalue of $L_0$. Hence, for sufficiently small positive $\varepsilon$, this branch is the smallest positive eigenvalue:

$$\lambda_{\mathrm{br}}(\varepsilon) = \varepsilon \left( \frac{1}{v_G} + \frac{1}{v_H} \right) + O(\varepsilon^2). \tag{15}$$

The expansion and the associated invariant-subspace limit follow from analytic perturbation theory (Kato, 1995).

For the nonzero first-order eigenvalue, take $c = (1/v_G, -1/v_H)^\top$. Transforming $Bc$ into normalized-Laplacian coordinates via $D_0^{1/2}$ and normalizing gives

$$w_0 = \begin{bmatrix} \sqrt{\alpha}\, q_G \\ -\sqrt{1-\alpha}\, q_H \end{bmatrix}, \qquad w_\varepsilon w_\varepsilon^\top \longrightarrow w_0 w_0^\top, \tag{16}$$

where $w_\varepsilon$ is a unit bridge eigenvector. The projector limit is independent of eigenvector signs. Restricting it to $G$ yields

$$R_G w_0 w_0^\top R_G^\top = \alpha q_G q_G^\top.$$

**Convergence of the remaining selected subspace.** When $k = 1$, there are no remaining selected modes. For $k > 1$, the positive eigenvalues outside the bridge branch converge to the merged component spectrum $\beta_1, \beta_2, \ldots$. The lower separation $\beta_1 > 0$ and the selection gap $\beta_{k-1} < \beta_k$ isolate the complete invariant subspace associated with $\beta_1, \ldots, \beta_{k-1}$.

Let $Q_\varepsilon$ be the projector onto the $k-1$ selected positive modes other than the bridge mode. A fixed contour $\Gamma$ separating this spectral cluster from the remaining spectrum gives

$$Q_\varepsilon = \frac{1}{2\pi \mathrm{i}} \int_\Gamma (zI - L_\varepsilon)^{-1}\, dz \longrightarrow Q_0.$$

Indeed, the resolvents converge uniformly on $\Gamma$ because $L_\varepsilon \to L_0$ and the contour remains separated from the spectrum. Internal eigenvalue multiplicities cause no ambiguity because the entire isolated subspace is retained.

The $G$ block of $Q_0$ is $\Pi^+_{r_+}(G)$. Combining it with the bridge contribution gives

$$R_G \Pi^+_k(J_\varepsilon) R_G^\top = R_G\big(w_\varepsilon w_\varepsilon^\top + Q_\varepsilon\big) R_G^\top \longrightarrow \Pi^+_{r_+}(G) + \alpha q_G q_G^\top.$$

For $k = 1$, the same identity holds with $Q_\varepsilon = 0$ and $r_+ = 0$. This proves Equation 10.

**Computing the discrepancy.** The matrix $\Pi^+_k(G) - \Pi^+_{r_+}(G)$ is an orthogonal projector of rank $k - r_+$. Its range is orthogonal to $q_G$, and $\|q_G q_G^\top\|_F^2 = 1$. Therefore

$$\begin{aligned}
&\left\|\Pi^+_k(G) - \Pi^+_{r_+}(G) - \alpha q_G q_G^\top\right\|_F^2 \\
&\quad = \left\|\Pi^+_k(G) - \Pi^+_{r_+}(G)\right\|_F^2 + \alpha^2 \left\|q_G q_G^\top\right\|_F^2 \\
&\quad = k - r_+ + \alpha^2.
\end{aligned}$$

Continuity of the norm and division by $k$ prove Equation 11. Since $r_+ \le k - 1$, the limit is strictly positive. This completes the proof of Theorem 2.

### A.3 An explicit four-node example

Take $G = H = K_2$, with each internal edge having unit weight, and join one endpoint from each graph by a bridge of weight $\varepsilon > 0$. The resulting graph is a four-node weighted path. Its normalized-Laplacian spectrum is

$$0, \qquad \frac{\varepsilon}{1+\varepsilon}, \qquad \frac{2+\varepsilon}{1+\varepsilon}, \qquad 2.$$

Thus its smallest positive eigenvalue is exactly

$$\lambda_{\mathrm{br}}(\varepsilon) = \frac{\varepsilon}{1+\varepsilon}. \tag{17}$$

For $k = 1$, the pre-bridge quantities are $v_G = v_H = 2$, $\alpha = 1/2$, and $r_+ = 0$. Theorem 2 gives

$$\lim_{\varepsilon \to 0^+} \Delta_1(G, J_\varepsilon) = \frac{\sqrt{5}}{2}. \tag{18}$$

The bridge eigenvalue therefore vanishes while the discrepancy from the original representation remains positive. A relative discrepancy exceeding one is permissible: the expanded projector restricted to $G$ is generally not an orthogonal projector of rank $k$.

### A.4 Proofs of consistency-preserving constructions

**Fixed spectral matrix functions.** Let $f : [0, 2] \to \mathbb{R}$ be fixed. Using full orthogonal eigendecompositions $L_G = U_G \Lambda_G U_G^\top$ and $L_H = U_H \Lambda_H U_H^\top$, define the matrix function by applying $f$ to the eigenvalues. Then

$$\begin{aligned}
f(L_G \oplus L_H) &= (U_G \oplus U_H)\big(f(\Lambda_G) \oplus f(\Lambda_H)\big)(U_G \oplus U_H)^\top \\
&= f(L_G) \oplus f(L_H).
\end{aligned}$$

Restriction to the $G$ block proves Equation 12. This finite-dimensional direct-sum property does not require continuity of $f$ (Higham, 2008).

**Heat-kernel recovery.** For fixed $t > 0$, the matrix-exponential difference satisfies

$$e^{-tL_\varepsilon} - e^{-tL_0} = -\int_0^t e^{-(t-s)L_\varepsilon}(L_\varepsilon - L_0)e^{-sL_0}\, ds.$$

Both Laplacians are symmetric positive semidefinite, so each exponential in the integrand has operator norm at most one. Consequently,

$$\left\|e^{-tL_\varepsilon} - e^{-tL_0}\right\|_2 \le t\|L_\varepsilon - L_0\|_2.$$

Since $\|R_G\|_2 = 1$ and $R_G e^{-tL_0} R_G^\top = e^{-tL_G}$, this proves Equation 13. The regularity established above gives $\|L_\varepsilon - L_0\|_2 = O(\varepsilon)$. For fixed $n_G$, the Frobenius discrepancy is at most $\sqrt{n_G}$ times the operator-norm discrepancy, and therefore also tends to zero.

**Finite-step random-walk returns.** Write $M_\varepsilon = D_\varepsilon^{-1} A_\varepsilon$ and $M_0 = M_G \oplus M_H$. Under disjoint extension,

$$M_0^j = M_G^j \oplus M_H^j,$$

so every old-node return probability is unchanged.

For a positive bridge, let $d_a, d_b$ be the pre-bridge endpoint degrees and let $\|\cdot\|_\infty$ denote the induced maximum absolute row-sum norm. Only the two endpoint rows change, giving

$$\|M_\varepsilon - M_0\|_\infty = 2\varepsilon \max\left\{\frac{1}{d_a + \varepsilon}, \frac{1}{d_b + \varepsilon}\right\} = O(\varepsilon).$$

For every positive integer $j$,

$$M_\varepsilon^j - M_0^j = \sum_{\ell=0}^{j-1} M_\varepsilon^\ell (M_\varepsilon - M_0) M_0^{j-1-\ell}.$$

Both transition matrices are row-stochastic, so their powers have induced infinity norm one. It follows that

$$\|M_\varepsilon^j - M_0^j\|_\infty \le j\|M_\varepsilon - M_0\|_\infty.$$

In particular,

$$\max_{\substack{v \in V_G \\ 1 \le j \le T}} \left|(M_\varepsilon^j)_{vv} - (M_G^j)_{vv}\right| \le T\|M_\varepsilon - M_0\|_\infty = O(T\varepsilon).$$

This proves exact preservation under disjoint extension and coordinatewise recovery along the weak-bridge path for fixed $T$.

**Source-preserving controls.** Independent encoding of unchanged connected components removes competition between their separate spectral budgets. A positive bridge instead joins the components, so this argument does not apply to decomposition by current connected components.

More generally, fix a deterministic predictor and its inference state. If every structural input required by a protected old-node computation is retained and all remaining computational dependencies are unchanged, the computation receives identical inputs before and after extension and therefore returns the same output. This statement concerns preservation of the specified old-node computations; it does not establish the accuracy of predictions on newly added nodes.

## B Comparison Conventions and Protected Computations

This appendix expands the comparison conventions in Section 2 and the computational safeguards used by the audit in Section 4.

### B.1 Persistent-node correspondence and comparison objects

The extension preserves the old induced graph, its edge weights, and the original raw features:

$$G'[V] = G, \qquad X'|_V = X. \tag{19}$$

These conditions do not require old total degrees to remain unchanged when edges to new nodes are added. In particular, a positive bridge changes its endpoint degrees, which can affect normalized propagation.

We order old nodes before new nodes and use

$$R_G = \begin{bmatrix} I_{|V|} & 0 \end{bmatrix}. \tag{20}$$

For a node-indexed matrix, left multiplication by $R_G$ selects the original nodes. For a node-pair relation, restriction acts on both indices: $R_G K_\Phi(G') R_G^\top$. Thus, encoding discrepancies compare the relation on all original nodes, whereas prediction discrepancies are evaluated on the declared protected subset $\mathcal{P}$.

For a spectral subspace with orthonormal basis $U_G$, the comparison object is

$$K_\Phi(G) = U_G U_G^\top, \qquad U_G^\top U_G = I. \tag{21}$$

For any orthogonal matrix $Q$ of the corresponding dimension,

$$(U_G Q)(U_G Q)^\top = U_G U_G^\top.$$

The projector therefore identifies the selected subspace without counting sign or basis changes as drift. Section 3 specifies the spectral gaps that make the selected projectors unique.

For general learned features, $Z_G Z_G^\top$ is a Gram relation and need not be an orthogonal projector. Its suitability as a comparison object depends on the encoding's declared symmetries. Moreover, equality of Gram relations or projectors alone does not imply equality of predictions when the predictor uses coordinates or additional structural inputs that the relation does not determine. This is why Equations equation 2 and equation 3 measure distinct properties.

### B.2 Coupling randomness across paired evaluations

For a randomized encoder, denote its random inputs by $\xi$ and write its output as $\Phi_\theta(G, X; \xi)$. Comparing independently sampled encodings would mix extension-induced changes with sampling variation. Our paired comparison instead preserves the realized shared random inputs and the draws assigned to persistent nodes.

For node-indexed draws, the coupling satisfies

$$\xi_{G'}|_V = \xi_G,$$

with additional draws generated for new nodes. Any shared random inputs retain the same realization in both evaluations. The R-PEARL audit uses this old-node coupling (Kanatsoulis et al., 2025). Along a weak-bridge path, the random realization is held fixed as the bridge weight varies.

These comparisons are conditional on the declared coupling. Using the same seed alone does not specify that coupling: the realized inputs must remain aligned with the same persistent nodes and shared computations.

### B.3 Protecting non-encoding computations

Protection is defined with respect to the computational dependencies of the audited predictor. An unchanged old induced graph alone does not guarantee unchanged downstream computation. We hold the learned parameters and inference state fixed, and isolate the non-encoding dependencies relevant to the evaluated nodes. Computations inside the structural encoder remain part of the encoding pathway being audited.

**Message passing.** For a finite-depth message-passing predictor, protection covers the raw features, edges, propagation coefficients, and normalization factors used by the evaluated computations. Preserving neighbor identities alone is insufficient when propagation coefficients depend on degrees that change under extension.

In the two-layer GCN weak-bridge audit, the evaluated old test nodes are at least three hops from the old bridge endpoint. The relevant normalized propagation operators are compared with those on the original graph. The distance condition excludes the changed endpoint from their two-layer computational neighborhoods; the operator comparison checks the propagation coefficients used within those neighborhoods. The protected set is held fixed across bridge weights.

**Global attention.** Disconnected graphs can still interact through global attention. The WIRE audit therefore uses a source-block mask that excludes attention between the original $G$ and $H$ groups (Reid et al., 2026). The partition remains fixed under both disconnected and weak-bridge extensions.

This mask fixes the permitted interaction domain. It does not fix the attention scores or weights within $G$: those quantities may change when the spectral coordinates are recomputed. Such within-source changes remain part of the encoding-dependent effect under audit.

**Other graph-dependent operations.** Pooling, normalization, or other operations outside the structural encoder can introduce additional dependencies on newly added nodes. A protected comparison requires those non-encoding dependencies to remain unchanged. Graph-dependent operations within the encoder are instead included in the recomputation pathway and examined through the method-specific interventions in Section 4.

### B.4 Task conditions and the interpretation of consistency

The consistency requirements apply to extensions declared to preserve the evaluated targets and to provide no new task-relevant evidence about them. Disconnectedness alone does not establish this task condition: targets defined by global rank, population composition, or total graph size may legitimately change after an extension.

DEC and WBES are defined separately at the encoding and prediction levels. An encoding-family guarantee ranges over the graph domain specified in its statement. A prediction-level guarantee additionally depends on the frozen predictor, its inference state, the protected nodes, and any declared randomness coupling. Observed agreement on a finite collection of audit cases does not by itself establish a family-wide guarantee.

For WBES, the graph pair and bridge endpoints remain fixed as $\varepsilon \to 0^+$. Prediction comparisons also retain the same protected set. The reference remains the original graph $G$ throughout the path; convergence to the disconnected expanded graph $G \sqcup H$ alone does not establish recovery relative to $G$.

### B.5 Source-preserving and cached controls

All paired controls retain the original predictor parameters, learned encoder parameters, and inference state. Their predictions are compared with the same model's output on the original graph.

Source-preserving recomputation applies the frozen encoder separately to the original $G$ and $H$ groups and combines their structural inputs for the expanded downstream computation. For node-wise encodings, this corresponds to concatenating the two groups' encodings. Under a positive bridge, the groups are identified by their original provenance, not by connected components of the expanded graph. Randomized encoders use the coupling of Appendix B.2.

The cached-encoding control directly reuses the original structural inputs required by protected computations. Both controls preserve the relevant old inputs while retaining the audit's downstream protection. Their interpretation concerns the evaluated old nodes, without a guarantee about new-node prediction quality.

Alternative encoders and no-encoding references use separately trained models. They establish how other pipelines behave under the same extension conditions, rather than serving as within-model interventions.

### B.6 Encoding-pathway interventions

The following diagnostics examine different parts of the recomputation pathway. They reuse frozen model parameters and do not assume that probability changes from different interventions add linearly.

**SPE.** For a spectral kernel channel, write

$$W_\ell = U \operatorname{diag}\big(\psi_\ell(\lambda)\big) U^\top.$$

The spectral-input replacement intervention substitutes the expanded kernel's old–old block while retaining the original node domain and graph computations. The domain-enlargement intervention instead restores the original old–old kernel block, retains the expanded new–new block, and sets cross-source blocks to zero, while evaluating the enlarged encoder domain. Joint recomputation uses the ordinary expanded inputs.

In the Brazil disjoint audit, spectral-input replacement gives median probability drift $4.25 \times 10^{-5}$. Domain enlargement gives 0.531466, compared with 0.531467 under joint recomputation, and reproduces all corresponding class flips. The enlarged domain adds zero-input columns to the old-node computation; nonzero encoder responses to those columns accumulate. The disjoint comparisons evaluate all old test nodes.

**BasisNet.** The fixed-domain intervention replaces the spectral-input package, including eigenvalues and old-node blocks of grouped spectral projectors, while retaining the original node domain and graph computations. The resulting median probability drift is $0.0887$ for Brazil $\leftarrow$ Europe. This intervention isolates sensitivity to the spectral-input package while holding the computational domain and downstream graph computations fixed.

**LLPE.** The full-spectrum diagnostic first controls eigenvector signs and exact repeated-eigenspace bases relative to component bases. It fixes the spectral mapping domain, using the normalized Laplacian map $\lambda \mapsto \lambda - 1$, and removes graph-level min–max preprocessing of eigenvectors by using the raw orthonormal basis.

All diagnostic changes are applied on both the original and extended graphs. Each pathway is compared against its own original-graph output under the same frozen model parameters. Thus, near-zero diagnostic drift does not assert equality with the author's original predictions or preservation of their task accuracy. Because several operations change together, the combined control does not provide a single-factor attribution.

The disjoint LLPE summary contains 150 intervention records: five directed graph pairs, ten training splits, and three bridge seeds. Even at zero bridge weight, evaluation uses the old test nodes at distance at least three from the preselected old endpoint. Different endpoint selections can therefore yield different evaluated subsets on the same disconnected expanded graph. These records are repeated evaluations, not 150 independent trained models.

# C Experimental Details and Additional Results

## C.1 Data, training, and audit configurations

**Datasets and fixed inputs.** Airports uses four activity classes, retaining the original regional label definitions after extension. For each original Airports graph, the six input features are a constant, $\log(1 + \text{degree})$, square-root degree, clustering coefficient, and the mean and standard deviation of neighbor degrees. Nonconstant columns are standardized on the original graph. These precomputed old-node features are retained after extension.

Airports splits allocate approximately 60%, 20%, and the remainder of each class to training, validation, and testing, with rounding performed within each class. Ten splits use split-generation seeds $20260805 + s$, $s = 0, \ldots, 9$. WebKB uses the release and splits of Pei et al. (2020), with five webpage classes.

**Main GCN comparison.** The main Airports models use two GCN layers, hidden width 64, dropout 0.35, and spectral budget $k = 16$ for BasisNet and SPE. Training uses Adam with learning rate 0.003 and weight decay $10^{-5}$, for at most 200 epochs. The main GCN training procedure uses patience 40 after a minimum of 100 epochs. Checkpoint selection maximizes validation accuracy, breaking ties by validation loss. Test labels are not used for checkpoint selection.

The weak-bridge audit uses both Airports directions, ten splits, three bridge seeds, and $\varepsilon \in \{1, 10^{-3}, 10^{-5}\}$. Protected nodes and propagation checks follow Appendix B.3.

**R-PEARL.** The configuration-matched reference is a local implementation following the paper and official design (Kanatsoulis et al., 2025). Its encoder uses eight GIN layers, three-layer MLPs, encoder hidden width 40, and output dimension 37. It uses 120 random samples during both training and evaluation. The downstream model has hidden width 64 and dropout 0.35; the learning rate is 0.003 and weight decay is $10^{-5}$, with a 200-epoch maximum and patience 40. Positive bridges use weighted messages. Three Monte Carlo banks are evaluated for each bridge location, with shared and old-node draws coupled across the paired inputs. The main table uses the bridge- and configuration-matched formal results.

**WIRE.** The controlled attention model uses two layers, hidden width 64, four heads, dropout 0.35, learning rate 0.001, and weight decay $10^{-5}$, with a 200-epoch maximum and patience 40. It uses normalized-Laplacian coordinates with budgets $m \in \{8, 16, 32\}$, including zero modes. The five directions are Brazil ← Europe, Europe ← Brazil, Cornell ← Texas, Texas ← Wisconsin, and Wisconsin ← Cornell. Each uses ten splits, three bridge seeds, and $\varepsilon \in \{0, 10^{-5}, 10^{-3}, 1\}$. The source-block mask is retained throughout.

**LLPE.** The full-spectrum audit uses the same five directions, ten splits, three bridge seeds, and four bridge weights as WIRE. Its configuration uses two GCN layers, hidden width 64, positional output dimension 16, Chebyshev order 64, dropout 0.35, learning rate 0.005, weight decay $10^{-5}$, and spectral regularization coefficient $10^{-4}$. Training has a 200-epoch maximum and patience 40. The disjoint summary in Table 3 uses only $\varepsilon = 0$ records. The diagnostic construction and its reference outputs are specified in Appendix B.6.

### C.2 Metrics and statistical aggregation

Probability drift is the mean $L_1$ distance between paired class-probability vectors on the evaluated nodes. Flip rate is the fraction whose predicted class changes, using fixed tie-breaking. Original and extended accuracies are computed against the same labels on the same evaluated nodes.

For Table 1, bridge locations are first aggregated within each training split by their median. R-PEARL additionally aggregates its Monte Carlo banks in the same within-split reduction. There are three within-split records for each main GCN method and nine for R-PEARL at a fixed bridge weight. The reported estimate is the median of the ten split-level estimates.

Descriptive 95% intervals use 10,000 bootstrap resamples of these ten split-level estimates and their 2.5th and 97.5th percentiles. The table-generation procedure uses base seed 20260810. The resampling unit is the training split; bridge locations and Monte Carlo banks are not treated as independent training replicates. These intervals are descriptive and are not multiplicity-corrected significance tests.

WIRE family summaries instead report medians over the intervention records within each family. The LLPE disjoint summary reports the median over its 150 intervention records. These pooling conventions differ from the two-stage aggregation of the main frozen-prediction table and should not be interpreted as identical statistical estimands.

### C.3 Main frozen-prediction intervals and accuracies

Table 4 gives the descriptive intervals corresponding to Table 1. All values use protected old test nodes at $\varepsilon = 10^{-5}$. Source-preserving BasisNet and SPE controls have zero probability drift and flip rate, with intervals $[0, 0]$.

Table 5 reports accuracies for the same records and aggregation rule. The number of evaluated protected test nodes per intervention ranges from 8 to 20 for Brazil and from 15 to 68 for Europe, with identical evaluated sets across the matched methods. Original and extended columns are separately aggregated medians; their difference is not the median paired accuracy change.

### C.4 WIRE coverage and implementation checks

The WIRE design contains $5 \times 3 \times 4 \times 10 \times 3 = 1{,}800$ intervention records across graph directions, spectral budgets, bridge weights, splits, and bridge locations. There are 720 Airports and 1,080

Table 4: Main frozen-prediction estimates with descriptive 95% split-bootstrap intervals. Flip rates and their interval endpoints are percentages. The aggregation follows Appendix C.2.

| Method / handling | Probability $L_1$ drift | Flip rate (%) |
|---|---|---|
| *Brazil ← Europe* | | |
| BasisNet (joint) | $0.0588$<br>$[0.0337,\ 0.0753]$ | $17.16$<br>$[8.33,\ 23.08]$ |
| Source-preserving | $0$<br>$[0,\ 0]$ | $0$<br>$[0,\ 0]$ |
| SPE (joint) | $0.4874$<br>$[0.3988,\ 0.5658]$ | $0.00$<br>$[0.00,\ 14.17]$ |
| Source-preserving | $0$<br>$[0,\ 0]$ | $0$<br>$[0,\ 0]$ |
| R-PEARL | $4.58 \times 10^{-7}$<br>$[1.13 \times 10^{-7},\ 1.78 \times 10^{-6}]$ | $0$<br>$[0,\ 0]$ |
| RWSE | $8.62 \times 10^{-9}$<br>$[2.19 \times 10^{-9},\ 1.24 \times 10^{-8}]$ | $0$<br>$[0,\ 0]$ |
| No PE | $0$<br>$[0,\ 0]$ | $0$<br>$[0,\ 0]$ |
| *Europe ← Brazil* | | |
| BasisNet (joint) | $0.0109$<br>$[0.0073,\ 0.0326]$ | $3.59$<br>$[3.06,\ 13.65]$ |
| Source-preserving | $0$<br>$[0,\ 0]$ | $0$<br>$[0,\ 0]$ |
| SPE (joint) | $0.0908$<br>$[0.0714,\ 0.1127]$ | $9.92$<br>$[4.26,\ 12.50]$ |
| Source-preserving | $0$<br>$[0,\ 0]$ | $0$<br>$[0,\ 0]$ |
| R-PEARL | $1.73 \times 10^{-8}$<br>$[4.97 \times 10^{-9},\ 2.73 \times 10^{-8}]$ | $0$<br>$[0,\ 0]$ |
| RWSE | $0$<br>$[0,\ 0]$ | $0$<br>$[0,\ 0]$ |
| No PE | $0$<br>$[0,\ 0]$ | $0$<br>$[0,\ 0]$ |

Table 5: Accuracy (%) on the protected nodes used in the main comparison at $\varepsilon = 10^{-5}$. Each column uses within-split medians followed by the median across ten splits. Source-preserving BasisNet and SPE controls retain their original predictions and corresponding accuracies.

| | Brazil ← Europe | | Europe ← Brazil | |
|---|---|---|---|---|
| Method | Original | Extended | Original | Extended |
| BasisNet | 55.00 | 60.00 | 42.23 | 43.54 |
| SPE | 54.20 | 50.00 | 46.77 | 42.82 |
| R-PEARL | 50.00 | 50.00 | 42.50 | 42.50 |
| RWSE | 61.25 | 61.25 | 45.60 | 45.60 |
| No PE | 63.07 | 63.07 | 45.63 | 45.63 |

WebKB records. These counts represent repeated interventions on trained models, not independent training runs.

In the recorded parity check, the audited attention layer matches the official layer's outputs and pre-softmax scores with maximum absolute error zero. The recorded local checks have zero cross-source attention and zero repeated-evaluation discrepancy. The maximum permutation-equivariance discrepancy is $5.96 \times 10^{-8}$. These checks validate the audited layer and masking setup; the experiment is a controlled wrapper rather than a reproduction of the complete original benchmark.

In the formal audit, cross-source attention remains zero. The maximum probability drift in both the source-preserving and no-WIRE controls is $4.31 \times 10^{-7}$. Joint WIRE recomputation changes

old-to-old scores in all 1,800 records, but only 16 records contain any class flip. The median flip rate is zero within every graph-pair, budget, and bridge-weight condition.

### C.5 Additional bridge-weight prediction comparisons

The representation sweep in Section 5.3 uses eleven positive weights from 1 to $10^{-5}$ and budgets $k \in \{8, 16, 32\}$. The separate frozen-prediction comparison evaluates $\varepsilon \in \{1, 10^{-3}, 10^{-5}\}$.

The predictor responses are heterogeneous. For Europe $\leftarrow$ Brazil, the median BasisNet probability drift remains approximately $0.01086$ as the bridge weakens. For Brazil $\leftarrow$ Europe, R-PEARL's corresponding medians decrease from $5.84 \times 10^{-3}$ to $4.45 \times 10^{-5}$ and then $4.58 \times 10^{-7}$.

These measurements distinguish non-recovery of the selected spectral representation from the response of a particular trained predictor. The finite bridge-weight grid does not independently prove a limit for prediction drift.

### C.6 Changing the external spectral load

The paired probability-drift differences are positive in all ten splits, with median $0.0238$ and descriptive 95% interval $[0.0144, 0.0717]$. This ordering is consistent with spectral competition in the BasisNet pipeline. However, external graph size, density, and spectrum change jointly, so the comparison does not isolate graph size or spectral load as a single causal factor.