

Graphons in Machine Learning

Sevvandi Kandanaarachchi, Cheng Soon Ong
CSIRO

NeurIPS Education Track 2026



Applications

- Transferability of GNNs
  - Graphon neural networks and the transferability of graph neural networks
- Neural Network Pruning
  - The Graphon Limit Hypothesis: Understanding Neural Network Pruning via Infinite Width Analysis

Ruiz, L., Chamon, L., & Ribeiro, A. (2020). Graphon neural networks and the transferability of graph neural networks. *Advances in Neural Information Processing Systems*, *33*, 1702-1712.

Pham, H., Jacobs, T., Burkholz, R., & Tran-Thanh, L. (2026). The graphon limit hypothesis: Understanding neural network pruning via infinite width analysis. *Advances in Neural Information Processing Systems*, *38*, 53283-53319.

Transferability of GNNs



Graphon NNs and Transferability

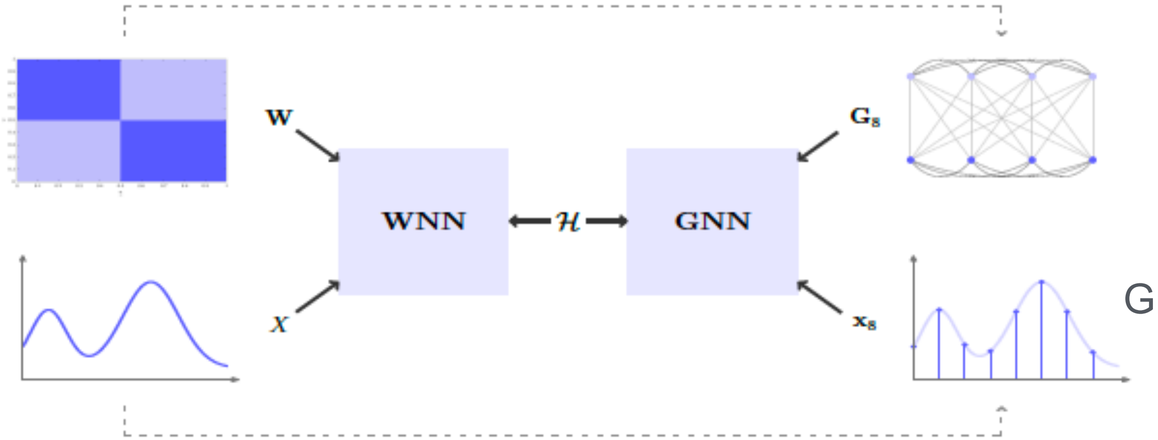
Theorem (GNN transferability, informal) Let $\mathbf{\Phi}(\mathbf{G})$ be a GNN with fixed parameters. Let $\mathbf{G}_{n_1}$ and $\mathbf{G}_{n_2}$ be deterministic graphs with n_1 and n_2 nodes obtained from a graphon $\mathbf{W}$. Under mild conditions, $\|\mathbf{\Phi}(\mathbf{G}_{n_1}) - \mathbf{\Phi}(\mathbf{G}_{n_2})\| = \mathcal{O}(n_1^{-0.5} + n_2^{-0.5})$.

Number of nodes	$n = 271$	$n = 542$	$n = 1354$
Accuracy on subnetwork $\mathbf{G}_n$	25.3%	29.5%	36.6%
Accuracy on full network	27.9%	35.4%	41.8%
Relative accuracy difference	23.1%	22.5%	15.4%

Accuracy on node classification using Cora citation network

Ruiz, L., Chamon, L., & Ribeiro, A. (2020). Graphon neural networks and the transferability of graph neural networks. *Advances in Neural Information Processing Systems*, *33*, 1702-1712.

Graphon NNs



Graph signal = data at node i

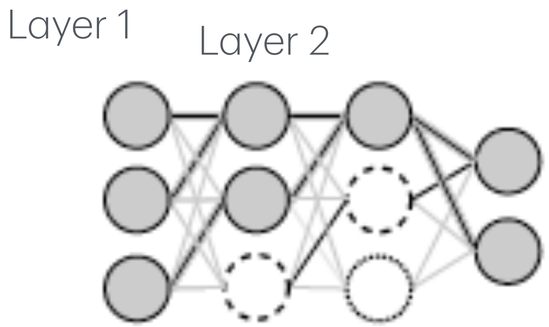
Figure 1: Example of a graphon neural network (WNN) given by $\mathbf{\Phi}(\mathcal{H}; \mathbf{W}; X)$, and of a graph neural network (GNN) instantiated from it as $\mathbf{\Phi}(\mathcal{H}; \mathbf{S}_8; \mathbf{x}_8)$. The graphon $\mathbf{W}$, shown on the top left corner, is a stochastic block model with intra-community probability $p_{c_i c_i} = 0.8$ and inter-community probability $p_{c_i c_j} = 0.2$, and the graphon signal X is plotted on the bottom left corner. The graph $\mathbf{G}_8$ (top right corner) and the graph signal $\mathbf{x}_8$ (bottom right corner) are obtained from $\mathbf{W}$ and X according to (11). Note that the parameter set $\mathcal{H}$ is shared between the WNN and the GNN.

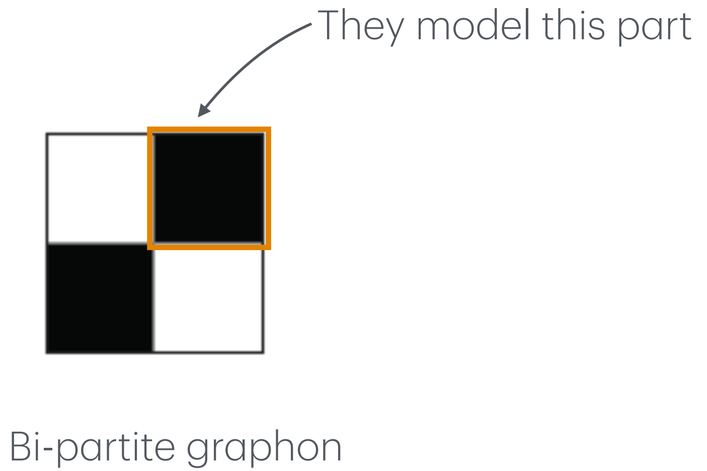
Ruiz, L., Chamon, L., & Ribeiro, A. (2020). Graphon neural networks and the transferability of graph neural networks. *Advances in Neural Information Processing Systems*, *33*, 1702-1712.

NN Pruning



NN Pruning at Initialization





- Bi-partite graph from layer 1 to layer 2

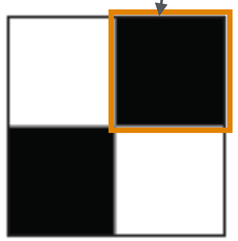
Given a sparsity level, each network pruning method defines sequences of binary masks $(M_n^{(l)})$ ***that converges layer-wisely to graphons*** $W^{(l)}$ ***in cut distance as the network width*** $n \to \infty$.

Pham, H., Liu, S., Xiang, L., Le, D., Wen, H., & Tran-Thanh, L. (2023). Towards data-agnostic pruning at initialization: What makes a good sparse mask?. *Advances in neural information processing systems*, *36*, 80044-80065.

Pham, H., Jacobs, T., Burkholz, R., & Tran-Thanh, L. (2026). The graphon limit hypothesis: Understanding neural network pruning via infinite width analysis. *Advances in Neural Information Processing Systems*, *38*, 53283-53319.

NN Pruning at Initialization

They model this part



Bi-partite graphon

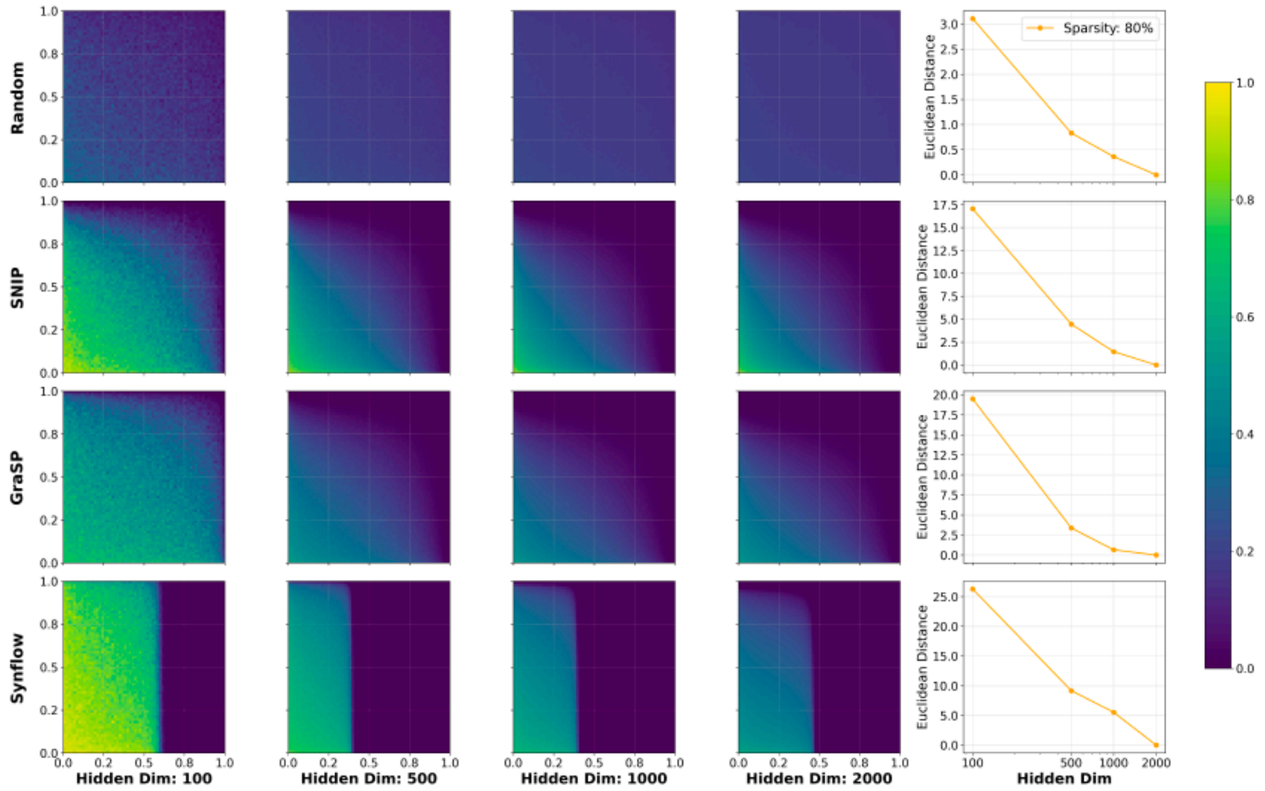


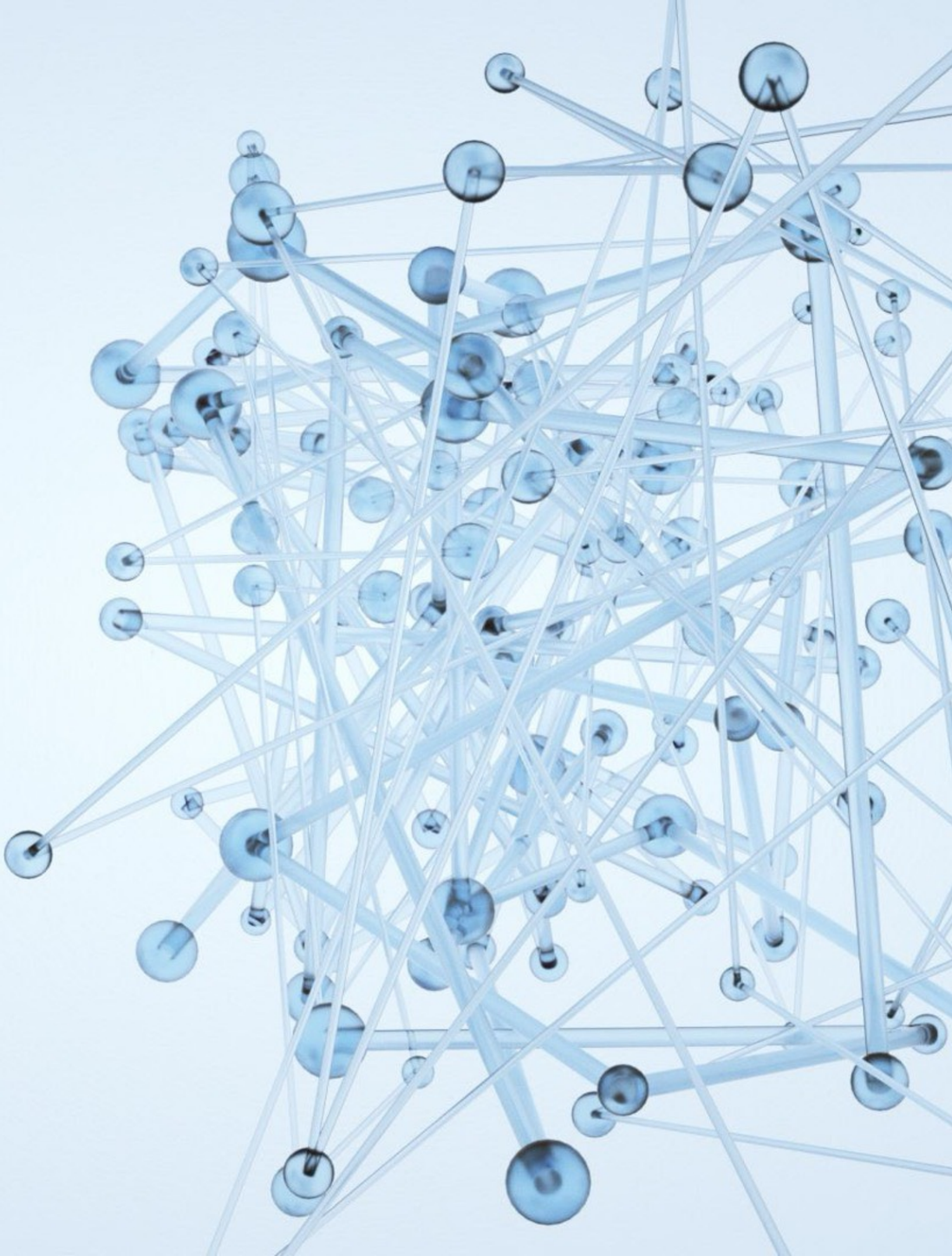
Figure 1: Graph limit of subnetworks' mask produced by PaI methods at 80% sparsity and the corresponding convergence of graphons via Euclidean distances.

Pham, H., Liu, S., Xiang, L., Le, D., Wen, H., & Tran-Thanh, L. (2023). Towards data-agnostic pruning at initialization: What makes a good sparse mask?. *Advances in neural information processing systems*, *36*, 80044-80065.

Pham, H., Jacobs, T., Burkholz, R., & Tran-Thanh, L. (2026). The graphon limit hypothesis: Understanding neural network pruning via infinite width analysis. *Advances in Neural Information Processing Systems*, *38*, 53283-53319.

CSIRO

Thank you!



NEURAL INFORMATION PROCESSING SYSTEMS