Locality-Aware Graph Rewiring in GNNs

GDL 2025
Group id: FBFRJT
Project id: LASER

Fabian Bosshard, Fanny Rorri, Laura Jimena Tagle Muñoz
{fabian.bosshard, fanny.rorri, jimena.tagle}@usi.ch

Abstract

Contents

List of Acronyms

GCN	Graph Convolutional Network
LASER	Locality-Aware Sequential Rewiring

1 Introduction

Locality-Aware Sequential Rewiring (LASER)

2 Background

Let $G = (V, E)$ be an undirected graph with the adjacency matrix $\underline{\boldsymbol{A}} \in \mathbb{R}^{n\times n}$

$$\underline{\boldsymbol{A}}_{uv} = \begin{cases} 1 & (u,v) \in E \\ 0 & (u,v) \notin E \end{cases} \tag{1}$$

The diagonal degree matrix $\underline{\boldsymbol{D}} \in \mathbb{R}^{n\times n}$ is defined by

$$\underline{\boldsymbol{D}}_{uv} = \begin{cases} d_u & u = v \\ 0 & u \neq v \end{cases} \tag{2}$$

i.e. $\underline{\boldsymbol{D}}$ simply places all node degrees on the diagonal.

2.1 normalized adjacency and multi-hop propagation

Definition 1. The symmetrically normalized adjacency matrix is

$$\boxed{\hat{\underline{\boldsymbol{A}}} = \underline{\boldsymbol{D}}^{-1/2}\underline{\boldsymbol{A}}\underline{\boldsymbol{D}}^{-1/2}} \tag{3}$$

or, entrywise,

$$\hat{\underline{\boldsymbol{A}}}_{uv} = \begin{cases} \dfrac{1}{\sqrt{d_u d_v}} & (u,v) \in E \\ 0 & (u,v) \notin E \end{cases}$$

✍

Proposition 1. multi-hop propagation. The entry $(\hat{\underline{\boldsymbol{A}}}^k)_{vu}$ can be computed explicitly as follows:

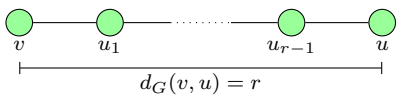
$$(\hat{\underline{\boldsymbol{A}}}^k)_{vu} = \sum_{\pi} \prod_{(x,y)\in E_\pi} \frac{1}{\sqrt{d_x d_y}} \tag{4}$$

the sum is over all walks $\pi = (v, \ldots, u)$ of length k from v to u and the product is over the edges $E_\pi = \{(v, u_1), \ldots, (u_{k-1}, u)\}$ on the walk. ◁

Corollary 2. Let $v, u \in V$ with $r = d_G(v, u)$, where $d_G(\cdot,\cdot)$ denotes the shortest-path distance. Assume there is exactly one path

$$(v, u_1, \ldots, u_{r-1}, u)$$

of length r between v and u:



Then

$$\begin{aligned}(\hat{\underline{\boldsymbol{A}}}^r)_{vu} &= \frac{1}{\sqrt{d_v d_{u_1}}} \cdot \prod_{i=1}^{r-2} \frac{1}{\sqrt{d_{u_i} d_{u_{i+1}}}} \cdot \frac{1}{\sqrt{d_{u_{r-1}} d_u}} \\ &= \frac{1}{\sqrt{d_v d_u}} \prod_{i=1}^{r-1} \frac{1}{d_{u_i}}\end{aligned} \tag{5}$$

◁

distance layers and layer degrees

Definition 2. For $\ell \in \mathbb{N}_0$, we define the distance-$(\ell+1)$ adjacency matrix $\underline{\boldsymbol{A}}_\ell \in \mathbb{R}^{n\times n}$ by

$$\left(\underline{\boldsymbol{A}}_\ell\right)_{uv} = \begin{cases} 1 & d_G(u,v) = \ell + 1 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $d_G(u,v)$ is the shortest-path distance. The corresponding layer degree of a node v at distance level ℓ is

$$d_{v,\ell} = \sum_{u \in V} (\underline{\boldsymbol{A}}_\ell)_{vu}, \tag{7}$$

i.e. the number of nodes at graph distance $\ell+1$ from v. Let $\underline{\boldsymbol{D}}_\ell$ be the diagonal matrix with $(\underline{\boldsymbol{D}}_\ell)_{vv} = d_{v,\ell}$. The normalized distance-$(\ell+1)$ adjacency is

$$\boxed{\hat{\underline{\boldsymbol{A}}}_\ell = \underline{\boldsymbol{D}}_\ell^{-1/2} \underline{\boldsymbol{A}}_\ell \underline{\boldsymbol{D}}_\ell^{-1/2}} \tag{8}$$

so that

$$(\hat{\underline{\boldsymbol{A}}}_\ell)_{uv} = \begin{cases} \frac{1}{\sqrt{d_{u,\ell} d_{v,\ell}}} & d_G(u,v) = \ell+1 \\ 0 & \text{otherwise} \end{cases}$$

Finally, we denote by

$$d_{\min} = \min_{v \in V} d_v \tag{9}$$

the minimum node degree in the graph.

2.2 graph Laplacian

Definition 3. The combinatorial graph Laplacian is

$$\underline{\boldsymbol{L}} = \underline{\boldsymbol{D}} - \underline{\boldsymbol{A}} \tag{10}$$

and the normalized graph Laplacian is

$$\boxed{\hat{\underline{\boldsymbol{L}}} = \underline{\boldsymbol{D}}^{-1/2} \underline{\boldsymbol{L}} \underline{\boldsymbol{D}}^{-1/2} \overset{(10)}{=} \underline{\boldsymbol{D}}^{-1/2} (\underline{\boldsymbol{D}} - \underline{\boldsymbol{A}}) \underline{\boldsymbol{D}}^{-1/2} \overset{(3)}{=} \underline{\boldsymbol{I}}_n - \hat{\underline{\boldsymbol{A}}}} \tag{11}$$

It is symmetric and positive semidefinite, and its eigenvalues satisfy

$$0 = \lambda_0 \le \lambda_1 \le \cdots \le \lambda_{n-1}$$

λ_1 is called the spectral gap. The number of zero eigenvalues (i.e., the multiplicity of the 0 eigenvalue) equals the number of connected components of the graph.

To understand Definition 3, consider a function $f : V \to \mathbb{R}$. Denote by $\vec{\boldsymbol{f}} \in \mathbb{R}^n$ the vector whose v-th entry is $f(v)$. Then

$$(\hat{\underline{\boldsymbol{L}}} \vec{\boldsymbol{f}})_v = f(v) - \frac{1}{\sqrt{d_v}} \sum_{(u,v) \in E} \frac{f(u)}{\sqrt{d_u}} \tag{12}$$

i.e., $(\hat{\underline{\boldsymbol{L}}} \vec{\boldsymbol{f}})_v$ is the value at v minus a degree-normalized average of the neighbors. This is why the Laplacian is often viewed as a discrete second derivative on the graph: it measures how much f at v deviates from its neighborhood. Another important identity is the quadratic form

$$\vec{\boldsymbol{f}}^\top \underline{\boldsymbol{L}} \vec{\boldsymbol{f}} = \frac{1}{2} \sum_{(u,v) \in E} \big(f(u) - f(v)\big)^2 \tag{13}$$

which shows that $\underline{\boldsymbol{L}}$ (and hence also $\hat{\underline{\boldsymbol{L}}}$) is positive semidefinite, since the right-hand side is always nonnegative. Moreover, (13) is small exactly when f varies slowly across edges, so the Laplacian encodes the smoothness of functions on the graph.

2.3 Cheeger inequality

The Cheeger inequality relates the spectral gap λ_1 to the Cheeger constant $h(G)$, which measures how difficult it is to separate the graph into two large pieces. It states, in particular, that

$$\tfrac{1}{2} h(G)^2 \le \lambda_1 \le 2h(G),$$

so a larger spectral gap implies that the graph is more "well-connected".

2.4 effective resistance

Definition 4. effective resistance. View each edge $(u,v) \in E$ as an electrical resistor of resistance $1\,\Omega$. The resulting network has a well-defined resistance between any two nodes.

For two nodes $s, t \in V$, the effective resistance $R(s,t)$ is defined as the voltage difference needed to send one unit of electrical current from s to t. It can be computed as

$$\boxed{R(s,t) = (\vec{\boldsymbol{e}}_s - \vec{\boldsymbol{e}}_t)^\top \underline{\boldsymbol{L}}^\dagger (\vec{\boldsymbol{e}}_s - \vec{\boldsymbol{e}}_t)} \tag{14}$$

where $\underline{\boldsymbol{L}}^\dagger$ is the Moore–Penrose pseudoinverse of $\underline{\boldsymbol{L}}$ and $\vec{\boldsymbol{e}}_v$ is the standard basis vector of vertex v.

Interpretation If the graph offers many short, parallel paths between s and t, then current can flow easily, so $R(s,t)$ is small. If there are few or long paths, the current is "bottlenecked" and $R(s,t)$ is large. Thus, effective resistance measures how "well-connected" two nodes are inside the global geometry of the graph.

Connection to random walks A random walk on G is the Markov chain that, from a node v, moves to a uniformly random neighbor of v. Its transition matrix is

$$\underline{\boldsymbol{P}} = \underline{\boldsymbol{D}}^{-1} \underline{\boldsymbol{A}} \tag{15}$$

so $\underline{\boldsymbol{P}}_{vu} = 1/d_v$ if $(v,u) \in E$.

For two nodes u, v, the commute time $\mathtt{CT}(u,v)$ is the expected number of steps for the random walk to start at u, reach v, and return to u again. It can be related to the effective resistance via

$$\mathtt{CT}(u,v) = 2|E| R(u,v) \tag{16}$$

giving a geometric interpretation of how "far apart" two nodes are in terms of random-walk behavior, i.e. two nodes have small commute time exactly when they have small effective resistance.

3 Theoretical analysis

3.1 Dynamic rewiring paradigm

The main contribution of [1] is the following idea: Instead of applying a single rewired graph $\mathcal{R}(G)$, consider a sequence of rewired graphs

$$G = G_0 \overset{\mathcal{R}}{\hookrightarrow} \cdots \overset{\mathcal{R}}{\hookrightarrow} G_{\ell-1} \overset{\mathcal{R}}{\hookrightarrow} G_\ell \overset{\mathcal{R}}{\hookrightarrow} \cdots \overset{\mathcal{R}}{\hookrightarrow} G_L \tag{17}$$

where G_ℓ is called the ℓ-snapshot, since we think of G_ℓ as the input graph evolved along a dynamical process for ℓ iterations.

The variant to derive the theoretical results in [1] is a Graph Convolutional Network (GCN)-LASER architecture

$$\begin{aligned} \vec{\boldsymbol{x}}_v^{(t)} = \mathrm{ReLU}\Bigg(& \sum_{u \in \mathcal{N}_1(v) \cup \{v\}} \frac{1}{\sqrt{d_v d_u}} \cdot \underline{\boldsymbol{W}}_0^{(t)} \vec{\boldsymbol{x}}_u^{(t-1)} \\ & + \sum_{\ell=1}^{L} \sum_{u \in \mathcal{N}_{\ell+1}^{\rho}(v)} \frac{1}{\sqrt{d_{v,\ell} d_{u,\ell}}} \cdot \underline{\boldsymbol{W}}_\ell^{(t)} \vec{\boldsymbol{x}}_u^{(t-1)} \Bigg) \end{aligned} \tag{18}$$

where $\underline{\boldsymbol{W}}_\ell^{(t)}$, for $\ell = 0, \ldots, L$, are learnable weight matrices and

$$\mathcal{N}_{\ell+1}^{\rho}(v) \subset \mathcal{N}_{\ell+1}(v) := \{u \in V \mid d_{G_0}(u,v) = \ell+1\} \tag{19}$$

is a ρ-fraction of nodes at distance exactly $\ell+1$ from v with the lowest connectivity score μ in the original graph $G = G_0$. So, for the ℓ^{th} snapshot G_ℓ, LASER considers $\rho \cdot d_{v,\ell+1}$ artificial edges from v to nodes at distance $d_{G_0}(u,v) = \ell+1$.

3.2 Jacobian sensitivity

Let $\underline{\boldsymbol{J}}^{(t)}(v,u) := \partial \vec{\boldsymbol{x}}_v^{(t)} / \partial \vec{\boldsymbol{x}}_u^{(0)}$ be the Jacobian of features after t layers of message passing on G using a GCN as described in [2], and similarly let $\tilde{\underline{\boldsymbol{J}}}^{(t)}(v,u)$ be the Jacobian after t layers using LASER, as defined in Equation 18.

Proposition 3. Let $v, u \in V$ with $r = d_{G_0}(v,u)$, where G_0 is the original graph before applying LASER. Assume there is exactly one path $\pi \subset G_0$ of length r connecting v and u (see Figure 1).

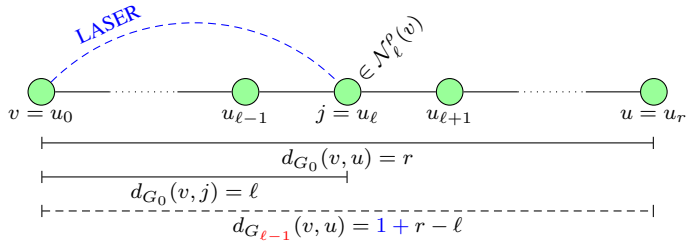


Figure 1. Unique shortest path $\pi \subset G_0$ between v and u of length r with an intermediate node j at distance ℓ from v. The dashed blue line indicates the edge added by LASER.

Assume LASER adds an edge between v and a node $j \in V_\pi$ with $d_{G_0}(v,j) = \ell < r$. Furthermore, assume that all products of weight matrices have unit norm. Then for all $m \leq r$ we have

$$\left\| \mathrm{E}\left[\tilde{\underline{\boldsymbol{J}}}^{(r-\ell+1)}(v,u) \right] \right\| \geq \frac{(d_{\min})^\ell \sqrt{\frac{d_j}{d_u}}}{\sqrt{d_{v,\ell-1} d_{j,\ell-1}}} \left\| \mathrm{E}\left[\underline{\boldsymbol{J}}^{(m)}(v,u) \right] \right\| \quad (20)$$

where $d_{\min} = \min_{v \in V} d_v$ is the minimum degree in G_0. ◁

Proof In the following, the expectations are with respect to ReLU' which is assumed to have probability of success p for all paths in the computational graph as in [3, 4]. The Jacobian after m layers of GCN is

$$\left\| \mathrm{E}\left[\underline{\boldsymbol{J}}^{(m)}(v,u) \right] \right\| = p \left\| \prod_{k=1}^{m} \underline{\boldsymbol{W}}^{(k)} \right\| (\hat{\underline{\boldsymbol{A}}}^m)_{vu} = p (\hat{\underline{\boldsymbol{A}}}^m)_{vu}$$

where we used the unit-norm assumption on the weight matrices and $\hat{\underline{\boldsymbol{A}}}$ denotes the normalized adjacency matrix from Definition 1.

For $m < r$, the expression above vanishes, satisfying (20), so we consider $m = r$ from now on.

With the same assumptions, the Jacobian after $r - \ell + 1$ layers of LASER is

$$\begin{aligned} \left\| \mathrm{E}\left[\tilde{\underline{\boldsymbol{J}}}^{(r-\ell+1)}(v,u) \right] \right\| &= p \left\| \prod_{k=2}^{r-\ell+1} \underline{\boldsymbol{W}}_0^{(k)} \underline{\boldsymbol{W}}_{\ell-1}^{(1)} \right\| (\hat{\underline{\boldsymbol{A}}}_{\ell-1})_{vj} (\hat{\underline{\boldsymbol{A}}}^{r-\ell})_{ju} \\ &= p (\hat{\underline{\boldsymbol{A}}}_{\ell-1})_{vj} (\hat{\underline{\boldsymbol{A}}}^{r-\ell})_{ju} \end{aligned}$$

As we can see in Figure 1, we have a unique path

$$(v = u_0, u_1, \ldots, u_{r-1}, u_r = u)$$

and we assume $j = u_\ell$. From Corollary 2, for the full path from v to u:

$$(\hat{\underline{\boldsymbol{A}}}^r)_{vu} = \frac{1}{\sqrt{d_v d_u}} \prod_{s=1}^{r-1} \frac{1}{d_{u_s}}$$

For the sub-path from $j = u_\ell$ to $u = u_r$ of length $r - \ell$, the same reasoning gives

$$(\hat{\underline{\boldsymbol{A}}}^{r-\ell})_{ju} = \frac{1}{\sqrt{d_j d_u}} \prod_{s=\ell+1}^{r-1} \frac{1}{d_{u_s}}.$$

Now we calculate the ratio

$$\begin{aligned} \frac{\left\| \mathrm{E}\left[\tilde{\underline{\boldsymbol{J}}}^{(r-\ell+1)}(v,u) \right] \right\|}{\left\| \mathrm{E}\left[\underline{\boldsymbol{J}}^{(r)}(v,u) \right] \right\|} &= \frac{(\hat{\underline{\boldsymbol{A}}}_{\ell-1})_{vj} (\hat{\underline{\boldsymbol{A}}}^{r-\ell})_{ju}}{\frac{1}{\sqrt{d_v d_u}} \prod_{s=1}^{r-1} \frac{1}{d_{u_s}}} \\ &= \frac{(\hat{\underline{\boldsymbol{A}}}_{\ell-1})_{vj}}{\frac{1}{\sqrt{d_v d_u}} \prod_{s=1}^{\ell-1} \frac{1}{d_{u_s}}} \cdot \frac{(\hat{\underline{\boldsymbol{A}}}^{r-\ell})_{ju}}{\prod_{s=\ell}^{r-1} \frac{1}{d_{u_s}}} \\ &\overset{(5)}{=} \frac{(\hat{\underline{\boldsymbol{A}}}_{\ell-1})_{vj}}{\frac{1}{\sqrt{d_v d_u}} \prod_{s=1}^{\ell-1} \frac{1}{d_{u_s}}} \cdot \frac{\frac{1}{\sqrt{d_j d_u}} \prod_{s=\ell+1}^{r-1} \frac{1}{d_{u_s}}}{\prod_{s=\ell}^{r-1} \frac{1}{d_{u_s}}} \\ &= \frac{(\hat{\underline{\boldsymbol{A}}}_{\ell-1})_{vj}}{\frac{1}{\sqrt{d_v d_u}} \prod_{s=1}^{\ell-1} \frac{1}{d_{u_s}}} \cdot \frac{\frac{1}{\sqrt{d_j d_u}} \prod_{s=\ell+1}^{r-1} \frac{1}{d_{u_s}}}{\frac{1}{d_{u_\ell}} \prod_{s=\ell+1}^{r-1} \frac{1}{d_{u_s}}} \\ &= \frac{(\hat{\underline{\boldsymbol{A}}}_{\ell-1})_{vj}}{\frac{1}{\sqrt{d_v d_u}} \prod_{s=1}^{\ell-1} \frac{1}{d_{u_s}}} \cdot \frac{d_{u_\ell}}{\sqrt{d_j d_u}} \\ &= \frac{(\hat{\underline{\boldsymbol{A}}}_{\ell-1})_{vj}}{\frac{1}{\sqrt{d_v d_u}} \prod_{s=1}^{\ell-1} \frac{1}{d_{u_s}}} \cdot \sqrt{\frac{d_j}{d_u}} \end{aligned}$$

Using

$$(\hat{\underline{\boldsymbol{A}}}_{\ell-1})_{vj} \overset{(8)}{=} \frac{1}{\sqrt{d_{v,\ell-1} d_{j,\ell-1}}}$$

and

$$\frac{1}{\sqrt{d_v d_u}} \prod_{s=1}^{\ell-1} \frac{1}{d_{u_s}} \leq \frac{1}{(d_{\min})^\ell}$$

we obtain

$$\frac{(\hat{\underline{\boldsymbol{A}}}_{\ell-1})_{vj}}{\frac{1}{\sqrt{d_v d_u}} \prod_{s=1}^{\ell-1} \frac{1}{d_{u_s}}} \geq \frac{(d_{\min})^\ell}{\sqrt{d_{v,\ell-1} d_{j,\ell-1}}}$$

which, combined with the previous equation, gives

$$\frac{\left\| \mathrm{E}\left[\tilde{\underline{\boldsymbol{J}}}^{(r-\ell+1)}(v,u) \right] \right\|}{\left\| \mathrm{E}\left[\underline{\boldsymbol{J}}^{(r)}(v,u) \right] \right\|} \geq \frac{(d_{\min})^\ell}{\sqrt{d_{v,\ell-1} d_{j,\ell-1}}} \cdot \sqrt{\frac{d_j}{d_u}}$$

concluding the proof. □

3.3 Preservation of locality

Proposition 4. Following [4], we compare how much a rewiring perturbs the shortest-path distances by means of the Frobenius norm

$$\left\| \mathcal{D}_G - \mathcal{D}_{G'} \right\|_F, \qquad (\mathcal{D}_G)_{uv} = d_G(u,v).$$

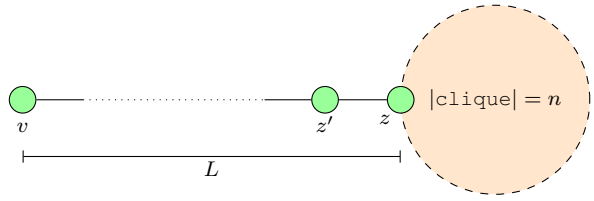


Figure 2. Lollipop graph with a chain of length L attached to a clique of size n.

◁

Setup. Let G be a lollipop graph consisting of a path of length L attached to a clique of size n. Let v be the endpoint of the path, z the junction between path and clique, and let the other $n-1$ clique nodes be "interior" nodes.

Spectral rewiring: lower bound Consider a spectral rewiring $\mathcal{R}$ that adds one edge between the pair with largest effective resistance, which in this graph is (v, u) for some interior clique node u. Before rewiring,

$$d_G(v, u) = L + 1, \qquad d_G(v, z) = L.$$

After adding (v, u) we have

$$d_{\mathcal{R}(G)}(v, u) = 1, \qquad d_{\mathcal{R}(G)}(v, w) \le 2 \text{ for all interior clique } w, \qquad d_{\mathcal{R}(G)}(v, z) \le 2.$$

Hence, for each of the $n - 1$ interior clique nodes

$$\left|d_G(v, w) - d_{\mathcal{R}(G)}(v, w)\right| \;\ge\; (L + 1) - 2 = L - 1,$$

and for z

$$\left|d_G(v, z) - d_{\mathcal{R}(G)}(v, z)\right| \;\ge\; L - 2.$$

Restricting the Frobenius sum to these n pairs yields

$$\left\|\mathcal{D}_G - \mathcal{D}_{\mathcal{R}(G)}\right\|_F^2 \;\ge\; (n - 1)(L - 1)^2 + (L - 2)^2, \tag{21}$$

and therefore

$$\left\|\mathcal{D}_G - \mathcal{D}_{\mathcal{R}(G)}\right\|_F \;\ge\; \sqrt{(n - 1)(L - 1)^2 + (L - 2)^2}. \tag{22}$$

LASER: upper bound Now consider LASER with a single snapshot restricted to distance-2 edges. On the path, every node except the penultimate one z' has at most two distance-2 neighbours; z' has the whole clique at distance 2. If we choose ρ such that

$$\rho = \frac{k}{2L}, \tag{23}$$

then the expected number of new edges incident to path nodes is k. Assuming that at most k such edges are added, we obtain uniform bounds on the distance changes.

Path–path pairs. Each new edge connects two nodes at distance 2, so using it on a shortest path can shorten that path by at most 1. Any path between two nodes x, y on the path can traverse at most k new edges, hence

$$\left|d_G(x, y) - d_{\text{LASER}}(x, y)\right| \le k \quad \text{for all path nodes } x, y.$$

There are at most L^2 ordered path–path pairs, so their contribution to the Frobenius norm satisfies

$$\sum_{x,y \text{ on path}} \left(d_G(x, y) - d_{\text{LASER}}(x, y)\right)^2 \;\le\; L^2 k^2. \tag{24}$$

Path–clique pairs. In the worst case, one of the k new edges connects z' directly to some clique node. Then any shortest path between a path node x and a clique node c can use at most k shortcuts along the path and one shortcut to enter the clique. Thus

$$\left|d_G(x, c) - d_{\text{LASER}}(x, c)\right| \;\le\; k + 1 \quad \text{for all } x \text{ on the path and } c \text{ in the clique.}$$

There are nL such ordered pairs, giving

$$\sum_{\substack{x \text{ on path} \\ c \text{ in clique}}} \left(d_G(x, c) - d_{\text{LASER}}(x, c)\right)^2 \;\le\; nL(k + 1)^2. \tag{25}$$

Distances between clique nodes are not increased by adding local edges from the path and are at most mildly decreased; their contribution can be bounded and is dominated by the two terms above. Hence we obtain the global upper bound

$$\left\|\mathcal{D}_G - \mathcal{D}_{\text{LASER}}\right\|_F^2 \;\le\; L^2 k^2 + nL(k + 1)^2, \tag{26}$$

and therefore

$$\left\|\mathcal{D}_G - \mathcal{D}_{\text{LASER}}\right\|_F \;\le\; \sqrt{L^2 k^2 + nL(k + 1)^2}. \tag{27}$$

Comparison and corrected parameter constraints A sufficient condition for LASER to be more locality-preserving than the spectral rewiring is

$$\left\|\mathcal{D}_G - \mathcal{D}_{\text{LASER}}\right\|_F \;\le\; \left\|\mathcal{D}_G - \mathcal{D}_{\mathcal{R}(G)}\right\|_F.$$

Using (22) and (27), this is guaranteed if

$$L^2 k^2 + nL(k + 1)^2 \;\le\; (n - 1)(L - 1)^2 + (L - 2)^2. \tag{28}$$

Rearranging (28) and solving for n gives

$$n \;\ge\; \frac{k^2 L^2 + 2L - 3}{L^2 - Lk^2 - 2Lk - 3L + 1}, \tag{29}$$

provided the denominator is positive. A sufficient condition for the denominator to be positive is

$$k^2 < \frac{L}{4}, \qquad L \ge 8, \tag{30}$$

which we verified analytically for $k \in \mathbb{N}$. Using the relation $\rho = k/(2L)$ from (23), the constraint (30) is equivalent to

$$\rho^2 < \frac{1}{16L} \quad \Longleftrightarrow \quad \rho < \frac{1}{4\sqrt{L}}. \tag{31}$$

In summary, on lollipop graphs with $L \ge 8$, sufficiently small sampling rate $\rho < 1/(4\sqrt{L})$ and n satisfying (29), we obtain

$$\left\|\mathcal{D}_G - \mathcal{D}_{\text{LASER}}\right\|_F \;\le\; \left\|\mathcal{D}_G - \mathcal{D}_{\mathcal{R}(G)}\right\|_F, \tag{32}$$

i.e. LASER perturbs the shortest-path distances less than the spectral effective-resistance-based rewiring.

Compared to the original derivation, we note that: (i) the path–clique bound should be interpreted as a worst-case inequality $\le k + 1$ rather than an equality; (ii) the algebraic manipulation leading to the bound on n in (29) fixes a sign error in the denominator; and (iii) the correct constraint on ρ is (31), which differs by a factor of 2 from the value stated in the paper.

4 Related works

5 Methodology

6 Implementation

6.1 Datasets

6.2 Hyperparameters

6.3 Experimental setup

6.4 Computational requirements

7 Results

8 Discussion and conclusion

References

[1] Federico Barbero, Ameya Velingker, Amin Saberi, Michael Bronstein, and Francesco Di Giovanni. Locality-aware graph-rewiring in gnns. 2024.

[2] Thomas N. Kipf and Max Welling. Semi-Supervised Classification with Graph Convolutional Networks. In International Conference on Learning Representations, 2017.

[3] Keyulu Xu, Chengtao Li, Yonglong Tian, Tomohiro Sonobe, Ken-ichi Kawarabayashi, and Stefanie Jegelka. Representation learning on graphs with jumping knowledge networks. In International Conference on Machine Learning, pages 5453–5462. PMLR, 2018.

[4] Francesco Di Giovanni, Lorenzo Giusti, Federico Barbero, Giulia Luise, Pietro Lio, and Michael Bronstein. On over-squashing in message passing neural networks: The impact of width, depth, and topology. arXiv preprint arXiv:2302.02941, 2023.