NSIA
Network Science
Informs AI

NetSci
Boston 2026

From Graphons to Graphlets: A Graph-Limit Lens on Sparse Neural Connectivity and Training

Siddhi Kanta Mishra and Ivan Garibay
University of Central Florida
Complex Adaptive Systems Laboratory

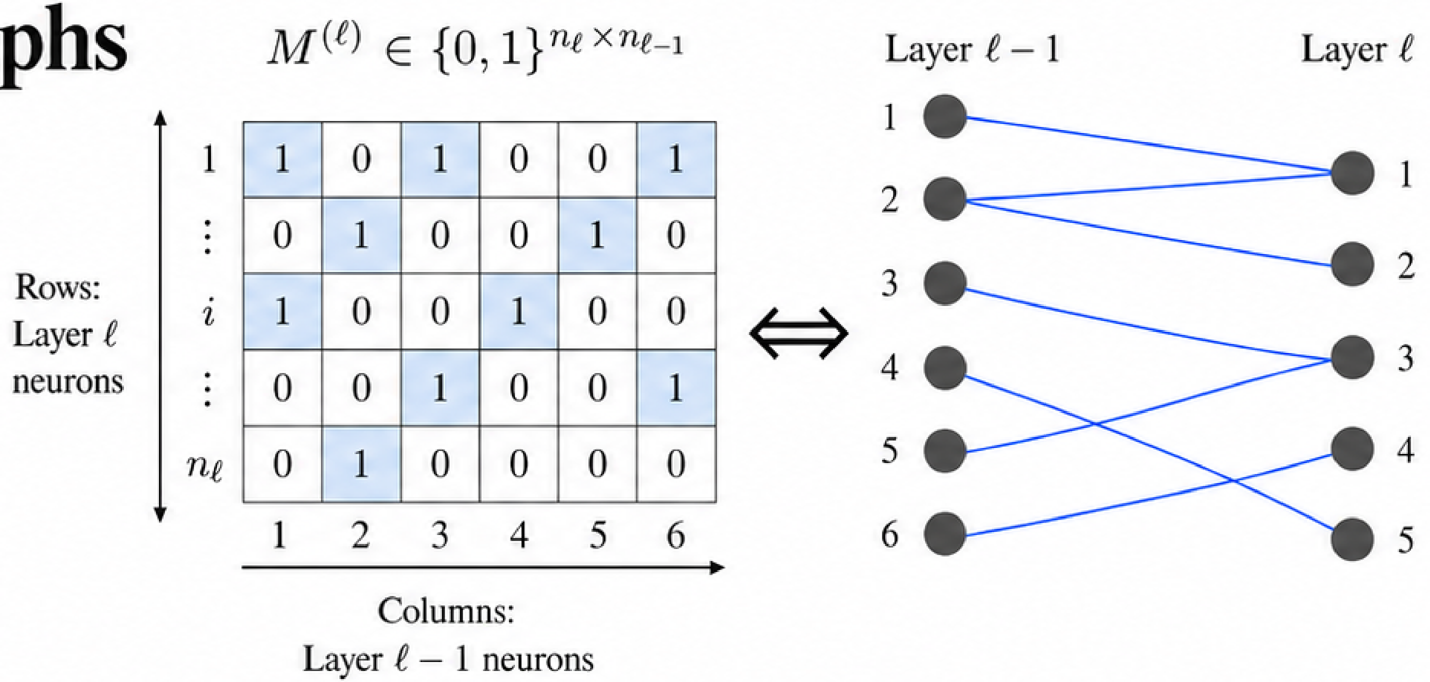
NSIA Network Science Informs AI

NetSci Boston 2026

1. Pruning masks are graphs

$$M^{(\ell)} \in \{0,1\}^{n_\ell \times n_{\ell-1}}$$

- A layerwise pruning mask is a bipartite graph.
- Source neurons: layer $\ell - 1$.
- Target neurons: layer ℓ.
- Retained weights are edges.



> *When we prune a neural network, we are not only removing parameters. We are creating a network.*

2. As width grows, what does a sparse neural mask converge to?

$$M_n \longrightarrow ?$$

Fixed retained fraction

$$p_n \to p > 0$$

Dense / fixed-fraction pruning

Graphon-friendly regime

Sparse fan-in

$$k_n = o(n)$$

$$p_n = \frac{k_n}{n} \to 0$$

Sublinear fan-in

Vanishing-density regime

- Fixed retained fraction: $p_n \to p > 0$.
- Sparse fan-in: $k_n = o(n)$, so $p_n = k_n/n \to 0$.
- These are different graph-limit regimes.

> *Existing Graphon NTK work assumes pruning masks converge to graphons at fixed sparsity level $p > 0$.*

NSIA
Network Science Informs AI

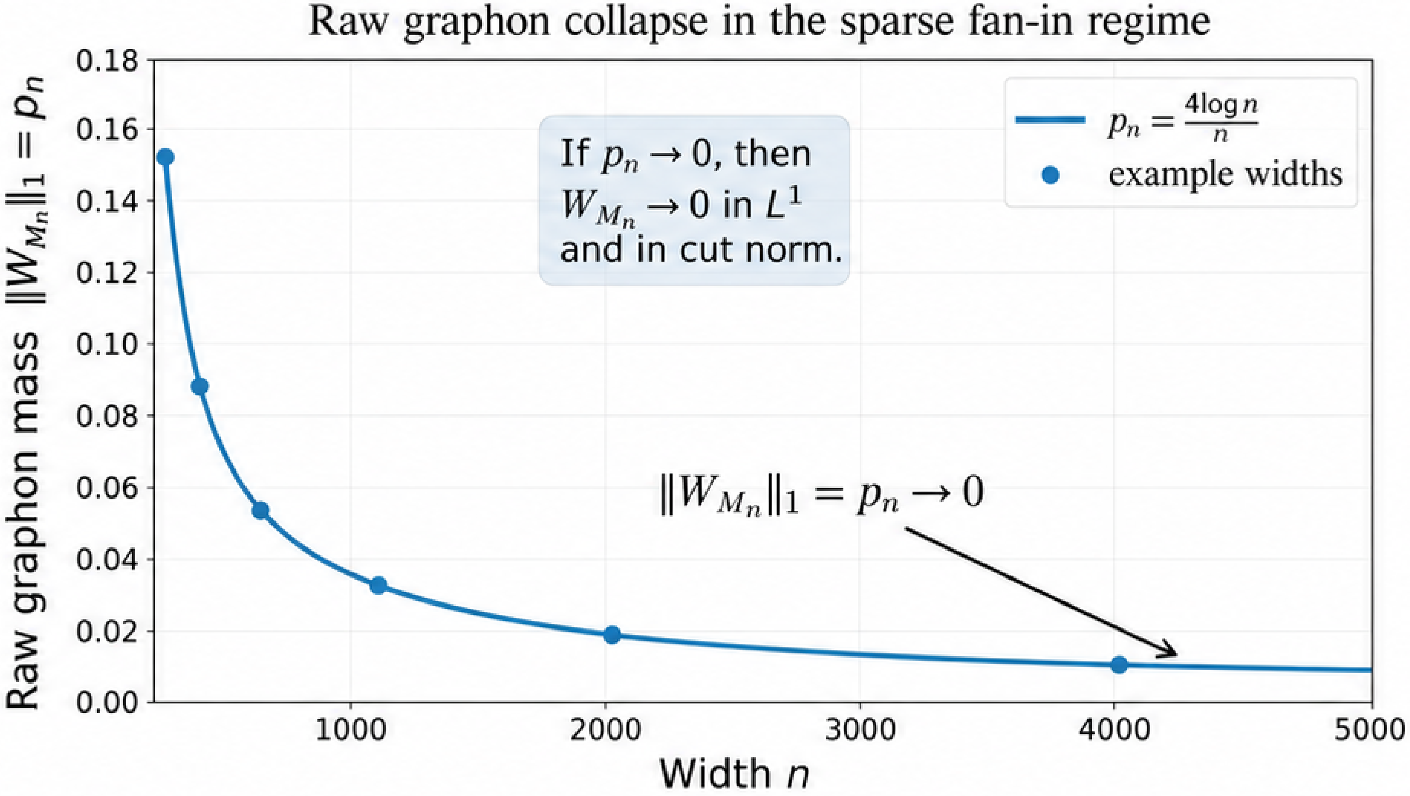
NetSci Boston 2026

3. Raw graphons collapse when density vanishes

$$\|W_{M_n}\|_1 = \frac{1}{n_L n_R}\sum_{i,j} M_{ij} = p_n.$$

$$p_n \to 0 \implies W_{M_n} \to 0 \text{ in } L^1 \text{ and cut norm.}$$

- This is not a failure of graphons in dense regimes.
- It is a density boundary.
- Raw adjacency mass disappears before normalized structure disappears.



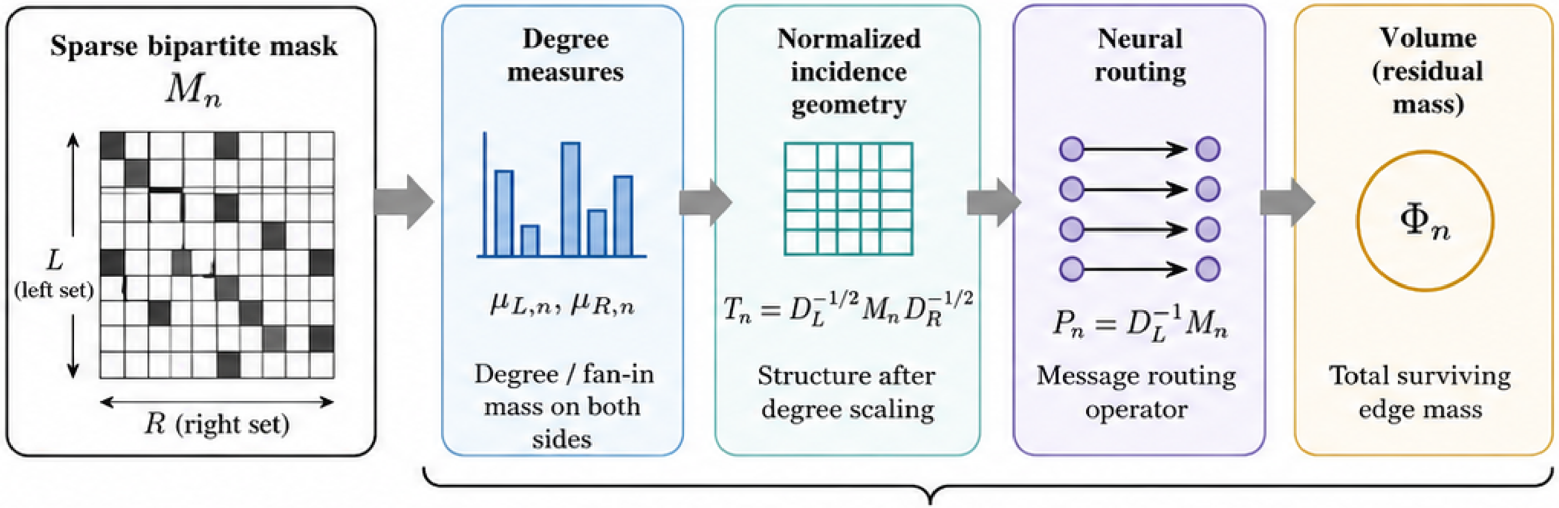
> *The theorem is almost embarrassingly simple, but it tells us exactly where the graphon coordinate loses information.*

4. Neural mask graphlets

$$G_n = (\mu_{L,n}, \mu_{R,n}, T_n, P_n, \Phi_n).$$

$$T_n = D_L^{-1/2} M_n D_R^{-1/2}, \qquad P_n = D_L^{-1} M_n.$$

- Φ_n: volume / total edge mass.
- μ_L, μ_R: degree mass on both sides.
- T_n: normalized incidence geometry.
- P_n: neural routing operator.



Graphlet decomposition: separates global sparsity, degree/fan-in mass, normalized geometry, routing, and residual structure beyond degree sequence.

> *The graphlet does not just ask how many edges survive. It asks how the surviving edge mass routes signals.*

NSIA Network Science Informs AI

NetSci Boston 2026

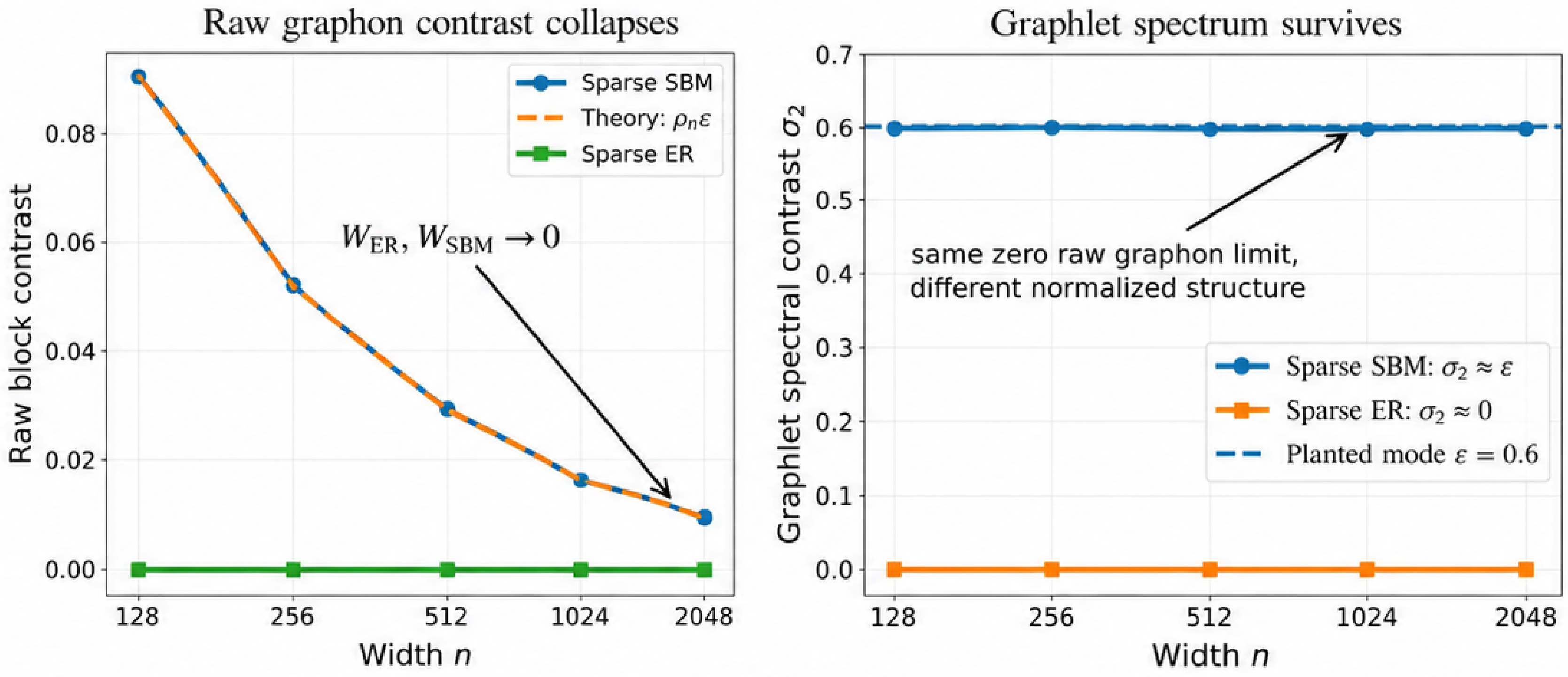
5. Same zero graphon limit, different sparse structure

sparse ER *vs.* sparse SBM

- Raw graphon contrast goes to 0 for both sparse ER and sparse SBM.
- Graphlet spectrum separates them.
- Raw graphons see both as zero.
- Graphlets retain the planted sparse mode.
- Normalization cancels the vanishing density factor.

$$\sigma_2(\mathrm{ER}) \approx 0, \quad \sigma_2(\mathrm{SBM}) \approx \varepsilon.$$

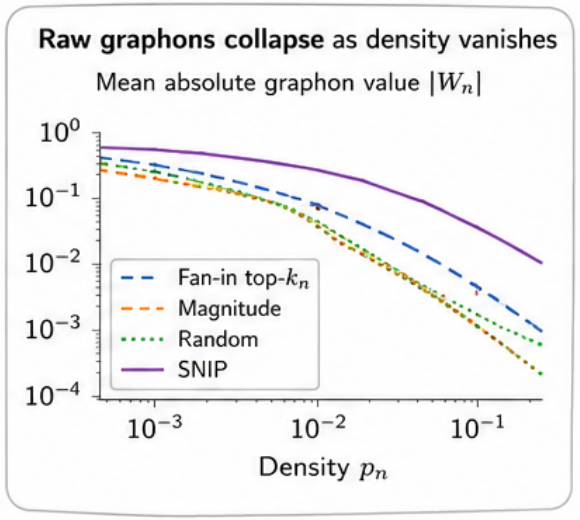
Sparse ER vs sparse SBM: graphon collapse, graphlet survival



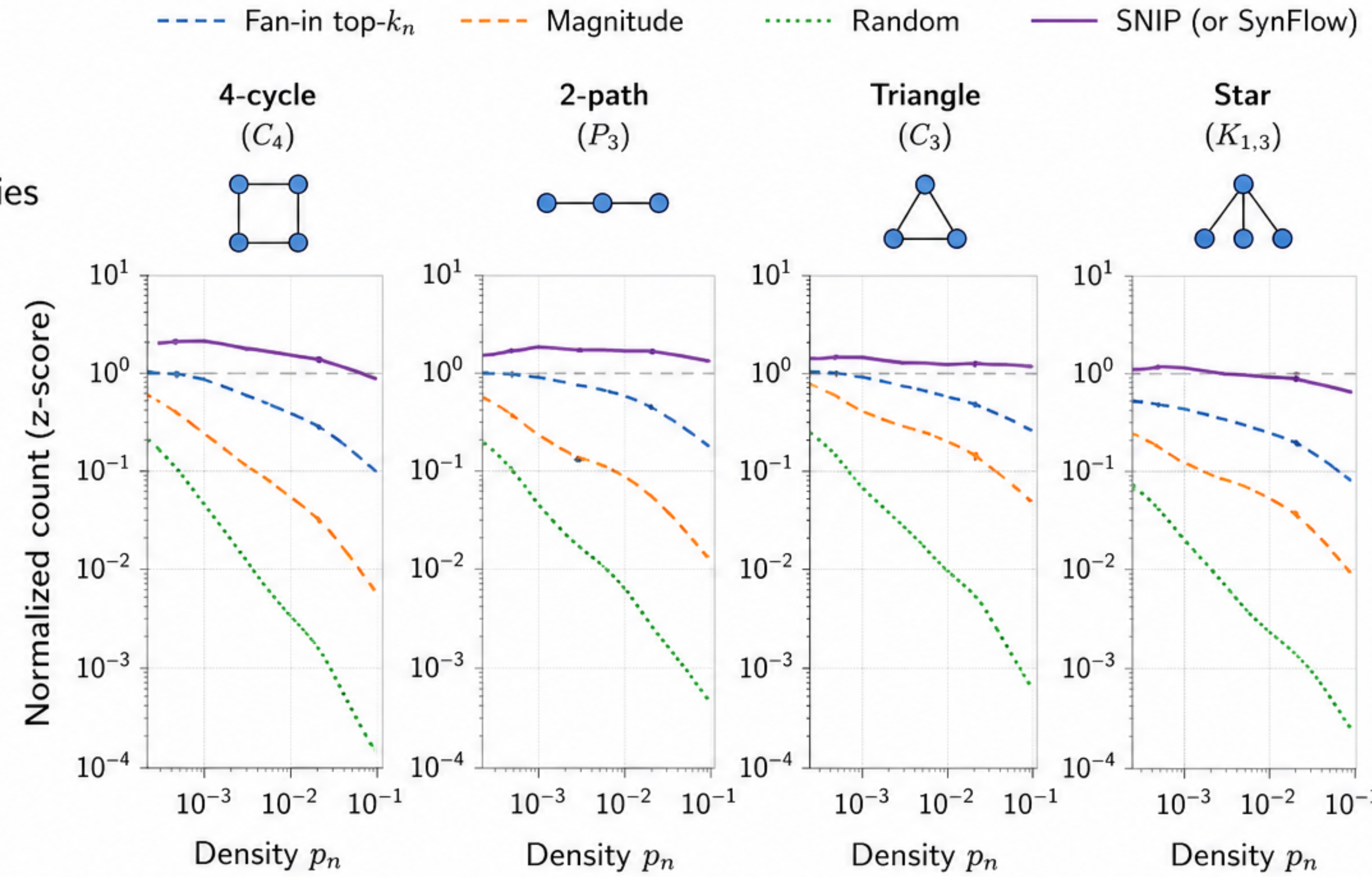
> *Raw graphons erase vanishing-density structure; graphlets retain the planted sparse mode.*

Neural masks remain distinguishable after normalization

- Fan-in top-k_n pruning gives $p_n = k_n/n \to 0$.
- Raw graphon summaries mostly shrink with density.
- Normalized graphlet summaries remain method-dependent.
- SNIP-like masks show stronger graphlet modes.



B. Normalized graphlet summaries (nontrivial as $p_n \to 0$)



Caution: These are finite-width diagnostics, not yet a full theorem for arbitrary learned pruning.

Toward sparse neural training dynamics

- Fan-in normalization makes **routing** central.
- Covariance propagation sees P, not raw M.
- This suggests a **Graphlet NTK** for $p_n \to 0$.

$$z_i(x) = \sum_j \frac{M_{ij}}{\sqrt{d_i}} \; W_{ij} \; h_j(x).$$

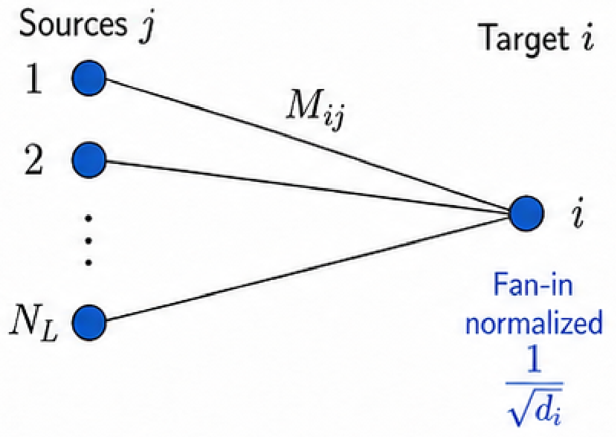
Covariance intuition:

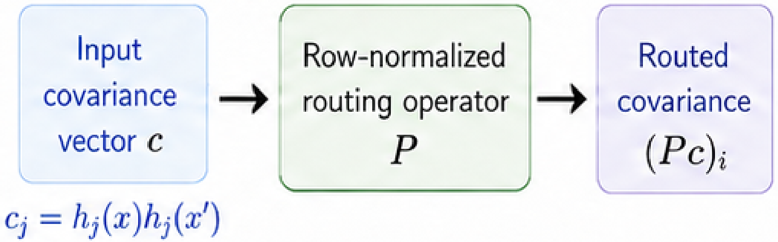
$$\mathrm{Cov}(z_i(x), z_i(x')) \sim \sum_j \frac{M_{ij}}{d_i} h_j(x) h_j(x') = (Pc)_i,$$

where $c_j = h_j(x) h_j(x')$.

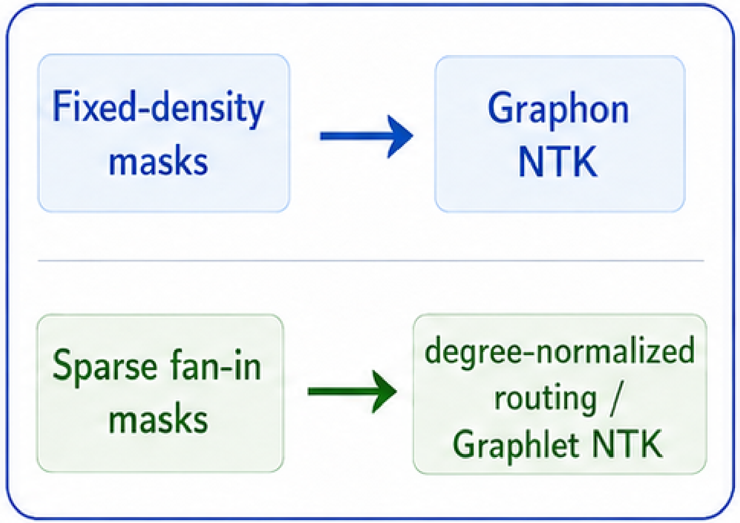
Therefore: $P = D_L^{-1} M$.

Fan-in normalized operator view





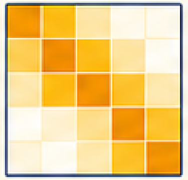
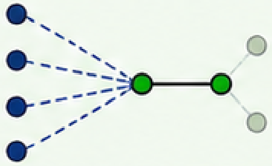
From graphon to sparse routing



Graphon NTK is natural for graphon-structured, fixed-density masks; sparse fan-in points instead to degree-normalized routing.

⚠ **Part of ongoing work.**

Graphons before density collapse. Graphlets after density collapse.

Regime	Condition $(n \to \infty)$		Interpretation
 Dense / fixed-fraction regime	$p_n \to p > 0$	$\Rightarrow$	graphons / Graphon NTK
 Sparse fan-in regime	$p_n \to 0$	$\Rightarrow$	graphlets / sparse routing operators

- Pruning masks are networks.
- Dense-style graphons are appropriate before density collapse.
- Sparse fan-in regimes require degree-aware graphlet coordinates.
- Training dynamics should be studied through normalized routing.

> *"Graphons tell us what survives before density collapse. Graphlets tell us what survives after density collapse."*

Thank You

Questions?

Key references

- Lovász, L. and Szegedy, B. (2006). Limits of Dense Graph Sequences.
- Lovász, L. (2012). Large Networks and Graph Limits.
- Chung, F. (2012). Graphlets: A Spectral Perspective for Graph Limits.
- Pham, H., Ta, T.-A., Jacobs, T., Burkholz, R., and Tran-Thanh, L. (2025).
  The Graphon Limit Hypothesis: Understanding Neural Network Pruning via Infinite Width Analysis.
- Pham, H., Ta, T.-A., and Tran-Thanh, L. (2026). Pruning at Initialisation through the Lens of Graphon Limit: Convergence, Expressivity, and Generalisation.