

Hamming Graph Metrics: A Multi-Scale Framework for Structural Redundancy and Uniqueness in Graphs

R. Scott Johnson

Abstract

Traditional graph centrality measures effectively quantify node importance but fail to capture the structural uniqueness of multi-scale connectivity patterns—critical for understanding network resilience and function. This paper introduces *Hamming Graph Metrics (HGM)*, a framework that represents a graph by its exact- k reachability tensor $\mathcal{B}_G \in \{0, 1\}^{N \times N \times D}$ with slices $(\mathcal{B}_G)_{::,1} = A$ and, for $k \geq 2$, $(\mathcal{B}_G)_{::,k} = \mathbf{1}[\sum_{t=1}^k A^t > 0] - \mathbf{1}[\sum_{t=1}^{k-1} A^t > 0]$ (shortest-path distance exactly k).

Guarantees. (i) *Permutation invariance*: $d_{\text{HGM}}(\pi(G), \pi(H)) = d_{\text{HGM}}(G, H)$ for all vertex relabelings π ; (ii) the *tensor Hamming distance*

$$d_{\text{HGM}}(G, H) := \|\mathcal{B}_G - \mathcal{B}_H\|_1 = \sum_{i,j,k} \mathbf{1}[(\mathcal{B}_G)_{ijk} \neq (\mathcal{B}_H)_{ijk}]$$

is a *true metric* on labeled graphs; and (iii) *Lipschitz stability* to edge perturbations with explicit degree-dependent constants (see “Graph-to-Graph Comparison” $\rightarrow$ “Tensor Hamming metric”; “Stability to edge perturbations”; Appendix A). For unlabeled graph comparison, one can apply HGM after graph canonization, or use an alignment-based variant (exponential worst-case cost).

We develop: (1) *per-scale spectral analysis* via classical MDS on double-centered Hamming matrices $D^{(k)}$, yielding spectral coordinates and explained variances; (2) *summary statistics* for node-wise and graph-level structural dissimilarity; (3) *graph-to-graph comparison* via the metric above; and (4) *analytic properties* including extremal characterizations, multi-scale limits, and stability bounds.

1 Introduction

1.1 The Research Gap

Graph centrality measures are fundamental tools for understanding network structure and identifying influential nodes across numerous domains including social networks, biological pathways, transportation infrastructure, and communication systems [5]. Traditional measures such as degree centrality, closeness centrality, and betweenness centrality have been extensively studied and applied for decades [1, 2]. These classical metrics typically emphasize frequency, reachability, and efficiency of traversal within a network. Specifically, betweenness centrality quantifies how often a node participates in shortest paths between other nodes [3], closeness centrality measures how quickly a node can reach other nodes [4], and degree centrality simply counts how many direct connections a node possesses [5].

However, these conventional measures fail to capture an important aspect of network structure: the structural diversity or redundancy of connectivity patterns. This limitation is particularly significant when analyzing complex networks where understanding the uniqueness of connection patterns is crucial. Local measures like degree centrality provide valuable information about immediate connections but offer limited insight into how these connections contribute to global structural patterns. Even path-based measures like betweenness centrality, while considering global connectivity, primarily quantify path frequency rather than structural uniqueness.

The structural uniqueness of connectivity patterns represents a fundamental property of networks that has remained largely unexplored. Two nodes with identical betweenness or closeness centrality values may differ substantially in how their connections are structured. One node might connect disparate regions of the network through unique paths that, if removed, would significantly alter the network’s topology. In contrast, another node with the same centrality values might have highly redundant paths that could be easily substituted if the node were removed. Traditional centrality measures cannot distinguish between these scenarios despite their differing implications for network resilience, information flow, and functional organization.

Many practical applications require a more nuanced structural fingerprint that can distinguish between nodes whose paths are structurally redundant and those whose paths offer unique connectivity patterns. For instance, in resilience analysis, nodes with structurally unique connectivity patterns may represent critical failure points [6, 7], while in anomaly detection [8], unusual path structures might signal deviations from expected network behavior

[7]. In communication networks, identifying nodes with diverse connectivity patterns can enhance routing strategies and improve network robustness [7]. These applications highlight the need for centrality measures that specifically quantify structural uniqueness and redundancy.

1.2 Contributions

This paper develops a rigorous mathematical framework for analyzing structural uniqueness in graphs, grounded in binary reachability patterns and their pairwise dissimilarities. Our key theoretical contributions are as follows:

1. **Hamming Graph Metrics Framework:** We define a comprehensive framework based on the distribution of Hamming distances between binary reachability vectors across all node pairs, capturing the complete spectrum of structural diversity within a graph.
2. **Multi-Scale Structural Profiles:** The framework decomposes connectivity into a spectrum of exact path lengths, with each scale k analyzed independently to reveal patterns invisible when distances are aggregated.
3. **Tensor Formulation and Properties:** We develop a family of convex functionals on binary dissimilarity distributions, including entropies, ℓ_1 /total-variation divergences, and spectral descriptors, enabling rich geometric analysis without transport distances.
4. **Graph-Level Aggregates:** We define dispersion via deviation from the mean profile (TV/ℓ_1) and entropy-based summaries, enabling structural comparisons across graphs.
5. **Theoretical Guarantees:** New theorems are proved for extremal bounds, monotonicity, and structural separation in canonical graph classes (complete, star, ring, regular, Erdős–Rényi, scale-free).
6. **Comparative Geometry of Graphs:** Hamming distributions offer a basis for comparing graphs structurally, independent of scale or density.
7. **Finite Sample Models:** We derive limiting behavior and finite-size approximations under synthetic conditions.

While the emphasis throughout is theoretical, we include a brief discussion of algorithmic strategies to compute Hamming Graph Metrics efficiently in Appendix B, and we show that the proposed measures can scale to real-world

networks with tens of thousands of nodes when implemented with bit-parallel operations. These observations indicate practical scalability and suggest future work in scalable approximation.

1.3 Preliminaries

1.3.1 Notational Conventions

We use superscript (k) to denote exact path length k , not cumulative distance. Thus $B^{(k)}$ indicates paths of length exactly k , and $\mathbf{b}_v^{(k)}$ is node v 's reachability vector at this specific distance.

Let $G = (V, E)$ be a finite, simple, undirected graph with vertex set V and edge set E , where $|V| = N$. Let $A \in \{0, 1\}^{N \times N}$ be the adjacency matrix of G , with entries [3]:

$$A_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E \\ 0 & \text{otherwise} \end{cases}$$

For any positive integer k , the matrix power A^k counts the number of walks of length k between nodes.

Exact vs. cumulative reachability. Let A be the adjacency matrix and $B_{\leq k} := \mathbf{1}[\sum_{t=1}^k A^t > 0]$ the cumulative reachability within k steps (element-wise). We use the *exact- k* convention throughout:

$$B^{(1)} = A, \quad B^{(k)} := B_{\leq k} - B_{\leq k-1} \quad (k \geq 2).$$

Thus the exact- k reachability tensor is $B^{(k)}$ for $k = 1, \dots, D$, with $\text{diag}(B^{(k)}) = 0$ for all k ; equivalently, $B_{ij}^{(k)} = \mathbf{1}\{\text{dist}(i, j) = k\}$. We henceforth *identify* the k -slice of the tensor with its matrix: $B^{(k)} \equiv B^{(k)}$.

1.3.2 Tensor Formulation and Cross-Scale Structure

We follow Kolda–Bader's tensor notation for unfoldings/matricization and mode products [10]; only the entries of $\mathcal{B}$ are nonnegative binary, while its unfoldings are real matrices used for spectral summaries.

While $\{B^{(k)}\}_{k=1}^D$ can be seen as a stack of matrices, the third index encodes *cross-scale constraints*: if (i, j) is reachable in exactly $k+1$ steps, then there exists ℓ with (i, ℓ) reachable in k steps and $(\ell, j) \in E$. These inter-slice implications (and their converses failing in general) make $\mathcal{B}$ a genuinely

third-order object. HGM measures geometry per scale and aggregates across k without collapsing them.

For each node v , define its exact- k reachability vector as the v -th row of $B^{(k)}$:

$$\mathbf{b}_v^{(k)} = (B_{v1}^{(k)}, B_{v2}^{(k)}, \dots, B_{vN}^{(k)}) \in \{0, 1\}^N, \quad \text{with } \text{diag}(B^{(k)}) = \mathbf{0}.$$

The Hamming distance between binary vectors $x, y \in \{0, 1\}^N$ is

$$\text{Ham}(x, y) = \sum_{i=1}^N |x_i - y_i| = \|x - y\|_1,$$

which counts the number of positions at which x and y differ. *Equivalently (since the vectors are binary), it equals the Hamming weight of $x \oplus y$.*

Additional notation.

- $\mu_v^{(k)}$: the empirical distribution of pairwise distances at scale k for node v ,

$$\mu_v^{(k)} = \frac{1}{N-1} \sum_{u \neq v} \delta_{\text{Ham}(\mathbf{b}_v^{(k)}, \mathbf{b}_u^{(k)})}.$$

- Higher moments at scale k (for node v) are taken with respect to $\mu_v^{(k)}$; we write variance $\sigma^2(\mu_v^{(k)})$, skewness $\gamma(\mu_v^{(k)})$, and kurtosis $\kappa(\mu_v^{(k)})$.
- The diameter D is the smallest integer such that $B_{\leq D} = \mathbf{1}_{N \times N} - I$ (in the connected case).

Unless otherwise stated, we assume G is connected.

Tensor Representation and Multi-Scale Hamming Distance We work with the exact- k reachability tensor $\mathcal{B} \in \{0, 1\}^{N \times N \times D}$ (Sec. 2.1). For a node i , the slice $\mathcal{B}[i, :, :] \in \{0, 1\}^{N \times D}$ stacks its per-scale neighborhoods. All symbols are summarized in the Notation Reference Table below.

Table 1: Notation Reference

Symbol	Type	Definition	First Use
Graph Structure			
$G = (V, E)$	Graph	Undirected simple graph with vertex set V and edge set E	§1.3

Symbol	Type	Definition	First Use
$N = V $	Integer	Number of vertices	§1.3
A	$N \times N$ matrix	Adjacency matrix; $A_{ij} = 1$ iff $(i, j) \in E$	§1.3
$\text{diam}(G)$	Integer	Graph diameter	§1.3
Reachability			
A^k	$N \times N$ matrix	k -walk matrix (counts walks of length k)	§1.3
$B_{\leq k}$	$N \times N$ binary matrix	Cumulative reachability: $\mathbf{1}[A^k > 0]$	§1.3
$B^{(k)}$	$N \times N$ binary matrix	Exact-k reachability: $B^{(k)} = B_{\leq k} - B_{\leq k-1}$ ($k \geq 2$, $B^{(1)} = A$)	§1.3.2
$\mathbf{b}_v^{(k)}$	Vector in $\{0, 1\}^N$	Row v of $B^{(k)}$ (exact- k reachability pattern)	§1.3
Tensors			
$\mathcal{A}$	$N \times N \times K$ tensor	k -walk <i>count</i> tensor: $\mathcal{A}(:, :, k) = A^k$ (integer entries)	§1.3.2
$\mathcal{B}$	$N \times N \times D$ tensor	Exact- k reachability (binary); $(i, j, k) = 1$ iff $\text{dist}(i, j) = k$; diagonal 0	§1.3.2
$\bar{\mathcal{B}}$	$N \times N \times D$ tensor	Mean reachability slab: $\bar{\mathcal{B}} = \frac{1}{N} \sum_{u=1}^N \mathcal{B}[u, :, :]$	§2.4
Distances & Centrality			
$\text{Ham}(x, y)$	Integer	Hamming distance between binary vectors	§1.3
$H(v, u)$	Integer	Tensorial Hamming distance (sum over k of per-scale Hamming)	§1.3.2
$\text{HC}^{(k)}(v)$	Real	Hamming centrality of node v at scale k	§2
$\text{HC}(v)$	Real	Multi-scale Hamming centrality (uniform average over k)	§2
$\text{HC}_{\text{tensor}}(v)$	Real	Tensor-based HC: $\ \mathcal{B}[v, :, :] - \bar{\mathcal{B}}\ _*$	§2.4
Distributions			

Symbol	Type	Definition	First Use
$\mu_v^{(k)}$	Probability mass function	Distribution of $\text{Ham}(\mathbf{b}_v^{(k)}, \mathbf{b}_u^{(k)})$ over $u \neq v$	§3.1
$\mu_G^{(k)}$	Probability mass function	Graph-level distance distribution at scale k (over unordered pairs)	§3.1
$D_v^{(k)}$	Multiset	$\{ H(v, u) : u \in V, u \neq v \}$	§3.1
$\mathcal{D}_G^{(k)}$	Multiset	All pairwise distances at scale k (ordered or unordered, as specified)	§3.1
Functionals			
Φ	Functional	Admissible functional on distributions (Def. 3.5)	§3.5
$\Psi^{(k)}(G)$	Real	TV-dispersion: $\frac{1}{N} \sum_v \ \mu_v^{(k)} - \bar{\mu}^{(k)}\ _1$	§3.5
$\Xi^{(k)}(G)$	Real	Information-divergence dispersion (e.g., Rényi/KL variants)	§3.5
Temporal Extension of HGM			
$\{G_t\}_{t=1}^T$	Sequence of graphs	Temporal snapshots on a fixed vertex set $[N]$	§5.1.1
$A^{(t)}$	$N \times N$ matrix	Adjacency at time t	§5.1.1
$B^{(k,t)}$	$N \times N$ binary matrix	Exact- k reachability at time t	§5.1.1
$\mathbb{B}$	$N \times N \times D \times T$ tensor	Temporal HGM tensor: $\mathbb{B}_{ijkt} = B_{ij}^{(k,t)}$	§5.1.1
d_{dyn}	Real	Labeled temporal metric $\ \mathbb{B}^G - \mathbb{B}^H\ _1$	§5.1.1
$d_{\text{dyn,iso}}$	Real	Orbit metric $\min_{\pi} \ \mathbb{B}^G - (\pi \cdot \mathbb{B}^H)\ _1$	§5.1.1
$E_k(t)$	Integer	Per-time per-scale energy $\ B^{(k,t)}\ _F^2$	§5.1.1

Notational Conventions:

- Superscript (k) denotes exact path length k , not cumulative distance
- Bold lowercase ($\mathbf{b}$) denotes vectors
- Roman uppercase ($\mathbf{B}$) denotes matrices
- Calligraphic ($\mathcal{B}$) denotes tensors

- $\mathbf{1}_{N \times N}$ is the all-ones matrix
- δ_d denotes point mass at d

1.3.3 Full tensor-based Hamming distance

For any two nodes (i, j) define the **integer tensorial Hamming distance**

$$H(i, j) = \sum_{k=1}^D \text{Ham}(\mathcal{B}[i, :, k], \mathcal{B}[j, :, k]) \in \{0, 1, \dots, ND\}.$$

Equivalently, since the inputs are binary, $H(i, j) = \sum_{k=1}^D \|\mathcal{B}[i, :, k] - \mathcal{B}[j, :, k]\|_1$.

A **normalized** variant,

$$\bar{H}(i, j) = \frac{1}{D} \sum_{k=1}^D \text{Ham}(\mathcal{B}[i, :, k], \mathcal{B}[j, :, k]) \in [0, N],$$

is convenient for scale-invariant plots; all distributional results can be stated for H (integer support) or for $\bar{H}$ (rescaled).

Two useful settings:

- **Unweighted sum** (default): treats every scale equally in the integer H .
- **Geometric down-weighting**: use $\bar{H}_\alpha(i, j) = \frac{\sum_{k=1}^D \alpha^{k-1} \text{Ham}(\cdot)}{\sum_{k=1}^D \alpha^{k-1}}$ with $0 < \alpha < 1$ to emphasise shorter paths.

1.3.4 Cross-scale distance tensor

To capture interactions across *different* scales we introduce the fourth-order tensor

$$\boxed{\mathcal{D}_{i,j,k,\ell} = \text{Ham}(\mathcal{B}[i, :, k], \mathcal{B}[j, :, \ell])} \in \mathbb{N}^{N \times N \times D \times D}.$$

$\mathcal{D}$ stores **every pairwise cross-scale discrepancy** in a single object and underpins the graph-to-graph metrics developed in §4.6 below.

1.4 Path Reachability and Structural Patterns

The binary reachability matrix $B^{(k)}$ encodes fundamental structural information about the graph. Unlike the power A^k , which counts length- k walks, $B^{(k)}$ captures pure shortest-path reachability at *exact* distance k . Two nodes can

therefore have the same number of k -walks but different exact reachability patterns.

Consider the evolution as k increases:

- $k = 1$: $B^{(1)} = A$, immediate neighborhoods.
- $k = 2$: $B^{(2)}$ flags pairs at shortest-path distance exactly 2 (second-order neighborhoods).
- $1 < k < D$: intermediate scales reveal multi-scale organization.
- $k > D$ (connected G): $B^{(k)} \equiv 0$ by exact- k saturation (Lemma~.20).

The transition from local to global connectivity through intermediate scales $1 < k < D$ reveals the multi-scale organization of the graph.

Remark 1.1. By construction, $B^{(k)} = B_{\leq k} - B_{\leq k-1}$ with $B_{\leq k} = \mathbf{1}[\sum_{t=1}^k A^t > 0]$. In unweighted graphs, the walk-path reduction (Lemma~.15) justifies this *summed* form and, on bipartite graphs, enforces the usual parity constraint. Hence $B^{(k)}$ flags pairs at *exact* distance k , and the slices form a disjoint decomposition of off-diagonal connectivity:

$$\sum_{k=1}^D B^{(k)} = \mathbf{1}_{N \times N} - I, \quad B^{(k)} \equiv 0 \text{ for } k > D.$$

This multi-scale profile $\{B^{(1)}, \dots, B^{(D)}\}$ is the basis for per-scale analysis (e.g., the classical-MDS embedding of $D^{(k)}$) used later.

2 Hamming Centrality: Foundations and Properties

Proofs and pointers. Flagship results include full proofs in the main text; longer derivations and auxiliary lemmas are deferred to Appendix~.4.3 (with brief sketches inline). Computational details are in Appendix~A.

We begin by defining the foundational concept that motivates our broader framework: Hamming Centrality, a node-level index of structural distinctiveness based on binary path dissimilarity. While our primary focus is on graph-level distributions, understanding individual node contributions provides essential intuition for the comprehensive framework that follows.

2.1 Definition

The Hamming Graph Metrics framework treats path lengths as a spectrum of **distance layers** analyzed separately (not an eigen-spectrum). Whereas the

power A^k counts length- k walks, the binary slice $B^{(k)}$ encodes shortest-path reachability at exact distance k . Two nodes can have the same number of k -walks yet reach different node sets at exact distance k .

Let $G = (V, E)$ be a connected graph on N nodes, let $D := \text{diam}(G)$, and let $\mathbf{b}_v^{(k)} \in \{0, 1\}^N$ denote the exact- k reachability row of $B^{(k)}$ for node v (see §2).

Then the Hamming centrality of node v at layer k is

$$\text{HC}^{(k)}(v) = \frac{1}{N-1} \sum_{\substack{u \in V \\ u \neq v}} \text{Ham}(\mathbf{b}_v^{(k)}, \mathbf{b}_u^{(k)}),$$

the average number of reachability discrepancies between v and the rest of the graph at depth k .

Convention. We use the term *structural uniqueness* at scale k to denote this first moment, i.e. $\text{SU}^{(k)}(v) \equiv \text{HC}^{(k)}(v)$ (see Def.~2.1).

We define the multi-scale Hamming centrality as

$$\text{HC}(v) = \frac{1}{K} \sum_{k=1}^K \text{HC}^{(k)}(v), \quad 1 \leq K \leq D,$$

and, more generally, a weighted version

$$\text{HC}_w(v) = \sum_{k=1}^K w_k \text{HC}^{(k)}(v), \quad w_k \geq 0, \quad \sum_{k=1}^K w_k = 1,$$

to emphasize early or late scales when desired. In practice, $K = D$ yields a complete analysis of all slices $B^{(1)}, \dots, B^{(D)}$, whereas smaller K captures local uniqueness.

Definition 2.1 (Structural uniqueness (canonical choice).).

At scale k , the structural uniqueness of a node v is the first moment of its distance distribution:

$$\text{SU}^{(k)}(v) := \mathbb{E}_{u \neq v} [\text{Ham}(b_v^{(k)}, b_u^{(k)})] = \frac{1}{N-1} \sum_{u \neq v} \text{Ham}(b_v^{(k)}, b_u^{(k)}) = \text{HC}^{(k)}(v).$$

Graph-level uniqueness at scale k is any *admissible* functional (Def.~3.5) of $\{\mu_v^{(k)}\}_{v \in V}$, e.g.

$$\Psi^{(k)}(G) = \frac{1}{N} \sum_v \|\mu_v^{(k)} - \bar{\mu}^{(k)}\|_1, \quad \bar{\mu}^{(k)} = \frac{1}{N} \sum_v \mu_v^{(k)}.$$

Multi-scale uniqueness aggregates over k (uniformly or with weights).

2.2 Examples

Let us examine HC in basic graph topologies:

Complete graph K_N : At $k = 1$, all pairwise distances equal 2 so $\text{HC}^{(1)}(v) = 2$ for all v . For $k \geq 2$ (exact- k), $B^{(k)} \equiv 0$ and $\text{HC}^{(k)}(v) = 0$.

Star graph S_N : At $k = 1$, leaf-leaf vectors are identical (distance 0), while center-leaf pairs have distance N .

Ring graph C_N : By rotation invariance, $\text{HC}^{(k)}(v)$ is constant across v (yet nonzero); per-pair distances take a small set of even values.

Note that in each case, we examine patterns at exact distance k , not cumulative patterns up to distance k . This spectral separation is what allows us to detect structural features at specific scales.

2.3 Theoretical Properties

We now present several formal results characterizing Hamming Centrality.

Proposition 2.2 (Zero Centrality in Complete Graphs at Saturation). *Let $G = K_N$. Then:*

For $k = 1$: $\text{HC}^{(1)}(v) = 2$ for all $v \in V$ For all $k \geq 2$: $\text{HC}^{(k)}(v) = 0$ for all $v \in V$

Proof. For $k = 1$, in K_N we have $B^{(1)} = A$. Each node v has reachability vector $\mathbf{b}_v^{(1)}$ with 1s everywhere except position v . For any two nodes $v \neq u$, their vectors differ at exactly positions v and u , giving $\text{Ham}(\mathbf{b}_v^{(1)}, \mathbf{b}_u^{(1)}) = 2$. Thus $\text{HC}^{(1)}(v) = 2$. For $k \geq 2$, using the exact- k convention, $B^{(2)} = \mathbf{0}$ (and likewise for all higher k), so all pairwise distances at $k = 2$ are 0 and $\mu_G^{(2)} = \delta_0$. We say the slice “saturates” at scale k when $B^{(k)} = \mathbf{0}$, i.e., no pair has shortest-path distance exactly k . ■

Proposition 2.3 (Star graph asymmetry (exact- k at $k=1$)). *The star’s center c has $\text{HC}^{(1)}(c) = N$, while each leaf has $\text{HC}^{(1)}(\ell) = \frac{N}{N-1}$. Counting all ordered pairs shows that most leaf-leaf distances vanish, while the small fraction involving the center has distance N . This yields a distribution supported on $\{0, N\}$ with the weights derived in Appendix A.1.*

Proposition 2.4 (Upper Bound). *For all graphs G , nodes v , and any step k : $\text{HC}^{(k)}(v) \leq N$.*

Proof. The Hamming distance between any two binary vectors in $\{0, 1\}^N$ is at most N . No additional constraint forces a zero at the same coordinate

for both vectors, so the tight worst case is N (e.g., $K_{m,n}$ at $k = 1$ gives $\text{Ham} = m + n = N$ across parts). ■

2.3.1 Proof of Proposition .1

Proof. Let G be connected with diameter D . Then for all $k \geq D$ and all v ,

$$\text{HC}^{(k+1)}(v) \leq \text{HC}^{(k)}(v),$$

with equality for every $k \geq D + 1$ (both sides = 0). ■

2.4 Tensor-Based Hamming Centrality

Let

$$\bar{\mathcal{B}} := \frac{1}{N} \sum_{u=1}^N \mathcal{B}[u, :, :]$$

denote the **mean reachability slab**. The **tensor Hamming centrality** of a node v is

$$\boxed{\text{HC}_{\text{tensor}}(v) = \|\mathcal{B}[v, :, :] - \bar{\mathcal{B}}\|_*},$$

where $\|\cdot\|_*$ is any admissible tensor norm (Frobenius, weighted Hamming, or an $\ell_{2,1}$ mixed norm). The original slice-wise centrality $\text{HC}^{(k)}(v)$ is recovered by choosing $\|X\|_* = \text{Ham}(X_{:,k})$ and fixing k .

3 Hamming Graph Metrics: Tensor Formulation and Properties

Building on the node-level foundation (where $\text{SU}^{(k)}(v) = \text{HC}^{(k)}(v)$), we now pass to the graph level and the full family $\{\mu_G^{(k)}\}_{k=1}^D$, which strictly contains HC as the first-moment special case.

Soundness at a glance. For labeled graphs on $[N]$, the *tensor Hamming distance* $d_{\text{ten}}(G, H) = \|\mathcal{B}_G - \mathcal{B}_H\|_1$ is a true metric and is permutation-invariant: $d_{\text{ten}}(\pi(G), \pi(H)) = d_{\text{ten}}(G, H)$. A normalized form $\bar{d}_{\text{ten}} = \|\cdot\|_1 / (N(N-1)D) \in [0, 1]$ aids scaling. For unlabeled comparison, one may canonize graphs or use the alignment variant $d_{\text{iso}}([G], [H]) = \min_{\pi} \|\mathcal{B}_G - \mathcal{B}_{\pi(H)}\|_1$ (metric on isomorphism classes; exponential worst-case).

3.1 Graph-Level Distributions and Functionals

We now pass from node-wise distances to graph-level distributions. Let $G = (V, E)$ be a connected graph on $(N = |V|)$ vertices and recall the reachability tensor $(\mathcal{B} \in \{0, 1\}^{N \times N \times D})$ from §2.1.1. For any two nodes $(v, u \in V)$ define the **multi-scale (tensorial) Hamming distance**

$$H(v, u) = \|\mathcal{B}[v, :, :] - \mathcal{B}[u, :, :]\|_H,$$

where $(\|\cdot\|_H)$ is the weighted tensor Hamming norm defined above. When a single slice (k) is required we simply write $(H^{(k)}(v, u) = \text{Ham}(\mathbf{b}_v^{(k)}, \mathbf{b}_u^{(k)}))$.

3.1.1 Node-level distance multiset and distribution

For a fixed node $(v \in V)$ the empirical **tensorial distance multiset**

$$D_v = \{H(v, u) : u \in V, u \neq v\}$$

collects the dissimilarities between (v) and every other node across *all* path scales simultaneously. Normalising by $(N - 1)$ yields the probability mass function

$$\mu_v(d) = \frac{1}{N - 1} |\{u \neq v : H(v, u) = d\}|, \quad d \in \{0, \dots, ND\}.$$

Remark 3.1. Setting the norm weights to $(w_k = \delta_{k\ell})$ recovers the slice-specific distribution $(\mu_v^{(\ell)})$ used in the original formulation, so all node-level results derived there remain valid as special cases.

Remark 3.2 (HC as a special case of HGM.). Choosing the admissible functional $\Phi(\mu) = \mathbb{E}_{d \sim \mu}[d]$ recovers $\text{SU}^{(k)}(v) = \text{HC}^{(k)}(v)$ and its multi-scale average. Thus HC is the first-moment summary within the broader distributional framework of HGM.

From a modeling standpoint, node-wise Hamming centrality $\text{HC}^{(k)}(v)$ explains **how** a single vertex differs from its peers at a fixed distance layer k . Many global questions, however, depend not on a single node but on the **distribution** of these differences across all node pairs. This motivates passing from $\text{HC}^{(k)}(v)$ to the graph-level family $\{\mu_G^{(k)}\}_{k=1}^D$, which records the full spectrum of per-scale disagreements and supports permutation-invariant summaries and comparison between graphs. The next section formalizes these distributions and their admissible functionals.

3.1.2 Graph-level distance multiset and distribution

Aggregating over all ordered pairs gives the **global multiset**

$$D_G = \{ H(v, u) : v, u \in V, v \neq u \},$$

which contains $(N(N-1))$ values and encodes the complete multi-scale dissimilarity structure of (G) . Its normalised histogram is the **tensorial distance distribution**

$$\mu_G(d) = \frac{1}{N(N-1)} |\{(v, u) : v \neq u, H(v, u) = d\}|.$$

For analyses that require scale resolution we still track the family $\{\mu_G^{(k)}\}_{k=1}^D$ obtained from the frontal slices $\mathcal{B}_{:, :, k}$.

Remark 3.3. Throughout, distributions $\mu_G^{(k)}$ are formed over **unordered** pairs $\{u < v\}$, whereas energies $E_k(G) = \|B_G^{(k)}\|_F^2$ count **ordered** pairs. Thus $E_k(G)/2$ equals the number of unordered pairs at distance k , and normalizations reflect this choice.

3.2 Multi-Scale Hamming Profile

The multi-scale profile $\{\mu_G^{(k)}\}_{k=1}^D$ can be understood as analyzing slices of the connectivity tensor $\mathcal{B}$. Each slice $\mathcal{B}_{:, :, k}$ yields a distribution $\mu_G^{(k)}$, and the complete tensor encodes all structural information without premature aggregation.

This connects to classical spectral graph theory: while the heat kernel $e^{\alpha A} = \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} A^k$ aggregates all scales with exponential weighting, our framework maintains full resolution by treating each tensor slice independently.

Theorem 3.4. *Let G be connected with diameter D . Then $B^{(k)} \equiv 0$ for all $k \geq D+1$. Consequently, for $k \geq D+1$ every row $b_v^{(k)}$ is the zero vector and $\mu_G^{(k)} = \delta_0$. In particular, there exists $k_0 \leq D+1$ such that for $k \geq k_0$ the slice-wise means are nonincreasing and equal 0 for all $k \geq D+1$.*

Proof. By definition of diameter, every ordered pair (i, j) has shortest-path distance at most D . Hence no pair has exact distance k once $k \geq D+1$, i.e., $B^{(k)} \equiv 0$ for all $k \geq D+1$. Thus $b_v^{(k)} = \mathbf{0}$ for each v and $\text{Ham}(b_v^{(k)}, b_u^{(k)}) = 0$ for all u , giving $\mu_G^{(k)} = \delta_0$ for $k \geq D+1$. Taking $k_0 := D$ yields nonincreasing slice means for all $k \geq k_0$ (they drop to 0 at $k = D+1$). ■

3.3 Individual Node Contributions

While our primary focus is the graph-level distribution, individual nodes contribute differently to this distribution. For a node v , define its contribution to the distribution at scale k as:

$$\text{HC}^{(k)}(v) = \frac{1}{N-1} \sum_{u \neq v} \text{Ham}(\mathbf{b}_v^{(k)}, \mathbf{b}_u^{(k)})$$

This measures the average dissimilarity between node v 's reachability pattern and those of all other nodes. We can similarly define the average across tensor slices:

$$\text{HC}(v) = \frac{1}{K} \sum_{k=1}^K \text{HC}^{(k)}(v)$$

where $K \leq \text{diam}(G)$ is the maximum path length of interest.

However, $\text{HC}(v)$ is merely the first moment of node v 's contribution to the distribution. The full distribution $\mu_v^{(k)}$ of distances from v contains much richer information about v 's structural role.

Note that we examine individual node contributions only to better understand the graph-level distribution $\mu_G^{(k)}$, which remains our primary object of study.

3.4 Examples

Let us examine how these distributions manifest in basic graph topologies:

Complete graph K_N : At $k = 1$, all pairwise distances equal 2 so $\mu_G^{(1)} = \delta_2$. For $k \geq 2$, $B^{(k)} \equiv 0$ and $\mu_G^{(k)} = \delta_0$.

Star graph S_N : At $k = 1$, the central hub's reachability vector differs from each leaf's vector in all N positions, while any two leaves have identical vectors (distance 0). This creates a distribution with most mass at distance 0 (leaf-leaf comparisons) and mass $2/N$ at distance N coming from the $2(N-1)$ ordered center-leaf pairs.

Ring graph C_N : Due to rotation invariance, all nodes contribute equally to the distribution, but the distribution itself is non-trivial. At $k = 1$, distances concentrate on a small set of even values (e.g., 2, 4; the exact set depends on N). The distribution evolves predictably with k .

Path graph P_N : Unlike the ring, the path graph lacks rotational symmetry. End nodes contribute differently than central nodes, creating a more complex distribution that reflects the linear structure.

These examples highlight how Hamming Graph Metrics capture structural patterns rather than just topological properties.

3.5 Structural–Dissimilarity Functionals

We turn the per-node/per-scale distance distributions into scalar descriptors via *admissible* functionals (Def.~3.5) at the node and graph levels.

Definition 3.5 (Admissible functionals.).

Let $\mathcal{P}(\{0, \dots, M\})$ be the set of probability measures on a finite alphabet (here $M = N - 1$). A map $\Phi : \mathcal{P}(\{0, \dots, M\}) \rightarrow \mathbb{R}$ is **admissible** if:

- (i) **Permutation invariance:** $\Phi(\mu)$ depends only on the measure (relabeling the support does not change Φ);
- (ii) **TV–continuity:** Φ is continuous in the ℓ_1 (total-variation) topology on $\mathcal{P}$;
- (iii) **Finite on extremals:** $\Phi(\delta_x)$ is finite for each point mass δ_x .

When quantitative stability is needed, assume a TV–Lipschitz constant L_Φ so that $|\Phi(\mu) - \Phi(\nu)| \leq L_\Phi \|\mu - \nu\|_1$ for all $\mu, \nu \in \mathcal{P}$.

Remark 3.6. Typical admissible choices include Shannon entropy, Rényi entropies ($\alpha > 0, \alpha \neq 1$), total-variation dispersion, Wasserstein-1 on $\{0, \dots, M\}$ (with fixed ground metric), and Gini-type indices.

Theorem 3.7. *Let $D := \text{diam}(G)$. For any admissible $\Phi : \mathcal{P}(\{0, \dots, N - 1\}) \rightarrow \mathbb{R}$, define*

$$\Phi_v := \frac{1}{D} \sum_{k=1}^D \Phi(\mu_v^{(k)}),$$

where $\mu_v^{(k)}$ is the empirical distribution of $\{\text{Ham}(b_v^{(k)}, b_u^{(k)}) : u \in V, u \neq v\}$. Then:

- 1) (**Automorphism invariance**) For any graph automorphism σ , $\Phi_{\sigma(v)} = \Phi_v$ for all v .
- 2) (**TV–continuity**) If Φ is TV–Lipschitz with constant L_Φ , then for two graphs G, H on the same vertex set,

$$|\Phi_v(G) - \Phi_v(H)| \leq \frac{L_\Phi}{D} \sum_{k=1}^D \|\mu_v^{(k)}(G) - \mu_v^{(k)}(H)\|_1.$$

Proof. 1) Let P_σ be the permutation matrix of σ . For every k , $B^{(k)}(\sigma(G)) = P_\sigma B^{(k)}(G) P_\sigma^\top$, so $\text{Ham}(b_{\sigma(v)}^{(k)}, b_{\sigma(u)}^{(k)}) = \text{Ham}(b_v^{(k)}, b_u^{(k)})$ for all u , hence $\mu_{\sigma(v)}^{(k)} = \mu_v^{(k)}$ and $\Phi(\mu_{\sigma(v)}^{(k)}) = \Phi(\mu_v^{(k)})$; averaging over k gives $\Phi_{\sigma(v)} = \Phi_v$.

2) Apply TV-Lipschitzness to $\mu_v^{(k)}(G)$ vs $\mu_v^{(k)}(H)$ and average over k . ■

Theorem 3.8. *Let $D := \text{diam}(G)$. Define*

$$\bar{\Phi}(G) := \frac{1}{D} \sum_{k=1}^D \Phi(\mu_G^{(k)}),$$

where $\mu_G^{(k)}$ is the empirical distribution over unordered pairs $\{u < v\}$ of the distances $\text{Ham}(b_u^{(k)}, b_v^{(k)})$. Then $\bar{\Phi}$ is invariant under vertex relabeling, and if Φ is TV-Lipschitz with constant L_Φ ,

$$|\bar{\Phi}(G) - \bar{\Phi}(H)| \leq \frac{L_\Phi}{D} \sum_{k=1}^D \|\mu_G^{(k)} - \mu_H^{(k)}\|_1.$$

Random-graph separation (sketch). For fixed $p \in (0, 1)$, two independent $G, H \sim G(n, p)$ satisfy $\bar{\Phi}(G) \neq \bar{\Phi}(H)$ with probability $\rightarrow 1$ as $n \rightarrow \infty$ for any non-constant admissible Φ ; see Proposition~.6,(a) (Appendix~.4).

Rather than fix a statistic *a priori*, any admissible functional Φ can serve as a uniqueness descriptor. Important examples include:

Per-scale spectral analysis (classical MDS) For each k , form $D^{(k)} \in \mathbb{R}^{N \times N}$ with $D_{uv}^{(k)} = \text{Ham}(b_u^{(k)}, b_v^{(k)})$. Let $J = I - \frac{1}{N} \mathbf{1}\mathbf{1}^\top$ and define the (double-centered) Gram matrix

$$G^{(k)} = -\frac{1}{2} J D^{(k)} J,$$

where, for binary vectors, $\text{Ham}(x, y) = \|x - y\|_2^2$, so $D^{(k)}$ is already a squared-distance matrix (no elementwise squaring). With the eigendecomposition $G^{(k)} = Q^{(k)} \Lambda^{(k)} (Q^{(k)})^\top$, the spectral coordinates are

$$X^{(k)} = Q_+^{(k)} (\Lambda_+^{(k)})^{1/2},$$

and the total explained variance is $\text{tr}(\Lambda_+^{(k)})$.

Table 2: Examples of admissible functionals on per-scale distance distributions.

Functional	Definition on μ	Interpretation
Expectation $\mathbb{E}_\mu[f]$	$\sum_d f(d) \mu(d)$	Recovers linear stats (e.g. classical HC with $f(d) = d$)
Cumulant GF $K_\mu(t)$	$\log \mathbb{E}_\mu[e^{td}]$	Generates all cumulants
Rényi entropy $H_\alpha(\mu)$	$\frac{1}{1-\alpha} \log \sum_d \mu(d)^\alpha$	Measures spread/uncertainty
Spectral radius of moment matrix	$\rho(M_{kl} = \sum_d d^{k+l} \mu(d))$	Governs tail heaviness & concentration

3.5.1 Tensor Fingerprints via Unfolding Spectra (Permutation-Invariant, Non-Metric)

We define a graph fingerprint from the exact- k tensor $\mathcal{B} \in \{0, 1\}^{N \times N \times D}$ that is invariant to vertex relabeling, stable to small perturbations, and empirically discriminative.

Mode spectra and per-scale energies. Let $\mathcal{B}_{(m)}$ denote the mode- m matricization (unfolding) of $\mathcal{B}$ (notation as in [10]):

$$\mathcal{B}_{(1)} \in \mathbb{R}^{N \times (ND)}, \quad \mathcal{B}_{(2)} \in \mathbb{R}^{N \times (ND)}, \quad \mathcal{B}_{(3)} \in \mathbb{R}^{D \times (N^2)}.$$

Let $\sigma^{(m)} = (\sigma_1^{(m)} \geq \dots)$ be the singular values of $\mathcal{B}_{(m)}$. Define the per-scale energies (ordered pairs at exact distance k)

$$E_k(G) := \|B_G^{(k)}\|_F^2 = \sum_{i,j} (B_G^{(k)})_{ij}.$$

The *HGM tensor fingerprint* of G is

$$\text{FP}(G) := (\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}, (E_1(G), \dots, E_D(G))).$$

Proposition 3.9. *Permutation invariance.* If $H = \pi(G)$ for a relabeling π with permutation matrix P , then $\text{FP}(H) = \text{FP}(G)$.

Proof. Let P be the permutation matrix of the relabeling. For every k ,

$$\mathcal{B}_H(:, :, k) = P B_G^{(k)} P^\top.$$

Hence

$$E_k(H) = \|\mathcal{B}_H(:, :, k)\|_F^2 = \|P B_G^{(k)} P^\top\|_F^2 = \|B_G^{(k)}\|_F^2 = E_k(G),$$

by Frobenius-norm invariance under left/right multiplication by orthogonal (permutation) matrices.

For the unfoldings, there exist permutation matrices Π_1, Π_2, Π_3 (from the unfolding convention) such that

$$\mathcal{B}_{H,(1)} = P \mathcal{B}_{G,(1)} \Pi_1, \quad \mathcal{B}_{H,(2)} = P \mathcal{B}_{G,(2)} \Pi_2, \quad \mathcal{B}_{H,(3)} = \mathcal{B}_{G,(3)} \Pi_3.$$

Left/right multiplication by orthogonal matrices preserves singular values, so $\sigma^{(m)}(H) = \sigma^{(m)}(G)$ for $m \in \{1, 2, 3\}$. ■

Proposition 3.10 (Stability.). *If G' is obtained from G by toggling one edge and Δ is the max degree of $G \cup G'$, then for $M_r = \max_x |B_r(x)|$ (balls in graph distance):*

$$|E_k(G) - E_k(G')| \leq 2 M_{k-1}^2 \quad (1 \leq k \leq D),$$

and hence

$$\|\mathcal{B}(G) - \mathcal{B}(G')\|_F^2 \leq 2 \sum_{k=1}^D M_{k-1}^2, \quad \|\mathcal{B}_{(m)}(G) - \mathcal{B}_{(m)}(G')\|_2 \leq \|\mathcal{B}(G) - \mathcal{B}(G')\|_F$$

for $m \in \{1, 2, 3\}$. In particular, for $\Delta \geq 3$,

$$M_r \leq \frac{\Delta}{\Delta - 2} (\Delta - 1)^r \implies \|\mathcal{B}(G) - \mathcal{B}(G')\|_F \leq \frac{\sqrt{2} \Delta}{\Delta - 2} \left(\sum_{k=1}^D (\Delta - 1)^{2(k-1)} \right)^{1/2}.$$

Proof. The bound on E_k is the exact- k edge-flip bound (Proposition~.7). Summing over k gives the Frobenius bound because $\|\mathcal{X}\|_F^2$ counts the number of flipped 1's across slices for binary tensors. For singular values, the Mirsky bound gives $|\sigma_r(A) - \sigma_r(B)| \leq \|A - B\|_2$, and Hoffman–Wielandt yields $\sum_r (\sigma_r(A) - \sigma_r(B))^2 \leq \|A - B\|_F^2$ [14, 15]. Unfolding preserves the Frobenius norm, so $\|\mathcal{B}_{(m)}(G) - \mathcal{B}_{(m)}(G')\|_2 \leq \|\mathcal{B}(G) - \mathcal{B}(G')\|_F$. The $\Delta \geq 3$ bound is the standard branching estimate for ball sizes. ■

Random-graph separation (sketch). For fixed $p \in (0, 1)$, two independent $G, H \sim G(n, p)$ satisfy $\text{FP}(G) \neq \text{FP}(H)$ with probability $\rightarrow 1$ as $n \rightarrow \infty$; see Proposition~.6, (b) (Appendix~.4).

Scope. $\text{FP}(G)$ is a *graph invariant* and a stable, compact *structural fingerprint*. It does not replace HGM's metric; rather, it complements it: the metric compares labeled tensors directly, while **FP** summarizes cross-scale structure in a permutation-invariant way for indexing, retrieval, or visualization.

Corollary 3.11. *Let $E_k(G) = \|B_G^{(k)}\|_F^2 = \#\{(i, j) : \text{dist}_G(i, j) = k\}$ and $\text{FP}(G) = (\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}, (E_1, \dots, E_D))$ be the HGM tensor fingerprint.*

(a) *If G is vertex-transitive, let $n_k = |S(v, k)|$ for any v (independent of v). Then*

$$E_k(G) = \sum_{i \in V} |S(i, k)| = N n_k,$$

so $(E_1, \dots, E_D)$ is N times the classical distance distribution of G . In particular, if G and H are distance-regular with different intersection arrays (hence different $\{n_k\}$), then $\text{FP}(G) \neq \text{FP}(H)$.

(b) *More generally, for any graphs G, H , if their ordered-pair distance histograms differ, then $(E_k(G))_k \neq (E_k(H))_k$ and hence $\text{FP}(G) \neq \text{FP}(H)$.*

Proof. (a) Vertex-transitivity implies $|S(i, k)| = n_k$ for all i , hence $E_k(G) = \sum_i n_k = N n_k$. In distance-regular graphs, the sequence $(n_k)_{k=0}^D$ is determined by the intersection array via the standard three-term recurrence; distinct arrays yield distinct (n_k) , so (E_k) differs and thus the fingerprints differ.

(b) By definition $E_k(G)$ counts ordered pairs at distance k ; different histograms force (E_k) to differ. ■

Remark 3.12. Since $E_k(G)/2$ equals the number of *unordered* pairs at distance k , the Wiener index is

$$W(G) = \sum_{i < j} \text{dist}(i, j) = \sum_{k=1}^D k \frac{E_k(G)}{2}.$$

which is refined by the full vector (E_k) retaining distance multiplicities.

3.5.2 Graph-to-Graph Comparison Metrics

Graphs that share similar node-level signatures can still differ in global organization. We compare graphs using the exact- k tensor $\mathcal{B}$ via a *tensor Hamming metric* on labeled graphs and, for unlabeled comparison, a brief alignment remark.

Tensor Hamming metric (labeled graphs). For graphs on a fixed labeled vertex set $[N]$, define

$$d_{\text{ten}}(G, H) = \|\mathcal{B}_G - \mathcal{B}_H\|_1 = \sum_{i,j,k} \mathbf{1}[(\mathcal{B}_G)_{ijk} \neq (\mathcal{B}_H)_{ijk}].$$

Proposition 3.13. *For graphs on a fixed labeled vertex set $[N]$,*

$$d_{\text{ten}}(G, H) = \|\mathcal{B}_G - \mathcal{B}_H\|_1 = \sum_{i,j,k} \mathbf{1}[(\mathcal{B}_G)_{ijk} \neq (\mathcal{B}_H)_{ijk}]$$

is a metric. The normalized form $\bar{d}_{\text{ten}} = \|\mathcal{B}_G - \mathcal{B}_H\|_1 / (N(N-1)D) \in [0, 1]$ aids cross-size comparison.

Proof. $\|\cdot\|_1$ on tensors satisfies nonnegativity, symmetry, and the triangle inequality; positivity holds because $\mathcal{B}_G$ is determined by G (exact- k slices), so $\mathcal{B}_G = \mathcal{B}_H$ iff $G = H$ on the common label set. ■

Unlabeled graphs. For isomorphism classes $[G]$, define

$$d_{\text{iso}}([G], [H]) = \min_{\pi \in S_N} \|\mathcal{B}_G - \mathcal{B}_{\pi(H)}\|_1.$$

Then d_{iso} is a metric on isomorphism classes: $d_{\text{iso}}([G], [H]) = 0$ iff $G \cong H$; symmetry is immediate; the triangle inequality follows by composing near-minimizers for $([G], [H])$ and $([H], [F])$. (Worst-case evaluation is exponential due to the permutation minimization.)

3.5.3 Theoretical Properties

We now present several formal results characterizing Hamming distributions.

Proposition 3.14 (Minimal Structural Diversity in Complete Graphs). *Let $G = K_N$. Then:*

For $k = 1$: $\mu_G^{(1)} = \delta_2$ (point mass at 2) For $k \geq 2$: $\mu_G^{(k)} = \delta_0$ if we consider saturation effects

Proof. At $k = 1$, any two adjacency rows of K_N differ only at their two diagonal positions, so all pairwise Hamming distances equal 2. For $k \geq 2$, exact- k reachability is empty in K_N and $B^{(k)} \equiv 0$ (Lemma~.20), hence $\mu_G^{(k)} = \delta_0$. ■

Proposition 3.15 (Distribution Convergence). *For connected G and $k \rightarrow \infty$: $\mu_G^{(k)} \rightarrow \delta_0$ in total variation distance*

Proof. Under the exact- k convention, $B^{(k)} \equiv 0$ for all $k \geq D + 1$ in a connected graph of diameter D . Thus every $b_v^{(k)}$ is the zero vector and all pairwise Hamming distances are 0, i.e., $\mu_G^{(k)} = \delta_0$ for $k \geq D + 1$. ■

3.6 Extremal-Class Results

The following sharpen earlier bounds within the functional setting.

Proposition 3.16 (Star-graph separation (entropy).).

For the star S_N at $k = 1$, the center has $\mu_c^{(1)} = \delta_N$ while any leaf ℓ has

$$\mu_\ell^{(1)} = \frac{N-2}{N-1} \delta_0 + \frac{1}{N-1} \delta_N.$$

Hence $H(\mu_c^{(1)}) = 0$ and $H(\mu_\ell^{(1)}) > 0$ for $N \geq 3$, so H distinguishes center vs. leaves at $k = 1$.

Proposition 3.17 (Star Graph Separation). *For S_N and any strictly convex functional Φ : $\Phi(\mathcal{P}_0^{(1)}) \neq \Phi(\mathcal{P}_i^{(1)})$ for every leaf i , capturing structural non-equivalence beyond mean distance.*

Proof. For $k = 1$, the center's distribution is a point mass at N , while a leaf's distribution has mass $(N-2)/(N-1)$ at 0 and mass $1/(N-1)$ at N . ■

Proposition 3.18 (TV dispersion bound (sharp).).

Let $\bar{\mu}^{(k)} = \frac{1}{N} \sum_v \mu_v^{(k)}$ and $\Psi^{(k)}(G) = \frac{1}{N} \sum_v \|\mu_v^{(k)} - \bar{\mu}^{(k)}\|_1$. Then

$$0 \leq \Psi^{(k)}(G) \leq 2 \left(1 - \sum_d (\bar{\mu}^{(k)}(d))^2 \right) < 2,$$

with equality in the upper bound iff each $\mu_v^{(k)}$ is a point mass (Dirac). In particular, if all $\mu_v^{(k)}$ are Dirac and split between two distances with proportions p and $1-p$, then $\Psi^{(k)}(G) = 4p(1-p) \leq 1$ (max at $p = \frac{1}{2}$).

Proof. For each distance value d , by convexity the average absolute deviation $\frac{1}{N} \sum_v |\mu_v^{(k)}(d) - \bar{\mu}^{(k)}(d)|$ is maximized when each coordinate takes values in $\{0, 1\}$; summing over d yields

$$\Psi^{(k)} = 2 \sum_d \bar{\mu}^{(k)}(d)(1 - \bar{\mu}^{(k)}(d)) = 2(1 - \sum_d \bar{\mu}^{(k)}(d)^2).$$

Strict inequality (<2) holds since $\sum_d \bar{\mu}^{(k)}(d)^2 > 0$. Equality in the bound occurs exactly when every $\mu_v^{(k)}$ is a Dirac. The two-group formula follows by plugging $\bar{\mu}^{(k)} = p \delta_{d_1} + (1-p) \delta_{d_2}$. ■

3.6.1 Stability to edge perturbations

Let G' be obtained from G by toggling a single edge $e = \{u, v\}$ and let Δ be the maximum degree of $G \cup G'$. For $r \geq 0$, write $B_r(x) = \{y : \text{dist}(x, y) \leq r\}$ and $M_r = \max_x |B_r(x)|$. A shortest k -path that changes status due to e must traverse e , hence has the form $i \rightsquigarrow u$ (length a), then $u-v$, then $v \rightsquigarrow j$ (length b) with $a + b + 1 = k$ (or the symmetric $u \leftrightarrow v$ case). Therefore, at scale k the set of ordered pairs (i, j) that can flip is contained in $B_{k-1}(u) \times B_{k-1}(v) \cup B_{k-1}(v) \times B_{k-1}(u)$, so the number of flips at scale k is at most

$$F_k \leq 2 |B_{k-1}(u)| |B_{k-1}(v)| \leq 2 M_{k-1}^2.$$

Summing over k gives the *exact- k* tensor bound

$$\|\mathcal{B}_G - \mathcal{B}_{G'}\|_1 \leq \sum_{k=1}^D F_k \leq 2 \sum_{k=1}^D M_{k-1}^2.$$

Using degree growth, for $\Delta \geq 3$ we have $M_r \leq 1 + \Delta \sum_{t=0}^{r-1} (\Delta - 1)^t \leq \frac{\Delta}{\Delta-2} (\Delta - 1)^r$ for $r \geq 1$, hence

$$\|\mathcal{B}_G - \mathcal{B}_{G'}\|_1 \leq \frac{2\Delta^2}{(\Delta - 2)^2} \sum_{k=1}^D (\Delta - 1)^{2(k-1)}.$$

For $\Delta = 2$ (paths/cycles), $M_r \leq 2r + 1$, yielding the quadratic bound

$$\|\mathcal{B}_G - \mathcal{B}_{G'}\|_1 \leq 2 \sum_{k=1}^D (2(k-1) + 1)^2.$$

Normalizing,

$$\bar{d}_{\text{ten}}(G, G') = \frac{\|\mathcal{B}_G - \mathcal{B}_{G'}\|_1}{N(N-1)D} \leq \begin{cases} \frac{2\Delta^2}{N(N-1)D(\Delta-2)^2} \sum_{k=1}^D (\Delta-1)^{2(k-1)}, & \Delta \geq 3, \\ \frac{2}{N(N-1)D} \sum_{k=1}^D (2(k-1)+1)^2, & \Delta = 2. \end{cases}$$

3.7 Connections to classical invariants

The Hamming distribution encodes classical graph invariants:

Proposition 3.19. *For any graph G :*

1. *For vertex-transitive graphs (e.g., cycles, hypercubes) and fixed k , all $|S(v, k)|$ are equal, so every pairwise per-scale distance is even; in particular $\text{supp}(\mu_G^{(k)}) \subseteq \{0, 2, 4, \dots\}$.*
2. *In general graphs, $\text{supp}(\mu_G^{(k)}) \subseteq \{0, 1, \dots, N\}$; parity constraints need not hold when shell sizes vary across vertices.*
3. *The mode of $\mu_G^{(k)}$ reflects the typical overlap structure at scale k (e.g., relates to k -shells/cores in many ensembles), though precise identification is graph-class dependent.*

These connections allow HGM to subsume and extend classical structural analysis.

3.8 Brief Computational Remark

While our focus is theoretical, HGM can be evaluated efficiently on large sparse graphs using bit-parallel primitives (bit-packing, XOR, and hardware `popcount`). Implementation details—`popcount`-based XOR kernels, min-hash sketching for approximate summaries, and blockwise parallelism—are given in Appendix~A. These techniques scale to graphs with $N \sim 10^5$ vertices in practice on multicore/GPU systems (see also [9]). All theoretical guarantees above are algorithm-independent.

4 Theoretical Analysis of Graph Classes

To further ground Hamming Graph Metrics in structural graph theory, we now derive and summarize their behavior across classical graph families. These results follow directly from the definitions without simulation or

measurement, and serve to illustrate how uniqueness, dissimilarity, and dispersion vary with symmetry, modularity, and degree heterogeneity.

Let Φ be any structural descriptor derived from Hamming distance distributions computed from tensor slices $\mathcal{B}_{:, :, k}$ (as introduced in Section 4), and let $\Phi_v^{(k)}$ denote the value of this descriptor at node v and path scale k .

4.1 Regular and Vertex-Transitive Graphs

Let G be a connected d -regular graph.

Theorem 4.1 (Uniformity under Symmetry). *If G is vertex-transitive, then for all $v, w \in V$: $\mu_v^{(k)} = \mu_w^{(k)}$ and thus $\Phi_v^{(k)} = \Phi_w^{(k)}$*

Proof. Vertex transitivity implies the existence of an automorphism mapping any vertex to any other. Such automorphisms preserve Hamming distances between reachability vectors, hence preserve the distributions. ■

Corollary 4.2. *Under Theorem~4.1, complete graphs K_N , cycles C_N , and hypercubes Q_n are vertex-transitive; hence for each fixed k , the distributions $\mu_v^{(k)}$ (and any admissible $\Phi_v^{(k)}$) are identical for all v . Multi-scale behavior can still differ across k .*

Remark 4.3. For circulant graphs, analytic expressions for $\mu_v^{(k)}$ can be derived using modular arithmetic on adjacency shifts. Specifically, for the cycle C_N , the distance between nodes i and j at scale k depends only on $|i - j| \bmod N$ and whether k is sufficient to traverse that arc length.

Theorem 4.4 (Spectral Characterization). *For d -regular graphs with adjacency eigenvalues $\lambda_1 = d > \lambda_2 \geq \dots \geq \lambda_N$:*

$$\text{Var} \left[\mu_G^{(k)} \right] \leq \frac{d^{2k} \cdot \left(1 - \left(\frac{\lambda_2}{d} \right)^{2k} \right)}{N \cdot \left(1 - \left(\frac{\lambda_2}{d} \right)^2 \right)}$$

This connects expansion properties to uniqueness dispersion [9].

4.2 Trees and Star Graphs

Trees exhibit hierarchical expansion and strong local asymmetry.

Theorem 4.5 (Star Graph Asymmetry, Generalized). *In the star graph S_N , the center node maximizes:*

- Structural dissimilarity $\Phi_c^{(k)}$
- Deviation from the mean profile $\|\mu_c^{(k)} - \bar{\mu}^{(k)}\|_1$
- Entropy $H(\mu_c^{(k)})$ among all nodes

Each leaf has identical minimal distributions. This is the maximal variance configuration among all trees.

Proof. By direct counting of reachability patterns at $k = 1$ (see Appendix A.1), we have explicit forms for $\mu_c^{(1)}$ and $\mu_{\ell_i}^{(1)}$. At $k = 1$, the center's distribution is δ_N , while each leaf's distribution places mass $(N - 2)/(N - 1)$ at 0 and $1/(N - 1)$ at N . The claims follow by direct computation. ■

Theorem 4.6 (Height-Monotonicity in Trees). *Let T be a tree rooted at node r . Then for any node v :*

$$\text{depth}(v) \uparrow \Rightarrow \Phi_v^{(k)} \downarrow \text{ for small } k$$

Proof. Nodes at greater depth have fewer descendants and more similar neighborhoods. Their reachability vectors at small k overlap more with their siblings, reducing average dissimilarity. ■

Proposition 4.7 (Binary Tree Regularity). *In a complete binary tree of height h :*

- Nodes at the same level have identical distributions
- $H(\mu_{\text{level}}^{(k)})$ decreases monotonically with level for $k < h$
- The root maximizes entropy at all scales

This stratification by height is a general feature of trees.

4.3 Random Graphs: Erdős–Rényi Model

Let $G \sim G(N, p)$, with $p \in (0, 1)$.

Proposition 4.8 (Expected Uniqueness Peak). *Heuristic outline. The regime analysis follows the usual $G(N, p)$ thresholds: uniqueness is minimal for $p \ll \frac{\log N}{N}$ or $p \rightarrow 1$, and peaks near $p_c \approx \frac{\log N}{N}$ as the giant component emerges and diameters are still large.*

In $G(N, p)$, uniqueness is low when p is small (fragmented graph) or large (distances collapse), but peaks near the connectivity threshold $p_c \approx \log N/N$.

- **Subcritical:** small isolated components, minimal diversity.
- **Critical:** giant component emerges, producing maximum path-length diversity.
- **Supercritical:** diameter shrinks, reachability vectors homogenize.

Proposition 4.9 (Concentration of Distributions). *Heuristic outline. Above the connectivity and diameter-collapse thresholds (typ. $O(\log N)$), rows of $B^{(k)}$ become nearly identical for fixed k , forcing $\mu_G^{(k)}$ to concentrate.*

4.4 Scale-Free Networks: Barabási–Albert Model

Let $G \sim \text{BA}(N, m)$, the preferential attachment graph with initial degree m .

In BA graphs, hubs connect to a wide range of degree classes, producing many distinct distances and high variance in their reachability vectors. Low-degree nodes connect mostly through hubs, yielding more uniform patterns. The support size for hub distances scales as $\Theta(\sqrt{N})$. Appendix A.9 contains the variance comparison and scaling argument.

Proposition 4.10. *In preferential-attachment graphs $G \sim \text{BA}(N, m)$, structural uniqueness correlates positively with degree and exhibits super-linear growth at hubs due to path diversity. See Appendix for variance comparison and scaling arguments.*

4.5 Small-World Networks: Watts–Strogatz Model

Let $G \sim \text{WS}(N, k, \beta)$, a rewiring of the k -regular ring lattice [6].

Proposition 4.11 (Shortcut-Induced Uniqueness). *For rewiring probability $\beta \in (0, 1)$, let $S \subset V$ be the set of shortcut endpoints. Then:*

1. *Nodes in S have significantly elevated $\Phi_v^{(k)}$ for small k*
2. *Their distance distributions $\mu_v^{(k)}$ deviate maximally from the lattice background*

Proof. Shortcuts create asymmetric reachability patterns that propagate locally. As k increases, the regular lattice structure dominates, diminishing the shortcut effect. ■

Theorem 4.12. *Consider the WS model on N vertices starting from a ring lattice where each vertex has degree d (even), so $m = Nd/2$ undirected edges. Each edge is rewired independently with probability $\beta = \beta(N)$ to a uniformly random new endpoint (avoiding loops/multi-edges). Then:*

- (a) (**Onset of shortcuts**) The expected number of rewired edges is $\mathbb{E}[X] = \beta m = \Theta(\beta N)$. The “first-shortcut” threshold satisfies

$$\beta_c \asymp \frac{1}{N},$$

in the sense that if $N\beta \rightarrow 0$ then $X \xrightarrow{p} 0$ (no shortcuts whp), while if $N\beta \rightarrow \infty$ then $X \xrightarrow{p} \infty$ (many shortcuts whp).

- (b) (**Distance regime split**) If $N\beta \rightarrow 0$, typical distances are ring-like (mean distance $\Theta(N/d)$). If $N\beta \rightarrow \infty$ with d fixed, the added shortcut set induces long-range connections whose coarse-grained effect is a random-sparse overlay; the average distance drops to $O(\log N)$ (small-world regime).

Proof. (a) Let $X \sim \text{Binomial}(m, \beta)$ be the number of rewired edges. With $m = \Theta(N)$, $\mathbb{E}[X] = \Theta(\beta N)$. A standard second-moment/Chernoff argument gives: if $N\beta \rightarrow 0$, then $X = 0$ with probability $1 - o(1)$; if $N\beta \rightarrow \infty$, then $X \rightarrow \infty$ in probability. Hence the onset occurs at $\beta_c = \Theta(1/N)$.

- (b) When $N\beta \rightarrow 0$, whp no shortcuts appear; the graph is the original ring lattice, so mean distance is $\Theta(N/d)$. When $N\beta \rightarrow \infty$, the shortcut set has $\Theta(\beta N)$ random long edges. Coarse-graining the ring into arcs of length $\ell = \ell(N)$ with $1 \ll \ell \ll N$, the induced “supergraph” on N/ℓ arcs receives $\Theta(\beta N)$ random edges, i.e., average super-degree $\Theta(\beta \ell)$. Choosing ℓ so that $\beta \ell \rightarrow c > 0$ yields a sparse random overlay whose giant-component/expander-like behavior drops average distance to $O(\log N)$ between arcs; lifting back to vertices gives $O(\log N)$ for the original graph up to constants. (This is the standard random-shortcut argument.) Appendix~4.1 derives $\beta_c \sim 1/N$ and the scaling. ■

4.6 Summary of Hamming profile behavior derived from tensor slices $\mathcal{B}_{:, :, k}$ across graph classes

4.7 Extended Results and Corollaries

We now present additional theoretical results that deepen our understanding of Hamming Graph Metrics.

Proposition 4.13 (Uniqueness Flatness in Distance-Regular Graphs). *Let G be a distance-regular graph, i.e., the number of nodes at each distance from a given node depends only on the distance, not the node itself. Then:*

Table 3: Summary of multi-scale Hamming profile behavior across graph classes.

Graph Class	Hamming Profile Behavior	Dispersion $\Psi^{(k)}(G)$	Entropy Peak
Complete Graph K_N	Uniform at $k = 1$ (HC=2)	0 (point mass)	$k = 1$
Star Graph S_N	Maximal asymmetry at $k = 1$	$\mathcal{O}(1)$	$k = 1$
Cycle C_N	Uniform, periodic pattern	$\mathcal{O}(1)$	$k \approx N/4$
Binary Tree	Level-stratified	$\mathcal{O}(\log N)$	$k \approx \log N$
Erdős–Rényi $G(N, p^*)$	Critical behavior	$\Theta(\sqrt{N})$	$k \approx \log N$
Barabási–Albert	Degree-correlated	$\Theta(\log N)$	$k \approx 2$
Watts–Strogatz	Shortcut spikes	$\mathcal{O}(\log N)$	Varies with β

- For all $v \in V$, $\mu_v^{(k)} = \mu_w^{(k)}$ for all $w \in V$
- Hence $\Phi_v^{(k)} = \text{const}$ for all v , for any Φ

Examples include: cycles C_n , complete graphs, and hypercubes Q_n . This extends Theorem~4.4 by identifying a larger class of graphs where uniqueness is structurally flat due to distance symmetry, not just vertex-transitivity.

Proof. Distance-regularity implies that the number of nodes at distance d from any node is constant. Combined with the fact that reachability at scale k depends only on distance relationships, the claim follows. ■

Proposition 4.14 (Extremal Support Collapse in Clique Chains). *Let G*

be a clique chain of r fully connected components $K_{n_1}, K_{n_2}, \dots, K_{n_r}$ joined sequentially by single bridges. Then:

- Nodes within the same clique have highly overlapping reachability vectors
- Bridge nodes exhibit maximal uniqueness support, with $|\text{Support}(\mu_v^{(k)})| = \{0, 1, \dots, d_{\max}\}$
- The Gini coefficient $G_\mu(G)$ increases linearly with r

Proof. Within cliques, all nodes reach the same set at each scale. Bridge nodes uniquely connect components, creating maximal diversity in their distance distributions. The Gini coefficient captures this inequality. ■

4.8 Extension: Spectral Interpretation of HC Dispersion

Let Δ denote the graph Laplacian of G , and let λ_2 be the algebraic connectivity (i.e., the second-smallest eigenvalue).

Theorem 4.15 (Spectral Lower Bound on Uniqueness Dispersion). *Let $D_\mu(G)$ be the variance of node-level uniqueness (using $\mu_v = \text{HC}(v)$). Then:*

$$D_\mu(G) \geq \frac{1}{N} \cdot \frac{(\sum_v \deg(v) \cdot \mu_v)^2}{\lambda_2 \cdot \sum_v \deg(v)^2} \cdot \text{Var}[\mu]$$

Proof. Apply the Poincar'e inequality to the function $f(v) = \mu_v$ on the graph:

$$\sum_{(u,v) \in E} (f(u) - f(v))^2 \geq \lambda_2 \sum_v \deg(v) \cdot (f(v) - \bar{f})^2$$

where $\bar{f}$ is the degree-weighted mean of f . Rearranging yields the stated bound. ■

Interpretation: Graphs with small spectral gap (i.e., loosely connected) allow greater variation in structural uniqueness, while tight expanders constrain nodes to similar roles.

4.9 Extension: Robustness Under Edge Perturbation

Let $G' = G + \Delta E$ be a graph obtained by inserting or deleting a small set ΔE of edges. Define:

$$\delta = \max_u \left\| \mu_u^{(k)}(G) - \mu_u^{(k)}(G') \right\|_1$$

Proposition 4.16 (Lipschitz Continuity of Hamming Distribution under Edge Noise). *There exists a constant C_k depending only on k and graph size such that:*

$$\left| \Phi_v^{(k)}(G') - \Phi_v^{(k)}(G) \right| \leq C_k \cdot \delta$$

for all admissible Φ that are 1-Lipschitz under total variation (ℓ_1).

Proof. Edge modifications affect reachability vectors only for nodes within distance k of the modified edges. The number of affected entries in any reachability vector is bounded by $(2d_{\max})^k$. The ℓ_1 (TV) Lipschitz property of Φ completes the proof. ■

Proposition 4.17 (Shortcut Bias in WS Graphs is Localised). *Let $G \sim \text{WS}(N, k, \beta)$, and let $S \subset V$ be nodes affected by rewired edges. Then for small β , the set:*

$$A_\varepsilon = \left\{ v : \left\| \mu_v^{(k)} - \bar{\mu}^{(k)} \right\|_1 > \varepsilon \right\}$$

has $|A_\varepsilon| = \mathcal{O}\left(\frac{\beta N k}{2}\right)$ with high probability. Thus, uniqueness deviations are sparse and concentrated near structural irregularities.

Proof. Each rewired edge affects $\mathcal{O}(1)$ nodes directly. The total number of rewired edges is approximately $\beta N k / 2$. Concentration inequalities for the rewiring process yield the result. ■

4.10 Extension: Graph Classes with Controlled Uniqueness Gradient

Define a uniqueness gradient as the discrete Laplacian applied to the field μ_v :

$$(\nabla^2 \mu)_v := \sum_{u \sim v} (\mu_u - \mu_v)$$

Theorem 4.18. *Let G be a connected d -regular graph on N vertices with random-walk matrix $P = A/d$ and spectral gap $\gamma := 1 - \lambda_2(P) > 0$. For any fixed scale $k \geq 1$, let*

$$f(v) := \text{HC}^{(k)}(v) = \frac{1}{N-1} \sum_{u \neq v} \text{Ham}(b_v^{(k)}, b_u^{(k)}).$$

Then

$$\text{Var}(f) \leq \frac{2 M_k^2}{\gamma}, \quad M_k := \max_x |S(x, k)|.$$

In particular, larger spectral gap γ (better expansion) forces the uniqueness field f to vary smoothly across the graph.

Proof. (Edgewise Lipschitz bound.) For any u, v, z in a Hamming space, the triangle inequality yields

$$|\text{Ham}(b_v^{(k)}, b_z^{(k)}) - \text{Ham}(b_u^{(k)}, b_z^{(k)})| \leq \text{Ham}(b_v^{(k)}, b_u^{(k)}).$$

Averaging over $z \neq v, u$ gives

$$|f(v) - f(u)| \leq \text{Ham}(b_v^{(k)}, b_u^{(k)}) \leq |S(v, k)| + |S(u, k)| \leq 2M_k.$$

Hence for each edge (u, v) , $(f(u) - f(v))^2 \leq 4M_k^2$.

(Poincaré on d -regular graphs.) The Dirichlet form is $\mathcal{E}(f, f) = \frac{1}{2N} \sum_{(u,v) \in E} (f(u) - f(v))^2$, and the Poincaré (spectral-gap) inequality reads

$$\text{Var}(f) \leq \frac{1}{\gamma} \mathcal{E}(f, f).$$

Using the edgewise bound and $|E| = dN/2$,

$$\mathcal{E}(f, f) \leq \frac{1}{2N} \cdot \frac{dN}{2} \cdot 4M_k^2 = 2M_k^2,$$

whence $\text{Var}(f) \leq \frac{2M_k^2}{\gamma}$. ■

Appendix~4.2 contains the spectral and Laplacian calculations

Remark 4.19. For nonregular graphs, replace γ by the spectral gap of the lazy random walk or use the normalized Laplacian; the same argument yields $\text{Var}(f) \lesssim M_k^2/\gamma$ up to degree factors.

Corollary 4.20. *Under Theorem~4.18, d -regular Ramanujan graphs (whose nontrivial spectrum lies in $[-2\sqrt{d-1}, 2\sqrt{d-1}]$) have optimal spectral gap; hence the uniqueness field varies smoothly across vertices with the strongest bound among d -regular expanders since $\lambda_2 \leq 2\sqrt{d-1}$, the bound from Thm.~4.18 is minimized.*

4.11 Detailed Algebraic Examples

We provide rigorous calculations for several graph families to illustrate the theoretical results.

Example 1: Complete Bipartite Graph $K_{m,n}$

Let $G = K_{m,n}$ with partitions A (size m) and B (size n), where $m \leq n$.

At $k = 1$:

- Nodes in A : $\mathbf{b}_v^{(1)}$ has 1s in all positions corresponding to B
- Nodes in B : $\mathbf{b}_v^{(1)}$ has 1s in all positions corresponding to A

Thus:

$$\text{-Ham}\left(\mathbf{b}_a^{(1)}, \mathbf{b}_{a'}^{(1)}\right) = 0 \text{ for } a, a' \in A \quad \text{-Ham}\left(\mathbf{b}_b^{(1)}, \mathbf{b}_{b'}^{(1)}\right) = 0 \text{ for } b, b' \in B$$

$$\text{-Ham}\left(\mathbf{b}_a^{(1)}, \mathbf{b}_b^{(1)}\right) = m + n \text{ for } a \in A, b \in B$$

The distribution is:

$$\mu_{K_{m,n}}^{(1)} = \frac{m(m-1) + n(n-1)}{(m+n)(m+n-1)} \delta_0 + \frac{2mn}{(m+n)(m+n-1)} \delta_{m+n}$$

At $k = 2$ (exact- k), the support is $\{2, N-2\}$: pairs within the same part have distance 2, and cross-part pairs have distance $N-2$. Thus

$$\mu_{K_{m,n}}^{(2)} = \frac{m(m-1) + n(n-1)}{N(N-1)} \delta_2 + \frac{2mn}{N(N-1)} \delta_{N-2}.$$

(Cumulatively, $B_{\leq 2}$ is fully connected.)

Example 2: Hypercube Q_n

The n -dimensional hypercube has $N = 2^n$ vertices, each of degree n . Vertices are binary strings of length n , with edges between strings that differ in exactly one bit.

At $k = 1$:

- Each node's reachability vector has weight exactly n (its neighbors).
- For two nodes u, v at Hamming distance $h = \text{dist}(u, v)$, using Appendix~4.4 (Hypercube sphere intersection),

$$\text{Ham}(b_u^{(1)}, b_v^{(1)}) = 2\left(n - \mathbf{1}_{\{h=2\}} \cdot 2\right) = \begin{cases} 0, & h = 0, \\ 2n - 4, & h = 2, \\ 2n, & h \in \{1\} \cup \{3, 4, \dots, n\}. \end{cases}$$

In particular, adjacent vertices ($h = 1$) have $\text{Ham} = 2n$ (not $2(n-1)$).

If v is chosen uniformly from $V \setminus \{u\}$, then

$$\mathbb{P}[\text{Ham}(b_u^{(1)}, b_v^{(1)}) = 2n - 4] = \frac{\binom{n}{2}}{2^n - 1}, \quad \mathbb{P}[\text{Ham} = 2n] = 1 - \frac{\binom{n}{2}}{2^n - 1},$$

and (for distinct pairs) $\mathbb{P}[\text{Ham} = 0] = 0$. Thus $\mu_{Q_n}^{(1)}$ is supported on $\{2n - 4, 2n\}$ for distinct pairs. (*See Appendix — Additional Technical Lemmas for the intersection counts underpinning these expressions.*)

Example 3: Petersen Graph

The Petersen graph is a 3-regular, vertex-transitive graph on 10 vertices with diameter 2 and girth 5.

At $k = 1$:

- Each node reaches exactly 3 neighbors (so $|S(v, 1)| = 3$ for all v).
- By vertex transitivity, $\text{HC}^{(1)}(v)$ is constant across v .

At $k = 2$ (exact- k):

- Each node has $|S(v, 2)| = 6$; thus $B^{(2)} \neq 0$.
- The per-scale distribution $\mu_G^{(2)}$ is supported on a small set of even values (not δ_0).

This illustrates that even in small-diameter, highly symmetric graphs, the exact- k slice at $k = 2$ remains informative, although the cumulative matrix $B_{\leq 2}$ is fully connected.

Example 4: Grid Graph $G_{m \times n}$

Consider the 2D grid with m rows and n columns.

- **Corner nodes** (degree 2): At $k = 1$, reach 2 neighbors $\rightarrow$ Have maximum average dissimilarity
- **Edge nodes** (degree 3): At $k = 1$, reach 3 neighbors $\rightarrow$ Intermediate dissimilarity
- **Interior nodes** (degree 4): At $k = 1$, reach 4 neighbors $\rightarrow$ Minimum average dissimilarity due to regular neighborhoods

The distribution $\mu_{G_{m \times n}}^{(k)}$ can be computed exactly using the Manhattan distance structure, revealing how boundary effects create structural heterogeneity even in regular lattices.

4.12 Tensor-Theoretic Properties

The tensor representation reveals additional structure:

Theorem 4.21 (Low-complexity slice span in distance-regular graphs). *Let G be a connected distance-regular k -regular graph with diameter D and adjacency matrix A . For each $i = 0, 1, \dots, D$, let A_i be the distance- i matrix, i.e., $(A_i)_{uv} = 1$ iff $\text{dist}(u, v) = i$ (so $A_1 = A$ and, under our exact- k convention, $B^{(i)} = A_i$ for $i \geq 1$). Then for each i there exists a polynomial p_i of degree i such that*

$$A_i = p_i(A).$$

Consequently, $\text{span}\{A_0, I, A_1, \dots, A_D\}$ has dimension at most $D + 1$; in particular, all exact-distance slices $\{B^{(i)}\}_{i=1}^D$ lie in a $(D + 1)$ -dimensional commutative algebra and are simultaneously diagonalizable with A .

Proof. In a distance-regular graph there are intersection numbers (a_i, b_i, c_i) such that for all $i = 0, \dots, D$,

$$A A_i = b_{i-1} A_{i-1} + a_i A_i + c_{i+1} A_{i+1},$$

with the conventions $A_{-1} = A_{D+1} = 0$ and $b_{-1} = c_{D+1} = 0$. (Combinatorially: multiplying by A moves you one step in graph distance, and the coefficients count how many neighbors land at distances $i - 1, i, i + 1$.) This three-term recurrence shows inductively that A_i lies in the polynomial algebra generated by A : set $p_0 \equiv 1$, $p_1(x) = x$, and use the recurrence to define $p_{i+1}(x)$ from $x p_i(x) = b_{i-1} p_{i-1}(x) + a_i p_i(x) + c_{i+1} p_{i+1}(x)$. Hence $A_i = p_i(A)$ with $\deg p_i = i$. The matrices $\{A_i\}_{i=0}^D$ form a basis of the Bose–Mesner algebra of the graph’s association scheme, which is a $(D + 1)$ -dimensional commutative algebra; therefore all A_i commute and are simultaneously diagonalizable with A . ■

Corollary 4.22. Hypercube Q_n . *For the n -dimensional hypercube Q_n ($N = 2^n$, diameter $D = n$), each exact-distance slice $B^{(i)} = A_i$ equals a degree- i polynomial $p_i(A)$ and all slices lie in an $(n + 1)$ -dimensional commutative algebra. In particular, the family $\{B^{(i)}\}_{i=1}^n$ admits an $O(\log N)$ -dimensional linear parametrization through A .*

4.12.1 Comparison with Traditional Metrics

Key insight. Traditional metrics summarize *importance*; HGM summarizes *how node-level structures differ* at each exact scale k , providing the full distribution $\mu_G^{(k)}$ rather than a single scalar per node or per graph.

Graph Edit Distance (GED): algorithmic vs. analytic Graph edit distance gives an **algorithmic** measure of discrepancy: it is the minimum total cost of a sequence of discrete edits (vertex/edge insertions, deletions, relabelings) that transforms one graph into another. GED thus captures *how to* align graphs procedurally, but it does not by itself yield **analytic invariants** or closed-form structure theorems about the distribution of connectivity patterns across scales.

By contrast, HGM provides an **analytic** account: the exact- k tensor $\mathcal{B}$ induces per-scale distributions, spectral summaries, and a labeled **metric** $d_{\text{ten}}(G, H) = \|\mathcal{B}_G - \mathcal{B}_H\|_1$, with an orbit metric d_{iso} for unlabeled comparison. These objects support stability bounds, extremal characterizations, and links to classical graph invariants.

Proposition 4.23 (Quantitative link under edge-only edits (labeled case).). *Let $\text{GED}_{\pm E}(G, H)$ be the minimum number of **edge toggles** needed to transform G into H on a fixed label set. With $M_r := \max_x |B_r(x)|$ (ball size at radius r),*

$$d_{\text{ten}}(G, H) = \|\mathcal{B}_G - \mathcal{B}_H\|_1 \leq 2 \text{GED}_{\pm E}(G, H) \sum_{k=1}^D M_{k-1}^2,$$

and, in particular for maximum degree $\Delta \geq 3$,

$$M_r \leq \frac{\Delta}{\Delta - 2} (\Delta - 1)^r \implies d_{\text{ten}}(G, H) \leq \frac{2\Delta}{\Delta - 2} \text{GED}_{\pm E}(G, H) \sum_{k=1}^D (\Delta - 1)^{2(k-1)}.$$

Proof. Each edge toggle affects only entries within $(k-1)$ steps of its endpoints in slice k , flipping at most $2M_{k-1}^2$ tensor entries (the exact- k edge-flip bound). Summing over k and over the $\text{GED}_{\pm E}(G, H)$ toggles yields the inequality. The degree-based estimate follows from the branching bound on M_r . ■

Takeaway. GED is a powerful **procedural** measure (edit programs), while HGM supplies **analytic** structure (per-scale distributions, spectra, and metrics) with stability guarantees. In regimes where an edit model is natural, the bound above shows how HGM’s tensor metric can be controlled by (edge-only) GED; conversely, HGM can distinguish graphs with identical low-cost edit programs by exposing differences in their multi-scale reachability distributions.

5 Extensions and Future Work

While Hamming Graph Metrics (HGM) offer a principled and scalable approach to quantifying structural uniqueness, several directions remain open for further theoretical development, practical extension, and domain-specific adaptation. We highlight five major avenues, each grounded in existing mathematical or computational structures.

5.1 Weighted and Directed Graphs

Motivation: Many real-world systems (e.g., transportation, gene regulation, web links) are neither unweighted nor symmetric. Path significance depends on edge weights (e.g., capacity, cost) and directions.

Proposal:

- For directed graphs, replace undirected adjacency A with asymmetric adjacency A_{dir} , and define separate reachability matrices for in-paths and out-paths: $B_{\text{in}}^{(k)}$, $B_{\text{out}}^{(k)}$
- Compute Hamming distances over:

$$\mathbf{b}_v^{(k,\text{in})} := \text{row}_v \left(B_{\text{in}}^{(k)} \right), \quad \mathbf{b}_v^{(k,\text{out})} := \text{row}_v \left(B_{\text{out}}^{(k)} \right)$$

- For weighted graphs, apply edge-thresholding:

$$A_{ij}^{(w)} = \begin{cases} 1 & \text{if } W_{ij} \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

or generalize the Hamming distance to quantized or fuzzy distance kernels between real-valued vectors.

Open Question: What analogues of Theorems 1–17 hold when directionality and/or weighting are introduced? Can uniqueness still be cleanly characterized via discrete dissimilarity measures?

5.1.1 Temporal HGM (Dynamic Graphs)

We extend HGM to evolving graphs by adding a time mode. Let $G_{t=1}^T$ be snapshots on a common labeled set $V = [N]$, with adjacencies $A^{(t)}$ and diameters $D_t = \text{diam}(G_t)$. Define exact- k reachability per snapshot

$$B^{(1,t)} := A^{(t)}, \quad B^{(k,t)} := \mathbf{1} \left[\sum_{s=1}^k (A^{(t)})^s > 0 \right] - \mathbf{1} \left[\sum_{s=1}^{k-1} (A^{(t)})^s > 0 \right] \quad (2 \leq k \leq D_t),$$

and set $B^{(k,t)} \equiv 0$ for $k > D_t$ so a uniform $D := \max_t D_t$ works across time.

The **temporal HGM tensor** is the fourth-order binary tensor

$$\mathbb{B} \in \{0, 1\}^{N \times N \times D \times T}, \quad \mathbb{B}_{ijkt} := B_{ij}^{(k,t)}.$$

Proposition 5.1. For labeled sequences $G_{1:T}$ and $H_{1:T}$ of equal length T ,

$$d_{\text{dyn}}(G_{1:T}, H_{1:T}) := \|\mathbb{B}^G - \mathbb{B}^H\|_1 = \sum_{t=1}^T \sum_{k=1}^D \sum_{i,j} \mathbf{1}[\mathbb{B}_{ijkt}^G \neq \mathbb{B}_{ijkt}^H]$$

is a **metric** on labeled temporal graphs. A normalized form $\bar{d}_{\text{dyn}} = \|\mathbb{B}^G - \mathbb{B}^H\|_1 / (N(N-1)DT) \in [0, 1]$ aids cross-size/horizon comparison.

Proof. The ℓ_1 norm on tensors obeys nonnegativity, symmetry, and the triangle inequality; positivity holds because the exact- k slices at each t determine G_t on the common label set. ■

Unlabeled sequences. For isomorphism classes, act with a **single** permutation on all times:

$$d_{\text{dyn,iso}}([G_{1:T}], [H_{1:T}]) := \min_{\pi \in S_N} \|\mathbb{B}^G - (\pi \mathbb{B}^H)\|_1, \quad (\pi \mathbb{B})_{ijkl} := \mathbb{B}_{\pi(i)\pi(j)kt}.$$

Then $d_{\text{dyn,iso}}$ is a **metric on isomorphism classes of temporal graphs** (zero only for timewise isomorphic sequences; triangle by composing near-minimizers). *Remark.* Allowing a different π_t per time gives a permutation-invariant dissimilarity but is not a metric on time-consistent orbits.

Temporal centrality and change diagnostics For each t , per-scale/node Hamming centrality is as before:

$$\text{HC}^{(k,t)}(v) = \frac{1}{N-1} \sum_{u \neq v} \text{Ham}(b_v^{(k,t)}, b_u^{(k,t)}), \quad b_v^{(k,t)} := \text{row } v \text{ of } B^{(k,t)}.$$

Define **temporal variation** and **trend** of structural uniqueness:

$$\text{TV}^{(k)}(v) := \sum_{t=2}^T |\text{HC}^{(k,t)}(v) - \text{HC}^{(k,t-1)}(v)|, \quad \text{trend}^{(k)}(v) := \frac{1}{T-1} \sum_{t=2}^T (\text{HC}^{(k,t)}(v) - \text{HC}^{(k,t-1)}(v)).$$

Streaming/online updates (implementation note) For small edge updates between G_t and G_{t+1} , update only rows/columns of $B^{(k,t)}$ whose entries can flip (frontier reuse across k). Popcount-based XOR kernels on packed bitboards keep pairwise Hamming costs at $\tilde{O}(N^2/w)$ per affected scale (Appendix B).

Stability across time (edge updates) Let G_t and G_{t+1} differ by r edge toggles and let Δ be the max degree in $G_t \cup G_{t+1} + 1$. Writing $M_s := \max_x |B_s(x)|$ for balls in graph distance, for each k ,

$$|E_k(G_{t+1}) - E_k(G_t)| \leq 2r M_{k-1}^2, \quad E_k(G_t) := \|B^{(k,t)}\|_F^2,$$

hence

$$\|\mathbb{B}^{(\cdot,t+1)} - \mathbb{B}^{(\cdot,t)}\|_F^2 \leq 2r \sum_{k=1}^D M_{k-1}^2.$$

For $\Delta \geq 3$, $M_s \leq \frac{\Delta}{\Delta-2}(\Delta-1)^s$ gives

$$\|\mathbb{B}^{(\cdot,t+1)} - \mathbb{B}^{(\cdot,t)}\|_F \leq \frac{\sqrt{2r} \Delta}{\Delta-2} \left(\sum_{k=1}^D (\Delta-1)^{2(k-1)} \right)^{1/2}.$$

Proof. Each edge toggle can only flip exact- k entries within $(k-1)$ steps of its endpoints (as in the static edge-flip analysis); this gives $2M_{k-1}^2$ flips per k . Summing over k and over r toggles yields the bounds; the degree-based estimate follows from the branching bound on s . ■

Remark 5.2 (Time-respecting variant.). For edge-timestamped temporal networks, one may replace per-snapshot reachability with **time-respecting paths** (nondecreasing timestamps). Let $\text{dist}_{\text{temp}}(i, j; \tau)$ be the minimum *elapsed* time to reach j from i under time-respecting walks; an exact-elapsed-time tensor $\tilde{\mathbb{B}}_{ijk\tau}$ (with k hops and elapsed time τ) yields a parallel HGM construction. We leave the temporal-path variant's bounds and algorithms to future work.

5.2 Sketching and Approximation

Motivation: Despite tractability, computing full $\mathcal{O}(N^2)$ Hamming distances becomes prohibitive at the million-node scale.

Proposal:

- Maintain the row weights $w_v^{(k)} := \|b_v^{(k)}\|_0$ alongside an s -sample Min-Hash sketch per row.

- Estimate Jaccard similarity $\hat{J}(\mathbf{b}_v^{(k)}, \mathbf{b}_u^{(k)})$ from the s hashes; then recover an unbiased intersection estimate

$$\hat{I} = \frac{\hat{J}}{1 + \hat{J}}(w_v^{(k)} + w_u^{(k)}),$$

and define

$$\widehat{\text{Ham}}(\mathbf{b}_v^{(k)}, \mathbf{b}_u^{(k)}) = (w_v^{(k)} + w_u^{(k)}) - 2\hat{I}.$$

- Since $\hat{J}$ is the mean of s i.i.d. Bernoulli indicators, for any $\varepsilon > 0$,

$$\Pr[|\hat{J} - J| > \varepsilon] \leq 2 \exp(-2s\varepsilon^2),$$

and $\widehat{\text{Ham}}$ inherits concentration via the above linear transform.

Open Problem: What is the minimal sketch dimension s required to preserve graph-level dispersion $\Psi^{(k)}(G)$ within δ -error with high probability?

5.3 Cross-Graph Comparison and Alignment

Motivation: Comparing two networks (e.g., ontologies, brain graphs, social networks) requires a notion of inter-graph correspondence beyond isomorphism [8].

Proposal:

- Use uniqueness distributions $\{\mu_v(G)\}$, $\{\mu_u(G')\}$ as embedding signatures
- Define matching via optimal transport:

$$\min_{\pi \in \Pi(V, V')} \sum_{(v, u)} \pi(v, u) \cdot W_1(\mu_v, \mu_u)$$

where $\Pi(V, V')$ is the set of doubly stochastic maps. This enables alignment without topological isomorphism-matching by role rather than identity.

Potential Applications: Cross-species connectome comparison; multi-lingual knowledge graph alignment; adversarial network mapping.

5.4 Theoretical Characterization of Graph Classes

Motivation: We have seen that Hamming profiles behave predictably on extremal graph families. But the taxonomy is incomplete.

Proposal:

- Define **Hamming-stable classes**: families for which node-wise dissimilarity is invariant under class-preserving transformations (e.g., adding self-loops in regular graphs)
- Investigate relationships between uniqueness spectra and known invariants:
 - Degree sequences
 - Spectral signatures (Laplacian, adjacency) [9]
 - Treewidth, genus
- Explore **inverse problems**: Given a uniqueness spectrum $\{\mu_v\}$, can one reconstruct a graph up to automorphism?

5.5 Tensor Methods for Structural Analysis

The tensor representation opens several research directions:

1. **Multi-way Spectral Analysis**: Apply tensor eigendecomposition to $\mathcal{B}$ to identify multi-scale communities
2. **Compressed Sensing**: Use tensor completion to infer missing scales from partial observations
3. **Cross-Graph Alignment**: Use tensor factorization for multi-graph matching problems

5.6 Summary

The Hamming Graph Metrics framework opens new pathways for analyzing structural differentiation, redundancy, and singularity in graphs. Beyond their immediate applications, HGMs suggest a broader research program:

- Establishing structural information geometry on graphs
- Characterizing dynamics via dissimilarity flows
- Embedding graphs in uniqueness-induced metric spaces

These directions integrate ideas from information theory, algebraic graph theory, approximate algorithms, and optimal transport [12, 13]—and promise fertile ground for future work in both theoretical and applied settings.

6 Conclusion

We have introduced Hamming Graph Metrics (HGM) as a theoretically grounded, tensor-based framework for measuring structural uniqueness in graphs. Unlike classical centrality measures, which quantify node importance through frequency, distance, or flow, HGM focuses on dissimilarity in structural configuration, using Hamming distances between binary path reachability vectors as its foundational primitive.

At the core of this framework is the empirical distribution of pairwise Hamming distances across all node pairs, which we extend across scales, define general functionals over dissimilarity distributions, introduce aggregation mechanisms at the graph level, and establish several new theorems characterizing extremal behavior in canonical graph families.

The theoretical foundation rests on several pillars:

- **Formal generalization:** Binary reachability distributions are treated as elements of probability space, allowing the application of convex functionals, entropy measures, and information divergence
- **Graph-theoretic bounds:** New inequalities and monotonicity theorems clarify how uniqueness behaves under connectivity, regularity, and symmetry constraints

Though the focus was not algorithmic, we showed that computing HGMs is tractable on real-world graphs of size $N \sim 10^5$ using bitwise operations and early termination strategies. This addresses prior critiques that dissimilarity-based metrics may be computationally prohibitive.

We also proposed several concrete directions for future work, including:

- Extension to directed, weighted, and evolving graphs
- Approximation via sketching and sampling
- Cross-graph matching via uniqueness alignment
- Inverse problems and structural reconstruction from uniqueness fields

Overall, Hamming Graph Metrics offer a multi-scale, intrinsic, and interpretable geometry over the space of structural patterns within graphs. By quantifying how structural patterns are distributed throughout a network rather than merely identifying central or connected nodes, HGM complements existing graph tools and opens the door to finer-grained structural analysis across domains.

The framework’s emphasis on complete distributions rather than summary statistics provides a richer view of network organization, revealing patterns

like bimodality in community structure, scale-dependent organization, and structural phase transitions that are invisible to traditional approaches. This distributional perspective, combined with rigorous theoretical foundations and demonstrated scalability, positions Hamming Graph Metrics as a valuable addition to the toolkit for understanding complex networks.

.1 Proofs and Technical Details

Notation. We write $\text{Ham}(\cdot, \cdot)$ for Hamming distance. Throughout, Hamming centrality is **normalized**:

$$\text{HC}^{(k)}(v) := \frac{1}{N-1} \sum_{u \neq v} \text{Ham}(b_v^{(k)}, b_u^{(k)}).$$

(Older unscaled variants are denoted $\text{HC}_{\text{raw}}^{(k)} = (N-1) \text{HC}^{(k)}$.)

.2 Proof of Proposition 2.3 (Star Graph Asymmetry)

Proof. At $k = 1$, $b_c^{(1)}$ has ones in all leaf positions and zero at c ; each leaf $b_\ell^{(1)}$ has a single one at c and zeros elsewhere. Thus $\text{Ham}(b_c^{(1)}, b_\ell^{(1)}) = N$ and $\text{Ham}(b_\ell^{(1)}, b_{\ell'}^{(1)}) = 0$ for distinct leaves $\ell \neq \ell'$. Hence

$$\text{HC}^{(1)}(c) = N, \quad \text{HC}^{(1)}(\ell) = \frac{N}{N-1}.$$

f For the distribution over ordered pairs,

$$\mu_{S_N}^{(1)} = \frac{(N-1)(N-2)}{N(N-1)} \delta_0 + \frac{2(N-1)}{N(N-1)} \delta_N = \frac{N-2}{N} \delta_0 + \frac{2}{N} \delta_N.$$

■

.3 Proofs on Monotonicity

Proposition .1. *Let G be connected with diameter D . Then for all $k \geq D$ and all $v \in V$,*

$$\text{HC}^{(k+1)}(v) \leq \text{HC}^{(k)}(v),$$

with equality for every $k \geq D+1$ (both sides equal 0).

Proof. By definition of exact- k slices, $B_{ij}^{(k)} = \mathbf{1}\{\text{dist}(i, j) = k\}$. If $k \geq D+1$, no ordered pair (i, j) has $\text{dist}(i, j) = k$, so $B^{(k)} \equiv 0$ and thus $b_v^{(k)} = \mathbf{0}$ for

every v . Therefore $\text{HC}^{(k)}(v) = \sum_{u \neq v} \text{Ham}(b_v^{(k)}, b_u^{(k)}) = 0$ for all $k \geq D + 1$. For $k = D$, $B^{(D+1)} \equiv 0$ while $B^{(D)}$ may be nonnegative; hence $\text{HC}^{(D+1)}(v) = 0 \leq \text{HC}^{(D)}(v)$. The claimed inequality for all $k \geq D$ follows. $\blacksquare$

Remark .2. The statement above is tight in general: without additional structure, $\text{HC}^{(k)}$ need not be monotone for $k < D$; exact- k shells can grow and shrink before saturation (e.g., on paths/cycles).

.3.1 A nontrivial tail monotonicity under mild structure

We first relate the **mean** pairwise Hamming at scale k to column sums of $B^{(k)}$.

Lemma .3. *Let $B^{(k)} \in \{0, 1\}^{N \times N}$ have column sums $s_j(k) = \sum_u B_{uj}^{(k)}$. The average over unordered row pairs of $\text{Ham}(\cdot, \cdot)$ equals*

$$\overline{H}^{(k)} = \frac{2}{N(N-1)} \sum_{j=1}^N s_j(k) (N - s_j(k)).$$

Proof. For a fixed column j , exactly $s_j(k) (N - s_j(k))$ unordered row pairs disagree in that coordinate; summing over j and dividing by the number of unordered pairs gives the formula. $\blacksquare$

We can now state a sufficient condition that is met in many graph families (trees, distance-regular graphs past the mode, vertex-transitive graphs past the mode, many expander families).

Theorem .4. *Assume there exists $k_0 < D$ such that for every vertex j the sphere sizes*

$$s_j(k) := |\{u : \text{dist}(u, j) = k\}|$$

satisfy $s_j(k+1) \leq s_j(k)$ and $s_j(k) \leq N/2$ for all $k \geq k_0$. Then the graph-average mean pairwise Hamming $\overline{H}^{(k)}$ is nonincreasing for $k \geq k_0$. Consequently, the average node centrality $\frac{1}{N} \sum_v \text{HC}^{(k)}(v) = (N-1) \overline{H}^{(k)}$ is nonincreasing for $k \geq k_0$.

Proof. By Lemma~.3,

$$\overline{H}^{(k)} = \frac{2}{N(N-1)} \sum_{j=1}^N f(s_j(k)) \quad \text{with} \quad f(s) = s(N-s).$$

On $[0, N/2]$, f is increasing. For $k \geq k_0$, $s_j(k+1) \leq s_j(k) \leq N/2$ for each j , so $f(s_j(k+1)) \leq f(s_j(k))$. Summation and scaling preserve the inequality, hence $\overline{H}^{(k+1)} \leq \overline{H}^{(k)}$ for all $k \geq k_0$. The identity $\frac{1}{N} \sum_v \text{HC}^{(k)}(v) = (N-1) \overline{H}^{(k)}$ gives the second claim. ■

Corollary .5. *If G is vertex-transitive and the common sphere sizes $n_k = |S(v, k)|$ become nonincreasing for $k \geq k_0$ with $n_k \leq N/2$ (e.g., past the mode of (n_k)), then $\overline{H}^{(k)}$ and the average $\text{HC}^{(k)}$ are nonincreasing for $k \geq k_0$.*

.4 Separation on $G(n, p)$

Proposition .6. *Fix $p \in (0, 1)$. Let $G, H \sim G(n, p)$ be independent. Then as $n \rightarrow \infty$, with probability $\rightarrow 1$:*

(a) *For any non-constant admissible Φ , the graph-level descriptor*

$$\overline{\Phi}(G) := \frac{1}{D(G)} \sum_{k=1}^{D(G)} \Phi(\mu_G^{(k)})$$

differs from $\overline{\Phi}(H)$; i.e., $\overline{\Phi}(G) \neq \overline{\Phi}(H)$.

(b) *The tensor fingerprint differs: $\text{FP}(G) \neq \text{FP}(H)$. Equivalently, at least one k has $E_k(G) \neq E_k(H)$ (and mode spectra need not be invoked).*

Proof. Sketch. For each fixed $k \leq c \log n$ (for any fixed $c > 0$), the unordered-pair count vector of exact- k distances is a Lipschitz function of the $\binom{n}{2}$ independent edges; bounded-difference/McDiarmid inequalities give concentration around the mean. For two independent graphs G, H , anti-concentration implies

$$\Pr [\mu_G^{(k)} = \mu_H^{(k)}] = o(1)$$

(and likewise for the ordered counts E_k ; the ordered/unordered choice only changes a factor of 2). A union bound over all $k \leq D(n) = O(\log n)$ yields

$$\Pr [\forall k \leq D(n) : \mu_G^{(k)} = \mu_H^{(k)}] = o(1),$$

so for some k we have $\mu_G^{(k)} \neq \mu_H^{(k)}$ and $E_k(G) \neq E_k(H)$ a.a.s. For (a), TV-continuity and non-constancy of Φ imply $\overline{\Phi}(G) \neq \overline{\Phi}(H)$ a.a.s. For (b), differing (E_k) forces $\text{FP}(G) \neq \text{FP}(H)$. ■

Proposition .7. *Let G and G' be graphs on the same labeled vertex set that differ by toggling a single undirected edge $\{x, y\}$. Let $B^{(k)}$ and $B'^{(k)}$ be their exact- k reachability matrices, and define*

$$M_r := \max_v |B_r(v)|, \quad B_r(v) := \{u : \text{dist}_{G \cup G'}(u, v) \leq r\}.$$

Then for every $k \geq 1$,

$$\|B'^{(k)} - B^{(k)}\|_1 = \|B'^{(k)} - B^{(k)}\|_F^2 \leq 2 M_{k-1}^2,$$

hence, writing $E_k(G) := \|B^{(k)}\|_F^2$,

$$|E_k(G') - E_k(G)| \leq 2 M_{k-1}^2.$$

Consequently, for the HGM tensor slices,

$$\|\mathcal{B}_G(:, :, k) - \mathcal{B}_{G'}(:, :, k)\|_1 \leq 2 M_{k-1}^2, \quad \sum_{k=1}^D \|\mathcal{B}_G(:, :, k) - \mathcal{B}_{G'}(:, :, k)\|_1 \leq 2 \sum_{k=1}^D M_{k-1}^2.$$

If G and G' differ by r edge toggles, the right-hand sides multiply by r .

Proof. Any newly created (or destroyed) exact- k connection (i, j) must have all shortest $i \rightarrow j$ paths in G' use the toggled edge $\{x, y\}$ exactly once; otherwise the shortest length is unchanged. Such a path decomposes as

$$i \xrightarrow{\leq k-1} x \xrightarrow{1} y \xrightarrow{\leq k-1} j \quad \text{or} \quad i \xrightarrow{\leq k-1} y \xrightarrow{1} x \xrightarrow{\leq k-1} j,$$

with the two “legs” having lengths summing to $k-1$. Thus the set of ordered pairs (i, j) whose exact- k status can change is contained in

$$B_{k-1}(x) \times B_{k-1}(y) \cup B_{k-1}(y) \times B_{k-1}(x),$$

which has size at most $2 |B_{k-1}(x)| |B_{k-1}(y)| \leq 2 M_{k-1}^2$. Since entries of $B^{(k)}$ are binary, the number of flips equals both the ℓ_1 and the squared Frobenius norm of the difference, proving the first two inequalities. Summing over k and using linearity over r toggles gives the remaining bounds. ■

Corollary .8. *If the maximum degree in $G \cup G'$ is $\Delta \geq 3$, then for all $r \geq 0$,*

$$M_r \leq 1 + \Delta \sum_{t=0}^{r-1} (\Delta - 1)^t \leq \frac{\Delta}{\Delta - 2} (\Delta - 1)^r,$$

and therefore

$$\sum_{k=1}^D \|\mathcal{B}_G(:, :, k) - \mathcal{B}_{G'}(:, :, k)\|_1 \leq \frac{2 \Delta^2}{(\Delta - 2)^2} \sum_{k=1}^D (\Delta - 1)^{2(k-1)}.$$

Remark .9. For directed graphs, the same argument yields $\|B'^{(k)} - B^{(k)}\|_1 \leq M_{k-1}^{\text{out}}(x) M_{k-1}^{\text{in}}(y) + M_{k-1}^{\text{out}}(y) M_{k-1}^{\text{in}}(x)$, with obvious definitions of in/out balls.

.4.1 Watts–Strogatz phase transition (derivation)

Theorem .10. *Let a Watts–Strogatz (WS) graph on N vertices start from a ring lattice with even degree d (so $m = Nd/2$ undirected edges). Each edge is independently **rewired** with probability $\beta = \beta(N)$ to a uniformly random new endpoint (avoiding loops/multi-edges). Then:*

1. (**Onset of shortcuts**) *If β_c denotes the threshold for the appearance of any rewired edge (shortcut), then*

$$\beta_c \asymp \frac{1}{N}.$$

More precisely, if $N\beta \rightarrow 0$ then with high probability (whp) there are no shortcuts; if $N\beta \rightarrow \infty$ then whp there are $\rightarrow \infty$ shortcuts.

2. (**Distance regimes**)

- *If $N\beta \rightarrow 0$, whp the graph coincides with the base ring lattice, so average distance is $\Theta(N/d)$.*
- *If $N\beta \rightarrow \infty$ with fixed d , then whp the random rewires form a sparse long-range overlay comparable to $G(N, p_{\text{eff}})$ with*

$$p_{\text{eff}} \approx \frac{2\beta d}{N},$$

and the average distance drops to $O(\log N)$ (small-world regime).

(1) *Onset.* Let $X \sim \text{Binomial}(m, \beta)$ be the number of rewired edges; $m = Nd/2 = \Theta(N)$. Then

$$\mathbb{E}[X] = \beta m = \Theta(\beta N), \quad \mathbb{P}[X = 0] = (1 - \beta)^m \leq \exp(-\beta m).$$

If $N\beta \rightarrow 0$, then $\beta m \rightarrow 0$ and $\mathbb{P}[X = 0] \rightarrow 1$ (no shortcuts whp). If $N\beta \rightarrow \infty$, then $\beta m \rightarrow \infty$ and $\mathbb{P}[X = 0] \rightarrow 0$, while Chernoff bounds give $X \rightarrow \infty$ in probability. Hence $\beta_c \asymp 1/N$.

(2) Distances. When $N\beta \rightarrow 0$, whp $X = 0$ and the graph is the base ring lattice with average distance $\Theta(N/d)$.

When $N\beta \rightarrow \infty$, the rewired endpoints are uniform over vertices (up to constant factors from local exclusions), so the rewires approximate an Erdős–Rényi overlay with edge-probability

$$p_{\text{eff}} = \frac{\text{expected \# rewired edges}}{\binom{N}{2}} \approx \frac{\beta(Nd/2)}{\binom{N}{2}} \approx \frac{2\beta d}{N}.$$

With $N\beta \rightarrow \infty$ and fixed d , we have $Np_{\text{eff}} \rightarrow \infty$; the random overlay alone has logarithmic average distance via standard branching-process heuristics (BFS grows by factor $\approx Np_{\text{eff}}$ per layer until covering N). Adding the ring edges only helps, so the combined graph has $O(\log N)$ average distance. ■

Remark .11. The threshold $\beta_c \asymp 1/N$ is the *first-shortcut* threshold. Logarithmic distances require a diverging number of shortcuts ($N\beta \rightarrow \infty$); for constant $\beta > 0$, the overlay has $\Theta(N)$ long edges and typical distances are $O(\log N)$.

.4.2 Smoothing via spectral gap (expander calculation)

We make precise the “uniqueness smoothing” statement using the Poincaré (spectral-gap) inequality. We treat the **normalized** Hamming centrality

$$f(v) := \text{HC}^{(k)}(v) = \frac{1}{N-1} \sum_{u \neq v} \text{Ham}(b_v^{(k)}, b_u^{(k)}),$$

and then note the unnormalized variant.

Theorem .12. *Let G be a connected d -regular graph on N vertices with random-walk matrix $P = A/d$ and spectral gap $\gamma := 1 - \lambda_2(P) > 0$. For any fixed scale $k \geq 1$, writing $M_k := \max_x |S(x, k)|$ (size of the distance- k sphere),*

$$\text{Var}(f) \leq \frac{d}{\gamma} \frac{M_k^2}{(N-1)^2}.$$

*Equivalently, for the **unnormalized** centrality $F(v) := \sum_{u \neq v} \text{Ham}(b_v^{(k)}, b_u^{(k)}) = (N-1)f(v)$,*

$$\text{Var}(F) \leq \frac{d}{\gamma} M_k^2.$$

Proof. (**Edgewise Lipschitz.**) For any u, v, z ,

$$|\text{Ham}(b_v^{(k)}, b_z^{(k)}) - \text{Ham}(b_u^{(k)}, b_z^{(k)})| \leq \text{Ham}(b_v^{(k)}, b_u^{(k)}) \leq |S(v, k)| + |S(u, k)| \leq 2M_k.$$

Averaging over $z \neq u, v$ and dividing by $N - 1$ gives

$$|f(v) - f(u)| \leq \frac{2M_k}{N - 1} \quad \text{for every edge } (u, v).$$

(**Dirichlet form and Poincaré.**) For d -regular G ,

$$\mathcal{E}(f, f) := \frac{1}{2N} \sum_{(u, v) \in E} (f(u) - f(v))^2, \quad \text{Var}(f) \leq \frac{1}{\gamma} \mathcal{E}(f, f).$$

Using the edgewise bound and $|E| = dN/2$,

$$\mathcal{E}(f, f) \leq \frac{1}{2N} \cdot \frac{dN}{2} \cdot \left(\frac{2M_k}{N - 1} \right)^2 = \frac{d M_k^2}{(N - 1)^2}.$$

Combine with Poincaré to obtain the stated variance bound. For $F = (N - 1)f$, variances scale by $(N - 1)^2$. ■

Corollary .13. *If the maximum degree Δ of G is at most $\Delta \geq 3$, then*

$$M_k \leq 1 + \Delta \sum_{t=0}^{k-1} (\Delta - 1)^t \leq \frac{\Delta}{\Delta - 2} (\Delta - 1)^k,$$

hence

$$\text{Var}(f) \leq \frac{d}{\gamma} \frac{1}{(N - 1)^2} \left(\frac{\Delta}{\Delta - 2} \right)^2 (\Delta - 1)^{2k}.$$

For F , remove the $(N - 1)^{-2}$ factor.

Remark .14. For **non-regular** graphs, replace P by the lazy random walk or use the normalized Laplacian $\mathcal{L} = I - D^{-1/2} A D^{-1/2}$; the same argument yields $\text{Var}(f) \lesssim \gamma^{-1} \cdot \frac{1}{|V|} \sum_{(u, v) \in E} (f(u) - f(v))^2$, and the edgewise Lipschitz bound now depends on local sphere sizes near the edge endpoints.

4.3 Technical Lemmas

Lemma .15. *For any simple unweighted graph,*

$$\mathbf{1}\left[\sum_{t=1}^k A^t > 0\right]_{ij} = 1 \iff \text{dist}(i, j) \leq k.$$

Moreover, on bipartite graphs, every walk between i and j has length congruent to $\text{dist}(i, j) \pmod{2}$.

Proof. If $\text{dist}(i, j) \leq k$, a simple path of length $\leq k$ exists; its length $t \leq k$ contributes $(A^t)_{ij} > 0$, so the sum $\sum_{t=1}^k (A^t)_{ij}$ is positive. Conversely, if $\sum_{t=1}^k (A^t)_{ij} > 0$, then for some $t \leq k$ there is a walk of length t ; shortcutting repeated vertices yields a simple path of length $\leq t \leq k$. The parity clause follows because walks on bipartite graphs alternate sides; thus all i – j walks have the same parity as $\text{dist}(i, j)$. ■

Lemma .16. *Let G be distance-regular with intersection numbers $\{a_i, b_i, c_i\}_{i=0}^D$, adjacency A , and distance matrices A_i (so $A_0 = I$, $A_1 = A$, and $A_i = B^{(i)}$ for $i \geq 1$). Then:*

1. *(Three-term matrix recurrence)*

$$A A_i = b_{i-1} A_{i-1} + a_i A_i + c_{i+1} A_{i+1} \quad (0 \leq i \leq D),$$

with $b_{-1} = c_{D+1} = 0$.

2. *(Polynomial dependence)*

There exist polynomials p_i with $\deg p_i = i$ such that $A_i = p_i(A)$, with $p_0 = 1$, $p_1 = x$, and

$$x p_i(x) = b_{i-1} p_{i-1}(x) + a_i p_i(x) + c_{i+1} p_{i+1}(x).$$

3. *(Bose–Mesner algebra)*

The matrices $\{A_0, \dots, A_D\}$ span a $(D+1)$ -dimensional commutative algebra (the Bose–Mesner algebra); in particular, all A_i commute and are simultaneously diagonalizable. See [11].

Proof. By distance-regularity, for any vertex at distance i from a basepoint, the numbers of neighbors at distances $i-1, i, i+1$ depend only on i , giving the matrix identity $A A_i = b_{i-1} A_{i-1} + a_i A_i + c_{i+1} A_{i+1}$. Inductively, this produces polynomials p_i with $A_i = p_i(A)$ and the stated scalar recurrence. Since each

A_i is a polynomial in A , we have $A_i A_j = p_i(A) p_j(A) = p_j(A) p_i(A) = A_j A_i$, so the span of $\{A_i\}$ is a $(D+1)$ -dimensional commutative algebra containing $I = A_0$, and all A_i are simultaneously diagonalizable. ■

Lemma .17. *For any graph G and scale k ,*

$$0 \leq \text{Ham}(b_v^{(k)}, b_u^{(k)}) \leq N.$$

The upper bound is tight (e.g., $K_{m,n}$ at $k = 1$ yields $\text{Ham} = m + n = N$ across parts).

.4.4 Hypercube vs. Sphere

Lemma .18. *Let Q_n be the n -cube and fix u, v with Hamming distance $h = \text{dist}(u, v)$. Let $S(u, k) = \{w : \text{dist}(u, w) = k\}$ and $b_u^{(k)} = \mathbf{1}_{S(u, k)} \in \{0, 1\}^{2^n}$. Then*

$$|S(u, k)| = \binom{n}{k}, \quad |S(u, k) \cap S(v, k)| = \begin{cases} \binom{h}{h/2} \binom{n-h}{k-h/2}, & \text{if } h \text{ is even and } k \geq h/2, \\ 0, & \text{if } h \text{ is odd or } k < h/2. \end{cases}$$

Consequently,

$$\text{Ham}(b_u^{(k)}, b_v^{(k)}) = 2 \left(\binom{n}{k} - \mathbf{1}_{\{h \text{ even}, k \geq h/2\}} \binom{h}{h/2} \binom{n-h}{k-h/2} \right).$$

Proof. Write $u = (0, \dots, 0)$, and let v differ from u in the first h coordinates. A node w lies in $S(u, k)$ iff it differs from u in exactly k coordinates. Among the h differing coordinates, let t equal v ; among the remaining $n - h$ coordinates, choose $k - t$ to flip, giving $\binom{h}{t} \binom{n-h}{k-t}$ options with $\text{dist}(u, w) = k$. We have $\text{dist}(v, w) = (h - t) + (k - t)$, so requiring $\text{dist}(v, w) = k$ forces $h - 2t = 0$, i.e., $t = h/2$ (hence h even) and $k \geq h/2$. The intersection count follows. Finally,

$$\text{Ham}(b_u^{(k)}, b_v^{(k)}) = |S(u, k) \triangle S(v, k)| = |S(u, k)| + |S(v, k)| - 2|S(u, k) \cap S(v, k)|,$$

and $|S(u, k)| = |S(v, k)| = \binom{n}{k}$. ■

Lemma .19. *Fix $k \geq 1$. If $|S(v, k)|$ is constant over all v (e.g., in vertex-transitive graphs), then for all u, v ,*

$$\text{Ham}(b_u^{(k)}, b_v^{(k)}) \in \{0, 2, 4, \dots, 2|S(\cdot, k)|\}.$$

Proof. If every row $b_v^{(k)}$ has weight $w := |S(\cdot, k)|$, then

$$\|x - y\|_1 = 2(w - |\text{supp}(x) \cap \text{supp}(y)|)$$

is even. ■

Lemma .20. *If G is connected with diameter D , then $B^{(k)} \equiv 0$ for all $k \geq D + 1$, hence $\mu_G^{(k)} = \delta_0$ for $k \geq D + 1$.*

Proof. By definition of diameter, no pair has shortest-path distance exactly k once $k \geq D + 1$. ■

Appendix A Computational Details

Although this work is primarily theoretical, it is important that Hamming Graph Metrics (HGM) admit efficient evaluation. We summarize asymptotic costs and the implementation choices that make the framework practical on large sparse graphs.

A.1 Complexity Overview

Let $(G=(V,E))$ be an unweighted, undirected graph with $(|V|=N)$, $(|E|=M)$, and diameter (D) .

A.1.1 1) Distances and exact-(k) slices

Compute all-pairs shortest-path **distances** by running BFS from each source:

Time = $\mathcal{O}(N(N + M))$, Space = $\mathcal{O}(N)$ (working).

Define exact-(k) slices by

$$B_{ij}^{(k)} = \mathbf{1}\{\text{dist}(i, j) = k\}, \quad k = 1, \dots, D,$$

which can be populated in $\mathcal{O}(N^2)$ total once distances are known.

A.1.2 2) Node-level summaries in $\mathcal{O}(D N^2)$ bit-parallel portions $\mathcal{O}(D N^2/w)$

For a fixed scale (k), let $B = B^{(k)}$ and let $s \in \mathbb{N}^N$ be its column sums, $s_j = \sum_u B_{uj}$. Then for all nodes simultaneously,

$$\text{HC}^{(k)}(v) = \sum_{u \neq v} \text{Ham}(b_v^{(k)}, b_u^{(k)}) = \left(\sum_{j=1}^N s_j \right) + [B(N\mathbf{1} - 2s)]_v.$$

Thus:

- compute (s) via bit-packed popcounts in $\mathcal{O}(N^2/w)$;
- form $c = N\mathbf{1} - 2s$ in $\mathcal{O}(N)$;
- multiply (B,c) in $\mathcal{O}(N^2)$ (upper bound), or $\mathcal{O}(\text{nnz}(B))$ if sparsity permits.

Per scale: $\mathcal{O}(N^2)$ (with the popcount portions $\mathcal{O}(N^2/w)$); across all (k): $\mathcal{O}(D N^2)$.

A.1.3 3) Graph-to-graph HGM distance in $\mathcal{O}(D N^2/w)$

For labeled graphs (G,H),

$$d_{\text{HGM}}(G, H) = \sum_{k=1}^D \|B_G^{(k)} - B_H^{(k)}\|_1$$

is evaluated by XOR+(POPCOUNT) over bit-packed slices in $\mathcal{O}(D N^2/w)$.

Equivalent distance-matrix formulation. Because for each ordered pair ((i,j)) exactly one (k) satisfies $(B^{\wedge}\{(k)\})_{ij}=1$,

$$\sum_{k=1}^D |B_G^{(k)}(i, j) - B_H^{(k)}(i, j)| = \begin{cases} 0, & \text{dist}_G(i, j) = \text{dist}_H(i, j), \\ 2, & \text{otherwise,} \end{cases}$$

hence

$$d_{\text{HGM}}(G, H) = 2 \# \{(i, j) : i \neq j, \text{dist}_G(i, j) \neq \text{dist}_H(i, j)\}.$$

Thus, once the two distance matrices are computed, a single $\mathcal{O}(N^2)$ pass suffices without materializing all slices.

A.1.4 4) Optional pairwise matrices (when explicitly needed)

If one forms the full pairwise Hamming matrix $(D^{\wedge\{(k)\}})$ with entries $D_{uv}^{(k)} = \text{Ham}(b_u^{(k)}, b_v^{(k)})$, the best straightforward bitset method costs $\mathcal{O}(N^3/w)$ per (k) (XOR+(POPCOUNT) for all pairs). This is **not** required for the node-level summaries or d_{HGM} computations above.

A.2 Bit-Parallel Representation

We work in the word-RAM model with machine word size (w) (e.g., $(w=64)$) and hardware POPCOUNT. Each row $b_v^{(k)} \in \{0, 1\}^N$ is stored in $\lceil N/w \rceil$ words. For bitsets (r, s) ,

$$\text{Ham}(r, s) = \sum_{t=1}^{\lceil N/w \rceil} \text{popcount}(r_t \oplus s_t).$$

This turns all bitwise portions of the algorithms above into $\mathcal{O}(N^2/w)$ passes per scale.

A.3 Practical Notes

- **Streaming over (k) :** to avoid storing $\mathcal{B}$ explicitly, accumulate (s) and the required functionals per scale while streaming rows produced by BFS.
- **Sparsity:** when many $(B^{\wedge\{(k)\}})$ are sparse (typical for small (k)), exploit $\text{nnz}(B^{(k)})$ in the (B, c) multiplication.
- **Parallelism:** BFS sources, per-scale passes, and (POPCOUNT) loops parallelize naturally across cores/GPUs.

A.4 Summary

Distances: $\mathcal{O}(N(N + M))$
Node summaries (all k): $\mathcal{O}(D N^2)$ (bitwise parts $\mathcal{O}(D N^2/w)$)
Graph–graph d_{HGM} : $\mathcal{O}(D N^2/w)$ (or $\mathcal{O}(N^2)$ via distances)

No Boolean matrix powers are used; bit-parallel XOR+(POPCOUNT) yields the (N^2/w) speedups on the bitwise portions.

References

- [1] Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, 31(4), 581-603.
- [2] Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5), 1170-1182.
- [3] Newman, M. E. J. (2010). *Networks: An Introduction*. Oxford University Press.
- [4] Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- [5] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Networks*, 1(3):215-239, 1979.
- [6] Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440-442.
- [7] Albert, R., Jeong, H., & Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794), 378-382.
- [8] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, 41(3), 15:1-15:58.
- [9] Vempala, S., & Wang, G. (2020). Spectral gap and vertex expansion. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, 1350-1363.
- [10] Kolda, T. G., & Bader, B. W. (2009). "Tensor decompositions and applications." *SIAM Review*, 51(3), 455-500.
- [11] A. E. Brouwer and W. H. Haemers, *Spectra of Graphs*. Universitext. Springer, New York, 2012. ISBN 978-1-4614-1938-9 (print); 978-1-4614-1939-6 (eBook).
- [12] C. Villani. *Optimal Transport: Old and New*. Grundlehren der mathematischen Wissenschaften, vol. 338. Springer, Berlin-Heidelberg, 2009. xix+973 pp. ISBN 978-3-540-71049-3. doi:10.1007/978-3-540-71050-9.
- [13] G. Peyré and M. Cuturi. *Computational Optimal Transport*. Foundations and Trends in Machine Learning, vol. 11, nos. 5-6, pp. 355-607. Now Publishers, 2019. ISBN 978-1-68083-551-9. doi:10.1561/22000000073.

- [14] R. Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics, vol. 169. Springer-Verlag, New York, 1997. xi+347 pp. ISBN 0-387-94846-5. doi:10.1007/978-1-4612-0653-8.
- [15] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Computer Science and Scientific Computing. Academic Press, Boston, 1990. 374 pp. ISBN 0-12-670230-6.

Table 4: What traditional graph metrics capture vs. what HGM distributions add.

Metric	Typical Output	What it Reveals	What HGM Distribution Adds
Degree	integers / histogram	Local connectivity of nodes	Distribution of <i>multi-scale</i> pattern differences across nodes and scales k
Betweenness	real in $[0, 1]$	Brokerage along shortest paths	How brokerage patterns differ across nodes at fixed k (disagreements of exact- k shells)
Closeness	real in $(0, 1]$	Average geodesic proximity	Whether “central” nodes have <i>similar</i> or <i>different</i> exact- k neighborhoods
Clustering coefficient	real in $[0, 1]$	Local triangle density	How triangle-rich regions appear as lower per-pair Hamming at $k = 2$
Modularity	real in $[0, 1]$	Community separability (global)	Whether separability manifests as <i>bimodality</i> or heavy tails in $\mu_G^{(k)}$ for some k
HGM entropy	bits	Diversity of exact- k structures	A scalar admissible functional of $\mu_G^{(k)}$; peaks indicate informative scales
HGM “bimodality”	modes of $\mu_G^{(k)}$	Natural partitions at scale k	Strength/sharpness of separation (collapse length