Empirical study of graph spectra and their limitations

Pierre Miasnikof[1] ⋆, Alexander Y. Shestopaloff[2,3], Cristián Bravo[4], and Yuri Lawryshyn[1]

[1] University of Toronto, Toronto, ON, Canada,
[2] Queen Mary University of London, London, United Kingdom,
[3] Memorial University of Newfoundland, St, John's, NL, Canada,
[4] University of Western Ontario, London, ON, Canada

Abstract. We examine the sensitivity of community-structured graph spectra to graph size, block size and inter-block edge probability. We use the Planted Partition Model because of its transparency. While this generative model may seem simplistic, it allows us to isolate the effects of graph and block size, edge probabilities and, consequently, vertex degree distribution on spectra. These sensitivities to key graph characteristics also generalize beyond Planted Partition Model graphs, because they are based on graph structure. Notably, our results show that eigenvalues converge to those of a complete graph, with increases in graph size or inter-block edge probability. Such convergence severely limits the use of spectral techniques.

Keywords: Spectral graph theory, eigengap, community detection

1 Introduction

Graphs have several matrix representations. The adjacency and the several Laplacian matrices are instances of these representations. Spectral decomposition of these matrices is used for many tasks in the study of graphs, especially those with community structure, and of complex networks [5]. For example, it is used for vertex clustering [4, 7, 10, 11, 19, 21, 22] and in the "eigengap heuristic" [4, 19]. In this short article, we identify the limitations of spectral decomposition of these graph matrix representations. In doing so, we also compare results obtained by decomposing the two most commonly used matrix representations, the adjacency and the symmetric normalized Laplacian matrices.

Our empirical investigations reveal that normalized Laplacian eigenvalues (eigenvectors) are extremely sensitive to noise and scale. This noise manifests itself in the form of edge randomness. This randomness is a consequence of inter-block edge probability or small block size. We find that increases in noise or graph size lead to eigenvalue uniformity, which renders spectral methods of limited use. Indeed, spectral techniques rely on eigenvalue differentiability. Obviously,

⋆ Corresponding author: p.miasnikof@mail.utoronto.ca

as eigenvalues move to uniformity, differentiability vanishes. Unfortunately, we also find the adjacency matrix does not provide a viable alternative in cases where Laplacian spectra are uninformative.

We investigate the behavior of the spectra of graphs displaying community structure, in order to help target the application of spectral techniques to instances where they are more likely be informative and to yield meaningful results. The work in this article is focused on the link between a few key graph properties and those of the corresponding symmetric normalized Laplacian matrix. Although we focus our attention on the widely used normalized symmetric Laplacian matrix, we also examine the eigenvalues of the adjacency matrix, for the sake of completeness.

Of course, spectral techniques are also limited by computational constraints (e.g., memory, precision, computation times, etc...). However, in this article, we do not examine these computational issues. Nevertheless, we do note that the limitations identified in this work are amplified by the various computational and tractability constraints. For the sake of brevity, the scope of this article is also restricted to simple graphs, unweighted undirected graphs with no self-loops or multiple edges and graph with only a single connected component.

2 Previous work

As mentioned earlier, spectral techniques are widely used in the study of graphs. The foundations of this area of study were laid by Chung [6] and later by Spielman [24]. In this short work, we focus on three areas of the spectral graph analysis literature. First, our experiments motivate us to examine past inquiries into the convergence of eigenvalues, under the stochastic block model (SBM) [23]. Then, in order to gain a better understanding of this convergence, we also survey studies in which authors have established a link between Laplacian eigenvalues and vertex degree [5, 27]. Finally, we are also motivated by authors who have highlighted the differences in the conclusions of analyses based on adjacency and Laplacian matrix representations [18, 21].

Asymptotic convergence of eigenvectors (and consequently eigenvalues) under the SBM was identified by Rohe et al. [23]. These authors posit the existence of a "population Laplacian" towards which the empirical Laplacian converges, as the graph grows. While our results do not agree with these authors' conclusions, we also document convergence with increased graph size and noise in connectivity.

Under random graph models, like the SBM or the Erdös-Rényi-Gilbert (ERG) model [9, 13], edge probabilities are independent of each other and only depend on the nodes they are connecting. These models have often been described as too simplistic to represent real world networks [1, 2, 15, 20, 23]. In particular, the degree uniformity yielded by these generative models, its lack of skewness or heavy right tail, has been identified as a weakness as models of real world networks.

Nevertheless, random graph models have been found to be adequate in many empirical cases [10, 11, 16, 20, 22]. For example, Newman et al. [20] state that *"in some cases random graphs with appropriate distributions of vertex degree*

predict with surprising accuracy the behavior of the real world". In closing, while a detailed examination of this debate on realistic models of real world networks is beyond the scope of our work, we do note that some authors claim that networks with power-law degree distributions are rare (e.g., [3]). We also note that the SBM is still used as a model of real world networks in the recent literature (e.g., [12])

3 Mathematical background

As discussed earlier, there are several matrix representations of graphs. The two most commonly used are the adjacency matrix (A) and the symmetric normalized Laplacian ($\mathcal{L}$). Because of its symmetry and specific properties, we use the symmetric normalized Laplacian, instead of the unnormalized or random walk Laplacians.

In order to study the link between graph characteristics and spectra, we generate several synthetic random graphs with known structure. Indeed, by modifying the parameters of random graph generative models, we are able to isolate and unambiguously observe the sensitivities of the spectra. We use the Planted Partition Model (PPM), a special case of the SBM, for its clarity. Using the PPM also allows us to compare our conclusions to those reported in the literature (e.g., [23]). We use the Python NetworkX library [14, 26, 25] to generate our graphs.

For the remainder of this article, we will use the following naming conventions:

- $N = |V|$ is the total number of vertices,
- n_k is the number of vertices in block k,
- d_i is the degree of the $i-$th vertex and
- $\lambda_0 = 0 < \lambda_1 \leq \ldots \leq \lambda_N$ are the eigenvalues of the $(N \times N)$ normalized symmetric matrix $\mathcal{L}$,
- Because we only consider graphs with one connected component, only the first eigenvalue is equal to 0, all others are strictly positive.

3.1 Matrices under consideration

As mentioned earlier, a graph can be represented by various matrices. In this work, we focus on two, the adjacency matrix (A) and the normalized symmetric Laplacian ($\mathcal{L}$). The latter is a simple transformation of the former.

Adjacency matrix The adjacency matrix A for a graph with N vertices is defined as

$$\begin{aligned} \underset{(N\times N)}{A} &= [w_{ij}] \\ &\text{with,} \\ w_{ij} &\in \{0,1\} \text{ and} \\ w_{ii} &= 0, \forall i\,. \end{aligned}$$

Throughout this article, the matrix A is symmetric and all matrix elements (edge weights w_{ij}) are binary (i.e., graphs are unweighted & undirected). Also, because we only consider simple graphs, all diagonal elements are equal to zero (i.e., no self-loops $\Leftrightarrow A_{ii} = w_{ii} = 0$).

Normalized symmetric Laplacian The (un-normalized) Laplacian matrix is defined as

$$L = D - A$$

where,

$$D = \begin{bmatrix} d_1 & 0 & \dots \\ 0 & \ddots & \dots \\ 0 & \dots & d_n \end{bmatrix} .$$

Here, d_i is the degree of the $i-th$ node. It is obtained by summing the $i-th$ row of the adjacency matrix:

$$d_i = \sum_j A_{ij} .$$

The corresponding symmetric normalized Laplacian $\mathcal{L}$ is defined as

$$\begin{aligned} \mathcal{L} &= D^{-1/2} L D^{-1/2} \\ &= D^{-1/2}(D - A)D^{-1/2} \\ &= I - D^{-1/2} A D^{-1/2} . \end{aligned}$$

(by convention, $\frac{1}{\sqrt{d_i}} = 0$, if $d_i = 0$)

The spectra of normalized symmetric Laplacian matrices have several appealing properties. Symmetry, which guarantees real-number eigenvalues, is the most obvious. However, there are several other very informative features. In the next section, we will review the most important of these properties.

3.2 Eigenvalues of normalized symmetric Laplacian matrices

While the topic of spectral graph theory is very vast, here we only review the few properties of the eigenvalues which we use in this work. For a thorough treatment of this topic, we refer the reader to Fan Chung's seminal text "Spectral graph theory" [6]. Here, we assume a graph with only one connected component.

- Naturally, all eigenvalues of the normalized symmetric Laplacian are real.
- The normalized symmetric Laplacian is positive semidefinite, all its eigenvalues are non-negative: $\lambda_i \geq 0,\ \forall i$.
- Since all graphs considered in this work have only one connected component, in this work the following properties hold:
  - $\diamond\ \sum_i \lambda_i = |V| = N$,
  - $\diamond\ \lambda_0 = 0 < \lambda_1 \leq \ldots \leq 2$ (i.e., only one zero eigenvalue).
- In a complete graph, $\lambda_i = N/(N-1) \gtrapprox 1,\ \forall i > 0$.

3.3 Eigengap heuristic

Very commonly, the spectra of normalized symmetric Laplacians are used to obtain the so-called "eigengap heuristic" [4, 19]. This heuristic provides a reliable estimate of the number of clusters in a graph. Many graph clustering techniques that require the number of clusters as input rely on this technique.

The eigengap heuristic consists of the following steps: (one connected component case)

- compute eigenvalues of the normalized symmetric Laplacian matrix,
- sort the eigenvalues in ascending order, i.e., $\lambda_0 < \lambda_1 \leq \ldots \leq \lambda_{N-1}$,
- the number of communities on the graph (K) is approximately equal to the index at which the eigenvalues (of the Laplacian) display their first spike.

In other words, this heuristic tells us: $\lambda_i \ll \lambda_{i+1} \Rightarrow K \approx i$
(where, $i \in \{1, \ldots, N-1\}$).

In our empirical experiments, we use PPM graphs with a known number of blocks (i in the case above). To examine the sensitivities of the eigengap, we compute the ratios $\left(\frac{\lambda_i}{\lambda_{i+1}}\right)$.

3.4 Graphs under consideration

While the ERG model is the prototypical random graph generative model, there exist other models that yield less trivial structures. The Stochastic Block Model (SBM) and a special case of it called the Planted Partition Model (PPM) [8, 10, 11] are two such instances. These two models are often used to generate examples of networks with community structure [10, 11]. Indeed, such graphs are composed of blocks (communities, clusters) of vertices that are densely connected while only being sparsely connected to the remaining graph.

The PPM is a special case of the more general SBM. Under the SBM each block of vertices has its own within-block edge probability, each block pair has its own inter-block edge probability and blocks have varying numbers of vertices. In the PPM, all blocks have the same number of nodes, while within-block and between-block edge probabilities are fixed for all blocks and block pairs.

Vertex degrees under SBM For an arbitrary vertex i belonging to a block k containing n_k nodes, the expected degree is

$$\mathbf{E}(d_i) = \underbrace{P_k \times (n_k - 1)}_{\text{expected intra-block edges}} + \underbrace{\sum_{m \neq k} P_{km}(n_k \times n_m)}_{\text{expected inter-block edges}} . \tag{1}$$

Here, P_k denotes the probability that two arbitrary nodes in block k are connected by an edge. Similarly, P_{km} denotes the probability that an arbitrary node in block k is connected to an arbitrary node in block m. (With, $P_k > P_{km}$, typically.)

Vertex degree under PPM For the remainder of this document, we will use the following variable naming conventions to describe the PPM graphs:

- $d_i = d_i^{in} + d_i^{out}$ is the degree of a node i.
- d_i is the sum of connections to nodes within the same block (d_i^{in}) and to nodes in other blocks (d_i^{out}).
- P_{in}, P_{out} are the within/between-block edge probabilities,
- N is the total number of vertices,
- n is the number of vertices within any given block and, finally,
- $K = N/n$ is the total number of blocks or partitions (under the PPM, all blocks have the same number of nodes).

$$\begin{aligned}\mathbf{E}(d_i) &= \underbrace{P_{in} \times (n-1)}_{\mathbf{E}(d_i^{in})} + \underbrace{P_{out} \times (N-n)}_{\mathbf{E}(d_i^{out})} \\ &= \underbrace{P_{in} \times (n-1)}_{\mathbf{E}(d_i^{in})} + \underbrace{P_{out} \times N\left(1 - \frac{1}{K}\right)}_{\mathbf{E}(d_i^{out})} \end{aligned} \tag{2}$$

In Equations 1 and 2, we clearly see how the (expected) degree of a vertex can be partitioned. Degree can be understood as the cardinality of the union of a set of connections to nodes within the same block and to nodes on the remaining graph. We use this partition to examine the sensitivity of graph spectra to overall degree, but also specific generative model parameters. Specifically, we examine the relationships between spectra and graph size, block size and inter-block edge probability. The analysis of these relationships is a useful tool in understanding the applicability and limitations of spectral graph techniques. While we use the PPM for its transparency, our conclusions reveal critical information about the relationship between graph structures and spectra. This relationship transcends generative models, because they study the links between graph structure (esp. degree) and spectra.

4 Empirical tests

To examine the sensitivity of graph spectra to graph size, block size and to noise from increased inter-block edge probability, we conduct four sets of experiments using the Planted Partition Model. In all experiments, we begin with a graph with small block size, number of blocks or inter-block probability. We then gradually increase these parameters and observe the effect on spectra. We stop our sensitivity tests, when a pattern appears (or disappears, e.g., eigengap) in the spectra. For completeness, we examine the spectra of both the symmetric normalized Laplacian and adjacency matrices. While most studies of graph spectra examine the normalized symmetric Laplacian (e.g., [6, 19, 24]), the adjacency matrix remains the most basic matrix representation of a graph.

In our experiments, growth in size occurs in two different ways. In the first case, growth occurs in the sizes of blocks, while the number of blocks remains constant. We start with a graph containing 50 blocks of five nodes each and then expand to graphs with 50 blocks of 50 and 500 nodes. In the second case, growth occurs in the numbers of blocks, while sizes of blocks remains constant. We begin with a graph of five blocks of 50 nodes and expand to 50 blocks and 500 blocks of 50 nodes. These results are presented in Section 4.1. We also isolate the effect of block size, within a fixed size graph. The goal of these numerical experiments is to identify the (in)ability of spectra to detect the presence of densely connected blocks of varying sizes, within a sparser graph with identical characteristics (edge probability, number of block and size). These results are presented in Section 4.2. In all the above-mentioned experiments, intra-/inter–block edge probabilities are held constant ($P_{in} = 0.9, P_{out} = 0.1$).

Our last batch of experiments, is an examination sensitivity to edge probability. In these experiments, we generate PPM graphs with intra-block edge probability of 0.9 but with varying inter-block edge probabilities. Here again, we isolate the effect of block size. We begin by generating graphs with a relatively large block size ($n = 500$) and relatively small number of blocks ($K = 50$). We then repeat the same experiments with graphs containing a relatively large number ($K = 500$) of relatively small ($n = 50$) blocks.

4.1 Sensitivity to graph size

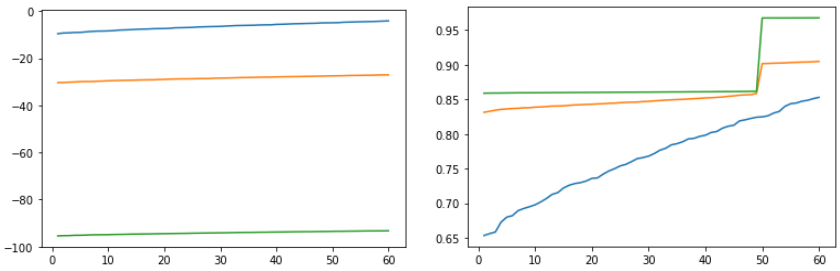
In this first set of experiments, we vary graph size by increasing block size (n), while keeping the number of blocks constant ($K = 50$). We compute the eigenvalues for the adjacency and normalized Laplacian matrices for graphs with:

- $n = 5 \Rightarrow N = 5 \times 50 = 250$,
- $n = 50 \Rightarrow N = 50 \times 50 = 2,500$,
- $n = 500 \Rightarrow N = 500 \times 50 = 25,000$,
- $P_{in} = 0.9, P_{out} = 0.1$.

Results are shown in Figure 1. The blue curve shows the $(N - 1)$ non-zero eigenvalues, sorted in ascending order, for the graph with 250 nodes. The orange curve shows the same, for the graph with $2,500$ nodes. Finally, the green curve shows the eigenvalues of the graph with $25,000$ nodes. In order to focus on the theoretical location of the eigengap, we adjust the x-axis accordingly.

In a separate set of experiments, we generate graphs with the same characteristics as those in Figure 1. We then record the range of eigenvalues and the eigengap of the normalized Laplacian as the graph grows in size. Results are reported in Table 1.

In our second set of experiments, we vary graph size by increasing the number of blocks (K), while keeping the block size constant ($n = 50$). It has been argued that this growth model is more realistic and consistent with real world networks [17, 23] We compute the eigenvalues for the adjacency and normalized Laplacian matrices for graphs with:



(a) Adj mat (b) Lap mat

Fig. 1: Varying N, number of blocks is constant ($K = 50$)
(blue 250 nodes, orange 2, 500 nodes, green 25, 000 nodes)

Table 1: Eigenvalue range, number of blocks is constant ($K = 50$)

	$N = 250$	$N = 2500$	$N = 25K$
Min	0.66	0.83	0.86
Max	1.34	1.11	1.03
Eigengap	1.00	1.05	1.12

- $K = 5 \Rightarrow N = 5 \times 50 = 250$,
- $K = 50 \Rightarrow N = 50 \times 50 = 2,500$,
- $K = 500 \Rightarrow N = 500 \times 50 = 25,000$,
- $P_{in} = 0.9, P_{out} = 0.1$.

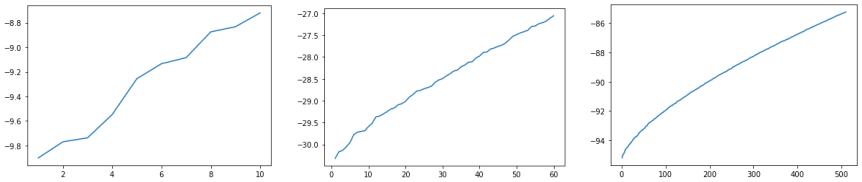
Results are shown in Figure 2. Here too, in order to focus on the theoretical location of the eigengap, we adjust the x-axis accordingly.

Once more, we also generate a new set of graphs with the same characteristics as those in Figure 2. We record the range of eigenvalues and the eigengap of the normalized Laplacian as the graph grows in size. Results are reported in Table 2.

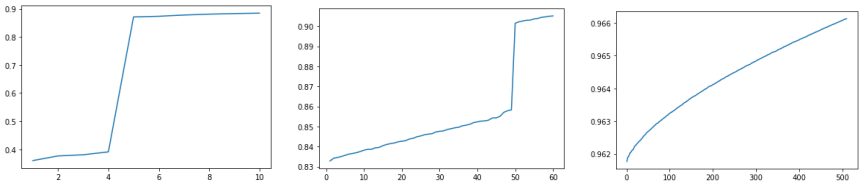
Table 2: Eigenvalue range, block size is constant ($n = 50$)

	$N = 250$	$N = 2500$	$N = 25K$
Min	0.35	0.83	0.96
Max	1.16	1.11	1.04
Eigengap	2.31	1.05	1.00

Results from all four experiments in this section highlight the relationships between Laplacian eigenvalues and graph and block sizes. Our observations are consistent with and extend prior work that has linked vertex degree and spectra [27, 5]. In particular, we isolate the effect of increases in graph and block sizes on



(a) Adj mat, 250 nodes (b) Adj mat, 2500 nodes (c) Adj mat, 25K nodes



(d) Lap mat, 250 nodes (e) Lap mat, 2500 nodes (f) Lap mat, 25K nodes

Fig. 2: Varying N, number of blocks varies, block size constant ($n = 50$)

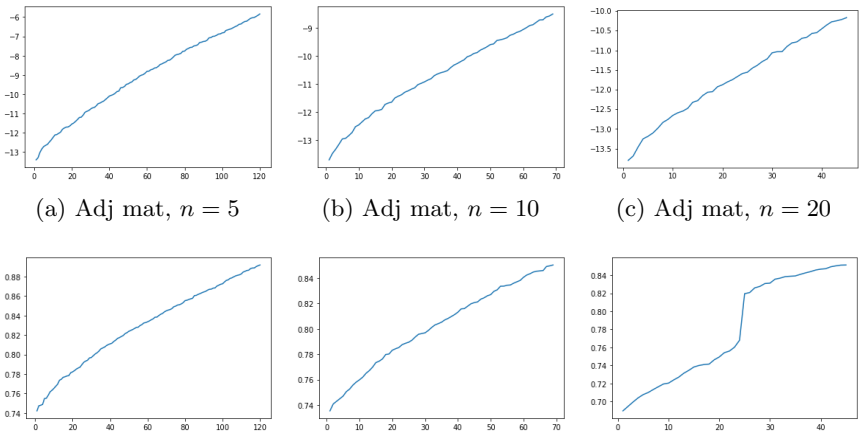
degree and, consequently, eigenvalues, as shown in Equations 1 and 2. Indeed, Figure 1 and Table 1 reveal that while block size may remain a constant proportion of graph size, blocks with a small number of nodes are undetectable. For example, in the first column of Table 1 blocks have only five nodes. Meanwhile, in the third column, blocks have 500 nodes. In the first column, no eigengap is detected, while in the second and third columns, where blocks are larger in absolute terms, a small eigengap is present. Nevertheless, we do note a significant monotonic narrowing of the range of eigenvalues with increases in graph size.

The narrowing of the range of eigenvalues is marked even more in the case of a graph with constant block size, as seen in Figure 2 and Table 2. In these same experiments, we also observe a vanishing eigengap with increases in graph sizes. Here, it is important to note that these experiments follow a pattern of network growth has been found to be more realistic and consistent with real world networks [17, 23]. Clearly, these results highlight the limitations of spectral techniques in the case of large real world networks.

4.2 Sensitivity to block size

To further isolate the effect of block size, we keep the number of nodes constant ($N = 500$), but vary block size ($n \in \{5, 10, 20\}$). Once again, we adjust the x-axis to focus on the eigengap. Results are shown in Figure 3.

Once again, we observe that smaller block sizes lead to increased uniformity in eigenvalues. In fact, the eigengap is non-existent, except in the very last experiment ($n = 20$).



(a) Adj mat, $n = 5$ (b) Adj mat, $n = 10$ (c) Adj mat, $n = 20$

(d) Lap mat, $n = 5$ (e) Lap mat, $n = 10$ (f) Lap mat, $n = 20$

Fig. 3: Sensitivity to block size (n), graph size constant ($N = 500$)

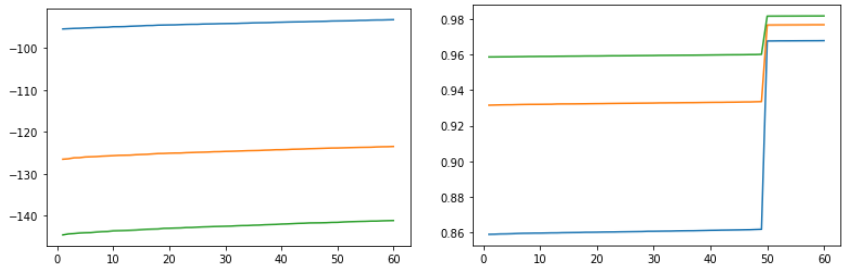
4.3 Sensitivity to inter-block edge probability

In this set of experiments, we keep the intra-block edge probability fixed ($P_{in} = 0.9$). To simulate a noisy network, we vary inter-block edge probability ($P_{out} \in \{0.1, 0.2, 0.3\}$). In the experiments shown in Figure 4, the graphs contain $K = 50$ blocks of $n = 500$ vertices. Meanwhile, for the ones in Figure 5, the graphs contain $K = 500$ blocks of $n = 50$ nodes. Arguably, this latter case is more representative of real world networks, which typically have smaller blocks (clusters, communities) [17, 23]. In both figures, the blue curve represents the eigenvalues of graph with $P_{out} = 0.1$, the orange curve is for a graph with $P_{out} = 0.2$ and the green curve is for a graph with $P_{out} = 0.3$. Finally, as in Section 4.1, we also reproduce the experiments in our figures (Figures 4 and 5) and record the eigenvalue ranges and eigengap ratios. These results are shown is Tables 3 and 4.

Once again, these experiments highlight the link between degree and spectra. Specifically, these experiments highlight the link between increases in P_{out} and consequently inter-block degree and spectra. In all experiments, we observe narrowing eigenvalue ranges and vanishing eigengaps.

5 Conclusion and future work

In this article, we identify the limitations of spectral graph analysis. We show that eigenvalues are sensitive to noise in connectivity and converge to uniformity as graph size increases. This noise is a function of both block size and inter-block edge probability. While we use the PPM to illustrate these sensitivities, we argue that our conclusions extend to other generative models as well. Indeed,

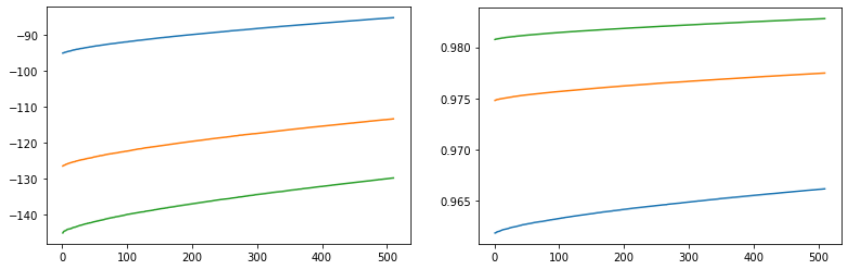


(a) Adj mat (b) Lap mat

Fig. 4: Increasing inter-block edge probability ($N = 25,000, K = 50, n = 500$)
(blue $P_{out} = 0.1$, orange $P_{out} = 0.2$, green $P_{out} = 0.3$)

Table 3: Eigenvalue range, with increasing inter-block edge probability ($N = 25,000, K = 50, n = 500$)

	$P_{out} = 0.1$	$P_{out} = 0.2$	$P_{out} = 0.3$
Min	0.86	0.93	0.96
Max	1.03	1.02	1.02
Eigengap	1.12	1.05	1.02



(a) Adj mat (b) Lap mat

Fig. 5: Increasing inter-block edge probability ($N = 25,000, K = 500, n = 50$)
(blue $P_{out} = 0.1$, orange $P_{out} = 0.2$, green $P_{out} = 0.3$)

Table 4: Eigenvalue range, with increasing inter-block edge probability ($N = 25,000, K = 500, n = 50$)

	$P_{out} = 0.1$	$P_{out} = 0.2$	$P_{out} = 0.3$
Min	0.96	0.97	0.98
Max	1.04	1.03	1.02
Eigengap	1.00	1.00	1.00

vertex degree, edge probability and block (community) sizes are variables that are present in all graphs with clustered structure, regardless of generative model. Naturally, graphs with power law degree distributions are less sensitive to the variations discussed in this paper. However, we reserve a detailed examination of their spectra for future work.

On the basis of our numerical experiments, we recommend against the use of spectral techniques for large graphs or in cases where blocks (communities) are expected to be small. Our future work will focus on identifying clear thresholds for the applicability of spectral techniques.

Acknowledgments

– The work of P.M. was funded by a MITACS grant (grant# IT33832).
– The authors thank Prof. Valery A. Kalyagin of the Laboratory of Algorithms and Technologies for Networks Analysis in Nizhny Novogorod, Russia, for his comments on the early stages of this work.

References

1. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Reviews of Modern Physics **74**, 47–97 (2002). DOI 10.1103/RevModPhys.74.47
2. Barabási, A.L., Albert, R.: Emergence of Scaling in Random Networks. Science **286**, 509–512 (1999)
3. Broido, A.D., Clauset, A.: Scale-free networks are rare. Nature communications **10**(1), 1017 (2019)
4. Bruneau, P., Parisot, O., Otjacques, B.: A heuristic for the automatic parametrization of the spectral clustering algorithm. In: 2014 22nd International Conference on Pattern Recognition, pp. 1313–1318 (2014). DOI 10.1109/ICPR.2014.235
5. Chen, J., Lu, J., Zhan, C., Chen, G.: Laplacian Spectra and Synchronization Processes on Complex Networks, pp. 81–113. Springer US, Boston, MA (2012). DOI 10.1007/978-1-4614-0754-6_4. URL https://doi.org/10.1007/978-1-4614-0754-6_4
6. Chung, F.R.K.: Spectral graph theory. American Mathematical Soc. (1997)
7. Coja-Oghlan, A., Goerdt, A., Lanka, A.: Spectral partitioning of random graphs with given expected degrees. In: G. Navarro, L. Bertossi, Y. Kohayakawa (eds.) Fourth IFIP International Conference on Theoretical Computer Science- TCS 2006, pp. 271–282. Springer US, Boston, MA (2006)
8. Condon, A., Karp, R.: Algorithms for graph partitioning on the planted partition model. Random Struct. Algorithms **18**(2), 116–140 (2001). DOI 10.1002/1098-2418(200103)18:2⟨116::AID-RSA1001⟩3.0.CO;2-2. URL http://dx.doi.org/10.1002/1098-2418(200103)18:2⟨116::AID-RSA1001⟩3.0.CO;2-2
9. Erdös, P., Rényi, A.: On Random Graphs I. Publicationes Mathematicae Debrecen **6**, 290–297 (1959)
10. Fortunato, S.: Community detection in graphs. Physics Reports **486**, 75–174 (2010). DOI 10.1016/j.physrep.2009.11.002
11. Fortunato, S., Hric, D.: Community detection in networks: A user guide. arXiv (2016)

12. Gan, L., Wan, X., Ma, Y., Lev, B.: Efficiency evaluation for urban industrial metabolism through the methodologies of emergy analysis and dynamic network stochastic block model. Sustainable Cities and Society p. 104396 (2023)
13. Gilbert, E.: Random graphs. Ann. Math. Statist. **30**(4), 1141–1144 (1959). DOI 10.1214/aoms/1177706098. URL https://doi.org/10.1214/aoms/1177706098
14. Hagberg, A., Schult, D., Swart, P.: Exploring Network Structure, Dynamics, and Function using NetworkX. In: G. Varoquaux, T. Vaught, J. Millman (eds.) Proceedings of the 7th Python in Science Conference, pp. 11–15. Pasadena, CA USA (2008)
15. Lancichinetti, A., Fortunato, S., Radicchi, F.: Benchmark graphs for testing community detection algorithms. arXiv **78**(4), 046110 (2008). DOI 10.1103/PhysRevE.78.046110
16. Lee, C., Wilkinson, D.J.: A review of stochastic block models and extensions for graph clustering. Applied Network Science **4**(1), 1–50 (2019)
17. Leskovec, J., Lang, K.J., Dasgupta, A., Mahoney, M.W.: Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. Internet Mathematics **6**(1), 29–123 (2009). DOI 10.1080/15427951.2009.10129177
18. Lutzeyer, J.F., Walden, A.T.: Comparing Graph Spectra of Adjacency and Laplacian Matrices. arXiv e-prints arXiv:1712.03769 (2017). DOI 10.48550/arXiv.1712.03769
19. von Luxburg, U.: A tutorial on spectral clustering. Statistics and computing **17**, 395–416 (2007)
20. Newman, M.E.J., Strogatz, S., Watts, D.J.: Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E **64**, 026,118 (2001). DOI 10.1103/PhysRevE.64.026118. URL https://link.aps.org/doi/10.1103/PhysRevE.64.026118
21. Priebe, C.E., Park, Y., Vogelstein, J.T., Conroy, J.M., Lyzinski, V., Tang, M., Athreya, A., Cape, J., Bridgeford, E.: On a two-truths phenomenon in spectral graph clustering. Proceedings of the National Academy of Sciences **116**(13), 5995–6000 (2019)
22. Rao Nadakuditi, R., Newman, M.E.J.: Graph spectra and the detectability of community structure in networks. arXiv e-prints arXiv:1205.1813 (2012). DOI 10.48550/arXiv.1205.1813
23. Rohe, K., Chatterjee, S., Yu, B.: Spectral clustering and the high-dimensional stochastic blockmodel. The Annals of Statistics **39**(4), 1878–1915 (2011). DOI 10.1214/11-AOS887. URL https://doi.org/10.1214/11-AOS887
24. Spielman, D.A.: Spectral graph theory and its applications. In: 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07), pp. 29–38 (2007). DOI 10.1109/FOCS.2007.56
25. documentaton page (author unknown), O.: Planted partition model. https://networkx.org/documentation/stable/reference/generated/networkx.generators.community.planted_partition_graph.html
26. documentaton page (author unknown), O.: Stochastic block model. https://networkx.org/documentation/stable/reference/generated/networkx.generators.community.stochastic_block_model.html
27. Zhan, C., Chen, G., Yeung, L.F.: On the distributions of Laplacian eigenvalues versus node degrees in complex networks. Physica A: Statistical Mechanics and its Applications **389**(8), 1779–1788 (2010). DOI https://doi.org/10.1016/j.physa.2009.12.005. URL https://www.sciencedirect.com/science/article/pii/S0378437109010012