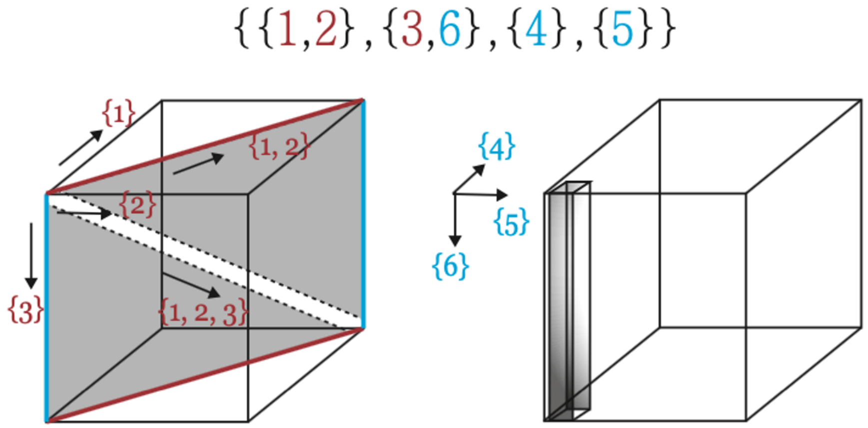
Convergence of Invariant Graph Networks

Chen Cai (UCSD) & Yusu Wang (UCSD)

ICML 2022

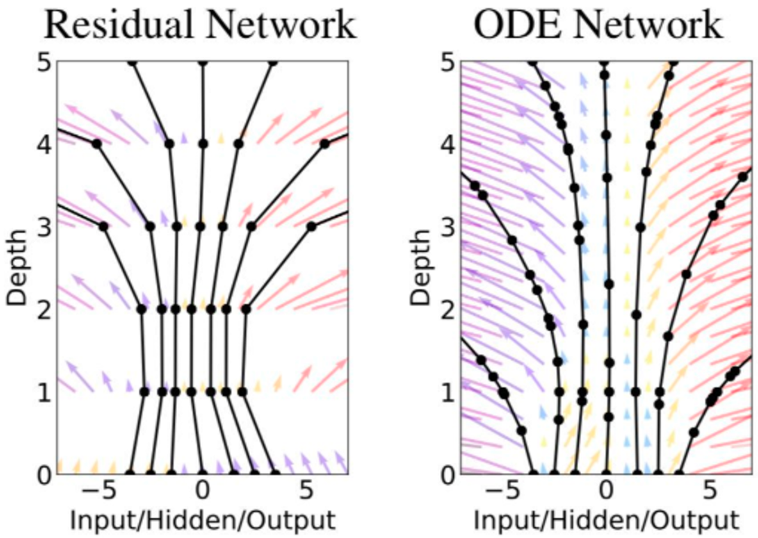


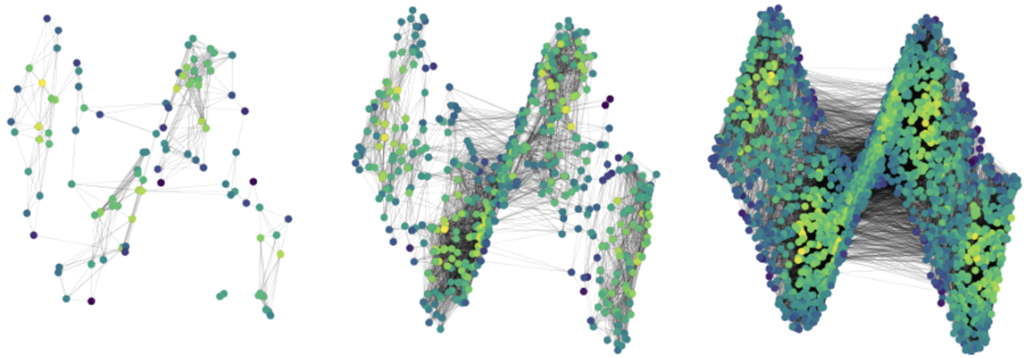
Convergence in Deep Learning

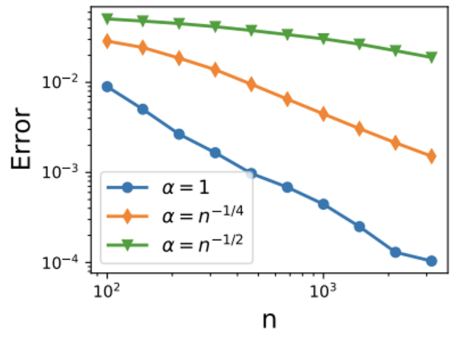
- Increasing width: Neural Tangent Kernel
- Increase depth: Neural ODE
- Increase input size? Convergence of graph neural network!

$$\Theta^{(L)}(x,y) = \sum_{p=1}^{P} \frac{d}{d\theta_p} f_\theta(x) \, \frac{d}{d\theta_p} f_\theta(y)$$

depth — two samples — all parameters

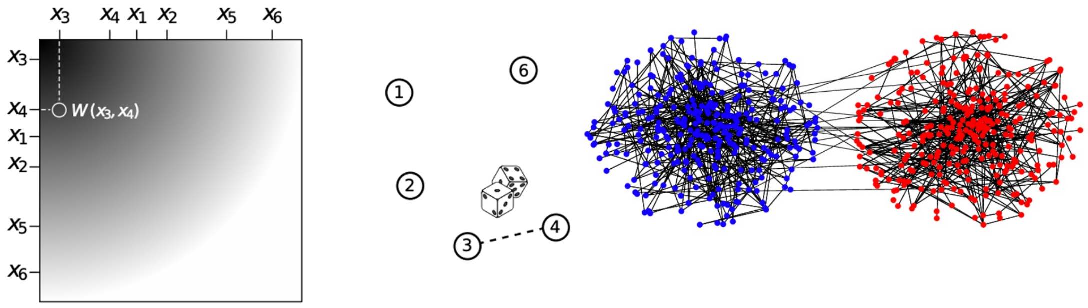




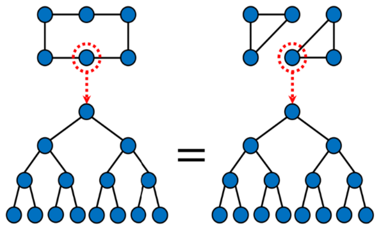


Setup & Existing work

- Model
  - graphon $W: [0,1]^2 \to [0,1]$
  - edge probability discrete model
  - edge weight continuous model
- Mainly study spectral GNN, which has limited expressive power
- What about more powerful GNN?



Connect nodes 3 and 4 with probability $W(x_3, x_4)$.

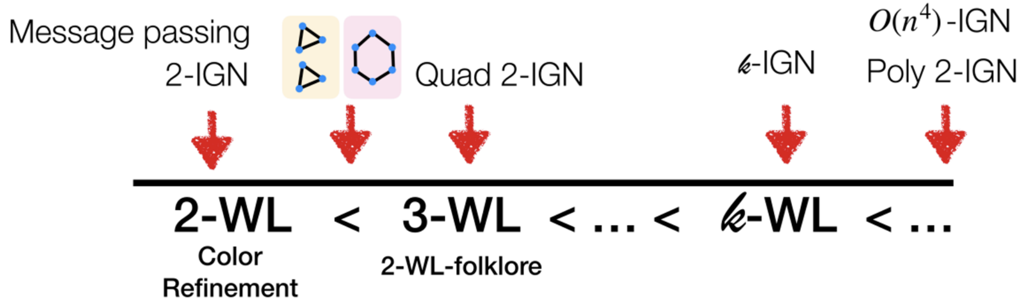


Study the Convergence of Invariant Graph Networks (IGN)

Invariant Graph Network (IGN)

- $F = h \circ L^{(T)} \circ \sigma \cdots \circ \sigma \circ L^{(1)}$
- GNN needs to be permutation equivaraint
- Characterize *linear permutation equivariant* functions
- 15 functions for $\mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$

Theorem [Maron et al 2018]: The space of linear permutation equivariant functions $\mathbb{R}^{n^l} \rightarrow \mathbb{R}^{n^m}$ is of dimension $bell(l+m)$ (number of partitions of set $\{1, 2, \ldots, l+m\}$



Summary

A novel interpretation of basis of the space of equivariant maps in k-IGN

Edge weight continuous model

- Convergence of 2-IGN and k-IGN
- For both deterministic and random sampling

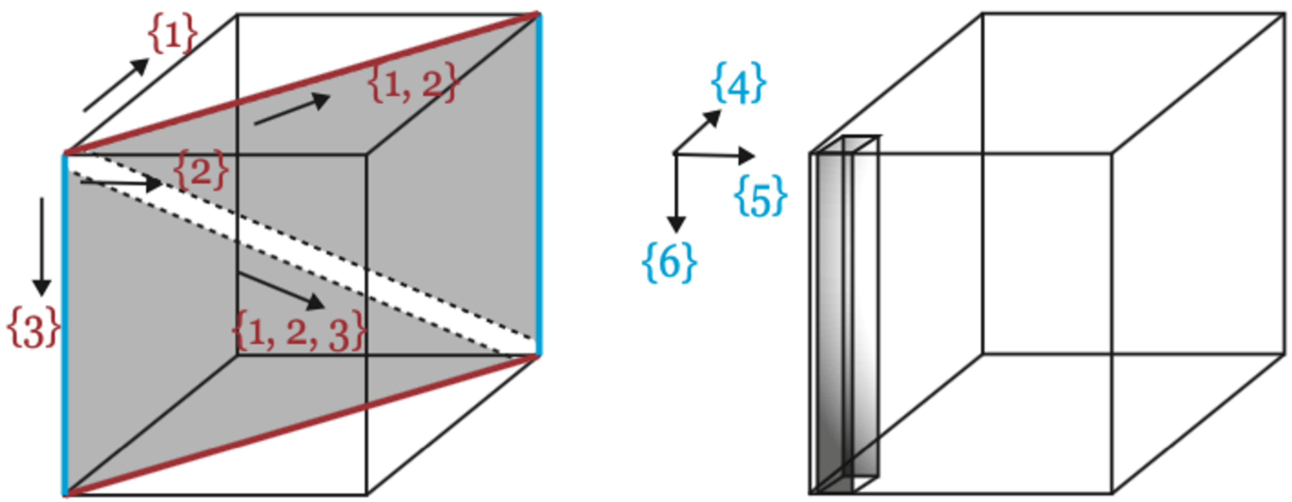
Edge probability discrete model

- Negative result in general
- Convergence of IGN-small after edge probability estimation
- IGN-small approximates spectral GNN arbitrarily well

Space of linear (permutation) equivariant maps

- from l-tensor to m-tensor
- dimension is $bell(l + m)$

$$\{\{1,2\},\{3,6\},\{4\},\{5\}\}$$



$$\underbrace{S_1 = \{\{1,2\}\}}_{\text{Only has input axis}} \cup \underbrace{S_2 = \{\{3,6\}\}}_{\text{has both input and output axis}} \cup \underbrace{S_3 = \{\{4\},\{5\}\}}_{\text{only has output axis}}$$

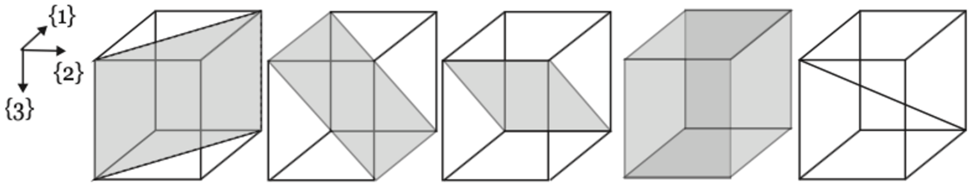
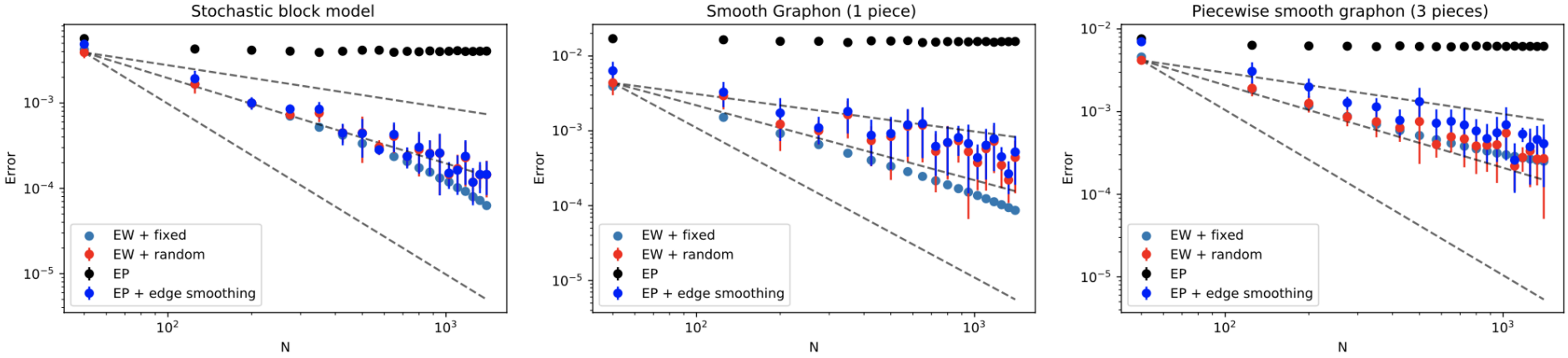


Figure 1: Five possible "slices" of a 3-tensor, corresponding to bell(3) = 5 paritions of [3]. From left to right: a) $\{\{1,2\},\{3\}\}$ b) $\{\{1\},\{2,3\}\}$ c) $\{\{1,3\},\{2\}\}$ d) $\{\{1\},\{2\},\{3\}\}$ e) $\{\{1,2,3\}\}$.

Experiments



Thank You for Listening!

Hall E #411 6pm-8pm