Community Detection in Stochastic Block Model Variations

Allison Gunby-Mann
Dartmouth College
allison.mann.th@dartmouth.edu

Peter Chin
Dartmouth College
pc@dartmouth.edu

***Abstract*—In this paper we expand upon a spectral algorithm for the stochastic block model first presented by Chin et. al. with additional variations of construction. Our algorithm works with graphs that contain unequal sized blocks and non-uniform densities which was not considered by the original algorithm. It builds upon the spectral algorithm and remains robust and simple while capable of solving additional cases.**

***Index Terms*—stochastic block model, spectral theory, machine learning**

I. Introduction

The Stochastic Block Model (SBM) is a widely studied model for community detection in graphs. Its appeal stems from its ability to represent latent community structures within networks, where vertices exhibit higher likelihoods of connection within their respective blocks than between blocks. It provides a probabilistic framework that models the uncertainty inherent in real-world block networks. Through the specification of block parameters governing intra- and inter-community connectivity probabilities, the SBM enables inference about a variety of tasks, including underlying community structures, facilitating tasks such as community detection, link prediction, and network generation.

More formally, in the classic stochastic block model, there are k blocks each of size n where edges are generated randomly based on the following distribution: Pairwise, there will be an edge between nodes (u, v) with probability $\frac{a}{n}$ if u and v belong to the same block, and with probability $\frac{b}{n}$ otherwise. Our paper will consider the general case of this scenario, with some variations in construction. For simplicity we will focus on the $k = 2$ case to start.

If we denote the true blocks (sets of users within the same community) to be V_1 and V_2 in the $k = 2$ case, we want to find a partition of the vertex set V_1' and V_2' such that V_1 and V_1' are close to each other, and as are V_2 and V_2'. The metric that is commonly used in literature to evaluate these reconstructions of blocks V_1' and V_2' is γ-correctness.

Definition 1. *Given a set of k true blocks V_i for $i \in [1, k]$, a reconstruction V_i' is γ-correct if $|V_i \cap V_i| \geq (1-\gamma)n \ \forall i$*

In [1], Chin et.al. showed that a simple spectral algorithm can find a γ-correct partition with high probability under certain constraints on a and b, namely that they have to have a large enough gap to guarantee high fidelity community detection. This work was an improvement on previous work proved by [2].

Theorem 1. *There are constants C_0 and C_1 such that the following holds. For any constantsa $> b > C_0$ and $\gamma > 0$ satisfying $\frac{(a-b)^2}{(a+b)} \geq C_1 log\frac{1}{\gamma}$, we can find a γ-correct partition with probability 1-o(1) using a simple spectral algorithm.*

Abbe [3] provided an extensive survey of recent results for SBMs and other block models, including an overview of algorithms for community detection, including spectral and probabilistic MAP techniques and precisely scope the limits of these approaches. Variations of the SBM are also studied in works such as [4] which considers community detection in a sparse hypergraph stochastic block model. [5] proposes a modeling system of dynamic social networks that is a temporal extension of the SBM.

In this paper, we will focus on variations of the simple spectral algorithm described by [1] for alternative constructions of the SBM. We consider three main cases:

- Blocks of unequal size, where n is not uniform between blocks
- Blocks of unequal density, where a is not uniform between blocks
- A combination, where blocks have both different sizes and densities

These proposed variations represent a more realistic variation of the SBM with the goal of enhancing the community detection algorithm's fidelity to alternate cases that are more aligned with real-world networks. Traditional SBM assumes homogeneous block sizes and densities, which may oversimplify the structure of many real networks which are characterized by heterogeneous communities. By accommodating such variability, a modified SBM can better capture the organization of communities within networks, thereby enabling more accurate modeling and inference. Thus, adapting the spectral algorithm to this more realistic variation of SBM promises to enrich both theoretical understanding and practical applications in graph theory.

II. Background

The spectral algorithm of [1] is motivated by the fact that the second eigenvector of the expectation of the adjacency matrix $\mathbb{E}(A_0)$ of the standard SBM separates the blocks. It is a rank 2 matrix with eigenvalues and eigenvectors

$$\lambda_1 = a + b, u_1(i) = \frac{1}{\sqrt{2n}} \forall i \in kn$$

$$\lambda_2 = a - b, u_2(i) = \begin{cases} \frac{1}{\sqrt{2n}} & \text{if } i \in V_1 \\ -\frac{1}{\sqrt{2n}} & \text{if } i \in V_2 \end{cases}$$

They observed that if they can find an approximation of the second eigenvector of $\mathbb{E}(A_0)$ then they can determine which nodes belong to which block based on its entry in the eigenvector. Therefore, the goal of the algorithm is to approximate this second eigenvector u_2 of the expectation of the adjacency matrix $\mathbb{E}(A_0)$. The spectral algorithm consists of first finding the vector space spanned by the top two eigenvectors of the adjacency matrix A_0. Since the first eigenvector of $\mathbb{E}(A_0)$ is the all ones vector, they take the projection of the all ones vector onto this space and approximate the second eigenvector by taking the vector that is orthogonal to this projection. Then, they sort the top k values and assign the corresponding nodes to V_1' and the bottom k to V_2'. Finally, there is a local correction step where nodes are reassigned to the block designated by the majority of their neighbors.

In this community detection problem, this local correction step is so powerful that typically any result that is close enough to the correct partitions can be vastly improved with this extra step. In their original paper [1] showed that this simple method achieves the result proposed in Theorem 1.

For the remainder of this paper, we will consider extensions of this algorithm for variations where the basic assumptions of the SBM are broken, and we will explore how the spectral algorithm can be adapted to work in these cases.

III. Variation 1: Blocks of unequal size

The first variation we consider in this paper is the case where the blocks are not the same size. In this modified SBM, the network is still partitioned into distinct blocks each representing a community. However, the sizes of these blocks can differ, reflecting the heterogeneous nature of real-world networks. We will require an additional parameter to define the network beyond the standard model, in this case the extra size variable. First we can consider the $k = 2$ case where the block sizes will be defined as n and m. The edge probabilities can no longer be defined as $\frac{a}{n}$ and $\frac{b}{n}$ as a result of the change, and we would like to keep the edge densities within communities consistent. Therefore, we introduce p as the probability of within-block edges and q as the probability of between block edges, where $p > q$. See Figure 1 for a visual representation of the expectation matrix in this variation.

Motivated by the previous result of [1] we will first consider the eigenvectors of $\mathbb{E}(A_0)$. Immediately, it becomes clear that the first eigenvector is no longer the all ones vector, and in fact, the first eigenvector separates the blocks of this matrix.

In this case, we will get an eigenvector of $\mathbb{E}(A_0)$ where values associated with vertices in V_1 will have value x and vertices in V_2 will have value y: $\begin{bmatrix} pmx + qny \\ qmx + pny \end{bmatrix} \sim \begin{bmatrix} x \\ y \end{bmatrix}$. Solving for $[x, y]$:

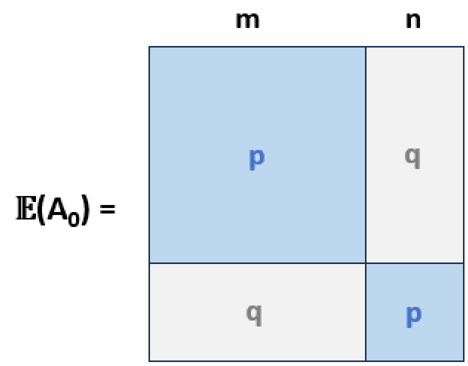

Fig. 1. The expectation of the adjacency matrix in variation 1 where block sizes vary; one block is size m and the other is size n.

$$y(pmx + qny) = x(qmx + pny)$$

Then, without loss of generality we can set $y = 1$ and after some algebra:

$$qmx^2 + p(n - m)x - qn = 0$$

$$x = \frac{p(m - n) \pm \sqrt{p^2(n - m)^2 + 4q^2mn}}{2qm}$$

The eigenvector corresponding to the largest eigenvalue will be the addition case. Since all p, q, n, m are defined by the problem, this eigenvector can be explicitly computed. Note that it will never be 1, unless $n = m$ or $p = q$, both of which we assumed to not be the case.

Now, we don't need to find the vector space spanned by the top two eigenvectors of A_0, we can simply compute the top eigenvector of A_0 and the Davis Kahan sinΘ theorem from [6] [7] allows us to bound the angle between the first eigenvector of $\mathbb{E}(A_0)$ and the first eigenvector of A_0.

The new strategy thus becomes:

1) Remove all high degree rows and columns in the adjacency matrix
2) Compute the first eigenvector of the adjacency matrix
3) Take the top n entries of the eigenvector and assign its corresponding vertices to V_1' and the rest to V_2'
4) Apply local correction by polling the neighborhood of each vertex. If it has more neighbors in the other block, move it to that block.

To prove that this will work with high probability, we define error $E_0 = A_0 - \mathbb{E}(A_0)$. First we must delete rows/columns in A_0 that are outliers with degree $> 20(p + q)(n + m)$ because they disproportionately contribute to the operator norm, causing $||E_0||$ to be too large. Let $A, \mathbb{E}(A)$, and E be the matrices obtained from deleting high degree rows and columns from $A_0, \mathbb{E}(A_0)$, and E_0 respectively. Define $\Delta = \mathbb{E}(A) - \mathbb{E}(A_0)$; then we have:

$$A = \mathbb{E}(A) + E$$

$$A = \mathbb{E}(A_0) + \Delta + E$$

The goal now is to show that Δ and E have a small contribution to the operator norm. It is a simple process to bound Δ, because very few nodes will have high degree.

We use the following theorem, from [1] via a Chernoff bound:

Theorem 2. *There exit a constant d_0 such that if $d \geq d_0$ then with probability $1 - exp(-\Omega((p+q)^{-2}(n+m)))$ not more than $(p+q)^{-3}(n+m)$ vertices have degree $\geq 20d$*

We take $d = (p+q)(n+m)$. From the lemma, there are at most $(p+q)^{-3}(n+m)$ vertices with degree $\geq 20(p+q)(n+m)$, which are precisely the vertices who's corresponding rows and columns in the original adjacency matrix will be recorded in Δ, by definition. Then, Δ has at most $2(p+q)^{-3}(n+m)^2$ non-zero entries, $2(n+m)$ for each vertex, because we zero out the entire row and column in A_0. The magnitude of each entry is at most $p+q$ so taking the norm of Δ, we get:

$$|\Delta|_{HS} \leq \sqrt{\frac{2(n+m)^2}{p+q}} = O(1)$$

This is sufficient because the Operator norm is bounded by the Hilbert-Schmidt norm.

Now, we will bound $|E|$ using the following lemma from [1] adapted from [8] and [9]:

Lemma 3. *Suppose M is a random symmetric matrix with zero on the diagonal whose entries above the diagonal are independent with the following distribution:*

$$M_{ij} = \begin{cases} 1 - p_{ij} & w.p.\ p_{ij} \\ -p_{ij} & w.p.\ 1 - p_{ij} \end{cases}$$

Let σ be a quantity such that $p_{ij} \leq \sigma^2$ and M_1 be the matrix obtained from M by zeroing out all the rows and columns having more than $20\sigma^2 n$ positive entries. Then with probability 1-o(1), $||M_1|| \leq C\sigma\sqrt{n}$ for some constant $C > 0$.

We can see that this structure is exactly the construction for the matrix E. It directly follows that $|E| \leq C\sqrt{(p+q)(n+m)}$ for some C with high probability.

Now we can apply the Davis-Kahan sinΘ Theorem to get:

$$sin(< u_1, u_1') \leq \frac{E + \Delta}{\lambda}$$

If we bound $\lambda = (p-q)(m-n)$ by $C_2\sqrt{(p+q)(n+m)}$

$$sin(< u_1, u_1') \leq \frac{C\sqrt{(p+q)(n+m)}}{C_2\sqrt{(p+q)(n+m)}} = c$$

Assuming this condition on λ, and thus that the gap between $pm+qn$ and $qm+pn$ is sufficiently large, we can see that the angle between the two vectors: the first eigenvector of $\mathbb{E}(A)$ and the first eigenvector of A is small. Therefore, we can use it to approximate the blocks.

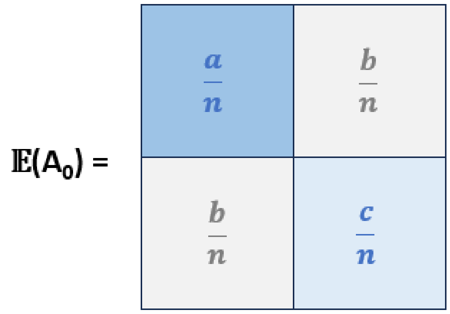

Fig. 2. The expectation of the adjacency matrix in variation 2 where block sizes are the same but the probabilities of edges within blocks differ by block; one block has intra-block edge probability $\frac{a}{n}$ and the other has intra-block edge probability $\frac{c}{n}$.

The rest of the proof follows directly from the [1] paper, once this bound is achieved we can show a high probability of a γ-correct partition reconstruction.

IV. Variation 2: Blocks of Unequal Edge Density

For the second variation, we will focus on blocks of equal size that differ in edge density within blocks. Here, we will denote the size of the blocks as n, just like in the classic SBM setup. We will also return to the $\frac{a}{n}$ and $\frac{b}{n}$ notation for intra- and inter-edge probabilities respectively, but $\frac{a}{n}$ will be the edge probability for just vertices within V_1, and for V_2 we will use probability $\frac{c}{n}$. See Figure 2 for a figure of the expectation matrix for this variation.

In this case, we follow similar steps to compute the eigenvector of the expectation of the adjacency matrix $\mathbb{E}(A_0)$ as we see that:

$$\begin{bmatrix} ax + by \\ bx + cy \end{bmatrix} \sim \begin{bmatrix} x \\ y \end{bmatrix}$$

$$y = 1$$

$$x = \frac{(a-c) \pm \sqrt{(a-c)^2 + 4b^2}}{2b}$$

Which we can easily verify would be 1 in the case that $a = c$. Again, we have that the all ones vector is no longer the first eigenvector, and indeed, the first eigenvector separates the blocks. We can follow a similar proof to the first variation to show that using the same alteration to the spectral algorithm will result in community detection with high probability of correctness.

V. Variation 3: Combined Case

In the combined case, we consider a scenario where the blocks have both unequal sizes and edge distributions. A visual representation of the expectation of the adjacency matrix in this situation is shown in Figure 3. Here, we define our two blocks to have sizes m and n respectively, similar to the first

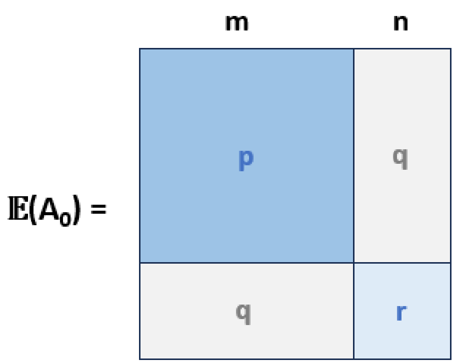

Fig. 3. The expectation of the adjacency matrix in variation 3 where both block sizes and edge probabilities differ. One block is size m with intra-block edge probability p and the other is size n with intra-block edge probability r.

variation. Additionally, similar to the second variation, block one will have edge probability within blocks p and the second block will have probability r within the block. For any two vertices not within the same block, the edge probability will be denoted q.

Following similar steps to the previous two cases, we can compute the eigenvectors of $\mathbb{E}(A)$ to observe that:

$$\begin{bmatrix} pmx + rny \\ qmx + pny \end{bmatrix} \sim \begin{bmatrix} x \\ y \end{bmatrix}$$

$$y = 1$$

$$x = \frac{(pm - rn) \pm \sqrt{(rn - pm)^2 + 4q^2mn}}{2qm}$$

Assuming that $x \neq 1$ then we can show perfect reconstruction using the first eigenvector approximation:

Theorem 4. *If* $mp + qn >> log(n + m)$ *and* $\sqrt{(mp + nq)log(m + n)} << \frac{|(mp+qn)-(mq+rn)|}{2}$ *then the proposed algorithm can partition communities with high probability.*

Proof. We want to show that we classify nodes to communities incorrectly with low probability. Where X_i is an entry in the adjacency matrix in row i, i will be misaligned with probability:

$$Pr(\sum_i X_i \geq \frac{(mp + qn) + (mq + rn)}{2}) =$$

$$Pr(\sum_i X_i \geq (mp + qn) + \frac{-(mp + qn) + (mq + rn)}{2})$$

From the assumption in the theorem statement, this is bounded by the following:

$$<< Pr(\sum_i X_i \geq (mp + qn) + \sqrt{(mp + nq)log(m + n)})$$

Take $\lambda = mp+nq$ and $C = 100\sqrt{log(m + n)}$, for example

$$Pr(\sum_i X_i \geq \lambda + C\sqrt{\lambda})$$

Then, from the Chernoff Inequality stated in [10], we see that this is bounded:

$$Pr(\sum_i X_i \geq \lambda + C\sqrt{\lambda}) \leq exp(-\frac{C^2\lambda}{2(\lambda + C\sqrt{\lambda}/3})$$

Since $C = 100\sqrt{log(m + n)} < \lambda^2$ by our first assumption:

$$Pr(\sum_i X_i \geq \lambda + C\sqrt{\lambda}) \leq exp(\frac{-C^2}{4})$$

$$= exp(-100^2 log(n + m)/4) = (n + m)^{-2500}$$

This probability is very small compared to $n + m$, so it is very probable that perfect reconstruction can be achieved in the general case under the theorem's assumptions.

□

If we additionally employ the correction step, this should remove the log term in the bound via the union bound over all sets of vertices, similarly to the result of [1]. The intuition behind this is that C can be a much larger factor since more vertices are allowed to fail at the step prior to local correction to achieve close to perfect reconstruction guarantees.

However, this general case is more complicated than the previous two, because it is possible to observe a situation where $x = 1$ if the equality: $pm + qn = qm + rn$ holds. In this situation, we would indeed need to explore the second eigenvector algorithm proposed originally by [1].

Therefore, in this combined case we need to employ an adaptive strategy that is robust to this equality becoming close to true. We implemented the algorithm in software and found experimentally that as expected, when the equality is not satisfied, the blocks are easily reconstructed, as shown in Fig 6. Similarly, when the equality is satisfied, the blocks cannot be reconstructed with the first eigenvector approximation and we must use the second eigenvector instead, see Fig 9.

An interesting case arises when the equality is close to true, but not completely satisfied, as shown in the example in Fig 12. If the first eigenvector algorithm is able to find any partition, even if it is incomplete, we can apply iterated correction steps until convergence to further improve the community recovery.

Therefore, the adaptive strategy becomes:

1) Run the approximation algorithm of the first eigenvector.
2) If the algorithm predicts any two communities, no matter their sizes, run the correction step repeatedly until the communities converge.
3) If the algorithm predicts only one community, then we approximate the second eigenvector of the expectation of the adjacency matrix to separate the blocks.

This strategy is robust to general variations of unequal block sizes and densities within the Stochastic Block Model, filling in a gap from the original paper [1].

[b].49

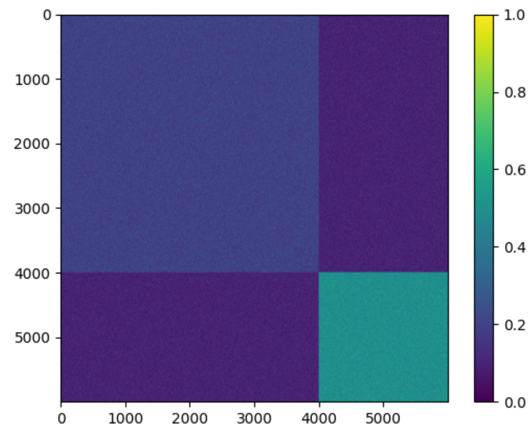

Fig. 4.

[b].49

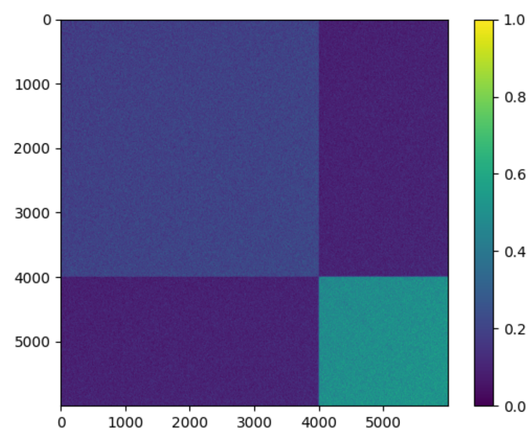

Fig. 5.

Fig. 6. Variation 3 (unequal block size and density) example: a) Original Adjacency Matrix, b) Reconstructed Adjacency Matrix demonstrating a full reconstruction with the first eigenvector approximation method. This example was generated with m=4000, n=2000, q=.1, p=.2, r=.5

VI. Acknowledgment

Distribution Statement 'A' (Approved for Public Release, Distribution Unlimited): The views, opinions, and/or findings expressed are those of the author(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. This Research was developed with funding from the Defense Advanced Research Projects Agency (DARPA).

VII. Conclusion

In this paper, we have provided an overview of a simple spectral algorithm for community detection of the stochastic block model (SBM) and its variants. We have underscored the importance of considering realistic variations of the SBM,

[b].49

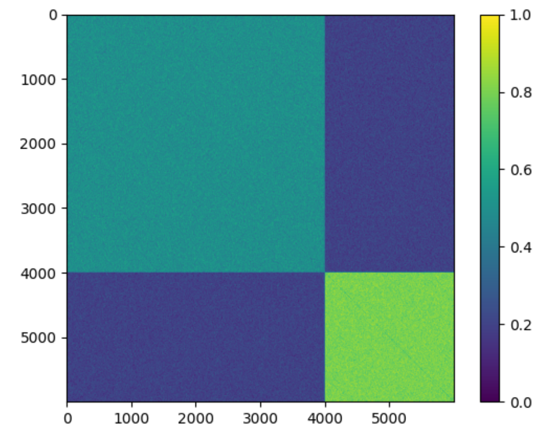

Fig. 7.

[b].49

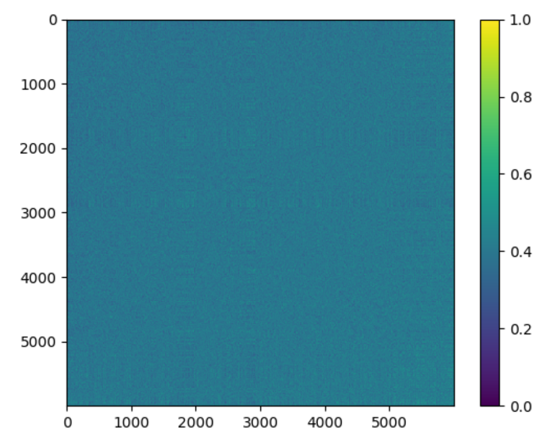

Fig. 8.

Fig. 9. Variation 3 (unequal block size and density) example: a) Original Adjacency Matrix, b) Reconstructed Adjacency Matrix demonstrating a failed reconstruction with the first eigenvector approximation method as the first eigenvector of the expectation of A is the all ones vector. In this case, we will need to use the second eigenvector approximation to separate the blocks. This example was generated with m=4000, n=2000, q=.2, p=.5, r=.8

such as those accommodating unequal block sizes and edge densities. By embracing such natural extensions, we hope to make the classic spectral algorithm more versatile, and to better capture the nuanced structure of real-world networks which often exhibit heterogeneous community organization. This refinement is motivated by the enhancement of the fidelity of SBM-based models to empirical data and also enriches our understanding of the limits of this approach.

We focused on the $k = 2$ case specifically but the algorithms discussed here can easily be extended to the general case. We have not implemented the general case in software yet, but this is something that we would like to complete as a next

[b].49

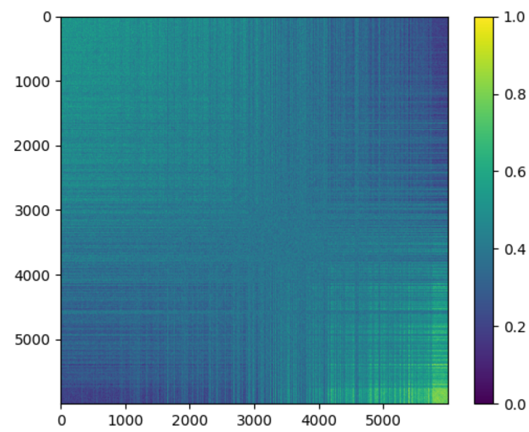

Fig. 10.

[b].49

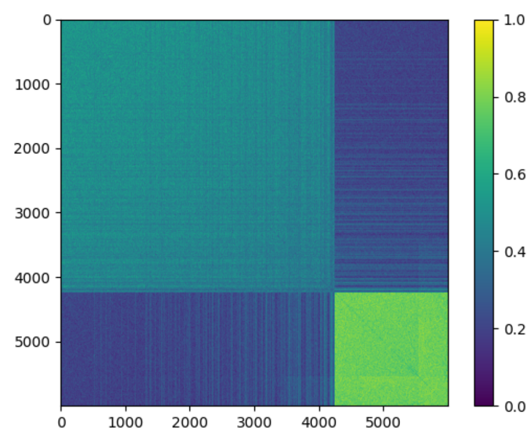

Fig. 11.

Fig. 12. Variation 3 (unequal block size and density) example: a) Reconstruction after one Correction step, b) Reconstruction after two Corrections. In this case, we observe an intermediate stage where the first eigenvector approximation is able to identify a portion of the blocks but not completely successfully. With repeated corrections, the reconstruction improves. This example was generated with m=4000, n=2000, q=.2, p=.5, r=.775

step in this project.

Some potential future work involves considering additional cases, such as when we do not have complete information about the graph. In this paper, we assumed that all of the parameters are known but this is not realistic for community detection tasks over real-world networks. We could consider a case like the third variant proposed here, but where the values of p, q, r, m, and n are not known. The algorithm at its core is still sound under this scenario, but it does have repercussions for specific implementation and will likely affect experimental results.

Additionally, it would be illuminating to explore further practical applications of this method. We implemented and tested the algorithm for the combined case on synthetic graphs, but it would be a natural next step to empirically examine how well this method performs on real-world networks, as the algorithm extension presented here is more aligned than other previous iterations of methods for this problem.

VIII. Citations and References

References

[1] P. Chin, A. Rao, and V. Vu, "Stochastic block model and community detection in sparse graphs: A spectral algorithm with optimal rate of recovery," in *Proceedings of The 28th Conference on Learning Theory*, ser. Proceedings of Machine Learning Research, P. Grünwald, E. Hazan, and S. Kale, Eds., vol. 40. Paris, France: PMLR, 03–06 Jul 2015, pp. 391–423. [Online]. Available: https://proceedings.mlr.press/v40/Chin15.html

[2] A. Coja-Oghlan, "Graph partitioning via adaptive spectral techniques," *Combinatorics, Probability and Computing*, vol. 19, no. 2, p. 227–284, 2010.

[3] E. Abbe, "Community detection and stochastic block models," 2023.

[4] S. Pal and Y. Zhu, "Community detection in the sparse hypergraph stochastic block model," *Random Structures & Algorithms*, vol. 59, no. 3, p. 407–463, Mar. 2021. [Online]. Available: http://dx.doi.org/10.1002/rsa.21006

[5] K. S. Xu and A. O. Hero, "Dynamic stochastic blockmodels for time-evolving social networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 4, p. 552–562, Aug. 2014. [Online]. Available: http://dx.doi.org/10.1109/JSTSP.2014.2310294

[6] C. Davis, "The rotation of eigenvectors by a perturbation," *Journal of Mathematical Analysis and Applications*, vol. 6, no. 2, pp. 159–173, 1963. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0022247X63900015

[7] R. Bhatia, *Matrix Analysis*. Springer, 1997, vol. 169.

[8] J. Friedman, J. Kahn, and E. Szemerédi, "On the second eigenvalue of random regular graphs," in *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing*, ser. STOC '89. New York, NY, USA: Association for Computing Machinery, 1989, p. 587–598. [Online]. Available: https://doi.org/10.1145/73007.73063

[9] U. Feige and E. Ofek, "Spectral techniques applied to sparse random graphs," *Random Structures & Algorithms*, vol. 27, no. 2, pp. 251–275, 2005. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.20089

[10] S. Janson, T. Luczak, and A. Rucinski, *Theory of random graphs*. New York; Chichester: John Wiley & Sons, 2000. [Online]. Available: http://www.amazon.com/Random-Graphs-Svante-Janson/dp/0471175412

[11] M. Heimann, H. Shen, T. Safavi, and D. Koutra, "Regal: Representation learning-based graph alignment," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, ser. CIKM '18. ACM, Oct. 2018. [Online]. Available: http://dx.doi.org/10.1145/3269206.3271788

[12] S. Zhang and H. Tong, "Final: Fast attributed network alignment," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 1345–1354. [Online]. Available: https://doi.org/10.1145/2939672.2939766

[13] R. Singh, J. Xu, and B. Berger, "Global alignment of multiple protein interaction networks with application to functional orthology detection," *Proceedings of the National Academy of Sciences*, vol. 105, no. 35, pp. 12 763–12 768, 2008. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.0806627105

[14] D. Koutra, H. Tong, and D. Lubensky, "Big-align: Fast bipartite graph alignment," in *2013 IEEE 13th International Conference on Data Mining*, 2013, pp. 389–398.

[15] K. Deng, L. Xing, L. Zheng, H. Wu, P. Xie, and F. Gao, "A user identification algorithm based on user behavior analysis in social networks," *IEEE Access*, vol. 7, pp. 47 114–47 123, 2019.

[16] F. Buccafurri, G. Lax, A. Nocera, and D. Ursino, "Discovering links among social networks," in *Machine Learning and Knowledge Discovery in Databases*, P. A. Flach, T. De Bie, and N. Cristianini, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 467–482.

[17] X. Zhou, X. Liang, H. Zhang, and Y. Ma, “Cross-platform identification of anonymous identical users in multiple social media networks,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 2, pp. 411–424, 2016.
[18] T. Man, H. Shen, S. Liu, X. Jin, and X. Cheng, “Predict anchor links across social networks via an embedding approach,” in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, ser. IJCAI’16. AAAI Press, 2016, p. 1823–1829.
[19] L. Liu, W. K. Cheung, X. Li, and L. Liao, “Aligning users across social networks using network embedding,” in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, ser. IJCAI’16. AAAI Press, 2016, p. 1774–1780.
[20] F. Zhou, L. Liu, K. Zhang, G. Trajcevski, J. Wu, and T. Zhong, “Deeplink: A deep learning approach for user identity linkage,” in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 2018, pp. 1313–1321.
[21] Y. Zhang, J. Tang, Z. Yang, J. Pei, and P. S. Yu, “Cosnet: Connecting heterogeneous social networks with local and global consistency,” in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD ’15. New York, NY, USA: Association for Computing Machinery, 2015, p. 1485–1494. [Online]. Available: https://doi.org/10.1145/2783258.2783268
[22] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial networks,” 2014.
[23] S. Abu-El-Haija, B. Perozzi, and R. Al-Rfou, “Learning edge representations via low-rank asymmetric projections,” in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, ser. CIKM ’17. ACM, Nov. 2017. [Online]. Available: http://dx.doi.org/10.1145/3132847.3132959
[24] A. Grover and J. Leskovec, “node2vec: Scalable feature learning for networks,” 2016.
[25] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” 2013.
[26] B. Perozzi, R. Al-Rfou, and S. Skiena, “Deepwalk: online learning of social representations,” in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, ser. KDD ’14. ACM, Aug. 2014. [Online]. Available: http://dx.doi.org/10.1145/2623330.2623732
[27] W. L. Hamilton, R. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” 2018.
[28] C.-Y. Li and S.-D. Lin, “Matching users and items across domains to improve the recommendation quality,” in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD ’14. New York, NY, USA: Association for Computing Machinery, 2014, p. 801–810. [Online]. Available: https://doi.org/10.1145/2623330.2623657