Fundamental Limits of Local Graph Neural Networks on High-Girth Graphs

Kalyan Cherukuri
Illinois Mathematics and Science Academy
Aurora, IL
kcherukuri@imsa.edu

Abstract

We determine exact limits of local, message-passing GNNs (Graph neural networks) on graphs of high girth. We prove upper bounds for any GNN with a receptive radius of L on graphs with a maximum degree of d and girth exceeding $2L + 1$. The central contribution is a finite-dimensional linear program that characterizes the optimal performance of any L-local randomized algorithm on the infinite d-regular tree. We demonstrate that this value serves as a hard asymptotic ceiling for local GNNs on large, high-girth graphs and prove the tightness with a GNN construction that achieves this bound.

1 Introduction

Message-passing GNNs are widely used for node-level and combinatorial tasks [23]. The expressive power of such GNNs is limited, dependent upon their finite receptive radius, much like other neural architectures that depend on locality (eg. CNNs) [20, 12, 18, 16], and hierarchical residual designs [3]. Thus, it becomes quite natural to raise the question: given an architecture that exchanges information only within a radius L, what is the maximal information that training and arbitrary supervision can achieve on sparse graphs that are locally tree-like at a scale L?

This question resonates with broader themes in deep learning, where the tension between locality, expressiveness, and scalability has driven major architectural innovations. Recurrent networks [11], convolutional networks [20, 12], and U-Nets [18] each exploit locality in distinct ways, while residual [12], normalization [2], and distillation methods [13] have enhanced their scalability. At the same time, stochastic optimization [15, 9] has been crucial for training these models effectively.

The shift to attention-based models, initiated by the introduction of the transformer [22], further extended receptive fields beyond locality, enabling breakthroughs in language modeling [7, 4, 1, 21], vision [8, 6], and multimodal learning [17]. Such architectures underpin state-of-the-art systems in vision-language alignment [17], large-scale representation learning [5], architecture search [25], and reinforcement learning [19, 24]. These advances highlight the central importance of expressiveness and receptive fields in neural networks.

In this research, we wish to answer the locality-expressiveness tradeoff for GNNs with precision. We show that the maximal achievable performance with any L-local GNN on bounded-degree high-girth graphs is equivalent to the value of a finite linear program. That linear program can be solved numerically for any degree bound d and any radius L, producing a rigorous and generalizable upper bound theory. We further show the tightness of this bound by constructing randomized L-local GNNs that approximate the LP optimizer arbitrarily.

The reduction from local GNNs to factor-of-i.i.d. processes and local algorithms is known informally in the literature. The contribution is to render this reduction fully quantitative, to give an

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Differentiable Learning of Combinatorial Algorithms.

explicit finite-dimensional convex program that computes the optimal value, and to prove matching approximation theorems that show the LP value is realizable. This connects the expressiveness of local algorithms to the broader arc of deep learning research—from adversarial learning [10], to amortized inference [14], to foundation models that unify modalities and tasks [21, 1].

2 Contributions

1. We introduce a finite-dimensional linear program that, for any maximum degree d and receptive radius L, exactly computes the optimal performance achievable by any L-local randomized algorithm on the infinite d-regular tree.
2. We rigorously prove that this value, $C(d, L)$, serves as a hard asymptotic performance ceiling for any L-local GNN on large graphs with girth greater than $2L + 1$, formally bounding the GNN's approximation ratio.
3. We provide a constructive proof that this bound is tight and apply the framework to derive concrete, numerically-solvable linear programs for Max-Cut and Min-Vertex-Cover for any given pair (d, L).

2.1 Notation

Let $d \geq 2$ and $L \geq 1$ be integers fixed throughout unless stated otherwise. For a finite undirected graph $G = (V, E)$ and vertex $v \in V$ we denote by $B_L(G, v)$ the rooted induced subgraph of G consisting of all vertices at graph distance at most L from v, together with the root v. The infinite rooted d-regular tree is T_d with root o. For a random variable X we write $\mathbb{P}[X]$ for its law and $\mathbb{E}[X]$ for its expectation. For a finite set S, $\Delta(S)$ denotes the simplex of probability distributions over S. (This just means the set of all probability distributions on S)

3 Modeling Local Randomized GNNs and Factor-of-i.i.d. Maps

We define the class of radius-L randomized message-passing rules and show that any fixed-parameter GNN in the standard synchronous setting belongs to this class. This formalization isolates the only relevant property: locality.

Definition 3.1 (Radius-L randomized rule)**.** A radius-L randomized rule is a measurable map

$$\Phi : \mathcal{B}_L \times \Xi^{V(B)} \to \mathcal{Y},$$

where $\mathcal{B}_L$ is the set of rooted graph isomorphism classes of radius-L balls with maximum degree at most d, B denotes a canonical representative of an element of $\mathcal{B}_L$, Ξ is a seed space equipped with the product measure $\mu_\xi^{\otimes |V(B)|}$, and $\mathcal{Y}$ is the decision space. The interpretation is that, given a rooted ball type $b \in \mathcal{B}_L$ and independent random seeds on the vertices of the ball, the decision at the root is $\Phi(b, \xi)$.

Every synchronously executed message-passing GNN with L layers and i.i.d. node-level randomness is an instance of a radius-L randomized rule: the per-node output depends only on the isomorphism class of the radius-L ball and the seeds inside that ball. We therefore work with the abstract class of all measurable Φ.

4 Reduction theorem

Let $\{G_n\}_{n \geq 1}$ be a sequence of finite graphs with maximum degree at most d and $|V(G_n)| \to \infty$. We will sample a uniform vertex $v_n \in V(G_n)$.

Theorem 4.1 (Reduction to tree factors)**.** *Fix integers $d \geq 2$ and $L \geq 1$. Let Φ be any radius-L randomized rule measurable as above. Let μ_ξ be the law of i.i.d. node seeds. For each n let $G_n = (V_n, E_n)$ be a finite graph with maximum degree at most d. For v_n sampled uniformly from V_n and $\xi^{(n)} \sim \mu_\xi^{\otimes |V_n|}$ i.i.d. seeds, set*

$$Y_n \;=\; \Phi\big(B_L(G_n, v_n), \{\xi_u^{(n)} : u \in B_L(G_n, v_n)\}\big).$$

Let $p_n \in \Delta(\mathcal{B}_L)$ be the distribution of the random rooted ball $B_L(G_n, v_n)$. Let $p_\infty \in \Delta(\mathcal{B}_L)$ be the distribution of $B_L(T_d, o)$. Then for every bounded measurable function $f : \mathcal{Y} \to \mathbb{R}$ we have

$$\left|\mathbb{E}[f(Y_n)] - \mathbb{E}_{b\sim p_\infty}\mathbb{E}_{\xi\sim\mu_\xi^{\otimes|V(b)|}}\left[f(\Phi(b,\xi))\right]\right| \le \|f\|_\infty \cdot \|p_n - p_\infty\|_1. \tag{1}$$

In particular, if $\|p_n - p_\infty\|_1 \to 0$ then $\mathbb{E}[f(Y_n)]$ converges to the expectation obtained by applying Φ to the truncated tree law.

Proof. Because $\mathcal{B}_L$ is finite the distribution p_n is well defined. Conditioning on the rooted ball yields

$$\mathbb{E}[f(Y_n)] \;=\; \sum_{b\in\mathcal{B}_L} p_n(b)\,\mathbb{E}_\xi[f(\Phi(b,\xi))].$$

The same equality holds with p_∞. Subtracting the two and applying the triangle inequality gives

$$\left|\mathbb{E}[f(Y_n)] - \mathbb{E}_{b\sim p_\infty}\mathbb{E}_\xi[f(\Phi(b,\xi))]\right| \le \sum_{b\in\mathcal{B}_L} |p_n(b) - p_\infty(b)| \cdot \|\mathbb{E}_\xi[f(\Phi(b,\xi))]\|_\infty,$$

and the integrand is bounded by $\|f\|_\infty$. This yields inequality (1). The remainder follows immediately. ☐

5 Finite-dimensional LP characterization

Let $\mathcal{B}_L$ be the set of rooted radius-L ball types with maximum degree at most d. We consider combinatorial objectives that can be written as expectations of bounded local reward functions. These include node-additive or edge-additive objectives.

Definition 5.1 (Local objective)**.** A local objective consists of a measurable reward function

$$\ell : \mathcal{B}_L \times \Xi^{V(b)} \times \mathcal{Y} \to \mathbb{R},$$

together with a normalization factor ρ, which is either $|V|$ or $|E|$ for finite graphs. For a radius-L randomized rule Φ, the expected normalized reward is

$$\mathcal{R}(\Phi) = \mathbb{E}_{b\sim p_\infty}\,\mathbb{E}_\xi\left[\,\ell(b,\xi,\Phi(b,\xi))\,\right].$$

Our goal is to characterize the supremum of $\mathcal{R}(\Phi)$ over all radius-L randomized rules. The key idea is that the joint distribution of

$$(\text{radius-}L\text{ ball},\ \text{seed labels on the ball},\ \text{decision at the root})$$

takes values in a finite set and must satisfy natural consistency constraints inherited from the tree structure. Maximizing $\mathcal{R}(\Phi)$ over such distributions becomes a finite linear program.

Definition 5.2 (Labeled ball types)**.** After discretizing the seed and decision alphabets (if needed), let $\mathcal{T}_L$ denote the finite set of isomorphism classes of rooted radius-L balls whose vertices are labeled by the seed alphabet and whose root carries the decision label. Each type $t \in \mathcal{T}_L$ determines a local reward value, denoted $r(t)$.

Proposition 5.1 (LP characterization)**.** *Let $\mathcal{T}_L$ and $r(\cdot)$ be as above. The supremal expected reward $\sup_\Phi \mathcal{R}(\Phi)$ equals the optimal value of the following finite linear program:*

maximize $\sum_{t\in\mathcal{T}_L} r(t)\,q(t)$
subject to $q \in \Delta(\mathcal{T}_L)$,
marginal consistency constraints on overlapping radius-ℓ interfaces for all $\ell \le L$,
automorphism symmetry constraints induced by $\mathcal{B}_L$.

The program is feasible and its optimal value is attained.

Theorem 5.1 (LP exactness)**.** *Assume the seed and decision alphabets are finite. Then the supremum of $\mathcal{R}(\Phi)$ over measurable radius-L randomized rules equals the optimum value of the finite linear program described in Proposition 5.1. Moreover the LP optimum can be computed in time polynomial in $|\mathcal{T}_L|$ up to standard LP solver complexity.*

Proof. Let $\mathcal{T}_L$ denote the set of labeled radius-L balls (seeds + root decision). Any measurable radius-L randomized rule Φ induces a distribution $q_\Phi \in \Delta(\mathcal{T}_L)$ giving the joint law of the labeled ball and root decision. By construction, q_Φ satisfies the interface consistency constraints, and $\mathcal{R}(\Phi)$ is linear in q_Φ. Hence every randomized rule corresponds to a feasible point of the LP.

Conversely, let $q \in \Delta(\mathcal{T}_L)$ satisfy the LP constraints. One can sequentially sample labels and decisions along any finite ball of the tree, extending radius-L neighborhoods in a manner consistent with q; this produces a randomized local rule Φ_q whose induced distribution equals q and attains the LP objective.

Compactness of $\Delta(\mathcal{T}_L)$ ensures the supremum is attained. Since $|\mathcal{T}_L|$ is finite and the constraints are polynomial in $|\mathcal{T}_L|$, the LP optimum can be computed in polynomial time using any standard LP solver. $\square$

Remark 5.1. The finiteness assumption on seeds and outputs can be removed by approximation: discretize to arbitrary precision and bound the approximation error. As the discretization is refined the LP optimum converges to the true supremum.

6 Realization by randomized local GNNs

We show that any optimal LP solution can be implemented, up to arbitrary precision, by an L-local randomized GNN.

Theorem 6.1 (Realizability and tightness)**.** *Let $q^\star$ be an optimal solution to the LP of Theorem 5.1. For every $\varepsilon > 0$ there exists an L-local randomized GNN Φ such that*

$$\Big|\mathcal{R}(\Phi) - \sum_{t\in\mathcal{T}_L} r(t)\, q^\star(t)\Big| \le \varepsilon.$$

Hence the LP value is the exact performance ceiling for all L-local GNNs on graphs of girth exceeding $2L+1$.

Proof. By Theorem 5.1, the optimal point $q^\star$ corresponds to a feasible set of radius-L local distributions satisfying the LP consistency constraints. Such a distribution can be realized by a randomized radius-L rule: at each node, one samples local randomness and outputs a label according to the prescribed conditional distribution of $q^\star$ on its radius-L neighborhood.

An L-layer randomized GNN can implement any such rule. The seeds are given as node features, message passing exposes the entire radius-L neighborhood, and the final node-level MLP can approximate the desired output distribution arbitrarily well because the domain (the set of labeled radius-L balls) is finite. By standard universal approximation, the resulting GNN matches the reward $\sum_t r(t)q^\star(t)$ up to an additive error at most ε.

Thus every LP optimum is realizable to arbitrary precision by an L-local randomized GNN. $\square$

7 Max-Cut and Min-Vertex-Cover Upper Bounds

Both objectives fit into the local LP framework by taking $\mathcal{T}_L$ to encode the labeled radius-L ball together with the root decision; the LP variables are a distribution $q \in \Delta(\mathcal{T}_L)$ subject to interface-consistency (and any hard feasibility) constraints, and the objective is the expected local contribution expressed linearly in q.

7.1 Max-Cut

Label vertices by spins $\sigma_v \in \{-1, +1\}$. The cut density on a finite graph $G = (V, E)$ is

$$\mathrm{cut}(G, \sigma) = \frac{1}{|E|} \sum_{\{u,v\}\in E} \frac{1 - \sigma_u \sigma_v}{2}.$$

For the LP one encodes in each $t \in \mathcal{T}_L$ the labeled radius-L ball together with the root spin; the local reward $r(t)$ is the sum of contributions of edges incident to the root, normalized by the degree d (so that summing $r(t)$ over vertices recovers the per-edge normalization). Let $\alpha_{d,L}$ denote the optimal LP value.

Corollary 7.1 (Max-Cut ceiling)**.** *Let $\alpha_{d,L}$ be the LP optimum for Max-Cut over radius-L factors. Then for any sequence of graphs G_n with maximum degree $\leq d$ and girth $> 2L+1$, every L-local (randomized) GNN attains expected cut density $\leq \alpha_{d,L}$ for all sufficiently large n. Conversely, $\alpha_{d,L}$ is asymptotically achievable by randomized L-local GNNs.*

Sketch. Reduction to the infinite d-regular tree identifies any limit law of radius-L neighborhoods with a feasible q for the LP, so Theorem 5.1 yields the upper bound $\alpha_{d,L}$. Realizability follows from Theorem 6.1. □

7.2 Min-Vertex-Cover

Encode a cover by indicators $x_v \in \{0,1\}$ where $x_v = 1$ iff v is in the cover. The (normalized) cover size is

$$\mathrm{cover}(G,x) = \frac{1}{|V|}\sum_{v\in V} x_v.$$

In the LP take each $t \in \mathcal{T}_L$ to record the labeled radius-L ball and the root indicator x_{root}. The local cost is $c(t) = x_{\text{root}}$, so the objective is to minimize $\sum_t c(t)\, q(t)$. Enforce covering as a hard (almost-sure) constraint: for every possible adjacent-pair overlap pattern, the LP contains linear constraints forcing the probability that an edge is uncovered (both endpoints have $x = 0$ with matching overlap) to be zero. Concretely, the uncovered-edge probability can be written as a linear combination of the single-ball q-variables (summing over those t whose local patterns, when paired with compatible neighbors, yield both roots uncovered), and we set that linear expression to zero.

Let $\beta_{d,L}$ denote the optimum of this LP.

Corollary 7.2 (Vertex-cover ceiling)**.** *Let $\beta_{d,L}$ be the LP optimum for Min-Vertex-Cover with the interface and uncovered-edge constraints described above. Under the same hypotheses as in the Max-Cut corollary, every radius-L randomized GNN attains asymptotic expected cover fraction $\geq \beta_{d,L}$, and $\beta_{d,L}$ is asymptotically realizable.*

Sketch. Exactly as for Max-Cut: any limit of L-local computations on large-girth graphs yields a feasible q, so the LP lower-bounds achievable cover fraction. The hard covering constraint in the LP enforces feasibility (no uncovered edges) in the limit. Realizability is obtained from Theorem 6.1 (the LP-optimal local law can be implemented by an L-local randomized GNN up to arbitrary precision). □

8 Limitations

The theoretical results present in this work are subjected to several key assumptions. Our tight bounds that we present are derived for the class of local, message-passing GNNs and do not apply to more expressive non-local architectures (e.g., think about those involving positional encodings or transformer-like mechanics). Furthermore, our core analysis relies on the assumption of high-girth graphs (specifically, girth greater than $2L+1$), and the derived bounds may not hold for graphs with dense clusters or numerous short cycles. Finally, our results are asymptotic in nature, describing the performance limits on sequences of large graphs, and may not precisely characterize the behavior on a specific, small-graph instance.

9 Conclusion

We provided a fully constructive and theory of the limits of radius-L message-passing GNNs on bounded-degree high-girth graphs. The tight LP characterization converts the informal folklore into a practical tool: for any pair (d, L) a practitioner can compute an exact performance ceiling for local GNNs on locally tree-like graphs and then decide whether architectural changes are necessary to exceed that ceiling.

References

[1] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

[2] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016.

[3] Andy Brock, Soham De, Samuel L Smith, and Karen Simonyan. High-performance large-scale image recognition without normalization. In *International conference on machine learning*, pages 1059–1071. PMLR, 2021.

[4] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

[5] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PmLR, 2020.

[6] Timothée Darcet, Maxime Oquab, Julien Mairal, and Piotr Bojanowski. Vision transformers need registers. *arXiv preprint arXiv:2309.16588*, 2023.

[7] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pages 4171–4186, 2019.

[8] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[9] Ian Gemp, Luke Marris, and Georgios Piliouras. Approximating nash equilibria in normal-form games via stochastic optimization. *arXiv preprint arXiv:2310.06689*, 2023.

[10] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[11] Alex Graves. Sequence transduction with recurrent neural networks. *arXiv preprint arXiv:1211.3711*, 2012.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[13] Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*, 2015.

[14] Edward J Hu, Moksh Jain, Eric Elmoznino, Younesse Kaddar, Guillaume Lajoie, Yoshua Bengio, and Nikolay Malkin. Amortizing intractable inference in large language models. *arXiv preprint arXiv:2310.04363*, 2023.

[15] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[16] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. A convnet for the 2020s. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11976–11986, 2022.

[17] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PmLR, 2021.

[18] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[19] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[20] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[21] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.

[22] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[23] Bohang Zhang, Jingchu Gai, Yiheng Du, Qiwei Ye, Di He, and Liwei Wang. Beyond weisfeiler-lehman: A quantitative framework for gnn expressiveness. *arXiv preprint arXiv:2401.08514*, 2024.

[24] Hongming Zhang and Tianyang Yu. Alphazero. In *Deep reinforcement learning: fundamentals, research and applications*, pages 391–415. Springer, 2020.

[25] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V Le. Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8697–8710, 2018.

A Appendix for *Fundamental Limits of Local Graph Neural Networks on High-Girth Graphs*

Max-Cut and Min-Vertex-Cover Upper Bounds

Throughout, $\mathcal{T}_L$ denotes the set of labeled radius-L rooted balls (labels include seeds and the root decision). An L-local randomized GNN induces a distribution $q \in \Delta(\mathcal{T}_L)$ satisfying the interface-consistency constraints. The LP optimum is computed over all such feasible q. All convergence statements use the fact that for any sequence of d-regular graphs G_n with girth $> 2L+1$, the law of a uniformly rooted radius-L ball equals the radius-L truncation of the d-regular tree.

A.1 Max-Cut

Represent spins by $\sigma_v \in \{-1,+1\}$. The cut density is

$$\mathrm{cut}(G,\sigma) = \frac{1}{|E|} \sum_{\{u,v\}\in E} \frac{1-\sigma_u\sigma_v}{2}.$$

For the LP we use the local reward

$$r(t) = \frac{1}{d} \sum_{u\sim \mathrm{root}} \frac{1-\sigma_{\mathrm{root}}\sigma_u}{2}.$$

For any d-regular graph,

$$\frac{1}{|V|} \sum_{v\in V} r(B_L(G,v),\sigma) = \mathrm{cut}(G,\sigma).$$

Let $\alpha_{d,L}$ be the LP optimum.

Corollary A.1 (Max-Cut Upper Bound)**.** *Let G_n be d-regular with girth $> 2L+1$. For any L-local randomized GNN Φ,*

$$\limsup_{n\to\infty} \mathbb{E}[\mathrm{cut}(G_n,\sigma^{\Phi})] \le \alpha_{d,L}.$$

Moreover, for every $\varepsilon > 0$ there exists an L-local randomized GNN Φ with

$$\liminf_{n\to\infty} \mathbb{E}[\mathrm{cut}(G_n,\sigma^{\Phi})] \ge \alpha_{d,L} - \varepsilon.$$

Proof. Fix Φ. Let v be uniformly random and let $\tilde{r}_n = r(B_L(G_n,v),\sigma^{\Phi})$. Since r depends only on the radius-L ball,

$$\mathbb{E}[\mathrm{cut}(G_n,\sigma^{\Phi})] = \mathbb{E}[\tilde{r}_n].$$

Because G_n has girth $> 2L+1$, the rooted radius-L ball distribution equals that of the d-regular tree. Hence

$$\lim_{n\to\infty} \mathbb{E}[\tilde{r}_n] = \mathbb{E}_{t\sim q_\Phi}[r(t)],$$

where q_Φ is the induced feasible LP distribution. Therefore

$$\limsup_{n\to\infty} \mathbb{E}[\mathrm{cut}(G_n,\sigma^{\Phi})] = \sum_t r(t)\, q_\Phi(t) \le \alpha_{d,L}.$$

For achievability, let $q^\star$ be an LP optimizer. By the realizability theorem (Theorem 6.1), for any $\varepsilon > 0$ there exists an L-local randomized GNN Φ with induced distribution q_Φ satisfying

$$\left| \sum_t r(t)\, q_\Phi(t) - \sum_t r(t)\, q^\star(t) \right| \le \varepsilon.$$

Thus for all large n,

$$\mathbb{E}[\mathrm{cut}(G_n,\sigma^{\Phi})] = \sum_t r(t)\, q_\Phi(t) \ge \alpha_{d,L} - \varepsilon.$$

□

A.2 Min-Vertex-Cover

Represent covers by $x_v \in \{0, 1\}$, and define

$$\mathrm{cover}(G, x) = \frac{1}{|V|} \sum_{v \in V} x_v.$$

For the LP use the local cost

$$c(t) = x_{\mathrm{root}}.$$

To enforce that no edge is uncovered, include linear constraints ensuring that for every oriented edge type, the probability (determined linearly from q via interface-consistency) that both endpoints have $x = 0$ is zero. Let $\beta_{d,L}$ be the LP optimum.

Corollary A.2 (Vertex-Cover Upper Bound)**.** *Let G_n be d-regular with girth $> 2L + 1$. For any L-local randomized GNN Φ that always outputs a valid vertex cover,*

$$\liminf_{n\to\infty} \mathbb{E}[\mathrm{cover}(G_n, x^\Phi)] \geq \beta_{d,L}.$$

For every $\varepsilon > 0$ there exists an L-local randomized GNN Φ whose outputs are valid covers and satisfy

$$\limsup_{n\to\infty} \mathbb{E}[\mathrm{cover}(G_n, x^\Phi)] \leq \beta_{d,L} + \varepsilon.$$

Proof. Fix Φ. Let $\tilde{c}_n = c(B_L(G_n, v), x^\Phi) = x_v^\Phi$ for v uniform. For every n,

$$\mathbb{E}[\mathrm{cover}(G_n, x^\Phi)] = \mathbb{E}[\tilde{c}_n].$$

As before, the rooted-radius-L distribution converges to that of the d-regular tree, hence

$$\lim_{n\to\infty} \mathbb{E}[\tilde{c}_n] = \sum_t c(t)\, q_\Phi(t),$$

where q_Φ is the induced law. Since Φ always outputs a valid cover, q_Φ satisfies the linear uncovered-edge constraints, so q_Φ is feasible. Therefore

$$\sum_t c(t)\, q_\Phi(t) \geq \beta_{d,L}.$$

For achievability, let $q^\star$ be an LP optimizer. By Theorem 6.1, for any $\varepsilon > 0$ there exists an L-local randomized GNN Φ inducing q_Φ with

$$\left| \sum_t c(t)\, q_\Phi(t) - \sum_t c(t)\, q^\star(t) \right| \leq \varepsilon.$$

Since $q^\star$ satisfies the uncovered-edge constraints, the realized GNN also produces valid covers on all large-girth graphs. Thus

$$\mathbb{E}[\mathrm{cover}(G_n, x^\Phi)] = \sum_t c(t)\, q_\Phi(t) \leq \beta_{d,L} + \varepsilon.$$

□

B Extensions to Real-World Applications

The LP framework applies directly to other networked systems where decisions are made using only radius-L neighborhood information. Below are two concise examples with clear mathematical interpretations. Note that Random Regular Graphs (RRGs) are high-girth, a standard benchmark for combinatorics, which show up quite often in epidemiology models and traffic problems.

B.1 Epidemiology Modeling

Let each individual be a vertex and ea,ch contact an edge. Let seeds encode local states (e.g., infection indicator $s_v \in \{0, 1\}$). A radius-L rule Φ outputs an intervention $a_v = \Phi(B_L(G, v), s)$ such as "isolate" or "no action."

Define the local transmission probability on an edge (u, v) as

$$p_{\mathrm{trans}}(u, v) = \Pr[s_u = 1,\, a_u = 0] \cdot \Pr[s_v = 0 \mid B_L(G, v)].$$

The expected transmission rate under Φ satisfies

$$\mathrm{Trans}(\Phi, G) = \frac{1}{|E|} \sum_{\{u,v\} \in E} p_{\mathrm{trans}}(u, v),$$

which can be written as a linear functional $\sum_{t \in \mathcal{T}_L} r_{\mathrm{epi}}(t)\, q(t)$ of the LP variables. Thus the LP optimum gives the minimum possible transmission achievable by any L-local policy.

B.2 Traffic Modeling

Let intersections be vertices and road segments edges. Seeds encode local traffic states. A radius-L rule Φ outputs a control action a_v (e.g., signal phase).

Let the delay contribution on edge (u, v) be a known function $D(q_u, q_v, a_u, a_v)$. The mean delay under Φ is

$$\mathrm{Delay}(\Phi, G) = \frac{1}{|E|} \sum_{\{u,v\} \in E} \mathbb{E}[D(q_u, q_v, a_u, a_v)],$$

which again equals a linear form $\sum_{t \in \mathcal{T}_L} r_{\mathrm{traffic}}(t)\, q(t)$.

Therefore, the LP optimum yields the best achievable delay for all radius-L local controllers on high-girth networks.