HyperGCN: A New Method for Training Graph Convolutional Networks on Hypergraphs

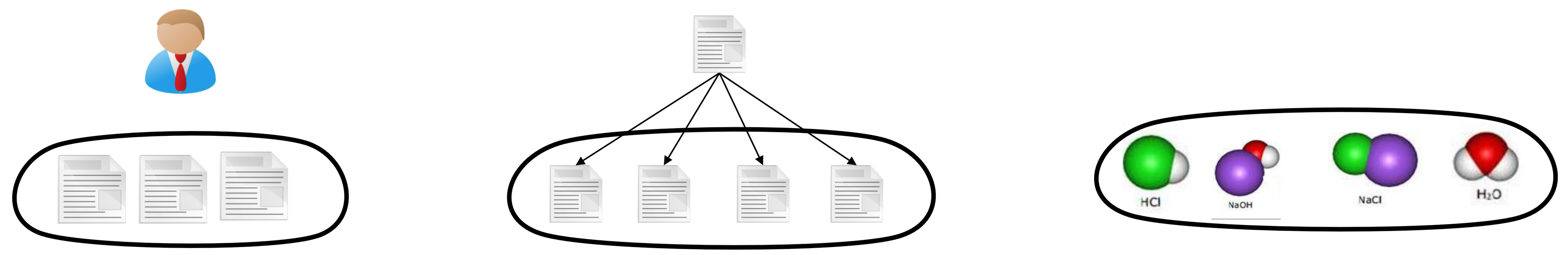
Naganand Yadati, Madhav Nimishakavi, Prateek Yadav, Vikram Nitin, Anand Louis, Partha Talukdar

Indian Institute of Science, Bangalore

naganand@iisc.ac.in

Hypergraph Motivation

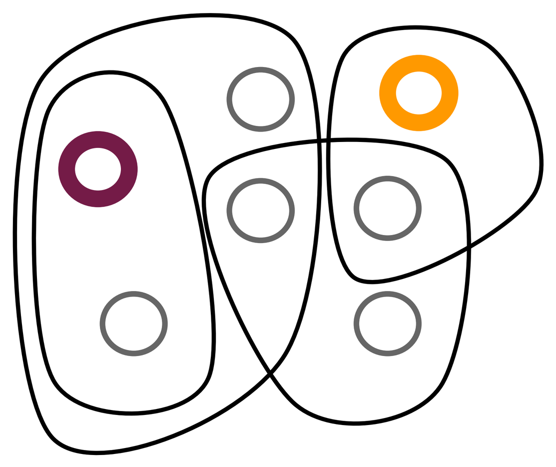
Networks have relationships beyond pairwise



co-authorship co-citation chemical reaction

Modelled flexibly by hypergraphs

Hypergraph Semi-Supervised Learning



$\mathcal{H} = (V, E, X)$

V : set of vertices

E : set of edges

X : matrix of vertex features

<u>Problem</u>

Label the unlabelled vertices in U

given labelled vertices in $V - U$ of hypergraph $\mathcal{H}$

<u>Challenges</u>

- Arbitrary size $E \subseteq 2^V$
- Low supervision $|V - U| \ll |U|$
- Noisy edges $|\{y_v : v \in e\}| > 1, \ e \in E$

<u>Approaches</u>

- Explicit regularisation [Zhou et al., Hein et al.]

$\mathcal{L} = \mathcal{L}_S + \lambda \cdot Q(\mathcal{H}, f)$ ✗ hyperedges encode similarity

- Implicit regularisation [Feng et al.]

$f_{Neural}\left(\mathcal{H}, X\right) = ?$ $\mathcal{L} = \mathcal{L}_S$ ✓ need not encode similarity

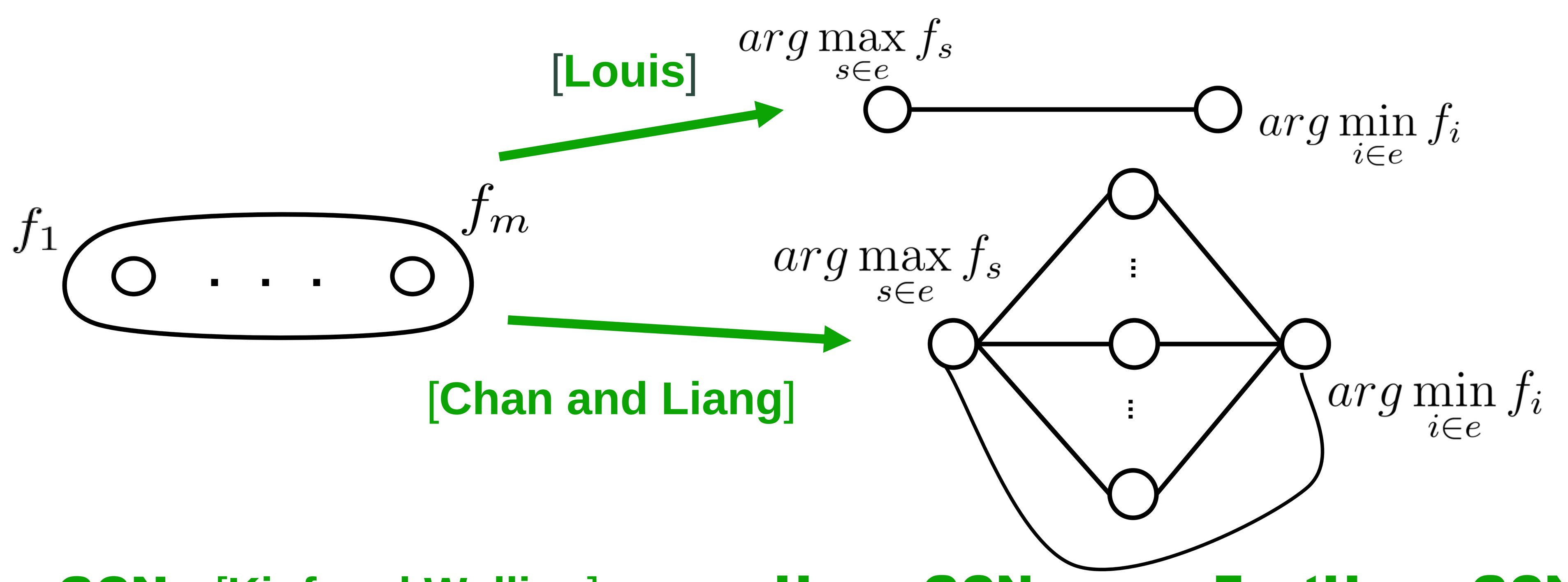
HyperGCN

<u>Idea</u>

If the maximally disparate vertices of a hyperedge are close then so are all vertices of the hyperedge

f : vertex signals

e.g., embeddings



GCN [Kipf and Welling]

$H^l = \sigma(A \cdot H^{l-1} \cdot W^l)$

HyperGCN

$f = H^{\{l-1\}} \cdot W^{\{l\}}$

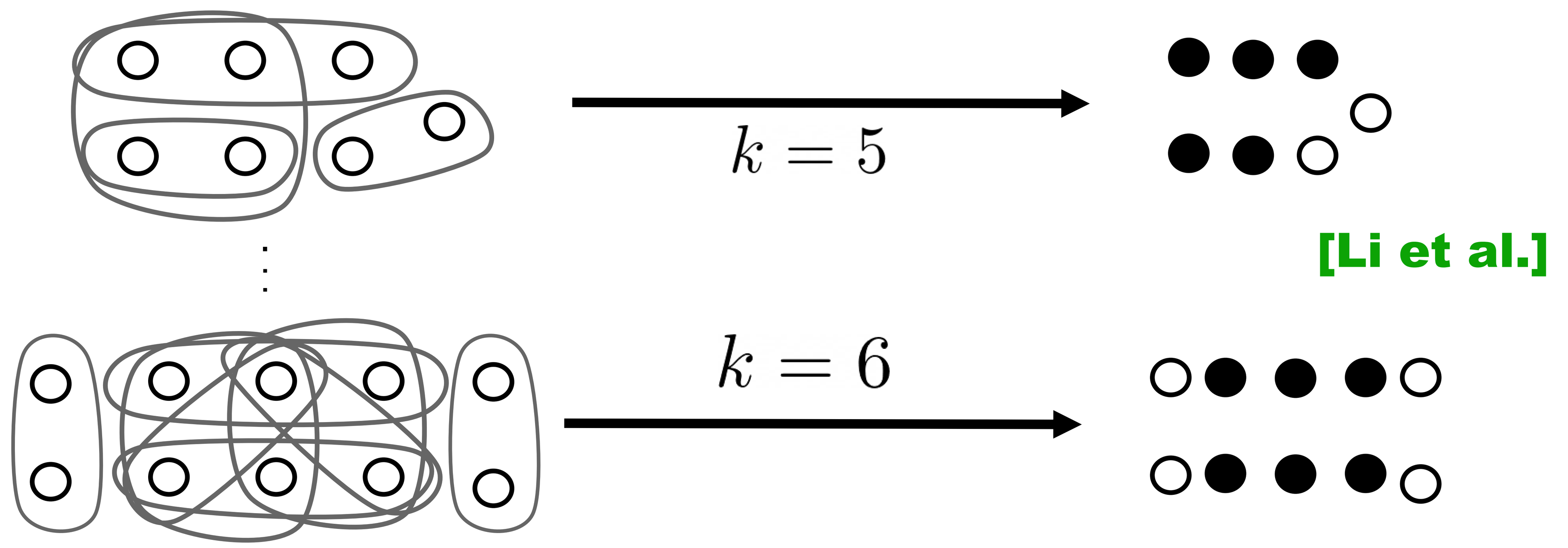
FastHyperGCN

$f = H^{\{0\}} = X$

HyperGCN for Combinatorial Optimisation

Hyperedges need not encode similarity e.g., Densest k-Subhypergraph problem

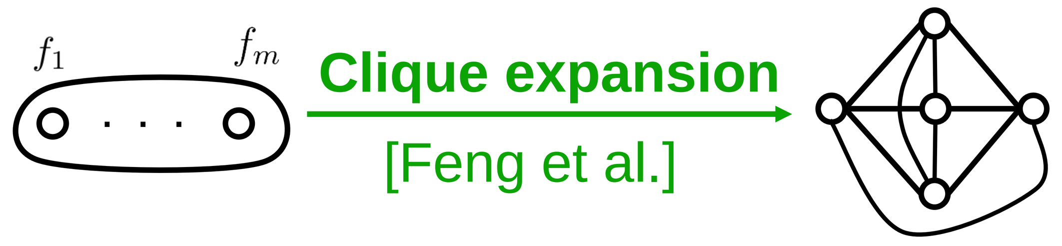
$$W \subseteq V, |W| = k, \ \text{maximise} \ |e \in E : \ e \subseteq W|$$



<u>Train</u>: Synthetic hypergraphs with optimal solutions, <u>Test</u>: Real-world hypergraphs

Results

Simple baseline: HyperGraph Neural Net



test error (lower is better)

Avg. size	3.0 ± 1.1	3.2 ± 2.0
Dataset	Cora	Citeseer
HGNN	32.41 ± 1.8	37.40 ± 1.6
HyperGCN	32.37 ± 1.7	37.35 ± 1.6
FastHyperGCN	32.42 ± 1.8	37.42 ± 1.7

comparable on small hyperedges

test error (lower is better)

Avg. size	8.5 ± 8.8	4.3 ± 5.7
Dataset	DBLP	Pubmed
HGNN	45.27 ± 2.48	29.41 ± 1.5
HyperGCN	$\mathbf{41.64 \pm 2.6}$	$\mathbf{25.56 \pm 1.6}$
FastHyperGCN	41.79 ± 2.8	29.48 ± 1.6

training time (lower is better)

Dataset	DBLP	Pubmed
HGNN	0.115s	0.019s
FastHyperGCN	**0.035s**	**0.016s**

more accurate, faster on large noisy hyperedges

density for k = 0.75*|V| (higher is better)

Dataset	DBLP	Pubmed	Cora	Citeseer
HGNN	6274	7865	437	**969**
HyperGCN	**7720**	**7928**	**504**	**971**
FastHyperGCN	7342	7893	452	**969**

training time (lower is better)

Type	Training Time	Density
HGNN	170s	337
FastHyperGCN	**143s**	**352**

Contributions Summary

- Bridge deep learning, spectral hypergraph theory
- Approximate hyperedge by a linear no. edges
- Show gains on large, noisy hyperedges

Future Work

- Unsupervised learning
- Label correlation
- Hypergraph pooling

Acknowledgement

Google MHRD GOVERNMENT OF INDIA

[Zhou et al.] D. Zhou, J. Huang, B. Scholkopf. Learning with Hypergraphs: Clustering, classification and embedding, *In NeurIPS 2006*

[Hein et al.] M. Hein, S. Setzer, L. Zost, S.S.Rangapuram. The Total Variation on Hypergraphs, *In NeurIPS 2013*

[Louis] A. Louis. Hypergraph Markov Operators, Eigenvalues and Approximation Algorithms, *In STOC 2015*

[Kipf and Welling] T. Kipf, M. Welling. Semi-supervised Classification with graph convolutional networks, *In ICLR 2017*

[Chan and Liang] T. H. Hubert Chan, Z. Liang. Generalizing the hypergraph laplacian via a diffusion process with mediators, *In COCOON 2018*

[Li et al.] Z. Li, Q. Chen, V. Koltun. Combinatorial optimization with graph convolutional networks and guided tree search, *In NeurIPS 2018*

[Feng et al.] Y. Feng, H. You, Z. Zhang, R. Ji, Y. Gao. Hypergraph Neural Networks, *In AAAI 2019*