Less is More: SlimG for Accurate, Robust, and Interpretable Graph Mining




*Jaemin YOO**

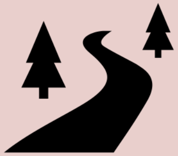
*Meng-Chieh (Jeremy) LEE**

Shubhranshu SHEKHAR

Christos FALOUTSOS

*Equal contributions.

Outline



- Problem Definition
- Q1: Reasons
- Q2: Method
- Q3: Sanity Checks
- Experiments
- Conclusions

Outline

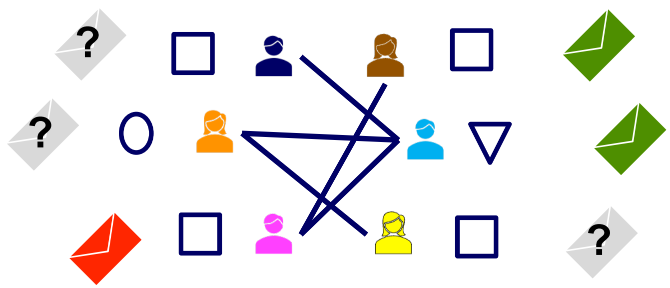
- Problem Definition
  - Motivation
  - Research Questions
- Q1: Reasons
- Q2: Method
- Q3: Sanity Checks
- Experiments
- Conclusions

Legend

→ Outline/Roadmap

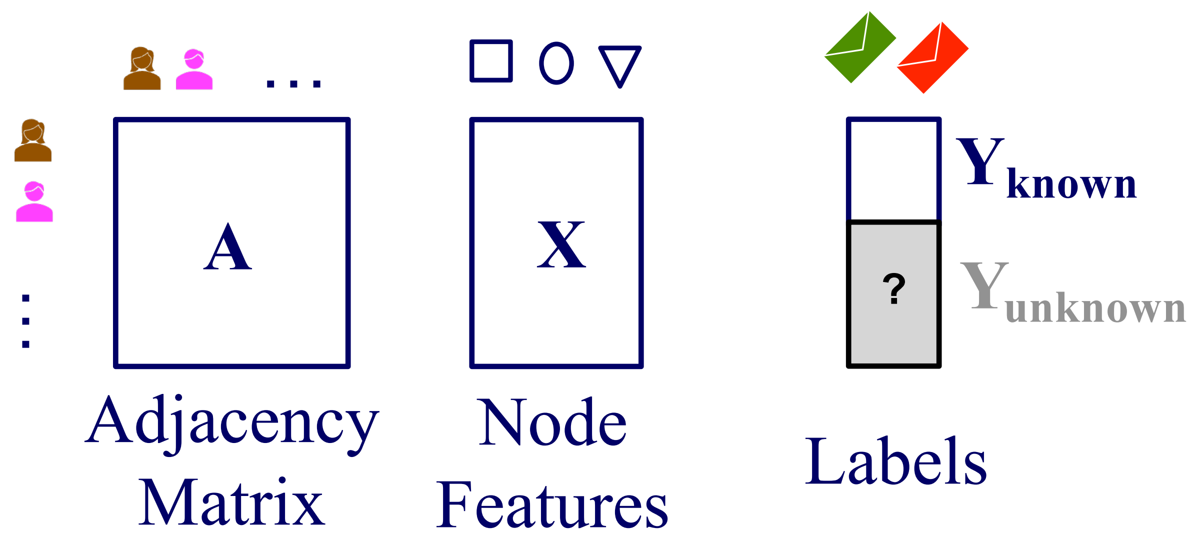
→ Reminder

Task Definition



Given:
- Graph **A**
- Node features **X**
- A few node labels

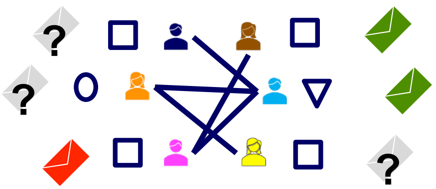
Predict: Rest labels



Task Definition

- Given
  - Graph (single edge-type)
  - Node features (black or green hair, etc.)
  - A few class labels (‘vote’, e.g., R/G)

- Predict
  - Rest labels



OUR Problem Definition

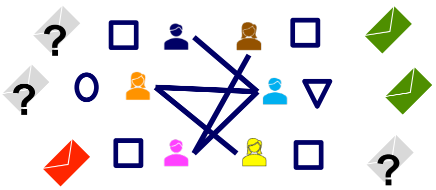
- Find good 'derived features' ≡ embeddings for each node
  - (Using ***A***, ***X***, ***AX***, etc.)
- That best predict the unknown labels $\boldsymbol{Y_{unknown}}$

Graph Neural Networks (GNNs)

- GNNs have shown their ability to handle different tasks in graph domain:
  - GCN
  - GAT
  - …
- Some GNNs remove the non-linear functions in their models:
  - SGC
  - DGC
  - …

Research Questions

- Q1) Reasons
- Q2) GNN that is
  - Accurate, Robust, Fast, Scalable, Interpretable
- Q3) Sanity Checks



Sneak Preview: Accurate & Robust

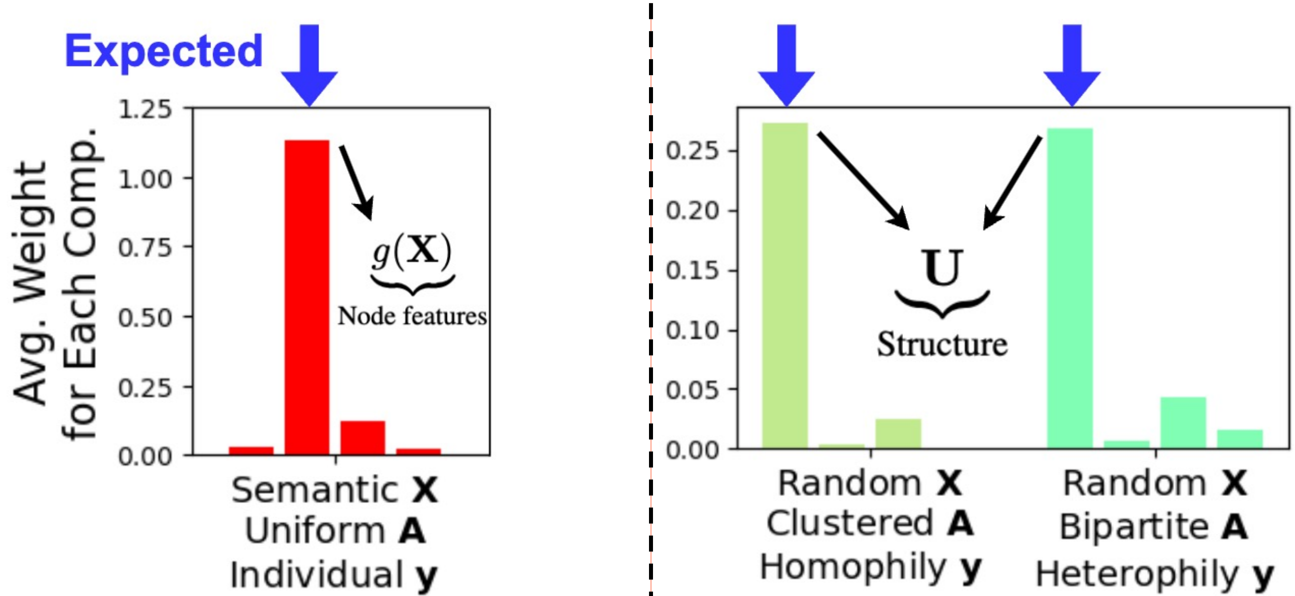
Real-World Datasets

7 Homophily | 6 Heterophily

Model	Cora	CiteSeer	PubMed	Comp.	Photo	ArXiv	Products	Cham.	Squirrel	Actor	Penn94	Twitch	Pokec	Avg. Rank
LR	51.5±1.2	52.9±4.5	79.9±0.5	73.9±1.2	79.3±1.5	48.3±1.9	56.4±0.5	24.9±1.7	26.7±1.9	27.8±0.8	63.5±0.5	53.0±0.1	61.3±0.0	11.7 (4.2)
Reg. Kernel	67.8±2.5	62.1±4.4	83.4±1.4	80.3±1.4	87.1±1.2	O.O.M.	O.O.M.	29.4±2.6	24.3±2.3	29.6±1.4	O.O.M.	O.O.M.	O.O.M.	12.2 (3.8)
Diff. Kernel	70.6±1.5	62.7±3.8	82.1±0.4	83.1±1.0	89.8±0.6	O.O.M.	O.O.M.	34.5±7.9	28.3±1.5	24.7±0.9	53.5±0.8	O.O.M.	O.O.M.	11.8 (2.5)
RW Kernel	72.7±1.7	64.1±3.9	83.1±0.7	84.2±0.7	90.6±0.7	63.2±0.2	74.2±0.0	34.9±3.5	25.0±1.6	26.4±1.1	63.1±0.7	57.6±0.1	59.5±0.0	8.3 (3.3)
SGC	76.2±1.1	65.8±3.9	84.1±0.8	83.7±1.6	90.1±0.9	65.0±3.4	74.6±5.1	38.1±4.5	33.1±1.0	24.6±0.8	64.0±1.1	56.5±0.1	69.8±0.0	6.6 (4.2)
DGC	77.8±1.4	66.1±4.2	84.3±0.6	83.9±0.7	90.4±0.2	65.2±4.0	68.7±13.	37.2±3.7	29.2±1.2	25.2±2.1	62.5±0.4	58.2±0.2	60.7±0.1	6.6 (3.2)
S^2GC	78.3±1.5	66.9±4.4	84.3±0.3	83.1±0.8	90.1±0.8	62.0±7.4	58.3±18.	34.9±4.9	27.6±1.8	26.7±1.8	63.1±0.5	58.7±0.1	61.2±0.0	6.6 (2.7)
G^2CN	76.6±1.5	64.2±3.3	81.4±0.6	82.8±1.6	88.8±0.5	O.O.M.	O.O.M.	40.7±2.9	32.1±1.5	24.3±0.5	O.O.M.	O.O.M.	O.O.M.	10.5 (4.5)
GCN	76.0±1.2	65.0±2.9	84.3±0.5	85.1±0.9	91.6±0.5	62.8±0.6	O.O.M.	38.5±3.0	31.4±1.8	26.8±0.4	62.9±0.7	57.0±0.1	63.9±0.4	6.3 (2.4)
SAGE	74.6±1.3	63.7±3.6	82.9±0.4	83.8±0.5	90.6±0.5	61.5±0.6	O.O.M.	39.8±4.3	27.0±1.3	27.8±0.9	O.O.M.	56.6±0.4	68.9±0.1	8.5 (3.5)
GCNII	77.8±1.7	63.4±3.0	84.9±0.8	82.3±1.8	90.8±0.6	45.7±0.5	O.O.M.	30.5±2.5	21.9±3.0	29.0±1.3	64.5±0.5	56.9±0.6	62.1±0.3	8.4 (4.6)
H^2GCN	77.6±0.9	64.7±3.8	85.4±0.4	49.5±16.	75.8±11.	O.O.M.	O.O.M.	31.9±2.6	25.0±0.5	28.9±0.6	63.9±0.4	58.7±0.0	O.O.M.	8.9 (4.9)
APPNP	80.0±0.6	67.1±2.8	84.6±0.5	84.2±1.7	92.5±0.3	53.4±1.3	O.O.M.	30.9±4.7	23.9±3.2	26.1±1.0	63.7±0.9	47.3±0.3	57.4±0.4	7.6 (4.8)
GPR-GNN	78.8±1.3	64.2±4.0	85.1±0.7	85.0±1.0	92.6±0.3	58.5±0.8	O.O.M.	31.7±4.7	26.2±1.6	29.5±1.1	64.5±0.4	57.6±0.2	67.6±0.1	5.4 (3.7)
GAT	78.2±1.2	65.8±4.0	83.6±0.2	85.4±1.4	91.7±0.5	58.2±1.0	O.O.M.	39.1±4.1	28.6±0.6	26.4±0.4	60.5±0.8	O.O.M.	O.O.M.	7.5 (3.7)
SlimG	77.8±1.1	67.1±2.3	84.6±0.5	86.3±0.7	91.8±0.5	66.3±0.3	84.9±0.0	40.8±3.2	31.1±0.7	30.9±0.6	68.2±0.6	59.7±0.1	73.9±0.1	1.9 (1.5)

2σ

Sneak Preview: Interpretable

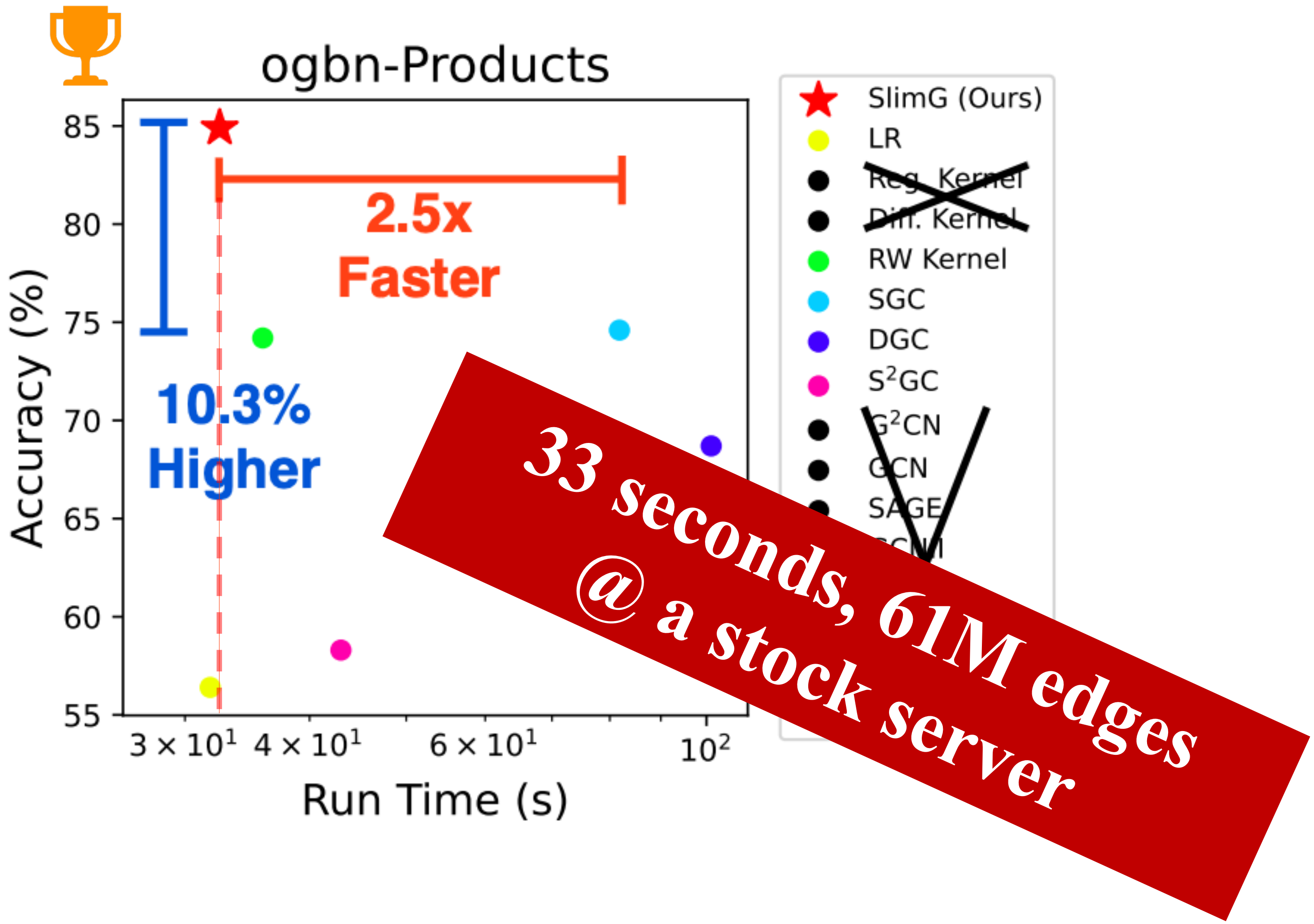


(a) No network effects **(b) Useless features**

$$\underbrace{\mathbf{U}}_{\text{Structure}} \,\|\, \boxed{\underbrace{g(\mathbf{X})}_{\text{Node features}}} \,\|\, \underbrace{g(\mathbf{A}^2_{\text{row}}\mathbf{X})}_{\text{2-step neighbors}} \,\|\, \underbrace{g(\tilde{\mathbf{A}}^2_{\text{sym}}\mathbf{X})}_{\text{Neighbors}}$$

$$\boxed{\underbrace{\mathbf{U}}_{\text{Structure}}} \,\|\, \underbrace{g(\mathbf{X})}_{\text{Node features}} \,\|\, \underbrace{g(\mathbf{A}^2_{\text{row}}\mathbf{X})}_{\text{2-step neighbors}} \,\|\, \underbrace{g(\tilde{\mathbf{A}}^2_{\text{sym}}\mathbf{X})}_{\text{Neighbors}}$$

Sneak Preview: Fast & Scalable

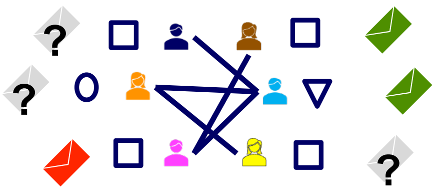


Research Questions

- Q1) Reasons
- Q2) Gl

"Careful Simplicity"

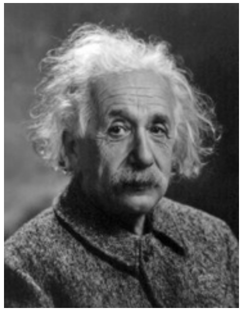
  - Accurate, Robust, Fast, Scalable, Interpretable
- Q3) Sanity Checks



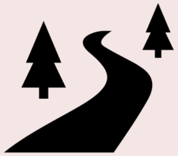
Over-Arching Principle

"Careful Simplicity"

"Everything should be made as simple as possible, but not simpler."



Outline



- Problem Definition
- Q1: Reasons
- Q2: Method
- Q3: Sanity Checks
- Experiments
- Conclusions

Q1: Reasons – Why?

- Why/when do GNNs succeed,
- … and why/when do they fail?

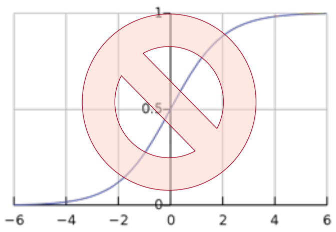
- Q: How should we proceed?

Model
LR
SGC
DGC
S^2GC
G^2CN
PPNP*
APPNP*
GDC*
GPR-GNN*
ChebNet*
GCN*
SAGE*
GCNII*
H_2GCN*
GAT**
DA-GNN**

Q1: Reasons – Why?

- Why/when do GNNs succeed,
- … and why/when do they fail?

- Q: How should we proceed?
- A: Drop non-linearities, to get the essence



GNNEXP

Jaemin Yoo

Model	Type	Propagator function $\mathcal{P}(\mathbf{A}, \mathbf{X})$
LR	Linear	$\mathbf{X}$
SGC	Linear	$\tilde{\mathbf{A}}_{\mathrm{sym}}^{K}\,\mathbf{X}$
DGC	Linear	$[(1-T/K)\mathbf{I} + (T/K)\tilde{\mathbf{A}}_{\mathrm{sym}}]^{K}\,\mathbf{X}$
S^2GC	Linear	$\sum_{k=1}^{K}(\alpha\mathbf{I} + (1-\alpha)\tilde{\mathbf{A}}_{\mathrm{sym}}^{k})\,\mathbf{X}$
G^2CN	Linear	$\Vert_{i=1}^{N}[\mathbf{I} - (T_i/K)((b_i-1)\mathbf{I} + \mathbf{A}_{\mathrm{sym}})^2]^{K}\,\mathbf{X}$
PPNP*	Decoupled	$(\mathbf{I} - (1-\alpha)\tilde{\mathbf{A}}_{\mathrm{sym}})^{-1}\,\mathbf{X}$
APPNP*	Decoupled	$[\sum_{k=0}^{K-1}\alpha(1-\alpha)^{k}\tilde{\mathbf{A}}_{\mathrm{sym}}^{k} + (1-\alpha)^{K}\tilde{\mathbf{A}}_{\mathrm{sym}}^{K}]\,\mathbf{X}$
GDC*	Decoupled	$\mathbf{S} = \mathrm{sparse}_{\epsilon}(\sum_{k=0}^{\infty}(1-\alpha)^{k}\tilde{\mathbf{A}}_{\mathrm{sym}}^{k})$ for $\tilde{\mathbf{S}}_{\mathrm{sym}}\,\mathbf{X}$
GPR-GNN*	Decoupled	$\Vert_{k=0}^{K}\,\tilde{\mathbf{A}}_{\mathrm{sym}}^{k}\,\mathbf{X}$
ChebNet*	Coupled	$\Vert_{k=0}^{K-1}\,\mathbf{A}_{\mathrm{sym}}^{k}\,\mathbf{X}$
GCN*	Coupled	$\tilde{\mathbf{A}}_{\mathrm{sym}}^{K}\,\mathbf{X}$
SAGE*	Coupled	$\Vert_{k=0}^{K}\,\mathbf{A}_{\mathrm{row}}^{k}\,\mathbf{X}$
GCNII*	Coupled	$\Vert_{k=0}^{K-2}\,\tilde{\mathbf{A}}_{\mathrm{sym}}^{k}\mathbf{X} \,\Vert\, ((1-\alpha)\tilde{\mathbf{A}}_{\mathrm{sym}}^{K} + \alpha\tilde{\mathbf{A}}_{\mathrm{sym}}^{K-1})\,\mathbf{X}$
H_2GCN*	Coupled	$\Vert_{k=0}^{2K}\,\mathbf{A}_{\mathrm{sym}}^{k}\mathbf{X}$
GAT**	Attention	$\prod_{k=1}^{K}[\mathrm{diag}(\mathbf{X}\mathbf{w}_{k,1})\tilde{\mathbf{A}} + \tilde{\mathbf{A}}\mathrm{diag}(\mathbf{X}\mathbf{w}_{k,2})]\,\mathbf{X}$
DA-GNN**	Attention	$\sum_{k=0}^{K}\mathrm{diag}(\tilde{\mathbf{A}}_{\mathrm{sym}}^{k}\mathbf{X}\mathbf{w})\tilde{\mathbf{A}}_{\mathrm{sym}}^{k}\,\mathbf{X}$

‘Derived’ Features ≡ Embedding

GNNEXP

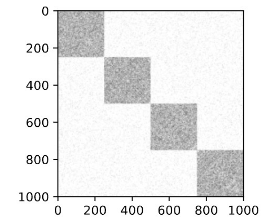
Jaemin Yoo

Model	Type	Propagator function $\mathcal{P}(\mathbf{A}, \mathbf{X})$
LR	Linear	$\mathbf{X}$
SGC	Linear	$\tilde{\mathbf{A}}_{\text{sym}}^{K}\,\mathbf{X}$
DGC	Linear	$[(1-T/K)\mathbf{I}+(T/K)\tilde{\mathbf{A}}_{\text{sym}}]^{K}\,\mathbf{X}$
S^2GC	Linear	$\sum_{k=1}^{K}(\alpha\mathbf{I}+(1-\alpha)\tilde{\mathbf{A}}_{\text{sym}}^{k})\,\mathbf{X}$
G^2CN	Linear	$\Vert_{i=1}^{N}[\mathbf{I}-(T_i/K)((b_i-1)\mathbf{I}+\mathbf{A}_{\text{sym}})^2]^{K}\,\mathbf{X}$
PPNP*	Decoupled	$(\mathbf{I}-(1-\alpha)\tilde{\mathbf{A}}_{\text{sym}})^{-1}\,\mathbf{X}$
APPNP*	Decoupled	$[\sum_{k=0}^{K-1}\alpha(1-\alpha)^{k}\tilde{\mathbf{A}}_{\text{sym}}^{k}+(1-\alpha)^{K}\tilde{\mathbf{A}}_{\text{sym}}^{K}]\,\mathbf{X}$
GDC*	Decoupled	$\mathbf{S}=\text{sparse}_{\epsilon}(\sum_{k=0}^{\infty}(1-\alpha)^{k}\tilde{\mathbf{A}}_{\text{sym}}^{k})$ for $\tilde{\mathbf{S}}_{\text{sym}}\,\mathbf{X}$
GPR-GNN*	Decoupled	$\Vert_{k=0}^{K}\,\tilde{\mathbf{A}}_{\text{sym}}^{k}\,\mathbf{X}$
ChebNet*	Coupled	$\Vert_{k=0}^{K-1}\,\mathbf{A}_{\text{sym}}^{k}\,\mathbf{X}$
GCN*	Coupled	$\tilde{\mathbf{A}}_{\text{sym}}^{K}\,\mathbf{X}$
SAGE*	Coupled	$\Vert_{k=0}^{K}\,\mathbf{A}_{\text{row}}^{k}\,\mathbf{X}$
GCNII*	Coupled	$\Vert_{k=0}^{K-2}\,\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}\,\Vert\,((1-\alpha)\tilde{\mathbf{A}}_{\text{sym}}^{K}+\alpha\tilde{\mathbf{A}}_{\text{sym}}^{K-1})\,\mathbf{X}$
H_2GCN*	Coupled	$\Vert_{k=0}^{2K}\,\mathbf{A}_{\text{sym}}^{k}\mathbf{X}$
GAT**	Attention	$\prod_{k=1}^{K}[\text{diag}(\mathbf{X}\mathbf{w}_{k,1})\tilde{\mathbf{A}}+\tilde{\mathbf{A}}\text{diag}(\mathbf{X}\mathbf{w}_{k,2})]\,\mathbf{X}$
DA-GNN**	Attention	$\sum_{k=0}^{K}\text{diag}(\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}\mathbf{w})\tilde{\mathbf{A}}_{\text{sym}}^{k}\,\mathbf{X}$

Q1: Reasons – Deep Dive

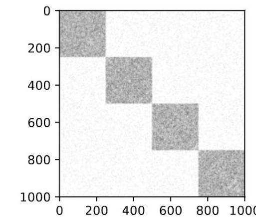
- Successes
  - S1: Homophily
  - S2: [Heterophily]
- Pain Points
  - PP1: Lack of Robustness
  - PP2: Vulnerability to Noisy Features
  - PP3: Efficiency and Effectiveness
  - PP4: Many Hyperparameters

Q1: Reasons – Deep Dive



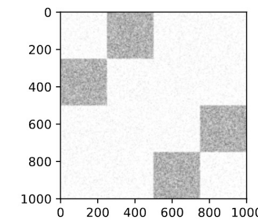
- Successes
  - S1: Homophily $AX; A^2X; \ldots \rightarrow$ neighbors' features
  - S2: [Heterophily]
- Pain Points
  - PP1: Lack of Robustness
  - PP2: Vulnerability to Noisy Features
  - PP3: Efficiency and Effectiveness
  - PP4: Many Hyperparameters

Q1: Reasons – Deep Dive



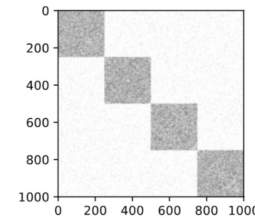
- Successes
  - S1: Homophily $AX; A^2X; \ldots \rightarrow$ neighbors' features
  - S2: [Heterophily] AX: HURTS

    A^2X: if concat., good

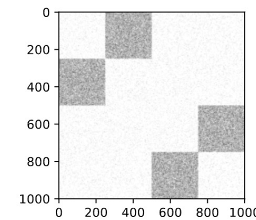


- Pain Points
  - PP1: Lack of Robustness
  - PP2: Vulnerability to Noisy Features
  - PP3: Efficiency and Effectiveness
  - PP4: Many Hyperparameters

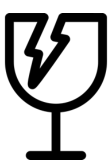
Q1: Reasons – Deep Dive



- Successes
  - ✓ S1: Homophily $AX; A^2X; \ldots \rightarrow$ neighbors' features
  - ✗ ✓ S2: [Heterophily] AX: HURTS
    A^2X: if concat., good



- Pain Points
  - ✗ PP1: Lack of Robustness $(\ldots)X$
  - ✗ PP2: Vulnerability to Noisy Features $(\ldots + I)X$
  - PP3: Efficiency and Effectiveness
  - PP4: Many Hyperparameters



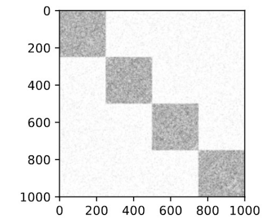
GnnExp

Model	Type	Propagator function $\mathcal{P}(\mathbf{A}, \mathbf{X})$
LR	Linear	$\mathbf{X}$
SGC	Linear	$\tilde{\mathbf{A}}_{\text{sym}}^{K}\mathbf{X}$
DGC	Linear	$[(1-T/K)\mathbf{I}+(T/K)\tilde{\mathbf{A}}_{\text{sym}}]^{K}\mathbf{X}$
$\mathrm{S}^2\mathrm{GC}$	Linear	$\sum_{k=1}^{K}(\alpha\mathbf{I}+(1-\alpha)\tilde{\mathbf{A}}_{\text{sym}}^{k})\mathbf{X}$
$\mathrm{G}^2\mathrm{CN}$	Linear	$\Vert_{i=1}^{N}[\mathbf{I}-(T_i/K)((b_i-1)\mathbf{I}+\mathbf{A}_{\text{sym}})^2]^{K}\mathbf{X}$
PPNP*	Decoupled	$(\mathbf{I}-(1-\alpha)\tilde{\mathbf{A}}_{\text{sym}})^{-1}\mathbf{X}$
APPNP*	Decoupled	$[\sum_{k=0}^{K-1}\alpha(1-\alpha)^{k}\tilde{\mathbf{A}}_{\text{sym}}^{k}+(1-\alpha)^{K}\tilde{\mathbf{A}}_{\text{sym}}^{K}]\mathbf{X}$
GDC*	Decoupled	$\mathbf{S}=\text{sparse}_{\epsilon}(\sum_{k=0}^{\infty}(1-\alpha)^{k}\tilde{\mathbf{A}}_{\text{sym}}^{k})$ for $\tilde{\mathbf{S}}_{\text{sym}}\mathbf{X}$
GPR-GNN*	Decoupled	$\Vert_{k=0}^{K}\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}$
ChebNet*	Coupled	$\Vert_{k=0}^{K-1}\mathbf{A}_{\text{sym}}^{k}\mathbf{X}$
GCN*	Coupled	$\tilde{\mathbf{A}}_{\text{sym}}^{K}\mathbf{X}$
SAGE*	Coupled	$\Vert_{k=0}^{K}\mathbf{A}_{\text{row}}^{k}\mathbf{X}$
GCNII*	Coupled	$\Vert_{k=0}^{K-2}\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}\parallel((1-\alpha)\tilde{\mathbf{A}}_{\text{sym}}^{K}+\alpha\tilde{\mathbf{A}}_{\text{sym}}^{K-1})\mathbf{X}$
$\mathrm{H}_2\mathrm{GCN}$*	Coupled	$\Vert_{k=0}^{2K}\mathbf{A}_{\text{sym}}^{k}\mathbf{X}$
GAT**	Attention	$\prod_{k=1}^{K}[\text{diag}(\mathbf{X}\mathbf{w}_{k,1})\tilde{\mathbf{A}}+\tilde{\mathbf{A}}\text{diag}(\mathbf{X}\mathbf{w}_{k,2})]\mathbf{X}$
DA-GNN**	Attention	$\sum_{k=0}^{K}\text{diag}(\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}\mathbf{w})\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}$

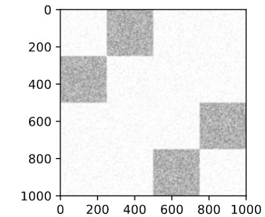
GNNEXP

Model	Type	Propagator function $\mathcal{P}(\mathbf{A}, \mathbf{X})$
LR	Linear	$\mathbf{X}$
SGC	Linear	$\tilde{\mathbf{A}}_{\text{sym}}^{K}\,\mathbf{X}$
DGC	Linear	$[(1-T/K)\mathbf{I} + (T/K)\tilde{\mathbf{A}}_{\text{sym}}]^{K}\,\mathbf{X}$
S^2GC	Linear	$\sum_{k=1}^{K}(\alpha\mathbf{I} + (1-\alpha)\tilde{\mathbf{A}}_{\text{sym}}^{k})\,\mathbf{X}$
G^2CN	Linear	$\Vert_{i=1}^{N}[\mathbf{I} - (T_i/K)((b_i-1)\mathbf{I} + \mathbf{A}_{\text{sym}})^2]^{K}\,\mathbf{X}$
PPNP*	Decoupled	$(\mathbf{I} - (1-\alpha)\tilde{\mathbf{A}}_{\text{sym}})^{-1}\,\mathbf{X}$
APPNP*	Decoupled	$[\sum_{k=0}^{K-1}\alpha(1-\alpha)^{k}\tilde{\mathbf{A}}_{\text{sym}}^{k} + (1-\alpha)^{K}\tilde{\mathbf{A}}_{\text{sym}}^{K}]\,\mathbf{X}$
GDC*	Decoupled	$\mathbf{S} = \text{sparse}_{\epsilon}(\sum_{k=0}^{\infty}(1-\alpha)^{k}\tilde{\mathbf{A}}_{\text{sym}}^{k})$ for $\tilde{\mathbf{S}}_{\text{sym}}\,\mathbf{X}$
GPR-GNN*	Decoupled	$\Vert_{k=0}^{K}\tilde{\mathbf{A}}_{\text{sym}}^{k}\,\mathbf{X}$
ChebNet*	Coupled	$\Vert_{k=0}^{K-1}\mathbf{A}_{\text{sym}}^{k}\,\mathbf{X}$
GCN*	Coupled	$\tilde{\mathbf{A}}_{\text{sym}}^{K}\,\mathbf{X}$
SAGE*	Coupled	$\Vert_{k=0}^{K}\mathbf{A}_{\text{row}}^{k}\,\mathbf{X}$
GCNII*	Coupled	$\Vert_{k=0}^{K-2}\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X} \Vert ((1-\alpha)\tilde{\mathbf{A}}_{\text{sym}}^{K} + \alpha\tilde{\mathbf{A}}_{\text{sym}}^{K-1})\,\mathbf{X}$
H_2GCN*	Coupled	$\Vert_{k=0}^{2K}\mathbf{A}_{\text{sym}}^{k}\mathbf{X}$
GAT**	Attention	$\prod_{k=1}^{K}[\text{diag}(\mathbf{X}\mathbf{w}_{k,1})\tilde{\mathbf{A}} + \tilde{\mathbf{A}}\text{diag}(\mathbf{X}\mathbf{w}_{k,2})]\,\mathbf{X}$
DA-GNN**	Attention	$\sum_{k=0}^{K}\text{diag}(\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}\mathbf{w})\tilde{\mathbf{A}}_{\text{sym}}^{k}\,\mathbf{X}$

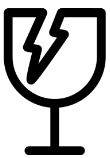
Q1: Reasons – Deep Dive



- Successes
  - ✓ S1: Homophily $AX; A^2X; \ldots \rightarrow$ neighbors' features
  - ✗ ✓ S2: [Heterophily] AX: HURTS
    A^2X: if concat., good



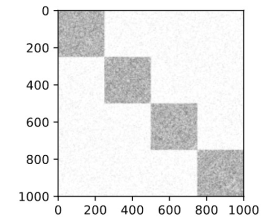
- Pain Points
  - ✗ PP1: Lack of Robustness $(\ldots)X$
  - ✗ PP2: Vulnerability to Noisy Features $(\ldots + I)X$
  - ✗ PP3: Efficiency and Effectiveness 'too many cooks'
  - PP4: Many Hyperparameters



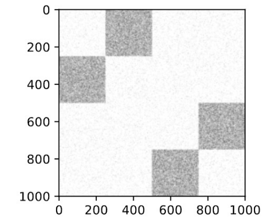
GNNEXP

Model	Type	Propagator function $\mathcal{P}(\mathbf{A}, \mathbf{X})$
LR	Linear	$\mathbf{X}$
SGC	Linear	$\tilde{\mathbf{A}}_{\text{sym}}^{K}\,\mathbf{X}$
DGC	Linear	$[(1-T/K)\mathbf{I} + (T/K)\tilde{\mathbf{A}}_{\text{sym}}]^{K}\,\mathbf{X}$
S^2GC	Linear	$\sum_{k=1}^{K}(\alpha\mathbf{I} + (1-\alpha)\tilde{\mathbf{A}}_{\text{sym}}^{k})\,\mathbf{X}$
G^2CN	Linear	$\boxed{\Vert_{i=1}^{N}}[\mathbf{I} - (T_i/K)((b_i-1)\mathbf{I} + \mathbf{A}_{\text{sym}})^2]^K\,\mathbf{X}$
PPNP*	Decoupled	$(\mathbf{I} - (1-\alpha)\tilde{\mathbf{A}}_{\text{sym}})^{-1}\,\mathbf{X}$
APPNP*	Decoupled	$[\sum_{k=0}^{K-1}\alpha(1-\alpha)^k\tilde{\mathbf{A}}_{\text{sym}}^{k} + (1-\alpha)^K\tilde{\mathbf{A}}_{\text{sym}}^{K}]\,\mathbf{X}$
GDC*	Decoupled	$\mathbf{S} = \text{sparse}_{\epsilon}(\sum_{k=0}^{\infty}(1-\alpha)^k\tilde{\mathbf{A}}_{\text{sym}}^{k})$ for $\tilde{\mathbf{S}}_{\text{sym}}\,\mathbf{X}$
GPR-GNN*	Decoupled	$\boxed{\Vert_{k=0}^{K}}\tilde{\mathbf{A}}_{\text{sym}}^{k}\,\mathbf{X}$
ChebNet*	Coupled	$\boxed{\Vert_{k=0}^{K-1}}\mathbf{A}_{\text{sym}}^{k}\,\mathbf{X}$
GCN*	Coupled	$\tilde{\mathbf{A}}_{\text{sym}}^{K}\,\mathbf{X}$
SAGE*	Coupled	$\boxed{\Vert_{k=0}^{K}}\mathbf{A}_{\text{row}}^{k}\,\mathbf{X}$
GCNII*	Coupled	$\boxed{\Vert_{k=0}^{K-2}}\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}\boxed{\Vert}((1-\alpha)\tilde{\mathbf{A}}_{\text{sym}}^{K} + \alpha\mathbf{A}_{\text{sym}}^{K-1})\,\mathbf{X}$
H_2GCN*	Coupled	$\boxed{\Vert_{k=0}^{2K}}\mathbf{A}_{\text{sym}}^{k}\mathbf{X}$
GAT**	Attention	$\prod_{k=1}^{K}[\text{diag}(\mathbf{X}\mathbf{w}_{k,1})\tilde{\mathbf{A}} + \tilde{\mathbf{A}}\text{diag}(\mathbf{X}\mathbf{w}_{k,2})]\,\mathbf{X}$
DA-GNN**	Attention	$\sum_{k=0}^{K}\text{diag}(\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}\mathbf{w})\tilde{\mathbf{A}}_{\text{sym}}^{k}\,\mathbf{X}$

Q1: Reasons – Deep Dive



- Successes
  - ✓ S1: Homophily $AX; A^2X; \ldots \rightarrow$ neighbors' features
  - ✗ ✓ S2: [Heterophily] AX: HURTS
    A^2X: if concat., good



- Pain Points
  - ✗ PP1: Lack of Robustness $(\ldots)X$
  - ✗ PP2: Vulnerability to Noisy Features $(\ldots + I)X$
  - ✗ PP3: Efficiency and Effectiveness 'too many cooks'
  - ✗ PP4: Many Hyperparameters

'6 degrees'; $A^{100} \rightarrow$ rank-1

GNNExp

Model	Type	Propagator function $\mathcal{P}(\mathbf{A}, \mathbf{X})$
LR	Linear	$\mathbf{X}$
SGC	Linear	$\tilde{\mathbf{A}}_{\text{sym}}^{K}\mathbf{X}$
DGC	Linear	$[(1-T/K)\mathbf{I}+(T/K)\tilde{\mathbf{A}}_{\text{sym}}]^{K}\mathbf{X}$
S^2GC	Linear	$\sum_{k=1}^{K}(\alpha\mathbf{I}+(1-\alpha)\tilde{\mathbf{A}}_{\text{sym}}^{k})\mathbf{X}$
G^2CN	Linear	$\Vert_{i=1}^{N}[\mathbf{I}-(T_i/K)((b_i-1)\mathbf{I}+\mathbf{A}_{\text{sym}})^2]^{K}\mathbf{X}$
PPNP*	Decoupled	$(\mathbf{I}-(1-\alpha)\tilde{\mathbf{A}}_{\text{sym}})^{-1}\mathbf{X}$
APPNP*	Decoupled	$[\sum_{k=0}^{K-1}\alpha(1-\alpha)^{k}\tilde{\mathbf{A}}_{\text{sym}}^{k}+(1-\alpha)^{K}\tilde{\mathbf{A}}_{\text{sym}}^{K}]\mathbf{X}$
GDC*	Decoupled	$\mathbf{S}=\text{sparse}_{\epsilon}(\sum_{k=0}^{\infty}(1-\alpha)^{k}\tilde{\mathbf{A}}_{\text{sym}}^{k})$ for $\tilde{\mathbf{S}}_{\text{sym}}\mathbf{X}$
GPR-GNN*	Decoupled	$\Vert_{k=0}^{K}\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}$
ChebNet*	Coupled	$\Vert_{k=0}^{K-1}\mathbf{A}_{\text{sym}}^{k}\mathbf{X}$
GCN*	Coupled	$\tilde{\mathbf{A}}_{\text{sym}}^{K}\mathbf{X}$
SAGE*	Coupled	$\Vert_{k=0}^{K}\mathbf{A}_{\text{row}}^{k}\mathbf{X}$
GCNII*	Coupled	$\Vert_{k=0}^{K-2}\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}\Vert((1-\alpha)\tilde{\mathbf{A}}_{\text{sym}}^{K}+\alpha\tilde{\mathbf{A}}_{\text{sym}}^{K-1})\mathbf{X}$
H_2GCN*	Coupled	$\Vert_{k=0}^{2K}\mathbf{A}_{\text{sym}}^{k}\mathbf{X}$
GAT**	Attention	$\prod_{k=1}^{K}[\text{diag}(\mathbf{X}\mathbf{w}_{k,1})\tilde{\mathbf{A}}+\tilde{\mathbf{A}}\text{diag}(\mathbf{X}\mathbf{w}_{k,2})]\mathbf{X}$
DA-GNN**	Attention	$\sum_{k=0}^{K}\text{diag}(\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}\mathbf{w})\tilde{\mathbf{A}}_{\text{sym}}^{k}\mathbf{X}$

Outline

- Problem Definition
- Q1: Reasons
- Q2: Method
- Q3: Sanity Checks
- Experiments
- Conclusions

Q2: Method – Desiderata

- Desiderata and Design Decisions
  - D0: Linear Classifier
  - D1: Concatenating Winning Components
  - D2: Structure Features
  - D3: Orthogonalization and Sparsification
  - D4: Multi-Level Neighborhood Aggregation

Q2: Method – Desiderata

- Desiderata and Design Decisions
  - D0: Linear Classifier
  - D1: Concatenating Winning Components
  - D2: Structure Features
  - D3: Orthogonalization and Sparsification
  - D4: Multi-Level Neighborhood Aggregation

Design Decision #0

Q: What classifier?

- Decision Trees?
- Random Forests?
- XGBoost?
- MLP?

Design Decision #0

A: Logistic regression (1-node MLP), after we find 'good embeddings'

It's simple and interpretable.

Logistic Regression (LR)

Design Decision #0

Q: # layers?

A: N/A

Q: Activation function?

A: Linear only (except the final LR step)

Q2: Method – Desiderata

- Desiderata and Design Decisions
  - D0: Linear Classifier
  - D1: Concatenating Winning Components
    - ‘||’ not ‘+’
    - Heterophily
  - D2: Structure Features
  - D3: Orthogonalization and Sparsification
  - D4: Multi-Level Neighborhood Aggregation

Design Decision #1 – Concat.

- Concatenate || ?

- Or add + ?

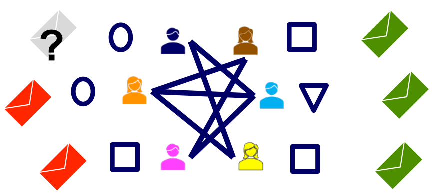
Design Decision #1 – Concat.

- Concatenate || ? -> Heterophily / Robustness

$$AX \parallel A^2X$$ ✓

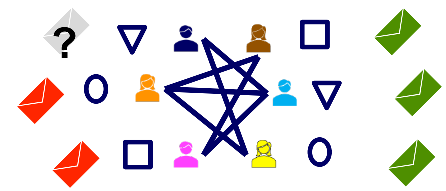
- Or add + ?

$$AX + A^2X$$



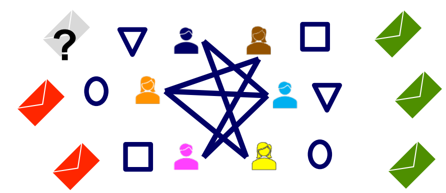
Design Decision D1 – Hom. / Het.

- How to handle both homophily, AND heterophily?

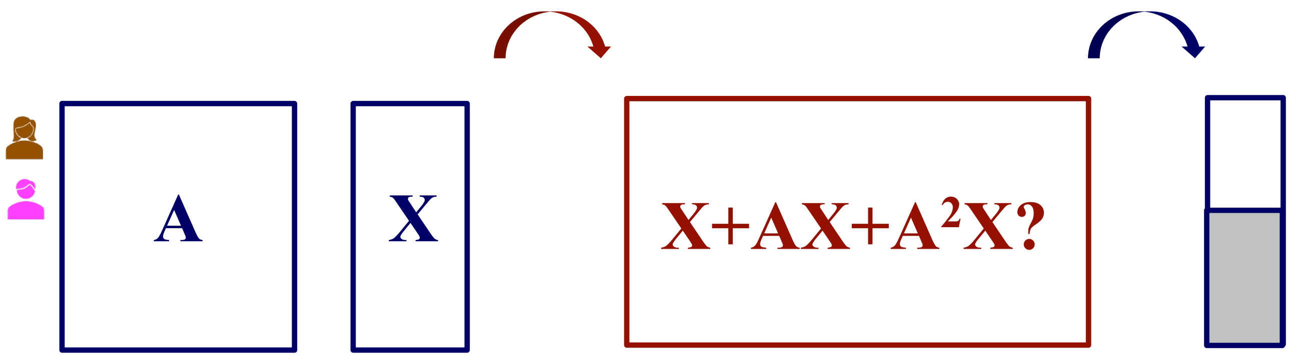


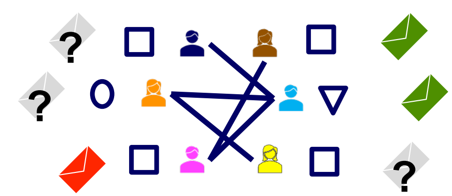
Design Decision D1 – Hom. / Het.

- How to handle both homophily, AND heterophily?



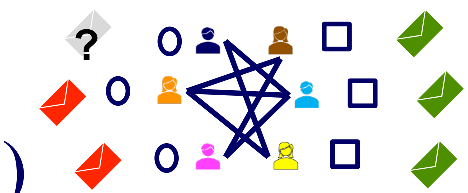
(If we add node features +
1-, 2-, … k-step neighbor features
→ mess and is hard to interpret)

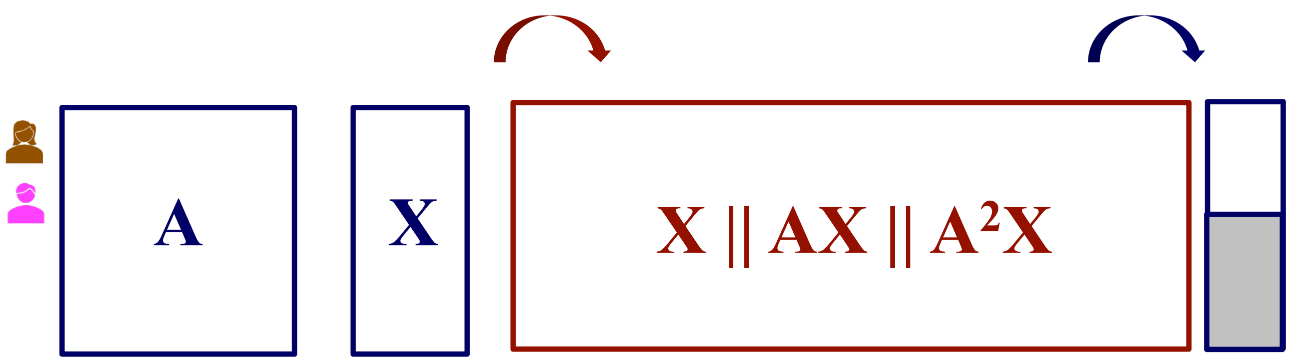


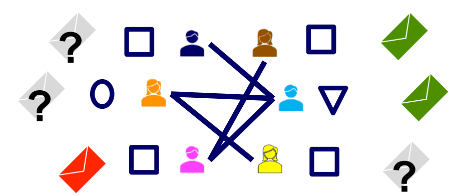


Design Decision D1 – Hom. / Het.

- How to handle both homophily, AND heterophily?
- Concatenate best homophily matrix,
- AND grand-neighbors (even powers)



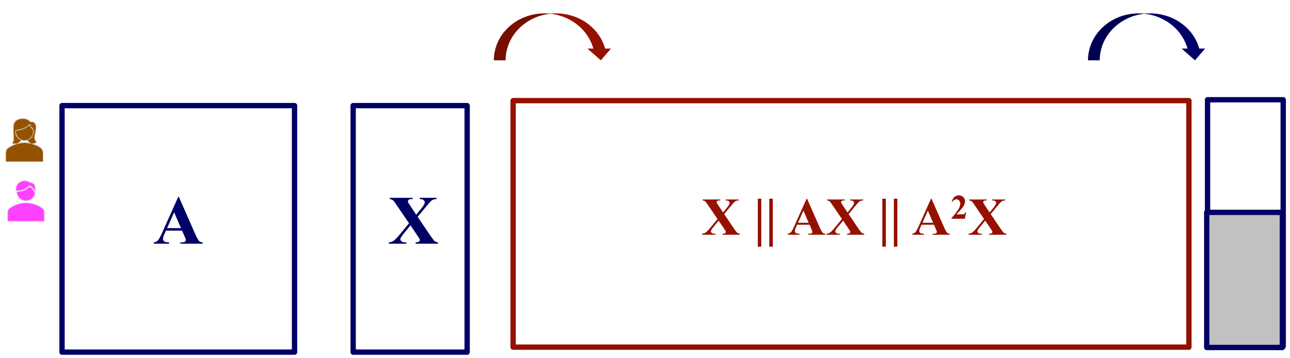


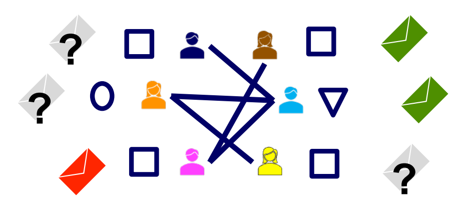


Design Decision D1 – Hom. / Het.

- Is **A** the best, for homophily?

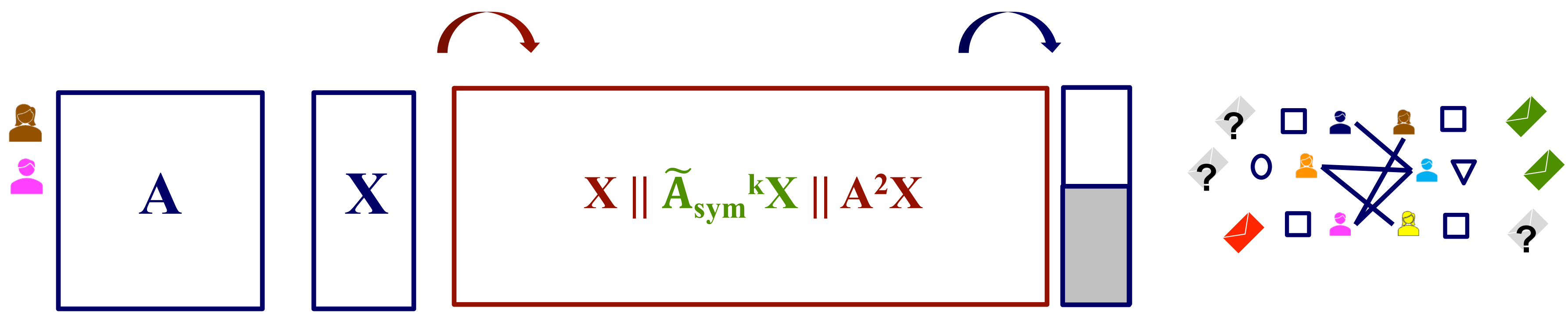
- Is **A²** …….., ….. heterophily?





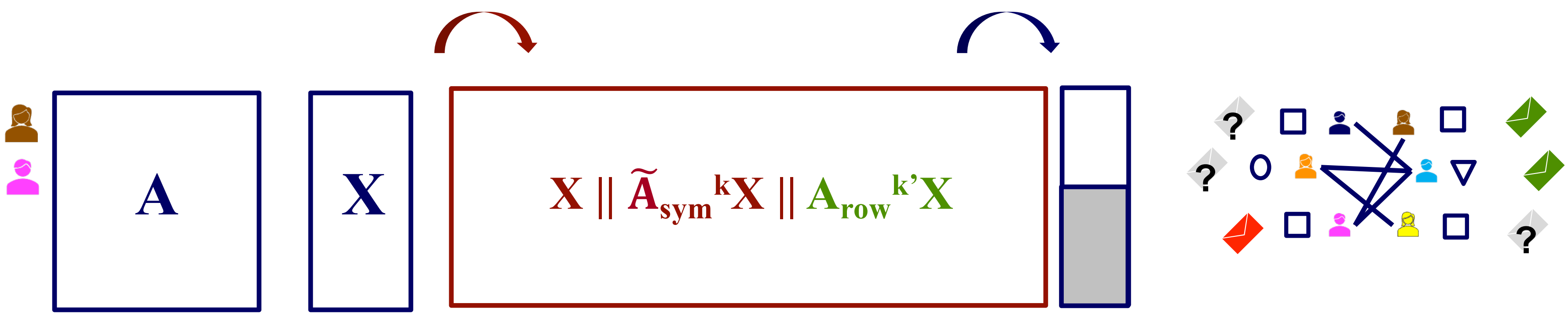
Design Decision D1 – Hom. / Het.

- Is **A** the best, for homophily?
- $\tilde{\mathbf{A}}_{sym}$ (symmetric norm. with self loops)
- Is $\mathbf{A}^2$ ...., ... heterophily?



Design Decision D1 – Hom. / Het.

- Is **A** the best, for homophily?
- $\tilde{\mathbf{A}}_{sym}$ (symmetric norm. with self loops)
- Is $\mathbf{A}^2$ ………, ….. heterophily?
- $\mathbf{A}_{row}$ (row-norm; NO self loops)

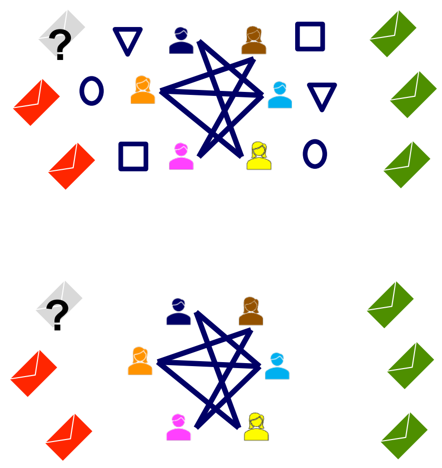


Q2: Method – Desiderata

- Desiderata and Design Decisions
  - D0: Linear Classifier
  - D1: Concatenating Winning Components
  - D2: Structure Features
  - D3: Orthogonalization and Sparsification
  - D4: Multi-Level Neighborhood Aggregation

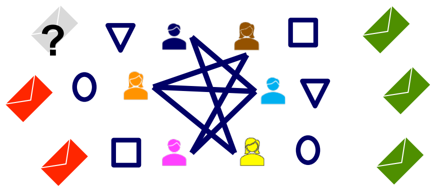
Design Decision D2 – Structure

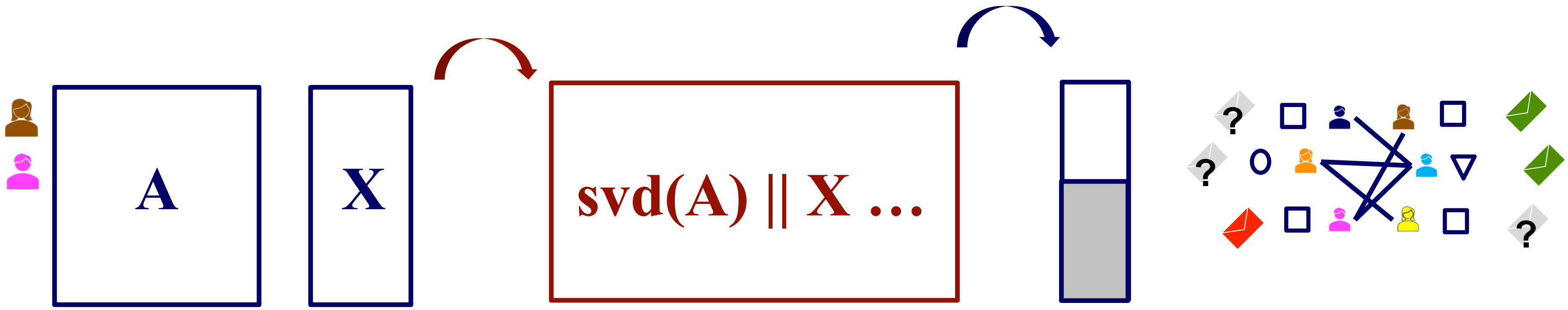
- Q: What if (raw) features are useless?
  - i.e., uncorrelated to labels
  - E.g., hair color vs. voting pattern
  - Or, feature-less?



Design Decision D2 – Structure

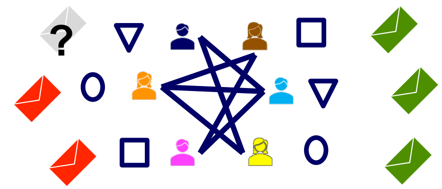
- Q: What if (raw) features are useless?
- A: Add structure info., ***A***





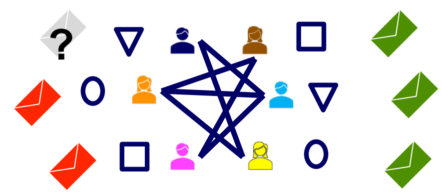
Design Decision D2 – Structure

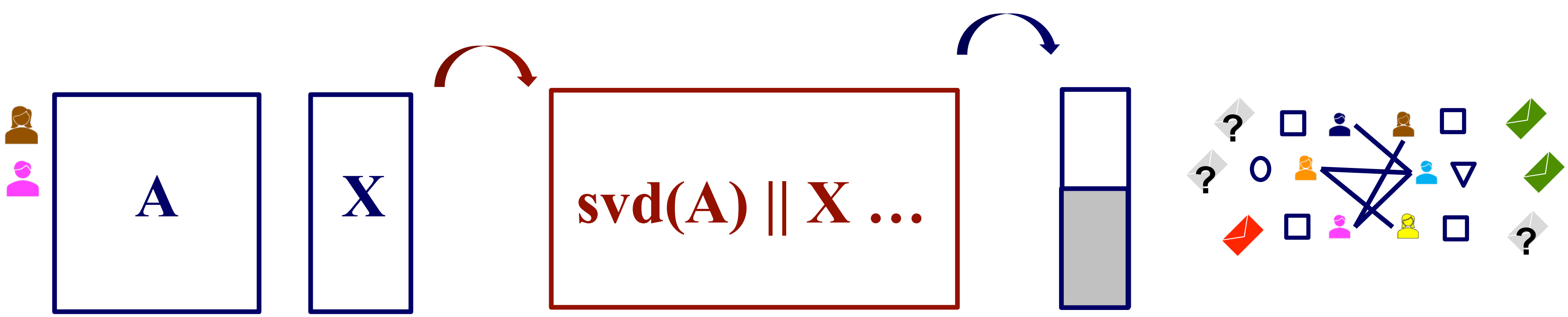
- Q: What if (raw) features are useless?
- A: Add structure info., ***A***
- (Q: but ***A*** is too wide, 1M×1M)



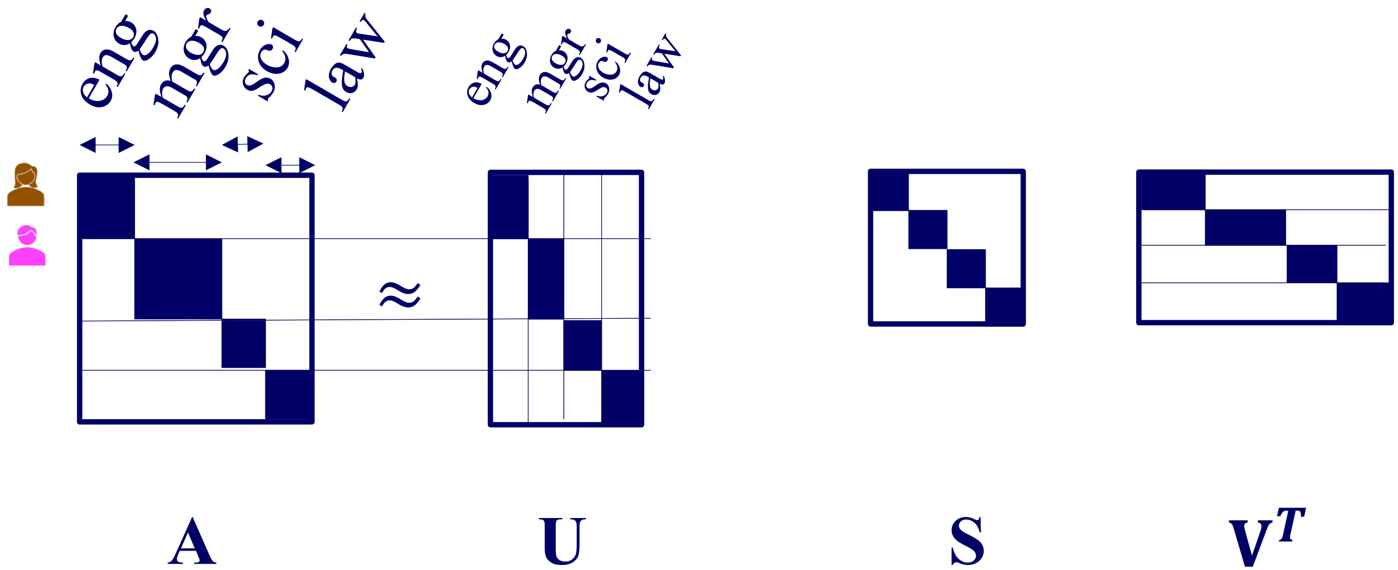
Design Decision D2 – Structure

- Q: What if (raw) features are useless?
- A: Add structure info., $\boldsymbol{A}$
- (Q: but $\boldsymbol{A}$ is too wide, 1M×1M)
- (A: $\boldsymbol{U}$, where $\boldsymbol{U\Sigma V}^T \approx \boldsymbol{A}$ by SVD)



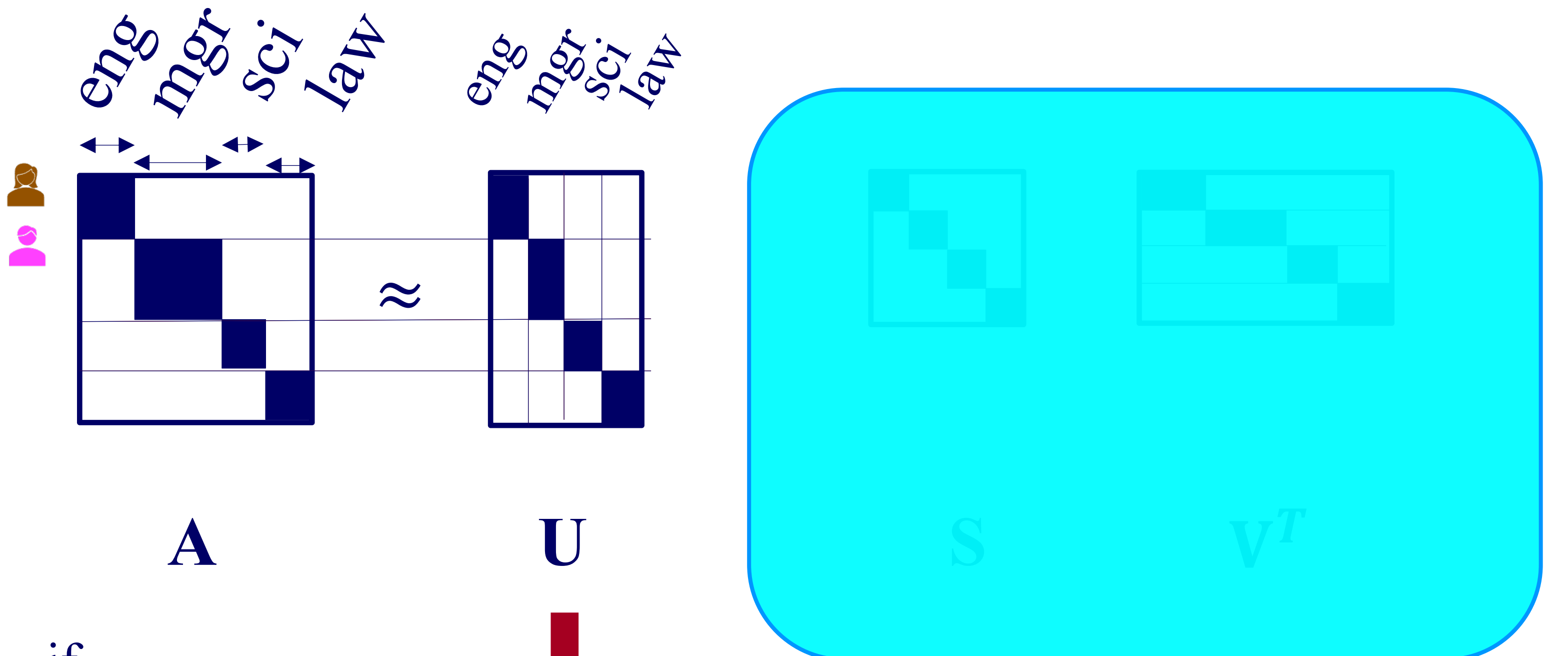


Reminder: SVD finds communities



Eg., if
block-diagonal =
communities ~
homophily

Reminder: SVD finds communities



Eg., if
block-diagonal =
communities ~
homophily

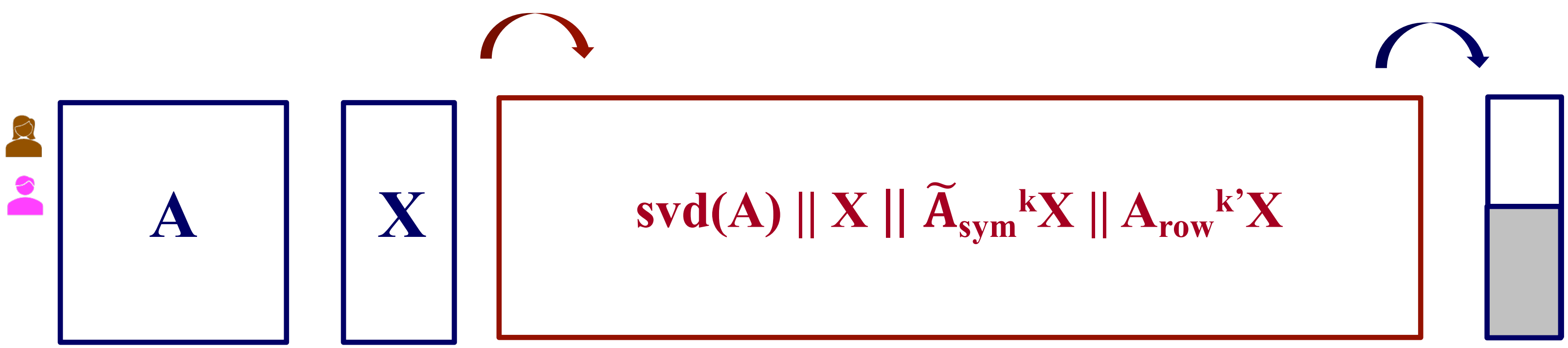
Person x community

Q2: Method – Desiderata

- Desiderata and Design Decisions
  - D0: Linear Classifier
  - D1: Concatenating Winning Components
  - D2: Structure Features
  - D3: Orthogonalization and Sparsification
    - Improve consistency
    - Remove redundant dimension(s)/component(s)
  - D4: Multi-Level Neighborhood Aggregation

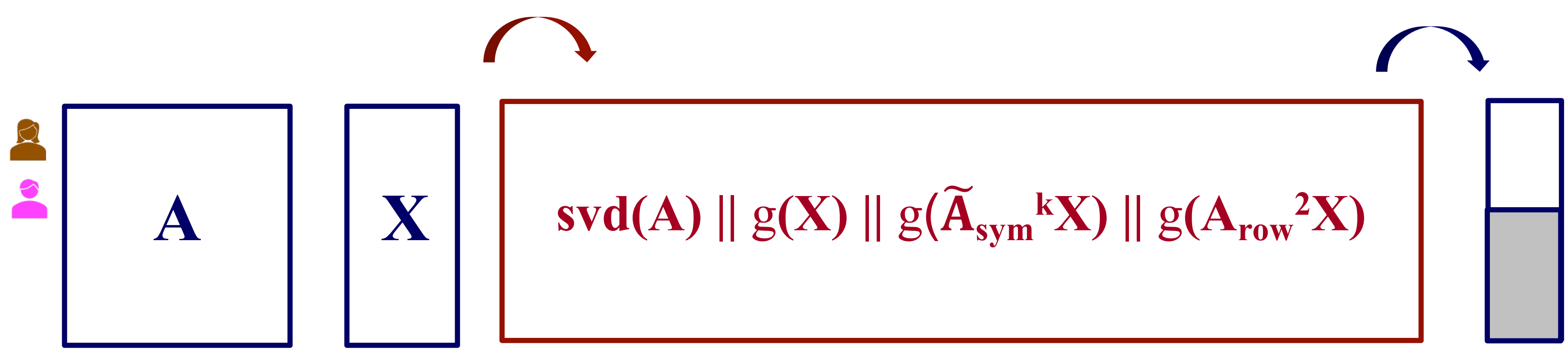
Design Decision D3 – Redundancy

- Q: Do we need all these 'derived features'?



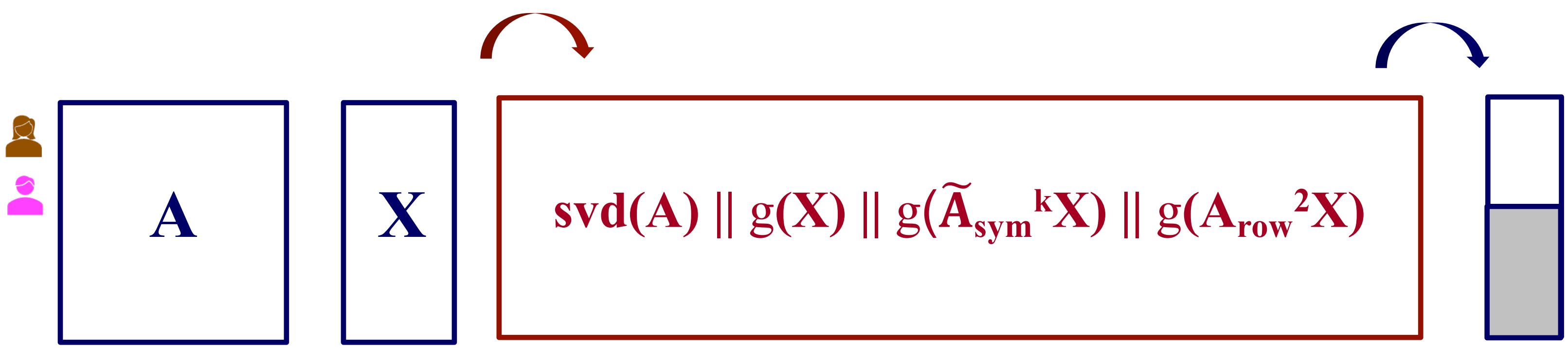
Design Decision D3 – Redundancy

- Q: Do we need all these ‘derived features’?
- A: No, dim. reduction *g()* → SVD/PCA on each component
  - – Reduce correlated features
  - – Fewer, good features → more accurate



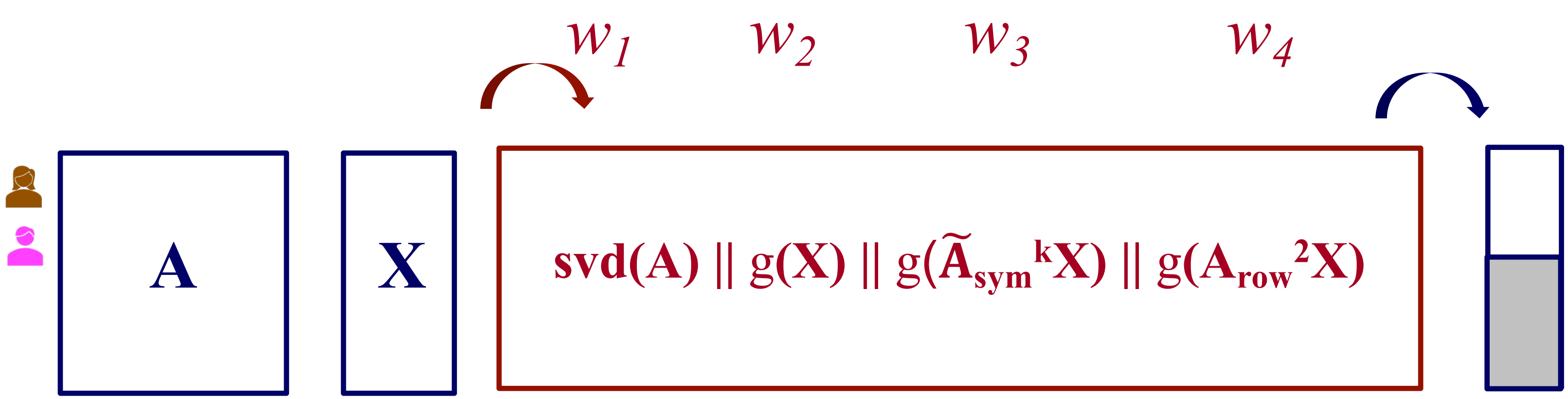
Design Decision D3 – More?

- Q: Do we always need all the components?

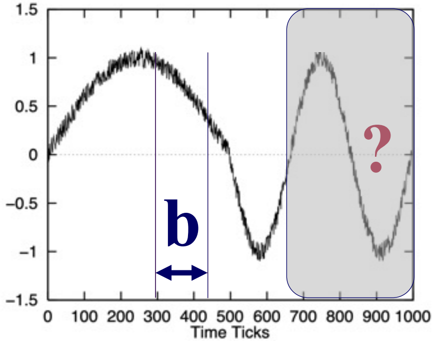


Design Decision D3 – More?

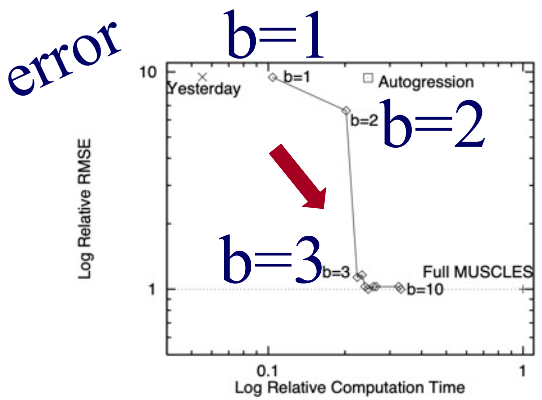
- Q: Do we always need all the components?
- A: Not necessarily, an entire component can be useless as well sometime



‘Muscles’ ICDE 2000



Common sense: ‘*the longer the window, the lower the error*’

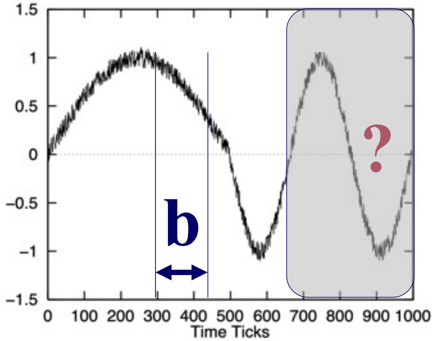


(a) US Dollar (CURRENCY)

run time

[*Online Data Mining for Co-Evolving Time Sequences*]() B.-K. Yi, N.D. Sidiropoulos, T. Johnson, H.V. Jagadish, C. Faloutsos, A. Biliris, ICDE 2000, San Diego, CA, 2000.

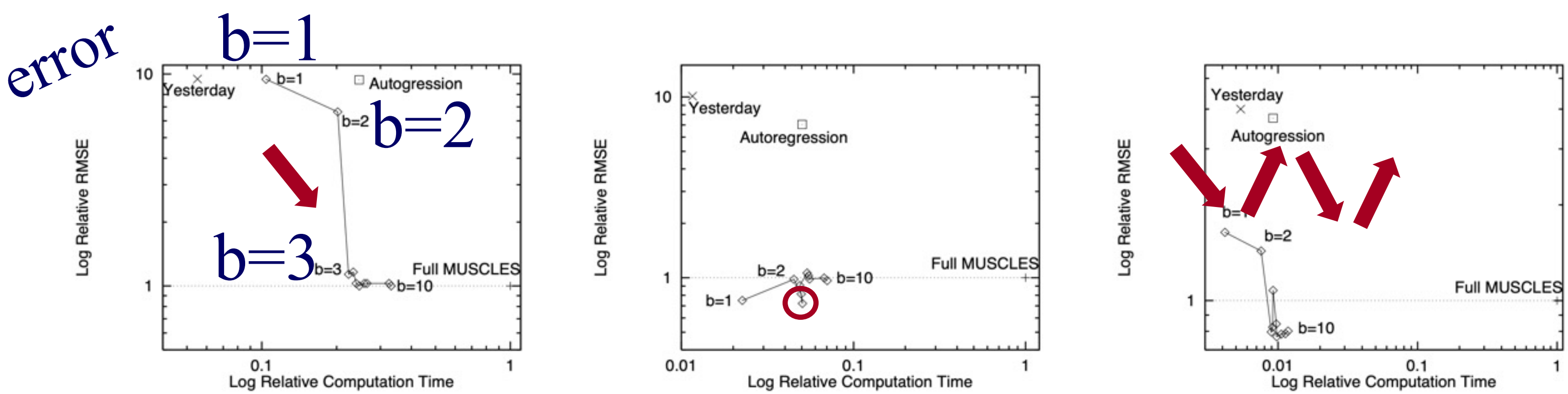
‘Muscles’ ICDE 2000



Common sense: *‘the longer the window, the lower the error’*

NOT always:

‘too many cooks spoil the broth’

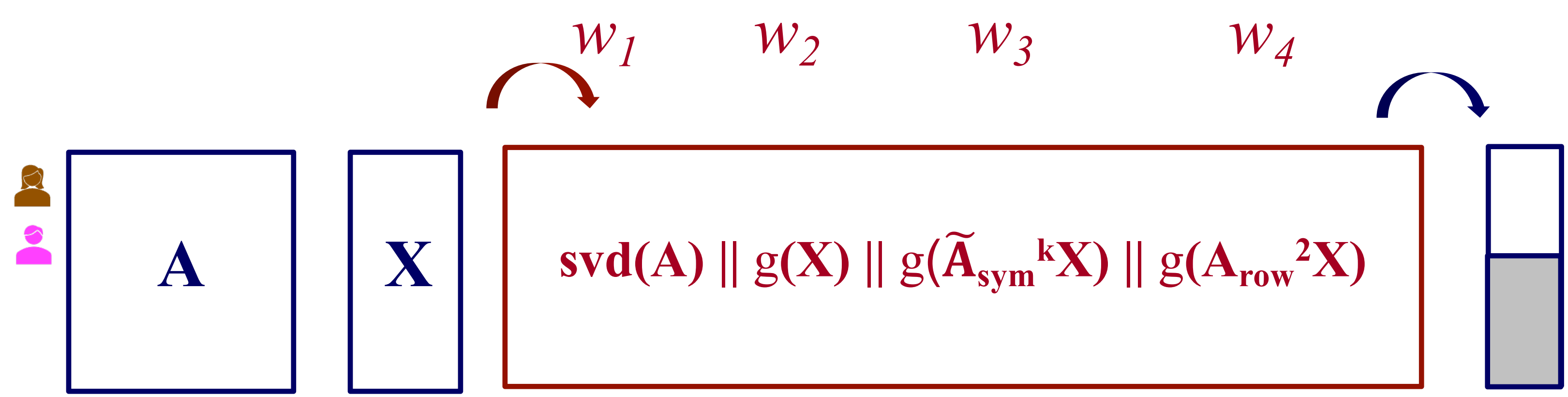


(a) US Dollar (**CURRENCY**) (b) 10-th modem (**MODEM**) (c) 10-th stream (**INTERNET**)

Online Data Mining for Co-Evolving Time Sequences B.-K. Yi, N.D. Sidiropoulos, T. Johnson, H.V. Jagadish, C. Faloutsos, A. Biliris, ICDE 2000, San Diego, CA, 2000.

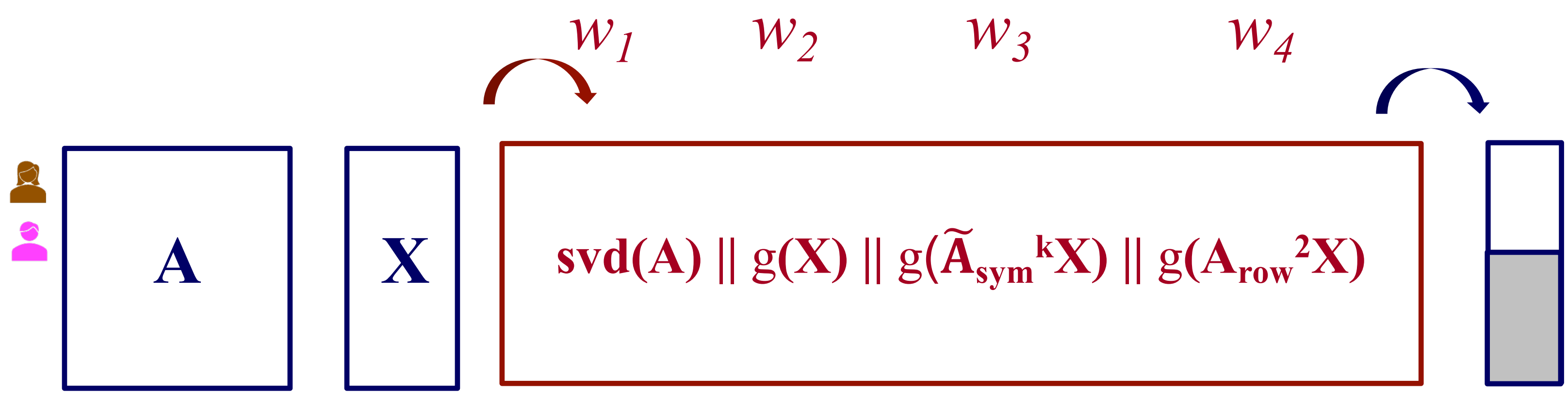
Design Decision D3 – More?

- Q: Do we always need all the components?
- A: Not necessarily, an entire component can be useless as well sometime
- A': Thus, 'group LASSO'



Design Decision D3 – More?

- Q: Do we always need all the components?
- A: Not necessarily, an entire component can be useless as well sometime
- A’: Thus, ‘group LASSO’
  - **Interpretable**

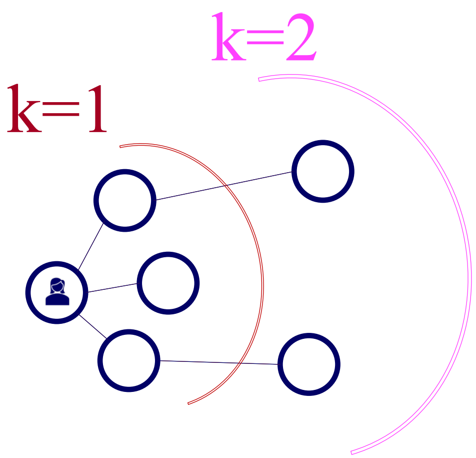


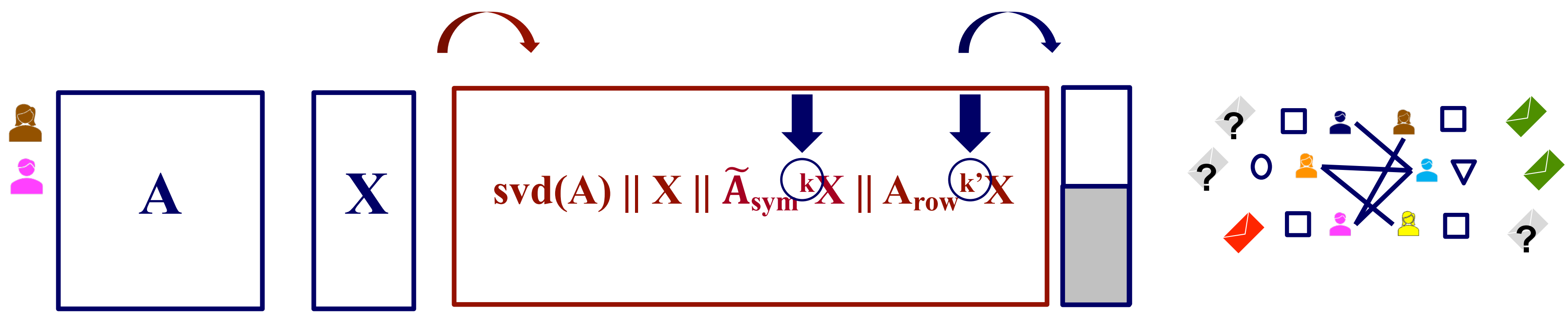
Q2: Method - Desiderata

- Desiderata and Design Decisions
  – D0: Linear Classifier
  – D1: Concatenating Winning Components
  – D2: Structure Features
  – D3: Orthogonalization and Sparsification
  – D4: Multi-Level Neighborhood Aggregation

Design Decision D4 – Exponent

- Q: Influence horizon k = ??





Design Decision D4 – Exponent

- Q: Influence horizon k = ??
- A: Propose: k = 2 (up to grand-neighbors)

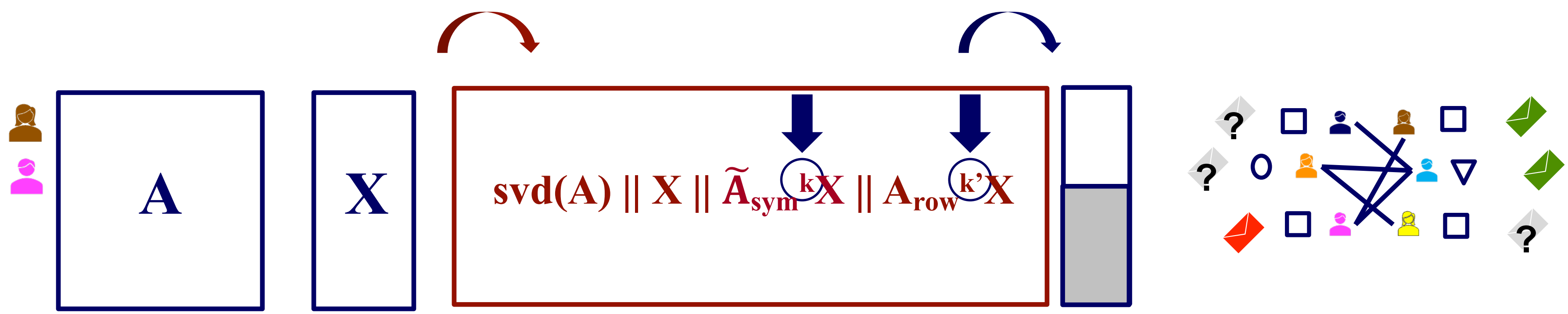
Design Decision D4 – Exponent

- Q: Influence horizon k = ??
- A: Propose: k = 2 (up to grand-neighbors)

→ 0-, 1-, and 2-hop neighborhood

→ b/c influence dissipates with distance

→ b/c no over-smoothing



Q2: Method – SLIMG

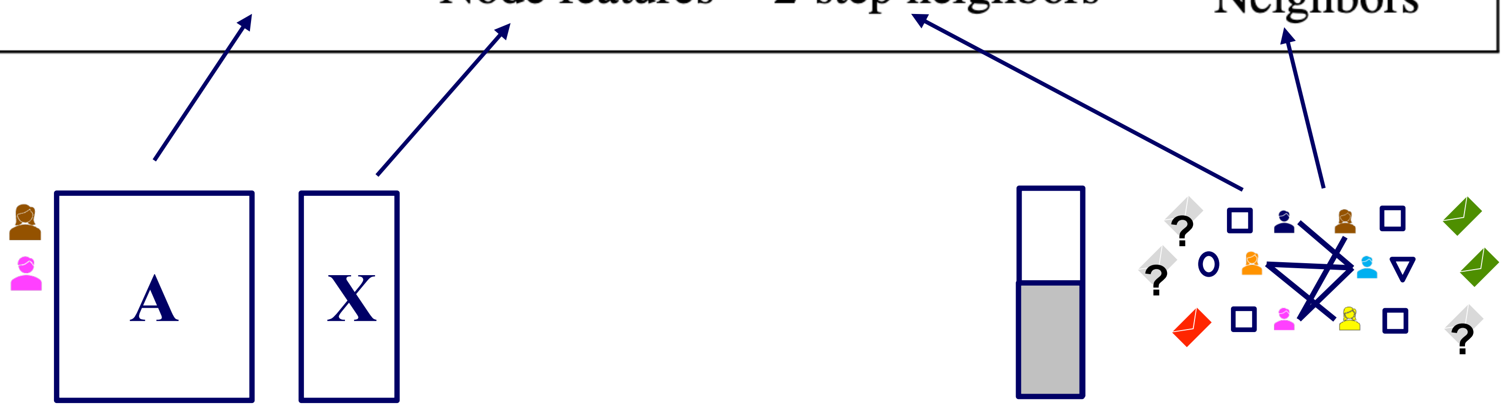
4 components, carefully designed

$$\mathcal{P}(\mathbf{A},\mathbf{X}) = \underbrace{\mathbf{U}}_{\text{Structure}} \parallel \underbrace{g(\mathbf{X})}_{\text{Node features}} \parallel \underbrace{g(\mathbf{A}^2_{\text{row}}\mathbf{X})}_{\text{2-step neighbors}} \parallel \underbrace{g(\tilde{\mathbf{A}}^2_{\text{sym}}\mathbf{X})}_{\text{Neighbors}}$$

Q2: Method – SLIMG

4 components, carefully designed

$$\mathcal{P}(\mathbf{A},\mathbf{X}) = \underbrace{\mathbf{U}}_{\text{Structure}} \;\|\; \underbrace{g(\mathbf{X})}_{\text{Node features}} \;\|\; \underbrace{g(\mathbf{A}^2_{\text{row}}\mathbf{X})}_{\text{2-step neighbors}} \;\|\; \underbrace{g(\tilde{\mathbf{A}}^2_{\text{sym}}\mathbf{X})}_{\text{Neighbors}}$$



Q2: Method – SLIMG

4 components, carefully designed

$$\mathcal{P}(\mathbf{A},\mathbf{X}) = \underbrace{\mathbf{U}}_{\text{Structure}} \parallel \underbrace{g(\mathbf{X})}_{\text{Node features}} \parallel \underbrace{g(\mathbf{A}^2_{\text{row}}\mathbf{X})}_{\text{2-step neighbors}} \parallel \underbrace{g(\tilde{\mathbf{A}}^2_{\text{sym}}\mathbf{X})}_{\text{Neighbors}}$$

Orthogonalization:
PCA

Q2: Method – SLIMG

"Careful Simplicity"

4 components, carefully designed

$$\mathcal{P}(\mathbf{A}, \mathbf{X}) = \underbrace{\mathbf{U}}_{\text{Structure}} \;\|\; \underbrace{g(\mathbf{X})}_{\text{Node features}} \;\|\; \underbrace{g(\mathbf{A}_{\mathrm{row}}^2 \mathbf{X})}_{\text{2-step neighbors}} \;\|\; \underbrace{g(\tilde{\mathbf{A}}_{\mathrm{sym}}^2 \mathbf{X})}_{\text{Neighbors}}$$

Outline

- Problem Definition
- Q1: Reasons
- Q2: Method
- Q3: Sanity Checks
- Experiments
- Conclusions

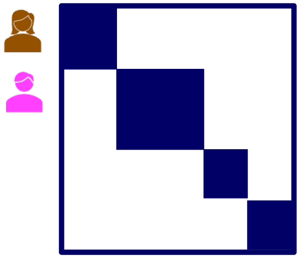
Q3: Sanity Checks

- Q: What are possible settings?
- A: Cross-product of scenarios

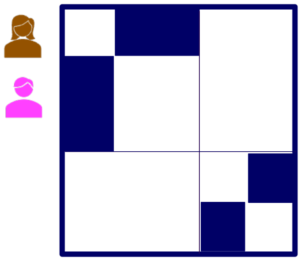
Q3: Sanity Checks

- Q: What are possible settings?
- A: Cross-product of scenarios
  - S1: connectivity
    - Block-diagonal
    - Block-off-diagonal
    - Random

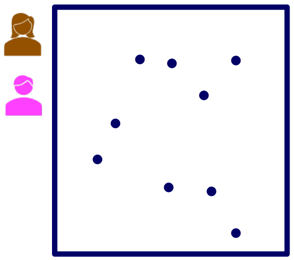
Possible (self/cross-) correlations?



block-diagonal =
communities ~
homophily



block-off-diagonal =
'familiar strangers'~
heterophily

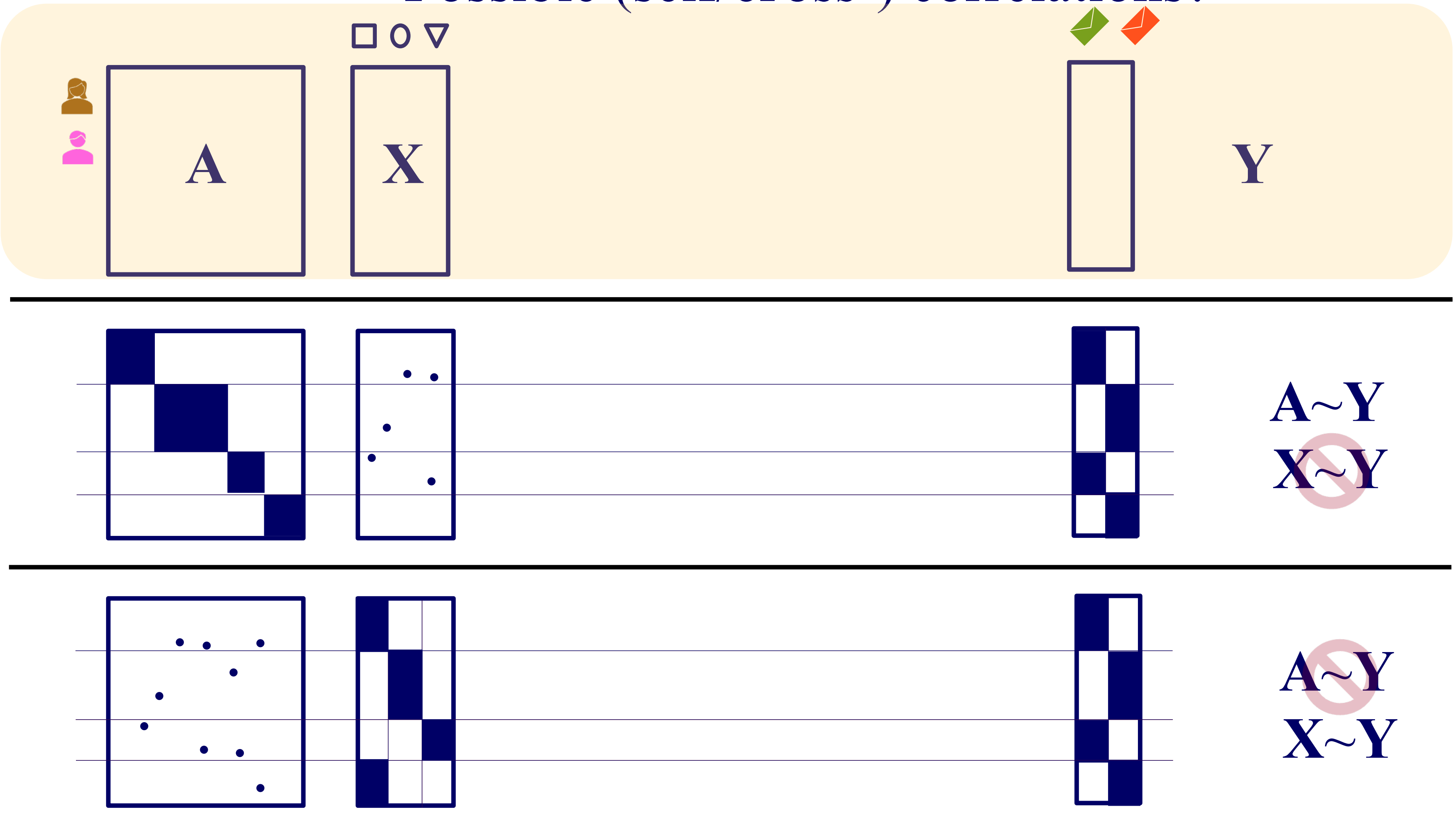


Random

Q3: Sanity Checks

- Q: What are possible settings?
- A: Cross-product of scenarios
  - S1: connectivity
  - S2: cross-correlations – labels are correlated w/
    - Connectivity only (**A**)
    - Features only (**X**)
    - Both
    - (None)

Possible (self/**cross-**) correlations?



(omitted – obvious – everybody does well)

A~Y
X~Y

Q3: Sanity Checks

Only **X** helps | Only **A** helps | Both **A, X** help

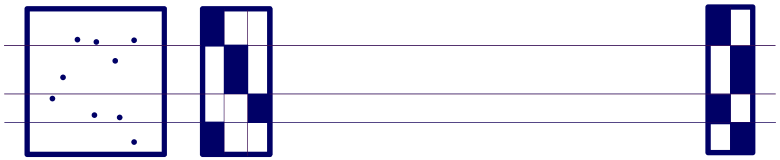
Model	Only X helps Semantic **X** Uniform **A**	Only A helps Random **X** Homophily	Random **X** Heterophily	Both X and A help Structural **X** Homophily	Structural **X** Heterophily	Semantic **X** Homophily	Semantic **X** Heterophily	**Avg. Acc**	**Avg. Rank**
LR	83.7±0.6	24.2±0.7	24.2±0.7	71.4±0.9	66.8±2.2	83.4±0.6	83.4±0.6	62.4 (26.9)	10.7 (5.4)
Reg. Kernel	82.7±0.5	27.9±0.4	24.3±1.0	75.7±0.2	65.3±1.6	91.5±0.5	79.5±0.3	63.8 (27.0)	10.4 (4.3)
Diff. Kernel	26.8±1.7	38.0±8.7	37.6±7.5	79.5±0.3	73.5±0.6	70.9±23.	56.1±27.	54.6 (20.7)	10.6 (4.0)
RW Kernel	72.2±0.7	37.0±0.4	24.5±1.3	81.3±1.2	51.0±1.1	94.5±0.9	57.8±0.7	59.8 (24.7)	10.4 (3.6)
SGC	44.6±9.8	64.3±0.7	50.2±14.	87.1±0.6	84.3±0.5	93.9±0.9	91.5±0.5	73.7 (20.4)	5.7 (3.1)
DGC	63.8±1.0	50.5±13.	26.0±0.9	88.6±1.0	45.3±1.3	96.2±0.4	54.0±0.6	60.6 (24.6)	8.3 (5.9)
S^2GC	79.9±0.6	38.5±12.	25.4±0.9	88.4±1.0	67.9±1.5	95.9±0.6	78.0±0.5	67.7 (26.2)	7.4 (3.4)
G^2CN	25.2±0.3	24.2±1.1	25.0±0.1	88.5±1.0	88.6±1.2	24.3±1.1	50.7±31.	46.6 (30.2)	11.6 (6.3)
GCN	36.3±3.5	46.7±8.0	43.7±1.9	83.3±1.3	72.2±1.7	91.2±1.2	80.3±3.9	64.8 (22.1)	8.1 (3.0)
SAGE	80.3±1.1	31.1±0.7	34.6±2.1	83.9±0.8	81.3±0.7	94.4±0.5	94.4±0.9	71.4 (27.0)	5.7 (2.9)
GCNII	73.5±1.2	30.7±0.7	27.1±1.3	84.2±0.8	69.0±1.4	90.6±0.9	80.4±1.2	65.1 (25.7)	8.7 (1.8)
H^2GCN	80.2±1.5	27.0±1.0	27.5±0.8	78.0±0.9	74.6±1.3	91.9±0.7	92.2±0.9	67.3 (28.2)	8.0 (3.9)
APPNP	66.0±2.6	30.3±1.2	25.2±0.7	71.2±4.9	43.8±2.0	83.2±3.8	58.7±4.5	54.1 (21.6)	12.9 (2.0)
GPR-GNN	73.4±0.4	74.6±0.7	65.9±2.1	89.9±0.6	87.6±1.2	95.0±1.1	91.9±1.1	82.6 (11.2)	3.3 (2.1)
GAT	32.7±5.5	42.6±4.8	36.8±5.7	64.0±5.7	55.6±6.8	68.5±7.1	67.0±12.	52.5 (15.0)	11.6 (4.1)
SlimG (Ours)	81.0±1.1	87.1±1.4	89.2±1.2	88.1±0.5	88.9±0.7	94.4±0.6	93.9±0.5	88.9 (4.5)	2.6 (1.8)

2σ

Q3: Sanity Checks – Part 1/3

Only **X** helps

Model	Only **X** helps Semantic **X** Uniform **A**
LR	83.7±0.6
Reg. Kernel	82.7±0.5
Diff. Kernel	26.8±1.7
RW Kernel	72.2±0.7
SGC	44.6±9.8
DGC	63.8±1.0
S^2GC	79.9±0.6
G^2CN	25.2±0.3
GCN	36.3±3.5
SAGE	80.3±1.1
GCNII	73.5±1.2
H^2GCN	80.2±1.5
APPNP	66.0±2.6
GPR-GNN	73.4±0.4
GAT	32.7±5.5
SlimG (Ours)	81.0±1.1

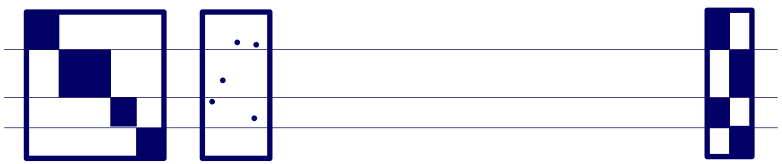


1) Logistic Reg. does great
2) … as well as SlimG
3) Rest: confused
    1) … often 1/4 (random)

Q3: Sanity Checks – Part 2/3

Only **A** helps

Model	Only **A** helps Random **X** Homophily	 Random **X** Heterophily
LR	24.2±0.7	24.2±0.7
Reg. Kernel	27.9±0.4	24.3±1.0
Diff. Kernel	38.0±8.7	37.6±7.5
RW Kernel	37.0±0.4	24.5±1.3
SGC	64.3±0.7	50.2±14.
DGC	50.5±13.	26.0±0.9
S^2GC	38.5±12.	25.4±0.9
G^2CN	24.2±1.1	25.0±0.1
GCN	46.7±8.0	43.7±1.9
SAGE	31.1±0.7	34.6±2.1
GCNII	30.7±0.7	27.1±1.3
H^2GCN	27.0±1.0	27.5±0.8
APPNP	30.3±1.2	25.2±0.7
GPR-GNN	74.6±0.7	65.9±2.1
GAT	42.6±4.8	36.8±5.7
SLIMG (Ours)	87.1±1.4	89.2±1.2



As expected:

1) **Only** SLIMG works
2) All else fail

$$(\mathbf{A}^k+\mathbf{A}^j+\cdots)\mathbf{X} \,||\, (\mathbf{A}^{k'}\ldots)\mathbf{X}$$

Q3: Sanity Checks – Part 3/3

Both **A, X** help

Model	Both **X** and **A** help Structural **X** Homophily	Structural **X** Heterophily	Semantic **X** Homophily	Semantic **X** Heterophily
LR	71.4±0.9	66.8±2.2	83.4±0.6	83.4±0.6
Reg. Kernel	75.7±0.2	65.3±1.6	91.5±0.5	79.5±0.3
Diff. Kernel	79.5±0.3	73.5±0.6	70.9±23.	56.1±27.
RW Kernel	81.3±1.2	51.0±1.1	94.5±0.9	57.8±0.7
SGC	87.1±0.6	84.3±0.5	93.9±0.9	91.5±0.5
DGC	88.6±1.0	45.3±1.3	96.2±0.4	54.0±0.6
S^2GC	88.4±1.0	67.9±1.5	95.9±0.6	78.0±0.5
G^2CN	88.5±1.0	88.6±1.2	24.3±1.1	50.7±31.
GCN	83.3±1.3	72.2±1.7	91.2±1.2	80.3±3.9
SAGE	83.9±0.8	81.3±0.7	94.4±0.5	94.4±0.9
GCNII	84.2±0.8	69.0±1.4	90.6±0.9	80.4±1.2
H^2GCN	78.0±0.9	74.6±1.3	91.9±0.7	92.2±0.9
APPNP	71.2±4.9	43.8±2.0	83.2±3.8	58.7±4.5
GPR-GNN	89.9±0.6	87.6±1.2	95.0±1.1	91.9±1.1
GAT	64.0±5.7	55.6±6.8	68.5±7.1	67.0±12.
SLIMG (Ours)	88.1±0.5	88.9±0.7	94.4±0.6	93.9±0.5

1) Several do OK on hom.
2) Fewer work on het.
3) SLIMG: never ‘red’

Outline

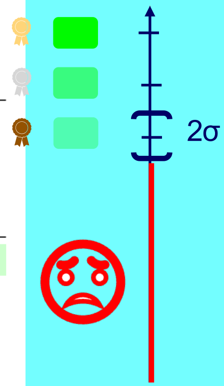
- Problem Definition
- Q1: Reasons
- Q2: Method
- Q3: Sanity Checks
- Experiments
  - E1: Accurate & Robust
  - E2: Interpretable
  - E3: Fast & Scalable
- Conclusions

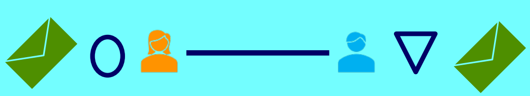
E1: Accurate & Robust

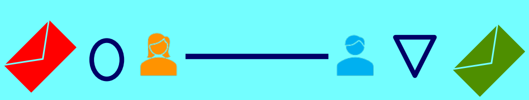
Real-World Datasets

7 Homophily | 6 Heterophily

Model	Cora	CiteSeer	PubMed	Comp.	Photo	ArXiv	Products	Cham.	Squirrel	Actor	Penn94	Twitch	Pokec	Avg. Rank
LR	51.5±1.2	52.9±4.5	79.9±0.5	73.9±1.2	79.3±1.5	48.3±1.9	56.4±0.5	24.9±1.7	26.7±1.9	27.8±0.8	63.5±0.5	53.0±0.1	61.3±0.0	11.7 (4.2)
Reg. Kernel	67.8±2.5	62.1±4.4	83.4±1.4	80.3±1.4	87.1±1.2	O.O.M.	O.O.M.	29.4±2.6	24.3±2.3	29.6±1.4	O.O.M.	O.O.M.	O.O.M.	12.2 (3.8)
Diff. Kernel	70.6±1.5	62.7±3.8	82.1±0.4	83.1±1.0	89.8±0.6	O.O.M.	O.O.M.	34.5±7.9	28.3±1.5	24.7±0.9	53.5±0.8	O.O.M.	O.O.M.	11.8 (2.5)
RW Kernel	72.7±1.7	64.1±3.9	83.1±0.7	84.2±0.7	90.6±0.7	63.2±0.2	74.2±0.0	34.9±3.5	25.0±1.6	26.4±1.1	63.1±0.7	57.6±0.1	59.5±0.0	8.3 (3.3)
SGC	76.2±1.1	65.8±3.9	84.1±0.8	83.7±1.6	90.1±0.9	65.0±3.4	74.6±5.1	38.1±4.5	33.1±1.0	24.6±0.8	64.0±1.1	56.5±0.1	69.8±0.0	6.6 (4.2)
DGC	77.8±1.4	66.1±4.2	84.3±0.6	83.9±0.7	90.4±0.2	65.2±4.0	68.7±13.	37.2±3.7	29.2±1.2	25.2±2.1	62.5±0.4	58.2±0.2	60.7±0.1	6.6 (3.2)
S^2GC	78.3±1.5	66.9±4.4	84.3±0.3	83.1±0.8	90.1±0.8	62.0±7.4	58.3±18.	34.9±4.9	27.6±1.8	26.7±1.8	63.1±0.5	58.7±0.1	61.2±0.0	6.6 (2.7)
G^2CN	76.6±1.5	64.2±3.3	81.4±0.6	82.8±1.6	88.8±0.5	O.O.M.	O.O.M.	40.7±2.9	32.1±1.5	24.3±0.5	O.O.M.	O.O.M.	O.O.M.	10.5 (4.5)
GCN	76.0±1.2	65.0±2.9	84.3±0.5	85.1±0.9	91.6±0.5	62.8±0.6	O.O.M.	38.5±3.0	31.4±1.8	26.8±0.4	62.9±0.7	57.0±0.1	63.9±0.4	6.3 (2.4)
SAGE	74.6±1.3	63.7±3.6	82.9±0.4	83.8±0.5	90.6±0.5	61.5±0.6	O.O.M.	39.8±4.3	27.0±1.3	27.8±0.9	O.O.M.	56.6±0.4	68.9±0.1	8.5 (3.5)
GCNII	77.8±1.7	63.4±3.0	84.9±0.8	82.3±1.8	90.8±0.6	45.7±0.5	O.O.M.	30.5±2.5	21.9±3.0	29.0±1.3	64.5±0.5	56.9±0.6	62.1±0.3	8.4 (4.6)
H^2GCN	77.6±0.9	64.7±3.8	85.4±0.4	49.5±16.	75.8±11.	O.O.M.	O.O.M.	31.9±2.6	25.0±0.5	28.9±0.6	63.9±0.4	58.7±0.0	O.O.M.	8.9 (4.9)
APPNP	80.0±0.6	67.1±2.8	84.6±0.5	84.2±1.7	92.5±0.3	53.4±1.3	O.O.M.	30.9±4.7	23.9±3.2	26.1±1.0	63.7±0.9	47.3±0.3	57.4±0.4	7.6 (4.8)
GPR-GNN	78.8±1.3	64.2±4.0	85.1±0.7	85.0±1.0	92.6±0.3	58.5±0.8	O.O.M.	31.7±4.7	26.2±1.6	29.5±1.1	64.5±0.4	57.6±0.2	67.6±0.1	5.4 (3.7)
GAT	78.2±1.2	65.8±4.0	83.6±0.2	85.4±1.4	91.7±0.5	58.2±1.0	O.O.M.	39.1±4.1	28.6±0.6	26.4±0.4	60.5±0.8	O.O.M.	O.O.M.	7.5 (3.7)
SLIMG	77.8±1.1	67.1±2.3	84.6±0.5	86.3±0.7	91.8±0.5	66.3±0.3	84.9±0.0	40.8±3.2	31.1±0.7	30.9±0.6	68.2±0.6	59.7±0.1	73.9±0.1	1.9 (1.5)







E1: Accurate & Robust

Real-W[illegible]rld Datasets

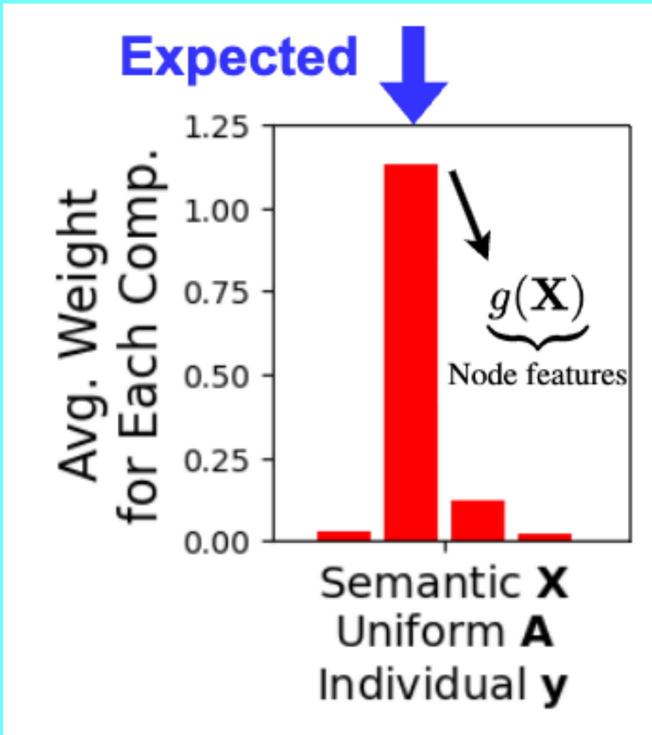
Many O.O.M.

7 Homophi[illegible] | 6 Heterophily

Model	Cora	CiteSeer	PubM[illegible]	[illegible]	[illegible]	ArXiv	Products	Cham.	Squirrel	Actor	Penn94	Twitch	Pokec	Avg. Rank
LR	51.5±1.2	52.9±4.5	[illegible]	[illegible]	[illegible]	48.3±1.9	56.4±0.5	24.9±1.7	26.7±1.9	27.8±0.8	63.5±0.5	53.0±0.1	61.3±0.0	11.7 (4.2)
Reg. Kernel	67.8±2.5	[illegible]	[illegible]	[illegible]	87.1±1.2	O.O.M.	O.O.M.	29.4±2.6	24.3±2.3	29.6±1.4	O.O.M.	O.O.M.	O.O.M.	12.2 (3.8)
Diff. Kernel	[illegible]	[illegible]	[illegible]	83.1±1.0	89.8±0.6	O.O.M.	O.O.M.	34.5±7.9	28.3±1.5	24.7±0.9	53.5±0.8	O.O.M.	O.O.M.	11.8 (2.5)
RW Kernel	[illegible]	[illegible]	[illegible]	84.2±0.7	90.6±0.7	63.2±0.2	74.2±0.0	34.9±3.5	25.0±1.6	26.4±1.1	63.1±0.7	57.6±0.1	59.5±0.0	8.3 (3.3)
SGC	[illegible]	[illegible]	84.1±0.8	83.7±1.6	90.1±0.9	65.0±3.4	74.6±5.1	38.1±4.5	33.1±1.0	24.6±0.8	64.0±1.1	56.5±0.1	69.8±0.0	6.6 (4.2)
[illegible]	[illegible]	[illegible]	84.3±0.6	83.9±0.7	90.4±0.2	65.2±4.0	68.7±13.	37.2±3.7	29.2±1.2	25.2±2.1	62.5±0.4	58.2±0.2	60.7±0.1	6.6 (3.2)
[illegible]	[illegible]	66.9±4.4	84.3±0.3	83.1±0.8	90.1±0.8	62.0±7.4	58.3±18.	34.9±4.9	27.6±1.8	26.7±1.8	63.1±0.5	58.7±0.1	61.2±0.0	6.6 (2.7)
[illegible]	[illegible]	64.2±3.3	81.4±0.6	82.8±1.6	88.8±0.5	O.O.M.	O.O.M.	40.7±2.9	32.1±1.5	24.3±0.5	O.O.M.	O.O.M.	O.O.M.	10.5 (4.5)
[illegible]	76.0±1.2	65.0±2.9	84.3±0.5	85.1±0.9	91.6±0.5	62.8±0.6	O.O.M.	38.5±3.0	31.4±1.8	26.8±0.4	62.9±0.7	57.0±0.1	63.9±0.4	6.3 (2.4)
SAGE	74.6±1.3	63.7±3.6	82.9±0.4	83.8±0.5	90.6±0.5	61.5±0.6	O.O.M.	39.8±4.3	27.0±1.3	27.8±0.9	O.O.M.	56.6±0.4	68.9±0.1	8.5 (3.5)
GCNII	77.8±1.7	63.4±3.0	84.9±0.8	82.3±1.8	90.8±0.6	45.7±0.5	O.O.M.	30.5±2.5	21.9±3.0	29.0±1.3	64.5±0.5	56.9±0.6	62.1±0.3	8.4 (4.6)
H^2GCN	77.6±0.9	64.7±3.8	85.4±0.4	49.5±16.	75.8±11.	O.O.M.	O.O.M.	31.9±2.6	25.0±0.5	28.9±0.6	63.9±0.4	58.7±0.0	O.O.M.	8.9 (4.9)
APPNP	80.0±0.6	67.1±2.8	84.6±0.5	84.2±1.7	92.5±0.3	53.4±1.3	O.O.M.	30.9±4.7	23.9±3.2	26.1±1.0	63.7±0.9	47.3±0.3	57.4±0.4	7.6 (4.8)
GPR-GNN	78.8±1.3	64.2±4.0	85.1±0.7	85.0±1.0	92.6±0.3	58.5±0.8	O.O.M.	31.7±4.7	26.2±1.6	29.5±1.1	64.5±0.4	57.6±0.2	67.6±0.1	5.4 (3.7)
GAT	78.2±1.2	65.8±4.0	83.6±0.2	85.4±1.4	91.7±0.5	58.2±1.0	O.O.M.	39.1±4.1	28.6±0.6	26.4±0.4	60.5±0.8	O.O.M.	O.O.M.	7.5 (3.7)
SlimG	77.8±1.1	67.1±2.3	84.6±0.5	86.3±0.7	91.8±0.5	66.3±0.3	84.9±0.0	40.8±3.2	31.1±0.7	30.9±0.6	68.2±0.6	59.7±0.1	73.9±0.1	1.9 (1.5)

2σ

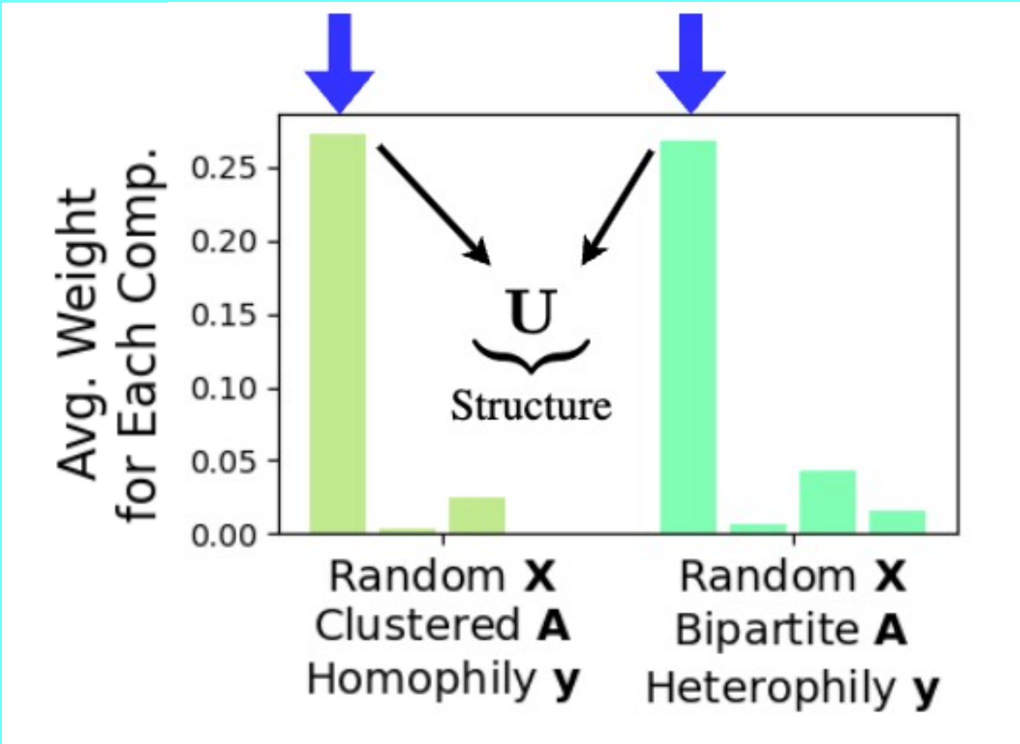
E2: Interpretable

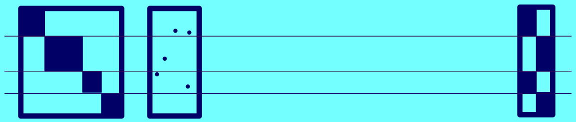


structure
features
neighbors
grand-neighb.

$$\underbrace{\mathbf{U}}_{\text{Structure}} \,\|\, \boxed{\underbrace{g(\mathbf{X})}_{\text{Node features}}} \,\|\, \underbrace{g(\mathbf{A}^2_{\text{row}}\mathbf{X})}_{\text{2-step neighbors}} \,\|\, \underbrace{g(\tilde{\mathbf{A}}^2_{\text{sym}}\mathbf{X})}_{\text{Neighbors}}$$

E2: Interpretable

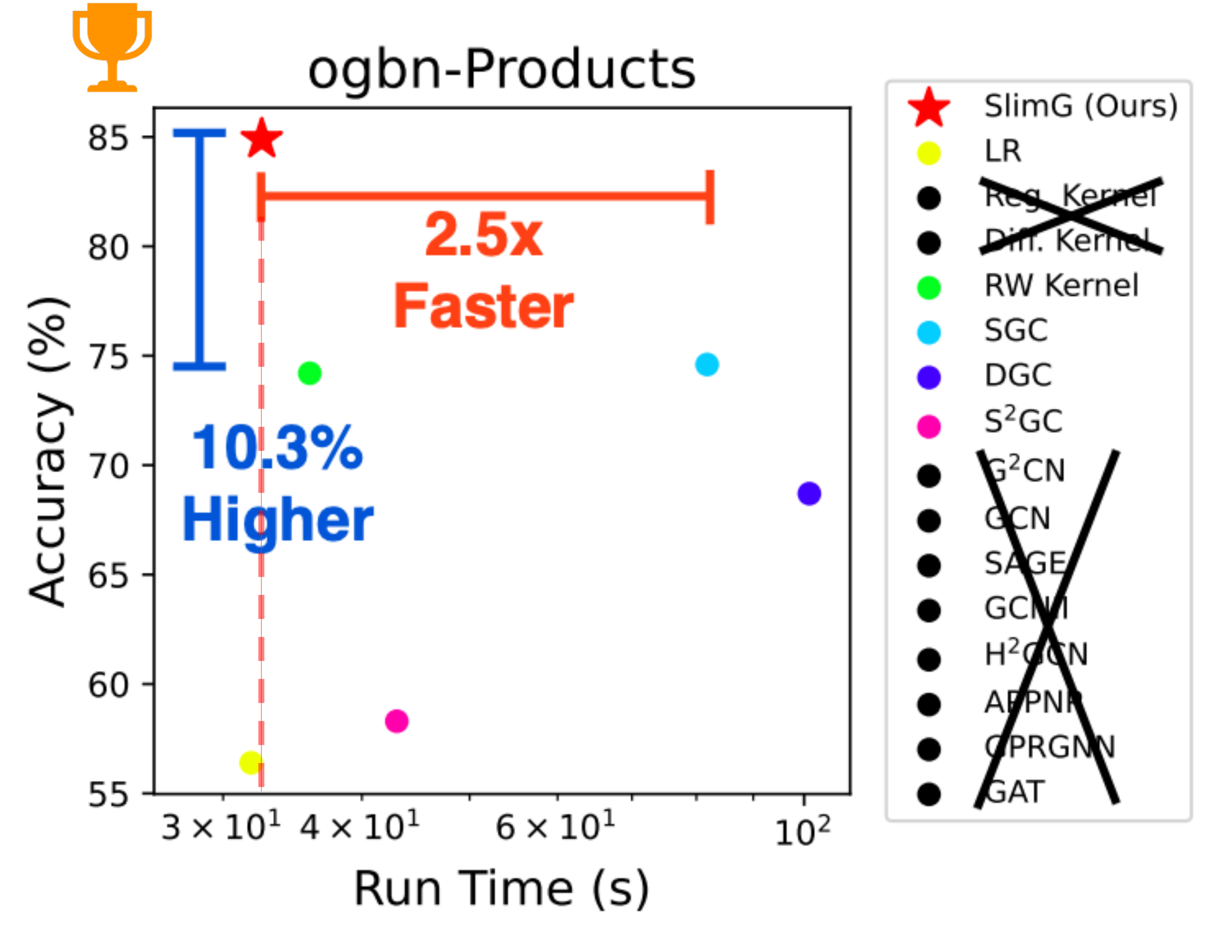




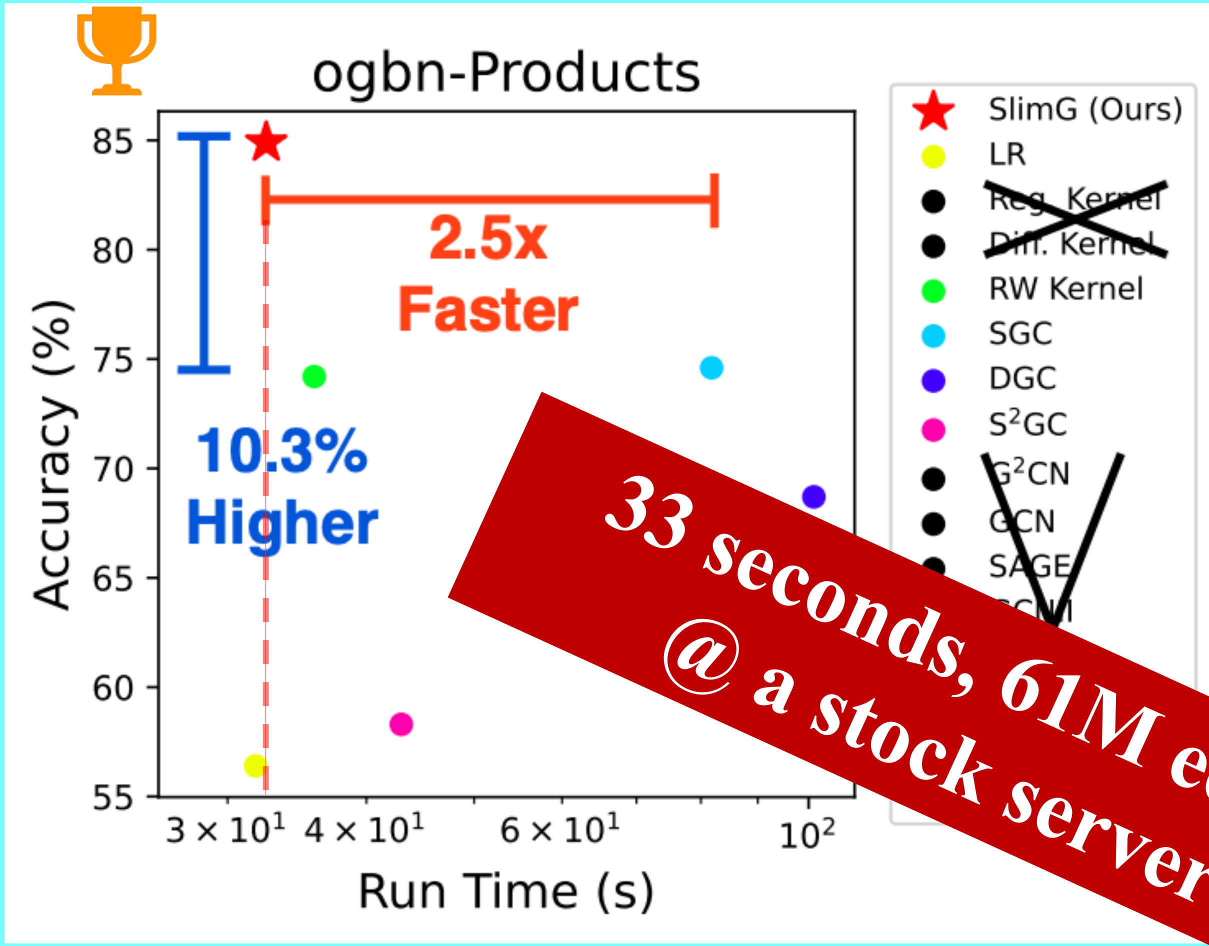
structure
features
neighbors
grand-neighb.

$$\boxed{\underbrace{\mathbf{U}}_{\text{Structure}}} \| \underbrace{g(\mathbf{X})}_{\text{Node features}} \| \underbrace{g(\mathbf{A}_{\text{row}}^{2}\mathbf{X})}_{\text{2-step neighbors}} \| \underbrace{g(\tilde{\mathbf{A}}_{\text{sym}}^{2}\mathbf{X})}_{\text{Neighbors}}$$

E3: Fast & Scalable



E3: Fast & Scalable



33 seconds, 61M edges @ a stock server

Conclusions

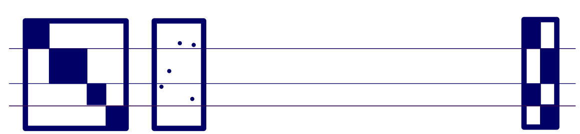
$$\underbrace{\mathbf{U}}_{\text{Structure}} \;\|\; \underbrace{g(\mathbf{X})}_{\text{Node features}} \;\|\; \underbrace{g(\mathbf{A}^2_{\text{row}}\mathbf{X})}_{\text{2-step neighbors}} \;\|\; \underbrace{g(\tilde{\mathbf{A}}^2_{\text{sym}}\mathbf{X})}_{\text{Neighbors}}$$

Model	Cora	CiteSeer	PubMed	Comp.	Photo	ArXiv	Products	Cham.	Squirrel	Actor	Penn94	Twitch	Pokec	Avg. Rank
LR	51.5±1.2	52.9±4.5	79.9±0.5	73.9±1.2	79.3±1.5	48.3±1.9	56.4±0.5	24.9±1.7	26.7±1.9	27.8±0.8	63.5±0.5	53.0±0.1	61.3±0.0	11.7 (4.2)
Reg. Kernel	67.8±2.5	62.1±4.4	83.4±1.4	80.3±1.4	87.1±1.2	O.O.M.	O.O.M.	29.4±2.6	24.3±2.3	29.6±1.4	O.O.M.	O.O.M.	O.O.M.	12.2 (3.8)
Diff. Kernel	70.6±1.5	62.7±3.8	82.1±0.4	83.1±1.0	89.8±0.6	O.O.M.	O.O.M.	34.5±7.9	28.3±1.5	24.7±0.9	53.5±0.8	O.O.M.	O.O.M.	11.8 (2.5)
RW Kernel	72.7±1.7	64.1±3.9	83.1±0.7	84.2±0.7	90.6±0.7	63.2±0.2	74.2±0.0	34.9±3.5	25.0±1.6	26.4±1.1	63.1±0.7	57.6±0.1	59.5±0.0	8.3 (3.3)
SGC	76.2±1.1	65.8±3.9	84.1±0.8	83.7±1.6	90.1±0.9	65.0±3.4	74.6±5.1	38.1±4.5	33.1±1.0	24.6±0.8	64.0±1.1	56.5±0.1	69.8±0.0	6.6 (4.2)
DGC	77.8±1.4	66.1±4.2	84.3±0.6	83.9±0.7	90.4±0.2	65.2±4.0	68.7±13.	37.2±3.7	29.2±1.2	25.2±2.1	62.5±0.4	58.2±0.2	60.7±0.1	6.6 (3.2)
S^2GC	78.3±1.5	66.9±4.4	84.3±0.3	83.1±0.8	90.1±0.8	62.0±7.4	58.3±18.	34.9±4.9	27.6±1.8	26.7±1.8	63.1±0.5	58.7±0.1	61.2±0.0	6.6 (2.7)
G^2CN	76.6±1.5	64.2±3.3	81.4±0.6	82.8±1.6	88.8±0.5	O.O.M.	O.O.M.	40.7±2.9	32.1±1.5	24.3±0.5	O.O.M.	O.O.M.	O.O.M.	10.5 (4.5)
GCN	76.0±1.2	65.0±2.9	84.3±0.5	85.1±0.9	91.6±0.5	62.8±0.6	O.O.M.	38.5±3.0	31.4±1.8	26.8±0.4	62.9±0.7	57.0±0.1	63.9±0.4	6.3 (2.4)
SAGE	74.6±1.3	63.7±3.6	82.9±0.4	83.8±0.5	90.6±0.5	61.5±0.6	O.O.M.	39.8±4.3	27.0±1.3	27.8±0.9	O.O.M.	56.6±0.4	68.9±0.1	8.5 (3.5)
GCNII	77.8±1.7	63.4±3.0	84.9±0.8	82.3±1.8	90.8±0.6	45.7±0.5	O.O.M.	30.5±2.5	21.9±3.0	29.0±1.3	64.5±0.5	56.9±0.6	62.1±0.3	8.4 (4.6)
H^2GCN	77.6±0.9	64.7±3.8	85.4±0.4	49.5±16.	75.8±11.	O.O.M.	O.O.M.	31.9±2.6	25.0±0.5	28.9±0.6	63.9±0.4	58.7±0.0	O.O.M.	8.9 (4.9)
APPNP	80.0±0.6	67.1±2.8	84.6±0.5	84.2±1.7	92.5±0.3	53.4±1.3	O.O.M.	30.9±4.7	23.9±3.2	26.1±1.0	63.7±0.9	47.3±0.3	57.4±0.4	7.6 (4.8)
GPR-GNN	78.8±1.3	64.2±4.0	85.1±0.7	85.0±1.0	92.6±0.3	58.5±0.8	O.O.M.	31.7±4.7	26.2±1.6	29.5±1.1	64.5±0.4	57.6±0.2	67.6±0.1	5.4 (3.7)
GAT	78.2±1.2	65.8±4.0	83.6±0.2	85.4±1.4	91.7±0.5	58.2±1.0	O.O.M.	39.1±4.1	28.6±0.6	26.4±0.4	60.5±0.8	O.O.M.	O.O.M.	7.5 (3.7)
SlimG	77.8±1.1	67.1±2.3	84.6±0.5	86.3±0.7	91.8±0.5	66.3±0.3	84.9±0.0	40.8±3.2	31.1±0.7	30.9±0.6	68.2±0.6	59.7±0.1	73.9±0.1	1.9 (1.5)

✓ Q1: Reasons

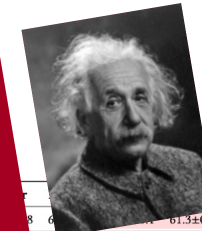
✓ Q2: Method 

– Accurate, Robust, Interpretable, Fast, Scalable

✓ Q3: Sanity Checks 

Conclusions

“Careful Simplicity”

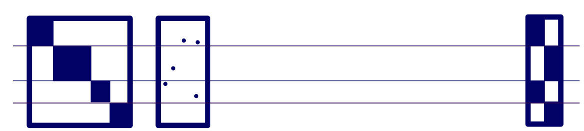


$$\underbrace{\mathbf{U}}_{\text{Structure}} \;\|\; \underbrace{g(\mathbf{X})}_{\text{No...}} \;\|\; g(\mathbf{A}^2_{\text{row}}\mathbf{X}) \;\|\; g(\tilde{\mathbf{A}}^2 \ldots$$

															Avg. Rank
														61.3±0.0	11.7 (4.2)
												O.O.M.	O.O.M.	O.O.M.	12.2 (3.8)
											24.7±0.9	53.5±0.8	O.O.M.	O.O.M.	11.8 (2.5)
								34.9±3.5	25.0±1.6	26.4±1.1	63.1±0.7	57.6±0.1	59.5±0.0		8.3 (3.3)
						65.0±3.4	74.6±5.1	38.1±4.5	33.1±1.0	24.6±0.8	64.0±1.1	56.5±0.1	69.8±0.0		6.6 (4.2)
				83.9±0.7	90.4±0.2	65.2±4.0	68.7±13.	37.2±3.7	29.2±1.2	25.2±2.1	62.5±0.4	58.2±0.2	60.7±0.1		6.6 (3.2)
		66.9±4.4	84.3±0.3	83.1±0.8	90.1±0.8	62.0±7.4	58.3±18.	34.9±4.9	27.6±1.8	26.7±1.8	63.1±0.5	58.7±0.1	61.2±0.0		6.6 (2.7)
	76.6±1.5	64.2±3.3	81.4±0.6	82.8±1.6	88.8±0.5	O.O.M.	O.O.M.	40.7±2.9	32.1±1.5	24.3±0.5	O.O.M.	O.O.M.	O.O.M.		10.5 (4.5)
GCN	76.0±1.2	65.0±2.9	84.3±0.5	85.1±0.9	91.6±0.5	62.8±0.6	O.O.M.	38.5±3.0	31.4±1.8	26.8±0.4	62.9±0.7	57.0±0.1	63.9±0.4		6.3 (2.4)
SAGE	74.6±1.3	63.7±3.6	82.9±0.4	83.8±0.5	90.6±0.5	61.5±0.6	O.O.M.	39.8±4.3	27.0±1.3	27.8±0.9	O.O.M.	56.6±0.4	68.9±0.1		8.5 (3.5)
GCNII	77.8±1.7	63.4±3.0	84.9±0.8	82.3±1.8	90.8±0.6	45.7±0.5	O.O.M.	30.5±2.5	21.9±3.0	29.0±1.3	64.5±0.5	56.9±0.6	62.1±0.3		8.4 (4.6)
H²GCN	77.6±0.9	64.7±3.8	85.4±0.4	49.5±16.	75.8±11.	O.O.M.	O.O.M.	31.9±2.6	25.0±0.5	28.9±0.6	63.9±0.4	58.7±0.0	O.O.M.		8.9 (4.9)
APPNP	80.0±0.6	67.1±2.8	84.6±0.5	84.2±1.7	92.5±0.3	53.4±1.3	O.O.M.	30.9±4.7	23.9±3.2	26.1±1.0	63.7±0.9	47.3±0.3	57.4±0.4		7.6 (4.8)
GPR-GNN	78.8±1.3	64.2±4.0	85.1±0.7	85.0±1.0	92.6±0.3	58.5±0.8	O.O.M.	31.7±4.7	26.2±1.6	29.5±1.1	64.5±0.4	57.6±0.2	67.6±0.1		5.4 (3.7)
GAT	78.2±1.2	65.8±4.0	83.6±0.2	85.4±1.4	91.7±0.5	58.2±1.0	O.O.M.	39.1±4.1	28.6±0.6	26.4±0.4	60.5±0.8	O.O.M.	O.O.M.		7.5 (3.7)
SlimG	77.8±1.1	67.1±2.3	84.6±0.5	86.3±0.7	91.8±0.5	66.3±0.3	84.9±0.0	40.8±3.2	31.1±0.7	30.9±0.6	68.2±0.6	59.7±0.1	73.9±0.1		1.9 (1.5)

✓ Q2: Method 

– Accurate, Robust, Interpretable, Fast, Scalable

✓ Q3: Sanity Checks 

Thank You!

*Jaemin YOO**

*Meng-Chieh (Jeremy) LEE**

Shubhranshu SHEKHAR

Christos FALOUTSOS

✓Q1: Reasons ✓Q2: Method ✓Q3: Sanity Checks

Model	Cora	CiteSeer	PubMed	Comp.	Photo	ArXiv	Products	Cham.	Squirrel	Actor	Penn94	Twitch	Pokec	Avg. Rank
LR	51.5±1.2	52.9±4.5	79.9±0.5	73.9±1.2	79.3±1.5	48.3±1.9	56.4±0.5	24.9±1.7	26.7±1.9	27.8±0.8	63.5±0.5	53.0±0.1	61.3±0.0	11.7 (4.2)
Reg. Kernel	67.8±2.5	62.1±4.4	83.4±1.4	80.3±1.4	87.1±1.2	O.O.M.	O.O.M.	29.4±2.6	24.3±2.3	29.6±1.4	O.O.M.	O.O.M.	O.O.M.	12.2 (3.8)
Diff. Kernel	70.6±1.5	62.7±3.8	82.1±0.4	83.1±1.0	89.8±0.6	O.O.M.	O.O.M.	34.5±7.9	28.3±1.5	24.7±0.9	53.5±0.8	O.O.M.	O.O.M.	11.8 (2.5)
RW Kernel	72.7±1.7	64.1±3.9	83.1±0.7	84.2±0.7	90.6±0.7	63.2±0.2	74.2±0.0	34.9±3.5	25.0±1.6	26.4±1.1	63.1±0.7	57.6±0.1	59.5±0.0	8.3 (3.3)
SGC	76.2±1.1	65.8±3.9	84.1±0.8	83.7±1.6	90.1±0.9	65.0±3.4	74.6±5.1	38.1±4.5	33.1±1.0	24.6±0.8	64.0±1.1	56.5±0.1	69.8±0.0	6.6 (4.2)
DGC	77.8±1.4	66.1±4.2	84.3±0.6	83.9±0.7	90.4±0.2	65.2±4.0	68.7±13.	37.2±3.7	29.2±1.2	25.2±2.1	62.5±0.4	58.2±0.2	60.7±0.1	6.6 (3.2)
S^2GC	78.3±1.5	66.9±4.4	84.3±0.3	83.1±0.8	90.1±0.8	62.0±7.4	58.3±18.	34.9±4.9	27.6±1.8	26.7±1.8	63.1±0.5	58.7±0.1	61.2±0.0	6.6 (2.7)
G^2CN	76.6±1.5	64.2±3.3	81.4±0.6	82.8±1.6	88.8±0.5	O.O.M.	O.O.M.	40.7±2.9	32.1±1.5	24.3±0.5	O.O.M.	O.O.M.	O.O.M.	10.5 (4.5)
GCN	76.0±1.2	65.0±2.9	84.3±0.5	85.1±0.9	91.6±0.5	62.8±0.6	O.O.M.	38.5±3.0	31.4±1.8	26.8±0.4	62.9±0.7	57.0±0.1	63.9±0.4	6.3 (2.4)
SAGE	74.6±1.3	63.7±3.6	82.9±0.4	83.8±0.5	90.6±0.5	61.5±0.6	O.O.M.	39.8±4.3	27.0±1.3	27.8±0.9	O.O.M.	56.6±0.4	68.9±0.1	8.5 (3.5)
GCNII	77.8±1.7	63.4±3.0	84.9±0.8	82.3±1.8	90.8±0.6	45.7±0.5	O.O.M.	30.5±2.5	21.9±3.0	29.0±1.3	64.5±0.5	56.9±0.6	62.1±0.3	8.4 (4.6)
H^2GCN	77.6±0.9	64.7±3.8	85.4±0.4	49.5±16.	75.8±11.	O.O.M.	O.O.M.	31.9±2.6	25.0±0.5	28.9±0.6	63.9±0.4	58.7±0.0	O.O.M.	8.9 (4.9)
APPNP	80.0±0.6	67.1±2.8	84.6±0.5	84.2±1.7	92.5±0.3	53.4±1.3	O.O.M.	30.9±4.7	23.9±3.2	26.1±1.0	63.7±0.9	47.3±0.3	57.4±0.4	7.6 (4.8)
GPR-GNN	78.8±1.3	64.2±4.0	85.1±0.7	85.0±1.0	92.6±0.3	58.5±0.8	O.O.M.	31.7±4.7	26.2±1.6	29.5±1.1	64.5±0.4	57.6±0.2	67.6±0.1	5.4 (3.7)
GAT	78.2±1.2	65.8±4.0	83.6±0.2	85.4±1.4	91.7±0.5	58.2±1.0	O.O.M.	39.1±4.1	28.6±0.6	26.4±0.4	60.5±0.8	O.O.M.	O.O.M.	7.5 (3.7)
SlimG	77.8±1.1	67.1±2.3	84.6±0.5	86.3±0.7	91.8±0.5	66.3±0.3	84.9±0.0	40.8±3.2	31.1±0.7	30.9±0.6	68.2±0.6	59.7±0.1	73.9±0.1	1.9 (1.5)

$$\underbrace{\mathbf{U}}_{\text{Structure}} \;\|\; \underbrace{g(\mathbf{X})}_{\text{Node features}} \;\|\; \underbrace{g(\mathbf{A}^2_{\text{row}}\mathbf{X})}_{\text{2-step neighbors}} \;\|\; \underbrace{g(\tilde{\mathbf{A}}^2_{\text{sym}}\mathbf{X})}_{\text{Neighbors}}$$