GLEE: Geometric Laplacian Eigenmap Embedding

Leo Torres
Kevin S. Chan and Tina Eliassi-Rad

DEVCOM ARMY RESEARCH LABORATORY

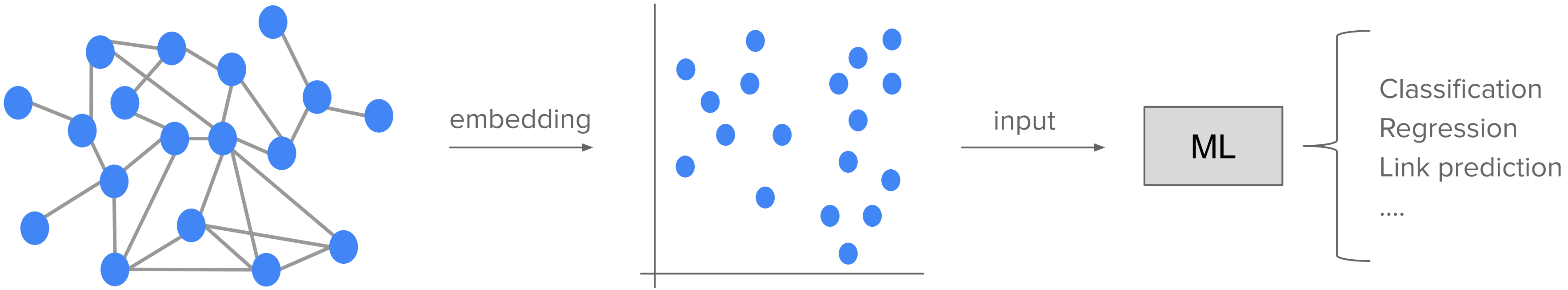
netsi
Northeastern University
Network Science Institute

Outline

1. The problem of graph embedding
2. Laplacian Eigenmaps
3. **GLEE**: Geometric Laplacian Eigenmap Embedding
4. Interpreting graph structure from GLEE
5. Experiments

The problem of embedding

- Seeks to represent the nodes of a graph as points in some space
- One can then use this representation for downstream ML tasks



[1] P. Goyal and E. Ferrara. Graph embedding techniques, applications, and performance: A survey. Knowledge-Based Systems, 151:78–94, 2018
[2] W. L. Hamilton, R. Ying, and J. Leskovec. Representation learning on graphs:Methods and applications.IEEE Data Eng. Bull., 40(3):52–74, 2017

Laplacian Eigenmaps

Originally from Belkin and Niyogi [3], it's based on the **spectral properties** of the quadratic form of the Laplacian. For a graph with ***n*** nodes, the Laplacian satisfies

$$L = D - A$$

$$tr(Y^t L Y) = \frac{1}{2} \sum_{ij} a_{ij} ||Y_i - Y_j||^2 \qquad Y \in \mathbb{R}^{n \times d}$$

[3] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation.Neural Computation, 15(6):1373–1396, 2003

Laplacian Eigenmaps

The Laplacian Eigenmap (LE) embedding of a graph is defined as

$$\min_Y tr(Y^t LY)$$
$$s.t. Y^T DY = I$$

[3] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation.Neural Computation, 15(6):1373–1396, 2003

Laplacian Eigenmaps

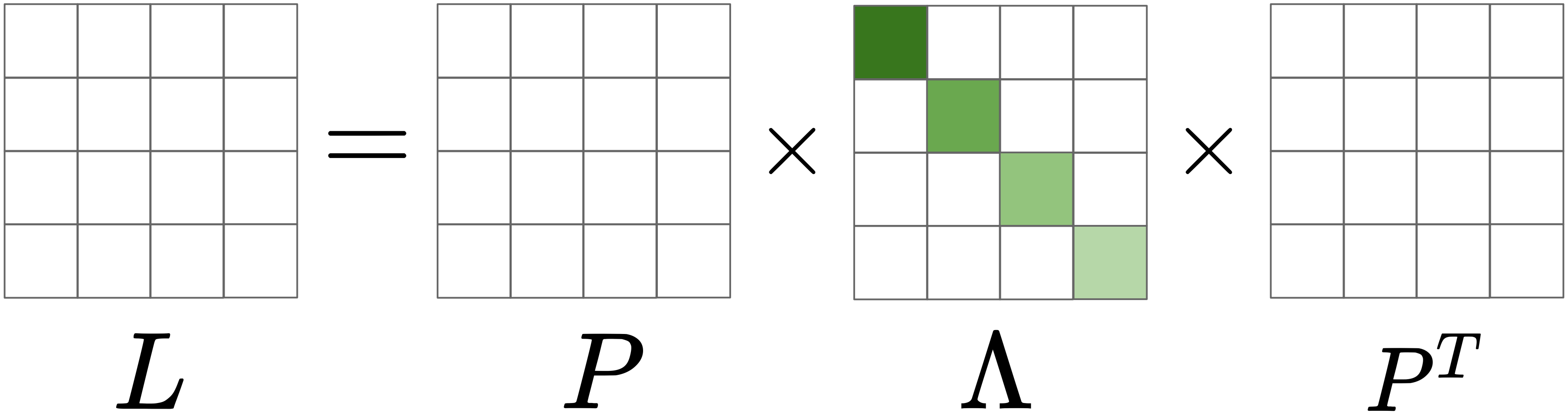
The Laplacian Eigenmap (LE) embedding of a graph is defined as

$$\min_Y tr(Y^t L Y)$$
$$s.t. \, Y^T D Y = I$$

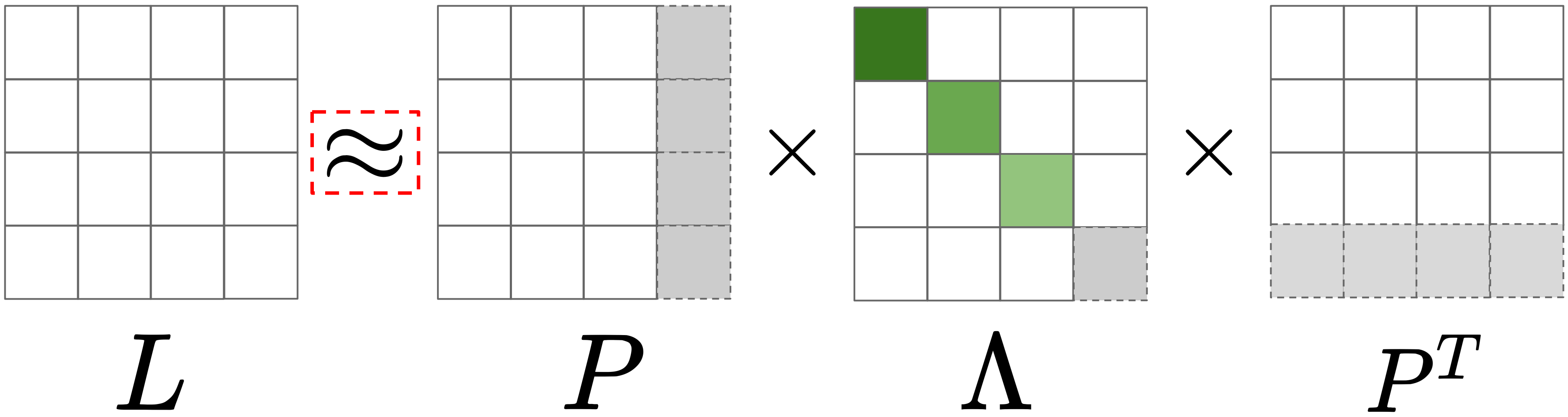
Lagrange multipliers shows that the columns of the solution **Y*** are given by the eigenvectors of the normalized Laplacian, **$D^{-1}L$**, corresponding to the **lowest eigenvalues**.

[3] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation.Neural Computation, 15(6):1373–1396, 2003

GLEE: Geometric Laplacian Eigenmap Embedding

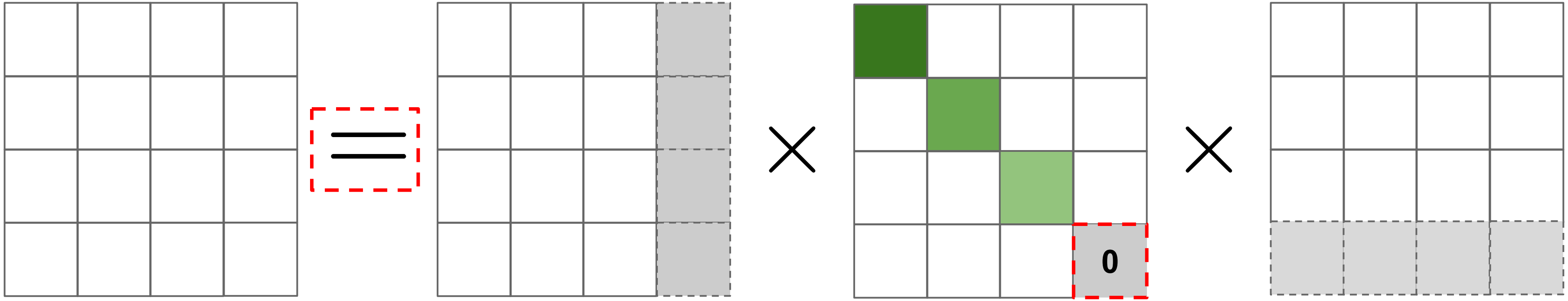


GLEE: Geometric Laplacian Eigenmap Embedding



Singular Value Decomposition says that eliminating the rows and columns corresponding to the lowest singular values give a good approximation of ***L***.

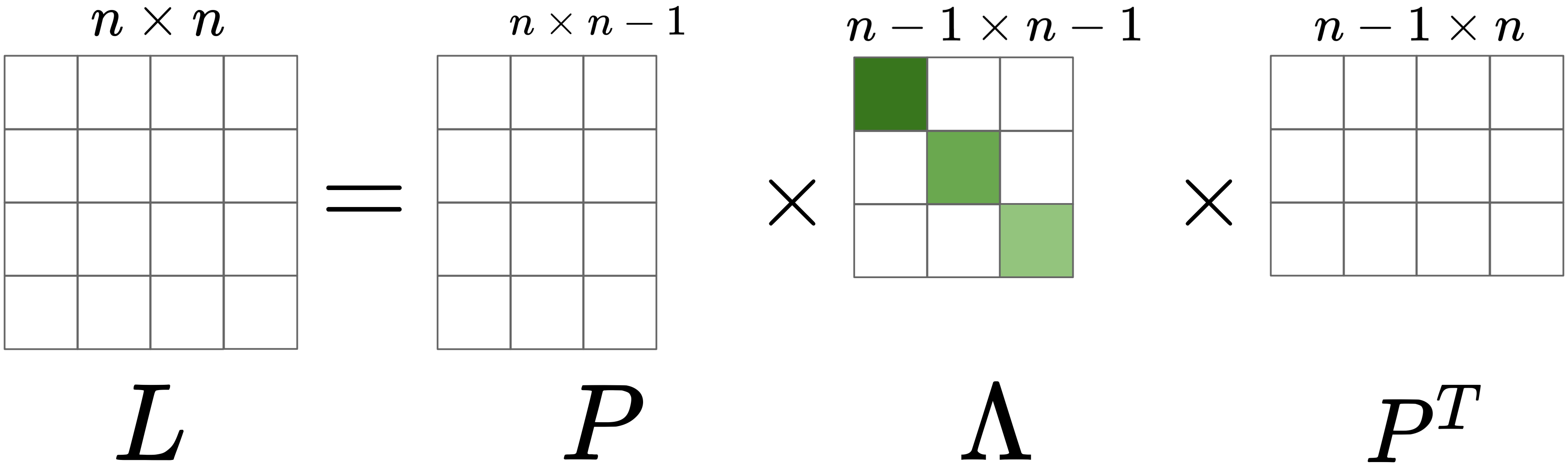
GLEE: Geometric Laplacian Eigenmap Embedding



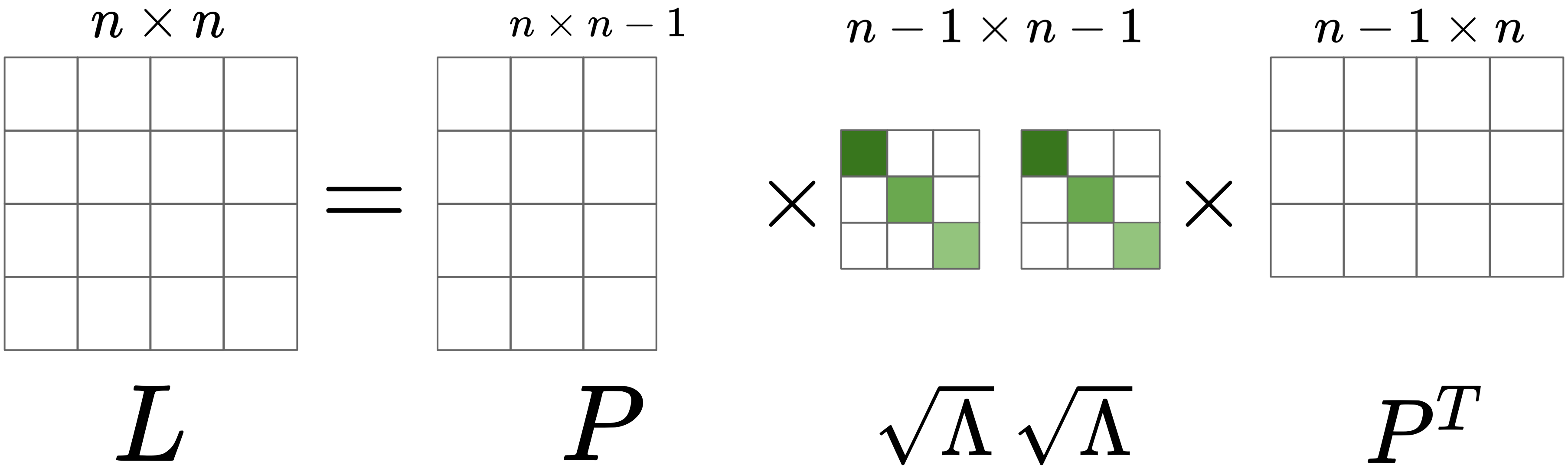
$$L = P \times \Lambda \times P^T$$

However, the last eigenvalue of ***L*** is always **0**, which implies exact equality.

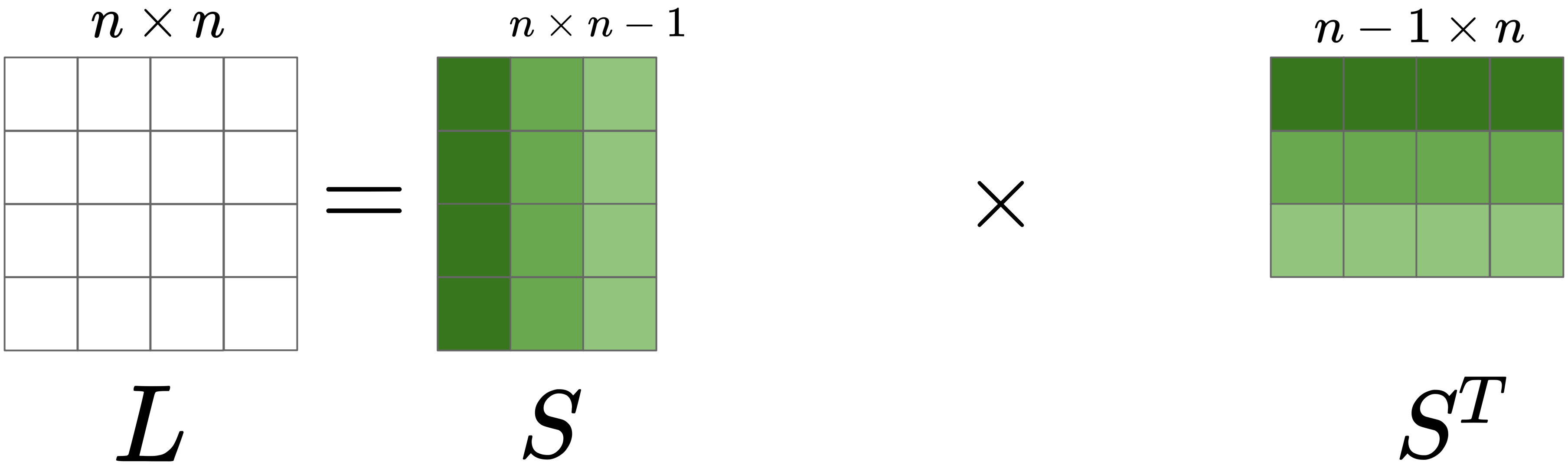
GLEE: Geometric Laplacian Eigenmap Embedding



GLEE: Geometric Laplacian Eigenmap Embedding

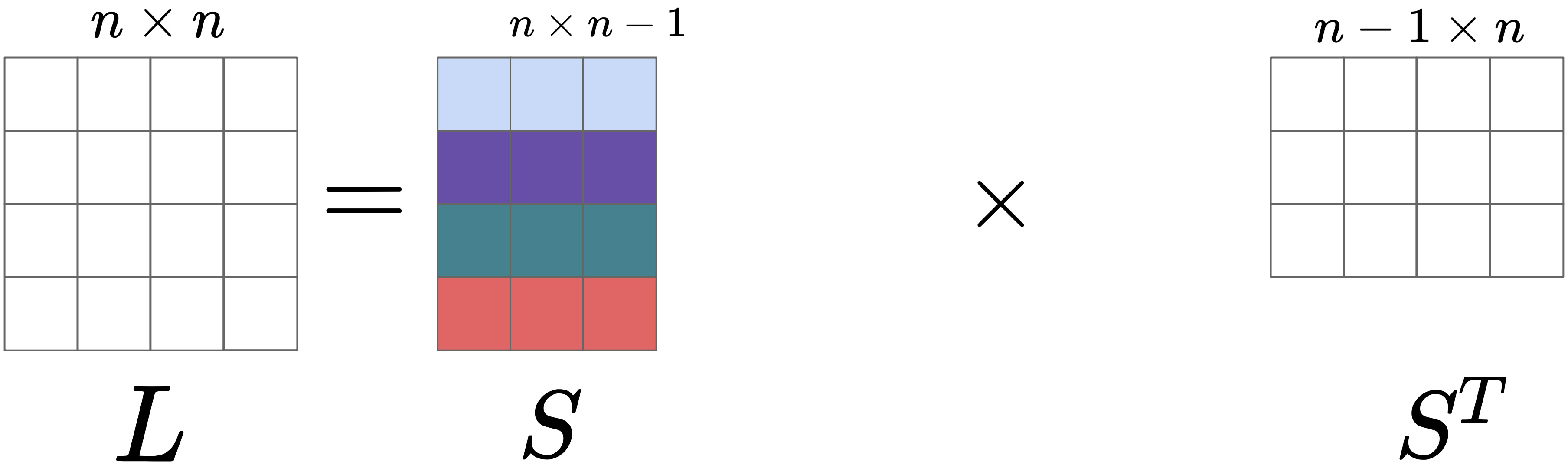


GLEE: Geometric Laplacian Eigenmap Embedding



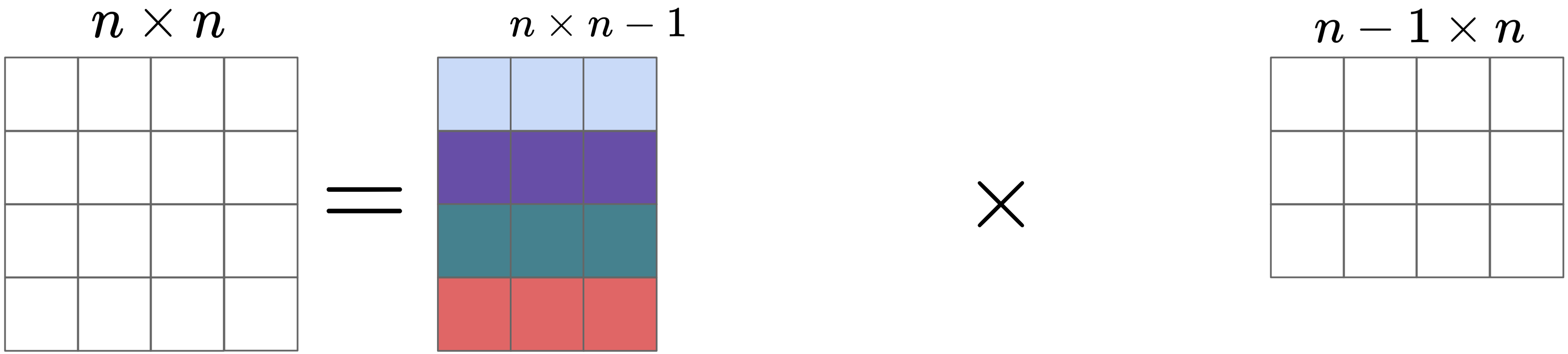
$$S = P\sqrt{\Lambda}$$

GLEE: Geometric Laplacian Eigenmap Embedding

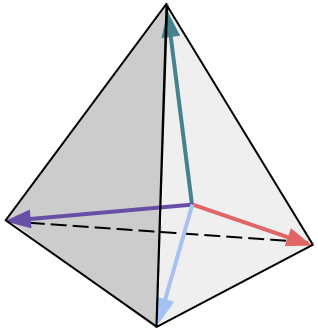


In a connected graph, ***L*** has rank ***n-1,*** and only one eigenvalue equal to ***0.***
This implies that ***S*** has full rank, i.e., rank ***n-1.***

GLEE: Geometric Laplacian Eigenmap Embedding

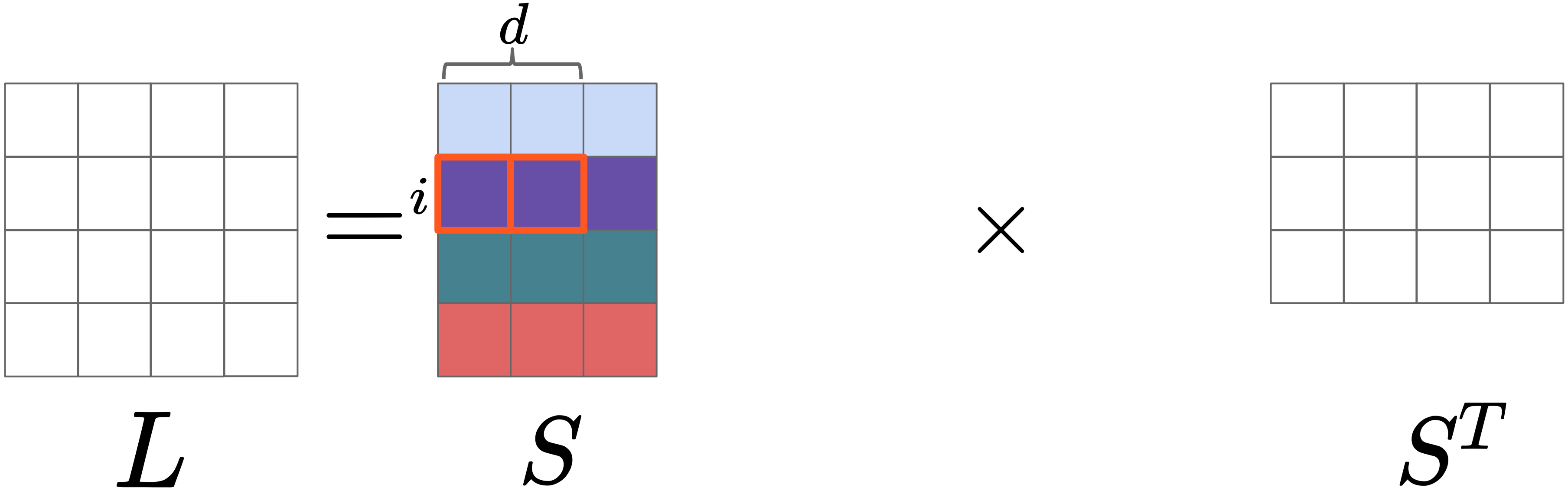


This implies that the rows of ***S*** point to the vertices of an **(n-1)**-D **simplex**.



[3] K. Devriendt and P. Van Mieghem. The simplex geometry of graphs.The Journal of Complex Networks, 2019.
[4] M. Fiedler. Matrices and graphs in geometry, volume 139 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2011.

GLEE: Geometric Laplacian Eigenmap Embedding



Given a graph **G = (V, E)**, define the ***d***-dimensional **GLEE** of a node ***i*** as the first ***d*** columns of the ***i***-th row of $S = P\Lambda^{1/2}$, and is denote it by s_i.

Similarity-based vs structure-based

Laplacian Eigenmaps

$$\min_Y tr(Y^t L Y)$$

$$s.t. \, Y^T D Y = I$$

Smallest eigenvalues give optimal **distance minimization** between **similar** nodes.

Similarity-based vs structure-based

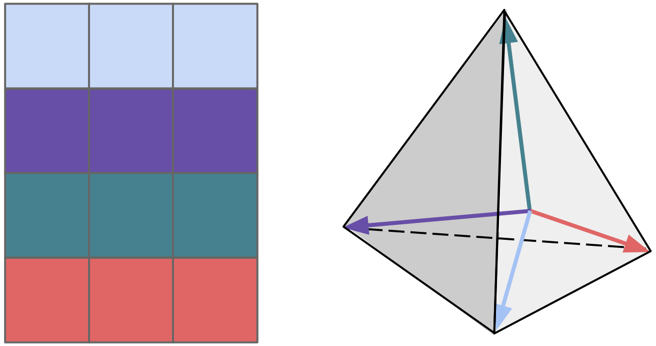
Laplacian Eigenmaps

$$\min_Y tr(Y^t L Y)$$

$$s.t.\, Y^T D Y = I$$

Smallest eigenvalues give optimal **distance minimization** between **similar** nodes.

GLEE



Largest eigenvalues give a **geometric encoding** of the graph's structure and the **best low-rank approximation**.

Outline

1. The problem of graph embedding
2. Laplacian Eigenmaps
3. **GLEE**: Geometric Laplacian Eigenmap Embedding
4. Interpreting graph structure from GLEE
5. Experiments

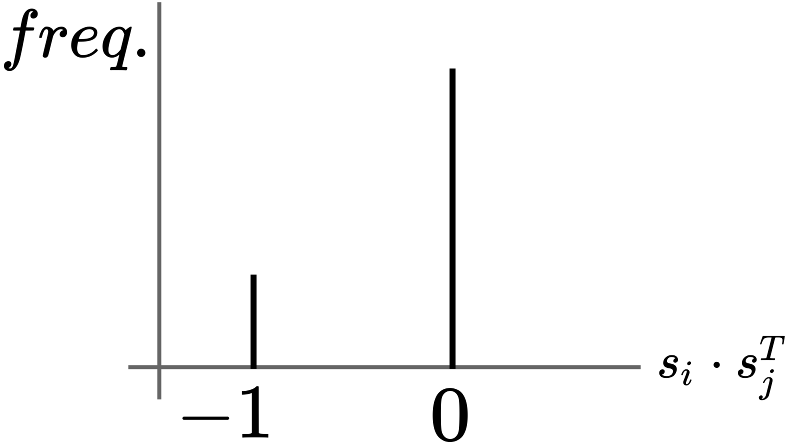
Graph Reconstruction

Given the matrix $\boldsymbol{S}$ whose rows are $\boldsymbol{s}_i$, how do we reconstruct the graph?

Graph Reconstruction

Given the matrix $\boldsymbol{S}$ whose rows are $\boldsymbol{s}_i$, how do we reconstruct the graph?

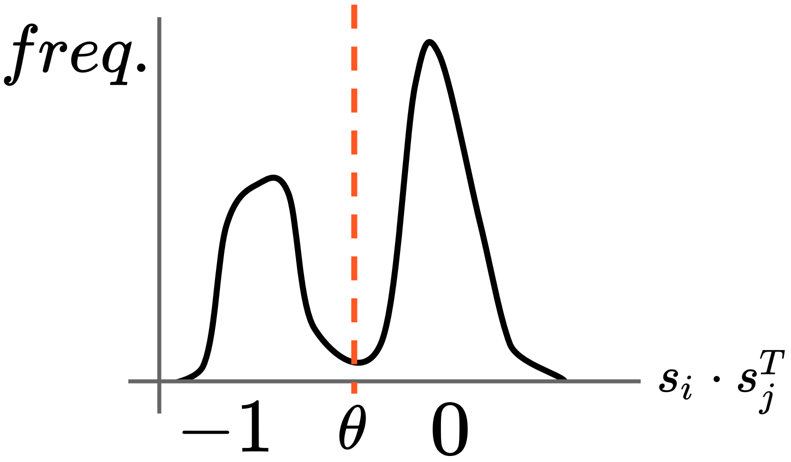
- Assume $\boldsymbol{d} = \boldsymbol{n-1}$. In this case, we simply have $\boldsymbol{L} = \boldsymbol{S}\,\boldsymbol{S}^T$.



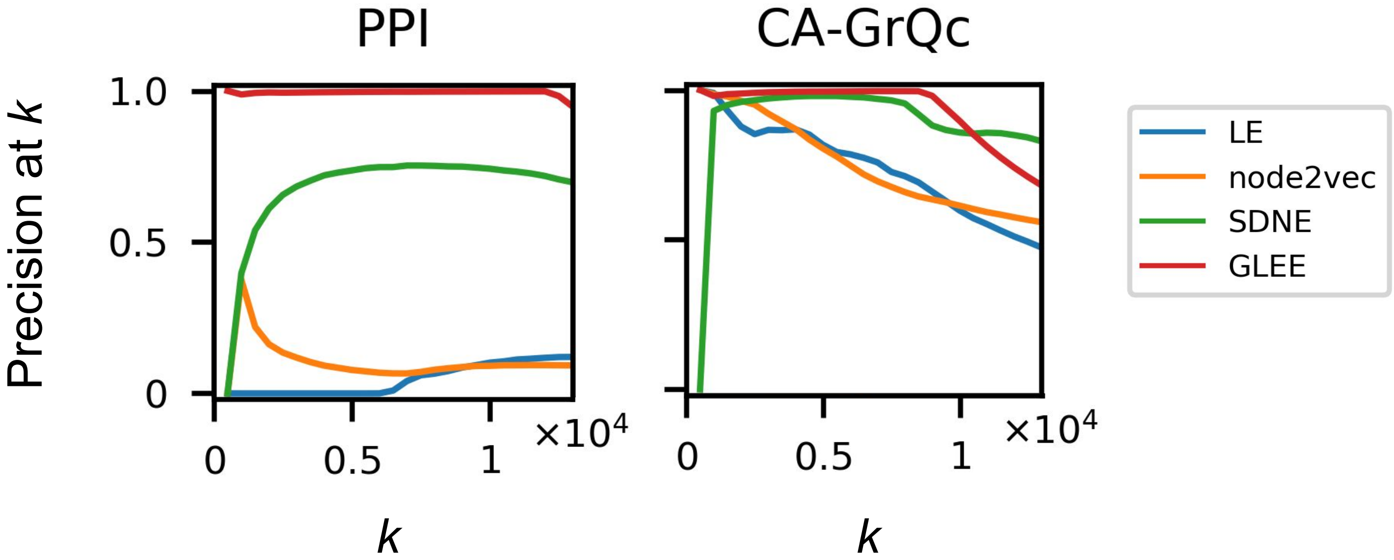
Graph Reconstruction

Given the matrix ***S*** whose rows are s_i, how do we reconstruct the graph?

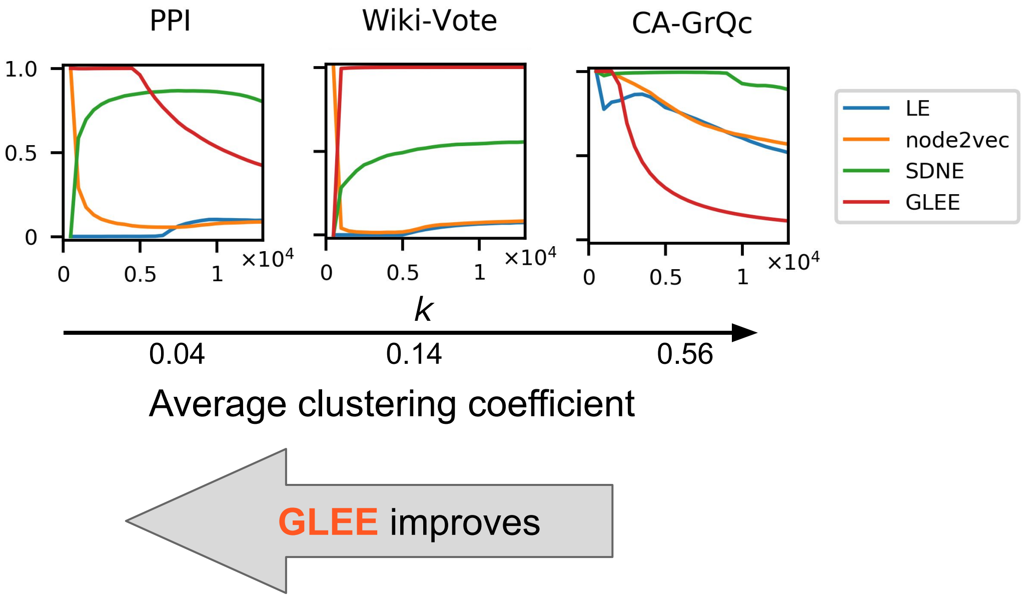
- Assume ***d = n-1***. In this case, we simply have $L = S\,S^T$.
- If ***d < n***, then $S\,S^T$ is the best rank-d approximation of ***L***.

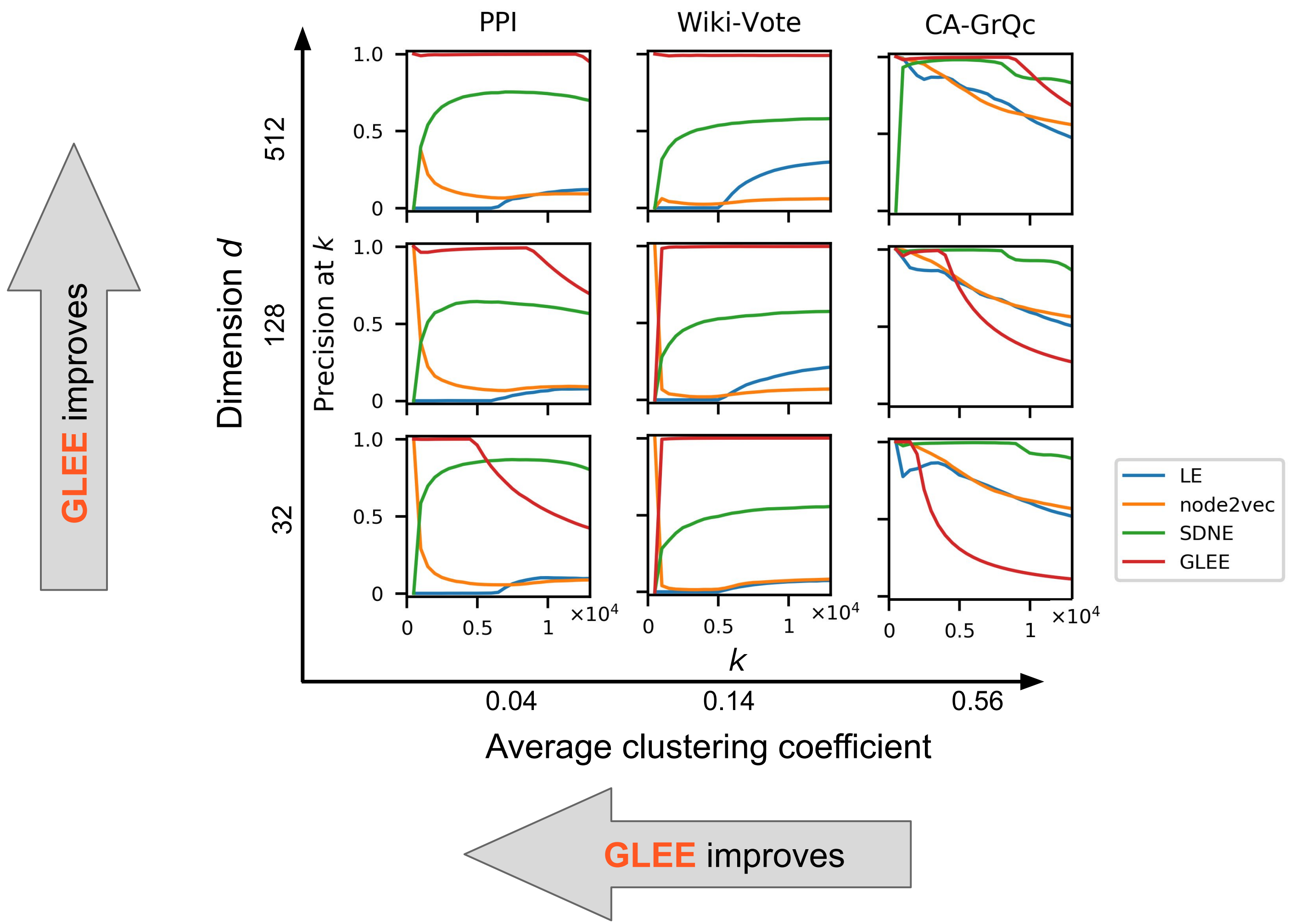


Graph Reconstruction: results



Same embedding dimension, similar network size, but **different average clustering**.

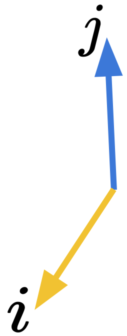

PPI
Wiki-Vote
CA-GrQc
1.0
0.5
0
0
0.5
1
×10^4
k
LE
node2vec
SDNE
GLEE
0.04
0.14
0.56
Average clustering coefficient
GLEE improves



PPI
Wiki-Vote
CA-GrQc
512
128
32
Dimension d
Precision at k
1.0
0.5
0
0.5
1
×10^4
k
LE
node2vec
SDNE
GLEE
0.04
0.14
0.56
Average clustering coefficient
GLEE improves
GLEE improves

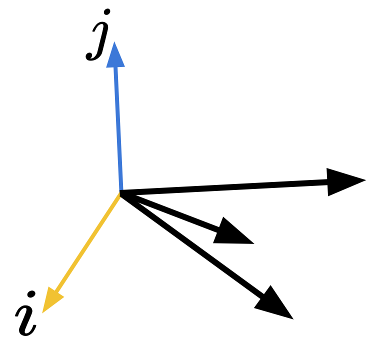
Link Prediction: common neighbors

In many networks (e.g. social networks), the number of common neighbors is an excellent predictor of links because of **triadic closure**.



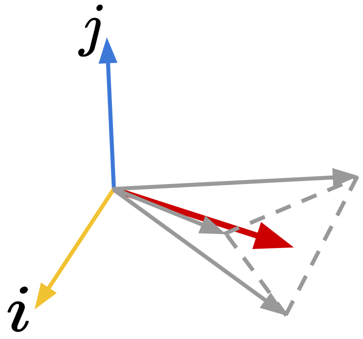
Link Prediction: common neighbors

In many networks (e.g. social networks), the number of common neighbors is an excellent predictor of links because of **triadic closure**.



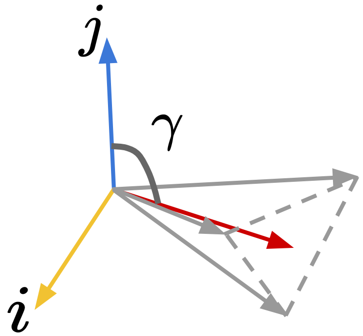
Link Prediction: common neighbors

In many networks (e.g. social networks), the number of common neighbors is an excellent predictor of links because of **triadic closure**.



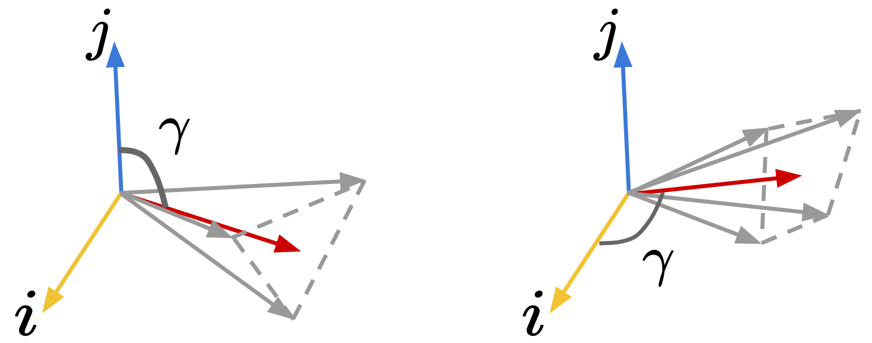
Link Prediction: common neighbors

In many networks (e.g. social networks), the number of common neighbors is an excellent predictor of links because of **triadic closure**.



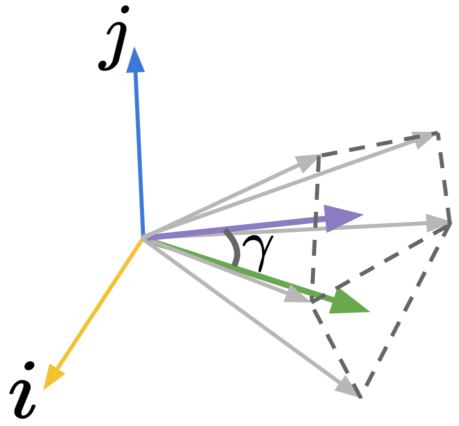
Link Prediction: common neighbors

In many networks (e.g. social networks), the number of common neighbors is an excellent predictor of links because of **triadic closure**.

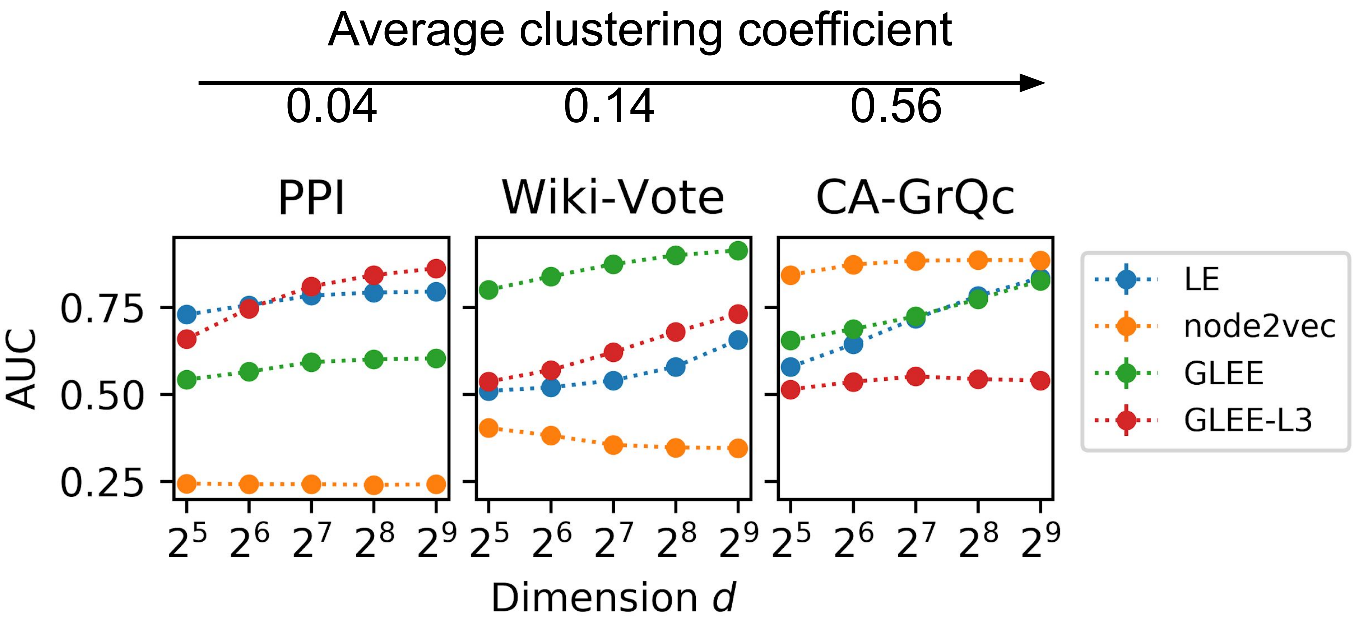


Link Prediction: 3-paths

In other networks with **low clustering** (e.g. PPI networks), a better predictor is the number of paths of length 3.

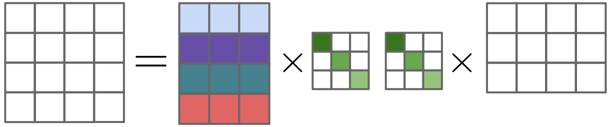


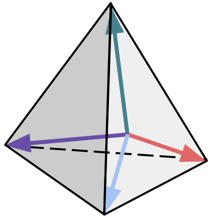
Link prediction: results

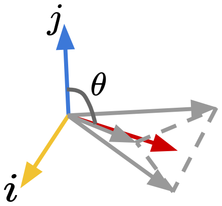


Thank You!

1. **GLEE** replaces distance-minimization with the direct geometric encoding of graph structure.
2. **GLEE** performs best when the clustering coefficient is low and the embedding dimension is high.
3. What other **geometric properties** of embeddings can we exploit?







Code: github.com/**leotrs**/glee

Paper: arxiv.org/abs/**1905.09763**

Contact: leo@**leotrs**.com

Supported by **NSF** CNS-1314603, **NSF** IIS-1741197, Army Research Laboratory
Cooperative Agreement W911NF-13-2-0045.