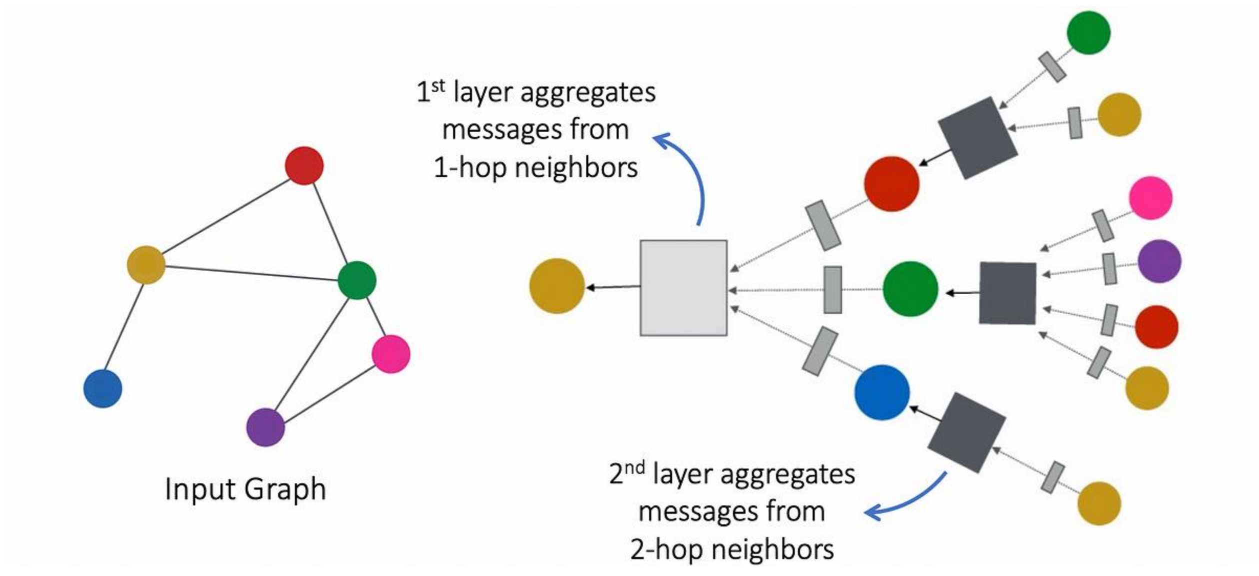
How Universal Polynomial Bases Enhance Spectral Graph Neural Networks: Heterophily, Over-smoothing, and Over-squashing

1 NUS National University of Singapore
2 Shanghai Jiao Tong University
3 Zhejiang Normal University
4 University of Cambridge

Keke Huang[1], Yu Guang Wang[2], Ming Li[3], Pietro Liò[4]

Introduction

- Message Passing in Graph Neural Networks (GNNs)



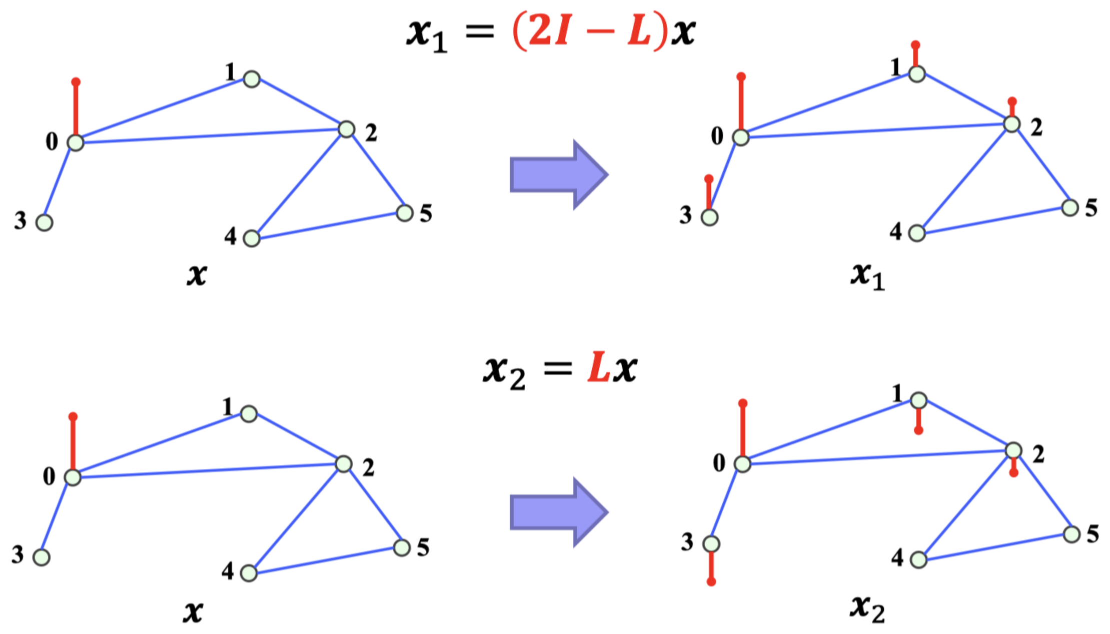
- Spectral Graph Neural Networks
  - Spectral Graph Filters on Graph Signals

$$L = I - D^{-1/2}AD^{-1/2} = U\Lambda U^T = U\begin{pmatrix}\lambda_1 & & \\ & \dots & \\ & & \lambda_n\end{pmatrix}U^T$$

  - Laplacian Matrix on Graph Signals

$$\boldsymbol{x}_1 = (2\boldsymbol{I} - \boldsymbol{L})\boldsymbol{x}$$

$$\boldsymbol{x}_2 = \boldsymbol{L}\boldsymbol{x}$$



- Learnable Spectral Graph Filters
  - Spectral filter $g_w(\cdot)$ parameterized by parameter $w \in R^n$

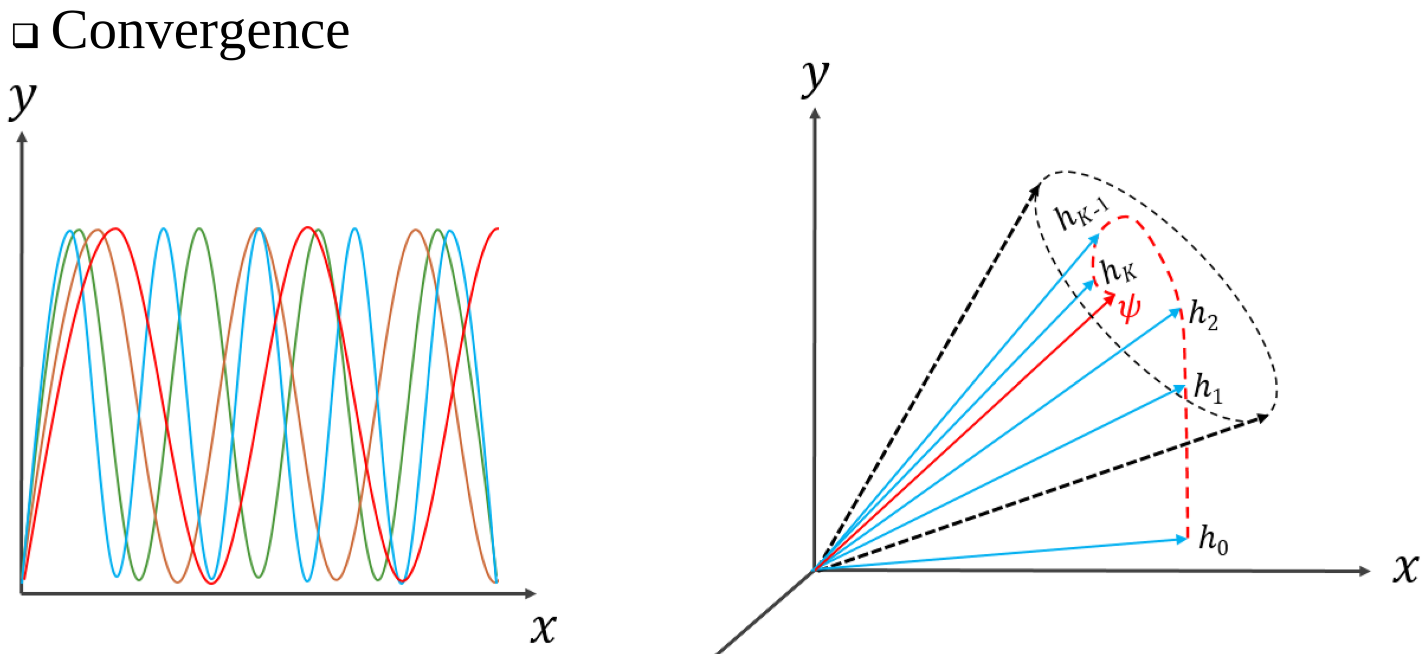
$$g_w(L)\cdot \mathrm{x} = Ug_w(\Lambda)U^T \cdot \mathrm{x}$$

$$g_w(\Lambda) = \text{ diag }[g_w(\lambda_1), g_w(\lambda_2), \dots, g_w(\lambda_n)]$$

  - Challenge
    - Time complexity: $O(n^3)$
- Polynomial Approximation
  - A graph signal x, an integer K, and propagation matrix $P \in R^{n\times n}$

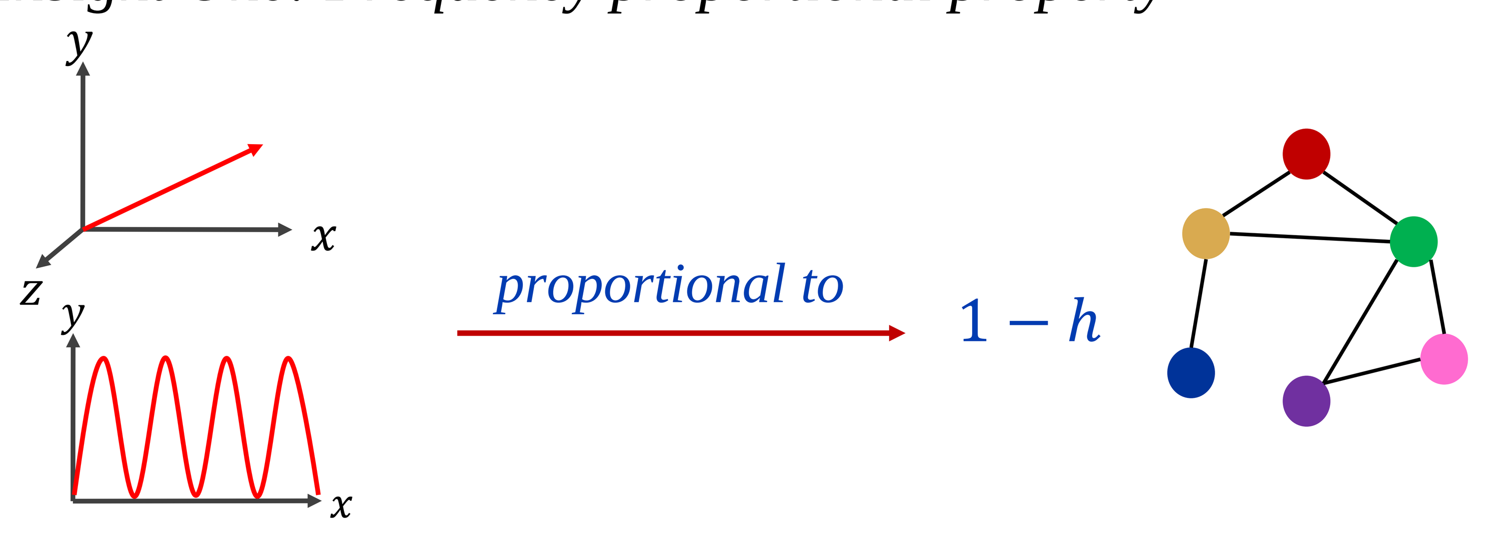
$$\mathrm{z} = Ug_{w(\Lambda)}U^T \cdot \mathrm{x} \approx \sum_{k=0}^{K} w_k P^k \cdot \mathrm{x}$$

- Limitations
  - Lack of adaptability due to the *fixed* polynomial bases
  - Convergence

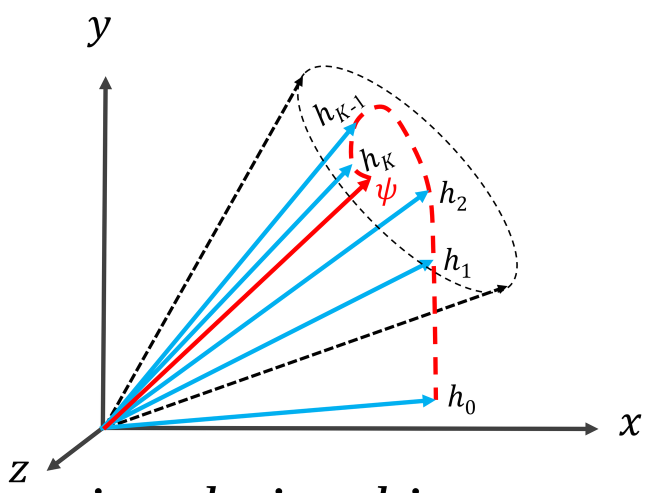


Universal Polynomial Bases

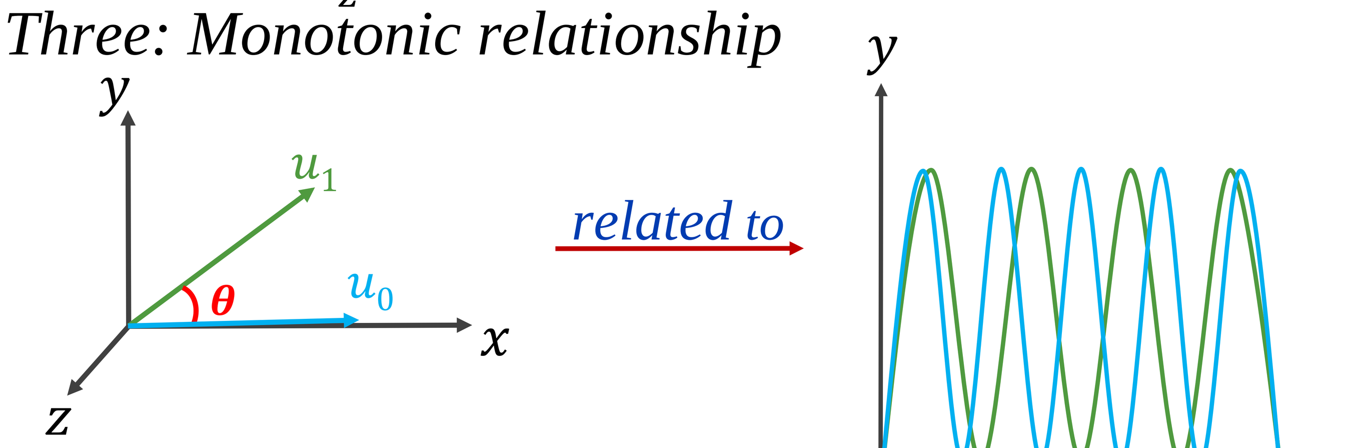
- *Insight One: Frequency proportional property*



- *Insight Two: Asymptotical convergence*

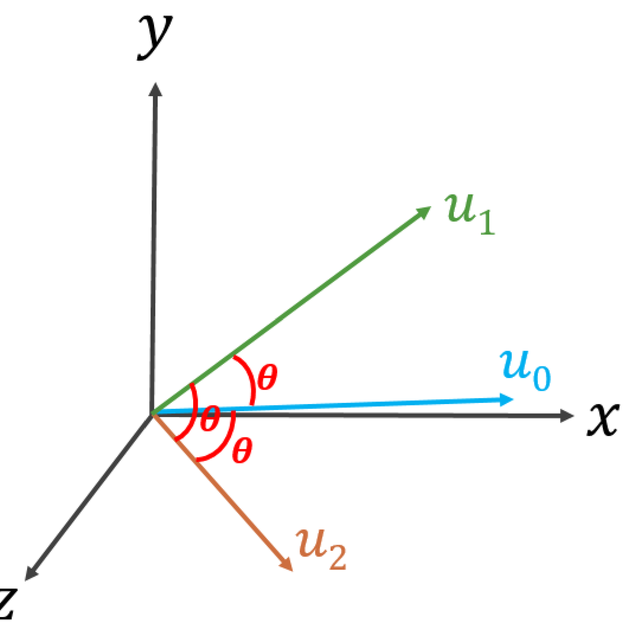


- *Insight Three: Monotonic relationship*



- *UniBasis & UniFilter*

$$\boldsymbol{\theta} = \frac{(1-h)\pi}{2}$$



  - *UniBasis*

$$\{\mathrm{x} + u_0, P\mathrm{x} + u_1, P^2\mathrm{x} + u_2, \dots, P^K\mathrm{x} + u_K\}$$

  - *UniFilter*

$$z^K = \sum_{k=0}^{K} w_k(\tau P^k\mathrm{x} + (1-\tau)u_k) \quad (\tau \in [0,1])$$

- Over-Smoothing

$$\lim_{K\to\infty} E(G, z^K) = (1-\tau)E(G, \mathrm{x})$$

- Over-Squashing

$|(\partial z_u^k)/(\partial z_v^k)|$ independent of propagation step k.

Experiments

Table 1: Accuracy (%) compared with polynomial filters.

Methods	Cora	Citeseer	Pubmed	Actor	Chameleon	Squirrel
SGC	86.83 ± 1.28	79.65 ± 1.02	87.14 ± 0.90	34.46 ± 0.67	44.81 ± 1.20	25.75 ± 1.07
SIGN	87.70 ± 0.69	80.14 ± 0.87	89.09 ± 0.43	41.22 ± 0.96	60.92 ± 1.45	45.59 ± 1.40
ASGC	85.35 ± 0.98	76.52 ± 0.36	84.17 ± 0.24	33.41 ± 0.80	71.38 ± 1.06	57.91 ± 0.89
GPR-GNN	88.54 ± 0.67	80.13 ± 0.84	88.46 ± 0.31	39.91 ± 0.62	67.49 ± 1.38	50.43 ± 1.89
EvenNet	87.77 ± 0.67	78.51 ± 0.63	90.87 ± 0.34	40.36 ± 0.65	67.02 ± 1.77	52.71 ± 0.85
ChebNet	87.32 ± 0.92	79.33 ± 0.57	87.82 ± 0.24	37.42 ± 0.58	59.51 ± 1.25	40.81 ± 0.42
ChebNetII	88.71 ± 0.93	80.53 ± 0.79	88.93 ± 0.29	41.75 ± 1.07	71.37 ± 1.01	57.72 ± 0.59
BernNet	88.51 ± 0.92	80.08 ± 0.75	88.51 ± 0.39	41.71 ± 1.12	68.53 ± 1.68	51.39 ± 0.92
JacobiConv	88.98 ± 0.72	80.78 ± 0.79	89.62 ± 0.41	41.17 ± 0.64	74.20 ± 1.03	57.38 ± 1.25
OptBasisGNN	87.00 ± 1.55	80.58 ± 0.82	90.30 ± 0.19	**42.39 ± 0.52**	74.26 ± 0.74	63.62 ± 0.76
Specformer	88.57 ± 1.01	81.49 ± 0.94	87.73 ± 0.58	41.93 ± 1.04	74.72 ± 1.29	64.64 ± 0.81
UniFilter	**89.49 ± 1.35**	**81.39 ± 1.32**	**91.44 ± 0.50**	40.84 ± 1.21	**75.75 ± 1.65**	**67.40 ± 1.25**

Table 2: Accuracy (%) compared with model-optimized methods.

Methods	Cora	Citeseer	Pubmed	Actor	Chameleon	Squirrel
GCN	86.98 ± 1.27	76.50 ± 1.36	88.42 ± 0.50	27.32 ± 1.10	64.82 ± 2.24	53.43 ± 2.01
GCNII	88.37 ± 1.25	77.33 ± 1.48	90.15 ± 0.43	37.44 ± 1.30	63.86 ± 3.04	38.47 ± 1.58
GAT	87.30 ± 1.10	76.55 ± 1.23	86.33 ± 0.48	27.44 ± 0.89	60.26 ± 2.50	40.72 ± 1.55
MixHop	87.61 ± 0.85	76.26 ± 1.33	85.31 ± 0.61	32.22 ± 2.34	60.50 ± 2.53	43.80 ± 1.48
H_2GCN	87.87 ± 1.20	77.11 ± 1.57	89.49 ± 0.38	35.70 ± 1.00	60.11 ± 2.15	36.48 ± 1.86
LINKX	84.64 ± 1.13	73.19 ± 0.99	87.86 ± 0.77	36.10 ± 1.55	68.42 ± 1.38	61.81 ± 1.80
WRGAT	88.20 ± 2.26	76.81 ± 1.89	88.52 ± 0.92	36.53 ± 0.77	65.24 ± 0.87	48.85 ± 0.78
ACM-GCN	87.91 ± 0.95	77.32 ± 1.70	90.00 ± 0.52	36.28 ± 1.09	66.93 ± 1.85	54.40 ± 1.88
GloGNN++	88.33 ± 1.09	77.22 ± 1.78	89.24 ± 0.39	37.70 ± 1.40	71.21 ± 1.84	57.88 ± 1.76
UniFilter	**89.12 ± 0.87**	**80.28 ± 1.31**	**90.19 ± 0.41**	**37.79 ± 1.11**	**73.66 ± 2.44**	**64.26 ± 1.46**

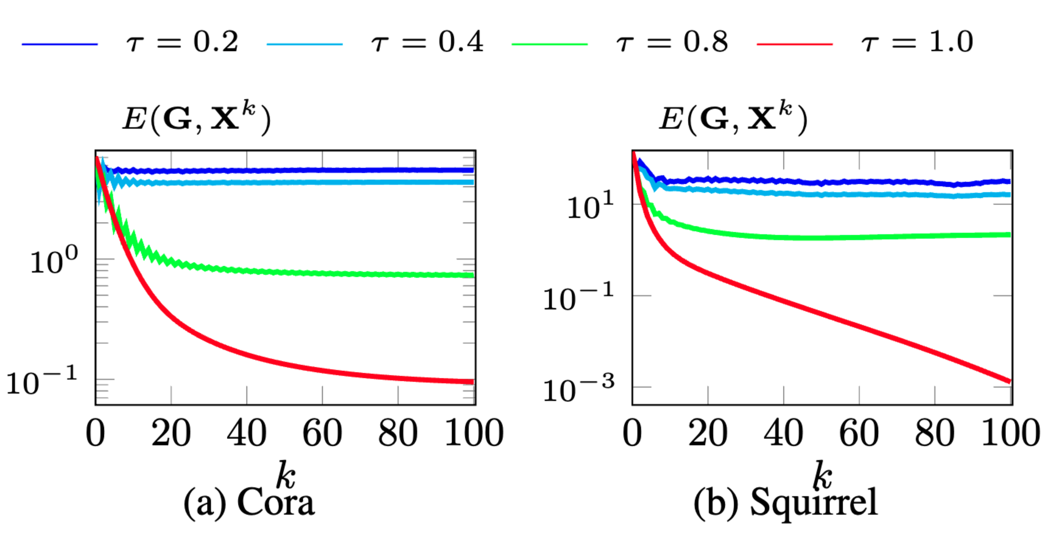


Figure 5: Dirichlet energy $E(\mathbf{G}, \mathbf{X}^k)$ with varying k.

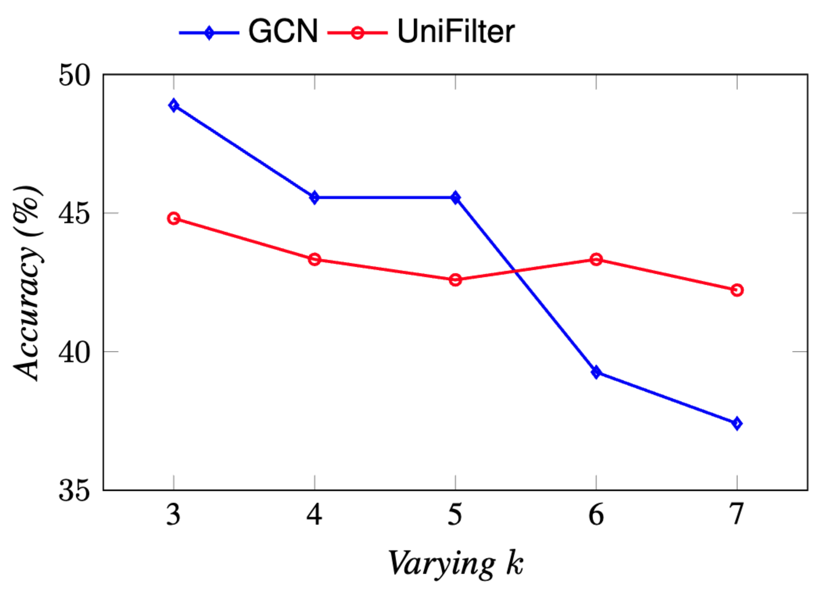


Figure 6: Accuracy (%) with varying k.