Appendix for STABLE: Identifying and Mitigating Instability in Embeddings of the Degenerate Core

David Liu, Tina Eliassi-Rad

A Appendix

A.1 Synthetic Networks Below we specify the configurations used to generate the synthetic networks used in our experiments.

ER We utilize the `erdos_renyi_graph` generator in the networkx package and call the generator twice with the following parameters: $(n = 5000, p = 0.002), (n = 5000, p = 0.004)$.

BA We utilize the `barabasi_albert_graph` generator in the networkx package and call the generator twice with the following parameters: $(n = 5000, m = 5), (n = 5000, m = 10)$.

BTER We generated the BTER graphs using the Matlab `FEASTPACK` software package which was downloaded from `http://www.sandia.gov/~tgkolda/feastpack/`. The software expects an input degree distribution. For the "BA (from BA)" graph, we provided the edgelist from the "BA ($m = 5$)" graph detailed above. For the "BA (Arbitrary)" graph we utilized `FEASTPACK` to generate an arbitrary degree distribution based on the following specifications: the maximum degree is < 1000, the target average degree is 15, the target maximum clustering coefficient is 0.95, and the target global clustering coefficient is 0.15.

A.2 Graph Dataset Diversity Figure 8 shows that the selected graphs span a variety of k-core structures, as defined by the graph degeneracy and the maximum-core link entropy [15], which is high when the degenerate core is well-connected with the outer shells.

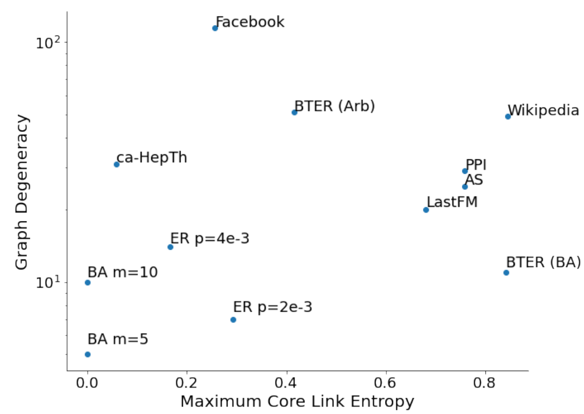


Figure 8: We chose graph datasets that have diverse k-core structures. Each graph is plotted based on its degeneracy vs. its maximum-core link entropy ($\in [0, 1]$). Liu et al. [15] define maximum-core link entropy, where higher values correspond with degenerate cores that are well-connected with the outer shells; and graphs with low maximum-core link entropy have degenerate cores isolated from the rest of the graph. The Y-axis is the graph's degeneracy (the largest k in the graph's k-cores).

A.3 Dense Erdös-Rényi Graph Theorem Proof of Theorem 3.1

Proof. For a weighted adjacency matrix W where w_{ij} is the weight of the edge between nodes i and j and embeddings $X \in \mathbb{R}^{nxd}$, the Laplacian Eigenmap loss function is:

$$l_G(X) = \sum_{ij} w_{ij} \|X(i) - X(j)\|^2 = 2X^T LX \tag{A.1}$$

Where the row-normalized Laplacian $L = D^{-1}(D - A)$ and $X(i)$ is the i^{th} row of X.

To avoid arbitrary scaling and weight nodes according to their degrees, the columns of X are constrained to be orthonormal. If $x_1, ..., x_n$ are the columns of X, the Laplacian Eigenmap loss can be expressed as:

$$l_G(X) = 2X^T L X \tag{A.2}$$

$$= 2\sum_{j=1}^{d} x_j^T L x_j \tag{A.3}$$

If $\lambda_1, ..., \lambda_n$ are the non-zero eigenvalues of L in increasing order, where it is assumed that G is connected, the minimum value of $l_G(X)$ is $2\sum_{i=1}^{d}\lambda_i$ and the maximum value is $2\sum_{i=n-d+1}^{n}\lambda_i$ where the ratio of the two can be lower-bounded as:

$$\frac{\min_X l_G(X)}{\max_X l_G(X)} = \frac{\sum_{i=1}^{d}\lambda_i}{\sum_{i=n-d+1}^{n}\lambda_i} \tag{A.4}$$

$$\geq \frac{\lambda_1}{\lambda_n} \tag{A.5}$$

In the case of an Erdös-Rényi graph G, Chung et al. establish that the eigenvalues of G are almost surely bounded as:

$$\max_i \|1 - \lambda_i\| \leq [1 + o(1)]\frac{2}{\sqrt{p(n-1)}}$$

More specifically, as n approaches infinity the density of eigenvalues converges in probability to a semi-circle distribution centered at 1 with radius $[1+o(1)]\frac{2}{\sqrt{p(n-1)}}$. Substituting the end points of the eigenvalue distribution into the loss function we have:

$$\frac{\min_X l_G(X)}{\max_X l_G(X)} \geq \frac{\lambda_1}{\lambda_n} \tag{A.6}$$

$$= \frac{\sqrt{p(n-1)} - 2[1+o(1)]}{\sqrt{p(n-1)} + 2[1+o(1)]} \tag{A.7}$$

As p approaches 1, the ratio approaches 1, showing that for dense Erdös-Rényi graphs the gap between the best set of embeddings and the worst diminishes. ∎

A.4 Empirical Validation of Theorem 3.1 We present simulation results that empirically validate Theorem 3.1, which was proven in Appendix A.3. For a given value of $n \in \{20, 50, 100, 200, 500\}$, we sampled Erdös-Rényi graphs with increasing values of edge-density p. Then, for each graph, we calculated the ratio between the Laplacian Eigenmap loss for the optimal embeddings where $d = 5$ and the loss for the least optimal, as defined in Equation A.5. Figure 9 shows that for all values of n, the ratio approaches 1 as edge density increases. Further, for fixed p, the ratios are larger for larger values of n, which is consistent with the bound in Theorem 3.1.

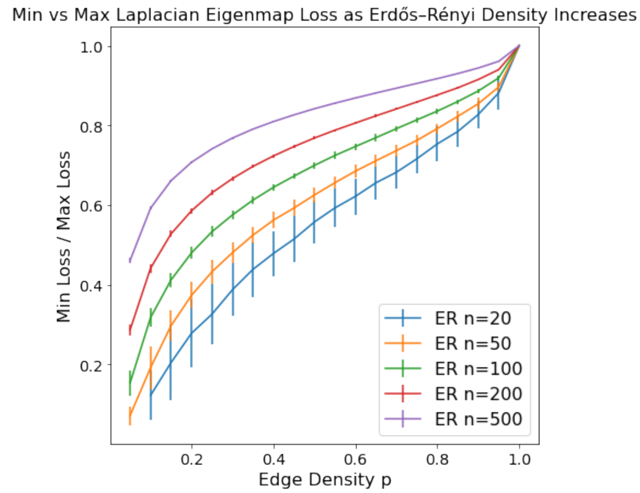


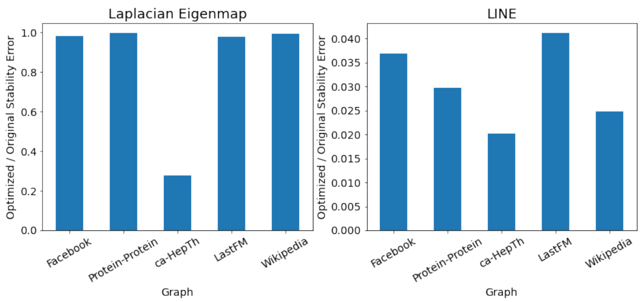
Figure 9: The above figure empirically validates Theorem 3.1, which states that as the density of an Erdös-Rényi graph increases, the gap between the best and worst embeddings diminishes. The y-axis is the ratio between the Laplacian Eigenmap loss for the optimal embeddings and the loss for the least optimal embeddings. We repeat for multiple values of n, and the error bars indicate 95% confidence intervals following 100 trials for each value of p.

A.5 Experimental Setup For our experimental results, we used the following hyperparameter settings: ($\alpha = 10^5, \beta = 0.1$) for `STABLE` Laplacian Eigenmaps (except in the case of the Facebook graph for which $\alpha = 10^6$) and ($\alpha = 10$) for `STABLE` LINE. These values were chosen so that the orders of magnitude for $\mathcal{L}_b$ and $\mathcal{L}_s$ are similar. For both instantiations, we use an initial learning rate of $\eta = 0.025$ that decreases linearly with each epoch until the rate reaches zero at the final epoch. Furthermore, our link prediction tests withhold 10% of links for the test set, and for each graph and algorithm configuration we ran five trials each with randomly sampled edge sets. The labels for link prediction are determined by sorting the cosine similarity scores for all pairs of nodes in the test set and all scores above a set threshold are labels as positive predictions. The threshold is set such that the number for positive predictions matches the number of true positives.

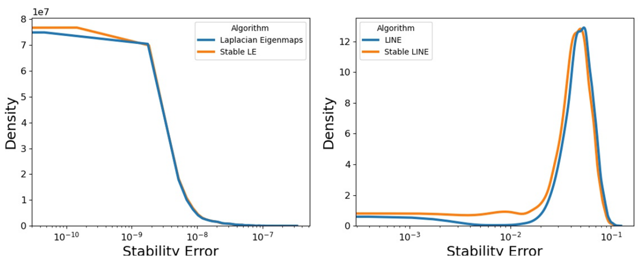
A.6 Further `STABLE` Results Figure 10a plots the ratio of `STABLE`'s average stability error to the base algorithm's, where the stability error for a pair $i, j \in \mathcal{D}$ is defined in Equation A.8, across real-world networks. For both instantiations, ratio is less than 1 indiciating an improvement in stability. The LINE instantiations exhibit much less stability error.

$$|p(\boldsymbol{u}_i, \boldsymbol{u}_j) - p(\hat{\boldsymbol{u}}_i, \hat{\boldsymbol{u}}_j)|^2 \tag{A.8}$$

Figure 10b provides a more detailed view of improved



(a) Reduced stability errors



(b) Detailed view for the Facebook graph

Figure 10: a) For both LINE and Laplacian Eigenmap instantiations, `STABLE` exhibits a decrease in stability error, as defined in Eq. A.8. The figure shows the ratio of `STABLE`'s average stability error to the base algorithm's average error. While all ratios are below 1, the LINE ratios are substantially smaller due to higher initial stability errors. b) A detailed view of the distribution of stability errors among all pairs of degenerate nodes in the Facebook graph. For both instantiations, the distribution of stability errors is closer to zero for the stable embeddings. That is, `STABLE` is able to find stable degenerate-core embeddings.

stability for the Facebook graph. The plots show the distribution of stability errors for all pairs of nodes in the degenerate core. The distributions for `STABLE` are more left-skewed than the base distributions indicating lower stability errors.

A.7 Reproducibility We have a GitHub repository for this work available at `https://github.com/dliu18/stable`.