

---

# SPONGE: A generalized eigenproblem for clustering signed networks

---

**Mihai Cucuringu**  
University of Oxford  
The Alan Turing Institute

**Peter Davies**  
University of Warwick

**Aldo Glielmo**  
King's College London

**Hemant Tyagi**  
University of Edinburgh  
The Alan Turing Institute

## Abstract

We introduce a principled and theoretically sound spectral method for  $k$ -way clustering in signed graphs, where the affinity measure between nodes takes either positive or negative values. Our approach is motivated by social balance theory, where the task of clustering aims to decompose the network into disjoint groups such that individuals within the same group are connected by as many positive edges as possible, while individuals from different groups are connected by as many negative edges as possible. Our algorithm relies on a generalized eigenproblem formulation inspired by recent work on constrained clustering. We provide theoretical guarantees for our approach in the setting of a signed stochastic block model, by leveraging tools from matrix perturbation theory and random matrix theory. An extensive set of numerical experiments on both synthetic and real data shows that our approach compares favorably with state-of-the-art methods for signed clustering, especially for large number of clusters.

**Keywords:** signed clustering, graph Laplacians, stochastic block models, spectral methods, time series, clustering correlation matrices

## 1 Introduction

Clustering is a popular unsupervised learning task aimed at extracting groups of nodes in a weighted graph in such a way that the average connectivity or similarity between pairs of nodes within the same group is larger than that of pairs of nodes from different groups. While most of the literature has focused on clustering graphs where the edge weights

are non-negative, the task of clustering signed graphs (whose edge weights can take negative values as well) remained relatively unexplored, and has recently become an increasingly important research topic [36].

The motivation for recent studies arose from a variety of examples from social networks where users express relationships of trust-distrust or friendship-enmity, online news and review websites such as Epinions [1] and Slashdot [3] that allow users to approve or denounce others [35], and shopping bipartite networks encoding like-dislike preferences between users and products [7].

Another application stems from time series analysis, in particular clustering time series [4], a task broadly used for analyzing gene expression data in biology [22], economic time series that capture macroeconomic variables [21], and financial time series corresponding to large baskets of instruments in the stock market [59, 44]. In such contexts, a popular similarity measure in the literature is given by the Pearson correlation coefficient that measures linear dependence between variables and takes values in  $[-1, 1]$ . By interpreting the correlation matrix as a weighted network whose (signed) edge weights capture the pairwise correlations, we cluster the multivariate time series by clustering the underlying signed network. To increase robustness, tests of statistical significance are often applied to individual pairwise correlations, leading to sparse networks after thresholding on the p-value associated to each individual sample correlation [26]. We refer the reader to the popular work of Smith et al. [50] for a detailed survey and comparison of various methods for turning time series into networks, where the authors explore the interplay between fMRI time series and the network generation process. Most importantly, they conclude that in general correlation-based approaches can be quite successful at estimating the connectivity of brain networks from fMRI time series.

**Contributions.** Our contributions are as follows.

- We propose a regularized spectral algorithm for clustering signed graphs that is based on solving a generalized eigenproblem. Our approach is scalable and compares favorably to state-of-the-art methods.

- We provide a detailed theoretical analysis of our algorithm with respect to its robustness against sampling sparsity and noise level, under a Signed Stochastic Block Model (SSBM).
- To our knowledge, we provide the first theoretical guarantees – in the SSBM framework – for the Signed Laplacian method introduced in the popular work of Kunegis et al. [34] for clustering signed graphs.
- Finally, we provide extensive numerical experiments on both synthetic and real data, showing that our algorithm compares favourably to state-of-art methods. In particular, it is able to recover partitions in the regime where the graph is sparse and the number of clusters  $k$  is large, where existing methods completely fail.

**Paper outline.** The remainder of this paper is organized as follows. Section 2 is a summary of related work from the signed clustering literature. Section 3 formulates our SPONGE (Signed Positive Over Negative Generalized Eigenproblem) algorithm for clustering signed graphs. Section 4 introduces the Signed Stochastic Block Model (SSBM) and contains our theoretical analysis of SPONGE in the SSBM. Section 5 contains a similar theoretical analysis for signed spectral clustering via the Signed Laplacian. Section 6 contains numerical experiments on various synthetic and real data sets. Finally, Section 7 summarizes our findings along with future research directions.

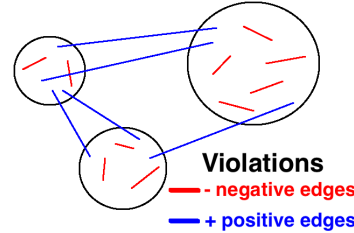
**Notation.** For a matrix  $A \in \mathbb{R}^{n \times n}$ , we denote its eigenvalues and eigenvectors by  $\lambda_i(A)$  and  $v_i(A)$  respectively,  $\forall i = 1, \dots, n$ . For symmetric  $A$ , we assume the ordering  $\lambda_1(A) \geq \dots \geq \lambda_n(A)$ . For  $A \in \mathbb{R}^{m \times n}$ ,  $\|A\|_2$  denotes its spectral norm, i.e., the largest singular value of  $A$ . We denote  $\mathbf{1}$  to be the all one’s column vector. For a matrix  $U$ ,  $\mathcal{R}(U)$  denotes the range space of its columns. Throughout the paper,  $G = (V, E)$  denotes the signed graph with vertex set  $V$ , edge set  $E$ , and adjacency matrix  $A \in \{0, \pm 1\}^{n \times n}$ . We let  $G^+ = (V, E^+)$ , respectively  $G^- = (V, E^-)$ , denote the edge-disjoint subgraphs of positive, respectively negative, edges with adjacency matrices (with non-negative entries)  $A^+$ , respectively  $A^-$ , such that  $A = A^+ - A^-$ . Here,  $A_{ij}^+ = \max\{A_{ij}, 0\}$  and  $A_{ij}^- = \max\{-A_{ij}, 0\}$ . Moreover  $E^+ \cap E^- = \emptyset$ , and  $E^+ \cup E^- = E$ .

## 2 Related literature

The problem of clustering signed graphs can be traced back to the work of Cartwright and Harary from the 1950s on social balance theory [27, 10], who explored the concept of *balance* in signed graphs. A signed graph is said to be  $k$ -balanced if there exists a partition into  $k$  disjoint sets such that positive edges exist only within clusters, and negative edges are only present across clusters. The “weak balance theory” of Davis [18] relaxed the balanced relationship by allowing an

enemy of one’s enemy to also be an enemy, laying the ground for the later interplay with graph clustering.

Motivated by this theory, the  $k$ -way clustering problem in signed graphs amounts to finding a partition into  $k$  clusters such that most edges within clusters are positive, and most edges across clusters are negative.



Alternatively, one may seek a partition such that the number of **violations** is minimized, i.e., negative edges within the cluster and positive edges across clusters, as depicted in

the figure above. In order to avoid partitions where clusters contain only a few nodes, one often wishes to also incentivize clusters of large size or volume.

A wide range of algorithms have been proposed to perform clustering of signed graphs. Doieran and Mrvar [20] proposed a local search approach in the spirit of the Kernighan-Lin algorithm [30]. Yang et al. [57] introduced an agent-based approach by considering a certain random walk on the graph. In recent years, several efforts for the analysis of signed graphs have lead to novel extensions for various tasks, including edge prediction [33, 35], node classification [9, 52], node embeddings [12, 19, 31, 54], node ranking [14, 49], and clustering [11, 34, 39]. We refer the reader to [53] for a recent survey on the topic.

**Spectral methods** on signed networks began with Anchuri et al. [5], whose proposed approach optimizes modularity and other objective functions in signed graphs, and demonstrated its effectiveness in detecting communities in real world networks. Kunegis et al. [34], proposed new spectral tools for clustering, link prediction, and visualization of signed graphs, by solving a signed version of the 2-way ratio-cut problem via the (combinatorial) Signed Laplacian [28]

$$\bar{L} = \bar{D} - A, \quad (2.1)$$

where  $\bar{D}$  is the diagonal matrix with  $\bar{D}_{ii} = \sum_{j=1}^n |A_{ij}|$ . Similar signed extensions also exist for the random-walk Laplacian and the symmetric graph Laplacian

$$\bar{L}_{\text{rw}} = I - \bar{D}^{-1}A, \quad \bar{L}_{\text{sym}} = I - \bar{D}^{-1/2}A\bar{D}^{-1/2}, \quad (2.2)$$

the latter of which is particularly suitable for skewed degree distributions. Dhillon et al. [13] put forth the claim that the Signed Laplacian  $\bar{L}$  faces a fundamental weakness when directly extended to  $k$ -way clustering for  $k > 2$ . They proposed a formulation based on the *Balanced Normalized Cut* (BNC) objective

$$\min_{\{x_1, \dots, x_k\} \in I} \left( \sum_{c=1}^k \frac{x_c^T (D^+ - A)x_c}{x_c^T \bar{D} x_c} \right). \quad (2.3)$$

Here,  $D^+$  denotes the diagonal matrix with degrees  $D_{ii}^+ = \sum_{j=1}^n A_{ij}^+$ ;  $C_1, \dots, C_k$  denote the  $k$  clusters, and  $I$  denotes a  $k$ -cluster indicator set, where  $(x_t)_i = 1$ , if node  $i \in C_t$ , and 0 otherwise. The same authors also consider the closely related *Balanced Ratio Cut*, which replaces  $\bar{D}$  in the denominator by  $I$ . We remark that spectral clustering algorithms (for signed/unsigned graphs) typically have a common pipeline wherein a suitable graph operator is considered (for eg. Laplacian), its  $k$  (or  $k-1$ ) extremal eigenvectors are found, and the resulting points in  $\mathbb{R}^k$  (or  $\mathbb{R}^{k-1}$ ) are then clustered using  $k$ -means.

Hsieh et al. [29] propose a low-rank model in which state-of-the-art matrix completion algorithms are first leveraged, before clustering via  $k$ -means using the top  $k$  eigenvectors of the completed matrix. Mercado et al. [39] present an extended spectral method based on the geometric mean of Laplacians. For  $k=2$ , [15] showed that signed clustering can be formulated as an instance of the group synchronization problem over  $\mathbb{Z}_2$ , for which spectral, semidefinite programming relaxations, and message passing algorithms have been studied. Finally, we refer the reader to [23] for a recent survey on clustering signed and unsigned graphs.

### 3 SPONGE: a signed generalized eigenproblem formulation

Given an unsigned graph  $H$  with adjacency matrix  $W$  with non-negative entries, for any cluster  $C \subset V$  define  $\text{cut}_H(C, \bar{C}) := \sum_{i \in C, j \in \bar{C}} W_{ij}$  as the total weight of edges crossing from  $C$  to  $\bar{C}$ . Also define the volume of  $C$ ,  $\text{vol}_H(C) := \sum_{i \in C} \sum_{j=1}^n W_{ij}$  as the sum of degrees of nodes in  $C$ . Motivated by the approach of [16] in the context of constrained clustering, we aim to minimize the following two measures of *badness*

$$\frac{\text{cut}_{G^+}(C, \bar{C})}{\text{vol}_{G^+}(C)}, \quad (3.1)$$

$$\left( \frac{\text{cut}_{G^-}(C, \bar{C})}{\text{vol}_{G^-}(C)} \right)^{-1} = \frac{\text{vol}_{G^-}(C)}{\text{cut}_{G^-}(C, \bar{C})}. \quad (3.2)$$

Ideally,  $C$  is such that both (3.1) and (3.2) are small. To this end, we first consider “merging” the objectives (3.1) and (3.2), and would like to solve

$$\min_{C \subset V} \frac{\text{cut}_{G^+}(C, \bar{C}) + \tau^- \text{vol}_{G^-}(C)}{\text{cut}_{G^-}(C, \bar{C}) + \tau^+ \text{vol}_{G^+}(C)}, \quad (3.3)$$

with  $\tau^+, \tau^- > 0$  denoting trade-off or regularization parameters. While at first sight this may seem rather ad-hoc in nature, we provide a sound theoretical justification for our approach in later sections. A natural extension to  $k > 2$  disjoint clusters  $C_1, \dots, C_k$  leads to the following discrete optimization problem

$$\min_{C_1, \dots, C_k} \sum_{i=1}^k \frac{\text{cut}_{G^+}(C_i, \bar{C}_i) + \tau^- \text{vol}_{G^-}(C_i)}{\text{cut}_{G^-}(C_i, \bar{C}_i) + \tau^+ \text{vol}_{G^+}(C_i)}. \quad (3.4)$$

For a subset  $C_i \subset V$ , the normalized indicator vector

$$(x_{C_i})_j = \begin{cases} (\text{cut}_{G^-}(C_i, \bar{C}_i) + \text{vol}_{G^+}(C_i))^{-1/2}; & v_j \in C_i \\ 0; & v_j \notin C_i \end{cases} \quad (3.5)$$

renders (3.4) as the discrete optimization problem

$$\min_{C_1, \dots, C_k} \sum_{i=1}^k \frac{x_{C_i}^T (L^+ + \tau^- D^-) x_{C_i}}{x_{C_i}^T (L^- + \tau^+ D^+) x_{C_i}}, \quad (3.6)$$

which is NP-hard. Here  $L^+$  (resp.  $L^-$ ) denotes the Laplacian of  $G^+$  (resp.  $G^-$ ), and  $D^+$  (resp.  $D^-$ ) denotes a diagonal matrix with the degrees of  $G^+$  (resp.  $G^-$ ). A common approach in this situation is to drop the discreteness constraint and allow each  $x_{C_i}$  to take values in  $\mathbb{R}^n$ . To this end, we introduce a new set of vectors  $z_1, \dots, z_k \in \mathbb{R}^n$  such that they are orthonormal with respect to  $L^- + \tau^+ D^+$ , i.e.,

- $z_i^T (L^- + \tau^+ D^+) z_i = 1$ , and
- $z_i (L^- + \tau^+ D^+) z_j = 0$ , for  $i \neq j$ .

This leads to the following modified version of (3.6)

$$\min_{z_i (L^- + \tau^+ D^+) z_j = \delta_{ij}} \sum_{i=1}^k \frac{z_i^T (L^+ + \tau^- D^-) z_i}{z_i^T (L^- + \tau^+ D^+) z_i}. \quad (3.7)$$

The above choice of  $(L^- + \tau^+ D^+)$ -orthonormality of vectors  $z_1, \dots, z_k$  is not – strictly speaking – a relaxation of (3.6). But it leads to a suitable eigenvalue problem. Indeed, assuming  $L^- + \tau^+ D^+$  is full rank, consider the change of variables

$$y_i = (L^- + \tau^+ D^+)^{1/2} z_i, \quad (3.8)$$

which changes the orthonormality constraints of (3.6) to  $y_i^T y_j = \delta_{ij}$ . Furthermore, by considering the matrix  $Y = [y_1, \dots, y_k] \in \mathbb{R}^{n \times k}$ , one can rewrite (3.7) as

$$\min_{Y^T Y = I} \text{Tr} \left( Y^T (L^- + \tau^+ D^+)^{-1/2} (L^+ + \tau^- D^-) (L^- + \tau^+ D^+)^{-1/2} Y \right). \quad (3.9)$$

The solution to (3.9) is given by the eigenvectors corresponding to the  $k$ -smallest eigenvalues of the matrix  $(L^- + \tau^+ D^+)^{-1/2} (L^+ + \tau^- D^-) (L^- + \tau^+ D^+)^{-1/2}$  (see for eg. [48, Theorem 2.1]). One can also verify<sup>1</sup> that  $(\lambda, v)$  is an eigenpair of the previous matrix if and only if  $(\lambda, (L^- + \tau^+ D^+)^{-1/2} v)$  is a *generalized* eigenpair of  $(L^+ + \tau^- D^-, L^- + \tau^+ D^+)$ .

Our complete algorithm SPONGE (Signed Positive Over Negative Generalized Eigenproblem) first finds the smallest  $k$  generalized eigenvectors of  $(L^+ +$

<sup>1</sup> Let  $A, B$  be symmetric matrices with  $A \succ 0$ . Then  $(\lambda, v)$  is an eigenpair of  $A^{-1/2} B A^{-1/2}$  iff  $(\lambda, A^{-1/2} v)$  is a generalized eigenpair of  $(B, A)$ . Indeed, for  $w = A^{-1/2} v$ ,  $A^{-1/2} B A^{-1/2} v = \lambda v \Leftrightarrow B w = \lambda A w$ .

$\tau^- D^-, L^- + \tau^+ D^+$ ) for suitably chosen  $\tau^+, \tau^- > 0$ . We then cluster the resulting embedding of the vertices in  $\mathbb{R}^k$  using  $k$ -means++. We also consider a variant of SPONGE, namely SPONGE<sub>sym</sub>, where the embedding is generated using the smallest  $k$  generalized eigenvectors of  $(L_{sym}^+ + \tau^- I, L_{sym}^- + \tau^+ I)$ , wherein  $L_{sym}^+ = (D^+)^{-1/2} L^+ (D^+)^{-1/2}$  is the so-called symmetric Laplacian of  $G^+$  (similarly for  $L_{sym}^-$ ).

**Remark 1.** *Solving (3.9) is computationally expensive in practice as it involves first finding a matrix-inverse. This is not the case if we solve the generalized eigenproblem version of (3.9). In our experiments, we use LOBPCG[32] – a preconditioned eigensolver<sup>2</sup> for solving large positive definite generalized eigenproblems.*

## 4 Analysis of SPONGE under SSBM

We begin by introducing the signed stochastic block model (SSBM) in Section 4.1 and then theoretically analyze the performance of SPONGE in Section 4.2.

### 4.1 Signed stochastic block model

For ease of exposition, we assume  $n$  is a multiple of  $k$ , and partition the vertices of  $G$  into  $k$ -equally sized clusters  $C_1, \dots, C_k$ . In particular, we assume w.l.o.g that  $C_l = \left\{ \frac{(l-1)n}{k} + 1, \dots, \frac{ln}{k} \right\}$  for  $l = 1, \dots, k$ . The graph  $G$  follows the Erdős-Rényi random graph model  $G(n, p)$  wherein each edge takes value  $+1$  if both its endpoints are contained in the same cluster, and  $-1$  otherwise. To model noise, we flip the sign of each edge independently with probability  $\eta \in [0, 1/2)$ .

Let  $A \in \{0, \pm 1\}^{n \times n}$  denote the adjacency matrix of  $G$ , then  $(A_{ij})_{i \leq j}$  are independent random variables. Recall that  $A = A^+ - A^-$ , where  $A^+, A^- \in \{0, 1\}^{n \times n}$  are the adjacency matrices of the unsigned graphs  $G^+, G^-$  respectively. Then,  $(A_{ij}^+)_{i \leq j}$  are independent and similarly  $(A_{ij}^-)_{i \leq j}$  are also independent. But clearly, for given  $i, j \in [n]$  with  $i \neq j$ ,  $A_{ij}^+$  and  $A_{ij}^-$  are dependent.

**Remark 2.** *Contrary to stochastic block models for unsigned graphs, we do not require the intra-cluster edge probabilities to be different from those of inter-cluster edges. While this is necessary in the unsigned case for detecting clusters (eg. [42, 43]), it is not the case for signed networks since the sign of the edge already achieves this purpose implicitly. In fact, as one would expect, it is the parameter  $\eta$  that is crucial for identifiability, as shown formally in our analysis.*

### 4.2 Theoretical results for SPONGE

We now theoretically analyze the performance of SPONGE under the SSBM. In particular, we analyze

<sup>2</sup>Locally Optimal Block Preconditioned Conjugate Gradient method.

the embedding given by the smallest  $k$  eigenvectors of

$$T = (L^- + \tau^+ D^+)^{-1/2} (L^+ + \tau^- D^-) (L^- + \tau^+ D^+)^{-1/2}, \quad (4.1)$$

for parameters  $\tau^-, \tau^+ > 0$ . Recall that  $(\lambda, v)$  is an eigenpair of  $T$  if and only if  $(\lambda, (L^- + \tau^+ D^+)^{-1/2} v)$  is a generalized eigenpair for the matrix pencil  $(L^+ + \tau^- D^-, L^- + \tau^+ D^+)$ . We assume throughout that both  $L^+ + \tau^- D^-$  and  $L^- + \tau^+ D^+$  are full rank. For ease of exposition we focus on the case  $k = 2$  but the results can be extended to the general  $k \geq 2$  setting (work in progress) using the same proof outline. Denote

$$\bar{T} = (\mathbb{E}[L^-] + \tau^+ \mathbb{E}[D^+])^{-1/2} (\mathbb{E}[L^+] + \tau^- \mathbb{E}[D^-]) (\mathbb{E}[L^-] + \tau^+ \mathbb{E}[D^+])^{-1/2}, \quad (4.2)$$

and also denote

$$V_2(T) = [v_n(T) \ v_{n-1}(T)], \quad V_2(\bar{T}) = [v_n(\bar{T}) \ v_{n-1}(\bar{T})], \quad (4.3)$$

to be  $n \times 2$  matrices consisting of the smallest two eigenvectors of  $T, \bar{T}$  respectively. Let

$$w = \frac{1}{\sqrt{n}} (\mathbf{1}^T \quad -\mathbf{1}^T)^T \in \mathbb{R}^n \quad (4.4)$$

correspond to the ‘‘ground truth’’ or ‘‘planted clusters’’ we seek to recover. Our main result is the following.

**Theorem 1.** *For  $\eta \in [0, 1/2)$  let  $\tau^+, \tau^- > 0$  satisfy*

$$\tau^- < \tau^+ \left( \frac{\frac{n}{2} - 1 + \eta}{\frac{n}{2} - \eta} \right). \quad (4.5)$$

*Then it holds that  $\{v_{n-1}(\bar{T}), v_n(\bar{T})\} = \{\mathbf{1}, w\}$  where  $w$  is defined in (4.4). Moreover, assuming  $n \geq 6$ , for given  $0 < \varepsilon \leq 1/2$ ,  $\epsilon \in (0, 1)$  and  $\varepsilon_\tau \in (0, 1)$  let  $\tau^- \leq \varepsilon_\tau \tau^+ \left( \frac{\frac{n}{2} - 1 + \eta}{\frac{n}{2} - \eta} \right)$  and  $p \geq c'_1(\varepsilon, \tau^+, \tau^-, \varepsilon_\tau, \eta, \epsilon) \frac{\log n}{n}$  where  $c'_1(\varepsilon, \tau^+, \tau^-, \varepsilon_\tau, \eta, \epsilon) > 0$  depends only on  $\varepsilon, \tau^+, \tau^-, \varepsilon_\tau, \eta, \epsilon$ . Then there exists a constant  $c_\varepsilon > 0$  depending only on  $\varepsilon$  such that with probability at least  $\left(1 - \frac{\epsilon}{n} - 2n \exp\left(\frac{-pn}{c_\varepsilon}\right)\right)$ , it holds that*

$$\|(I - V_2(T)V_2(T)^T)V_2(\bar{T})\|_2 \leq \frac{\epsilon}{1 - \epsilon}.$$

The theorem states that  $\mathcal{R}(V_2(T))$  is close to  $\mathcal{R}(V_2(\bar{T}))$  with high probability provided  $n, p$  are suitably large, and  $\tau^-$  is sufficiently small compared to  $\tau^+$ . The latter condition is required to ensure that the smallest two eigenvectors of  $\bar{T}$  are  $\{\mathbf{1}, w\}$ . Also note that the embedding generated by any orthonormal basis for  $\mathcal{R}(V_2(T))$  leads to the same clustering performance<sup>3</sup>. Since the embedding corresponding to  $V_2(\bar{T})$  leads to perfectly separated (ground truth) clusters, hence the closer  $\mathcal{R}(V_2(T))$  is to  $\mathcal{R}(V_2(\bar{T}))$ , the better is the clustering performance. Using standard tools, one can actually use bounds on subspace recovery to bound the misclustering rate of  $k$ -means (see for eg. [46])

<sup>3</sup>For a  $2 \times 2$  rotation matrix  $O$ , the rows of the matrix  $V_2(T)O$  are obtained via the same rotation transformation on the corresponding rows of  $V_2(T)$ .

**Proof sketch.** The proof is deferred to the appendix but the main steps involved are as follows. We first compute the spectra of  $\mathbb{E}[L^-]$ ,  $\mathbb{E}[L^+]$ ,  $\mathbb{E}[D^-]$ ,  $\mathbb{E}[D^+]$  by finding the eigenvalues and the corresponding *relevant* eigenvectors (i.e., associated to the smallest two eigenvalues). We then identify conditions on the parameters  $\tau^+, \tau^-$  under which the smallest two eigenvectors of  $\bar{T}$  are  $\{w, \mathbf{1}\}$ . As shown in the proof,  $w$  is always one of the smallest two eigenvectors (since  $\tau^+, \tau^- > 0$ ). The condition (4.5) leads to  $\mathbf{1} \in \{v_{n-1}(\bar{T}), v_n(\bar{T})\}$ . Next, we derive concentration bounds using tools from random matrix theory for  $A^-, A^+, D^-, D^+$  holding with high probability. This in turn leads to a bound on  $\|T - \bar{T}\|_2$ . Combining the above results and by controlling the perturbation term  $\|T - \bar{T}\|_2$ , we obtain via the Davis-Kahan theorem [17], a bound on  $\|\sin(\Theta(\mathcal{R}(V_2(T)), \mathcal{R}(V_2(\bar{T}))))\|_2$  which equals  $\|(I - V_2(T)V_2(T)^T)V_2(\bar{T})\|_2$ . Here,  $\Theta(\mathcal{R}(V_2(T)), \mathcal{R}(V_2(\bar{T})))$  is the diagonal matrix of canonical angles between  $\mathcal{R}(V_2(T))$  and  $\mathcal{R}(V_2(\bar{T}))$ .

**Selecting only  $v_n(T)$ .** Alternately, one could consider taking just the smallest eigenvector of  $T$ , i.e.  $v_n(T)$ , leading to a one dimensional embedding. The following theorem states that provided  $\tau^-$  is sufficiently larger than  $\tau^+$ , and if  $n, p$  are suitably large, then  $\mathcal{R}(v_n(T))$  is close to  $\mathcal{R}(w)$  with high probability.

**Theorem 2.** For  $\eta \in [0, 1/2)$  let  $\tau^+, \tau^- > 0$  satisfy

$$\tau^- > \left(\frac{\eta}{1-\eta}\right) \left(\frac{\frac{n}{2}-1+\eta}{\frac{n}{2}-\eta}\right) \tau^+. \quad (4.6)$$

Then it holds that  $v_n(\bar{T}) = w$  with  $w$  defined in (4.4).

Moreover, assuming  $n \geq 6$ , for given  $0 < \varepsilon \leq 1/2, \epsilon \in (0, 1)$  and  $\varepsilon_\tau \in (0, 1)$  let  $\tau^- \geq \frac{1}{\varepsilon_\tau} \left(\frac{\eta}{1-\eta}\right) \left(\frac{\frac{n}{2}-1+\eta}{\frac{n}{2}-\eta}\right) \tau^+$ ,

$$p \geq c'_2(\varepsilon, \tau^+, \tau^-, \varepsilon_\tau, \eta, \epsilon) \frac{\log n}{n},$$

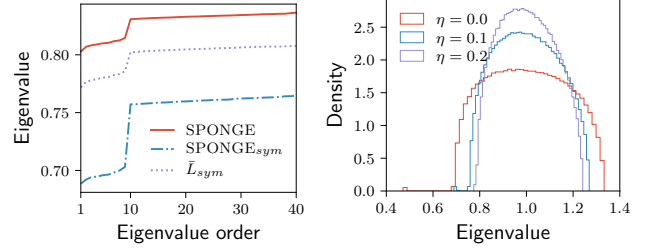
where  $c'_2(\varepsilon, \tau^+, \tau^-, \varepsilon_\tau, \eta, \epsilon) > 0$  depends only on the indicated parameters. Then there exists a constant  $c_\varepsilon > 0$  depending only on  $\varepsilon$  such that with probability at least  $\left(1 - \frac{4}{n} - 2n \exp\left(-\frac{pn}{c_\varepsilon}\right)\right)$ , it holds that

$$\|(I - v_n(T)v_n(T)^T)w\|_2 \leq \frac{\epsilon}{1-\epsilon}.$$

The proof is deferred to the appendix, being similar to that of Theorem 1. The main difference is in the conditions on  $\tau^-, \tau^+$  for ensuring that the smallest (two) eigenvector(s) of  $\bar{T}$  correspond to the ground truth clustering, these are clearly weaker in Theorem 2 compared to Theorem 1. For eg. if  $\eta = 0$ , then any  $\tau^+, \tau^- > 0$  imply  $v_n(\bar{T}) = w$  by Theorem 2, while the analogous statement is not true in Theorem 1.

Figure 1 (left) compares the 40 smallest eigenvalues of SPONGE, SPONGE<sub>sym</sub>, and  $\bar{L}_{sym}$  in the scenario  $n = 10,000$ ,  $p = 0.01$ ,  $\eta = 0.1$ , and  $k =$

10. SPONGE<sub>sym</sub> clearly exhibits the largest spectral gap after 10 eigenvalues. Figure 1 (right) also compares the spectral densities of SPONGE<sub>sym</sub> for several  $\eta \in \{0, 0.1, 0.2\}$ . As expected, the spectral gap decreases as the noise level increases. Figure 2



**Figure 1:** Left: Bottom spectrum of SPONGE<sub>sym</sub>, SPONGE, and  $\bar{L}_{sym}$ . Right: Spectrum of SPONGE<sub>sym</sub> for three values of noise  $\eta$ , ( $n = 10000$ ,  $p = 0.01$ , and  $k = 10$ ).

compares heatmaps of recovery rates for SPONGE and SPONGE<sub>sym</sub> for  $k = 2$  clusters and varying  $\tau^+, \tau^- > 0$ . Observe that (4.6) shows up when we consider only the smallest eigenvector for SPONGE. Figure 3 shows similar plots for  $k = 8$ , where we observe that SPONGE<sub>sym</sub> allows for a wider choice of  $\tau^+, \tau^- > 0$  for successful clustering.

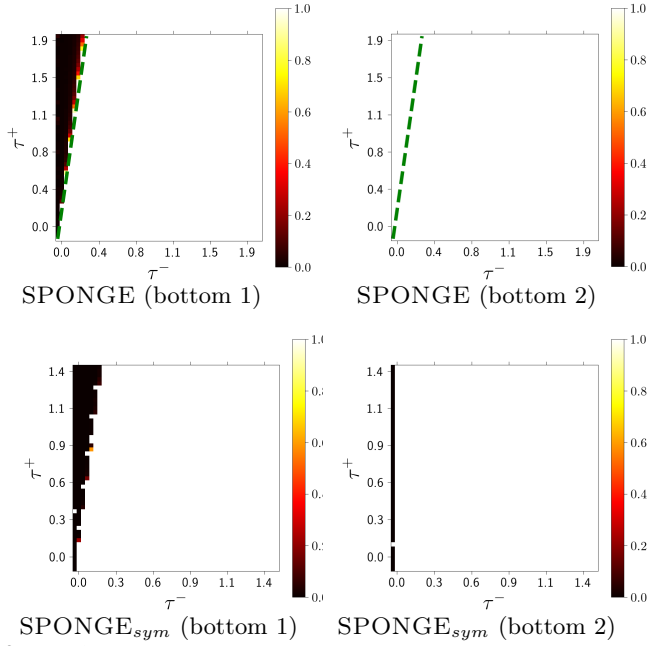
At a high level, our proof technique (tools from matrix perturbation, random matrices) has been used before for analyzing spectral methods for clustering unsigned graphs (eg. [46]). In the sparse regime where  $p \rightarrow 0$  as  $n \rightarrow \infty$ , Theorems 1, 2 state that  $p \gtrsim \frac{\log n}{n}$  ensures that the success probability tends to one. Similar scalings are known for unsigned graphs however there, the intra-cluster and inter-cluster edge probabilities necessarily need to be different (see Remark 2).

## 5 Theoretical analysis of the Signed Laplacian $\bar{L}$ under SSBM

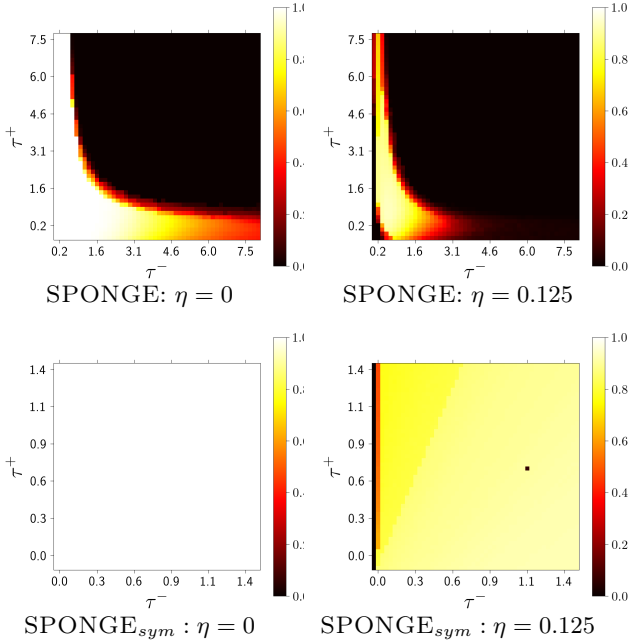
In this section, we theoretically analyze the popular Signed Laplacian based method of Kunegis et al. [34] for clustering signed graphs under the SSBM. This method is particularly appealing due to its simplicity, but to our knowledge, there do not exist any theoretical guarantees on the performance of this approach. We fill this gap by providing a detailed analysis for the  $k = 2$  case. This choice is for ease of exposition, but the proof outline clearly extends<sup>4</sup> to any  $k \geq 2$ .

Recall that for a signed graph  $G$  with adjacency matrix  $A \in \{0, \pm 1\}^{n \times n}$ , and with the diagonal matrix  $\bar{D}$  consisting of the degree terms defined as  $\bar{D}_{ii} := \sum_{j=1}^n |A_{ij}|$ , the Signed Laplacian of  $G$ , denoted by  $\bar{L} \in \mathbb{R}^{n \times n}$ , is given by  $\bar{L} = \bar{D} - A$ . Kunegis et al. [34] showed that  $\bar{L}$  is positive semi-definite for any graph

<sup>4</sup>This is part of work currently in progress.



**Figure 2:** Heatmap of recovery rates for  $k = 2$  clusters for SPONGE and SPONGE<sub>sym</sub>, with  $n = 5000, p = 0.012$  and  $\eta = 0.125$ , via the bottom one or two eigenvectors. The green dotted line is the condition (4.6).



**Figure 3:** Heatmap of recovery rates for SPONGE and SPONGE<sub>sym</sub>, using the bottom  $k - 1$  eigenvectors, as we vary  $\tau^+, \tau^-$ , with  $n = 5000, p = 0.012, k = 8$  and  $\eta \in \{0, 0.125\}$ .

(see [34, Theorem 4.1]). Moreover, they also showed that  $\bar{L}$  is positive definite iff the graph is unbalanced<sup>5</sup> [34, Theorem 4.4]. Kunegis et al. proposed using  $\bar{L}$  to first compute a lower dimensional embedding of the graph – obtained from the smallest  $k$  eigenvectors of  $\bar{L}$  (in fact, as we will see, taking only  $k - 1$  is sufficient and more effective in signed graphs) and then clustering the obtained points in  $\mathbb{R}^k$  (or  $\mathbb{R}^{k-1}$ ) using any standard clustering method (e.g.  $k$ -means).

Our main result for the Signed Laplacian based clustering approach of Kunegis et al. [34] is stated below, the proof is deferred to the appendix.

**Theorem 3.** *Assuming  $0 \leq \eta < 1/2$ , it holds that  $v_n(\mathbb{E}[\bar{L}]) = w$  where  $w$  is defined in (4.4). Moreover, let  $n \geq 2$  and for given  $0 < \epsilon < 1, 0 < \epsilon \leq 1/2$  let*

$$p \geq \frac{4((1 + \epsilon)2\sqrt{2} + 1)^2 \log n}{\epsilon^2(1 - 2\eta)^2 n}.$$

*Then there exists a constant  $c_\epsilon > 0$  depending only on  $\epsilon$  such that with probability at least*

$$1 - \frac{2}{n} - n \exp\left(-\frac{pn}{4c_\epsilon}\right),$$

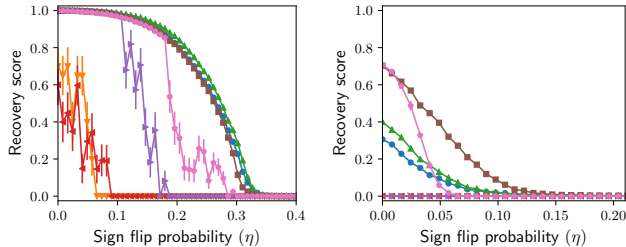
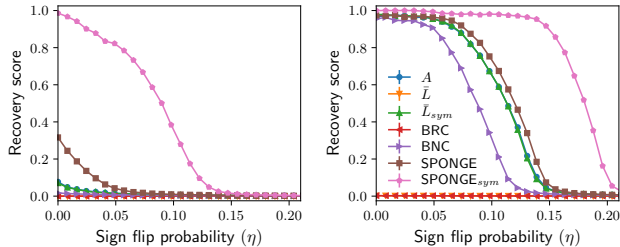
*it holds that  $\|(I - ww^T)v_n(\bar{L})\|_2 \leq \frac{\epsilon}{1 - \epsilon}$ .*

Theorem 3 states that for  $n, p$  suitably large,  $\mathcal{R}(v_n(\bar{L})) \approx \mathcal{R}(w)$  with high probability. In particular, if  $\eta$  is bounded away from  $1/2$ , then in the sparse regime where  $p \rightarrow 0$  as  $n \rightarrow \infty$ , the success probability approaches one if  $p \gtrsim \frac{\log n}{n}$ . As seen in the proof,  $\mathbb{E}[\bar{L}]$  is positive definite if  $\eta \neq 0$ , and positive semi-definite otherwise. This makes sense since for  $\eta = 0$ , the generated graph (under the SSBM) is balanced by construction and so is positive semi-definite [34, Theorem 4.4]. The fact that  $\mathbb{E}[\bar{L}]$  is positive definite for  $\eta \neq 0$  tells us that the resulting graph will be unbalanced with high probability. Finally, we note that as  $\eta$  approaches  $1/2$ , the condition on  $p$  becomes stricter since the expected number of intra-cluster positive edges is almost the same as the number of inter-cluster positive edges (similarly for negative edges). Hence, to get a non-trivial lower bound on  $p$ , we require  $n$  to be sufficiently large.

## 6 Numerical experiments

This section contains numerical experiments comparing our SPONGE and SPONGE<sub>sym</sub> algorithms (setting  $\tau^+ = \tau^- = 1$ ), with several existing spectral signed clustering techniques based on: the adjacency matrix  $A$ , the Signed Laplacian matrix  $\bar{L}$ , its symmetrically normalized version  $\bar{L}_{sym}$  [34], and the two algorithms introduced in [13] that optimize the Balanced

<sup>5</sup>A graph is unbalanced if it contains a cycle containing an odd number of negative edges [34, Definition 3]

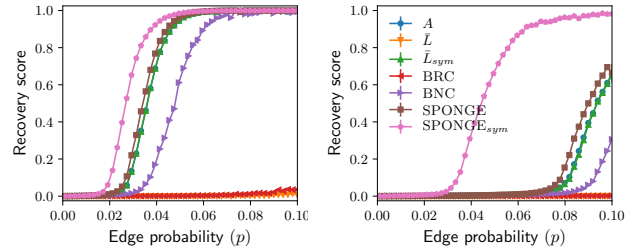
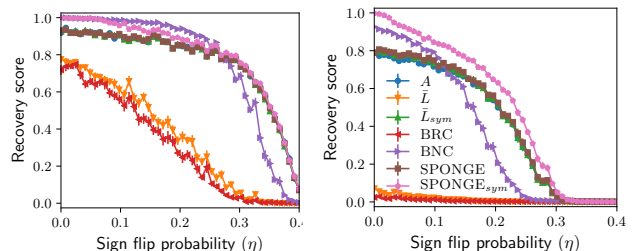
(a)  $k = 2, p = 0.001$ (b)  $k = 5, p = 0.001$ (c)  $k = 20, p = 0.01$ (d)  $k = 50, p = 0.1$ **Figure 4:** ARI recovery scores versus  $\eta$  for increasing  $k$ , with communities of equal size and  $n = 10000$ .

Ratio Cut and the Balanced Normalized Cut objectives. In all cases, the bottom  $k - 1$  (or top  $k - 1$  in the case of adjacency matrix  $A$ ) eigenvectors of the relevant matrix or generalized eigenvalue problem are considered as an embedding, and `kmeans++` is applied to obtain a  $k$ -clustering. Section 6.1 contains numerical experiments on synthetic graphs generated under the SSBM, while Section 6.2 details the results obtained on four different real-world data sets. An extensive set of simulations can be found in the appendix.

### 6.1 Signed stochastic block model

This section compares all algorithms on a variety of synthetic graphs generated from the SSBM. Since the ground truth partition is available, we measure accuracy by the Adjusted Rand Index (ARI) [24], an improved version of the popular Rand Index [47]. Both measures indicate how well the recovered partition matches ground truth, with a value close to 1, resp. 0, indicating an almost perfect recovery, resp. an almost random assignment of the nodes into clusters.

The SSBM considered here and introduced in Section 4.1 has four parameters:  $n$ ,  $k$ ,  $p$  and  $\eta$ . In our experiments, we fix  $n = 10000$ , and let  $k \in \{2, 5, 10, 20, 50\}$  with clusters chosen of equal size  $n/k$ . We analyze the performance of all algorithms by plotting mean and standard error, over 20 repetitions, of the ARI as a function of  $\eta$  for  $p \in \{0.001, 0.01, 0.1\}$ . The results are reported in Figure 4. When  $k = 2$  (Figure 4 (a)),  $\bar{L}_{sym}$  performs slightly better than all other algorithms. As  $k$  increases, the SPONGE algorithms start to significantly outperform all other methods. In

(a)  $k = 20, \eta = 0.2$ (b)  $k = 50, \eta = 0.1$ **Figure 5:** ARI recovery scores as a function of the edge probability  $p$ , for  $k = 20$  and  $k = 50$  at two different noise levels. The communities are of equal size, and  $n = 10000$ .(a)  $k = 20, p = 0.1$ (b)  $k = 50, p = 0.1$ **Figure 6:** ARI recovery scores of all algorithms, as a function of the noise  $\eta$ , for  $k \in \{20, 50\}$  clusters of randomly chosen sizes, and fixed edge density  $p = 0.1$ .

particular, while for intermediate values of  $k$  (Figure 4 (b)) SPONGE was the best performer, once  $k = 20$  or  $k = 50$  (Figure 4 (c) and (d)) SPONGE<sub>sym</sub> was greatly superior, being able to perfectly recover the cluster structure (ARI = 1) when all other methods completely fail (ARI  $\approx$  0). We remark that similar results, showing excellent recovery for large  $k$  via SPONGE<sub>sym</sub>, hold true over a wider range of values of the sparsity  $p$ , and are reported in the appendix.

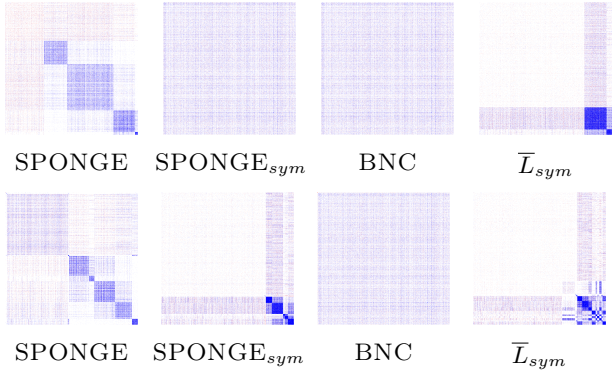
We also tested the algorithms on SSBM graphs with clusters of unequal sizes, with the probability of each node to be part of a given cluster being uniformly sampled in  $[0, 1]$ , and subsequently normalized, which typically lead to widely different cluster sizes. Under this setting (see Figure 6 and appendix), SPONGE<sub>sym</sub> was still the best performer, although the extent of the performance gap was less pronounced. Interestingly, the performance of BNC often matched (but rarely overcame) that of SPONGE<sub>sym</sub>. Overall, we find that for large enough  $k \geq 5$ , SPONGE and especially SPONGE<sub>sym</sub>, outperform all state-of-art algorithms across a broad range of values for  $p, \eta$ , and for  $n$  sufficiently large for a clustering to be recoverable.

### 6.2 Real data

This section details the outcomes of experiments on a variety of real-world signed network data sets. Due to space constraints, we show results for the four al-

gorithms that performed best on the synthetic experiments: SPONGE, SPONGE<sub>sym</sub>, BNC and  $\bar{L}_{sym}$ . Since we no longer have ground truth, we compare the output of the algorithms by plotting the network adjacency matrix sorted by membership of the clusters produced. For our time-series data applications, we also demonstrate visually that our algorithms have recovered meaningful information in their clusterings.

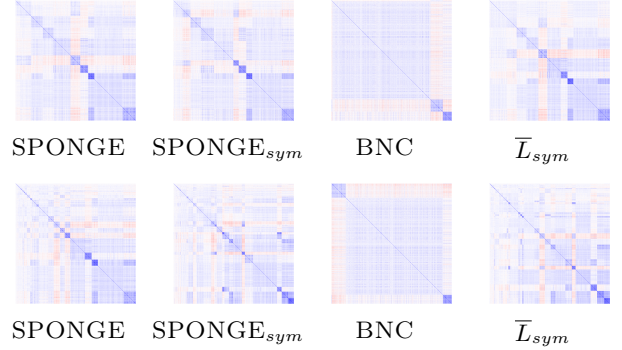
**Wikipedia elections.** We consider the classic data set of Wikipedia Requests for Adminship [55] from SNAP [37]; a network of positive, neutral, and negative votes between Wikipedia editors running in adminship elections. We construct a signed, undirected, weighted graph using the sums of edge weights for each pair of nodes. We then discard 0-weighted edges and consider only the largest connected component of the resulting graph. Thus, we obtain a graph on  $n = 11,259$  nodes with 132,412 (resp. 37,423) positive (resp. negative) edges. Figure 7 shows the resulting adjacency matrix sorted by cluster membership with  $k = 6$ , where blue (resp. red) denotes positive (resp. negative) edges. Previous work on signed networks [39], also succeeded in finding clustering structure in this data. However, the majority of the nodes are placed in a single large cluster which is very sparse and does not exhibit discernible associations. A major advantage of the clustering in Figure 7 is that all clusters demonstrate a significantly higher ratio of positive to negative internal edges, compared to that of the graph as a whole.



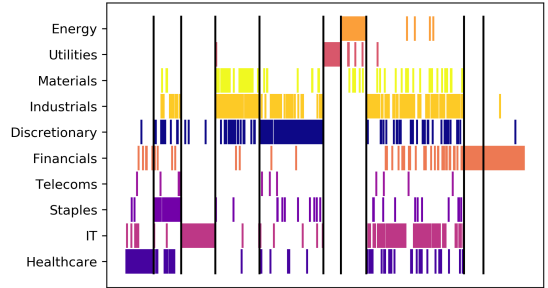
**Figure 7:** Sorted adjacency matrix of the Wikipedia graph for  $k = 6$  (top row) and  $k = 50$  (bottom row).

**Correlations of financial market returns.** We consider daily prices for  $n = 1500$  stocks in the S&P 1500 Index, during 2003-2015, and build correlation matrices from market excess returns. We refer the reader to the appendix, for a detailed overview of our steps. Figure 8 shows that, for  $k \in \{10, 30\}$ , we are able to find a meaningful segmentation of the market. In Figure 9, we interpret our results in light of the popular GICS sector decomposition [45].

**Correlations of Australian rainfalls.** We also consider time series of historical rainfalls in loca-

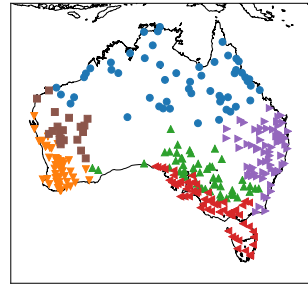


**Figure 8:** Adjacency matrix of the S&P 1500 data set sorted by cluster membership, for  $k \in \{10, 30\}$ .



**Figure 9:** GICS decomposition for SPONGE<sub>sym</sub> clusters.

tions throughout Australia. Edge weights are obtained from the pairwise Pearson correlation, leading to a complete signed graph on  $n = 306$  nodes.



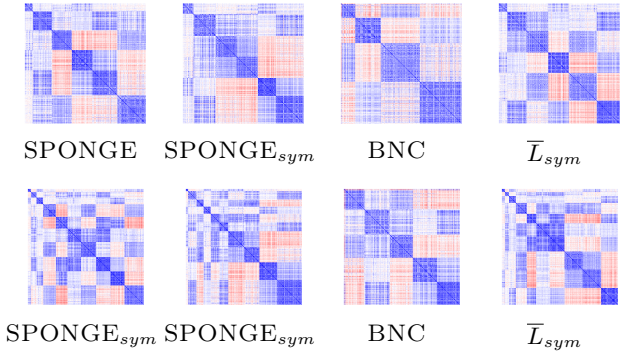
**Figure 10:** SPONGE:  $k = 6$ , Australian rainfalls data.

Figure 11 shows a clear clustering structure, for  $k = 6$  clusters, and Figure 10 plots the points onto the corresponding geographic locations. SPONGE has very effectively identified geographic regions with similar climate, based only on the correlations of the rainfall measurements.

## 7 Discussion and future directions

We introduced a principled spectral algorithm (SPONGE) for clustering signed graphs, that amounts to a generalized eigenvalue problem, and provided a theoretical analysis for  $k = 2$  clusters. Extensive numerical experiments demonstrate its robustness to noise and sampling sparsity. In particular, for sparse graphs and large  $k$ , we recover clusterings when all state-of-the-art methods completely fail.

There are several interesting avenues for possible fu-



**Figure 11:** Sorted adjacency matrix of the Australian rainfall data set, with  $k = 6$  (top) and  $k = 10$  (bottom).

ture investigation such as: (i) extending our SSBM to a more general setting that allows for different intra/inter-cluster edge sampling probabilities and noise levels; (ii) extending our current analysis to the challenging setting of very sparse graphs, where  $p = \Theta(\frac{1}{n})$ ; (iii) exploring the usefulness of the SPONGE embedding as a dimensionality reduction tool in multivariate time series analysis, and leveraging it for prediction and classification tasks; (iv) exploring semidefinite programming based approaches, inspired by recent work on community detection in sparse networks [25], and (v) investigating graph-based diffuse interface models utilizing the Ginzburg-Landau functionals, based on the MBO scheme [40] (work in progress).

## References

- [1] Epinions data set. <http://www.epinions.com>. Accessed: 2010-09-30.
- [2] Foreign exchange sdr valuation. [https://www.imf.org/external/np/fin/data/rms\\_sdrv.aspx](https://www.imf.org/external/np/fin/data/rms_sdrv.aspx).
- [3] Slashdot data set. <http://www.slashdot.com>. Accessed: 2010-09-30.
- [4] Saeed Aghabozorgi, Ali Seyed Shirkhorshidi, and Teh Ying Wah. Time-series clustering—a decade review. *Information Systems*, 53:16–38, 2015.
- [5] Pranay Anchuri and Malik Magdon-Ismael. Communities and balance in signed networks: A spectral approach. In *Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference on*, pages 235–242. IEEE, 2012.
- [6] Afonso S. Bandeira and Ramon van Handel. Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *Ann. Probab.*, 44(4):2479–2506, 07 2016.
- [7] Sujogya Banerjee, Kaushik Sarkar, Sedat Gokalp, Arunabha Sen, and Hasan Davulcu. Partitioning signed bipartite graphs for classification of individuals and organizations. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, pages 196–204. Springer, 2012.
- [8] R. Bhatia. *Matrix Analysis*. Springer New York, 1996.
- [9] J. Bosch, P. Mercado, and M. Stoll. Node classification for signed networks using diffuse interface methods. *ArXiv e-prints*, September 2018.
- [10] Dorwin Cartwright and Frank Harary. Structural balance: a generalization of Heider’s theory. *Psychological review*, 63(5):277, 1956.
- [11] K. Chiang, J. Whang, and I. Dhillon. Scalable clustering of signed networks using balance normalized cut. *CIKM*, 2012.
- [12] Kai-Yang Chiang, Nagarajan Natarajan, Ambuj Tewari, and Inderjit S. Dhillon. Exploiting longer cycles for link prediction in signed networks. *CIKM*, 2011.
- [13] Kai-Yang Chiang, Joyce Whang, and Inderjit S. Dhillon. Scalable Clustering of Signed Networks using Balance Normalized Cut. In *ACM Conference on Information and Knowledge Management (CIKM)*, oct 2012.
- [14] Fan Chung, Alexander Tsias, and Wensong Xu. Dirichlet pagerank and ranking algorithms based

- on trust and distrust. *Internet Mathematics*, 9(1):113–134, 2013.
- [15] M. Cucuringu. Synchronization over  $Z_2$  and community detection in multiplex networks with constraints. *Journal of Complex Networks*, 3:469–506, 2015.
- [16] M. Cucuringu, I. Koutis, S. Chawla, G. Miller, and R. Peng. Scalable Constrained Clustering: A Generalized Spectral Method. *Artificial Intelligence and Statistics Conference (AISTATS) 2016*, 2016.
- [17] Chandler Davis and W. M. Kahan. The rotation of eigenvectors by a perturbation. iii. *SIAM Journal on Numerical Analysis*, 7(1):1–46, 1970.
- [18] J. A. Davis. Clustering and structural balance in graphs. *Human Relations*, 20(2):181–187, 1967.
- [19] T. Derr, Y. Ma, and J. Tang. Signed Graph Convolutional Network. *ArXiv e-prints*, August 2018.
- [20] Patrick Doreian and Andrej Mrvar. Partitioning signed social networks. *Social Networks*, 31(1):1–11, 2009.
- [21] Sergio M Focardi. Clustering economic and financial time series: Exploring the existence of stable correlation conditions. *The Intertek Group*, 2005.
- [22] André Fujita, Patricia Severino, Kaname Kojima, João Ricardo Sato, Alexandre Galvão Patriota, and Satoru Miyano. Functional clustering of time series gene expression data by granger causality. *BMC systems biology*, 6(1):137, 2012.
- [23] Jean H. Gallier. Notes on elementary spectral graph theory. applications to graph clustering using normalized cuts. *CoRR*, abs/1311.2492, 2013.
- [24] Alexander J. Gates and Yong-Yeol Ahn. The impact of random models on clustering similarity. *Journal of Machine Learning Research*, 18(87):1–28, 2017.
- [25] Olivier Guédon and Roman Vershynin. Community detection in sparse networks via Grothendieck’s inequality. *Probability Theory and Related Fields*, 165(3-4):1025–1049, 2016.
- [26] Gyeong-Gyun Ha, Jae Woo Lee, and Ashadun Nobi. Threshold network of a financial market using the p-value of correlation coefficients. *Journal of the Korean Physical Society*, 66(12):1802–1808, Jun 2015.
- [27] Frank Harary. On the notion of balance of a signed graph. *Michigan Math. J.*, 2(2):143–146, 1953.
- [28] Jao Ping Hou. Bounds for the least Laplacian eigenvalue of a signed graph. *Acta Mathematica Sinica*, 21(4):955–960, 2005.
- [29] Cho-Jui Hsieh, Kai-Yang Chiang, and Inderjit S. Dhillon. Low-Rank Modeling of Signed Networks. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2012.
- [30] B. W. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. *Bell System Technical Journal*, 49(2):291–307, 1970.
- [31] Junghwan Kim, Haekyu Park, Ji-Eun Lee, and U Kang. Side: Representation learning in signed directed networks. In *WWW*, 2018.
- [32] A. Knyazev. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method. *SIAM Journal on Scientific Computing*, 23(2):517–541, 2001.
- [33] Srijan Kumar, Francesca Spezzano, V.S. Subrahmanian, and Christos Faloutsos. Edge weight prediction in weighted signed networks. In *ICDM*, 2016.
- [34] Jérôme Kunegis, Stephan Schmidt, Andreas Lommatzsch, Jürgen Lerner, Ernesto William De Luca, and Sahin Albayrak. Spectral analysis of signed graphs for clustering, prediction and visualization. *SDM*, 10:559–570, 2010.
- [35] J. Leskovec, D. Huttenlocher, and J. Kleinberg. Predicting positive and negative links in online social networks. In *WWW*, pages 641–650, 2010.
- [36] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Signed Networks in Social Media. In *CHI*, pages 1361–1370, 2010.
- [37] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford Large Network Dataset Collection. <http://snap.stanford.edu/data>, June 2014.
- [38] Ren-Cang Li. On perturbations of matrix pencils with real spectra. *Mathematics of Computation*, 62(205):231–265, 1994.
- [39] Pedro Mercado, Francesco Tudisco, and Matthias Hein. Clustering signed networks with the geometric mean of Laplacians. In *NIPS*. 2016.
- [40] Ekaterina Merkurjev, Tijana Kostic, and Andrea L Bertozzi. An MBO scheme on graphs for classification and image processing. *SIAM Journal on Imaging Sciences*, 6(4):1903–1930, 2013.
- [41] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, New York, NY, USA, 2005.
- [42] Elchanan Mossel, Joe Neeman, and Allan Sly. Consistency thresholds for the planted bisection model. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing, STOC ’15*, pages 69–75, 2015.

- [43] Elchanan Mossel, Joe Neeman, and Allan Sly. Reconstruction and estimation in the planted partition model. *Probability Theory and Related Fields*, 162(3):431–461, Aug 2015.
- [44] Nicos G Pavlidis, Vassilis P Plagianakos, Dimitris K Tasoulis, and Michael N Vrahatis. Financial forecasting through unsupervised clustering and neural networks. *Operational Research*, 6(2):103–127, 2006.
- [45] Ryan L. Phillips and Rita Ormsby. Industry classification schemes: An analysis and review. *Journal of Business & Finance Librarianship*, 21(1):1–25, 2016.
- [46] Tai Qin and Karl Rohe. Regularized spectral clustering under the degree-corrected stochastic blockmodel. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2*, NIPS’13, pages 3120–3128, 2013.
- [47] W.M. Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336):846–850, 1971.
- [48] Ahmed Sameh and Zhanye Tong. The trace minimization method for the symmetric generalized eigenvalue problem. *Journal of Computational and Applied Mathematics*, 123(1):155 – 175, 2000.
- [49] Moshen Shahriari and Mahdi Jalili. Ranking nodes in signed social networks. *Social Network Analysis and Mining*, 4(1):172, Jan 2014.
- [50] Stephen M. Smith, Karla L. Miller, Gholamreza Salimi-Khorshidi, Matthew Webster, Christian F. Beckmann, Thomas E. Nichols, Joseph D. Ramsey, and Mark W. Woolrich. Network modelling methods for fmri. *NeuroImage*, 54(2):875 – 891, 2011.
- [51] G.W. Stewart and Ji guang Sun. *Matrix Perturbation Theory*. Academic Press, 1990.
- [52] Jiliang Tang, Charu Aggarwal, and Huan Liu. Node classification in signed social networks. In *SDM*, 2016.
- [53] Jiliang Tang, Yi Chang, Charu Aggarwal, and Huan Liu. A survey of signed network mining in social media. *ACM Comput. Surv.*, 49(3):42:1–42:37, 2016.
- [54] Suhang Wang, Jiliang Tang, Charu Aggarwal, Yi Chang, and Huan Liu. *Signed Network Embedding in Social Media*. 2017.
- [55] Robert West, Hristo S. Paskov, Jure Leskovec, and Christopher Potts. Exploiting social network structure for person-to-person sentiment analysis. *Transactions of the Association for Computational Linguistics*, 2:297–310, 2014.
- [56] Hermann Weyl. Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung). *Mathematische Annalen*, 71(4):441–479, Dec 1912.
- [57] B. Yang, W. K. Cheung, and J. Liu. Community mining from signed social networks. *IEEE Trans Knowl Data Eng*, 19(10):1333–1348, 2007.
- [58] Y. Yu, T. Wang, and R. J. Samworth. A useful variant of the davis–kahan theorem for statisticians. *Biometrika*, 102(2):315–323, 2015.
- [59] Hartmut Ziegler, Marco Jenny, Tino Gruse, and Daniel A Keim. Visual market sector analysis for financial time series data. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on*, pages 83–90. IEEE, 2010.

## A Matrix perturbation analysis

Let  $A \in \mathbb{C}^{n \times n}$  be Hermitian with eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$  and corresponding eigenvectors  $v_1, v_2, \dots, v_n \in \mathbb{C}^n$ . Let  $\tilde{A} = A + W$  be a perturbed version of  $A$ , with the perturbation matrix  $W \in \mathbb{C}^{n \times n}$  being Hermitian. Let us denote the eigenvalues of  $\tilde{A}$  and  $W$  by  $\tilde{\lambda}_1 \geq \dots \geq \tilde{\lambda}_n$  and  $\epsilon_1 \geq \epsilon_2 \geq \dots \geq \epsilon_n$  respectively.

To begin with, we would like to quantify the perturbation of the eigenvalues of  $\tilde{A}$  with respect to the eigenvalues of  $A$ . Weyl's inequality [56] is a very useful result in this regard.

**Theorem 4** (Weyl's Inequality [56]). *For each  $i = 1, \dots, n$ , it holds that*

$$\lambda_i + \epsilon_n \leq \tilde{\lambda}_i \leq \lambda_i + \epsilon_1. \quad (\text{A.1})$$

*In particular, this implies that  $\tilde{\lambda}_i \in [\lambda_i - \|W\|_2, \lambda_i + \|W\|_2]$ .*

One can also quantify the perturbation of the subspace spanned by eigenvectors of  $A$ , this was established by Davis and Kahan [17]. Before introducing the theorem, we need some definitions. Let  $U, \tilde{U} \in \mathbb{C}^{n \times k}$  (for  $k \leq n$ ) have orthonormal columns respectively and let  $\sigma_1 \geq \dots \geq \sigma_k$  denote the singular values of  $U^* \tilde{U}$ . Also, let us denote  $\mathcal{R}(U)$  to be the range space of the columns of  $U$ , same for  $\mathcal{R}(\tilde{U})$ . Then the  $k$  principal angles between  $\mathcal{R}(U), \mathcal{R}(\tilde{U})$  are defined as  $\theta_i := \cos^{-1}(\sigma_i)$  for  $1 \leq i \leq k$ , with each  $\theta_i \in [0, \pi/2]$ . It is usual to define  $k \times k$  diagonal matrices  $\Theta(\mathcal{R}(U), \mathcal{R}(\tilde{U})) := \text{diag}(\theta_1, \dots, \theta_k)$  and  $\sin \Theta(\mathcal{R}(U), \mathcal{R}(\tilde{U})) := \text{diag}(\sin \theta_1, \dots, \sin \theta_k)$ . Denoting  $\|\cdot\|$  to be any unitarily invariant norm (Frobenius, spectral, etc.), the following relation holds (see for eg., [38, Lemma 2.1], [51, Corollary I.5.4]).

$$\|\sin \Theta(\mathcal{R}(U), \mathcal{R}(\tilde{U}))\| = \|(I - \tilde{U} \tilde{U}^*) U\|.$$

With the above notation in mind, we now introduce a version of the Davis-Kahan theorem taken from [58, Theorem 1] (see also [51, Theorem V.3.6]).

**Theorem 5** (Davis-Kahan). *Fix  $1 \leq r \leq s \leq n$ , let  $d = s - r + 1$ , and let  $U = (u_r, u_{r+1}, \dots, u_s) \in \mathbb{C}^{n \times d}$  and  $\tilde{U} = (\tilde{u}_r, \tilde{u}_{r+1}, \dots, \tilde{u}_s) \in \mathbb{C}^{n \times d}$ . Write*

$$\delta = \inf \left\{ |\hat{\lambda} - \lambda| : \lambda \in [\lambda_s, \lambda_r], \hat{\lambda} \in (-\infty, \tilde{\lambda}_{s+1}] \cup [\tilde{\lambda}_{r-1}, \infty) \right\}$$

*where we define  $\tilde{\lambda}_0 = \infty$  and  $\tilde{\lambda}_{n+1} = -\infty$  and assume that  $\delta > 0$ . Then*

$$\|\sin \Theta(\mathcal{R}(U), \mathcal{R}(\tilde{U}))\| = \|(I - \tilde{U} \tilde{U}^*) U\| \leq \frac{\|W\|}{\delta}.$$

For instance, if  $r = s = j$ , then by using the spectral norm  $\|\cdot\|_2$ , we obtain

$$\sin \Theta(\mathcal{R}(\tilde{v}_j), \mathcal{R}(v_j)) = \|(I - v_j v_j^*) \tilde{v}_j\|_2 \leq \frac{\|W\|_2}{\min \left\{ |\tilde{\lambda}_{j-1} - \lambda_j|, |\tilde{\lambda}_{j+1} - \lambda_j| \right\}}. \quad (\text{A.2})$$

## B Useful concentration inequalities

### B.1 Chernoff bounds

Recall the following Chernoff bound for sums of independent Bernoulli random variables.

**Theorem 6** ([41, Corollary 4.6]). *Let  $X_1, \dots, X_n$  be independent Bernoulli random variables with  $\mathbb{P}(X_i = 1) = p_i$ . Let  $X = \sum_{i=1}^n X_i$  and  $\mu = \mathbb{E}[X]$ . For  $\delta \in (0, 1)$ ,*

$$\mathbb{P}(|X - \mu| \geq \delta \mu) \leq 2 \exp(-\mu \delta^2 / 3).$$

### B.2 Spectral norm of random matrices

We will make use of the following result for bounding the spectral norm of symmetric matrices with independent, centered and bounded random variables.

**Theorem 7** ([6, Corollary 3.12, Remark 3.13]). *Let  $X$  be an  $n \times n$  symmetric matrix whose entries  $X_{ij}$  ( $i \leq j$ ) are independent, centered random variables. Then there exists for any  $0 < \varepsilon \leq 1/2$  a universal constant  $c_\varepsilon$  such that for every  $t \geq 0$ ,*

$$\mathbb{P}(\|X\|_2 \geq (1 + \varepsilon)2\sqrt{2}\tilde{\sigma} + t) \leq n \exp\left(-\frac{t^2}{c_\varepsilon \tilde{\sigma}_*^2}\right) \quad (\text{B.1})$$

where

$$\tilde{\sigma} := \max_i \sqrt{\sum_j \mathbb{E}[X_{ij}^2]}, \quad \tilde{\sigma}_* := \max_{i,j} \|X_{ij}\|_\infty.$$

Note that it suffices to employ upper bound estimates on  $\tilde{\sigma}, \tilde{\sigma}_*$  in (B.1). Indeed, if  $\tilde{\sigma} \leq \tilde{\sigma}^{(u)}$  and  $\tilde{\sigma}_* \leq \tilde{\sigma}_*^{(u)}$ , then

$$\mathbb{P}(\|X\|_2 \geq (1 + \varepsilon)2\sqrt{2}\tilde{\sigma}^{(u)} + t) \leq \mathbb{P}(\|X\|_2 \geq (1 + \varepsilon)2\sqrt{2}\tilde{\sigma} + t) \leq n \exp\left(-\frac{t^2}{c_\varepsilon \tilde{\sigma}_*^2}\right) \leq n \exp\left(-\frac{t^2}{c_\varepsilon (\tilde{\sigma}_*^{(u)})^2}\right).$$

## C Signed stochastic block model (SSBM)

Let  $A \in \{0, \pm 1\}^{n \times n}$  denote the adjacency matrix of  $G$ , with  $A_{ii} = 0$ , and  $A_{ij} = A_{ji}$ . Under the SSBM, we observe for each  $i < j$  that

<u><math>i, j</math> lie in <b>same</b> cluster</u>		<u><math>i, j</math> lie in <b>different</b> clusters</u>
$A_{ij} = \begin{cases} 1 & ; \text{ w. p } p(1 - \eta) \\ -1 & ; \text{ w. p } p\eta \\ 0 & ; \text{ w. p } (1 - p) \end{cases} \quad (\text{C.1})$	$A_{ij} = \begin{cases} 1 & ; \text{ w. p } p\eta \\ -1 & ; \text{ w. p } p(1 - \eta) \\ 0 & ; \text{ w. p } (1 - p) \end{cases} . \quad (\text{C.2})$	

In particular,  $(A_{ij})_{i \leq j}$  are independent random variables. Next, we recall that  $A$  can be decomposed as

$$A = A^+ - A^-, \quad (\text{C.3})$$

where  $A^+, A^- \in \{0, 1\}^{n \times n}$  are the adjacency matrices of the unsigned graphs  $G^+, G^-$  respectively. For any given  $i < j$ , we have

<u><math>i, j</math> lie in <b>same</b> cluster</u>		<u><math>i, j</math> lie in <b>different</b> clusters</u>
$A_{ij}^+ = \begin{cases} 1 & ; \text{ w. p } p(1 - \eta) \\ 0 & ; \text{ w. p } 1 - p(1 - \eta) \end{cases}, \quad (\text{C.4})$	$A_{ij}^+ = \begin{cases} 1 & ; \text{ w. p } p\eta \\ 0 & ; \text{ w. p } 1 - p\eta \end{cases}, \quad (\text{C.6})$	

$A_{ij}^- = \begin{cases} 1 & ; \text{ w. p } p\eta \\ 0 & ; \text{ w. p } 1 - p\eta \end{cases}, \quad (\text{C.5})$	$A_{ij}^- = \begin{cases} 1 & ; \text{ w. p } p(1 - \eta) \\ 0 & ; \text{ w. p } 1 - p(1 - \eta) \end{cases}. \quad (\text{C.7})$
---	---

Since  $A_{ij}^+ = \max\{A_{ij}, 0\}$ , therefore  $(A_{ij}^+)_{i \leq j}$  are independent random variables. Similarly  $(A_{ij}^-)_{i \leq j}$  are also independent. But clearly, for given  $i, j \in [n]$  with  $i \neq j$ , the entries  $A_{ij}^+$  and  $A_{ij}^-$  are dependent random variables.

## D Proof of Theorem 1

We will prove the following more precise version of Theorem 1 in this section.

**Theorem 8.** *Assuming  $\eta \in [0, 1/2)$  let  $\tau^+, \tau^- > 0$  satisfy  $\tau^- < \tau^+ \left(\frac{\frac{n}{2}-1+\eta}{\frac{n}{2}-\eta}\right)$ . Then it holds that  $\{v_{n-1}(\bar{T}), v_n(\bar{T})\} = \{1, w\}$  where  $w$  is defined in (4.4). Moreover, assuming  $n \geq 6$ , for given  $0 < \varepsilon \leq 1/2$ ,  $\varepsilon \in (0, 1)$  and  $\varepsilon_\tau \in (0, 1)$  let  $\tau^- \leq \varepsilon_\tau \tau^+ \left(\frac{\frac{n}{2}-1+\eta}{\frac{n}{2}-\eta}\right)$  and*

$$p \geq \max \left\{ 24, \frac{36\tilde{c}_\varepsilon^2}{(\tau^+)^2}, \frac{36\tilde{c}_\varepsilon^2}{(\tau^-)^2}, \left( \frac{\bar{c}(\varepsilon, \tau^+, \tau^-)}{\varepsilon \min \left\{ \frac{2}{3} \frac{(1-\varepsilon_\tau)}{(1+\tau^+)}, \frac{(1-2\eta)}{3} \frac{(3+\tau^++\tau^-)}{(1+\tau^+)^2} \right\}} \right)^4 \right\} \left( \frac{\log n}{n} \right),$$

where  $\tilde{c}_\varepsilon = (1 + \varepsilon)2\sqrt{2} + 1 + \sqrt{3}$ , and

$$\bar{c}(\varepsilon, \tau^+, \tau^-) = \frac{3^{3/2}\sqrt{2}\tilde{c}_\varepsilon^{1/2}(1 + \tau^-)}{(\tau^+)^{3/2}} + \frac{3\tilde{c}_\varepsilon}{\tau^+} + \frac{6^{3/2}\tilde{c}_\varepsilon^{3/2}}{(\tau^+)^{3/2}} + \frac{18\tilde{c}_\varepsilon^2}{(\tau^+)^2} + \frac{9\tilde{c}_\varepsilon(1 + \tau^-)}{(\tau^+)^2}.$$

Then for  $c_\varepsilon > 0$  depending only on  $\varepsilon$ , it holds with probability at least  $\left(1 - \frac{4}{n} - 2n \exp\left(\frac{-pn}{c_\varepsilon}\right)\right)$  that

$$\|(I - V_2(T)V_2(T)^T)V_2(\bar{T})\|_2 \leq \frac{\varepsilon}{1 - \varepsilon}.$$

The proof is outlined in the following steps.

### D.1 Step 1: Analysis of the spectra of $\mathbb{E}[L^-]$ , $\mathbb{E}[L^+]$ , $\mathbb{E}[D^-]$ and $\mathbb{E}[D^+]$

**Lemma 1.** *With  $w$  as defined in (4.4), the following holds true regarding the spectra of  $\mathbb{E}[L^+]$  and  $\mathbb{E}[D^+]$ .*

1.  $\mathbb{E}[D^+] = d^+I = p\left(\frac{n}{2} - 1 + \eta\right)I$ .
2.  $\lambda_n^+ = \lambda_n(\mathbb{E}[L^+]) = 0$ ,  $v_n^+ = v_n(\mathbb{E}[L^+]) = \mathbf{1}$ .
3.  $\lambda_{n-1}^+ = \lambda_{n-1}(\mathbb{E}[L^+]) = pn\eta$ ,  $v_{n-1}^+ = v_{n-1}(\mathbb{E}[L^+]) = w$ .
4.  $\lambda_l^+ = \lambda_l(\mathbb{E}[L^+]) = \frac{n}{2}p$ ,  $\forall l = 1, \dots, n-2$ .

Similarly, the following holds for the spectra of  $\mathbb{E}[L^-]$  and  $\mathbb{E}[D^-]$ .

1.  $\mathbb{E}[D^-] = d^-I = p\left(\frac{n}{2} - \eta\right)I$ .
2.  $\lambda_n^- = \lambda_n(\mathbb{E}[L^-]) = 0$ ,  $v_n^- = v_n(\mathbb{E}[L^-]) = \mathbf{1}$ .
3.  $\lambda_1^- = \lambda_1(\mathbb{E}[L^-]) = pn(1 - \eta)$ ,  $v_1^- = v_1(\mathbb{E}[L^-]) = w$ .
4.  $\lambda_l^- = \lambda_l(\mathbb{E}[L^-]) = \frac{n}{2}p$ ,  $\forall l = 2, \dots, n-1$ .

Before going to the proof, we can see from Lemma 1 that  $\mathbb{E}[L^+]$  and  $\mathbb{E}[L^-]$  have the same eigenspaces. In particular, the following decomposition holds true

$$\mathbb{E}[L^+] = \left[ \begin{array}{c|c|c} \underbrace{v_n^+}_{\mathbf{1}} & \underbrace{v_{n-1}^+}_{\mathbf{1}_{C,\bar{C}}} & \tilde{V}_{n \times (n-2)} \end{array} \right] \left[ \begin{array}{ccc} \lambda_n^+ & & \\ & \lambda_{n-1}^+ & \\ & & \ddots \end{array} \right] \left[ \begin{array}{c} (v_n^+)^T \\ (v_{n-1}^+)^T \\ \tilde{V}^T \end{array} \right] = U\Lambda_+U^T \quad (\text{D.1})$$

$$\mathbb{E}[L^-] = \left[ \begin{array}{c|c|c} \underbrace{v_n^-}_{\mathbf{1}} & \underbrace{v_1^-}_{\mathbf{1}_{C,\bar{C}}} & \tilde{V}_{n \times (n-2)} \end{array} \right] \left[ \begin{array}{ccc} \lambda_n^- & & \\ & \lambda_1^- & \\ & & \ddots \end{array} \right] \left[ \begin{array}{c} (v_n^-)^T \\ (v_1^-)^T \\ \tilde{V}^T \end{array} \right] = U\Lambda_-U^T. \quad (\text{D.2})$$

*Proof.* To begin with, let us note that for every  $i \neq j$ ,

$$\mathbb{E}[A_{ij}^+] = \begin{cases} p(1 - \eta) & ; \text{ if } i, j \text{ lie in same cluster} \\ p\eta & ; \text{ if } i, j \text{ lie in different clusters} \end{cases},$$

$$\text{and } \mathbb{E}[A_{ij}^-] = \begin{cases} p\eta & ; \text{ if } i, j \text{ lie in same cluster} \\ p(1 - \eta) & ; \text{ if } i, j \text{ lie in different clusters} \end{cases}.$$

This leads to the following block structure for the matrices  $\mathbb{E}[A^+], \mathbb{E}[A^-]$ .

$$\mathbb{E}[A^+] = \begin{array}{c} \begin{array}{cc} & \overbrace{\quad\quad\quad}^{n/2} \\ \left[ \begin{array}{cc} p(1-\eta)\mathbf{1}\mathbf{1}^T & p\eta\mathbf{1}\mathbf{1}^T \\ p\eta\mathbf{1}\mathbf{1}^T & p(1-\eta)\mathbf{1}\mathbf{1}^T \end{array} \right] & -p(1-\eta)I = M^+ - p(1-\eta)I, \end{array} \end{array}$$

and similarly

$$\mathbb{E}[A^-] = \left[ \begin{array}{c|c} p\eta\mathbf{1}\mathbf{1}^T & p(1-\eta)\mathbf{1}\mathbf{1}^T \\ \hline p(1-\eta)\mathbf{1}\mathbf{1}^T & p\eta\mathbf{1}\mathbf{1}^T \end{array} \right] - p\eta I = M^- - p\eta I.$$

We can observe that both  $M^+$  and  $M^-$  are rank-2 matrices.

**Computing  $\mathbb{E}[D^+], \mathbb{E}[D^-]$ .** It can be easily verified that

$$\begin{aligned} \mathbb{E}[A^+]\mathbf{1} &= \left\{ \frac{n}{2} [p(1-\eta) + p\eta] - p(1-\eta) \right\} \mathbf{1} = p \left[ \frac{n}{2} - \eta + 1 \right] \mathbf{1} \\ \text{and so, } \mathbb{E}[D^+] &= p \left( \frac{n}{2} - 1 + \eta \right) I. \end{aligned}$$

Similarly, one can also verify that

$$\begin{aligned} \mathbb{E}[A^-]\mathbf{1} &= \left\{ \frac{n}{2} [p\eta + p(1-\eta)] - p\eta \right\} \mathbf{1} = p \left[ \frac{n}{2} - \eta \right] \mathbf{1} \\ \text{and so, } \mathbb{E}[D^-] &= p \left( \frac{n}{2} - \eta \right) I. \end{aligned}$$

**Spectra of  $\mathbb{E}[A^+], \mathbb{E}[A^-]$ .** From the preceding calculations, we easily see that

$$\mathbb{E}[A^+]\mathbf{1} = p \left[ \frac{n}{2} - \eta + 1 \right] \mathbf{1} \quad \Rightarrow \quad \lambda_1(A^+) = p \left[ \frac{n}{2} - \eta + 1 \right], \quad v_1(A^+) = \mathbf{1}.$$

Recall that the informative vector is defined as  $w := \frac{1}{\sqrt{n}} \left( \underbrace{1, \dots, 1}_{n/2}, \underbrace{-1, \dots, -1}_{n/2} \right)^T$ . Clearly,

$$M^+ w = \left[ \begin{array}{c|c} p(1-\eta)\mathbf{1}\mathbf{1}^T & p\eta\mathbf{1}\mathbf{1}^T \\ \hline p\eta\mathbf{1}\mathbf{1}^T & p(1-\eta)\mathbf{1}\mathbf{1}^T \end{array} \right] w = \frac{n}{2} [p(1-2\eta)] w.$$

Therefore

$$\begin{aligned} \mathbb{E}[A^+]w &= \left( \frac{n}{2} [p(1-\eta) - p\eta] - p(1-\eta) \right) w \\ &= \left( \frac{n}{2} p(1-2\eta) - p(1-\eta) \right) w \\ &= p \left( \frac{n}{2} (1-2\eta) - (1-\eta) \right) \underbrace{w}_{v_2(\mathbb{E}[A^+])} \\ &\quad \underbrace{\hspace{10em}}_{\lambda_2(\mathbb{E}[A^+])} \end{aligned}$$

with  $\lambda_1(\mathbb{E}[A^+]) \geq \lambda_2(\mathbb{E}[A^+]) > 0$ . Since  $M^+$  is rank 2, therefore  $\lambda_3(\mathbb{E}[A^+]) = \dots = \lambda_n(\mathbb{E}[A^+]) = -p(1-\eta)$ .

Next, we repeat the above same procedure for  $A^-$ . Firstly,

$$\mathbb{E}[A^-]\mathbf{1} = p \left( \frac{n}{2} - \eta \right) \mathbf{1} \quad \Rightarrow \quad \lambda_1(\mathbb{E}[A^-]) = p \left( \frac{n}{2} - \eta \right), \quad v_1(\mathbb{E}[A^-]) = \mathbf{1}.$$

Moreover,

$$M^- w = \left[ \begin{array}{c|c} p\eta\mathbf{1}\mathbf{1}^T & p(1-\eta)\mathbf{1}\mathbf{1}^T \\ \hline p(1-\eta)\mathbf{1}\mathbf{1}^T & p\eta\mathbf{1}\mathbf{1}^T \end{array} \right] w = \frac{n}{2} [p\eta - p(1-\eta)] w = \frac{n}{2} p(2\eta - 1) w$$

which leads to

$$\begin{aligned}\mathbb{E}[A^-]w &= \left[ \frac{n}{2}p(2\eta - 1) - p\eta \right] w = p \left[ \frac{n}{2}(2\eta - 1) - \eta \right] w \\ \Rightarrow \lambda_n(\mathbb{E}[A^-]) &= p \underbrace{\left[ \frac{n}{2}(2\eta - 1) - \eta \right]}_{<0}, \quad v_n(\mathbb{E}[A^-]) = w.\end{aligned}$$

Since  $M^-$  is rank 2, hence

$$\lambda_2(\mathbb{E}[A^-]) = \dots = \lambda_{n-1}(\mathbb{E}[A^-]) = -p\eta.$$

**Spectra of  $\mathbb{E}[L^+], \mathbb{E}[L^-]$ .** Since  $\mathbb{E}[L^+] = \mathbb{E}[D^+] - \mathbb{E}[A^+] = \left(\frac{n}{2} - \eta + 1\right)pI - \mathbb{E}[A^+]$ , therefore the smallest two largest eigenvalue of  $\mathbb{E}[L^+]$  are given by

$$\begin{aligned}\lambda_n(\mathbb{E}[L^+]) &= p\left(\frac{n}{2} - 1 + \eta\right) - p\left(\frac{n}{2} - 1 + \eta\right) = 0, \\ \lambda_{n-1}(\mathbb{E}[L^+]) &= p\left(\frac{n}{2} - 1 + \eta\right) - p\left[\frac{n}{2}(1 - 2\eta) - (1 - \eta)\right] = np\eta.\end{aligned}$$

For  $1 \leq l \leq n - 2$ , the remaining eigenvalues are given by

$$\lambda_l(\mathbb{E}[L^+]) = p\left(\frac{n}{2} - 1 + \eta\right) + p(1 - \eta) = \frac{n}{2}p.$$

Note that, since  $\eta < \frac{1}{2}$ , it holds true that  $\lambda_l(\mathbb{E}[L^+]) > \lambda_{n-1}(\mathbb{E}[L^+])$ ,  $\forall 1 \leq l \leq n - 2$ . Also note that the eigenvectors of  $\mathbb{E}[L^+]$  are the same as for  $\mathbb{E}[A^+]$ .

Repeating the process for  $\mathbb{E}[L^-]$  using  $\lambda(\mathbb{E}[L^-]) = p\left(\frac{n}{2} - \eta\right) - \lambda\mathbb{E}[A^-]$ , we obtain

$$\begin{aligned}\lambda_n(\mathbb{E}[L^-]) &= 0 = p\left(\frac{n}{2} - \eta\right) - p\left(\frac{n}{2} - \eta\right), \\ \lambda_1(\mathbb{E}[L^-]) &= p\left(\frac{n}{2} - \eta\right) - p\left[\frac{n}{2}(2\eta - 1) - \eta\right] = np(1 - \eta), \\ \lambda_l(\mathbb{E}[L^-]) &= p\left(\frac{n}{2} - \eta\right) + p\eta = \frac{np}{2} \quad \left( < \lambda_1(\mathbb{E}[L^-]) \right), \quad \forall l = 2, \dots, n - 1.\end{aligned}$$

□

## D.2 Step 2: Analyzing the spectra of $\bar{T}$

**Lemma 2.** Let  $\tau^+, \tau^- > 0$  satisfy  $\tau^- < \tau^+ \left(\frac{\frac{n}{2}-1+\eta}{\frac{n}{2}-\eta}\right)$  and let  $\eta < \frac{1}{2}$ . Then the following is true.

1.  $\left\{ \lambda_n(\bar{T}), \lambda_{n-1}(\bar{T}) \right\} \in \left\{ \frac{\tau^- \left(\frac{n}{2}-\eta\right)}{\tau^+ \left(\frac{n}{2}-1+\eta\right)}, \frac{n\eta + \tau^- \left(\frac{n}{2}-\eta\right)}{n(1-\eta) + \tau^+ \left(\frac{n}{2}-1+\eta\right)} \right\}$  and  $\lambda_l(\bar{T}) = \frac{n+2\tau^- \left(\frac{n}{2}-\eta\right)}{n+2\tau^+ \left(\frac{n}{2}-1+\eta\right)}$ , for  $l = 1, \dots, n - 2$ .
2.  $\{v_n(\bar{T}), v_{n-1}(\bar{T})\} \in \{\mathbf{1}, w\}$ .

Moreover, if  $n \geq 6$  and  $\tau^- \leq \varepsilon_\tau \tau^+ \left(\frac{\frac{n}{2}-1+\eta}{\frac{n}{2}-\eta}\right)$ , then the spectral gap ( $\lambda_{gap}$ ) between  $\left\{ \lambda_n(\bar{T}), \lambda_{n-1}(\bar{T}) \right\}$  and  $\lambda_i(\bar{T}) (i = n - 2, \dots, 1)$  satisfies

$$\lambda_{gap} = \lambda_{n-2}(\bar{T}) - \lambda_{n-1}(\bar{T}) \geq \min \left\{ \frac{2(1 - \varepsilon_\tau)}{3(1 + \tau^+)}, \frac{(1 - 2\eta)(3 + \tau^+ + \tau^-)}{3(1 + \tau^+)^2} \right\}.$$

*Proof.* Using (D.1), (D.2) and Lemma 1, we can write  $\bar{T}$  as

$$\begin{aligned}\bar{T} &= (\mathbb{E}[L^-] + \tau^+ \mathbb{E}[D^+])^{-1/2} (\mathbb{E}[L^+] + \tau^- \mathbb{E}[D^-]) (\mathbb{E}[L^-] + \tau^+ \mathbb{E}[D^+])^{-1/2} \\ &= (U\lambda^- U^T + \tau^+ d^+ I)^{-1/2} (U\lambda^+ U^T + \tau^- d^- I) (U\lambda^- U^T + \tau^+ d^+ I)^{-1/2} \\ &= U \underbrace{(\Lambda^- + \tau^+ d^+ I)^{-1} (\Lambda^+ + \tau^- d^- I)}_{\Lambda_{\bar{T}}} U^T.\end{aligned}$$

$\Lambda_{\overline{T}}$  has at most three distinct values which we denote as

$$\begin{aligned} \underbrace{\Lambda_{\overline{T}}^{(1)}}_{\text{eigenvector } \mathbf{1}} &= \frac{\lambda_n^+ + \tau^- d^-}{\lambda_n^- + \tau^+ d^+} = \frac{\tau^- (\frac{n}{2} - \eta)}{\tau^+ (\frac{n}{2} - 1 + \eta)}, \\ \underbrace{\Lambda_{\overline{T}}^{(2)}}_{\text{eigenvector } w} &= \frac{\lambda_{n-1}^+ + \tau^- d^-}{\lambda_1^- + \tau^+ d^+} = \frac{n\eta + \tau^- (\frac{n}{2} - \eta)}{n(1 - \eta) + \tau^+ (\frac{n}{2} - 1 + \eta)}, \\ \Lambda_{\overline{T}}^{(3)} &= \frac{\lambda_l^+ + \tau^- d^-}{\lambda_{l'}^- + \tau^+ d^+} = \frac{n + 2\tau^- (\frac{n}{2} - \eta)}{n + 2\tau^+ (\frac{n}{2} - 1 + \eta)} \quad (\text{for any } 1 \leq l \leq n-2, 2 \leq l' \leq n-1). \end{aligned}$$

We would like to ensure that  $\Lambda_{\overline{T}}^{(1)}, \Lambda_{\overline{T}}^{(2)} < \Lambda_{\overline{T}}^{(3)}$  holds in order to obtain the right embedding. To this end, we consider next the following cases.

1.  $\Lambda_{\overline{T}}^{(1)} > \Lambda_{\overline{T}}^{(2)}$ . This is equivalent to

$$\frac{\tau^- (\frac{n}{2} - \eta)}{\tau^+ (\frac{n}{2} - 1 + \eta)} > \frac{n\eta + \tau^- (\frac{n}{2} - \eta)}{n(1 - \eta) + \tau^+ (\frac{n}{2} - 1 + \eta)} \iff \tau^- > \tau^+ \left( \frac{\eta (\frac{n}{2} - 1 + \eta)}{(1 - \eta) (\frac{n}{2} - \eta)} \right). \quad (\text{D.3})$$

2.  $\Lambda_{\overline{T}}^{(1)} < \Lambda_{\overline{T}}^{(3)}$ . This is equivalent to

$$\frac{\tau^- (\frac{n}{2} - \eta)}{\tau^+ (\frac{n}{2} - 1 + \eta)} < \frac{n + 2\tau^- (\frac{n}{2} - \eta)}{n + 2\tau^+ (\frac{n}{2} - 1 + \eta)} \iff \tau^- < \tau^+ \left( \frac{\frac{n}{2} - 1 + \eta}{\frac{n}{2} - \eta} \right). \quad (\text{D.4})$$

3.  $\Lambda_{\overline{T}}^{(2)} < \Lambda_{\overline{T}}^{(3)}$ . This is equivalent to

$$\begin{aligned} \frac{n\eta + \tau^- (\frac{n}{2} - \eta)}{n(1 - \eta) + \tau^+ (\frac{n}{2} - 1 + \eta)} &< \frac{n + 2\tau^- (\frac{n}{2} - \eta)}{n + 2\tau^+ (\frac{n}{2} - 1 + \eta)} \\ \iff n^2\eta + 2n\eta\tau^+ \left( \frac{n}{2} - 1 + \eta \right) + n\tau^- \left( \frac{n}{2} - \eta \right) &< n^2(1 - \eta) + 2n\tau^- (1 - \eta) \left( \frac{n}{2} - \eta \right) + n\tau^+ \left( \frac{n}{2} - 1 + \eta \right) \\ \iff n(1 - 2\eta) + \tau^+ \left( \frac{n}{2} - 1 + \eta \right) (1 - 2\eta) + \tau^- \left( \frac{n}{2} - \eta \right) (1 - 2\eta) &> 0, \end{aligned} \quad (\text{D.5})$$

which holds true since  $n \geq 2$  and  $\eta < \frac{1}{2}$ .

Therefore, we can conclude that if  $\eta < \frac{1}{2}$ , and if  $\tau^- < \tau^+ \left( \frac{\frac{n}{2} - 1 + \eta}{\frac{n}{2} - \eta} \right)$ , then  $\Lambda_{\overline{T}}^{(1)}, \Lambda_{\overline{T}}^{(2)} < \Lambda_{\overline{T}}^{(3)}$ . We would like to lower bound the following *spectral gap*,  $\lambda_{gap} := \min\{\Lambda_{\overline{T}}^{(3)} - \Lambda_{\overline{T}}^{(2)}, \Lambda_{\overline{T}}^{(3)} - \Lambda_{\overline{T}}^{(1)}\}$ , a quantity we analyze next.

1. Lower bounding  $\Lambda_{\bar{T}}^{(3)} - \Lambda_{\bar{T}}^{(1)}$ .

$$\begin{aligned}
\Lambda_{\bar{T}}^{(3)} - \Lambda_{\bar{T}}^{(1)} &= \frac{n + 2\tau^-(\frac{n}{2} - \eta)}{n + 2\tau^+(\frac{n}{2} - 1 + \eta)} - \frac{\tau^-(\frac{n}{2} - \eta)}{\tau^+(\frac{n}{2} - 1 + \eta)} \\
&= \frac{n \left[ \tau^+(\frac{n}{2} - 1 + \eta) - \tau^-(\frac{n}{2} - \eta) \right]}{\left[ n + 2\tau^+(\underbrace{\frac{n}{2} - 1 + \eta}_{\leq n/2}) \right] \tau^+(\underbrace{\frac{n}{2} - 1 + \eta}_{\leq n/2})} \\
&\geq \frac{n \left[ \tau^+(\frac{n}{2} - 1 + \eta) - \tau^-(\frac{n}{2} - \eta) \right]}{(n + \tau^+ n) \tau^+ \frac{n}{2}} \\
&= \frac{2 \left[ \tau^+(\frac{n}{2} - 1 + \eta) - \tau^-(\frac{n}{2} - \eta) \right]}{n(1 + \tau^+) \tau^+} \\
&\geq \frac{2 \left[ \tau^+(\frac{n}{2} - 1 + \eta)(1 - \varepsilon_\tau) \right]}{n(1 + \tau^+) \tau^+} \quad \left( \text{using } \tau^-(\frac{n}{2} - \eta) \leq \varepsilon_\tau \tau^+(\frac{n}{2} - 1 + \eta) \text{ for } \varepsilon_\tau \in (0, 1) \right) \\
&\geq \frac{2\tau^+ \frac{n}{3} (1 - \varepsilon_\tau)}{n(1 + \tau^+) \tau^+} \quad \left( \text{since } \frac{n}{2} - 1 + \eta \geq \frac{n}{3} \text{ if } n \geq 6 \right) \\
&= \frac{2(1 - \varepsilon_\tau)}{3(1 + \tau^+)}.
\end{aligned}$$

2. Lower bounding  $\Lambda_{\bar{T}}^{(3)} - \Lambda_{\bar{T}}^{(2)}$ .

$$\begin{aligned}
\Lambda_{\bar{T}}^{(3)} - \Lambda_{\bar{T}}^{(2)} &= \frac{n + 2\tau^-(\frac{n}{2} - \eta)}{n + 2\tau^+(\frac{n}{2} - 1 + \eta)} - \frac{n\eta + \tau^-(\frac{n}{2} - \eta)}{n(1 - \eta) + \tau^+(\frac{n}{2} - 1 + \eta)} \\
&= n \frac{n(1 - 2\eta) + \tau^+(\frac{n}{2} - 1 + \eta)(1 - 2\eta) + \tau^-(\frac{n}{2} - \eta)(1 - 2\eta)}{\left[ n + 2\tau^+(\frac{n}{2} - 1 + \eta) \right] \left[ n(1 - \eta) + \tau^+(\frac{n}{2} - 1 + \eta) \right]} \\
&= n(1 - 2\eta) \frac{n + \tau^+(\frac{n}{2} - 1 + \eta) + \tau^-(\frac{n}{2} - \eta)}{\left[ n + 2\tau^+(\underbrace{\frac{n}{2} - 1 + \eta}_{\leq n/2}) \right] \left[ \underbrace{n(1 - \eta)}_{\leq n} + \tau^+(\underbrace{\frac{n}{2} - 1 + \eta}_{\leq n}) \right]} \\
&\geq \frac{n(1 - 2\eta)(n + \frac{\tau^+ n}{3} + \frac{\tau^- n}{3})}{n^2(1 + \tau^+)^2} \quad (\text{if } n \geq 6) \\
&= \frac{(1 - 2\eta)}{3} \frac{(3 + \tau^+ + \tau^-)}{(1 + \tau^+)^2}.
\end{aligned}$$

We conclude that if  $\eta < 1/2$ ,  $n \geq 6$  and  $\tau^-(\frac{n}{2} - \eta) \leq \varepsilon_\tau \tau^+(\frac{n}{2} - 1 + \eta)$  for  $\varepsilon_\tau \in (0, 1)$ , then

$$\begin{aligned}
\lambda_{gap} &= \min \left\{ \Lambda_{\bar{T}}^{(3)} - \Lambda_{\bar{T}}^{(2)}, \Lambda_{\bar{T}}^{(3)} - \Lambda_{\bar{T}}^{(1)} \right\} \\
&= \lambda_{n-2}(\bar{T}) - \lambda_{n-1}(\bar{T}) \geq \min \left\{ \frac{(1 - 2\eta)}{3} \frac{(3 + \tau^+ + \tau^-)}{(1 + \tau^+)^2}, \frac{2(1 - \varepsilon_\tau)}{3(1 + \tau^+)} \right\}.
\end{aligned}$$

This completes the proof.  $\square$

### D.3 Step 3: Perturbation of $\bar{T}$

**Lemma 3** (Perturbation of  $\bar{T}$ ). *Let us denote*

$$\begin{aligned}
P &= L^- + \tau^+ D^+, & \bar{P} &= \mathbb{E}[L^-] + \tau^+ \mathbb{E}[D^+] \\
Q &= L^+ + \tau^- D^-, & \bar{Q} &= \mathbb{E}[L^+] + \tau^- \mathbb{E}[D^-].
\end{aligned}$$

Assume that  $\|A^\pm - \mathbb{E}[A^\pm]\|_2 \leq \Delta_A$  and  $\|D^\pm - \mathbb{E}[D^\pm]\|_2 \leq \Delta_D$  holds. Moreover, let the perturbation terms  $\Delta_A, \Delta_D$  satisfy

$$\underbrace{\Delta_A + \Delta_D(1 + \tau^+)}_{\Delta_{AD}^+} \leq \frac{\tau^+ p}{2} \left( \frac{n}{2} - 1 + \eta \right), \quad \underbrace{\Delta_A + \Delta_D(1 + \tau^-)}_{\Delta_{AD}^-} \leq \frac{\tau^- p}{2} \left( \frac{n}{2} - \eta \right).$$

Then,  $P, Q \succ 0$ , and the following holds true.

$$\begin{aligned} \|\underbrace{P^{-1/2} Q P^{-1/2}}_T - \underbrace{\bar{P}^{-1/2} \bar{Q} \bar{P}^{-1/2}}_{\bar{T}}\|_2 &\leq \frac{2\sqrt{2}(\Delta_{AD}^+)^{1/2}}{[\tau^+ p(\frac{n}{2} - 1 + \eta)]^{3/2}} \left( \frac{n}{2} p + \tau^- p(\frac{n}{2} - \eta) \right) + \frac{\Delta_{AD}^-}{\tau^+ p(\frac{n}{2} - 1 + \eta)} \\ &\quad + \frac{2\sqrt{2}\Delta_{AD}^-(\Delta_{AD}^+)^{1/2}}{[\tau^+ p(\frac{n}{2} - 1 + \eta)]^{3/2}} + \frac{2\Delta_{AD}^+ \Delta_{AD}^-}{[\tau^+ p(\frac{n}{2} - 1 + \eta)]^2} \\ &\quad + \frac{2\Delta_{AD}^+}{[\tau^+ p(\frac{n}{2} - 1 + \eta)]^2} \left[ \frac{n}{2} p + \tau^- p(\frac{n}{2} - \eta) \right]. \end{aligned}$$

*Proof.* To begin with, we have via triangle inequality that

$$\|L^+ - \mathbb{E}[L^+]\|_2 \leq \Delta_A + \Delta_D, \quad \|L^- - \mathbb{E}[L^-]\|_2 \leq \Delta_A + \Delta_D.$$

This in turn implies the bounds

$$\|P - \bar{P}\|_2 \leq \Delta_A + \Delta_D + \tau^+ \Delta_D = \Delta_A + \Delta_D(1 + \tau^+) \quad (=:\Delta_{AD}^+), \quad (\text{D.6})$$

$$\|Q - \bar{Q}\|_2 \leq \Delta_A + \Delta_D + \tau^- \Delta_D = \Delta_A + \Delta_D(1 + \tau^-) \quad (=:\Delta_{AD}^-). \quad (\text{D.7})$$

By Weyl's inequality [56] (see Theorem 4), it follows for each  $l = 1, \dots, n$  that

$$\lambda_l(P) \in [\lambda_l(\bar{P}) \pm (\Delta_A + \Delta_D(1 + \tau^+))], \quad \lambda_l(Q) \in [\lambda_l(\bar{Q}) \pm (\Delta_A + \Delta_D(1 + \tau^-))]. \quad (\text{D.8})$$

By inspection, the eigenvalues of  $\bar{P}, \bar{Q}$  are easily derived as detailed below.

1.  $\lambda_1(\bar{P}) = \lambda_1(\mathbb{E}[L^-]) + \tau^+ p \left( \frac{n}{2} - 1 + \eta \right) = pn(1 - \eta) + \tau^+ p \left( \frac{n}{2} - 1 + \eta \right)$ .
2.  $\lambda_n(\bar{P}) = 0 + \tau^+ p \left( \frac{n}{2} - 1 + \eta \right) = \tau^+ p \left( \frac{n}{2} - 1 + \eta \right)$ .
3.  $\lambda_l(\bar{P}) = \frac{n}{2} p + \tau^+ p \left( \frac{n}{2} - 1 + \eta \right), \quad \forall l = 2, \dots, n - 1$ .
4.  $\lambda_l(\bar{Q}) = \lambda_l(\mathbb{E}[L^+]) + \tau^- p \left( \frac{n}{2} - \eta \right) = \frac{n}{2} p + \tau^- p \left( \frac{n}{2} - \eta \right), \quad \forall l = 1, \dots, n - 2$ .
5.  $\lambda_{n-1}(\bar{Q}) = pn\eta + \tau^- p \left( \frac{n}{2} - \eta \right)$  and  $\lambda_n(\bar{Q}) = \tau^- p \left( \frac{n}{2} - \eta \right)$ .

Now using (D.8) we can bound the extremal eigenvalues of  $P, Q$  as follows.

1.

$$\begin{aligned} \lambda_n(P) &\geq \lambda_n(\bar{P}) - (\Delta_A + \Delta_D(1 + \tau^+)) \\ &= \tau^+ p \left( \frac{n}{2} - 1 + \eta \right) - (\Delta_A + \Delta_D(1 + \tau^+)) \geq \frac{\tau^+ p}{2} \left( \frac{n}{2} - 1 + \eta \right) > 0 \end{aligned}$$

if  $(\Delta_A + \Delta_D(1 + \tau^+)) \leq \frac{\tau^+ p}{2} \left( \frac{n}{2} - 1 + \eta \right)$  and  $n \geq 2$ .

2.

$$\lambda_1(Q) \leq \lambda_1(\bar{Q}) - (\Delta_A + \Delta_D(1 - \tau^-)) = \frac{n}{19} \frac{p}{2} + \tau^- p \left( \frac{n}{2} - \eta \right) + (\Delta_A + \Delta_D(1 - \tau^-)).$$

$$3. \lambda_n(Q) \geq \lambda_n(\bar{Q}) - (\Delta_A + \Delta_D(1 + \tau^-)) = \tau^- p \left(\frac{n}{2} - \eta\right) - (\Delta_A + \Delta_D(1 - \tau^-)) \geq \frac{\tau^- p}{2} \left(\frac{n}{2} - \eta\right) > 0 \text{ if } (\Delta_A + \Delta_D(1 + \tau^-)) \leq \frac{\tau^- p}{2} \left(\frac{n}{2} - \eta\right).$$

$$4. \lambda_1(P) \leq \lambda_1(\bar{P}) - (\Delta_A + \Delta_D(1 + \tau^+)) = pn(1 - \eta) + \tau^+ p \left(\frac{n}{2} - 1 + \eta\right) + (\Delta_A + \Delta_D(1 + \tau^+)).$$

Next, we would like to bound the following quantity

$$\| \underbrace{P^{-1/2} Q P^{-1/2}}_T - \underbrace{\bar{P}^{-1/2} \bar{Q} \bar{P}^{-1/2}}_{\bar{T}} \|_2$$

where  $P, Q, \bar{P}, \bar{Q} \succ 0$ . Before proceeding, let us observe that as a consequence of the bounds on the spectra of  $P, \bar{P}$ , we obtain

$$\|P^{-1/2}\|_2 \leq \left(\frac{2}{\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)}\right)^{1/2}, \quad \|\bar{P}^{-1/2}\|_2 = \left(\frac{1}{\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)}\right)^{1/2}. \quad (\text{D.9})$$

Moreover, since  $P, \bar{P} \succ 0$ , therefore

$$\|P^{1/2} - \bar{P}^{1/2}\|_2 \leq \|P - \bar{P}\|_2^{1/2} \quad (\text{D.10})$$

holds as  $(\cdot)^{1/2}$  is operator monotone (see [8, Theorem X.1.1]). With these observations in mind, we obtain the bound

$$\begin{aligned} \|P^{-1/2} - \bar{P}^{-1/2}\|_2 &= \|P^{-1/2}(P^{1/2} - \bar{P}^{1/2})\bar{P}^{-1/2}\|_2 \\ &\leq \|P^{-1/2}\|_2 \|P^{1/2} - \bar{P}^{1/2}\|_2 \|\bar{P}^{-1/2}\|_2 \quad (\text{submultiplicativity of } \|\cdot\|_2 \text{ norm}) \\ &\leq \|P^{-1/2}\|_2 \|P - \bar{P}\|_2^{1/2} \|\bar{P}^{-1/2}\|_2 \quad (\text{due to (D.10)}) \\ &\leq \left(\frac{2}{\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)}\right)^{1/2} \left(\frac{1}{\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)}\right)^{1/2} \underbrace{(\Delta_A + \Delta_D(1 + \tau^+))^{1/2}}_{\Delta_{AD}^+} \quad (\text{due to (D.9), (D.6)}) \\ &= \frac{\sqrt{2}(\Delta_{AD}^+)^{1/2}}{\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)}. \end{aligned}$$

Therefore, denoting  $P^{-1/2} = \bar{P}^{-1/2} + E_P$  and  $Q = \bar{Q} + E_Q$ , we have shown thus far

$$\|E_P\|_2 \leq \frac{\sqrt{2}}{\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)} (\Delta_{AD}^+)^{1/2}, \quad \|E_Q\|_2 \leq \Delta_{AD}^-.$$

Also, from the spectra of  $\bar{Q}$  computed earlier, we see that  $\|\bar{Q}\|_2 = \left(\frac{n}{2} p + \tau^- p \left(\frac{n}{2} - \eta\right)\right)$ . Using these bounds, we can now upper bound  $\|T - \bar{T}\|_2$  as follows.

$$\begin{aligned} \|T - \bar{T}\|_2 &= \|(\bar{P}^{-1/2} + E_P)(\bar{Q} + E_Q)(\bar{P}^{-1/2} + E_P) - \bar{P}^{-1/2} \bar{Q} \bar{P}^{-1/2}\|_2 \\ &\leq \|\bar{P}^{-1/2} \bar{Q} E_P\|_2 + \|\bar{P}^{-1/2} E_Q \bar{P}^{-1/2}\|_2 + \|\bar{P}^{-1/2} E_Q E_P\|_2 \\ &\quad + \|E_P \bar{Q} \bar{P}^{-1/2}\|_2 + \|E_P \bar{Q} E_P\|_2 + \|E_P E_Q \bar{P}^{-1/2}\|_2 + \|E_P E_Q E_P\|_2 \quad (\text{triangle inequality}) \\ &\leq 2 \|E_P\|_2 \|\bar{Q}\|_2 \|\bar{P}^{-1/2}\|_2 + \|\bar{P}^{-1/2}\|_2^2 \|E_Q\|_2 \\ &\quad + 2 \|\bar{P}^{-1/2}\|_2 \|E_P\|_2 \|E_Q\|_2 + \|E_P\|_2^2 \|E_Q\|_2 + \|E_P\|_2^2 \|\bar{Q}\|_2 \quad (\text{submultiplicativity of } \|\cdot\|_2 \text{ norm}) \\ &\leq \frac{2\sqrt{2}(\Delta_{AD}^+)^{1/2}}{[\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)]^{3/2}} \left(\frac{n}{2} p + \tau^- p \left(\frac{n}{2} - \eta\right)\right) + \frac{\Delta_{AD}^-}{\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)} + \frac{2\sqrt{2}\Delta_{AD}^- (\Delta_{AD}^+)^{1/2}}{[\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)]^{3/2}} \\ &\quad + \frac{2\Delta_{AD}^+ \Delta_{AD}^-}{[\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)]^2} + \frac{2\Delta_{AD}^+}{[\tau^+ p \left(\frac{n}{2} - 1 + \eta\right)]^2} \left[\frac{n}{2} p + \tau^- p \left(\frac{n}{2} - \eta\right)\right]. \end{aligned}$$

#### D.4 Step 4: Concentration bounds for $A^+, A^-, D^+, D^-$

**Lemma 4.** *The following holds true.*

1. For every  $0 < \varepsilon \leq 1/2$ , there is a constant  $c_\varepsilon > 0$  such that

$$\mathbb{P}\left(\|A^+ - \mathbb{E}[A^+]\|_2 \leq \underbrace{((1 + \varepsilon)2\sqrt{2} + 1)\sqrt{np}}_{\Delta_A}\right) \geq 1 - n \exp\left(\frac{-pn}{c_\varepsilon}\right).$$

2. For every  $0 < \varepsilon \leq 1/2$ , there is a constant  $c_\varepsilon > 0$  such that

$$\mathbb{P}\left(\|A^- - \mathbb{E}[A^-]\|_2 \leq \underbrace{((1 + \varepsilon)2\sqrt{2} + 1)\sqrt{np}}_{\Delta_A}\right) \geq 1 - n \exp\left(\frac{-pn}{c_\varepsilon}\right).$$

3. If  $p > \frac{6 \log n}{\frac{n}{2} - 1 + \eta}$  then

$$\mathbb{P}\left(\|D^+ - \mathbb{E}[D^+]\|_2 \leq \underbrace{\sqrt{3pn \log n}}_{\Delta_D}\right) \geq 1 - \frac{2}{n}.$$

4. If  $p > \frac{6 \log n}{\frac{n}{2} - \eta}$  then

$$\mathbb{P}\left(\|D^- - \mathbb{E}[D^-]\|_2 \leq \underbrace{\sqrt{3pn \log n}}_{\Delta_D}\right) \geq 1 - \frac{2}{n}.$$

*Proof. Bounding  $\|A^+ - \mathbb{E}[A^+]\|_2$ .* Recall that  $A^+$  is a symmetric matrix with  $A_{ii}^+ = 0$  and where the random variables  $(A_{ij}^+)_{i < j}$  are independent and defined in (C.4) (when  $i, j$  are in same cluster) and (C.6) (when  $i, j$  are in different clusters). Let us denote  $Z_{ij}^+ = A_{ij}^+ - \mathbb{E}[A_{ij}^+]$  so that  $Z_{ii}^+ = 0$ ,  $Z_{ij}^+ = Z_{ji}^+$  and  $(Z_{ij}^+)_{i < j}$  are independent centered random variables defined as follows.

$$\begin{array}{cc} \underline{i, j \text{ lie in same cluster}} & \underline{i, j \text{ lie in different clusters}} \\ Z_{ij}^+ = \begin{cases} 1 - p(1 - \eta) & ; \text{ w. p. } p(1 - \eta) \\ -p(1 - \eta) & ; \text{ w. p. } 1 - p(1 - \eta) \end{cases} & , \quad Z_{ij}^+ = \begin{cases} 1 - p\eta & ; \text{ w. p. } p\eta \\ -p\eta & ; \text{ w. p. } 1 - p\eta \end{cases} . \end{array}$$

For  $i, j$  in the same cluster, we have

$$\begin{aligned} \mathbb{E}[(Z_{ij}^+)^2] &= p(1 - \eta)[1 - p(1 - \eta)]^2 + [1 - p(1 - \eta)]p^2(1 - \eta)^2 \\ &= p(1 - \eta)[1 - p(1 - \eta)][1 - p(1 - \eta) + p(1 - \eta)] \\ &= p(1 - \eta)[1 - p(1 - \eta)]. \end{aligned}$$

For  $i, j$  in different clusters, we have

$$\mathbb{E}[(Z_{ij}^+)^2] = p\eta(1 - p\eta)^2 + p^2\eta^2(1 - p\eta) = p\eta(1 - p\eta).$$

Hence for each  $i = 1, \dots, n$  we have that

$$\begin{aligned} \sqrt{\sum_{j=1}^n \mathbb{E}[(Z_{ij}^+)^2]} &= \sqrt{p(1 - \eta)[1 - p(1 - \eta)]\left(\frac{n}{2} - 1\right) + \frac{n}{2}p\eta(1 - p\eta)} \\ &\leq \sqrt{\frac{n}{2}p \left[ \underbrace{(1 - \eta)[1 - p(1 - \eta)]}_{\leq 1} + \underbrace{\eta(1 - p\eta)}_{\leq 1} \right]} \leq \sqrt{np}. \end{aligned}$$

Hence,  $\tilde{\sigma}^+ := \max_i \sqrt{\sum_{j=1}^n \mathbb{E}[(Z_{ij}^+)^2]} \leq \sqrt{np}$ . Moreover,  $\tilde{\sigma}_*^+ := \max_{i,j} \|Z_{ij}^+\|_\infty \leq 1$ . Therefore we can bound  $\|Z^+\|_2 = \|A^+ - \mathbb{E}[A^+]\|_2$  using Theorem 7 (with  $t = \sqrt{np}$ ) which tells us that for any given  $0 < \varepsilon \leq 1/2$ ,

$$\|A^+ - \mathbb{E}[A^+]\|_2 \leq ((1 + \varepsilon)2\sqrt{2} + 1)\sqrt{np}$$

with probability at least  $1 - n \exp\left(\frac{-pn}{c_\varepsilon}\right)$ . Here  $c_\varepsilon > 0$  depends only on  $\varepsilon$ .

**Bounding  $\|A^- - \mathbb{E}[A^-]\|_2$ .** Using the mixture model defined in (C.5) and (C.7) for the subgraph of negative edges, we can proceed in an identical fashion as above by replacing  $\eta$  with  $1 - \eta$ . We then obtain for any given  $0 < \varepsilon \leq 1/2$  that

$$\|A^- - \mathbb{E}[A^-]\|_2 \leq ((1 + \varepsilon)2\sqrt{2} + 1)\sqrt{np}. \quad (\text{D.11})$$

with probability at least  $1 - n \exp\left(\frac{-pn}{c_\varepsilon}\right)$ .

**Bounding  $\|D^+ - \mathbb{E}[D^+]\|_2$ .** Note that for any given  $i$ ,  $D_{ii}^+ = \sum_{j=1}^n A_{ij}^+$  with  $(A_{ij}^+)_{j=1}^n$  being independent Bernoulli random variables. Denoting  $\mu = \mathbb{E}[D_{ii}^+] = p\left(\frac{n}{2} - 1 + \eta\right)$ , we then obtain via standard Chernoff bounds (see Theorem 6) that

$$\mathbb{P}(|D_{ii}^+ - \mu| \geq \delta\mu) \leq 2 \exp\left(-\frac{\mu\delta^2}{3}\right)$$

for any given  $\delta \in (0, 1)$ . Letting  $\delta = \sqrt{\frac{6 \log n}{\mu}}$  and assuming  $p > \frac{6 \log n}{\frac{n}{2} - 1 + \eta}$  (so  $\delta \in (0, 1)$ ), we have for any given  $i$  that

$$\mathbb{P}\left(|D_{ii}^+ - \mu| \geq \sqrt{6(\log n)\mu}\right) \leq 2 \exp(-2 \log n) = \frac{2}{n^2}.$$

Then by applying the union bound, we finally conclude that

$$\|D^+ - \mathbb{E}[D^+]\|_2 \leq \sqrt{6p\left(\frac{n}{2} - 1 + \eta\right) \log n} \leq \sqrt{3np \log n}$$

with probability at least  $1 - \frac{2}{n}$ .

**Bounding  $\|D^- - \mathbb{E}[D^-]\|_2$ .** For this quantity, we obtain the same bound as above with  $\eta$  replaced with  $1 - \eta$ . So we have that

$$\|D^- - \mathbb{E}[D^-]\|_2 \leq \sqrt{6p\left(\frac{n}{2} - \eta\right) \log n} \leq \sqrt{3np \log n}$$

with probability at least  $1 - \frac{2}{n}$  if  $p > \frac{6 \log n}{\frac{n}{2} - \eta}$ . This completes the proof.  $\square$

## D.5 Step 5: Putting it together

From Lemma 4, we can see via the union bound that all the events hold simultaneously with probability at least

$$1 - \frac{4}{n} - 2n \exp\left(\frac{-pn}{c_\varepsilon}\right), \quad (\text{D.12})$$

provided  $p > \frac{6 \log n}{\min\{\frac{n}{2} - 1 + \eta, \frac{n}{2} - \eta\}}$ . For  $0 \leq \eta < 1/2$ , we have  $\frac{n}{2} - 1 + \eta < \frac{n}{2} - \eta$ . Moreover, if  $n \geq 6$ , then  $\frac{n}{2} - 1 + \eta > \frac{n}{4}$ , and so the condition  $p \geq \frac{24 \log n}{n}$  clearly implies  $p > \frac{6 \log n}{\frac{n}{2} - 1 + \eta}$ .

Let us now look at the requirements in Lemma 3. Plugging

$$\Delta_A = ((1 + \varepsilon)2\sqrt{2} + 1)\sqrt{np}, \quad \Delta_D = \sqrt{3pn \log n}$$

from Lemma 4, and using the definition of  $\Delta_{AD}^+$ , we obtain

$$\begin{aligned} \Delta_{AD}^+ &= ((1 + \varepsilon)2\sqrt{2} + 1)\sqrt{np} + \sqrt{3pn \log n} \\ &\leq \left( ((1 + \varepsilon)2\sqrt{2} + 1)\sqrt{p} + \sqrt{3p} \right) \sqrt{n \log n} \\ &= \tilde{c}_\varepsilon \sqrt{np \log n} \end{aligned}$$

where  $\tilde{c}_\varepsilon = (1 + \varepsilon)2\sqrt{2} + 1 + \sqrt{3}$ . Now note that if  $n \geq 6$ , then

$$\frac{\tau^+ p}{2} \left( \frac{n}{2} - 1 + \eta \right) \geq \frac{\tau^+ p}{2} \left( \frac{n}{3} \right) = \frac{\tau^+ np}{6}. \quad (\text{D.13})$$

Therefore the condition  $\Delta_{AD}^+ \leq \frac{\tau^+ p}{2} \left( \frac{n}{2} - 1 + \eta \right)$  is satisfied if

$$\tilde{c}_\varepsilon \sqrt{np \log n} \leq \frac{\tau^+ np}{6} \Leftrightarrow \sqrt{\frac{np}{\log n}} \geq \frac{6\tilde{c}_\varepsilon}{\tau^+} \Leftrightarrow p \geq \left( \frac{36 \tilde{c}_\varepsilon^2}{(\tau^+)^2} \right) \frac{\log n}{n}.$$

In an identical fashion, one can readily verify that the condition

$$\Delta_{AD}^- \leq \frac{\tau^- p}{2} \left( \frac{n}{2} - \eta \right)$$

is satisfied if  $p \geq \left( \frac{36 \tilde{c}_\varepsilon^2}{(\tau^-)^2} \right) \frac{\log n}{n}$ . Since  $\frac{n}{2} - 1 + \eta \geq \frac{n}{3}$  holds if  $n \geq 6$ , then by using this along with the bounds  $\Delta_{AD}^+, \Delta_{AD}^- \leq \tilde{c}_\varepsilon \sqrt{np \log n}$  in Lemma 3, we obtain

$$\begin{aligned} \left\| \underbrace{P^{-1/2} Q P^{-1/2}}_T - \underbrace{\bar{P}^{-1/2} \bar{Q} \bar{P}^{-1/2}}_{\bar{T}} \right\|_2 &\leq \frac{2\sqrt{2}\tilde{c}_\varepsilon^{1/2} (np \log n)^{1/4} \left( \frac{n}{2} p (1 + \tau^-) \right)}{\left( \tau^+ p \frac{n}{3} \right)^{3/2}} + \frac{\tilde{c}_\varepsilon \sqrt{np \log n}}{\tau^+ p \frac{n}{3}} \\ &\quad + \frac{2\sqrt{2} \tilde{c}_\varepsilon^{3/2} (np \log n)^{3/4}}{\left( \tau^+ p \frac{n}{3} \right)^{3/2}} + \frac{2\tilde{c}_\varepsilon^2 (np \log n)}{\left( \tau^+ p \frac{n}{3} \right)^2} \\ &\quad + \frac{2\tilde{c}_\varepsilon \sqrt{np \log n} \left( \frac{n}{2} p (1 + \tau^-) \right)}{\left( \tau^+ p \frac{n}{3} \right)^{3/2}} \\ &= \frac{3^{3/2} \tilde{c}_\varepsilon^{1/2} (1 + \tau^-) (\log n)^{1/4}}{(\tau^+)^{3/2}} + \frac{3 \tilde{c}_\varepsilon (\log n)^{1/2}}{\tau^+} \\ &\quad + \frac{6^{3/2} \tilde{c}_\varepsilon^{3/2} (\log n)^{3/4}}{(\tau^+)^{3/2}} + \frac{18 \tilde{c}_\varepsilon^2 (\log n)}{(\tau^+)^2} + \frac{9\tilde{c}_\varepsilon (1 + \tau^-) (\log n)^{1/2}}{(\tau^+)^2} \end{aligned}$$

Since  $\frac{\log n}{np} \leq 1$  by assumption, the above bound simplifies to

$$\begin{aligned} \left\| P^{-1/2} Q P^{-1/2} - \bar{P}^{-1/2} \bar{Q} \bar{P}^{-1/2} \right\|_2 &\leq \left( \frac{3^{3/2} \sqrt{2} \tilde{c}_\varepsilon^{1/2} (1 + \tau^-)}{(\tau^+)^{3/2}} + \frac{3\tilde{c}_\varepsilon}{\tau^+} + \frac{6^{3/2} \tilde{c}_\varepsilon^{3/2}}{(\tau^+)^{3/2}} \right. \\ &\quad \left. + \frac{18 \tilde{c}_\varepsilon^2}{(\tau^+)^2} + \frac{9 \tilde{c}_\varepsilon (1 + \tau^-)}{(\tau^+)^2} \right) \left( \frac{\log n}{np} \right)^{1/4} \\ &= \bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4}. \end{aligned}$$

where  $\bar{c}(\varepsilon, \tau^+, \tau^-)$  is as defined in the statement of Theorem 8. To summarize, so far, we have seen that  $T = \bar{T} + R = \bar{P}^{-1/2} \bar{Q} \bar{P}^{-1/2} + R$  with  $\|R\|_2 \leq \bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4}$ .

Let  $(\lambda_i(T), v_i(T))$  denote the eigenpairs of  $T$  for  $i = 1, \dots, n$ . Using Weyl's inequality [56] (see Theorem 4), we obtain

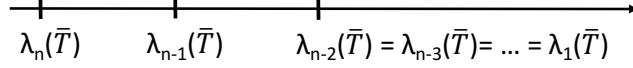
$$\lambda_i(T) \in \left[ \lambda_i(\bar{T}) \pm \bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4} \right] \quad \forall i = 1, \dots, n. \quad (\text{D.14})$$

In particular, this means that

$$\lambda_{n-2}(T) \geq \lambda_{n-2}(\bar{T}) - \bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4} > \lambda_{n-1}(\bar{T})$$

if the following condition holds.

$$\bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4} < \underbrace{\lambda_{n-2}(\bar{T}) - \lambda_{n-1}(\bar{T})}_{=\lambda_{gap}}. \quad (\text{D.15})$$



**Figure 12:** Spectrum of  $\bar{T}$ .

Now let  $V_2(T), V_2(\bar{T}) \in \mathbb{R}^{n \times 2}$  denote matrices whose columns are the eigenvectors corresponding to the smallest two eigenvalues of  $T, \bar{T}$  respectively. We then obtain from the Davis-Kahan theorem [17] (see Theorem 5) that

$$\|\sin \Theta(\mathcal{R}(V_2(T)), \mathcal{R}(V_2(\bar{T})))\|_2 = \|(I - V_2(T)V_2(T)^T)V_2(\bar{T})\|_2 \leq \frac{\|R\|_2}{\delta} \leq \frac{\bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4}}{\delta} \quad (\text{D.16})$$

holds, provided  $\delta := \min \left\{ |\hat{\lambda} - \lambda| : \lambda \in [\lambda_n(\bar{T}), \lambda_{n-1}(\bar{T})], \hat{\lambda} \in (-\infty, \underbrace{\lambda_{n+1}(T)}_{=-\infty}) \cup [\lambda_{n-2}(T), \infty) \right\} > 0$ . If (D.15) holds then we can see that

$$\begin{aligned} \delta &= \lambda_{n-2}(T) - \lambda_{n-1}(\bar{T}) \\ &\geq \lambda_{n-2}(\bar{T}) - \lambda_{n-1}(\bar{T}) - \bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4} \quad (\text{using (D.14)}) \\ &> 0. \end{aligned}$$

Recall the lower bound on  $\lambda_{n-2}(\bar{T}) - \lambda_{n-1}(\bar{T})$  from Lemma 2. Hence for any given  $\varepsilon \in (0, 1)$ , if the condition

$$\begin{aligned} \bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4} &\leq \varepsilon \min \left\{ \frac{2(1-\varepsilon_\tau)}{3(1+\tau^+)}, \frac{(1-2\eta)(3+\tau^++\tau^-)}{3(1+\tau^+)^2} \right\} \\ &\Leftrightarrow p \geq \left( \frac{\bar{c}(\varepsilon, \tau^+, \tau^-)}{\varepsilon \min \left\{ \frac{2(1-\varepsilon_\tau)}{3(1+\tau^+)}, \frac{(1-2\eta)(3+\tau^++\tau^-)}{3(1+\tau^+)^2} \right\}} \right)^4 \left( \frac{\log n}{n} \right) \end{aligned} \quad (\text{D.17})$$

holds, then  $\delta$  can be lower bounded as

$$\delta \geq (1-\varepsilon) \min \left\{ \frac{2(1-\varepsilon_\tau)}{3(1+\tau^+)}, \frac{(1-2\eta)(3+\tau^++\tau^-)}{3(1+\tau^+)^2} \right\}. \quad (\text{D.18})$$

Using (D.18), (D.17) in (D.16) we obtain the stated error bound on  $\|(I - V_2(T)V_2(T)^T)V_2(\bar{T})\|_2$ . This completes the proof.

## E Proof of Theorem 2

We will prove the following more precise version of Theorem 2 in this section.

**Theorem 9.** *Assuming  $\eta \in [0, 1/2)$  let  $\tau^+, \tau^- > 0$  satisfy*

$$\tau^- > \left( \frac{\eta}{1-\eta} \right) \left( \frac{\frac{n}{2} - 1 + \eta}{\frac{n}{2} - \eta} \right) \tau^+.$$

Then it holds that  $v_n(\bar{T}) = w$  with  $w$  defined in (4.4). 24

Moreover, assuming  $n \geq 6$ , for given  $0 < \varepsilon \leq 1/2$ ,  $\epsilon \in (0, 1)$  and  $\varepsilon_\tau \in (0, 1)$  let  $\tau^- \geq \frac{1}{\varepsilon_\tau} \left( \frac{\eta}{1-\eta} \right) \left( \frac{\frac{n}{2}-1+\eta}{\frac{n}{2}-\eta} \right) \tau^+$  and

$$p \geq \max \left\{ 24, \frac{36\tilde{c}_\varepsilon^2}{(\tau^+)^2}, \frac{36\tilde{c}_\varepsilon^2}{(\tau^-)^2}, \left( \frac{\bar{c}(\varepsilon, \tau^+, \tau^-)}{\epsilon \min \left\{ \frac{\eta(\frac{1}{\varepsilon_\tau}-1)}{1-\eta+\tau^+}, \frac{(1-2\eta)(3+\tau^++\tau^-)}{3(1+\tau^+)^2} \right\}} \right)^4 \right\} \left( \frac{\log n}{n} \right),$$

where  $\tilde{c}_\varepsilon$  and  $\bar{c}(\varepsilon, \tau^+, \tau^-)$  are as defined in Theorem 8. Then for  $c_\varepsilon > 0$  depending only on  $\varepsilon$ , it holds with probability at least  $\left( 1 - \frac{4}{n} - 2n \exp\left(\frac{-pn}{c_\varepsilon}\right) \right)$  that

$$\|(I - v_n(T)v_n(T)^T)w\|_2 \leq \frac{\epsilon}{1-\epsilon}.$$

*Proof of Theorem 9.* The proof is identical to that of Theorem 8 barring minor changes, hence we only highlight the differences.

To begin with, Lemma 1 holds as it is. Lemma 2 changes however as we now seek conditions under which  $w$  is the smallest eigenvector. This is stated in the following Lemma.

**Lemma 5** (Analogue of Lemma 2). *Let  $\tau^+, \tau^- > 0$  satisfy  $\tau^- > \frac{\eta}{1-\eta} \left( \frac{\frac{n}{2}-1+\eta}{\frac{n}{2}-\eta} \right) \tau^+$ . Then for  $\eta \in [0, \frac{1}{2})$ , the following is true.*

1.  $\lambda_n(\bar{T}) = \frac{n\eta+\tau^-(\frac{n}{2}-\eta)}{n(1-\eta)+\tau^+(\frac{n}{2}-1+\eta)}$  and  $\lambda_l(\bar{T}) \in \left\{ \frac{\tau^-(\frac{n}{2}-\eta)}{\tau^+(\frac{n}{2}-1+\eta)}, \frac{n+2\tau^-(\frac{n}{2}-\eta)}{n+2\tau^+(\frac{n}{2}-1+\eta)} \right\}$ , for  $l = 1, \dots, n-1$ .
2.  $v_n(\bar{T}) = w$ .

Moreover, if  $n \geq 6$  and for a given  $\varepsilon_\tau \in (0, 1)$ ,  $\tau^- \geq \frac{1}{\varepsilon_\tau} \frac{\eta}{1-\eta} \left( \frac{\frac{n}{2}-1+\eta}{\frac{n}{2}-\eta} \right) \tau^+$  holds, then the spectral gap ( $\lambda_{gap}$ ) between  $\lambda_n(\bar{T})$  and  $\{\lambda_{n-1}(\bar{T}), \dots, \lambda_1(\bar{T})\}$  satisfies

$$\lambda_{gap} = \lambda_{n-1}(\bar{T}) - \lambda_n(\bar{T}) \geq \min \left\{ \frac{\eta(\frac{1}{\varepsilon_\tau}-1)}{1-\eta+\tau^+}, \frac{(1-2\eta)(3+\tau^++\tau^-)}{3(1+\tau^+)^2} \right\}.$$

The proof of the Lemma is deferred to end of this section. Carrying on, Lemma's 3, 4 remain unchanged so we now just need to combine the results of Lemma's 1, 5, 3, 4.

Recall the bounds on the eigenvalues of  $T$  as in (D.14). This in particular implies that

$$\lambda_{n-1}(T) \geq \lambda_{n-1}(\bar{T}) - \bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4} > \lambda_n(\bar{T})$$

if the following condition holds.

$$\bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4} < \underbrace{\lambda_{n-1}(\bar{T}) - \lambda_n(\bar{T})}_{=\lambda_{gap}}. \quad (\text{E.1})$$

We then obtain from the Davis-Kahan theorem [17] (see Theorem 5) that

$$\|(I - ww^T)v_n(T)\|_2 \leq \frac{\bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4}}{\delta} \quad (\text{E.2})$$

holds provided  $\delta := \min \left\{ |\hat{\lambda} - \lambda| : \lambda = \lambda_n(\bar{T}), \hat{\lambda} \in [\lambda_{n-1}(T), \infty) \right\} > 0$ . Note that if (E.1) holds, then

$$\delta = \lambda_{n-1}(T) - \lambda_n(\bar{T}) \geq \lambda_{n-1}(\bar{T}) - \lambda_n(\bar{T}) - \bar{c}(\varepsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4} > 0.$$

Recall the lower bound on  $\lambda_{n-1}(\bar{T}) - \lambda_n(\bar{T})$  from Lemma 5. We then see for any given  $\epsilon \in (0, 1)$  that if the condition

$$\begin{aligned} \bar{c}(\epsilon, \tau^+, \tau^-) \left( \frac{\log n}{np} \right)^{1/4} &\leq \epsilon \min \left\{ \frac{\eta(\frac{1}{\epsilon_\tau} - 1)}{1 - \eta + \tau^+}, \frac{(1 - 2\eta)(3 + \tau^+ + \tau^-)}{3(1 + \tau^+)^2} \right\} \\ \Leftrightarrow p &\geq \left( \frac{\bar{c}(\epsilon, \tau^+, \tau^-)}{\epsilon \min \left\{ \frac{\eta(\frac{1}{\epsilon_\tau} - 1)}{1 - \eta + \tau^+}, \frac{(1 - 2\eta)(3 + \tau^+ + \tau^-)}{3(1 + \tau^+)^2} \right\}} \right)^4 \frac{\log n}{n} \end{aligned} \quad (\text{E.3})$$

holds then  $\delta$  can be lower bounded as

$$\delta \geq (1 - \epsilon) \left\{ \frac{\eta(\frac{1}{\epsilon_\tau} - 1)}{1 - \eta + \tau^+}, \frac{(1 - 2\eta)(3 + \tau^+ + \tau^-)}{3(1 + \tau^+)^2} \right\}. \quad (\text{E.4})$$

Finally, using (E.3), (E.4) in (E.2), we obtain the stated bound on  $\|(I - ww^T)v_n(T)\|_2$ . This completes the proof.  $\square$

*Proof of Lemma 5.* In the proof of Lemma 2, recall that  $\Lambda_{\bar{T}}^{(2)}$  is the eigenvalue of  $\bar{T}$  corresponding to the eigenvector  $w$ . We can see from (D.3), (D.5) that  $\Lambda_{\bar{T}}^{(2)} < \Lambda_{\bar{T}}^{(1)}, \Lambda_{\bar{T}}^{(3)}$  holds if  $n \geq 2, \eta \in [0, 1/2)$  and  $\tau^-, \tau^+$  satisfy  $\tau^- > \tau^+ \left( \frac{\eta(\frac{n}{2} - 1 + \eta)}{(1 - \eta)(\frac{n}{2} - \eta)} \right)$ .

Now, recall that if  $n \geq 6$ , then  $\Lambda_{\bar{T}}^{(3)} - \Lambda_{\bar{T}}^{(2)} \geq \frac{(1 - 2\eta)(3 + \tau^+ + \tau^-)}{3(1 + \tau^+)^2}$ . If  $\tau^-, \tau^+$  additionally satisfy

$$\tau^- \geq \frac{1}{\epsilon_\tau} \tau^+ \left( \frac{\eta(\frac{n}{2} - 1 + \eta)}{(1 - \eta)(\frac{n}{2} - \eta)} \right), \quad \text{for } \epsilon_\tau \in (0, 1), \quad (\text{E.5})$$

then we can lower bound  $\Lambda_{\bar{T}}^{(1)} - \Lambda_{\bar{T}}^{(2)}$  as follows.

$$\begin{aligned} \Lambda_{\bar{T}}^{(1)} - \Lambda_{\bar{T}}^{(2)} &= \frac{\tau^-(\frac{n}{2} - \eta)}{\tau^+(\frac{n}{2} - 1 + \eta)} - \frac{n\eta + \tau^-(\frac{n}{2} - \eta)}{n(1 - \eta) + \tau^+(\frac{n}{2} - 1 + \eta)} \\ &= n \frac{\tau^-(1 - \eta)(\frac{n}{2} - \eta) - \tau^+\eta(\frac{n}{2} - 1 + \eta)}{\tau^+(\frac{n}{2} - 1 + \eta) \left[ n(1 - \eta) + \tau^+(\frac{n}{2} - 1 + \eta) \right]} \\ &\geq n \left[ \frac{\eta\tau^+(\frac{n}{2} - 1 + \eta)(\frac{1}{\epsilon_\tau} - 1)}{\tau^+(\frac{n}{2} - 1 + \eta) \left[ n(1 - \eta) + \tau^+(\frac{n}{2} - 1 + \eta) \right]} \right] \quad (\text{using (E.5)}) \\ &= \frac{n\eta(\frac{1}{\epsilon_\tau} - 1)}{n(1 - \eta) + \tau^+(\frac{n}{2} - 1 + \eta)} \geq \frac{\eta(\frac{1}{\epsilon_\tau} - 1)}{1 - \eta + \tau^+}. \end{aligned}$$

$\underbrace{\hspace{10em}}_{\leq n}$

Hence the stated lower bound on the spectral gap follows, which completes the proof.  $\square$

## F Proof of Theorem 3

The proof is divided into the following steps.

**Step 1: Spectrum of  $\mathbb{E}[\bar{\mathbf{L}}]$ .** To begin with, we first observe for any  $i, j$  that

$$\mathbb{E}A_{ij} = \begin{cases} p(1 - 2\eta) & ; \text{ if } i, j \text{ lie in same cluster} \\ -p(1 - 2\eta) & ; \text{ if } i, j \text{ lie in different clusters} \\ 0 & ; \text{ if }_{26} i = j \end{cases} \quad (\text{F.1})$$

Due to the construction of  $C_1, C_2$  as per the SSBM, this means that

$$\mathbb{E}[A] = \frac{n}{2} \left\{ \begin{array}{cc} p(1-2\eta)\mathbf{1}\mathbf{1}^T & \overbrace{-p(1-2\eta)\mathbf{1}\mathbf{1}^T}^{n/2} \\ -p(1-2\eta)\mathbf{1}\mathbf{1}^T & p(1-2\eta)\mathbf{1}\mathbf{1}^T \end{array} \right\} - p(1-2\eta)I = M - p(1-2\eta)I.$$

$M$  is clearly a rank 1 matrix, indeed,  $M = np(1-2\eta)ww^T$  where  $w$  is defined in (4.4). Therefore, we obtain

$$\lambda_i(\mathbb{E}[A]) = \begin{cases} p(n-1)(1-2\eta) & ; \quad i = 1 \\ -p(1-2\eta) & ; \quad i = 2, \dots, n \end{cases}, \quad (\text{F.2})$$

and also  $v_1(\mathbb{E}[A]) = v_1(M) = w$ . Moreover, one can easily check that  $\mathbb{E}[\bar{D}] = (n-1)pI$ . Therefore  $\mathbb{E}[\bar{L}] = \mathbb{E}[\bar{D}] - \mathbb{E}[A] = (n-1)pI - \mathbb{E}[A]$  and hence

$$\lambda_i(\mathbb{E}[\bar{L}]) = \begin{cases} 2\eta(n-1)p & ; \quad i = n \\ (n-1)p + p(1-2\eta) = (n-2\eta)p & ; \quad i = 1, \dots, n-1 \end{cases}, \quad (\text{F.3})$$

with  $v_n(\mathbb{E}[\bar{L}]) = v_n(\mathbb{E}[A]) = w$ .

**Step 2: Bounding  $\|\bar{L} - \mathbb{E}[\bar{L}]\|_2$ .** Next, we will like to bound  $\|\bar{L} - \mathbb{E}[\bar{L}]\|_2$ . Since

$$\|\bar{L} - \mathbb{E}[\bar{L}]\|_2 \leq \|\bar{D} - \mathbb{E}[\bar{D}]\|_2 + \|A - \mathbb{E}[A]\|_2,$$

we will bound the terms on the RHS individually starting with the first term.

Recall that  $\bar{D}_{ii} = \sum_{j \neq i} |A_{ij}| = \sum_{j \neq i} Z_{ij}$  where  $Z_{ij} = 1$  with probability  $p$  and is 0 with probability  $1-p$ . Also, for a given  $i$ , note that  $(Z_{ij})_{j=1, j \neq i}^n$  are i.i.d. Therefore from Chernoff bounds for sums of independent Bernoulli random variables (see Theorem 6), it follows for any given  $\delta \in (0, 1)$  that

$$\begin{aligned} \mathbb{P}(|\bar{D}_{ii} - (n-1)p| \geq \delta(n-1)p) &\leq 2 \exp\left(-\frac{(n-1)p\delta^2}{3}\right) \\ &\leq 2 \exp\left(-\frac{np\delta^2}{6}\right) \quad (\text{if } n \geq 2). \end{aligned}$$

If  $p > \frac{12 \log n}{n}$  then we can set  $\delta = \sqrt{\frac{12 \log n}{np}}$  and apply the union bound. We then have that

$$\|\bar{D} - \mathbb{E}[\bar{D}]\|_2 = \max_i |\bar{D}_{ii} - \mathbb{E}[\bar{D}_{ii}]| \leq \sqrt{12pn \log n} \quad (\text{F.4})$$

with probability at least  $1 - \frac{2}{n}$ .

We now look to bound  $\|A - \mathbb{E}[A]\|_2$ . Since  $A$  is a random symmetric matrix with  $(A_{ij})_{i \leq j}$  being independent, bounded random variables, we will use Theorem 7 to bound  $\|A - \mathbb{E}[A]\|_2$  with high probability. For given  $i, j$  with  $i \neq j$ , if  $i, j$  belong to the same cluster, then

$$A_{ij} - \mathbb{E}[A_{ij}] = \begin{cases} 1 - p(1-2\eta) & ; \quad \text{w. p. } p(1-\eta) \\ -1 - (1-2\eta)p & ; \quad \text{w. p. } p\eta \\ -p(1-2\eta) & ; \quad \text{w. p. } (1-p) \end{cases}, \quad (\text{F.5})$$

and if  $i, j$  belong to different clusters, then

$$A_{ij} - \mathbb{E}[A_{ij}] = \begin{cases} 1 + (1-2\eta)p & ; \quad \text{w. p. } p\eta \\ -1 + (1-2\eta)p & ; \quad \text{w. p. } p(1-\eta) \\ (1-2\eta)p & ; \quad \text{w. p. } (1-p) \end{cases}. \quad (\text{F.6})$$

In order to use Theorem 7, we need to compute (upper bounds on) the quantities

$$\tilde{\sigma} := \max_i \sqrt{\sum_{j=1}^n \mathbb{E}[(A_{ij} - \mathbb{E}[A_{ij}])^2]}, \quad \tilde{\sigma}_* := \max_{i,j} \|A_{ij} - \mathbb{E}[A_{ij}]\|_\infty.$$

Note that  $\tilde{\sigma}_* \leq 1 + (1 - 2\eta)p \leq 2$ . Moreover, for any  $i \neq j$  (irrespective of whether in same cluster or not), we obtain from (F.5),(F.6) that

$$\begin{aligned} \mathbb{E}[(A_{ij} - \mathbb{E}[A_{ij}])^2] &= (1 - p(1 - 2\eta)^2)p(1 - \eta) + (1 + (1 - 2\eta)p)^2p\eta + (1 - p)p^2(1 - 2\eta)^2 \\ &= [(1 + p^2(1 - 2\eta)^2 - 2p(1 - 2\eta))(1 - \eta) + (1 + (1 - 2\eta)^2p^22(1 - 2\eta)p)\eta]p \\ &\quad + (1 - p)p^2(1 - 2\eta)^2 \\ &= [1 - p(1 - 2\eta)^2(2 - p)]p + (1 - p)p^2(1 - 2\eta)^2 \\ &= (1 - p(1 - 2\eta)^2)p \leq p. \end{aligned}$$

This gives us  $\tilde{\sigma} \leq \sqrt{n-1}\sqrt{p} \leq \sqrt{np}$ . Then using Theorem 7 with  $t = \sqrt{np}$  we obtain for any  $0 < \varepsilon \leq 1/2$  that

$$\mathbb{P}(\|A - \mathbb{E}[A]\|_2 \geq ((1 + \varepsilon)2\sqrt{2} + 1)\sqrt{np}) \leq n \exp\left(-\frac{pn}{4c_\varepsilon}\right), \quad (\text{F.7})$$

where  $c_\varepsilon > 0$  depends only on  $\varepsilon$ . Using (F.4), (F.7) and applying the union bound, we have that

$$\begin{aligned} \|\bar{L} - \mathbb{E}[\bar{L}]\|_2 &\leq \underbrace{((1 + \varepsilon)2\sqrt{2} + 1)\sqrt{np}}_{\geq \sqrt{12}} + \sqrt{12pn \log n} \\ &\leq 2((1 + \varepsilon)2\sqrt{2} + 1)\sqrt{np \log n} \end{aligned} \quad (\text{F.8})$$

holds with probability at least  $1 - \frac{2}{n} - n \exp\left(-\frac{pn}{4c_\varepsilon}\right)$ .

**Step 3: Using Davis-Kahan theorem.** Say  $\|\bar{L} - \mathbb{E}[\bar{L}]\|_2 \leq \Delta$  holds. Then from Weyl's inequality [56] (see Theorem 4), it holds that  $|\lambda_i(\bar{L}) - \lambda_i(\mathbb{E}[\bar{L}])| \leq \|\bar{L} - \mathbb{E}[\bar{L}]\|_2 \leq \Delta$  for  $i = 1, \dots, n$ . Moreover, from the Davis-Kahan theorem (see Theorem 5), we have that

$$\|(I - v_n(\mathbb{E}[\bar{L}])v_n(\mathbb{E}[\bar{L}])^T)v_n(\bar{L})\|_2 \leq \frac{\Delta}{|\lambda_{n-1}(\bar{L}) - \lambda_n(\mathbb{E}[\bar{L}])|} \quad (\text{F.9})$$

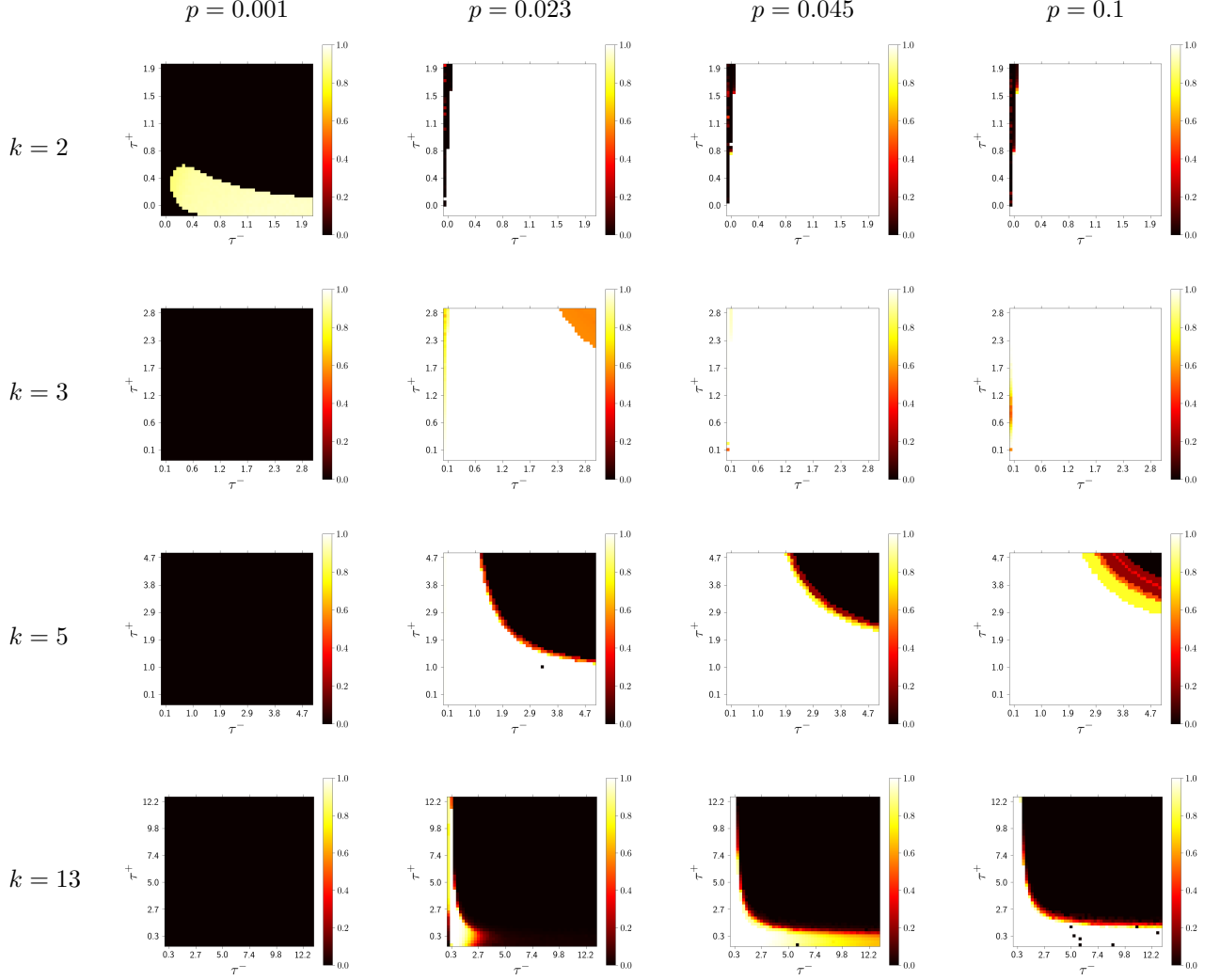
holds if  $|\lambda_{n-1}(\bar{L}) - \lambda_n(\mathbb{E}[\bar{L}])| > 0$ . Now,

$$\begin{aligned} \lambda_{n-1}(\bar{L}) - \lambda_n(\mathbb{E}[\bar{L}]) &\geq \lambda_{n-1}(\mathbb{E}[\bar{L}]) - \lambda_n(\mathbb{E}[\bar{L}]) - \Delta \\ &= (n - 2\eta)p - 2\eta(n - 1)p - \Delta \\ &= np(1 - 2\eta) - \Delta > 0 \end{aligned}$$

if  $\Delta < np(1 - 2\eta)$  holds. Therefore, for  $0 < \epsilon < 1$ , if  $\Delta \leq \epsilon np(1 - 2\eta)$  is satisfied, then from (F.9), we obtain the bound  $\|(I - v_n(\mathbb{E}[\bar{L}])v_n(\mathbb{E}[\bar{L}])^T)v_n(\bar{L})\|_2 \leq \frac{\epsilon}{1 - \epsilon}$ . Finally, from (F.8), we have that

$$\Delta < np(1 - 2\eta) \Leftrightarrow p \geq \frac{4((1 + \varepsilon)2\sqrt{2} + 1)^2 \log n}{\epsilon^2(1 - 2\eta)^2} \frac{1}{n}.$$

The above bound on  $p$  also implies  $p > 12 \log n/n$  which we required earlier for deriving (F.4). This completes the proof.



**Figure 13:** SSBM recovery of SPONGE with  $k - 1$  eigenvectors as a function of  $\tau^+$  and  $\tau^-$  with  $n = 5000$ ,  $\eta = 0.05$  and varying values of  $k$  and  $p$ .

## G Additional experiments

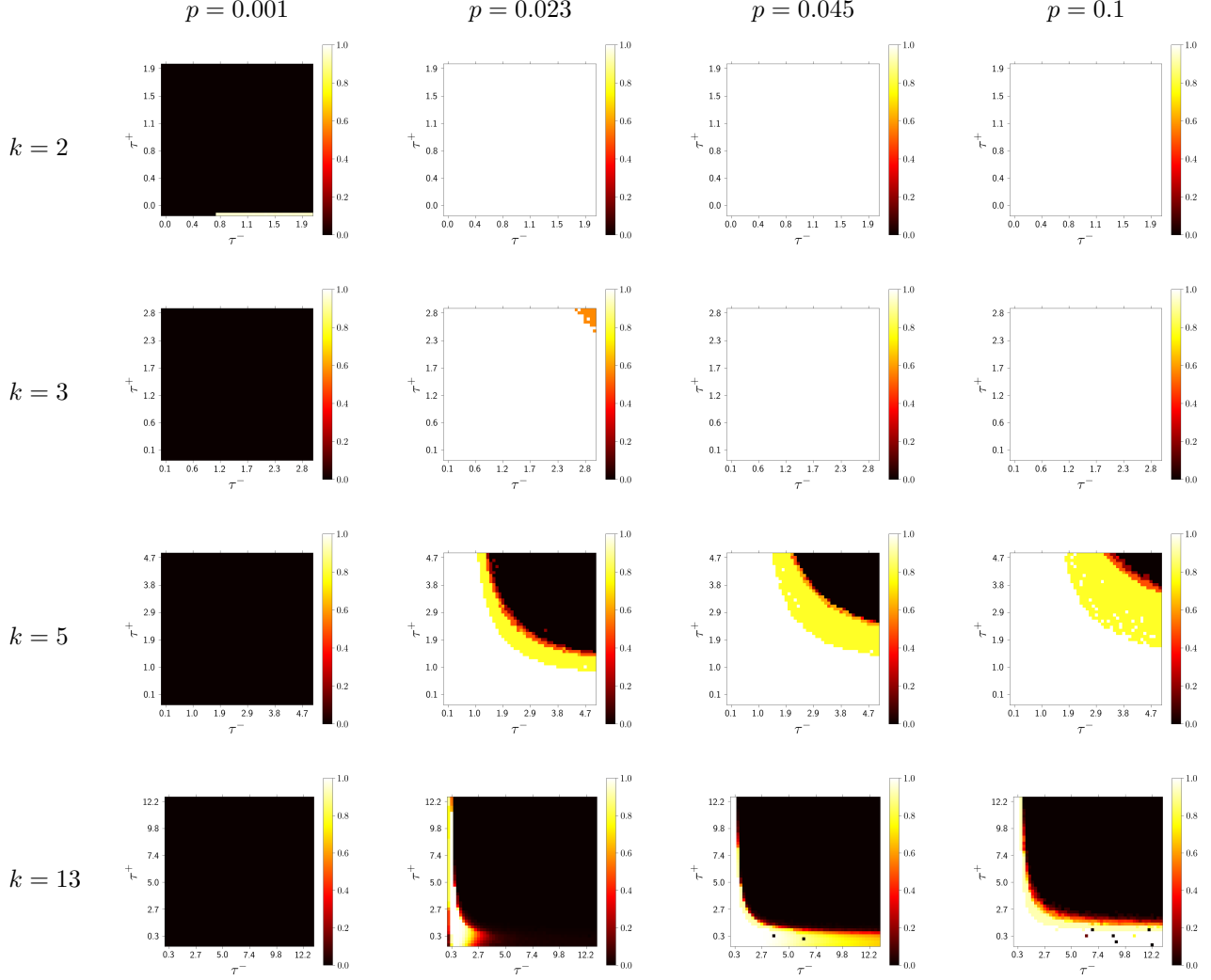
In this section we show the results of additional supporting experiments: plots demonstrating how the parameters  $\tau^+$  and  $\tau^-$  affect the performance of SPONGE and SPONGE<sub>sym</sub> in a range of regimes, and plots comparing the performance of all algorithms on the SSBM under a wide range of parameters  $p$  and  $\eta$ .

### G.1 Parameter analysis for $\tau^+$ and $\tau^-$

Here we plot performance of SPONGE and SPONGE<sub>sym</sub> while varying parameters  $\tau^+$  AND  $\tau^-$ , to motivate our choices thereof.

In Figure 13, we see the performance of SPONGE on SSBM graphs, taking  $k - 1$  eigenvectors, plotted for a range of values for  $\tau^+$  and  $\tau^-$ . In most cases, we see that performance is not too sensitive to the choice of these parameters. When there are regions of good and poor performance, we see that the region of good performance is concentrated around the axes, i.e. when either  $\tau^+$  and  $\tau^-$  are low. We note that the position  $\tau^+ = \tau^- = 1$  always falls within the region of maximum recovery when it is present (with the exception of the top left plot).

Figure 14 shows the same experiments but using the bottom  $k$  eigenvectors for the SPONGE algorithms instead of  $k - 1$ . The plots look similar to those of Figure 13 in most cases, but with two main differences: firstly, as we

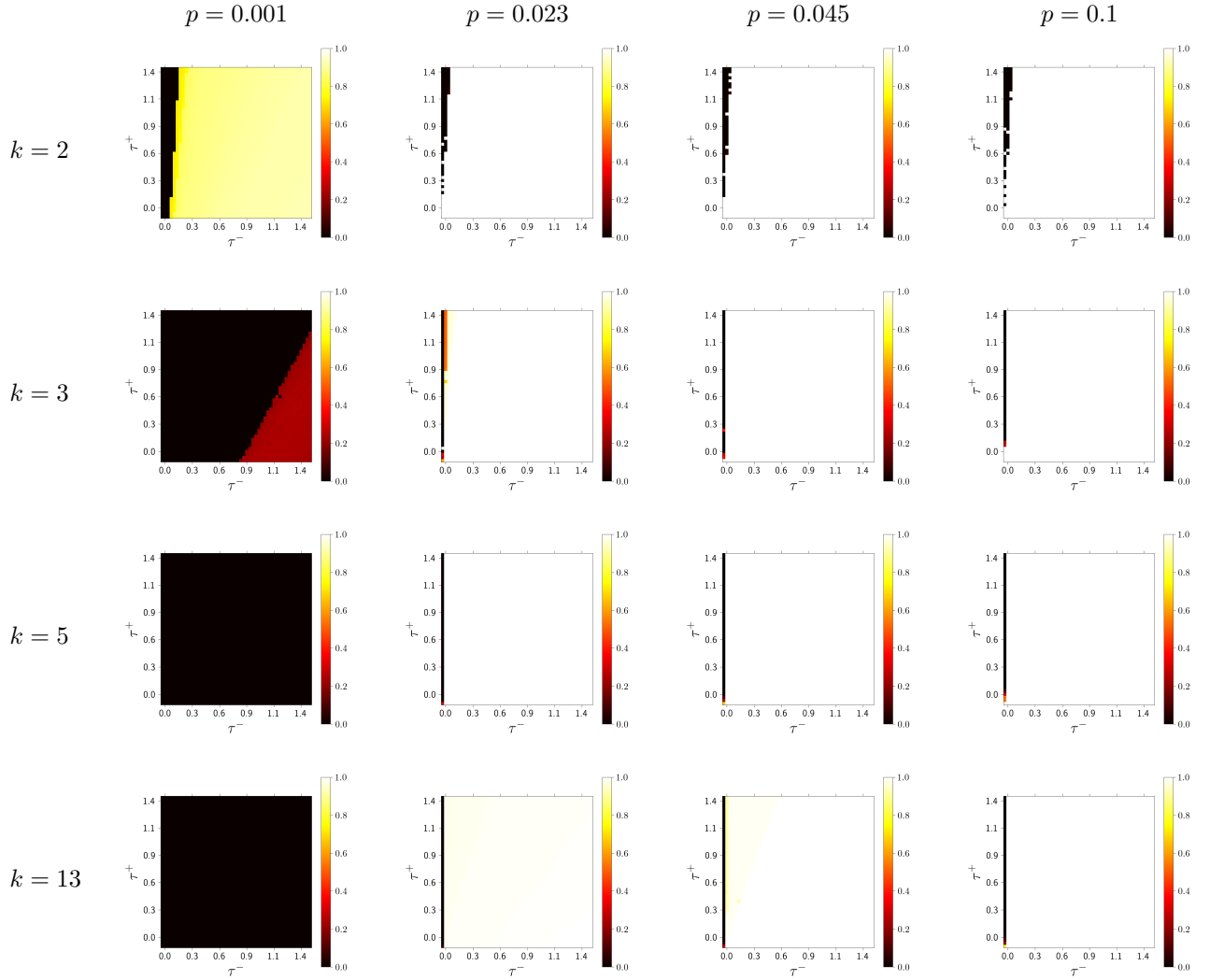


**Figure 14:** SSBM recovery of the SPONGE algorithm with  $k$  eigenvectors as a function of  $\tau^+$  and  $\tau^-$  with  $n = 5000$ ,  $\eta = 0.05$  and varying values of  $k$  and  $p$

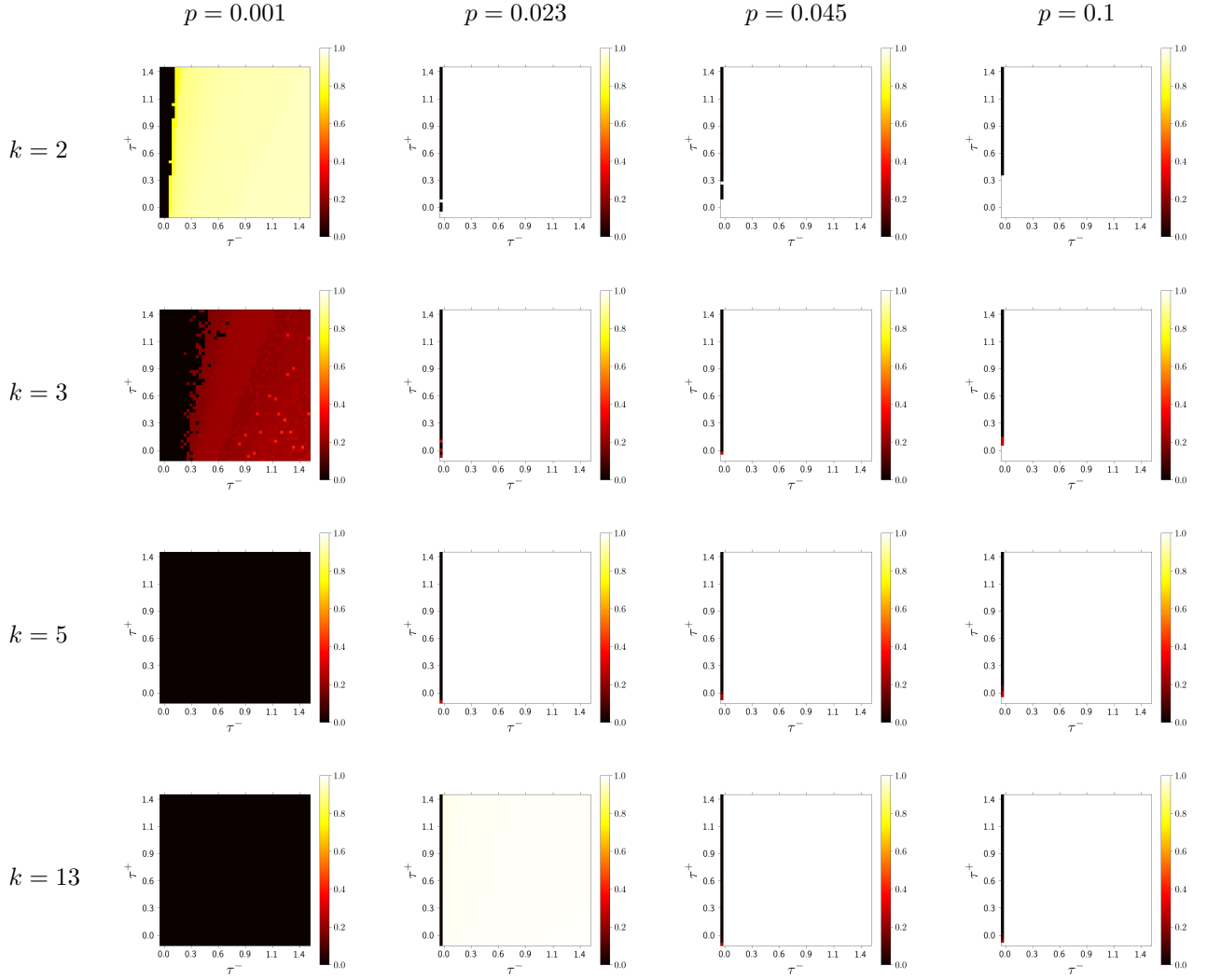
see most clearly from the rightmost three plots in the top row, there are some marginal regions where recovery with  $k - 1$  eigenvectors was poor, and with  $k$  eigenvectors is greatly improved. This is because previously an informative eigenvector was being displaced by a non-informative one, and so by taking the extra eigenvector we capture this useful information. However, there are also some regions (most prominent in the  $k = 5$  row) where recovery drops from near-perfect to mediocre. These are regions where we were already capturing all informative eigenvectors, so by taking another one we *dilute* the quality of our embedding with what is effectively an extra dimension of noise. In general, if we pick a natural parameter choice such as  $\tau^+ = \tau^- = 1$ , our recovery score is not improved, and in some cases is worsened, if we use  $k$  eigenvectors instead of  $k - 1$ .

The plots of Figure 15, for  $\text{SPONGE}_{sym}$  with  $k - 1$  eigenvectors, demonstrate that the values of  $\tau^+$  and  $\tau^-$  are even more flexible for this algorithm. In fact, as long as  $\tau^+$  is bounded sufficiently far away from zero, there is a large region where performance is equally good. In practice, we also want to bound  $\tau^-$  away from zero to avoid convergence issues with the eigenvalue solver. So again,  $\tau^+ = \tau^- = 1$  seems a reasonable choice.

When we use an additional eigenvector in our algorithm, shown in Figure 16, we again see (particularly in the  $k = 2$  row) some marginal regions where recovery is greatly improved, since we pick up an informative eigenvector we previously lost. Unlike for SPONGE, we do not see any decrease in performance due to *dilution* when we already recovered all informative eigenvectors, though this is perhaps because the parameter settings we plot do not seem to have captured ‘borderline’ regimes where such an effect might impact the recovery process.



**Figure 15:** SSBM recovery of the SPONGE<sub>sym</sub> algorithm with  $k - 1$  eigenvectors as a function of  $\tau^+$  and  $\tau^-$  with  $n = 5000$ ,  $\eta = 0.05$  and varying values of  $k$  and  $p$



**Figure 16:** SSBM recovery of the SPONGE<sub>sym</sub> algorithm with  $k$  eigenvectors as a function of  $\tau^+$  and  $\tau^-$  with  $n = 5000$ ,  $\eta = 0.05$  and varying values of  $k$  and  $p$

## G.2 Numerical experiments on the SSBM

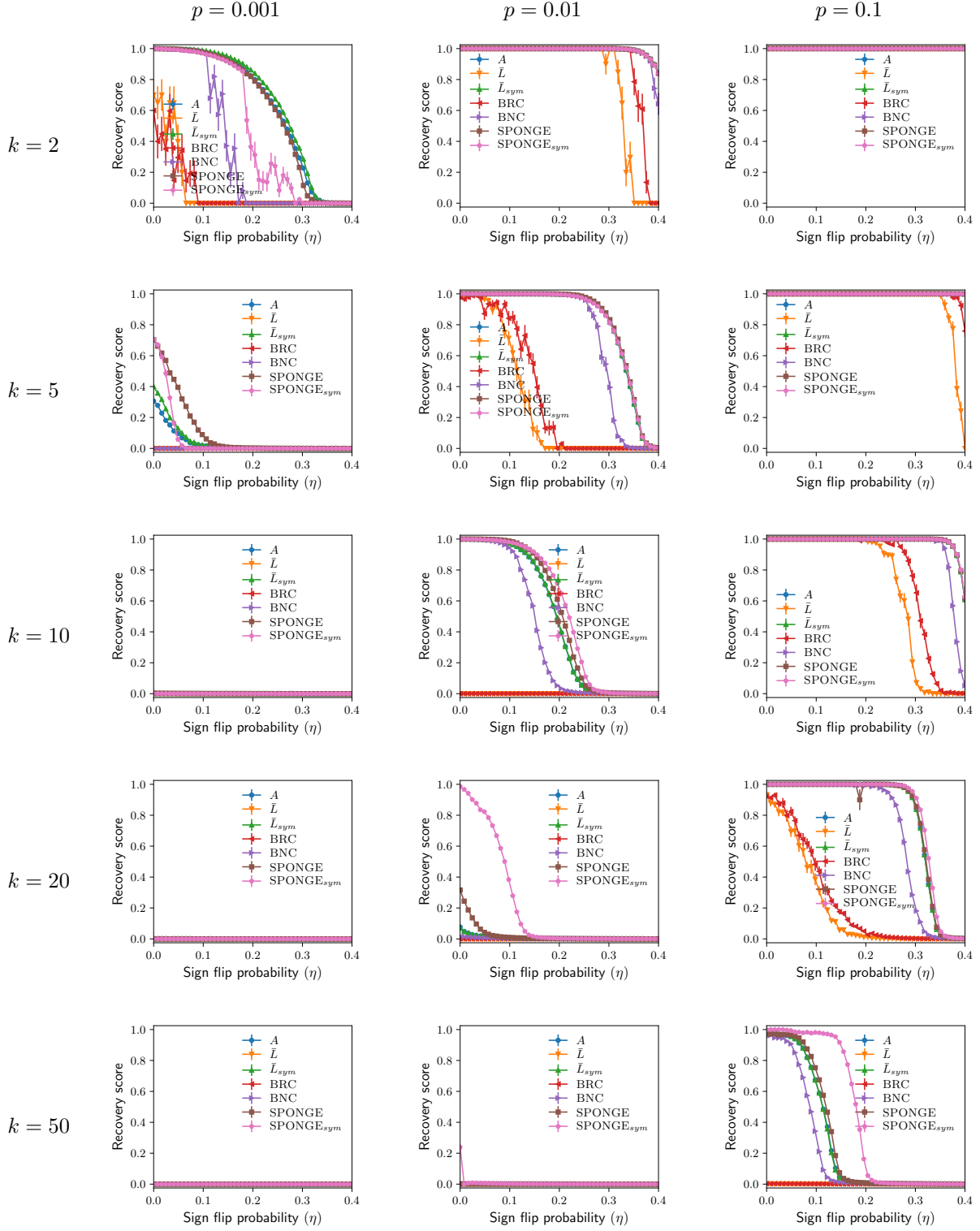
In the following figures we plot performance of SPONGE and SPONGE<sub>sym</sub> against five benchmark algorithms, on graphs generated from the SSBM, with equal size planted clusters, and with performance measured by the Adjusted Rand Index (ARI) against the ground truth.

In Figure 17 we fix some values for  $p$  and  $k$ , and plot the ARI score against the flip probability  $\eta$ . We see that for  $k = 2$ , the symmetric Signed Laplacian  $\bar{L}_{sym}$  of Kunegis et al. [34] can tolerate the highest  $\eta$ . However, as  $k$  increases, the previous state-of-the-art algorithms are quickly overtaken, first by SPONGE, and then SPONGE<sub>sym</sub>. When  $k$  is large (rows  $k = 20$  and  $k = 50$ ), SPONGE<sub>sym</sub> outperforms all other algorithms by a large margin.

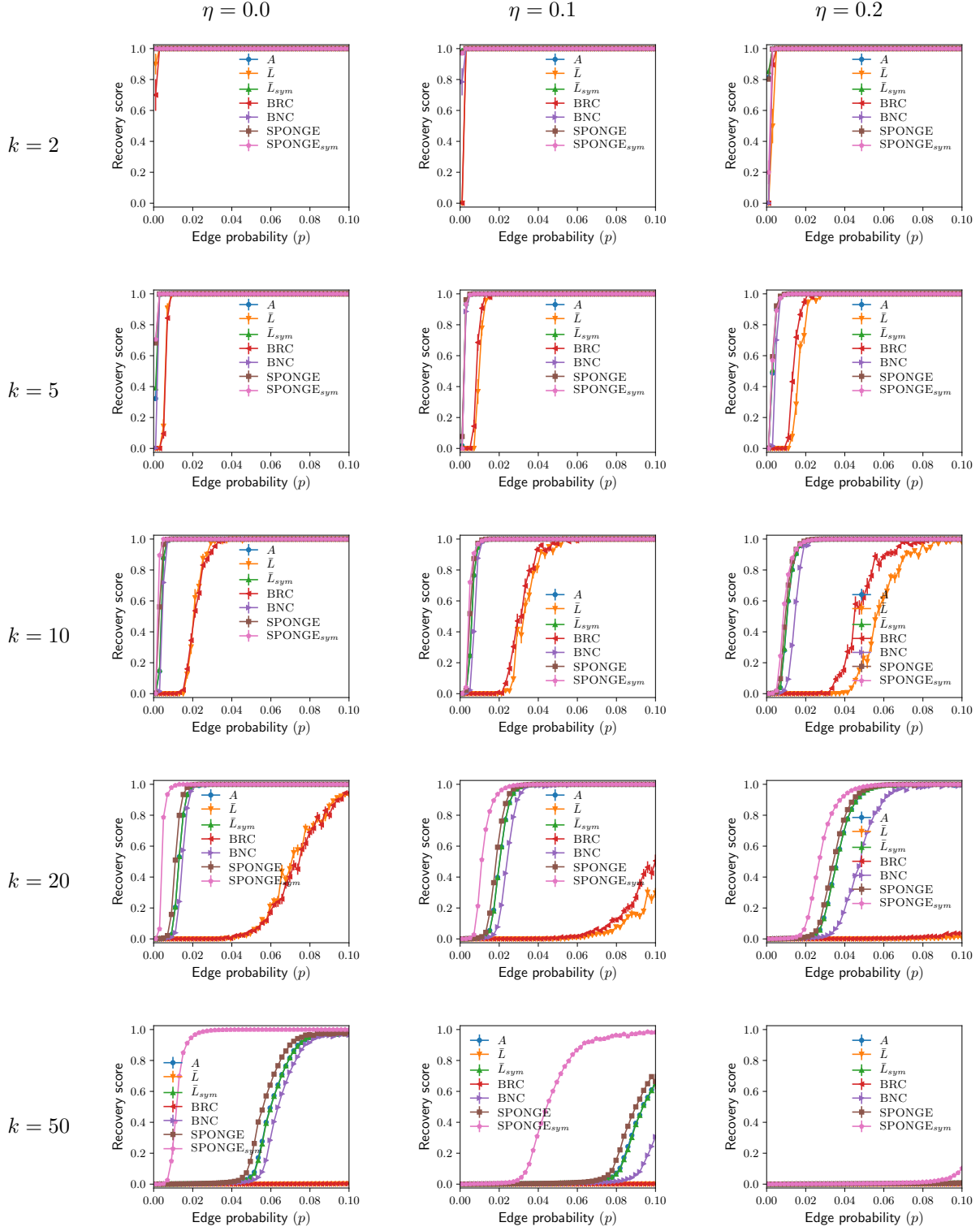
A similar story is shown by Figure 18, in which  $k$  and  $\eta$  are fixed and we plot the score against the sparsity parameter  $p$ . Here, the  $k = 2$  graphs are hard to infer comparative performance from, but we can again see that as  $k$  increases, SPONGE and particularly SPONGE<sub>sym</sub> quickly outstrip all other algorithms.

The purpose of Figure 19 is to ensure that we are not over-fitting to a synthetic regime where we know all communities are equally sized. We alter our graph generation process under the SSBM to produce unevenly-sized clusters as follows: first, each of the  $k$  (initially empty) clusters is given an *affinity* value, drawn uniformly at random from  $[0, 1]$ . Then, each node joins one of the  $k$  clusters at random, independently with probability proportional to the clusters’s affinity value. In this way, some clusters receive many nodes and some receive very few. Once cluster membership has been allocated, we generate edges as usual: they are present independently with probability  $p$ , initially have sign dependent on whether they fall within a community, and have their sign flipped with probability  $\eta$ .

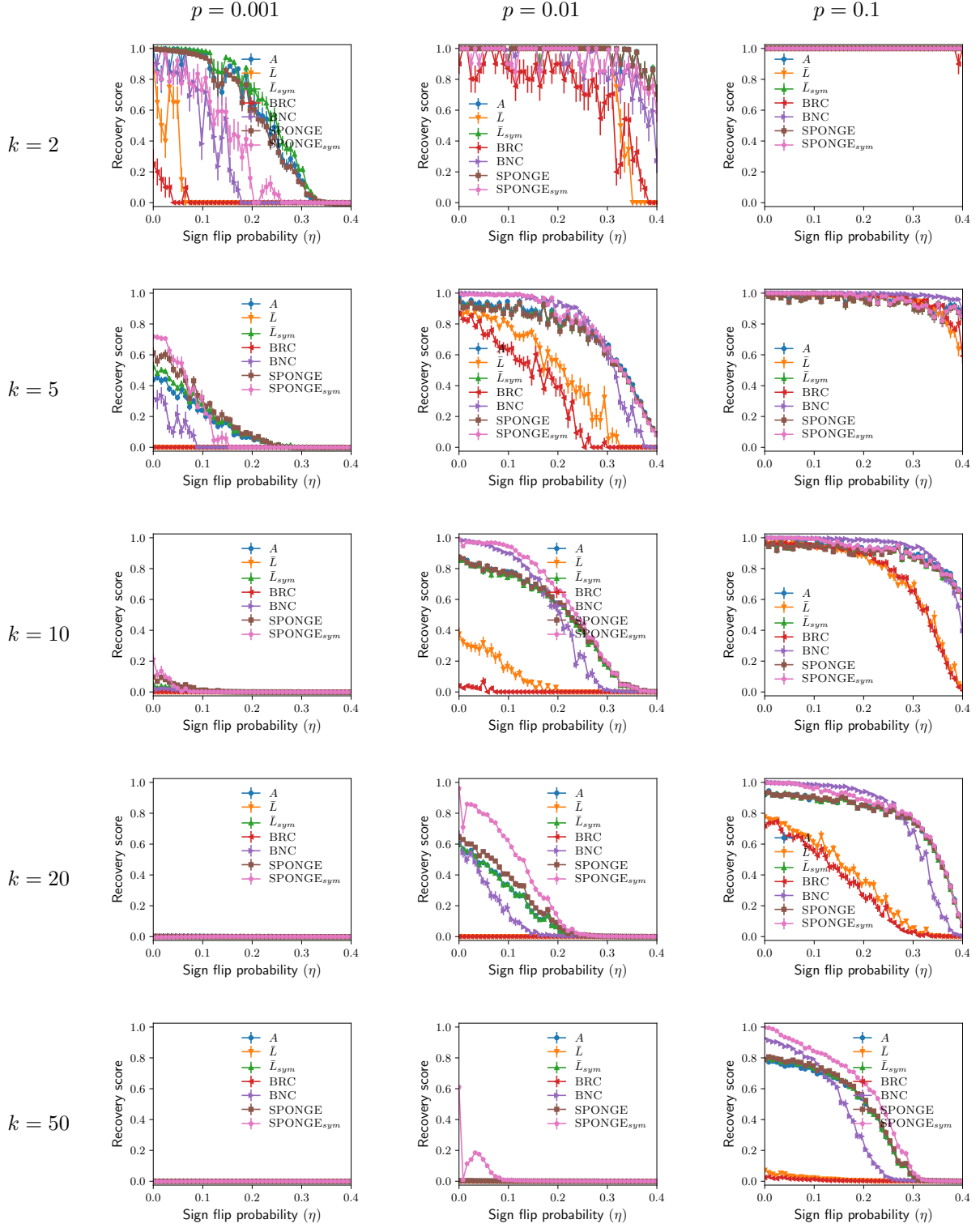
This set of experiments reveals that our scores have more variance, and for low  $k$  our proposed algorithms do comparatively worse than for uniform cluster sizes, and the Balance Normalized Cut (BNC) algorithm of Chiang et al. [13] performs comparatively better. As  $k$  increases, SPONGE and SPONGE<sub>sym</sub> once again begin to outperform the benchmark algorithms, though unlike in the uniform communities regime, the BNC algorithm remains competitive under some parameter settings.



**Figure 17:** Adjusted Rand Index achieved by several algorithms as a function of the noise level  $\eta$  for different values of  $p$  and  $k$ ,  $n = 10000$ , and clusters of fixed equal size.



**Figure 18:** Adjusted Rand Index achieved by several algorithms as a function of the edge probability  $p$  for different  $\eta$  and  $k$  values and  $n = 10000$  for clusters of fixed equal size.



**Figure 19:** Adjusted Rand Index achieved by several algorithms as a function of the flipping probability  $\eta$  for different  $p$  and  $k$  values and  $n = 10000$ . The size of each cluster was randomly chosen to be uneven.

### G.3 Numerical experiments on real data sets

This section details the results of our numerical experiments on two additional data sets.

**Correlations of financial market returns - S&P 500.** We detail here the results of our experiment on financial equity time series data, corresponding to constituents of S&P 500. The procedure for obtaining the signed network is the same as the one for S&P 1500, discussed in the main text.

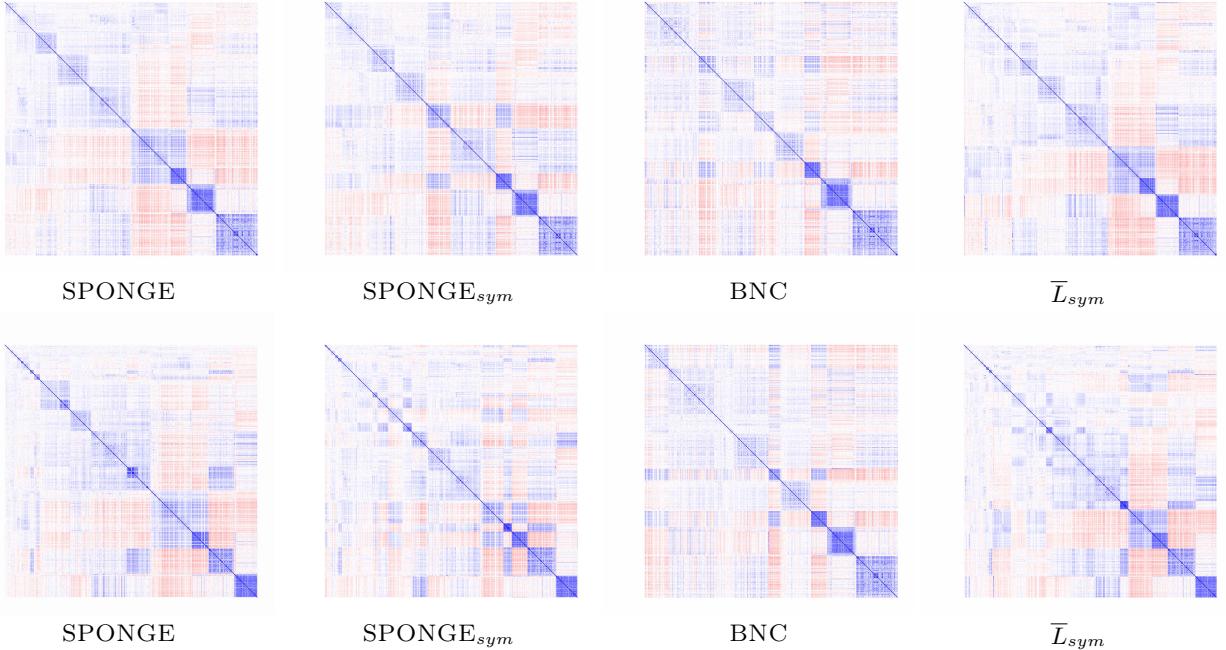
We consider time series price data for  $n = 500$  stocks in the S&P 500 Index, during 2003-2015, containing approximately  $n_d = 3000$  trading days. We work with daily log returns of the prices

$$R_{i,t} = \log \frac{P_{i,t}}{P_{i,t-1}}, \quad (\text{G.1})$$

where  $P_{i,t}$  denotes the market close price of instrument  $i$  on day  $t$ . Next, we compute the daily market excess return for each instrument,

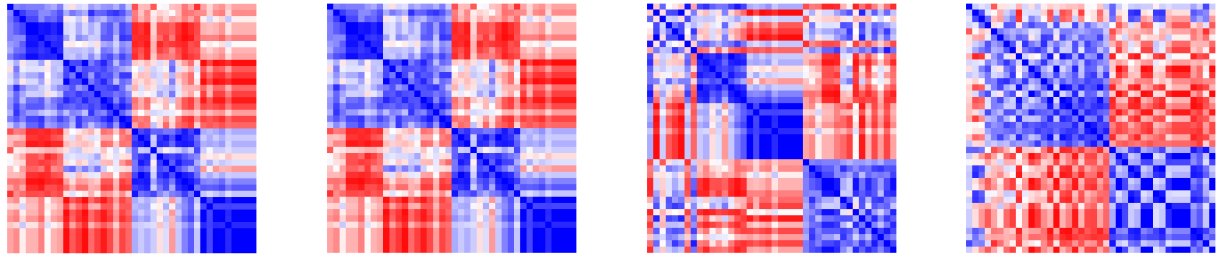
$$\tilde{R}_{i,t} = R_{i,t} - R_{\text{SPY},t}, \forall i = 1, \dots, n, t = 1, \dots, n_d \quad (\text{G.2})$$

where  $R_{\text{SPY},t}$  denotes the daily log return of SPY, the S&P 500 index ETF used as a proxy for the market. We then calculate the Pearson correlation coefficient between historical return for each pair of companies, and use that as an edge weight in our signed graph. Figure 20 shows that, for  $k = \{10, 20\}$ , we are able to recover the clustering structure covering the entire network.



**Figure 20:** Adjacency matrix of the S&P 500 data set sorted by cluster membership, for  $k = \{10, 20\}$ .

**Foreign Exchange correlations.** The SPONGE algorithms proved to work particularly well in the setting of clustering a Foreign Exchange matrix derived from daily Special Drawing Rights (SDR) exchange value rates [2]. Indeed, as shown in Figure 21, only SPONGE and SPONGE<sub>sym</sub> were able to recover four neat clusters associated respectively to the EURO €, US Dollar \$, UK Pound Sterling £, and Japanese Yen ¥. These are precisely the four currencies that, in certain percentage weights, define the value of the SDR reserve. Figure 22 shows that the recovered clusters align well with their geographic locations.



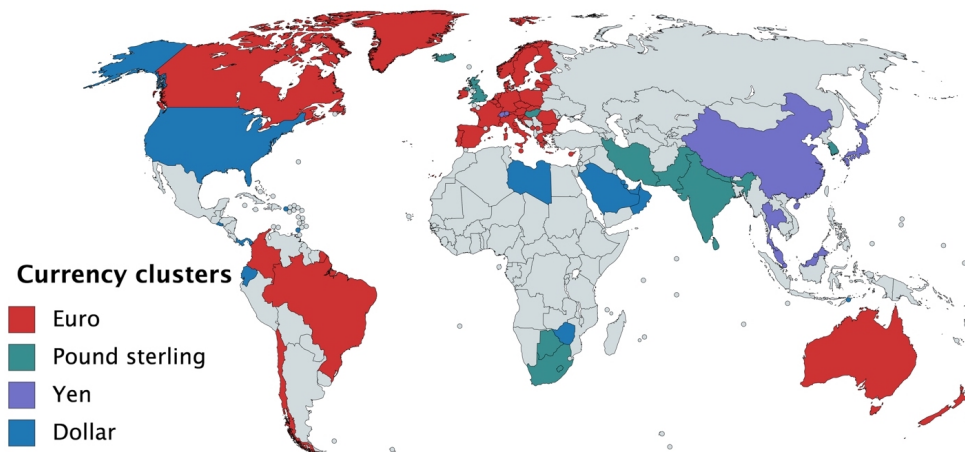
SPONGE

SPONGE<sub>sym</sub>

BNC

$\bar{L}_{sym}$

**Figure 21:** Adjacency matrix of the Forex data set, sorted by cluster membership, for  $k = 4$  clusters. We remark that BNC and especially  $\bar{L}_{sym}$  return less meaningful results.



**Figure 22:** SPONGE<sub>sym</sub> clustering of the 50 currencies with  $k = 4$  clusters.