Spectral Clustering of Signed Graphs via Matrix Power Means

Pedro Mercado[1,2], Francesco Tudisco[3], Matthias Hein[1]

[1]University of Tübingen [2]Saarland University [3]University of Strathclyde

Introduction

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

MAIN CONTRIBUTIONS:

1. We introduce the *Signed Power Mean Laplacian* as an alternative way to blend the information of different kinds of relations.
2. We show that in expectation under the Stochastic Block Model our method *outperforms* current approaches.
3. We show concentration bounds for eigenvalues and eigenvectors of the *Signed Power Mean Laplacian*

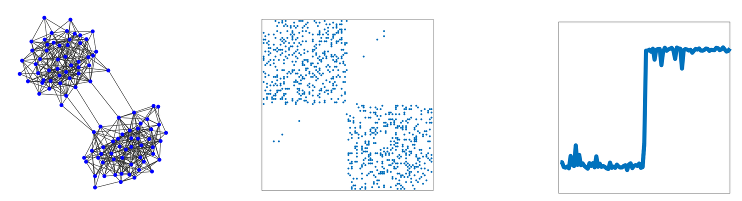
Clustering as Graph Partitioning

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

Assortative Case

$$\mathbf{L} = \mathbf{D} - \mathbf{W}$$

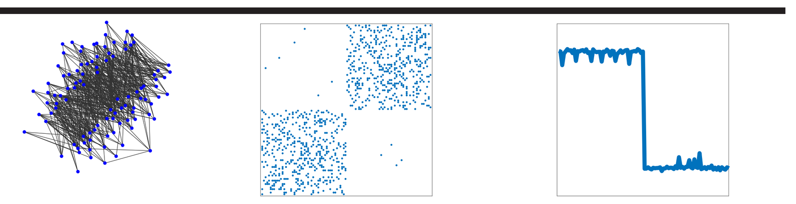
$$\mathbf{L}_{\text{sym}} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$$



Disassortative Case

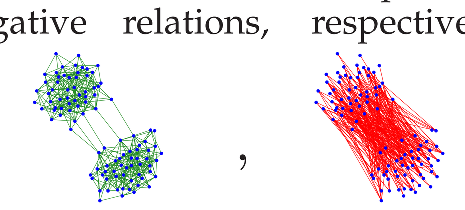
$$\mathbf{Q} = \mathbf{D} + \mathbf{W}$$

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



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.

$G^\pm = \Bigg($  $\Bigg)$

$$\mathbf{L_{BR}} = \mathbf{D}^+ - \mathbf{W}^+ + \mathbf{W}^- \text{ [1]}$$
$$= \mathbf{L}^+ + \mathbf{W}^-$$
$$\mathbf{L_{SR}} = \mathbf{D}^+ - \mathbf{W}^+ + \mathbf{D}^- + \mathbf{W}^- \text{ [2]}$$
$$= \mathbf{L}^+ + \mathbf{Q}^-$$
$$\mathbf{H} = \beta I + \alpha(\mathbf{W}^+ - \mathbf{W}^-) + \mathbf{D}^+ + \mathbf{D}^- \text{ [3]}$$

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

Scalar Power Mean

The scalar power mean is defined as $\boxed{m_p(x_1, \dots, x_T) = (\frac{1}{T}\sum_{i=1}^T x_i^p)^{1/p}}$

	$p \to -\infty$	$p = -1$	$p \to 0$	$p = 1$	$p \to \infty$
$m_p(x_1, \dots, x_T)$	$\min\{x_1, \dots, x_T\}$	$T\,(\sum_{i=1}^T x_i^{-1})^{-1}$	$(\prod_{i=1}^T x_i)^{1/T}$	$\frac{1}{T}\sum_{i=1}^T x_i$	$\max\{x_1, \dots, x_T\}$
name	minimum	harmonic mean	geometric mean	arithmetic mean	maximum

Matrix Power Mean

The matrix power mean of positive definite matrices $\mathbf{A}_1, \dots, \mathbf{A}_T$ is defined by [4]

$$M_p(\mathbf{A}_1, \dots, \mathbf{A}_T) = \left(\frac{1}{T}\sum_{i=1}^T \mathbf{A}_i^p\right)^{1/p}$$

OBSERVATION: Let $\mathbf{A}_t\mathbf{u} = \lambda_t\mathbf{u}$, for $t = 1, \dots, T$. Then,

$$M_p(\mathbf{A}_1, \dots, \mathbf{A}_T)\mathbf{u} = m_p(\lambda_1, \dots, \lambda_T)\mathbf{u}$$

Relative ordering of eigenvalues is different among matrix means.

Signed Power Mean Laplacian

We define the Signed Power Mean Laplacian of $G^\pm$ as

$$\mathbf{L_p} = M_p(\mathbf{L}^+_{\text{sym}}, \mathbf{Q}^-_{\text{sym}})$$

For $p \leq 0$ 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.

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

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

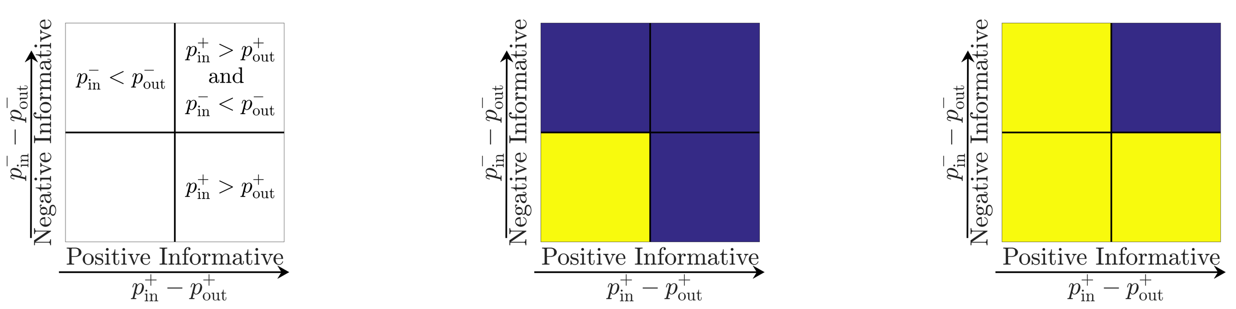
THEOREM: Let $\boldsymbol{\chi}_1 = \mathbf{1}$ and $\boldsymbol{\chi}_i = (k-1)\mathbf{1}_{\mathcal{C}_i} - \mathbf{1}_{\overline{\mathcal{C}_i}}$, $\tilde{k} = 2$ if $p \geq 1$ and $\tilde{k} = 1$ if $p < 1$. Let $\rho^+ = 1 - (p^+_{\text{in}} - p^+_{\text{out}})/(p^+_{\text{in}} + (k-1)p^+_{\text{out}})$ and $\rho^- = 1 + (p^-_{\text{in}} - p^-_{\text{out}})/(p^-_{\text{in}} + (k-1)p^-_{\text{out}})$. Then, $\{\boldsymbol{\chi}_i\}_{i=\tilde{k}}^k$ correspond to the smallest eigenvalues of $\mathcal{L}_\mathbf{p}$ if and only if $m_p(\rho^+, \rho^-) < 1$. In particular, for $p \to \pm\infty$,

- $\{\boldsymbol{\chi}_i\}_{i=1}^k$ are the k-smallest eigenvectors of $\mathcal{L}_\infty$ if and only if $p^+_{\text{in}} > p^+_{\text{out}}$ **AND** $p^-_{\text{in}} < p^-_{\text{out}}$
- $\{\boldsymbol{\chi}_i\}_{i=2}^k$ are the $(k$-$1)$-smallest eigenvectors of $\mathcal{L}_{-\infty}$ if and only if $p^+_{\text{in}} > p^+_{\text{out}}$ **OR** $p^-_{\text{in}} < p^-_{\text{out}}$

$\mathcal{L}_\infty$ and $\mathcal{L}_{-\infty}$ are related to the logical operators AND and OR.

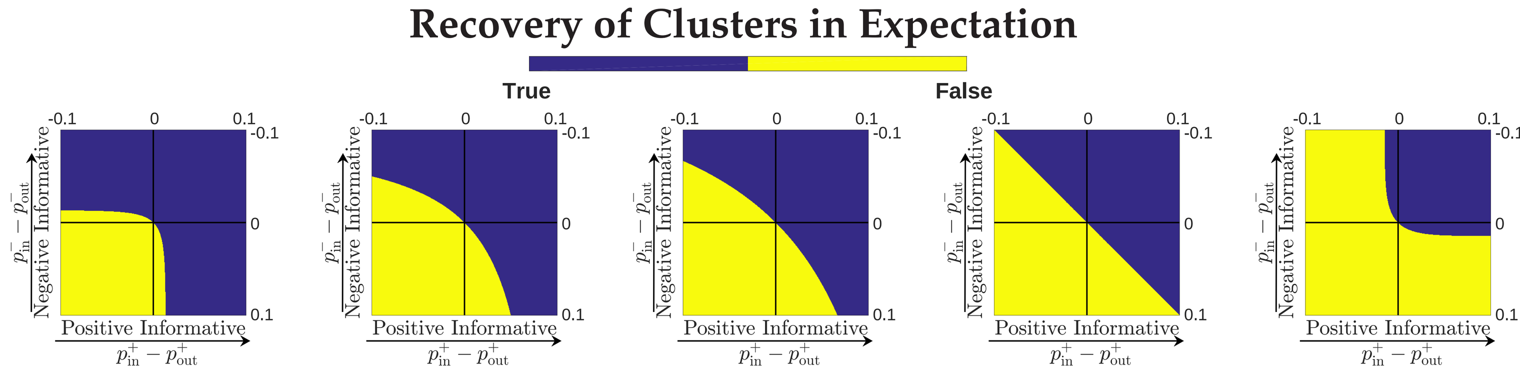
Analysis on expected graphs following the SBM. Setting:

- From every setting we show if conditions of our theorem holds.
- Fixed average degree: $p^+_{\text{in}} + p^+_{\text{out}} = 0.1$ and $p^-_{\text{in}} + p^-_{\text{out}} = 0.1$



SBM Diagram ($\mathcal{L}_{-\infty}$ (OR)) ($\mathcal{L}_\infty$ (AND))

Recovery of Clusters in Expectation

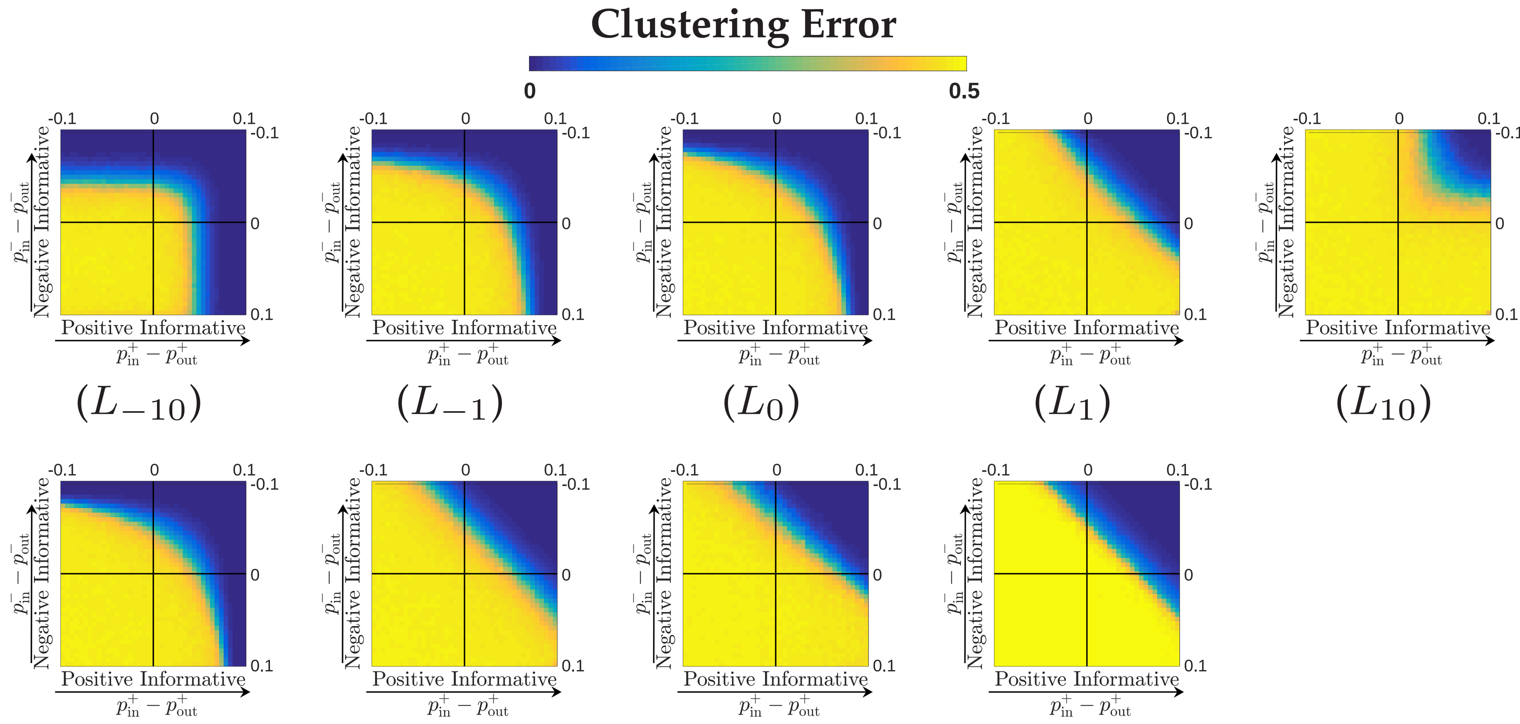


$(\mathcal{L}_{-10})$ $(\mathcal{L}_{-1})$ $(\mathcal{L}_0)$ $(\mathcal{L}_1)$ $(\mathcal{L}_{10})$

Analysis on random graphs following the SBM. Setting:

- From every setting we show the average clustering error out of 50 samples.
- Fixed average degree: $p^+_{\text{in}} + p^+_{\text{out}} = 0.1$ and $p^-_{\text{in}} + p^-_{\text{out}} = 0.1$

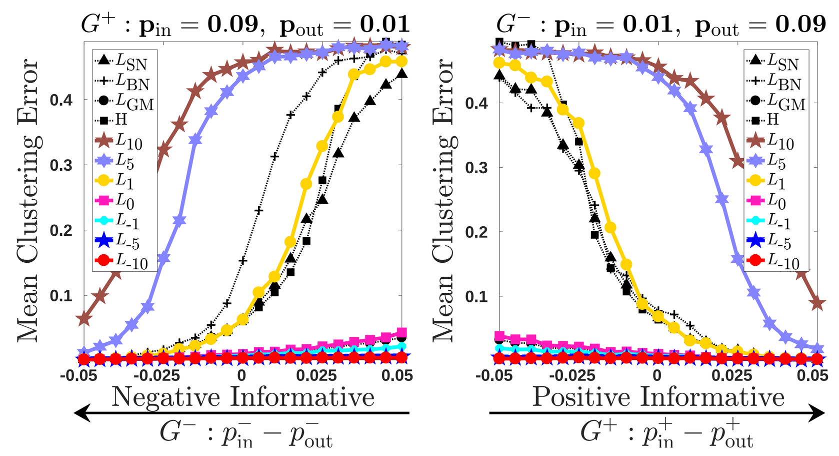
Clustering Error



(L_{-10}) (L_{-1}) (L_0) (L_1) (L_{10})

(L_{GM}) (L_{SN}) (L_{BN}) (H)

The smaller the value of p, the larger the coverage of $\mathcal{L}_\mathbf{p}$ under the SBM.

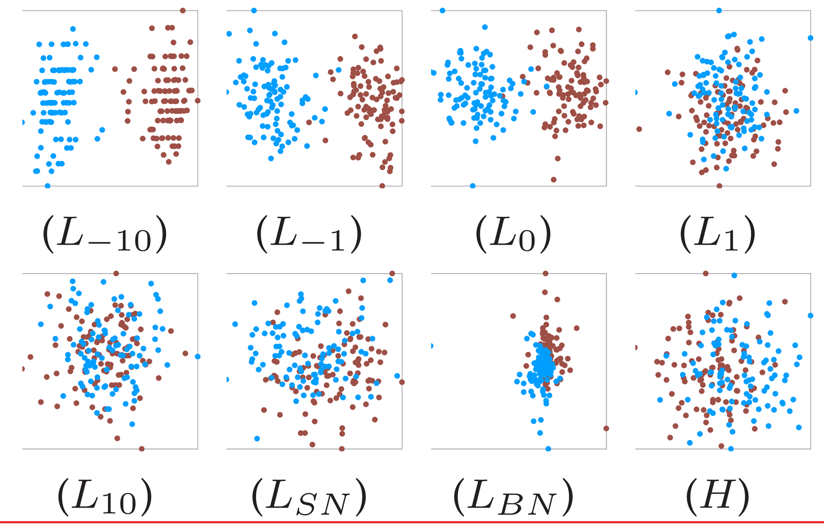


Setting: 50 runs on graphs with two clusters of 100 nodes.

- The larger the value of p, the larger the clustering error of $\mathbf{L_p}$.
- The smaller the value of p, the more robust is $\mathbf{L_p}$ against noise.
- $\mathbf{L_{-10}}$ presents the smallest clustering error.

The smaller the value of p, the better the performance of $\mathbf{L_p}$ under the SBM

SBM: Node Embeddings



(L_{-10}) (L_{-1}) (L_0) (L_1)

(L_{10}) (L_{SN}) (L_{BN}) (H)

$G^+ : p^+_{\text{in}} - p^+_{\text{out}} = $ -0.05 (not-informative)

$G^- : p^-_{\text{in}} - p^-_{\text{out}} = 0.08$ (informative)

- The larger the value of p, the better embedding of $\mathbf{L_p}$.
- The smaller the value of p, the more robust is the embedding of $\mathbf{L_p}$ against noise.
- $\mathbf{L_{-10}}$ presents the best node embedding.

The smaller the value of p, the better the embedding of $\mathbf{L_p}$ under the SBM.

SBM: Concentration Bounds

THEOREM: Let p be a non-zero integer, and

$$C_p = \begin{cases} (2p)^{1/p}(2+\varepsilon)^{1-1/p} & p \geq 1 \\ 2p^{1/p}\varepsilon^{-(3+1/p)} & p \leq -1 \end{cases}$$

Let $\epsilon > 0$, $\delta^+ := \frac{n}{k}(p^+_{\text{in}} + (k-1)p^+_{\text{out}})$ and $\delta^- := \frac{n}{k}(p^-_{\text{in}} + (k-1)p^-_{\text{out}})$.
If $\delta^+ > 3\ln(8n/\epsilon)$ and $\delta^- > 3\ln(8n/\epsilon)$, then with probability at least $1 - \epsilon$,

$$\|L_p - \mathcal{L}_p\| \leq C_p\, m_p^{1/p}\left(\sqrt{\frac{3\ln(8n/\epsilon)}{\delta^+}}, \sqrt{\frac{3\ln(8n/\epsilon)}{\delta^-}}\right)$$

THEOREM: Let $V_k, \mathcal{V}_k \in \mathbb{R}^{n\times k}$ be orthonormal matrices whose columns are the eigenvectors of the k smallest eigenvalues of L_p and $\mathcal{L}_p$, respectively. Let $\tilde{k} = k-1$, if $p \geq 1$, and $\tilde{k} = k$, if $p \leq -1$. If $m_p(\rho^+, \rho^-) < 1$, $\delta^+ > 3\ln(8n/\epsilon)$, and $\delta^- > 3\ln(8n/\epsilon)$, then there exists an orthogonal matrix $O_{\tilde{k}} \in \mathbb{R}^{\tilde{k}\times\tilde{k}}$ such that, with probability at least $1 - \epsilon$,

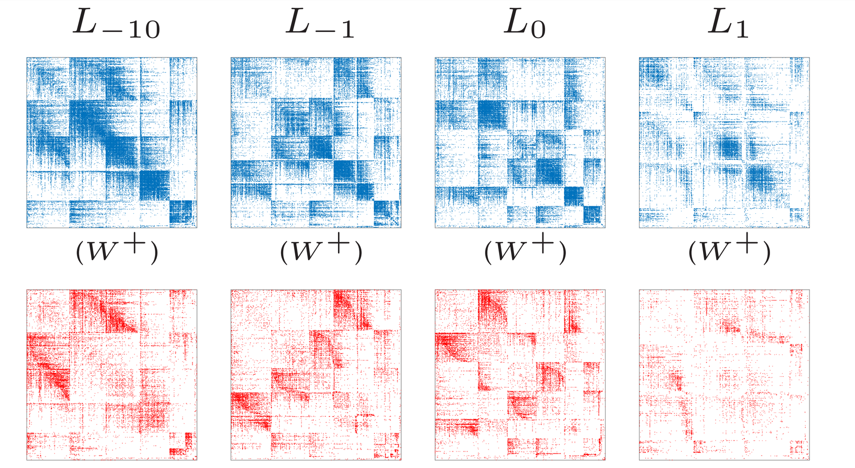
$$\|V_{\tilde{k}} - \mathcal{V}_{\tilde{k}}O_{\tilde{k}}\| \leq \frac{\sqrt{8\tilde{k}}C_p\, m_p^{1/p}\left(\sqrt{\frac{3\ln(8n/\epsilon)}{\delta^+}}, \sqrt{\frac{3\ln(8n/\epsilon)}{\delta^-}}\right)}{1 - m_p(\rho^+, \rho^-)}$$

First concentration bounds for Matrix Power Means under SBM

Experiments on Wikipedia Dataset

- We look for 30 clusters with $\mathbf{L_p}$.
- We present sorted adjacency matrices based on clusters obtained with $\mathbf{L_p}$.
- Our method consistently finds cluster structure in this network.

L_{-10} L_{-1} L_0 L_1



(W^+) (W^+) (W^+) (W^+)

(W^-) (W^-) (W^-) (W^-)

Our method consistently finds cluster structure in this network with $p \leq 0$

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] A. Saade, M. Lelarge, F. Krzakala, and L. ZdeborovÃą. Spectral detection in the censored block model. In *2015 IEEE International Symposium on Information Theory (ISIT)*, pages 1184–1188, June 2015.

[4] K. V. Bhagwat and R. Subramanian. Inequalities between means of positive operators. *Mathematical Proceedings of the Cambridge Philosophical Society*, 83(3):393âĂŞ401, 1978.