CLUSTERING SIGNED NETWORKS WITH THE GEOMETRIC MEAN OF LAPLACIANS

UNIVERSITÄT DES SAARLANDES

PEDRO MERCADO[1], FRANCESCO TUDISCO[2], MATTHIAS HEIN[1]

[1]SAARLAND UNIVERSITY [2]UNIVERSITY OF PADUA

INTRODUCTION

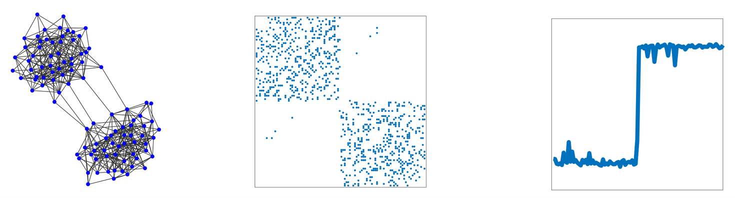
GOAL: Extend Spectral Clustering to networks that have positive and negative relationships, by defining a new Laplacian that blends both sources of information.

CONTRIBUTIONS:
1. We introduce the *geometric mean of Laplacians* as an alternative way to blend information of positive and negative relations.
2. We show that in expectation under the Stochastic Block Model our method *outperforms* current approaches.
3. We show that eigenvectors of the matrix geometric mean can be computed efficiently *without ever computing the matrix itself*.

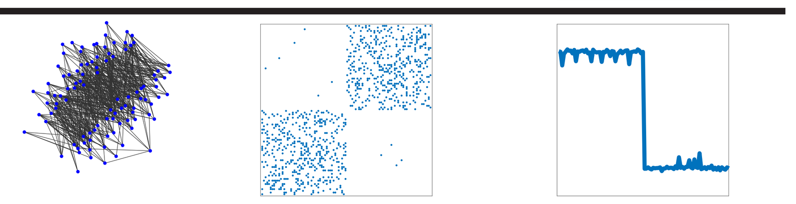
CLUSTERING AS GRAPH PARTITIONING

1. Get eigenvectors $\{\mathbf{u}_i\}_{i=1}^k$ corresponding to the k **smallest** eigenvalues of L.
2. Let $U = (\mathbf{u}_1, \ldots, \mathbf{u}_k)$.
3. Cluster the rows of U with k-means into clusters $C_1, \ldots, C_k$.

Assortative Case
$$\mathbf{L} = \mathbf{D} - \mathbf{W}$$
$$\mathbf{L}_{\text{sym}} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$$

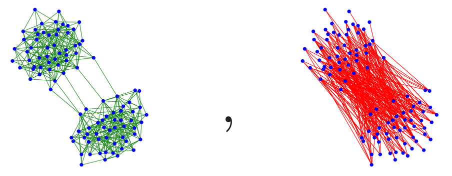


Disassortative Case
$$\mathbf{Q} = \mathbf{D} + \mathbf{W}$$
$$\mathbf{Q}_{\text{sym}} = \mathbf{D}^{-1/2}\mathbf{Q}\mathbf{D}^{-1/2}$$



CASE OF SIGNED NETWORKS:
A signed graph is the pair $G^{\pm} = (G^+, G^-)$ where G^+ and G^- encode positive and the negative relations, respectively.

$$G^{\pm} = \big(\quad , \quad \big)$$



$$\mathbf{L}_{\text{BR}} = \mathbf{D}^+ - \mathbf{W}^+ + \mathbf{W}^- \text{[1]}$$
$$= \mathbf{L}^+ + \mathbf{W}^-$$
$$\mathbf{L}_{\text{SR}} = \mathbf{D}^+ - \mathbf{W}^+ + \mathbf{D}^- + \mathbf{W}^- \text{[2]}$$
$$= \mathbf{L}^+ + \mathbf{Q}^-$$

Current signed Laplacians are some sort of arithmetic mean of Laplacians.

MATRIX GEOMETRIC MEAN

Let $\mathbf{A}, \mathbf{B}$ be p.d. matrices. The geometric mean of $\mathbf{A}$ and $\mathbf{B}$ is the matrix $\mathbf{A}\#\mathbf{B}$ defined by[3]
$$\mathbf{A}\#\mathbf{B} = \mathbf{A}^{1/2}(\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2})^{1/2}\mathbf{A}^{1/2}$$

OBSERVATION: Let $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$ and $\mathbf{B}\mathbf{u} = \mu\mathbf{u}$. Then,
$$(\mathbf{A} + \mathbf{B})\mathbf{u} = (\lambda + \mu)\mathbf{u}, \qquad (\mathbf{A}\#\mathbf{B})\mathbf{u} = (\sqrt{\lambda\mu})\mathbf{u}$$

Relative ordering of eigenvalues is different among matrix means.

GEOMETRIC MEAN LAPLACIAN

We define the normalized geometric mean Laplacian of $G^{\pm}$ as
$$\mathbf{L}_{\text{GM}} = \mathbf{L}^+_{\text{sym}}\#\mathbf{Q}^-_{\text{sym}}$$

We add a small diagonal shift to enforce Laplacians to be p.d.

STOCHASTIC BLOCK MODEL ANALYSIS

In the Stochastic Block Model (SBM), the edge W_{ij} exists with probability p if v_i and v_j are in the same cluster and q if they are in different clusters.

For signed networks we consider a SBM for each graph G^+ and G^-, with parameters $(p^+_{\text{in}}, p^+_{\text{out}})$ and $(p^-_{\text{in}}, p^-_{\text{out}})$ respectively, i.e.

$$P(W^+_{ij} = 1) = \begin{cases} p^+_{\text{in}} & \text{if } v_i, v_j \text{ are in the same cluster} \\ p^+_{\text{out}} & \text{if } v_i, v_j \text{ are in the different clusters} \end{cases}$$

$$P(W^-_{ij} = 1) = \begin{cases} p^-_{\text{in}} & \text{if } v_i, v_j \text{ are in the same cluster} \\ p^-_{\text{out}} & \text{if } v_i, v_j \text{ are in the different clusters} \end{cases}$$

COROLLARY (SBM IN EXPECTATION)

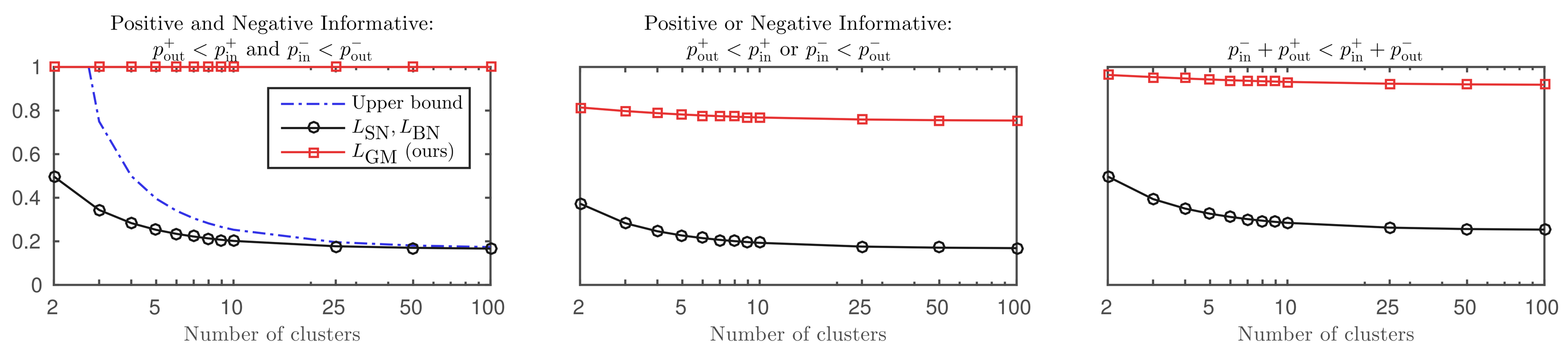
Let $\mathcal{D} = \{(p^+_{\text{in}}, p^+_{\text{out}}, p^-_{\text{in}}, p^-_{\text{out}}) \in [0,1]^4 \mid p^+_{\text{out}} < p^+_{\text{in}}, \text{ and } p^-_{\text{in}} < p^-_{\text{out}}\}$. Assume k clusters of same size. Let $\boldsymbol{\chi}_1 = \mathbf{1}$ and $\boldsymbol{\chi}_i = (k-1)\mathbf{1}_{\mathcal{C}_i} - \mathbf{1}_{\overline{\mathcal{C}_i}}$. Then, under condition $\mathcal{D}$,

- $\{\boldsymbol{\chi}_i\}_{i=1}^k$ **always** correspond to the k smallest eigenvalues of $\mathbb{E}[\mathbf{L}_{\text{GM}}]$.
- $\{\boldsymbol{\chi}_i\}_{i=1}^k$ correspond to the k smallest eigenvalues of $\mathbb{E}[\mathbf{L}_{\text{SN}}]$ in **at most** $\frac{1}{6} + \frac{2}{3(k-1)} + \frac{1}{(k-1)^2}$ proportion of cases.

Under condition $\mathcal{D}$, our method always identifies the informative eigenvectors, whereas the state of the art doesn't.

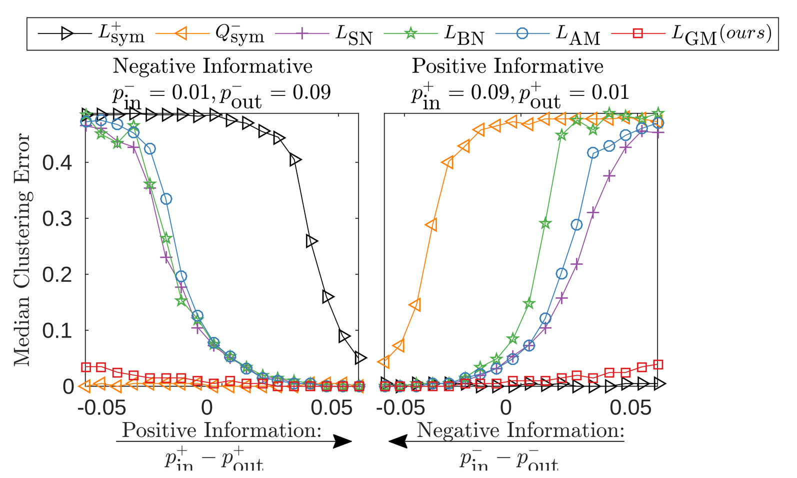
A more general result can be found in our paper.

We count the number of settings $(p^+_{\text{in}}, p^+_{\text{out}}, p^-_{\text{in}}, p^-_{\text{out}})$ where in expectation $\{\boldsymbol{\chi}_i\}_{i=1}^k$ correspond to the k smallest eigenvalues.



Our method identifies the informative eigenvectors in at least 75% of the cases, whereas the state of the art does it in at most 50% of the cases.

SBM IN SPARSE GRAPHS



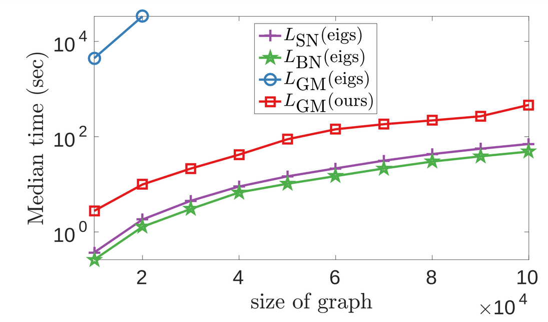
Setting: 50 runs on graphs with two clusters of 100 nodes.

- $\mathbf{L}_{\text{BN}}, \mathbf{L}_{\text{SN}}$ fail when either G^+ or G^- are informative about the cluster structure.
- Our method works when either G^+ or G^- are informative.

Our method works when either G^+ or G^- are informative, whereas the state of the art requires both G^+ and G^- to be informative.

POWER METHOD FOR $\mathbf{L}^+_{\text{SYM}}\#\mathbf{Q}^-_{\text{SYM}}$

We need the eigenvectors of $\mathbf{L}^+_{\text{sym}}\#\mathbf{Q}^-_{\text{sym}}$.



- Computation of $\mathbf{A}\#\mathbf{B}$ is expensive.
- $\mathbf{A}\#\mathbf{B}$ is in general a dense matrix, even if $\mathbf{A}$ and $\mathbf{B}$ are sparse.

We compute eigenvectors of $\mathbf{A}\#\mathbf{B}$ without ever computing $\mathbf{A}\#\mathbf{B}$.

Inverse Power Method (IPM) for $\mathbf{A}\#\mathbf{B}$:
Alternative formulation[4]: $\mathbf{A}\#\mathbf{B} = \mathbf{A}(\mathbf{A}^{-1}\mathbf{B})^{1/2}$

$$(\mathbf{A}\#\mathbf{B})x_{k+1} = x_k \iff \mathbf{A}u_k = x_k \text{ and } (\mathbf{A}^{-1}\mathbf{B})^{1/2}v_k = u_k$$

For $(\mathbf{A}^{-1}\mathbf{B})^{1/2}x = y$ we use the Extended Krylov Subspace Method (EKSM)[5]:
Let $M = A^{-1}B$ and $f(X) = X^{-1/2}$. Project M onto the subspace

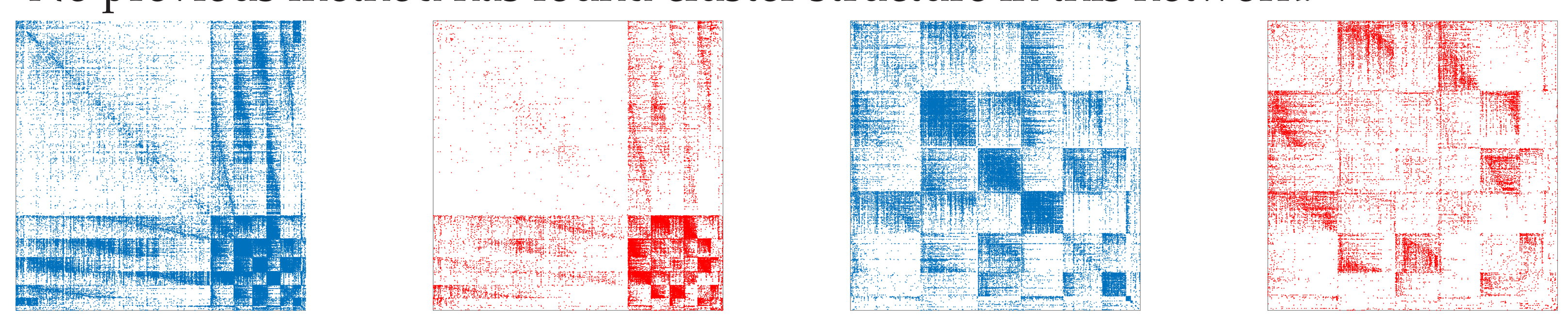
$$\mathbb{K}^s(M, \mathbf{y}) = \text{span}\{\mathbf{y}, M\mathbf{y}, M^{-1}\mathbf{y}, \ldots, M^{s-1}\mathbf{y}, M^{1-s}\mathbf{y}\},$$

At each step we have $MV_s = V_sH_s + [\mathbf{u}_{s+1}, \mathbf{v}_{s+1}][\mathbf{e}_{2s+1}, \mathbf{e}_{2s+2}]^T$, where,
- H_s is $2s \times 2s$ symmetric tridiagonal,
- $\mathbf{u}_{s+1}$ and $\mathbf{v}_{s+1}$ are orthogonal to V_s

Solution $\mathbf{x}$ is approximated by $\mathbf{x}_s = V_sH_s^{-1/2}\mathbf{e}_1\|\mathbf{y}\| \approx M^{-1/2}\mathbf{y}$

EXPERIMENTS IN WIKIPEDIA

- We look for 30 clusters with $\mathbf{L}_{\text{GM}}$.
- We present sorted adjacency matrices based on clusters obtained with $\mathbf{L}_{\text{GM}}$.
- No previous method has found cluster structure in this network.



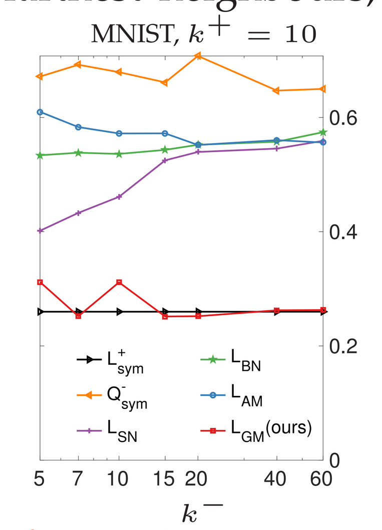
(W^+) (W^-) $(W^+$ (Zoom)$)$ $(W^-$ (Zoom)$)$

Our method is the first one finding cluster structure in this network.

EXPERIMENTS ON UCI DATASETS

We build W^+ and W^- through k^+-nearest neighbours and k^--farthest neighbours, respectively, with $k^+, k^- \in \{5, 7, 10, 15, 20, 40, 60\}$.

		iris	wine	ecoli	optdig	USPS	pendig	MNIST
# vertices		150	178	310	5620	9298	10992	70000
# classes		3	3	3	10	10	10	10
$\mathbf{L}_{\text{SN}}$	Best (%)	23.4	40.6	18.8	28.1	10.9	10.9	12.5
	Str. best (%)	10.9	21.9	14.1	28.1	9.4	10.9	12.5
$\mathbf{L}_{\text{BN}}$	Best (%)	17.2	21.9	7.8	0.0	1.6	3.1	0.0
	Str. best (%)	7.8	4.7	6.3	0.0	1.6	3.1	0.0
$\mathbf{L}_{\text{AM}}$	Best (%)	12.5	28.1	14.1	0.0	0.0	1.6	0.0
	Str. best (%)	10.9	14.1	12.5	0.0	0.0	1.6	0.0
L_{GM}	Best (%)	**59.4**	**42.2**	**65.6**	**71.9**	**89.1**	**84.4**	**87.5**
	Str. best (%)	**57.8**	**35.9**	**60.9**	**71.9**	**87.5**	**84.4**	**87.5**



Our method is robust against misleading sources of information.

REFERENCES

[1] K. Chiang, J. Whang, and I. Dhillon. Scalable clustering of signed networks using balance normalized cut. CIKM, 2012.
[2] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. Luca, and S. Albayrak. Spectral analysis of signed graphs for clustering, prediction and visualization. In *ICDM*, 2010.
[3] Rajendra Bhatia. *Positive definite matrices*. Princeton University Press, 2009.
[4] Bruno Iannazzo. The geometric mean of two matrices from a computational viewpoint. *Numerical Linear Algebra with Applications*, 2016.
[5] V. Druskin and L. Knizhnerman. Extended Krylov subspaces: approximation of the matrix square root and related functions. *SIAM J. Matrix Anal. Appl.*, 1998.