PageRank Embedding
==
Spectral Embedding

Disha Shur

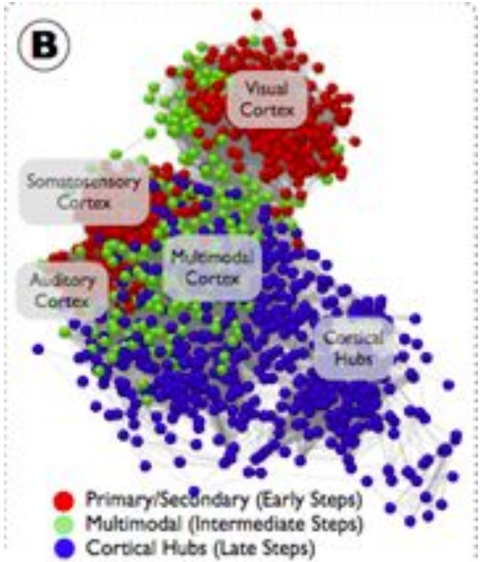
Purdue University

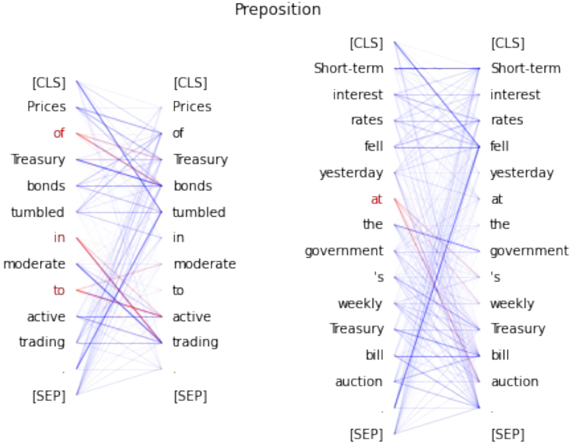
Thesis Defense - M.S.E.C.E Candidate

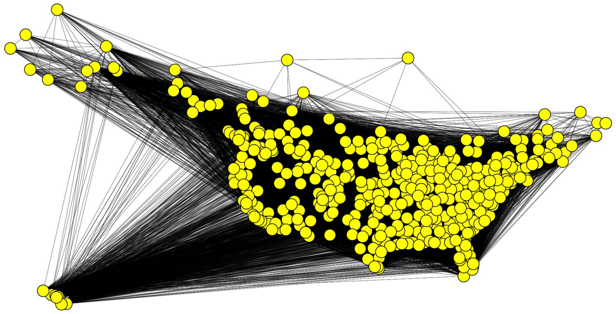
April 18, 2022

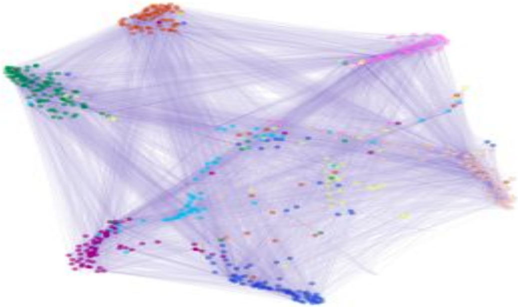
Low Dimensional Embeddings

Low dimensional representations are useful for learning and visualization

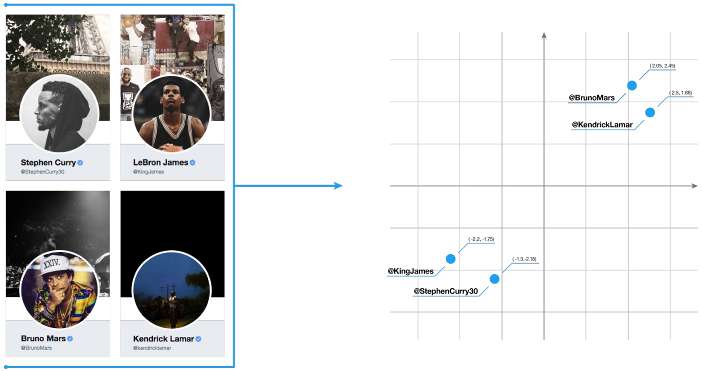






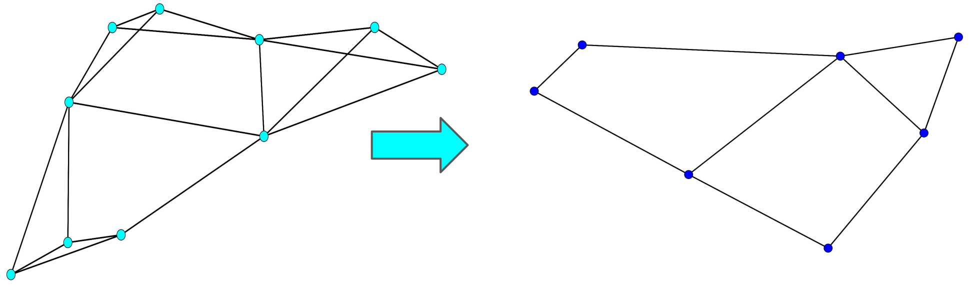


Two dimensional “user” embedding



Spectral Embedding

<u>Input: G</u> <u>Output: Vertex coordinates R^d</u>

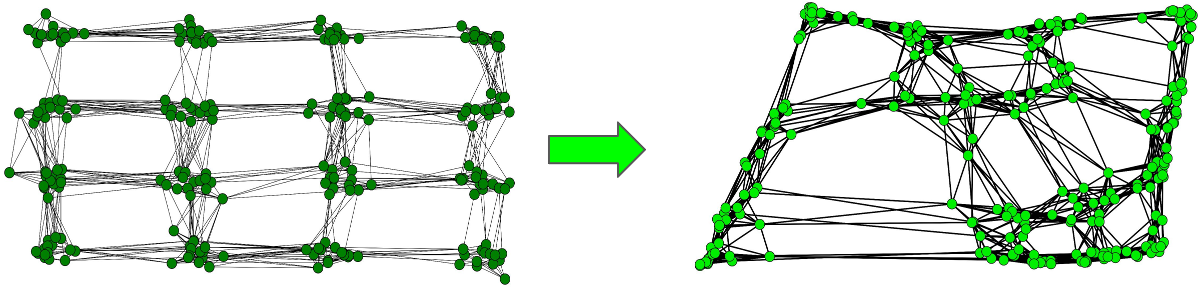


<u>Eigenvectors of the symmetric Laplacian</u>

$$\tilde{\mathbf{L}} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{G}\mathbf{D}^{-1/2} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$$

$$\mathbf{Z} = \mathbf{D}^{-1/2}\mathbf{Q}$$

$$i : (\mathbf{Z}[i,2], \mathbf{Z}[i,3])$$

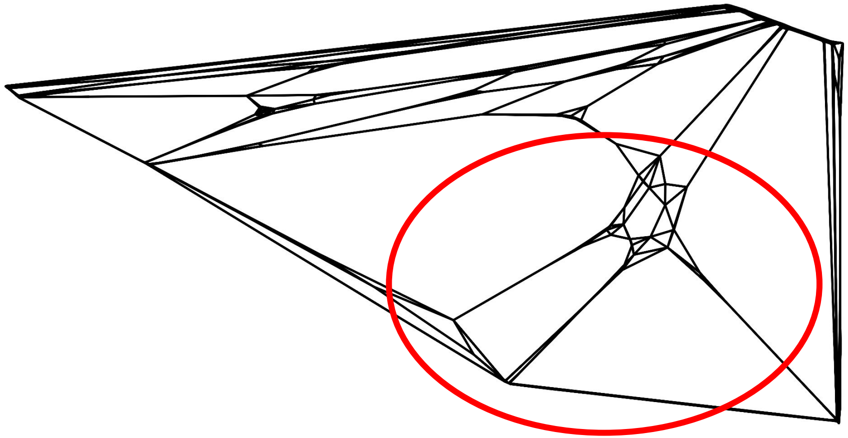


Spectral Embedding

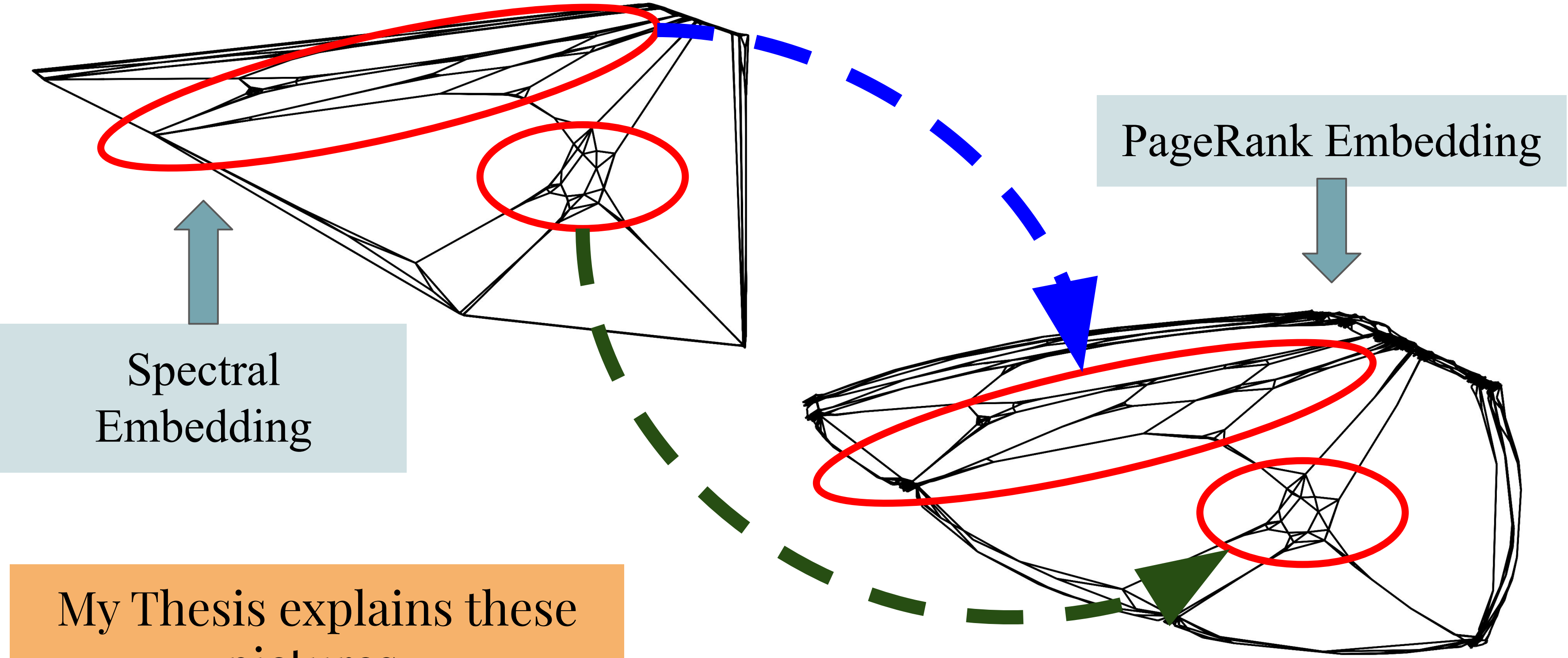
Original

$$\tilde{\mathbf{L}} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{G}\mathbf{D}^{-1/2} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^T$$
$$\mathbf{Z} = \mathbf{D}^{-1/2}\mathbf{Q}$$
$$i : (\mathbf{Z}[i,2], \mathbf{Z}[i,3])$$

Spectral Representation

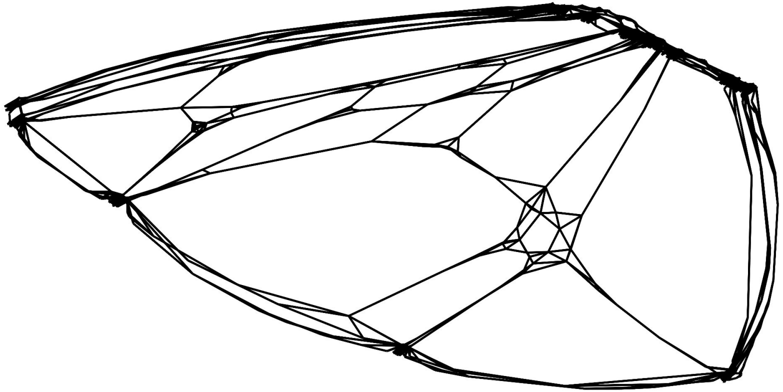


Similar Embeddings, Distinct Procedure



My Thesis explains these pictures.

PageRank Embeddings



$$\mathbf{P} = \mathbf{G}\mathbf{D}^{-1}$$

$$\boxed{(\mathbf{I} - \alpha\mathbf{P})\mathbf{x} = (1-\alpha)\mathbf{v} \quad \text{— (1)}}$$

$$[\alpha\mathbf{P} + (1-\alpha)\mathbf{v}\mathbf{e}^T]\mathbf{x} = \mathbf{x} \quad \text{— (2)}$$

$$\boxed{\mathbf{x} = (1-\alpha)(\mathbf{I} - \alpha\mathbf{P})^{-1}\mathbf{v}} \quad \text{— (3)}$$

$$\mathbf{v} = \mathbf{e}_m$$

$$\boxed{i : (\mathbf{U}[i,2], \mathbf{U}[i,3])}$$

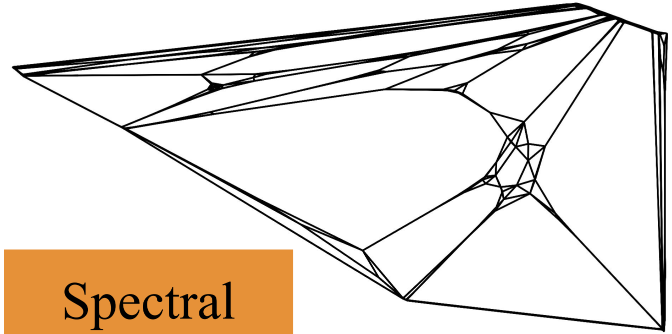
$$\boxed{\mathbf{x} = [(1-\alpha)(\mathbf{I} - \alpha\mathbf{P})^{-1}]_{(:,m)}} \quad \text{— (4)}$$

$$(\mathbf{I} - \alpha\mathbf{P})[\mathbf{x_1} \cdots \mathbf{x_k}] = (1-\alpha)[\mathbf{v_1} \cdots \mathbf{v_k}] \quad \text{— (5)} \quad \mathbf{k} << n$$

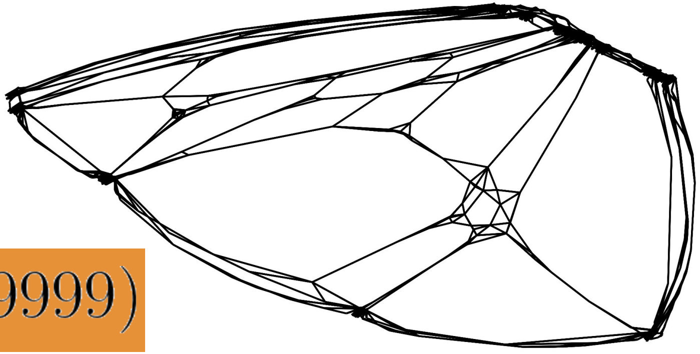
$$\mathbf{X} = [\mathbf{x_1} \cdots \mathbf{x_k}] = (1-\alpha)(\mathbf{I} - \alpha\mathbf{P})^{-1}[\mathbf{v_1} \cdots \mathbf{v_k}] \quad \text{— (6)}$$

$$\log.(\mathbf{X}) = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad \text{— (7)}$$

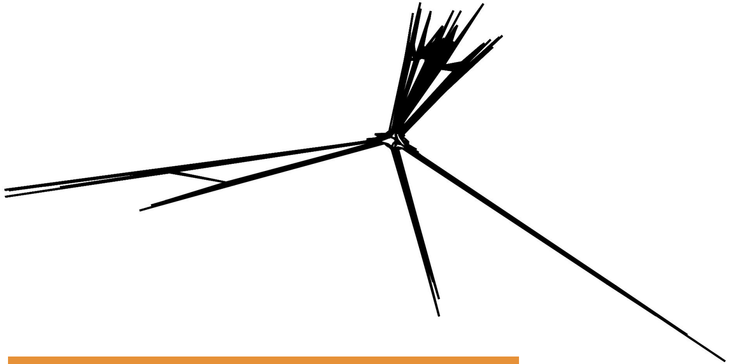
My Thesis!



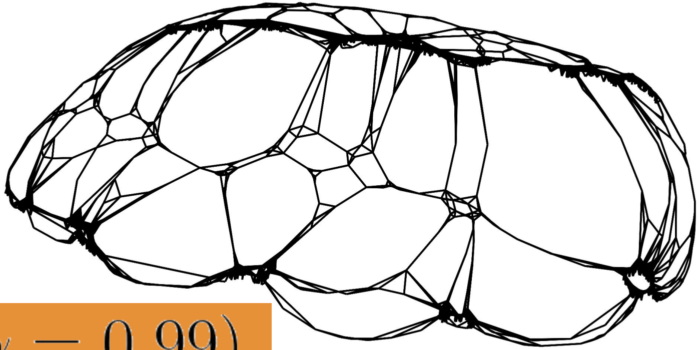
Spectral



$\log .\mathbf{X}(\alpha = 0.9999)$

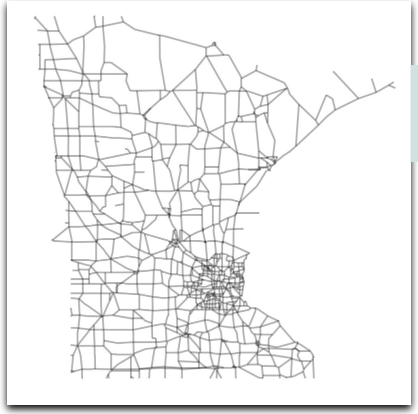


$\mathbf{X}(\alpha = 0.9999)$



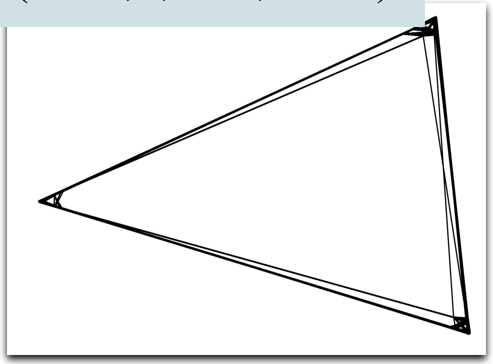
$\log .\mathbf{X}(\alpha = 0.99)$

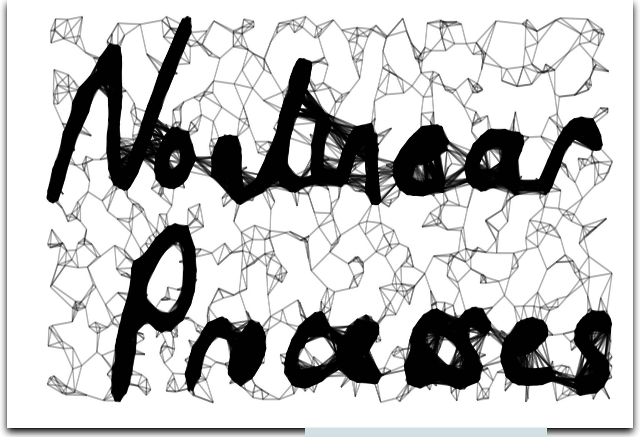
Generalizable?



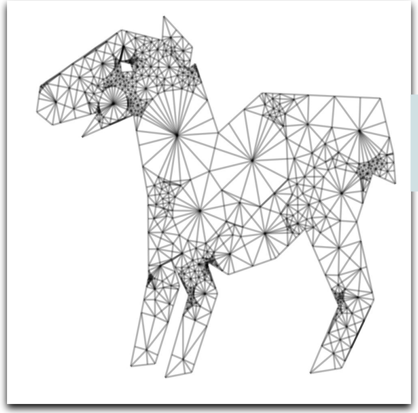
Minnesota

Sbm3000
(1000,3,0.25,0.001)

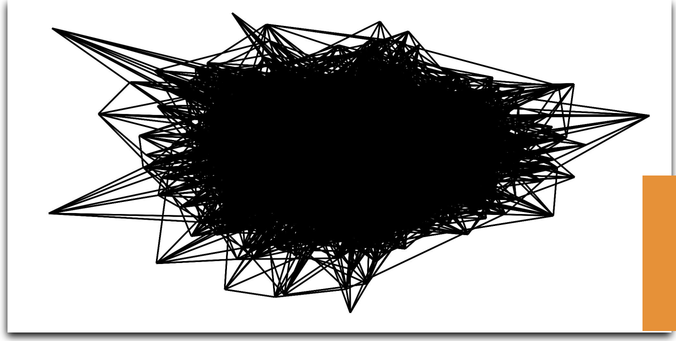




Original

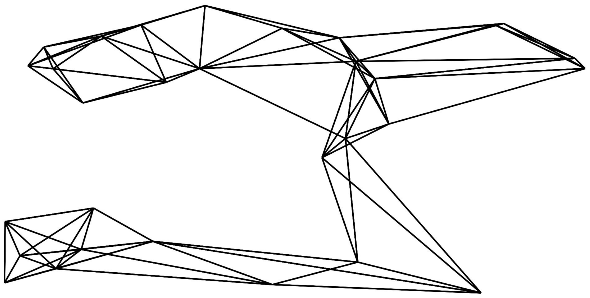


Tapir



Sbm3000
(50,60,0.001,0.005)

Generalizable?



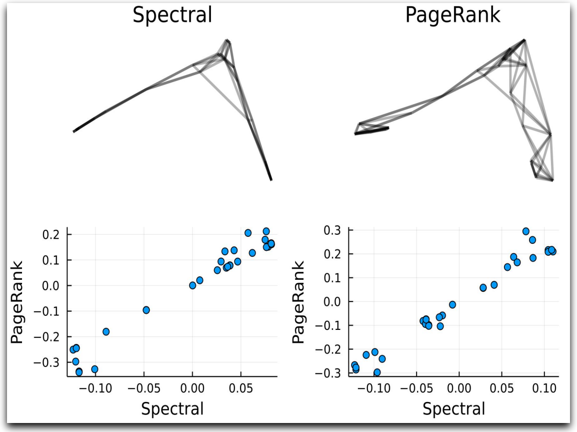
$$e = \frac{s - p}{s}$$

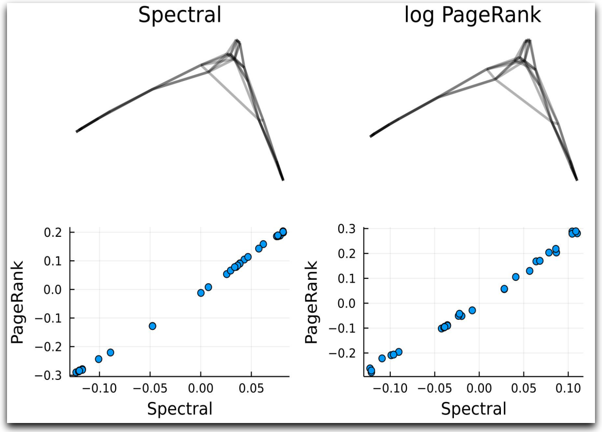
$$s = \frac{Z[:,2]^T \tilde{L} Z[:,2]}{Z[:,2]^T Z[:,2]}$$

$$p = \frac{U[:,2]^T \tilde{L} U[:,2]}{U[:,2]^T U[:,2]}$$

0.99

3.27%





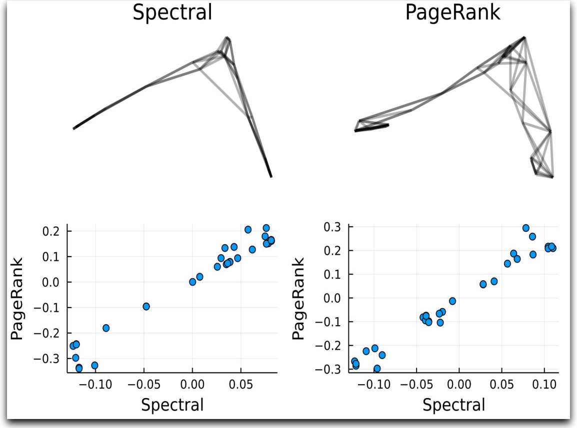
0.99-with log

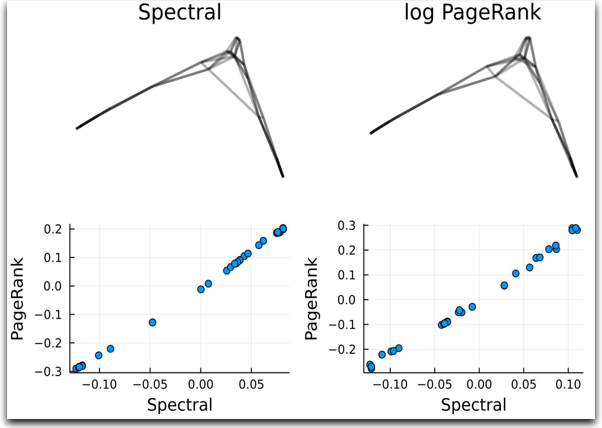
0.06%

Generalizable? - Small Nearest Neighbour graph

0.99

3.27%



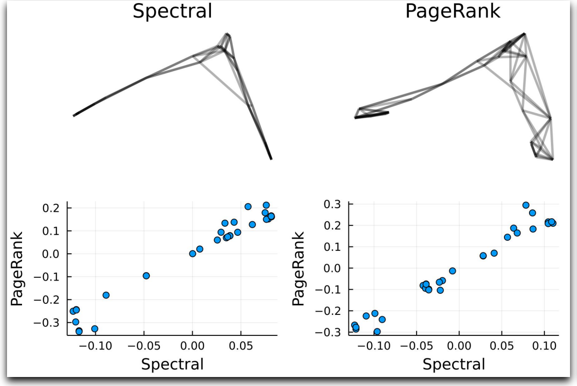


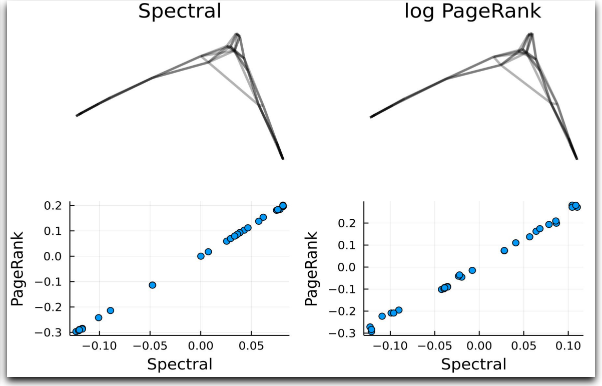
0.99-with log

0.06%

0.99999

2.89%





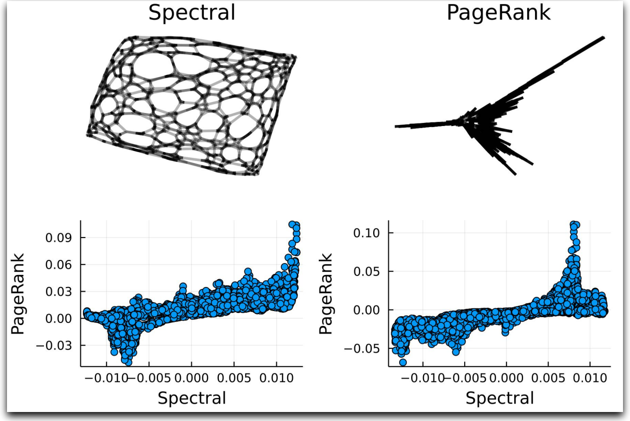
0.99999-with log

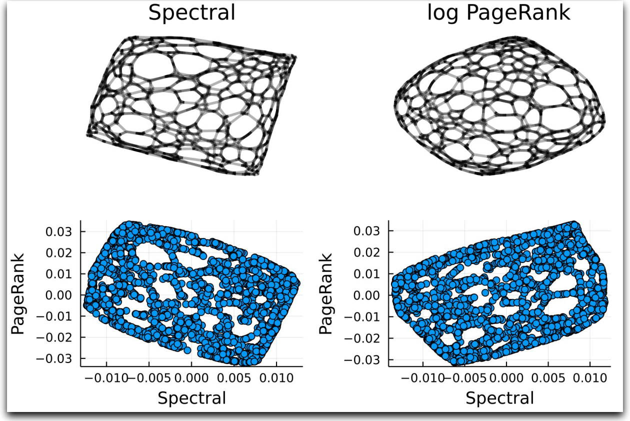
0.05%

Generalizable? - Large Nearest Neighbour Graph

0.99

47.6%



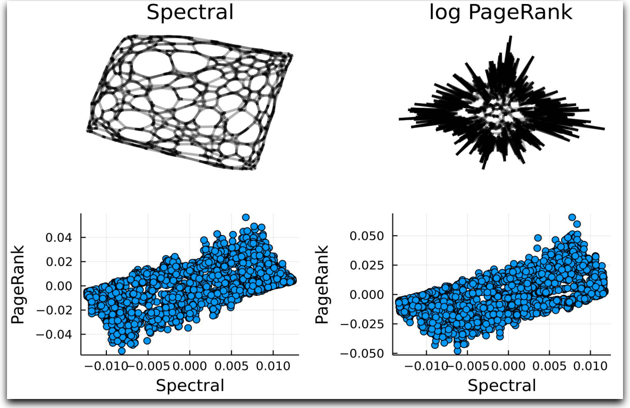


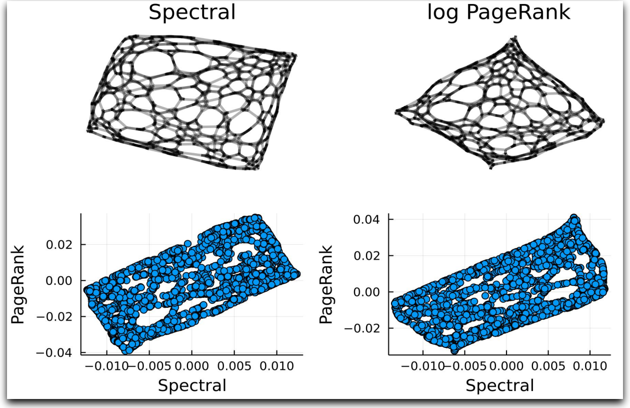
0.99-with log

0.37%

0.99999

5.06%



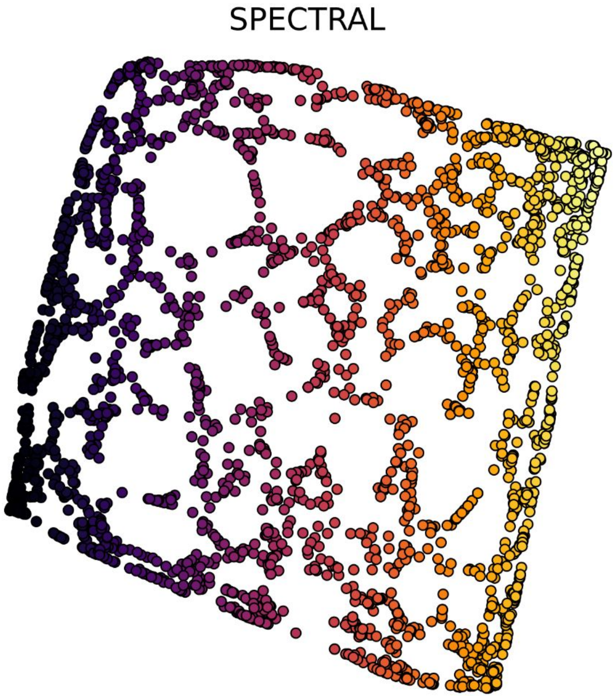


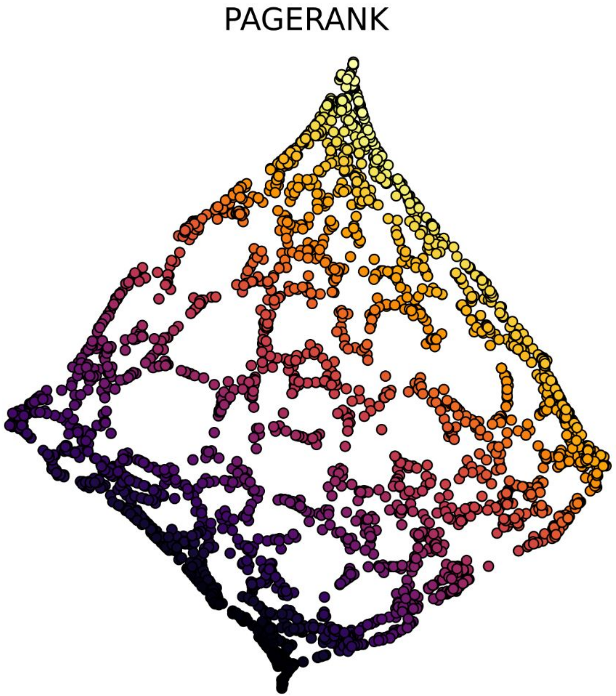
0.99999-with log

2.88%

Generalizable? - Large Nearest Neighbour Graph

$$\lambda = \begin{bmatrix} 0.0000 \\ 0.0005617 \\ 0.0005958 \\ 0.001243 \\ 0.002170 \end{bmatrix}$$

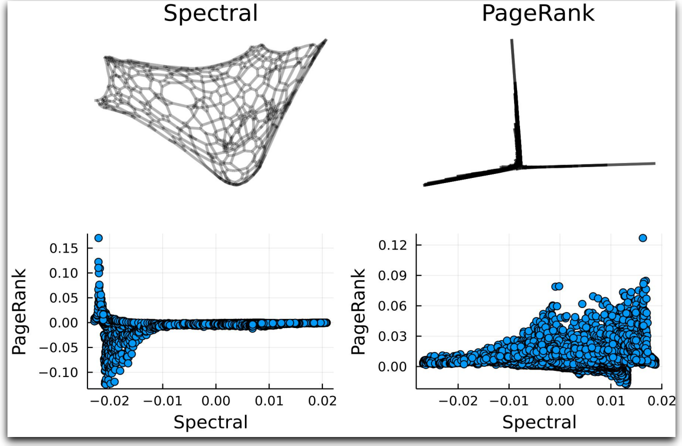


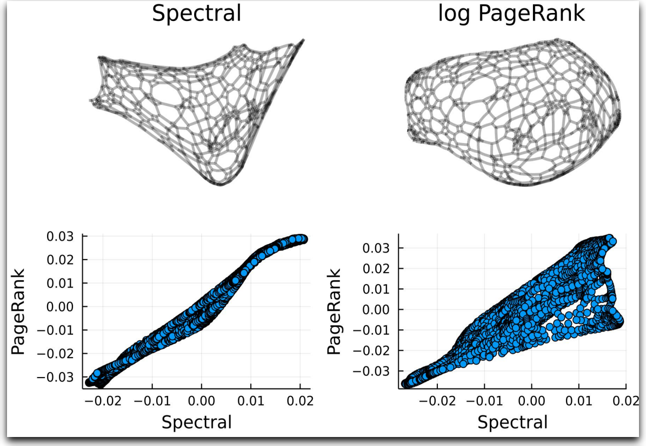


Generalizable? - Minnesota Network

0.99

16%



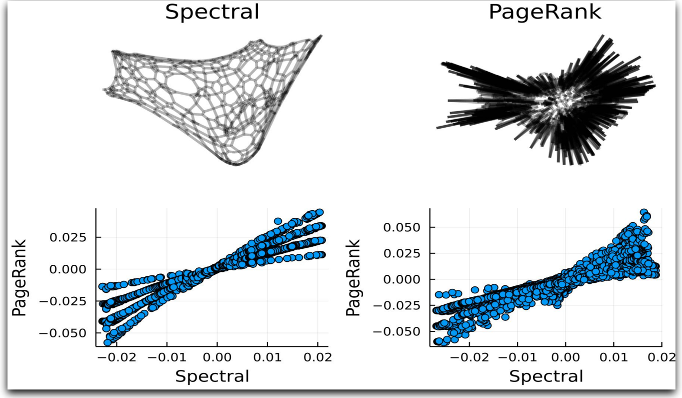


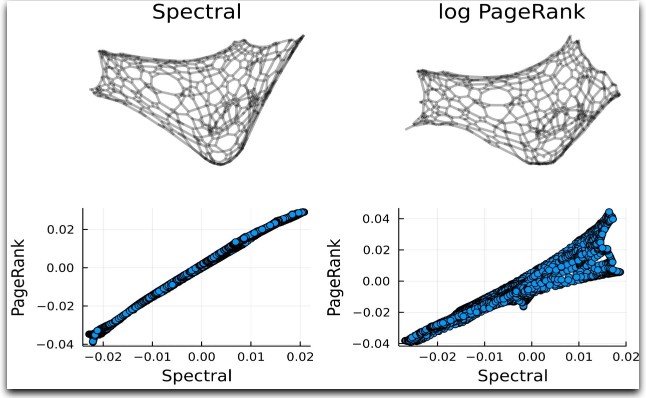
0.99-with log

1.97%

0.99999

11.15%





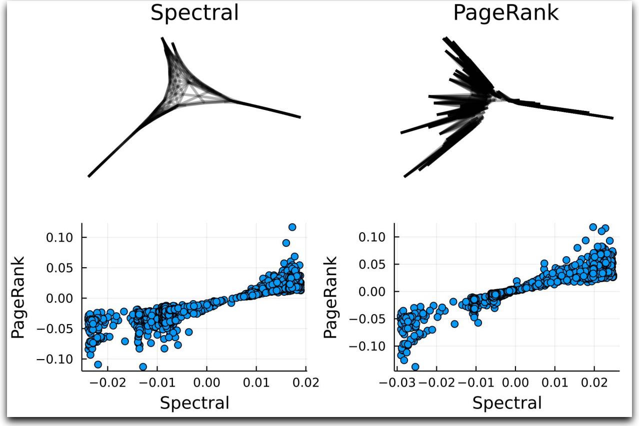
0.99999-with log

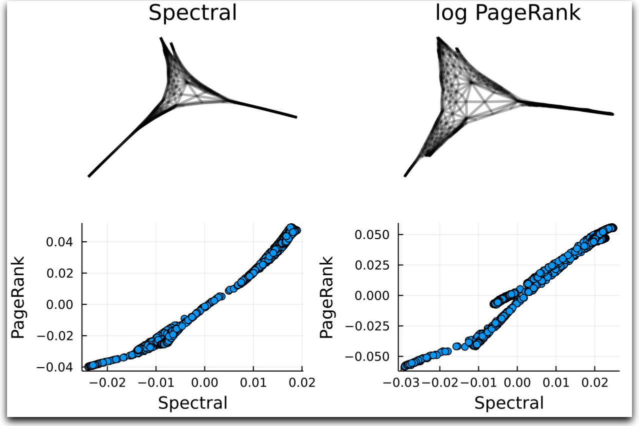
0.44%

Generalizable? - Tapir

0.99

10.2%



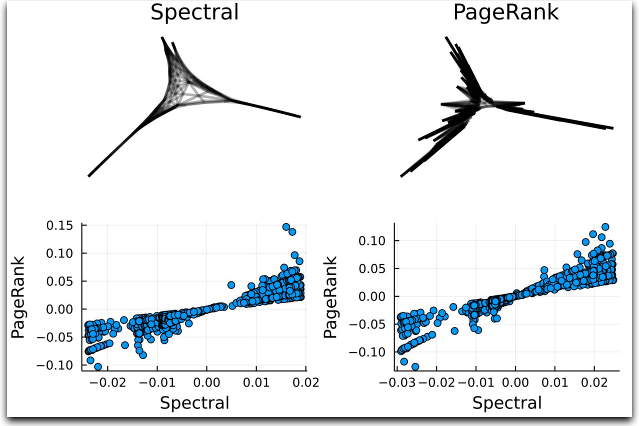


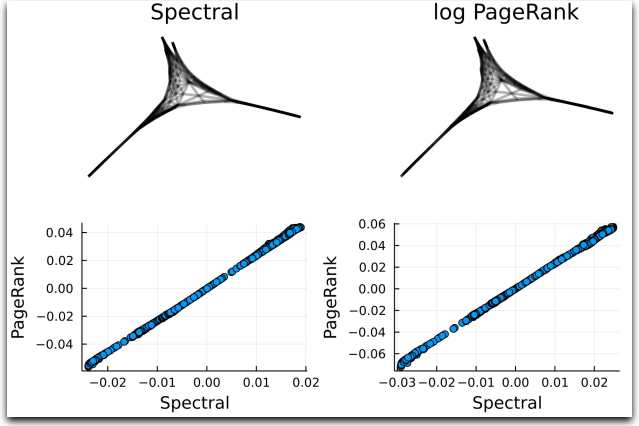
0.99-with log

1.13%

0.99999

15.1%



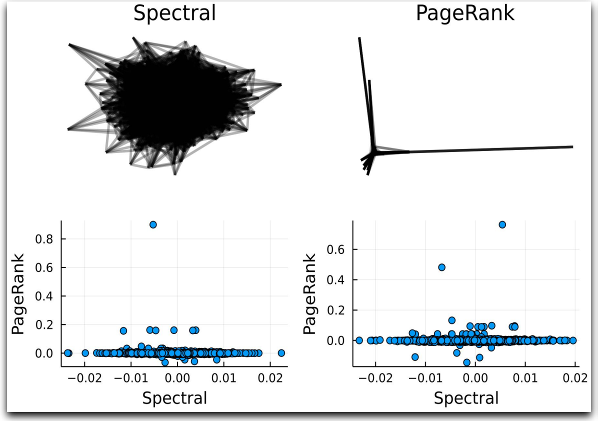


0.99999-with log

0.66%

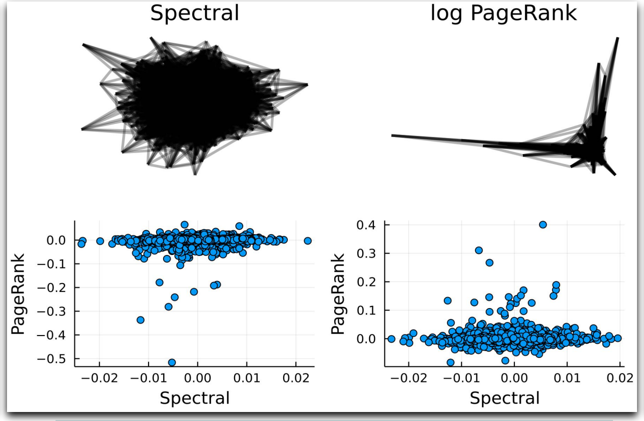
Generalizable? - Planted Partition Model

Sbm3000 (50,60,0.001,0.005)



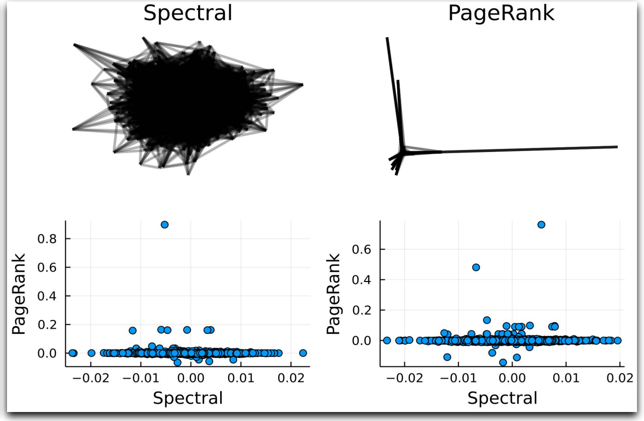
0.99

51.77%



0.99 with log

15.22%

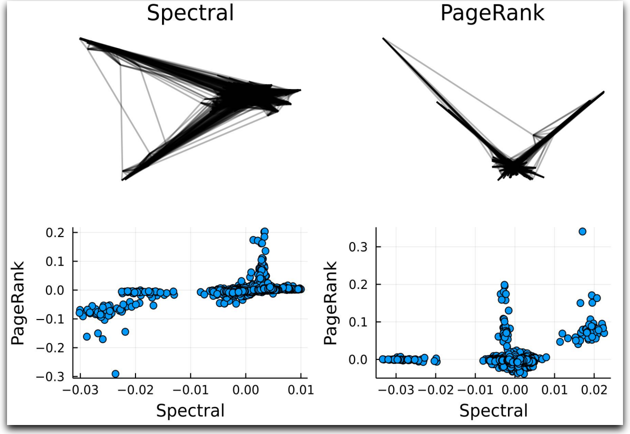


0.99999

51.32%

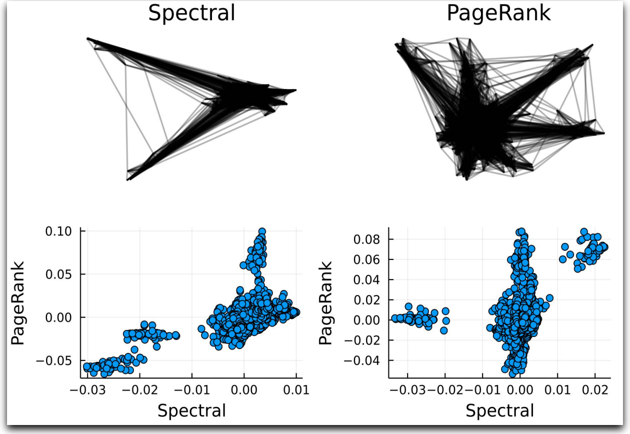
Generalizable? - Planted Partition Model

Sbm3000 (50,60,0.25,0.001)



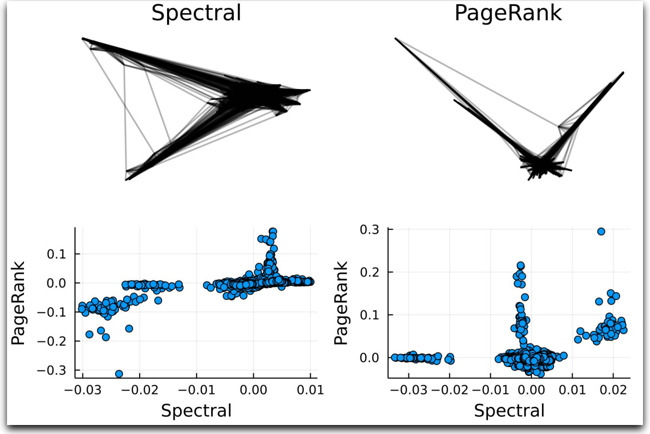
0.99

17.88%



0.99 with log

15.22%

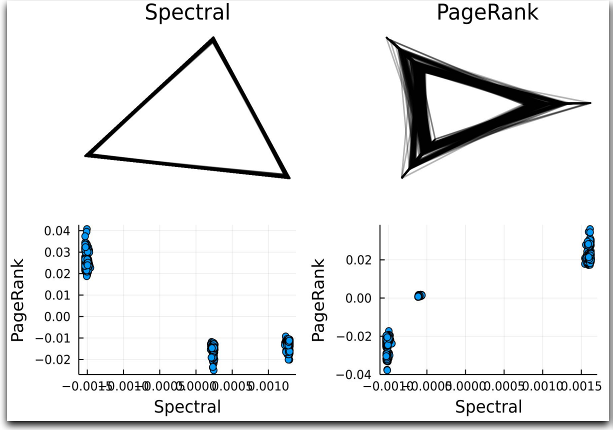


0.99999

90.13%

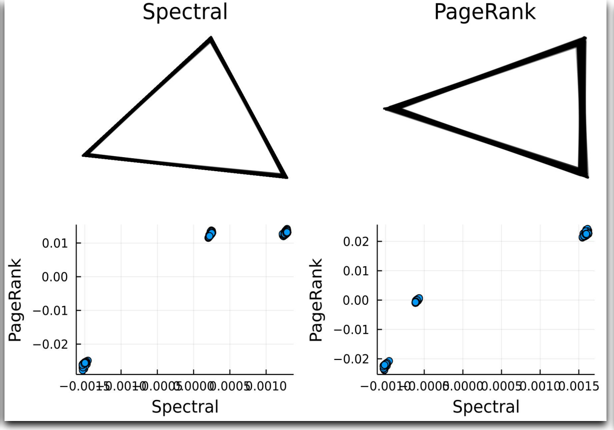
Generalizable? - Planted Partition Model

Sbm3000 (1000,3,0.25,0.001)



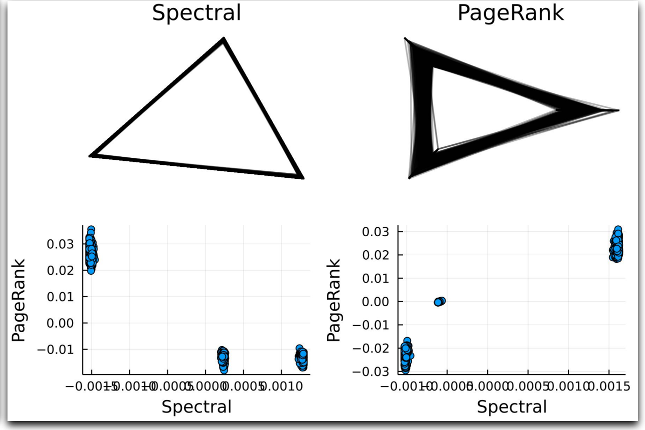
0.99

53.7%



0.99 with log

1.04%



0.99999

16.21%

Embedding with larger alpha

Graph	X(0.99)	X(0.99999)
Tapir	10.17%	15.41%
Minnesota	16.07%	11.15%
3000 - 6	47.6%	5.06%
Original	75.92%	81.14%
Sbm3000 (50,60,0.001,0.005)	51.77%	51.32%
Sbm3000(50,60,0.25,0.001)	17.88%	90.13%
Sbm3000 (1000,3,0.25,0.001)	53.7%	16.21%

Embedding with log

Graph	X(0.99)	log.X(0.99)
Tapir	10.17%	1.13%
Minnesota	16.07%	1.97%
3000 - 6	47.6%	0.37%
Original	75.92%	4.11%
Sbm3000(50,60,0.001,0.005)	51.77%	15.22%
Sbm3000(50,60,0.25,0.001)	17.88%	15.2%
Sbm3000 (1000,3,0.25,0.001)	53.7%	1.04%

Spectral & PageRank Embedding use the same information

$$\tilde{\mathbf{L}} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{\mathbf{T}}$$

$$\rightarrow (\mathbf{I} - \alpha\mathbf{A}\mathbf{D}^{-1})^{-1} = \mathbf{D}^{1/2}\mathbf{Q}(\mathbf{I} - \alpha\mathbf{E})^{-1}\mathbf{Q}^{\mathbf{T}}\mathbf{D}^{-1/2}$$, where $\mathbf{E} = \mathbf{I} - \boldsymbol{\Lambda}$

$$\mathbf{X} = (1-\alpha)(\mathbf{I} - \alpha\mathbf{A}\mathbf{D}^{-1})^{-1} = (1-\alpha)\mathbf{D}\mathbf{Z}(\mathbf{I} - \alpha\mathbf{E})^{-1}\mathbf{Z}^{\mathbf{T}}$$

$$= \left[\tfrac{1-\alpha}{1-\alpha\epsilon_1}\mathbf{d}\mathbf{z_1} \cdots \tfrac{1-\alpha}{1-\alpha\epsilon_n}\mathbf{d}\mathbf{z_n}\right]\begin{bmatrix}\mathbf{z_1}^T \\ \vdots \\ \mathbf{z_n}^T\end{bmatrix}$$, where $\mathbf{d} = \mathbf{D}\mathbf{e}$

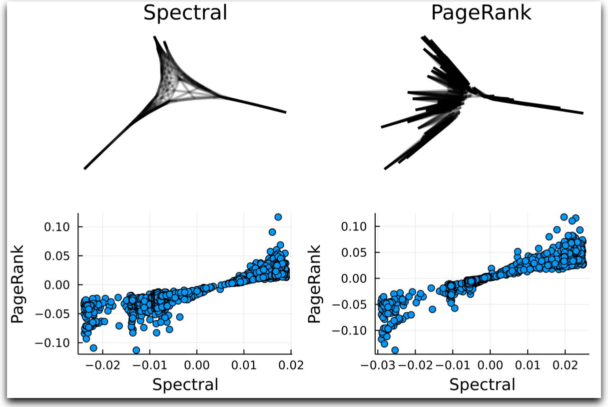
Spectral & PageRank Embedding use the same information

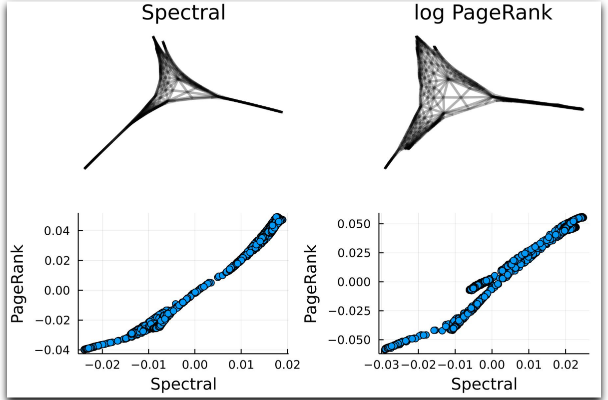
$$\mathbf{X} = \left[\tfrac{1-\alpha}{1-\alpha\epsilon_1}\mathbf{d}\mathbf{z_1} \cdots \tfrac{1-\alpha}{1-\alpha\epsilon_n}\mathbf{d}\mathbf{z_n}\right] \begin{bmatrix} \mathbf{z_1}^T \\ \vdots \\ \mathbf{z_n}^T \end{bmatrix}$$

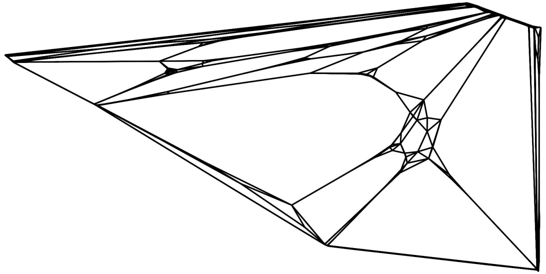
$$\epsilon_i = 1 - \lambda_i$$

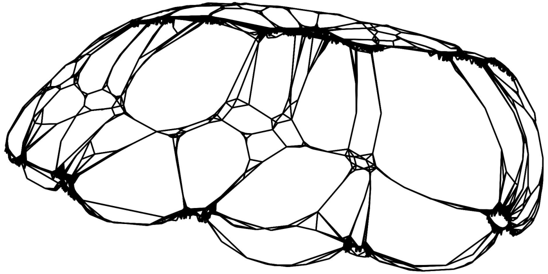
$$\mathbf{X} = \left[\mathbf{d}\mathbf{z}_1 \cdots \tfrac{1-\alpha}{1-\alpha\epsilon_n}\mathbf{d}\mathbf{z_n}\right] \begin{bmatrix} \mathbf{z_1}^T \\ \vdots \\ \mathbf{z_n}^T \end{bmatrix} \xrightarrow{\alpha \to 1} \mathbf{X} = \mathbf{d}\mathbf{z_1}\mathbf{z_1}^T$$

Spectral & PageRank Embedding use the same information







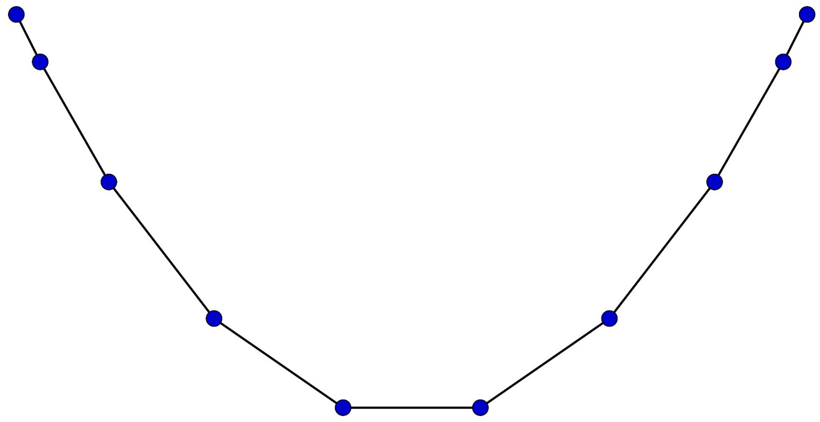


$$\rightarrow X_{i,j} = \sum_{t=1}^{n} \frac{1-\alpha}{1-\alpha\epsilon_t} z_{ti} z_{tj} d_i$$

$$\log(X_{i,j}) = \log(1-\alpha) + \log(d_i) + \log\left(\sum_{t=1}^{n} \frac{z_{ti} z_{tj}}{1-\alpha\epsilon_t}\right)$$

PageRank on Chain

$$(\alpha/2)x_2 = x_1$$
$$\alpha x_1 + (\alpha/2)x_3 = x_2$$
$$(\alpha/2)x_2 + (\alpha/2)x_4 = x_3$$
$$\vdots$$
$$(\alpha/2)x_{k-2} + (\alpha/2)x_k = x_{k-1}$$
$$(\alpha/2)x_{k-1} + (\alpha/2)x_{k+1} = x_k - (1-\alpha)$$
$$(\alpha/2)x_k + (\alpha/2)x_{k+1} = x_{k+1}$$
$$\vdots$$
$$(\alpha/2)x_{n-3} + (\alpha/2)x_{n-1} = x_{n-2}$$
$$(\alpha/2)x_{n-2} + \alpha x_n = x_{n-1}$$
$$(\alpha/2)x_{n-1} = x_n$$
$$\sum_{i=1}^{n} x_i = 1$$

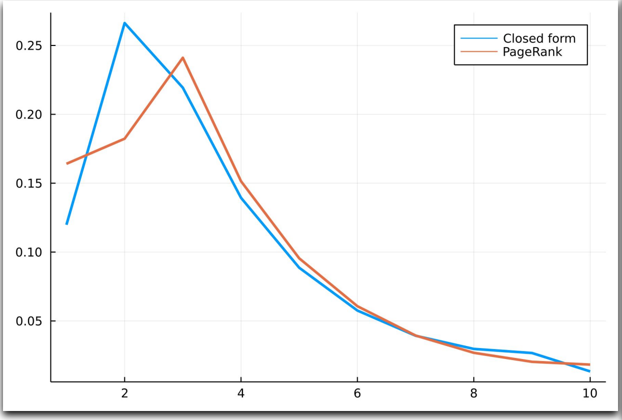


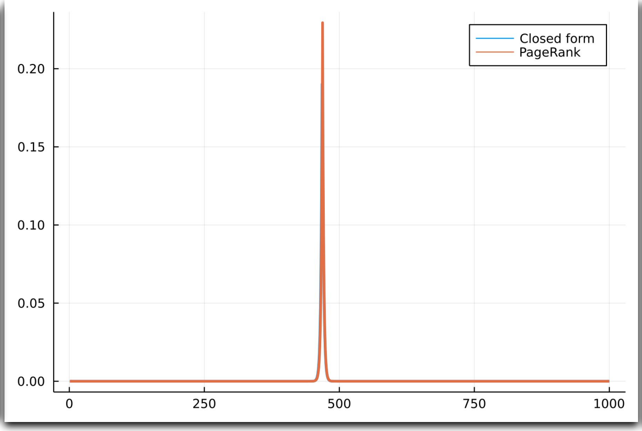
PageRank on Chain

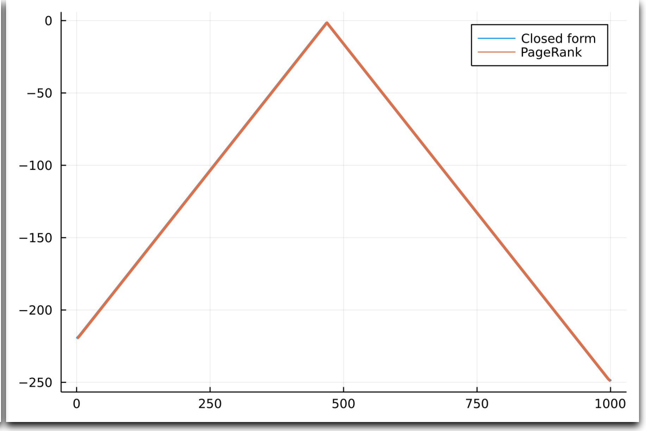
$$x_i = \begin{cases} \dfrac{\sqrt{(1-\alpha)/(1+\alpha)}}{(+)^{k-1}+(-)^{k-1}}\left((+)^{i-1}+(-)^{i-1}\right), i = 2, \cdots, k-1 \\ \dfrac{\alpha}{2}\sqrt{\dfrac{1-\alpha}{1+\alpha}}\left(\dfrac{(+)^{k-2}+(-)^{k-2}}{(+)^{k-1}+(-)^{k-1}} + \dfrac{(+)^{n-k-1}+(-)^{n-k-1}}{(+)^{n-k}+(-)^{n-k}}\right), i = k \\ \dfrac{\sqrt{(1-\alpha)/(1+\alpha)}}{(+)^{n-k}+(-)^{n-k}}\left((+)^{n-i}+(-)^{n-i}\right), i = k+1, \cdots, n-1 \end{cases}$$

$$(+) = \frac{1+\sqrt{1-\alpha^2}}{\alpha}$$

$$(-) = \frac{1-\sqrt{1-\alpha^2}}{\alpha}$$





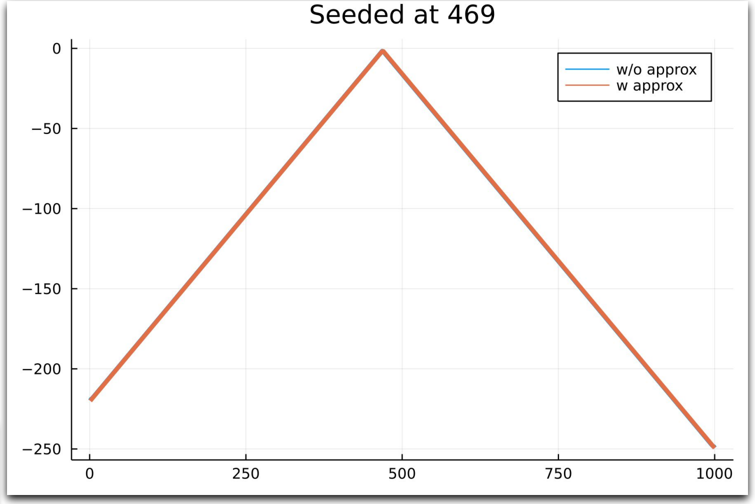


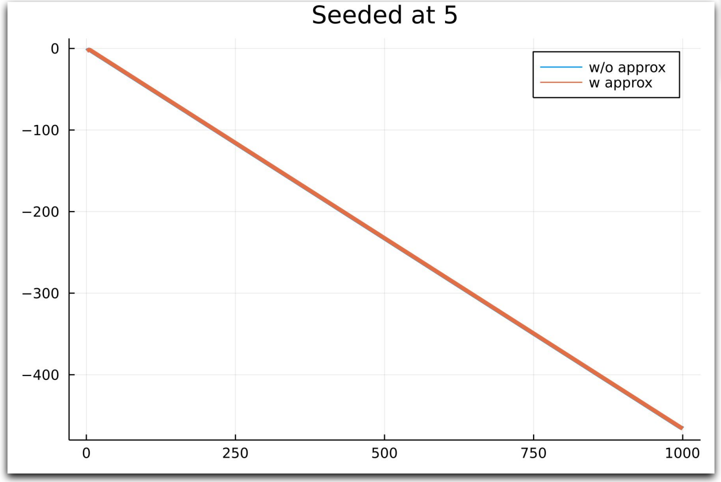
PageRank on Chain

$$\hat{x}_i = \sqrt{\frac{1-\alpha}{1+\alpha}}(+)^{-|i-k|}$$

$$\log(\hat{x}_i) = -|k-i|\log((+)) + \log(\sqrt{\frac{1-\alpha}{1+\alpha}})$$

$$(+) = \frac{1+\sqrt{1-\alpha^2}}{\alpha}$$

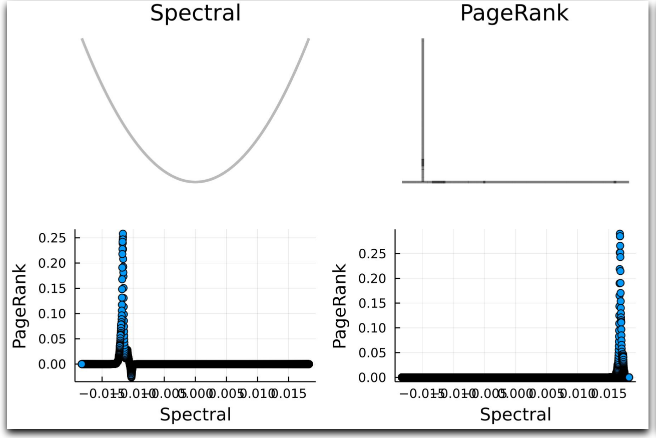




Generalizable? - Chain graph

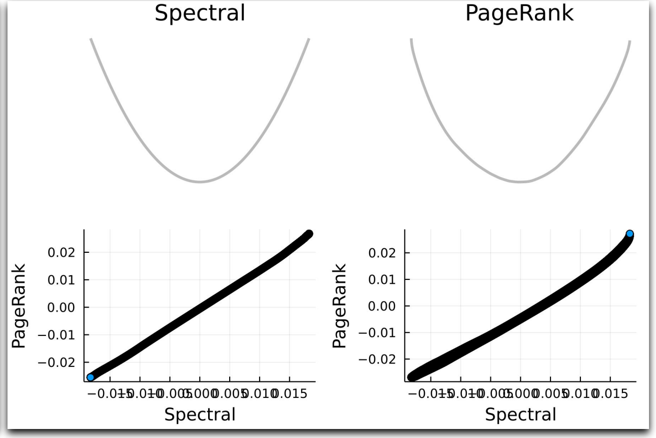
0.99

2858%



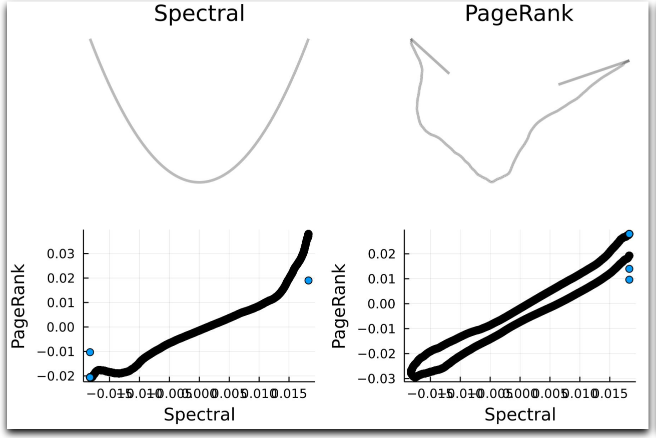
0.99-with log

1.06%



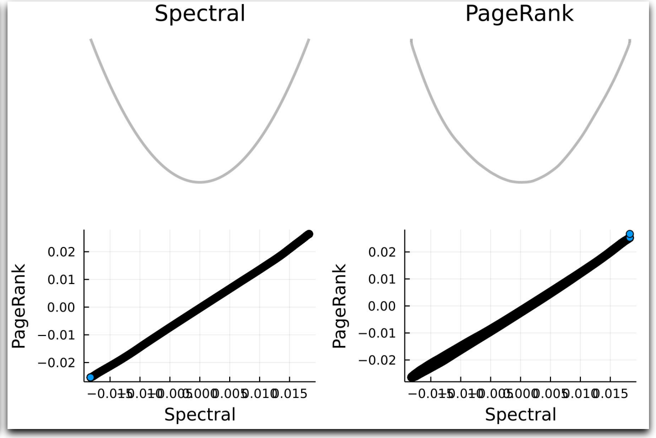
0.99999

30.38%



0.99999-with log

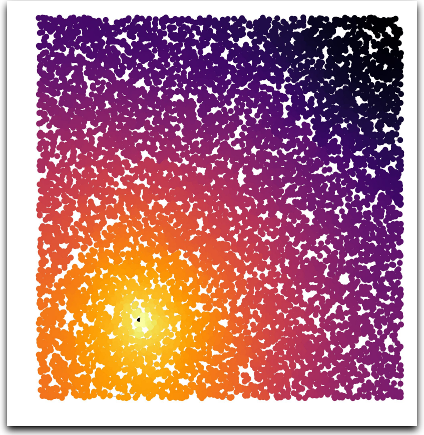
0.75%



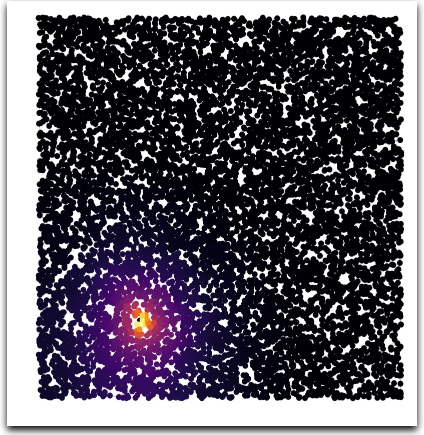
Why does this work?

$$\sum_{t=1}^{\infty}(\alpha \mathbf{G}\mathbf{D}^{-1})^t$$

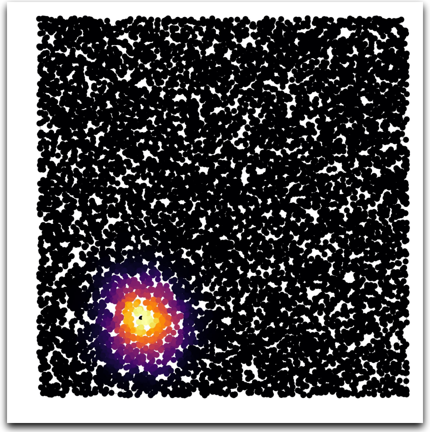
$$\left(\mathbf{D}^{-1/2}\mathbf{G}\mathbf{D}^{-1/2}\right)^r \left(\mathbf{D}^{-1/2}\mathbf{G}\mathbf{D}^{-1/2}\mathbf{v_k}\right)$$



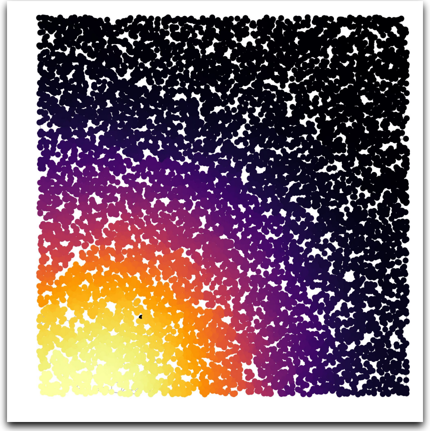
Log of PageRank(0.999)



PageRank(0.999)



r=100



r=2000

Why does this work?

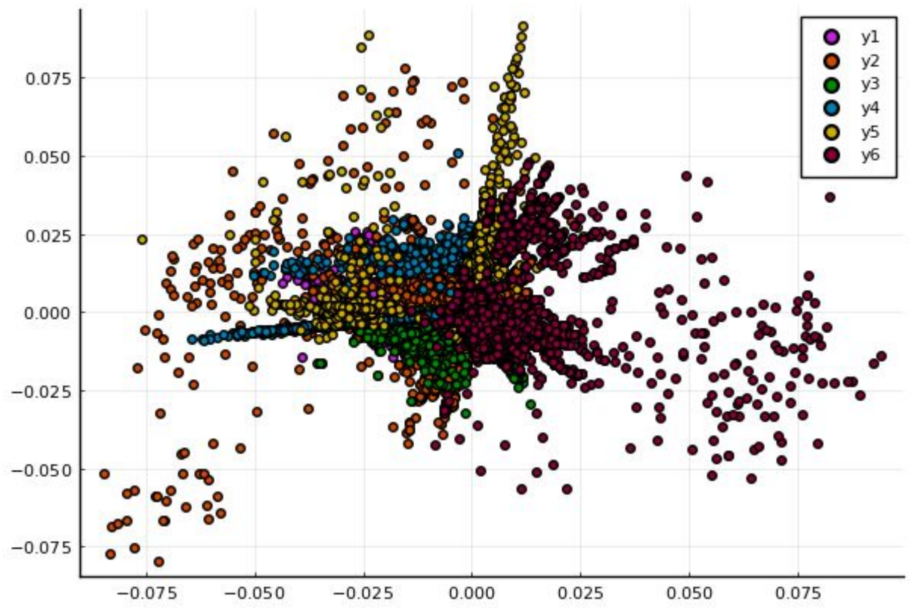
$$(\alpha = 0.9)^{1000} = 1.7e^{-46} \qquad (\alpha = 0.99999)^{1000} = 0.99$$
$$1000 \log(\alpha = 0.9) = -105.36$$

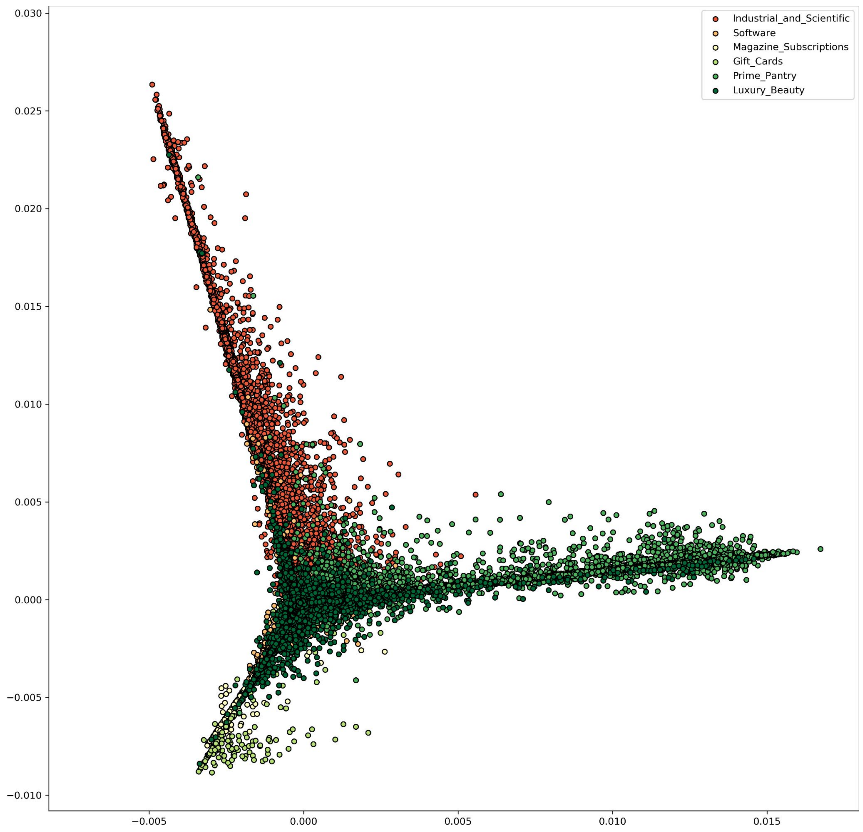
Graph	Matrix Power	Distance	Power attained
30 Chain	100	13	e^{-2}
3000 Chain	100	106	e^{-30}
3000 Chain	1000	106	e^{-3}
30 nearest neighbor	100	13	e^{-1}
3000 nearest neighbor	100	1792	e^{-20}
3000 nearest neighbor	1000	1792	e^{-4}

Summarizing

- Embedding procedure for the local environment of a seed node
- Embeddings created by using the left singular vectors of PageRank personalized on individual nodes resemble the spectral embeddings for graph where the log of PageRank is a function of the graph distance from the seed nodes.
- Reliable representation

Extension to higher order





Questions