## The Signed Stochastic Block Model

**1.** We analyse the spectrum in the Signed Stochastic Block Model (SSBM). We consider the following extension of the traditional stochastic block model, to the setting when both positive and negative edges are present. We focus on case when there are $k = 2$ clusters, $C_1, C_2$, of equal size $\frac{n}{2}$. Without loss of generality, assume the first half of the nodes are part of $C_1$, and the second half of the nodes are part of $C_2$. Let $A \in \{0, \pm 1\}^{n \times n}$ denote the adjacency matrix of the signed graph $G$, with $A_{ii} = 0$, and $A_{ij} = A_{ji}$. Under the SSBM, the entries of $A$ are given by the following mixture model. Each edge is present independently in the graph with probability $p$, and its weight is as follows:

-for a pair of nodes within the same cluster

- $A_{ij} = 1$ with probability $\mathbf{1} - \eta$, otherwise
- $A_{ij} = -1$ with probability $\eta$

-for a pair of nodes in different clusters

- $A_{ij} = 1$ with probability $\eta$
- $A_{ij} = -1$ with probability $1 - \eta$, otherwise

You may assume $0 \leqslant \eta < 1/2$.

(a) Compute the expected adjacency matrix $\mathbb{E}[A]$ and its eigenvalues.

(b) Definition: A signed graph is a graph whose edge weights may take either positive of negative values. If $H$ is the adjacency matrix of a signed graph, the corresponding graph Signed Laplacian is defined as $\overline{L} = \overline{D} - H$, where $\overline{D}$ is a diagonal matrix with $\overline{D}_{ii} = \sum_{j=1}^{n} |H_{ij}|$.
Question: Compute the graph Signed Laplacian corresponding to $\mathbb{E}[A]$, and its eigenvalues.

**Hint:** For each $i < j$, note that the above mixture model is given by

$i, j$ belong to the **same** cluster / $i, j$ belong to **different** clusters

$$A_{ij} = \begin{cases} 1 & ; \text{ w. p } & p(1-\eta) \\ -1 & ; \text{ w. p } & p\eta \\ 0 & ; \text{ w. p } & (1-p) \end{cases} \quad (1) \qquad A_{ij} = \begin{cases} 1 & ; \text{ w. p } & p\eta \\ -1 & ; \text{ w. p } & p(1-\eta) \\ 0 & ; \text{ w. p } & (1-p) \end{cases} . \quad (2)$$

In particular, $(A_{ij})_{i \leqslant j}$ are independent random variables. It is also helpful to consider the following vector of size $n$

$$w = \frac{1}{\sqrt{n}} (\underbrace{1\ 1 \ldots 1}_{n/2} \quad \underbrace{-1\ \ -1 \ldots -1}_{n/2})^T \in \mathbb{R}^n \tag{3}$$

which corresponds to the "ground truth", i.e., the two "planted clusters".

Solution:

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

$$\mathbb{E}A_{ij} = \begin{cases} p(1-2\eta) & ; \text{ if } & i, j \text{ lie in same cluster} \\ -p(1-2\eta) & ; \text{ if } & i, j \text{ lie in different clusters} \\ 0 & ; \text{ if } & i = j \end{cases} . \tag{4}$$

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

$$\mathbb{E}[A] = \underbrace{\left[\begin{array}{c|c} p(1-2\eta)J & -p(1-2\eta)J \\ \hline -p(1-2\eta)J & p(1-2\eta)J \end{array}\right]}_{M} - p(1-2\eta)I = M - p(1-2\eta)I, \quad (5)$$

where $I$ denotes the identity matrix, and $J$ the all ones matrix. $M$ is clearly a rank 1 matrix; indeed,

$$M = np(1-2\eta)ww^T \quad (6)$$

where $w$ is defined in (3). Therefore, we obtain eigenvalues

$$\lambda_i(\mathbb{E}[A]) = \left\{ \begin{array}{rcl} p(n-1)(1-2\eta) & ; & i = 1 \\ -p(1-2\eta) & ; & i = 2, \ldots, n \end{array} \right. , \quad (7)$$

The eigenvector $v_1$ is given by

$$v_1(\mathbb{E}[A]) = v_1(M) = w. \quad (8)$$

Moreover, one can check that the diagonal matrix with the degrees of $\mathbb{E}[A]$ is given by

$$\mathbb{E}[\overline{D}] = (n-1)pI. \quad (9)$$

Therefore, the expected associated graph Signed Laplacian is given by

$$\mathbb{E}[\overline{L}] = \mathbb{E}[\overline{D}] - \mathbb{E}[A] = (n-1)pI - \mathbb{E}[A], \quad (10)$$

and hence its eigenvalues are

$$\lambda_i(\mathbb{E}[\overline{L}]) = \left\{ \begin{array}{rcl} 2\eta(n-1)p & ; & i = n \\ (n-1)p + p(1-2\eta) = (n-2\eta)p & ; & i = 1, \ldots, n-1 \end{array} \right. , \quad (11)$$

The eigenvector $v_n$ is such that

$$v_n(\mathbb{E}[\overline{L}]) = v_n(\mathbb{E}[A]) = w. \quad (12)$$