Inference in large random graphs

Laurent Massoulié and Ludovic Stéphan

January 30, 2023

Contents

Introduction

Many inference problems amount to finding structure hidden in some large data set. Examples include:

- community detection, i.e. clustering of graph nodes into groups of nodes with statistically similar properties (applications: recommender systems for online social networks; functional groups of proteins in cell chemistry)
- graph alignment, i.e. finding a mapping of one graph's nodes to another one's that is at least approximately a graph isomorphism (applications: automatic translation; de-anonymization of databases)
- matrix completion, i.e. filling missing entries of large matrix so that the result has low rank (application: recommender systems)
- Hamiltonian cycle detection, i.e. finding a Hamiltonian cycle in a graph so that most edges connect nodes that are nearby in the cycle (application: fast sequencing of DNA).

To understand the hardness of the task, and assess the performance of candidate algorithms, one approach consists in considering random instances where the desired structure has been **planted** in some random background.

In the past few years, this approach has been attempted on all the above problems, revealing fascinating phase transition phenomena: For some problem-dependent notion of *signal-to-noise* ratio (SNR), an **information-theoretic** threshold SNR_{IT} exists such that the task is impossible below it, not enough signal being present in the observation, and feasible above it. There may further exist a **computational** threshold SNR_{Comp} such that no algorithm is known to succeed in polynomial time below it, while fast algorithms are known to succeed above it.

Community detection on data generated from the Stochastic Block Model provides a good illustration of these phenomena, and of the methods from probability, information theory and statistical physics allowing their characterization.

In these notes we will primarily consider the problems of community detection and tree reconstruction for large graphs drawn from probabilistic models. For community detection, we will consider the so-called Stochastic Block Model, which generalizes the Erdős-Rényi random graph. For tree reconstruction, we will consider Galton-Watson branching trees.

The organization of these notes is as follows.

In a first part we consider community detection in a "strong signal" regime, where it is possible to not only cluster all but a vanishing fraction of nodes into their correct communities, but it is also possible to estimate all parameters of the generative Stochastic Block Model from which the graph was drawn. The analysis for this strong signal regime relies on two complementary tools. Linear algebra provides bounds on perturbation of the eigenvalues and eigenvectors of Hermitian matrices. We complement it with probabilistic bounds on the spectral norm of random matrices.

In a second part we consider tree reconstruction. To put this problem in context, we introduce the framework of Gibbs Markov random fields. We review general properties, and introduce the Bethe free energy, the belief propagation (or product-sum) algorithm, its properties on tree Markov fields and its link to Bethe free energy minimization.

We then establish phase transitions on feasibility of the tree reconstruction problem. First we show that the so-called Kesten-Stigum threshold determines the phase transition for a particularly simple reconstruction named census reconstruction. Next we characterize the phase transition for optimal reconstruction from properties of a fixed point equation.

The third part covers results on community detection in the stochastic block model in a "weak signal" regime. We show the links between this problem and that of tree reconstruction, establish feasibility of community detection above the Kesten-Stigum threshold, impossibility of community detection below the tree reconstruction threshold, and the existence of "hard phase" below the Kesten-Stigum threshold where community detection is feasible, but known algorithms require exponential time to perform this detection.

Additional topics to be be covered:

-hypothesis testing on presence or not of planted structure in given graph;

-semi-definite algorithms and their application to community detection / hypothesis testing;

-phase transitions on spectra of low-rank deformations of Wigner matrices?

Part I

The strong signal case: matrix perturbation bounds

Chapter 1

Spectral bounds and perturbation inequalities

Spectral methods will be central in our treatment of graph clustering / community detection. In this chapter we thus introduce fundamental inequalities and perturbation bounds for eigenvalues and eigenvectors of matrices that will be needed in the sequel. For Hermitian matrices we shall in particular introduce inequalities of Weyl on perturbations of eigenvalues, and the famous "sin Θ" theorem of Davis and Kahan on perturbations of eigenvectors. For general (non-Hermitian) matrices we shall introduce the Bauer-Fike theorem on eigenvalue perturbations. Finally we shall discuss fundamental inequalities on spectra of graphs, namely the Cheeger inequality and the Alon-Boppana inequality, and their relationship to the concept of expander graphs.

1.1 Singular value decomposition and principal components analysis

Consider a rectangular $n \times p$ matrix $X = (X_{ij})_{i \in [n], j \in [p]} \in \mathbb{C}^{n \times p}$ with complex-valued entries X_{ij}. Its singular value decomposition is given by the factorization

$$X = U \Lambda V^*, \tag{1.1}$$

where $U \in \mathbb{C}^{n \times n}$ and $V \in \mathbb{C}^{p \times p}$ are unitary matrices, i.e. $U^*U = I_n$, $V^*V = I_p$, where $M^* = \overline{M}^\top$ is the transpose of the complex conjugate of a matrix M and I_n is the identity matrix in dimension n (in case X is real, U and V can be taken real and orthogonal), $\Lambda \in \mathbb{R}^{n \times p}$, with non-zero elements only on its diagonal. In addition, the diagonal elements σ_i of Λ are non-negative, non-increasing: $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{n \wedge p}$. They are known as the singular values of X and the number of non-zero singular values $\sigma_i > 0$ is the rank of matrix X. The columns of U (respectively, of V) are the left (respectively, right) singular vectors of X.

Denoting by u_i and v_i the i-th left and right singular vectors, the above factorization also reads

$$X = \sum_{i=1}^{n \wedge p} \sigma_i u_i v_i^*.$$

One can show (exercise!) existence of the singular value decomposition by considering the Hermitian $(n + p) \times (n + p)$-matrix

$$M = \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix},$$

and applying to it the spectral theorem, which says that Hermitian matrices have a real spectrum and an orthonormal basis of eigenvectors.

Recall that the **operator norm** of a matrix $X \in \mathbb{C}^{n\times p}$ is defined as

$$\|X\|_{op} = \sup_{u\in\mathbb{C}^p} \frac{\|Xu\|}{\|u\|}. \tag{1.2}$$

It is also equal to the largest singular value, σ_1, of X. Operator norm is often called spectral norm.

Recall also that the **Frobenius norm** of X is defined as

$$\|X\|_F = \sqrt{\sum_{i,j} |X_{ij}|^2}. \tag{1.3}$$

It also coincides with the ℓ_2-norm of the vector of its singular values, i.e.

$$\|X\|_F = \sqrt{\sum_i \sigma_i^2}.$$

Principal components analysis is a dimensionality reduction technique. It amounts to approximating matrix X, for any given $r \leq r_0 \leq n \wedge p$, where r_0 is the rank of X, by the rank r-matrix X_r defined as

$$X_r := \sum_{i=1}^{r} \sigma_i u_i v_i^*.$$

It enjoys the following properties.

Proposition 1.1. *Among all rank-r approximations of X, X_r minimizes the approximation error both in Frobenius norm and operator norm. Moreover one has*

$$\|X - X_r\|_{op} = \sigma_{r+1},\ \|X - X_r\|_F = \sqrt{\sum_{i=r+1}^{n\wedge p} \sigma_i^2}.$$

1.2 Perturbations of eigenvalues and eigenvectors

We will denote by convention, for a $n \times n$ matrix M with real spectrum, $\lambda_i(M)$ its i-th eigenvalue, sorted in decreasing order, i.e. $\lambda_1(M) \geq \cdots \geq \lambda_n(M)$.

The following inequalities are due to Weyl:

Theorem 1.1. *For two Hermitian $n \times n$ matrices H, W, and any $i \in [n]$, one has:*

$$|\lambda_i(H) - \lambda_i(H + W)| \leq \sigma_1(W). \tag{1.4}$$

Proof. The Courant-Fisher min-max characterization theorem states that for a Hermitian matrix H, one has

$$\lambda_i(H) = \sup_{V:i-\text{dimensional subspace}} \inf_{u\in V} \frac{\|Hu\|}{\|u\|}. \tag{1.5}$$

Use this to bound $\lambda_i(H + W)$, taking for V the subspace spanned by the i first eigenvectors of H, say $u_1, \ldots, u_i$, assumed orthonormal. Specifically, let $\theta = (\theta_1, \ldots, \theta_i)$, with $\|\theta\|^2 = \sum_{j=1}^i |\theta_j|^2$, and $u(\theta) = \sum_{j=1}^i \theta_j u_j$. This gives

$$\begin{aligned}
\lambda_i(H + W) &\leq \inf_{\theta:\|\theta\|=1} u(\theta)^*(H + W)u(\theta) \\
&\leq \sup_{u:\|u\|=1} u^*Wu + \inf_{\theta:\|\theta\|=1} u(\theta)^*Hu(\theta) \\
&\leq \sigma_1(W) + \inf_{\theta:\|\theta\|=1} \sum_{j=1}^i |\theta_j|^2 \lambda_j(H),
\end{aligned}$$

which gives $\lambda_i(H + W) \leq \lambda_i(H) + \sigma_1(W)$. The same reasoning gives $\lambda_i(H) \leq \lambda_i(H + W) + \sigma_1(-W) = \lambda_i(H + W) + \sigma_1(W)$, hence the result. □

The Courant-Fisher min-max characterization can also be used to prove the following

Theorem 1.2. *(Cauchy interlacing theorem) Let A be a $n \times n$ Hermitian matrix, and $P \in \mathbb{C}^{n\times m}$, with $m < n$ be such that $P^*P = I_m$, the identity matrix. Then the $m \times m$ matrix $B := P^*AP$ is such that*

$$\lambda_i(A) \geq \lambda_i(B) \geq \lambda_{n-m+i}(A).$$

Proof. Let $v_1, \ldots, v_i$ be an orthonormal collection of eigenvectors of B associated respectively with $\lambda_1(B), \ldots, \lambda_i(B)$. Let E_i denote the space spanned by $v_1, \ldots, v_i$. Thus denoting F_i the image by P of E_i, one has

$$\begin{aligned}\lambda_i(B) &= \inf_{\substack{u \in E_i \\ \|u\|=1}} u^*P^*APu \\ &= \inf_{\substack{v \in F_i \\ \|v\|=1}} v^*Av \\ &\leq \lambda_i(A),\end{aligned}$$

where the last inequality follows from the Courant-Fisher theorem and the fact that F_i has dimension i, which itself follows from the fact that P has full column rank.

For the converse inequality, the same reasoning can be applied to $-A$ and $-B$ to yield

$$\lambda_{m-i+1}(-B) = -\lambda_i(B) \leq \lambda_{m-i+1}(-A) = -\lambda_{n-m+i}(A).$$

□

We now state a theorem on the perturbation of eigenvectors of Hermitian matrices. It appears in Yu et al. [46], and is a variant of the celebrated Davis-Kahan "$\sin\Theta$" theorem. We refer to [46] for its proof.

Theorem 1.3. *Let $H, \hat{H} \in \mathbb{R}^{p\times p}$ be symmetric real matrices with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_p$ and $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_p$ respectively. Fix $1 \leq r \leq s \leq p$ and assume that $\min(\lambda_{r-1} - \lambda_r, \lambda_s - \lambda_{s+1}) > 0$, where $\lambda_0 = +\infty$ and $\lambda_{p+1} = -\infty$. Let $d = s - r + 1$ and let $V = (v_r, \ldots, v_s) \in \mathbb{R}^{p\times d}$, $\hat{V} = (\hat{v}_r, \ldots, \hat{v}_s) \in \mathbb{R}^{p\times d}$ have orthonormal columns of eigenvectors, $Hv_j = \lambda_j v_j$, $\hat{H}\hat{v}_j = \hat{\lambda}_j \hat{v}_j$.*

Then there exists an orthogonal matrix $O \in \mathbb{R}^{d\times d}$ such that

$$\|\hat{V}O - V\|_F \leq \frac{2^{3/2}\min(d^{1/2}\|H - \hat{H}\|_{op}, \|H - \hat{H}\|_F)}{\min(\lambda_{r-1} - \lambda_r, \lambda_s - \lambda_{s+1})}. \tag{1.6}$$

We give a proof of a version of this theorem when one only aims at controlling the perturbation of a single eigenvector:

Proposition 1.2. *Let H, $\hat{H}$ be two Hermitian $p \times p$ matrices, $r \in [p]$ and let*

$$\delta := \inf_{j:\lambda_j \neq \lambda_r} |\lambda_j - \lambda_r|.$$

Let $\hat{v}_r$ be a normed eigenvector of $\hat{H}$ associated with $\hat{\lambda}_r$. Then, provided that $\|\hat{H} - H\|_{op} < \delta$, there exists a normed eigenvector w of H associated with λ_r such that

$$\langle \hat{v}_r, w\rangle \geq \sqrt{1 - \frac{\|\hat{H} - H\|_{op}^2}{[\delta - \|\hat{H} - H\|_{op}]^2}}. \tag{1.7}$$

Proof. Decompose $\hat{v}_r$ on an orthonormal basis $v_1, \ldots, v_p$ of eigenvectors of v as

$$\hat{v}_r = \sum_{i=1}^{p} \theta_i v_i.$$

Apply $\hat{H}$ to both sides to obtain

$$\hat{\lambda}_r \sum_{i=1}^{p} \theta_i v_i = \sum_{i=1}^{p} \theta_i \lambda_i v_i + (\hat{H} - H)\hat{v}_r,$$

which in turn gives

$$\sum_{i=1}^{p} \theta_i [\hat{\lambda}_r - \lambda_i] v_i = (\hat{H} - H)\hat{v}_r.$$

Taking norms, this gives

$$\sum_{i=1}^{p} |\theta_i|^2 [\hat{\lambda}_r - \lambda_i]^2 = \|(\hat{H} - H)\hat{v}_r\|^2 \leq \|\hat{H} - H\|_{op}^2.$$

For $\lambda_i \neq \lambda_r$, one has

$$\begin{aligned} \delta &\leq |\lambda_i - \lambda_r| \\ &\leq |\hat{\lambda}_r - \lambda_i| + |\hat{\lambda}_r - \lambda_r| \\ &\leq |\hat{\lambda}_r - \lambda_i| + \|\hat{H} - H\|_{op}, \end{aligned}$$

where we used Weyl's inequality in the last step. Combined with the previous inequality this gives, assuming $\delta > \|\hat{H} - H\|_{op}$:

$$\sum_{i:\lambda_i \neq \lambda_r} |\theta_i|^2 [\delta - \|\hat{H} - H\|_{op}]^2 \leq \|\hat{H} - H\|_{op}^2.$$

Let then

$$w := \frac{1}{\sqrt{\sum_{i:\lambda_i = \lambda_r} |\theta_i|^2}} \sum_{i:\lambda_i = \lambda_r} \theta_i v_i.$$

It is easily verified that w is a normed eigenvector of H associated with λ_r.

Using the fact that $\sum_{i=1}^{p} |\theta_i|^2 = 1$, and $\langle \hat{v}_r, w \rangle = \sqrt{\sum_{i:\lambda_i = \lambda_r} |\theta_i|^2}$, this gives

$$\delta > \|\hat{H} - H\|_{op} \Rightarrow 1 - \langle \hat{v}_r, w \rangle^2 \leq \frac{\|\hat{H} - H\|_{op}^2}{[\delta - \|\hat{H} - H\|_{op}]^2},$$

the announced inequality (1.7). □

In contrast, an application of the theorem to the special case $d = 1$ would give:

$$2[1 - |\langle \hat{v}_r, v_r \rangle|] \leq \frac{2^3 \|\hat{H} - H\|_{op}^2}{\delta^2}. \tag{1.8}$$

In that particular case, both (1.7) and (1.8) imply that, when $\delta \gg \|\hat{H} - H\|_{op}$, eigenvector v_r of H associated with λ_r appears nearly unchanged, up to sign, as eigenvector $\hat{v}_r$ of $\hat{H}$ associated with $\hat{\lambda}_r$. We now give a version of the celebrated Bauer-Fike theorem, which can be used instead of Weyl's inequality when dealing with non-Hermitian matrices.

Theorem 1.4. *Let matrix $M \in \mathbb{C}^{n \times n} = T \Lambda T^{-1}$ where T is an invertible matrix and Λ is a diagonal matrix with diagonal entries $\{\lambda_i\}_{i \in [n]}$, so that the λ_i are the eigenvalues of M. Let $W \in \mathbb{C}^{n \times n}$ be (non-necessarily Hermitian) perturbation matrix. Then any eigenvalue μ of matrix $N := M + W$ verifies*

$$\inf_{i \in [n]} |\mu - \lambda_i| \leq \|W\|_{op} \|T\|_{op} \left\|T^{-1}\right\|_{op}. \tag{1.9}$$

Proof. Let μ be an eigenvalue of N and u an associated normed eigenvector. The result trivially holds if μ is an eigenvalue of M. Assume then that $\mu \neq \lambda_i$, $i \in [n]$. Write

$$\mu u = (M + W)u,$$

so that

$$u = (\mu I_n - M)^{-1} W u = T(\mu I_n - \Lambda)^{-1} T^{-1} W u.$$

Taking norms, this yields

$$\begin{aligned} \|u\| = 1 \quad &\leq \left\| T(\mu I_n - \Lambda)^{-1} T^{-1} W \right\|_{op} \\ &\leq \|T\|_{op} \sup_{i \in [n]} |\mu - \lambda_i|^{-1} \left\| T^{-1} \right\|_{op} \|W\|_{op}, \end{aligned}$$

which yields the announced result. □

Remark. *We will see in subsequent chapter associated results for perturbations of eigenvectors that can be used instead of Davis-Kahan's* $\sin\Theta$ *theorem for non-Hermitian matrices.*

1.3 Graph spectra, expansion and Cheeger inequality

We now look more specifically at spectral properties of matrices associated with graphs. Let then $G = (V, E)$ be an undirected graph, with vertex set $V = [n]$ and edge set E, E consisting of non-oriented pairs $\{i, j\}$ of distinct vertices. The two most fundamental matrices associated with such a graph are: its **adjacency matrix** $A = A(G)$ defined by

$$A_{ij} = \mathbf{1}_{\{i,j\} \in E},$$

and its **Laplacian matrix** $L = L(G)$, defined by

$$L_{ij} = \begin{cases} d_i := \sum_{k \neq i} A_{ik} & \text{if } j = i, \\ -A_{ij} & \text{if } j \neq i. \end{cases}$$

In the above we introduced **the degree** $d_i = d_i(G)$ **of node** i, defined as the number of neighbors of node i in the graph.

A graph G whose nodes all have the same degree d, i.e. $d_i(G) \equiv d$, is called a d**-regular graph**. For d-regular graphs on n vertices, it holds that $L(G) = dI_n - A(G)$.

We often write L for $L(G)$ and A for $A(G)$ to shorten notations. It is readily verified that for all $u \in \mathbb{R}^n$,

$$u^\top L u = \sum_{i<j} A_{ij}(u_i - u_j)^2.$$

This establishes that the Laplacian matrix L is **positive semi-definite**, a property we denote by

$$L \succeq 0.$$

In fact the above identity gives us more: it implies that $\lambda_n(L) = 0$, and that the constant vector $e = \{1\}$ is an associated eigenvector.

By definition, the **spectral gap** of graph G is given by its second smallest Laplacian eigenvalue, $\lambda_{n-1}(L(G))$.

We shall also need the following two quantities associated with graph G:

$$\Delta(G) := \sup_{i \in [n]} d_i(G)$$

is the largest degree of nodes in G. For any vertex set $S \subset V$, the corresponding partition of V into S and its complementary set $\overline{S} = V \setminus S$ has an associated **isoperimetric ratio** $\frac{|E(S,\overline{S})|}{\min(|S|,|\overline{S}|)}$, where $E(A, B)$ denotes the number of edges in E with one endpoint in set A and the other endpoint in set B.

We then define the **isoperimetric constant** of G as

$$I(G) := \min\left\{\frac{|E(S,\overline{S})|}{|S|}, S \subset V, 0 < |S| \leq \frac{n}{2}\right\}. \tag{1.10}$$

It is readily seen to be the smallest isoperimetric ratio of all partitions of node set V into two parts $S, \overline{S}$.

Definition 1.1. *An undirected graph $G = (V, E)$ is a γ-expander (respectively, a γ-spectral expander) for some constant $\gamma > 0$ if its isoperimetric constant $I(G)$ (respectively, its spectral gap $\lambda_{|V|-1}(L(G))$) is larger than γ.*

A family of undirected graphs $\{G_n = (V_n, E_n)\}_{n\in\mathbb{N}}$ is an expander (respectively, a spectral expander) if there exists some $\gamma > 0$ that is a lower bound of the isoperimetric constant $I(G_n)$ (respectively, of the spectral gap $\lambda_{|V_n|-1}(L(G_n))$) uniformly in $n \in \mathbb{N}$.

It turns out that such expansion and spectral expansion are closely related properties. Indeed we have the following inequality due to Cheeger:

Theorem 1.5. *The isoperimetric constant $I(G)$ of undirected graph G verifies*

$$I(G) \leq \sqrt{2\Delta(G)\lambda_{n-1}(L(G))}. \tag{1.11}$$

Before we turn to its proof, let us establish the easier result:

Proposition 1.3. *The isoperimetric constant $I(G)$ of undirected graph G verifies*

$$I(G) \geq \frac{\lambda_{n-1}(L(G))}{2}. \tag{1.12}$$

Proof. The Courant-Fisher variational characterization gives

$$\lambda_{n-1}(L) = \inf\left\{\frac{x^\top Lx}{\|x\|^2}, x : \langle x, e\rangle = 0\right\},$$

where we used that $\lambda_n(L) = 0$ admits e as an associated eigenvector.

Define then $x \in \mathbb{R}^n$ by $x_i = \mathbf{1}_{i\in S} - \frac{|S|}{n}$, $i \in [n]$, where S is the subset of V achieving the minimum in the definition of $I(G)$. The above characterization yields

$$\lambda_{n-1} \leq \frac{|E(S,\overline{S})|}{|S|(1-|S|/n)}.$$

The inequality (1.12) follows since necessarily $|S| \leq n/2$. □

It directly follows from the Proposition (Exercise!) that graph G is connected if and only if $\lambda_{n-1}(L) > 0$.

Proof. (of Cheeger's inequality). Let x be a non-constant eigenvector of L associated with its eigenvalue λ_{n-1}. Possibly after changing the sign of x, assume that $S := \{i \in [n] : x_i > 0\}$ verifies $|S| \leq n/2$. Let $y := \{x_i^+\}_{i\in[n]}$.

Write

$$\begin{aligned}
\lambda_{n-1}\|y\|^2 &= \textstyle\sum_{i\in[n]} x_i^+ \left((Lx)_i\right)^+ \\
&= \textstyle\sum_i x_i^+ \left[\sum_{j\neq i} A_{ij}(x_i - x_j)\right]^+ \\
&= \textstyle\sum_i x_i^+ \left[\sum_{j\neq i} A_{ij}(x_i^+ - x_j)\right]^+ \\
&\geq \textstyle\sum_i x_i^+ \left[\sum_{j\neq i} A_{ij}(x_i^+ - x_j)\right] \\
&\geq \textstyle\sum_i x_i^+ \left[\sum_{j\neq i} A_{ij}(x_i^+ - x_j^+)\right] \\
&= \textstyle\sum_{(ij)\in E} A_{ij}(y_i - y_j)^2,
\end{aligned}$$

which yields

$$\lambda_{n-1} \geq \frac{\sum_{(ij)\in E} A_{ij}(y_i - y_j)^2}{\|y\|^2}.$$

Note next that

$$\sum_{(ij)\in E} A_{ij}(y_i + y_j)^2 \leq 2 \sum_{(ij)\in E} A_{ij}(y_i^2 + y_j^2) \leq 2\Delta(G)\|y\|^2.$$

Use now Cauchy-Schwarz inequality to obtain

$$\left(\sum_{(ij)\in E} A_{ij}|y_i^2 - y_j^2| \right)^2 \leq \left[\sum_{(ij)\in E} A_{ij}(y_i - y_j)^2 \right] \left[\sum_{(ij)\in E} A_{ij}(y_i + y_j)^2 \right].$$

The three previous inequalities yield the bound

$$\lambda_{n-1} \geq \frac{\sum_{(ij)\in E} A_{ij}(y_i - y_j)^2 \sum_{(ij)\in E} A_{ij}(y_i + y_j)^2}{\|y\|^2 \sum_{(ij)\in E} A_{ij}(y_i + y_j)^2} \geq \frac{R^2}{2\Delta(G)\|y\|^4}, \tag{1.13}$$

where we introduced the term

$$R := \sum_{(ij)\in E} A_{ij}|y_i^2 - y_j^2|.$$

Finally, R can be lower-bounded in terms of the isoperimetric constant $I(G)$ as follows. Let $t_0 = 0 < t_1 < \ldots < t_m$ denote the distinct values taken by the y_i. For $k = 0, \ldots, m$, let $V_k := \{i \in V : y_i \geq t_k\}$. Necessarily, $|V_k| \leq |S| \leq n/2$ for all $k \geq 1$. Write then

$$\begin{aligned}
R &= \sum_{(ij)\in E} A_{ij}|y_i^2 - y_j^2| \\
&= \sum_{k=1}^{m} \sum_{(ij)\in E: y_j < y_i = t_k} A_{ij}(y_i^2 - y_j^2) \\
&= \sum_{k=1}^{m} \sum_{y_i = t_k, y_j = t_\ell, \ell < k} A_{ij}[(t_k^2 - t_{k-1}^2) + \cdots + (t_{\ell+1}^2 - t_\ell^2)] \\
&= \sum_{k=1}^{m} \sum_{i\in V_k} \sum_{j\in \overline{V}_k} A_{ij}(t_k^2 - t_{k-1}^2) \\
&= \sum_{k=1}^{m} |E(V_k, \overline{V}_k)|(t_k^2 - t_{k-1}^2) \\
&\geq I(G) \sum_{k=1}^{m} |V_k|(t_k^2 - t_{k-1}^2) \\
&= I(G) \sum_{k=1}^{m} t_k^2(|V_k| - |V_{k+1}|) \\
&= I(G)\|y\|^2.
\end{aligned}$$

Combined with (1.13), this inequality yields the annoucned result (1.11). □

Remark. *The proof given here is taken from [30], which in fact establishes a refined version, with* $(\Delta(G) - \lambda_{n-1})$ *in place of* $\Delta(G)$ *in the upper bound. The original Cheeger inequality is concerned with Laplace operators on manifolds, and is obtained through an integration-by-parts argument of which the above proof contains a discrete analogue.*

Remark. *Together, Cheeger's inequality* (1.11) *and* (1.12) *entail the following. A family of graphs* $\{G_n\}_{n\in\mathbb{N}}$ *with uniformly bounded degrees* $\Delta(G_n) \leq \Delta$, $n \in \mathbb{N}$ *is an expander if and only if it is a spectral expander.*

Remark. *Eigenvector x associated with the second smallest eigenvalue λ_{n-1} of the graph Laplacian $L(G)$ is sometimes referred to as the Fiedler vector of G. The proof of Cheeger's inequality shows in fact that there exists some $k \in [m]$ such that*

$$\frac{|E(V_k, \overline{V}_k)|}{|V_k|} \leq \sqrt{2\Delta(G)\lambda_{n-1}}.$$

In other words, there exists some threshold $t \in \mathbb{R}$ such that the partition of $[n]$ based according to the Fiedler vector x into $S = \{i : x_i \geq t\}$ and $\overline{S} = \{i; x_i < t\}$ necessarily satisfies

$$|S| \leq \frac{n}{2} \text{ and } \frac{|E(S, \overline{S})|}{|S|} \leq \sqrt{2\Delta(G)\lambda_{n-1}(L(G))}.$$

Combined with inequality (1.12)*, this implies*

$$\frac{|E(S, \overline{S})|}{|S|} \leq 2\sqrt{\Delta(G)I(G)}.$$

Thus for some sequence of graphs $G_n = (V_n, E_n)$ such that $I(G_n) \to 0$ as $n \to \infty$ and with uniformly bounded degrees $\Delta(G_n) \leq \Delta$, it follows that there exist graph partitions $(S_n, \overline{S}_n)$ for G_n, based on the Fiedler vectors of each G_n, with an asymptotically vanishing isoperimetric ratio, i.e.

$$\lim_{n\to\infty} \frac{|E_n(S_n, \overline{S}_n)|}{\min(|S_n|, |\overline{S}_n|)} = 0,$$

where $E_n(A, B)$ is the number of edges of G_n with one endpoint in A and the other in B. This provides a first justification that spectral partitioning of a graph, i.e. partitioning of a graph based on eigenvectors of associated matrices, produces interesting partitions, namely here a partition with vanishing isoperimetric ratio. We will see in the sequel other spectral partitioning methods, as well as desirable properties they enjoy for graphs sampled from probability distributions of interest.

1.4 Ramanujan graphs are the best expanders: Alon-Boppana inequality

We shall need the following notions. The **graph distance** $d_G(i, j)$ between two vertices $i, j \in V$ is defined as the length in edges of the shortest path connecting i to j in G. The **graph diameter** $D(G)$ is defined as the largest graph distance $d_G(i, j)$ between distinct vertices i, j. For given vertex $i \in V$ and integer $r \geq 0$, $\mathcal{B}_G(i, r) := \{j \in V, d_G(, j) \leq r\}$ is then the ball for d_G centered at i and of radius r.

We now focus on undirected d-regular graphs $G = (V, E)$ with node set $[n]$. In that case, the spectral gap $\lambda_{n-1}(L(G))$ also reads

$$\lambda_{n-1}(L(G)) = d - \lambda_2(A(G)).$$

It turns out that this spectral gap cannot be arbitrarily large. Indeed we have the following result of Alon and Boppana:

Theorem 1.6. *For a d-regular graph with diameter at least $2r + 1$, then necessarily*

$$\lambda_2(A(G)) \geq 2\sqrt{d-1}\cos\left(\frac{\pi}{r+2}\right). \tag{1.14}$$

A simple counting argument (Exercise!) guarantees that a graph $G = (V, E)$ with node degrees upper-bounded by Δ can have within distance r of any given vertex at most $\Delta[(\Delta-1)^r - 1]/(\Delta-2)$ vertices (we assume implicitly that $\Delta \geq 3$: simpler bounds hold for $\Delta < 3$). Thus if it comprises n vertices, its diameter $D(G)$ must satisfy

$$D(G) \geq \frac{\ln(1 + n(\Delta-2)/\Delta)}{\ln(\Delta)} \geq \log_\Delta(n/3) = \Theta(\log(n)).$$

A direct consequence of Theorem 1.6, obtained by taking a Taylor expansion of the term $\cos\left(\frac{\pi}{r+2}\right)$ in (1.14), is the following

Corollary 1.1. *For fixed d, a d-regular graph G on n nodes verifies*

$$\lambda_2(A(G)) \geq 2\sqrt{d-1}\left(1 - O(\log(n)^{-2})\right), \tag{1.15}$$

so that in the limit $n \to \infty$, its spectral gap $\lambda_{n-1}(L(G))$ is no larger than $d - 2\sqrt{d-1} - o_n(1)$.

It is interesting at this stage to compare this with the notion of **Ramanujan graph** that we now define:

Definition 1.2. *An undirected d-regular graph G is a Ramanujan graph if it verifies*

$$\max(\lambda_2(A(G)), |\lambda_n(A(G))|) \leq 2\sqrt{d-1}. \tag{1.16}$$

In light of the Alon-Boppana result, we see that there do not exist any d-regular graphs with a spectral gap larger than that of Ramanujan graphs by a non-vanishing margin $\Omega_n(1)$. In other words, Ramanujan graphs are the best possible expanders.

Proof. (of the Alon-Boppana theorem 1.6). Let i, j be two vertices in V at distance $d_G(i,j)$ at least $2r+1$, so that the balls $\mathcal{B}(i,r)$, $\mathcal{B}(j,r)$ are disjoint. Let $\mathcal{M} := \mathcal{B}(i,r) \cup \mathcal{B}(j,r)$ be the union of these two balls, and $m := |\mathcal{M}|$ its cardinal. Let G' denote the subgraph of G induced by the vertices in $\mathcal{M}$. Its adjacency matrix $A(G') \in \mathbb{R}^{m \times m}$ also reads

$$A(G') = P^* A(G) P,$$

where $P^* \in \mathbb{R}^{n \times m}$ denotes the matrix corresponding to projection onto coordinates in $\mathcal{M}$. It is readily seen that $P^*P = I_m$. Thus, by Cauchy's interlacing theorem, one has

$$\lambda_2(A(G)) \geq \lambda_2(A(G')).$$

Note that by construction, G' consists of two disjoint connected components $\mathcal{B}(i,r)$ and $\mathcal{B}(j,r)$. By Perron-Frobenius theorem, the adjacency matrix $A(\mathcal{B}(i,r))$ (respectively, $A(\mathcal{B}(j,r))$) of the subgraph of G induced by $\mathcal{B}(i,r)$ (respectively, $\mathcal{B}(j,r)$) admits a non-negative eigenvalue $\lambda_1(A(\mathcal{B}(i,r)))$ (respectively, $\lambda_1(A(\mathcal{B}(j,r)))$) with largest modulus among its eigenvalues. Both are also eigenvalues of $A(G')$, so that

$$\lambda_2(A(G)) \geq \min\left(\lambda_1(A(\mathcal{B}(i,r))), \lambda_1(A(\mathcal{B}(j,r)))\right). \tag{1.17}$$

For $q \geq 0$, denote by $\mathcal{W}_{2q}(i, \mathcal{B}(i,r))$ the collection of walks in $\mathcal{B}(i,r)$ of length $2q$ which start and end at node i, and by $w_{2q}(i, \mathcal{B}(i,r))$ its cardinal. Again by Perron-Frobenius theorem, it holds that

$$\lambda_1(A(\mathcal{B}(i,r))) = \lim_{q \to \infty} \left(w_{2q}(i, \mathcal{B}(i,r))\right)^{1/(2q)}.$$

To lower-bound this limit, we now introduce the notion of **graph cover**.

Definition 1.3. *Given an undirected graph $G = (V,E)$, graph $\mathcal{C} = (W,F)$ is a cover of G if there exists a cover map $\phi : W \to V$ such that:*

1) $\phi(W) = V$, and

2) for any $u \in W$, ϕ defines a bijection between the sets $\{v \in W : (uv) \in F\}$ and $\{j \in V : (ij) \in E\}$.

Let $\mathcal{T}_d$ denote the infinite d-regular tree. Specifically, its vertex set consists of sequences $(i_1, \ldots, i_\ell)$ in $[d] \times [d-1]^{\ell-1}$, $\ell \geq 1$, and of the empty sequence $()$. Its edges consist of all pairs of vertices of the form $\{(i_1, \ldots, i_{\ell-1}), (i_1, \ldots, i_\ell)\}$.

A fundamental property is then that any d-regular connected graph G admits $\mathcal{T}_d$ as a cover. This is easily shown by constructing the cover map ϕ iteratively, starting with some arbitrary choice $i_0 \in V$ for $\phi()$, then choosing for $\{\phi(k)\}$, $k \in [d]$ an arbitrary map with image set $\{j \in V, (i_0, j) \in E\}$. The construction can be carried on indefinitely, and must necessarily cover all of V if G is connected.

The infinite d-regular tree $\mathcal{T}_d$ is known as the **universal cover** of d-regular graphs.

We now use this construct to lower bound $w_{2q}(\mathcal{B}(i,r))$. We may assume that the cover map ϕ is such that $\phi() = i$.

We further denote by $\mathcal{T}_{d,r}$ the subgraph of $\mathcal{T}_d$ made of nodes at distance at most r from the root (). A key property is then that

$$w_{2q}(i, \mathcal{B}(i,r)) \geq w_{2q}((), \mathcal{T}_{d,r}). \tag{1.18}$$

To establish this, let us first verify that (i) closed walks $u_1, \ldots, u_{2q}$ in $\mathcal{W}_{2q}((), \mathcal{T}_{d,r})$ are mapped by ϕ to closed walks $(\phi(u_1), \ldots, \phi(u_{2q}))$ in $\mathcal{W}_{2q}(i, \mathcal{B}(i,r))$. This is readily established, arguing by induction on r that ϕ maps the vertices in $\mathcal{T}_{d,r}$ to those in $\mathcal{B}(i,r)$, then using the property that ϕ necessarily maps edges of $\mathcal{T}_{d,r}$ to edges of $\mathcal{B}(i,r)$, and finally since by construction $\phi() = i$.

Next we argue that (ii) two distinct walks in $\mathcal{W}_{2q}((), \mathcal{T}_{d,r})$ must be mapped by ϕ to distinct walks in $\mathcal{W}_{2q}(i, \mathcal{B}(i,r))$. Consider two walks $u_1, \ldots, u_{2q}$ and $v_1, \ldots, v_{2q}$ in $\mathcal{W}_{2q}((), \mathcal{T}_{d,r})$, and assume that the first index at which they differ is $\ell \leq 2q$. Then by the local bijection property of the cover map ϕ, it must be that $\phi(u_\ell) \neq \phi(v_\ell)$. This establishes (1.18).

We now evaluate the right-hand side in (1.18). Let $\mathcal{P}_{r+1}$ denote the path graph of length r and with $r+1$ vertices that we take as $\{0, \ldots, r\}$. A counting argument ensures that

$$w_{2q}((), \mathcal{T}_{d,r}) \geq (d-1)^q w_{2q}(0, \mathcal{P}_{r+1}). \tag{1.19}$$

Indeed we can project each walk in $\mathcal{W}_{2q}((), \mathcal{T}_{d,r})$ to a walk in $\mathcal{W}_{2q}(1, \mathcal{P}_{r+1})$ by mapping a vertex at distance s in $\mathcal{T}_{d,r}$ from the root to vertex s in $\mathcal{P}_{r+1}$. We then argue that to each walk $i_1, \ldots, i_{2q}$ in $\mathcal{W}_{2q}(0, \mathcal{P}_{r+1})$ we can associate at least $(d-1)^q$ distinct walks in $\mathcal{W}_{2q}((), \mathcal{T}_{d,r})$ that get projected onto $i_1, \ldots, i_{2q}$, since indeed we have at least $(d-1)$ choices for each move away from the root, and there must be q such moves.

Let us now show that

$$\lim_{q \to \infty} \left(w_{2q}(1, \mathcal{P}_{r+1})\right)^{1/(2q)} = 2\cos(\pi/(r+2)). \tag{1.20}$$

Combined with (1.17)–(1.19) and the fact that i and j are interchangeable, this will conclude our proof of (1.14).

The limit in the left-hand side of (1.20) coincides with the Perron-Frobenius eigenvalue of $A(\mathcal{P}_{r+1})$. The result will thus follow if we can exhibit $(x_0, \ldots, x_r)$ that is an eigenvector of $A(\mathcal{P}_{r+1})$ with non-negative entries, and associated with eigenvalue $2\cos(\pi/(r+2))$. Classical trigonometric identities can be used to verify that $x_i = \sin\left(\frac{(i+1)\pi}{r+2}\right)$ satisfies these conditions. □

Remark. *The above proof of the Alon-Boppana theorem is inspired from Mohar [32], which establishes stronger results. In particular the argument extends to show that many eigenvalues of $A(G)$ must exceed the left-hand side of* (1.14)*, and also extends to give a version of the Alon-Boppana lower bound applicable to certain classes of non-regular graphs.*

1.5 Notes

Good references on linear algebra, PCA and SVD: Horn and Johnson (and Golub and Van Loan). We only saw a subset of Weyl's inequalities, a complete treatment is in some blog of Terence Tao. Specific focus on inequalities for eigenvalues: Bhatia. Non-normal matrices more difficult, give rise to pseudo-spectra: Trefethen et al.

Reference on graph theory: Bollobas. Specifically on algebraic graph theory, Laplace eigenvalues etc: Dan Spielman, Fan Chung?

On the role of Cheeger inequality and its relation to graph clustering: see higher order Cheeger inequality paper [26].

On expander graphs, and construction of Ramanujan graphs: book by Lubotzky.

Recent breakthrough on showing existence of (bipartite) Ramanujan graphs of all sizes: Spielman et al.

On Alon-Boppana inequality and Ramanujan graphs: classical proof given in [36]. Mention Friedman's result, or perhaps at a later chapter.

Chapter 2

Bounding the spectral norm of random matrices

In this chapter we consider random symmetric matrices $W \in \mathbb{R}^{n\times n}$ such that $\{W_{ij}\}_{i\leq j}$ are independent random variables. We introduce bounds on the corresponding spectral, or operator norm $\|W\|_{op}$ that will be used in the next Chapter, in conjunction with the tools on perturbation of eigenstructure of matrices of the previous chapter.

2.1 The trace method

The basic building block of the trace method is the following inequality: for any $k \geq 1$,

$$\rho(W)^{2k} \leq \operatorname{Tr}\left(W^{2k}\right). \tag{2.1}$$

This inequality is asymptotically sharp, in the sense that

$$\rho(W) = \lim_{k\to\infty} \operatorname{Tr}\left(W^{2k}\right)^{\frac{1}{2k}},$$

and hence not much is lost in the approximation. On the other hand, the trace of W^{2k} admits the following combinatorial expansion

$$\operatorname{Tr}(W^{2k}) = \sum_{\substack{\gamma_0,\dots,\gamma_{2k}\in[n]\\ \gamma_0=\gamma_{2k}}} \prod_{j=0}^{2k-1} W_{\gamma_j\gamma_{j+1}}. \tag{2.2}$$

This combinatorial representation together with moment assumptions on the variables W_{ij} allows to establish bounds on the moments of $\rho(W)$ and in turn bounds on tail probabilities of its distribution. Here is a simple illustration of this approach.

Proposition 2.1. *Assume that the W_{ij} are independent centered random variables such that*

$$|W_{ij}| \leq 1 \text{ a.s.} \qquad \textit{and} \qquad \mathbb{E}\big[W_{ij}^2\big] \leq \frac{d}{n}$$

for some (possibly n-dependent) constant $d \geq 1$.

Then for any $\epsilon > 0$, one has

$$\lim_{n\to\infty} \mathbb{P}\Big(\rho(W) \geq \sqrt{d}n^{\epsilon}\Big) = 0. \tag{2.3}$$

Proof. Fix some $k \geq 1$. For a given path $\gamma = (\gamma_0, \ldots, \gamma_{2k})$ appearing in (2.2), we denote by $v(\gamma)$ (resp. $e(\gamma)$) the number of distinct vertices (resp. edges) appearing in γ. Since the W_{ij} are centered, for a term in (2.2) to be nonzero it is necessary that each edge be traversed at least twice. Additionally the paths γ considered are connected, so we can restrict our study to paths satisfying

$$v(\gamma) - 1 \leq e(\gamma) \leq k.$$

We first bound the contribution of a fixed path γ: for each edge (x, y) appearing in γ we denote its *multiplicity* k_{xy}, i.e. the number of times it is visited. Then

$$\mathbb{E} \prod_{j=0}^{2k-1} W_{\gamma_j \gamma_{j+1}} \leq \prod_{(x,y)\in\gamma} \mathbb{E}\big[|W_{xy}|^{k_{xy}}\big] \tag{2.4}$$

$$\leq \prod_{(x,y)\in\gamma} \mathbb{E}\big[W_{xy}^2\big] \tag{2.5}$$

$$\leq \left(\frac{d}{n}\right)^{e(\gamma)}, \tag{2.6}$$

where we used the fact that $|W_{xy}| \leq 1$ almost surely.

Now, let $C(v, e)$ be the set of paths with $v(\gamma) = v$ and $e(\gamma) = e$, and $C'(v, e)$ the set of *canonical* paths, i.e. paths in $C(v, e)$ with vertex set equal to $[v]$. Any path in $C(v, e)$ can be uniquely mapped to a canonical path by labeling its vertices by order of appearance, and hence

$$|C(v, e)| \leq n^v \, |C'(v, e)| \leq n^v \max_{v(\gamma)-1\leq e(\gamma)\leq k} C'(v, e), \tag{2.7}$$

and the above maximum depends only on k. Combining (2.6) and (2.7), we find

$$\begin{aligned}
\mathbb{E}\big[\mathrm{Tr}\left(W^{2k}\right)\big] &\leq \sum_{e=0}^{k} \sum_{v=1}^{e+1} |C(v, e)| \left(\frac{d}{n}\right)^e \\
&\leq \sum_{e=0}^{k} \sum_{v=1}^{e+1} n^v \, |C'(v, e)| \left(\frac{d}{n}\right)^e \\
&\leq C''(k) n d^k,
\end{aligned}$$

where the last constant only depends on k.

We are now in a position to apply a Markov bound: for $\epsilon > 0$, choose k such that $2k\epsilon > 1$, and write

$$\begin{aligned}
\mathbb{P}\Big(\rho(W) \geq \sqrt{d} n^{\epsilon}\Big) &\leq \frac{\mathbb{E}\big[\mathrm{Tr}\left(W^{2k}\right)\big]}{d^k n^{2k\varepsilon}} \\
&\leq C''(k) n^{1-2k\epsilon},
\end{aligned}$$

which goes to 0 by choice of k. □

A finer combinatorial analysis, performed in Anderson et al. [3] (Theorem 2.1.22 p.23), yields a sharper result:

Theorem 2.1. *Assume that the W_{ij} are centered, and such that for some $\sigma > 0$, $\mathbb{E}(W_{ij}^2) \leq \sigma^2$, and for each $k > 2$, one has:*

$$\mathbb{E}|W_{ij}|^k \leq \sigma^k k^{Ck}$$

where C is a constant that does not depend on n. Then for any $\epsilon > 0$, one has

$$\lim_{n\to\infty} \mathbb{P}(\rho(W) \geq \sigma\sqrt{n}(2 + \epsilon)) = 0. \tag{2.8}$$

Note that it requires stronger moment assumptions than the previous result. It is sharp, in that if $\mathbb{E}W_{ij}^2 \equiv \sigma^2$, then $\rho(W)/(\sigma\sqrt{n})$ converges in probability to 2. The trace method was introduced by Fűredi and Komlós in [16], where they established the result (2.8) under the assumptions that $\mathbb{E}W_{ij} = 0$, $|W_{ij}| \leq 1$ almost surely, and $\mathrm{Var}(W_{ij}) \leq \sigma^2$ for some fixed $\sigma > 0$.

Remark (Further discussion on random matrix theory)**.** *Variations on the above-mentioned result in Anderson et al. [3], Theorem 2.1.22 p. 23 can be found in Bai and Silverstein [4], Theorem 5.1 p. 92. More recently, Van Vu [44] and Péché and Soshnikov [38] also establish a similar scaling for the spectral radius. However these results require strong specific conditions on the distributions of the entries of the random matrices that are typically not satisfied in the random graph context we consider, hence the need for the alternative tools that we shall next consider.*

2.2 Bernstein inequality for sums of centered independent random matrices

In this section we establish the following inequality:

Theorem 2.2. *Let $X_1, \ldots, X_m$ be independent Hermitian random matrices such that:*

$$\mathbb{E}(X_k) = 0, \quad \|X_k\|_{op} \leq L \text{ almost surely, } k \in [m]. \tag{2.9}$$

Let $Y = \sum_{k\in[m]} X_k$, and

$$v(Y) := \|\mathbb{E}(Y^2)\|_{op} = \|\sum_{k\in[m]} \mathbb{E}X_k^2\|_{op}. \tag{2.10}$$

Then for all $t > 0$,

$$\mathbb{P}(\lambda_1(Y) \geq t) \leq n \exp\left(\frac{-t^2}{2(v(Y) + Lt/3)}\right). \tag{2.11}$$

This implies for all $t > 0$

$$\mathbb{P}(\|Y\|_{op} \geq t) \leq 2n \exp\left(\frac{-t^2}{2(v(Y) + Lt/3)}\right). \tag{2.12}$$

For a Hermitian matrix $X = U\Lambda U^*$, and a function $f : \mathbb{R} \to \mathbb{R}$, the matrix $f(X)$ is defined as

$$f(X) = U \,\mathrm{diag}(f(\Lambda_{ii}))U^*.$$

We will need the following lemmas, whose proof will be given after that of the Theorem:

Lemma 2.1. *For independent Hermitian matrices X_k, $k \in [m]$, and $Y = \sum_{k\in[m]} X_k$, one has*

$$\mathbb{E}\,\mathrm{Tr}\, e^{\theta Y} \leq \mathrm{Tr}\exp\left(\sum_{k\in[m]} \ln \mathbb{E}e^{\theta X_k}\right). \tag{2.13}$$

Lemma 2.2. *For Hermitian X such that $\mathbb{E}(X) = 0$ and $\|X\|_{op} \leq L$ almost surely, then*

$$\forall \theta \in (0, 3/L), \begin{cases} \mathbb{E}e^{\theta X} \preceq \exp\left(\frac{\theta^2/2}{1-\theta L/3}\mathbb{E}X^2\right), \\ \ln \mathbb{E}e^{\theta X} \preceq \frac{\theta^2/2}{1-\theta L/3}\mathbb{E}X^2, \end{cases} \tag{2.14}$$

where $\preceq$ represents the semi-definite order on Hermitian matrices.

Lemma 2.3. *For two Hermitian matrices A, B such that $A \preceq B$, then for all $i \in [n]$, $\lambda_i(A) \leq \lambda_i(B)$.*

For a non-decreasing function $f : \mathbb{R} \to \mathbb{R}$, it follows that

$$\mathrm{Tr}\, f(A) \leq \mathrm{Tr}\, f(B).$$

Proof. (of Theorem 2.2) The second property (2.12) follows from two applications of the first inequality (2.11), to Y and $-Y$, and a union bound.

To establish (2.11), first note that for arbitrary $\theta > 0$, one has

$$\mathbb{P}(\lambda_1(Y) \geq t) \leq e^{-\theta t}\mathbb{E}\, e^{\lambda_1(Y)} \leq e^{-\theta t}\mathbb{E}\,\mathrm{Tr}(e^{\theta Y}).$$

Optimized over $\theta > 0$ this gives

$$\mathbb{P}(\lambda_1(Y) \geq t) \leq \inf_{\theta \geq 0} e^{-\theta t}\mathbb{E}\,\mathrm{Tr}(e^{\theta Y}). \tag{2.15}$$

Next use Lemma 2.1 to bound the right-hand side of the previous inequality, and obtain

$$\mathbb{P}(\lambda_1(Y) \geq t) \leq \inf_{\theta \geq 0} e^{-\theta t}\,\mathrm{Tr}\exp\left(\sum_{k\in[m]} \ln \mathbb{E}e^{\theta X_k}\right).$$

Lemma 2.2 entails that the matrix argument in the exponential is less, for the semi-positive definite order, than $\theta^2/(1-\theta L/3)\mathbb{E}\sum_{k\in[m]} X_k^2$. Lemma 2.3, used with for f the exponential function, then provides an upper bound on the right-hand side of this expression, yielding

$$\mathbb{P}(\lambda_1(Y) \geq t) \leq \inf_{\theta\in(0,3/L)} e^{-\theta t}\,\mathrm{Tr}\exp\left(\frac{\theta^2/2}{1-\theta L/3}\mathbb{E}\sum_{k\in[m]} X_k^2\right).$$

This last expression is bounded by

$$\inf_{\theta\in(0,L/3)} n\exp\left(-\theta t + \frac{\theta^2/2}{1-\theta L/3}v(Y)\right).$$

Chosing θ so that $\theta t = \frac{\theta^2}{1-\theta L/3}v(Y)$, or equivalently taking $\theta = t/(v(Y)+tL/3)$ yields the announced result (2.11). □

To prove Lemma 2.1, we will exploit Lieb's theorem, that we now state:

Theorem 2.3. *(Lieb). For a fixed $n \times n$ Hermitian matrix H, the function*

$$A \to \mathrm{Tr}\exp\left(H + \ln(A)\right)$$

defines a concave mapping on the set of $n \times n$ Hermitian positive definite matrices A.

Proof. (of Lemma 2.1). Let $Y = \sum_{k\in[m]} X_k$. Taking expectations first conditionally on $X_1, \ldots, X_{m-1}$, one has

$$\begin{aligned}\mathbb{E}\,\mathrm{Tr}\exp(Y) &\leq \mathbb{E}\left\{\mathbb{E}\left[\mathrm{Tr}\exp\left(\textstyle\sum_{k=1}^{m-1} X_k + \ln(\exp(X_m))\right)|X_1^{m-1}\right]\right\}\\ &\leq \mathbb{E}\left\{\mathrm{Tr}\exp\left(\textstyle\sum_{k=1}^{m-1} X_k + \ln(\mathbb{E}\exp(X_m))\right)\right\},\end{aligned}$$

where we used Jensen's inequality together with Lieb's theorem, the latter being applicable because $\exp(X_m)$ is a positive definite matrix. Iterating this argument gives the desired inequality (2.13). □

Proof. (of Lemma 2.2). Let $\theta \in (0, 3/L)$ be fixed. Write

$$e\theta X = I + \theta X\left(e\theta X - I - \theta X\right) = I + \theta X + Xf(X)X,$$

where

$$f(x) = \frac{e^{\theta x} - 1 - \theta x}{x^2} \text{ for } x \neq 0, \text{ and } f(0) = \frac{\theta^2}{2}\cdot$$

The function f is non-decreasing (this can be seen from its series expansion). For $x \leq L$, one therefore has $f(x) \leq f(L)$. By assumption, the eigenvalues of X do not exceed L, so that $f(X) \preceq f(L)I$. Thus necessarily, for all $u \in \mathbb{C}^n$,

$$u^* X f(X) X u = (Xu)^* f(X)(Xu) \leq f(L) u^* X^2 u,$$

and thus $Xf(X)X \preceq f(L)X^2$. It follows that

$$e^\theta X \preceq I + \theta X + f(L)X^2. \tag{2.16}$$

To conclude, note that

$$f(L) = \frac{e^{\theta L} - 1 - \theta L}{L^2} = \frac{1}{L^2} \sum_{k \geq 2} \frac{(\theta L)^{k-2}}{k!} \leq \frac{\theta^2}{2} \sum_{k \geq 2} \frac{(\theta L)^{k-2}}{3^{k-2}} = \frac{\theta^2}{2} \frac{1}{1 - \theta L/3}.$$

This yields together with (2.16), using the fact that X^2 is positive semi-definite:

$$e^\theta X \preceq I + \theta X + \frac{\theta^2/2}{1 - \theta L/3} X^2.$$

Taking expectations, using the fact that X has zero mean, we obtain

$$\mathbb{E} e^\theta X \preceq I + \frac{\theta^2/2}{1 - \theta L/3} \mathbb{E} X^2.$$

Since for all $x \in \mathbb{R}$, $(1 + x) \leq e^x$, we can deduce from this the first inequality of Lemma 2.2, i.e.

$$\mathbb{E} e^\theta X \preceq \exp\left(\frac{\theta^2/2}{1 - \theta L/3} \mathbb{E} X^2 \right).$$

To deduce from this the second inequality of the Lemma, we exploit the fact that the logarithm is *operator-monotone*, stated in the following lemma. □

Lemma 2.4. *For two positive definite Hermitian matrices A, B such that $A \preceq B$, necessarily one has $\ln(A) \preceq \ln(B)$.*

Proof. We first establish that for all non-negative $u \in \mathbb{R}_+$,

$$-(A + uI)^{-1} \preceq -(B + uI)^{-1}. \tag{2.17}$$

To that end, let $A_u = A + uI$, $B_u = B + uI$. Note that $0 \prec A_u \preceq B_u$, so that

$$0 \prec B_u^{-1/2} A_u B_u^{-1/2} \preceq I.$$

This entails that

$$I \preceq (B_u^{-1/2} A_u B_u^{-1/2})^{-1} = B_u^{1/2} A_u^{-1} B_u^{1/2},$$

which in turn guarantees that $B_u^{-1} \preceq A_u^{-1}$, or equivalently (2.17). Thus for all $u \geq 0$,

$$(1 + u)^{-1} I - (A + uI)^{-1} \preceq (1 + u)^{-1} I - (B + uI)^{-1}.$$

Note that the logarithm $\ln(A)$ admits an integral representation as

$$\ln(A) = \int_0^\infty [(1 + u)^{-1} I - (A + uI)^{-1}] du.$$

Since the semidefinite order is preserved by integration against a positive measure, the previous inequality implies that $\ln(A) \preceq \ln(B)$ as desired. □

Proof. (of Lemma 2.3) For any subspace of dimension i, the fact that $A \preceq B$ implies

$$\inf_{x \in E} \frac{x^* A x}{x^* x} \leq \inf_{x \in E} \frac{x^* B x}{x^* x}.$$

Taking suprema over B, the Courant-Fisher theorem then guarantees that $\lambda_i(A) \leq \lambda_i(B)$. Since $\operatorname{Tr} f(A) = \sum_{i \in [n]} f(\lambda_i(A))$, for non-decreasing $f : \mathbb{R} \to \mathbb{R}$, the second inequality $\operatorname{Tr} f(A) \leq \operatorname{Tr} f(B)$ follows. □

For a proof of Lieb's theorem, and more in-depth discussion of matrix concentration inequalities, the reader can consult the monograph by Tropp [42], from which the above proof is taken.

2.3 Epsilon-nets and the Feige-Ofek bound

Definition 2.1. *Let $\mathcal{M}$ be a metric space, endowed with distance $d_{\mathcal{M}}$. For a subset C of $\mathcal{M}$ and any $\epsilon > 0$, an ϵ-net $\mathcal{N}$ of C is a collection of points x of C such that C is contained in the union of balls $\mathcal{B}(x, \epsilon)$ of radius ϵ whose centres x span $\mathcal{N}$:*

$$A \subset \bigcup_{x \in \mathcal{N}} \mathcal{B}(x, \epsilon).$$

The smallest cardinality of ϵ-nets of C is denoted by $N(C, d_{\mathcal{M}}, \epsilon)$.

Our use of ϵ-nets will be via the following Lemma, which essentially reduces the control of the spectral norm of a symmetric matrix A to the control of the supremum of $|x^\top A x|$ over x in an ϵ-net of the unit sphere $\mathcal{S}^{n-1}$.

Lemma 2.5. *Let A be a symmetric $n \times n$ real matrix and let $\mathcal{S}^{n-1}$ denote the unit sphere in $\mathbb{R}^n$ for the Euclidean distance. Then for any $\epsilon \in (0, 1/2)$ and any ϵ-net $\mathcal{N}$ of $\mathcal{S}^{n-1}$, one has*

$$\sup_{x \in \mathcal{N}} \left| x^\top A x \right| \leq \|A\| \leq \frac{1}{1 - 2\epsilon} \sup_{x \in \mathcal{N}} \left| x^\top A x \right|. \tag{2.18}$$

Proof. Let $\epsilon \in (0, 1/2)$ and $\mathcal{N}$ an ϵ-net of $\mathcal{S}^{n-1}$. By Courant-Fisher, it holds that

$$\|A\| = \max(\lambda_1(A), \lambda_n(A)) = \max_{y \in \mathcal{S}^{n-1}} |y^\top A y|.$$

This establishes the left-side inequality, since $\mathcal{N} \subset \mathcal{S}^{n-1}$.

For the right-side inequality, take any $y \in \mathcal{S}^{n-1}$, and an associated $x \in \mathcal{N}$ such that $\|x - y\| \leq \epsilon$. Write then

$$y^\top A y = x^\top A x + x^\top A (y - x) + (y - x)^\top A y,$$

so that

$$|y^\top A y| \leq |x^\top A x| + 2\epsilon \|A\| \leq \sup_{x \in \mathcal{N}} |x^\top A x| + 2\epsilon \|A\|.$$

The right-side inequality follows by taking the supremum over $y \in \mathcal{S}^{n-1}$. □

Remark. *It is readily seen that the conclusion* (2.18) *of the Lemma also holds when for $\mathcal{N}$ we take an ϵ-net of the ball $\mathcal{B}_n$ centered at 0 of radius 1 in $\mathbb{R}^n$ rather than a net of its boundary $\mathcal{S}^{n-1}$.*

The following is a typical volume argument to obtain bounds on minimal net sizes $n(C, d_{\mathcal{M}}, \epsilon)$:

Lemma 2.6. *For each $n \in \mathbb{N}$, $\epsilon \in (0, 1)$, letting d_n denote the Euclidean distance in $\mathbb{R}^n$, one has*

$$N(\mathcal{B}_n, d_n, \epsilon) \leq \left(1 + \frac{2}{\epsilon}\right)^n. \tag{2.19}$$

Proof. Fix $\epsilon \in (0,1)$ and $n > 0$. Construct the ϵ-net $\mathcal{N}$ in an iterative fashion, adding at each step t a new point $x(t)$ in $\mathcal{B}_n \setminus \cup_{s=1}^{t-1} \mathcal{B}(x(s), \epsilon)$, i.e. a point that is not yet covered. By construction, the points of the net are separated by distance at least ϵ so that the balls $\mathcal{B}(x, \epsilon/2)$ have disjoint interiors. Moreover, these balls are all included in $\mathcal{B}(0, 1+\epsilon/2)$. Thus

$$\mathrm{Vol}\left(\cup_{x\in\mathcal{N}} \mathcal{B}(x, \epsilon/2)\right) = |\mathcal{N}| \mathrm{Vol}(\mathcal{B}(0, \epsilon/2)) \leq \mathrm{Vol}(\mathcal{B}(0, 1+\epsilon/2)).$$

The conclusion follows be noting that the volume of a Euclidean ball of radius r in $\mathbb{R}^n$ is $r^n \mathrm{Vol}(\mathcal{B}_n)$. □

Using techniques of Feige and Ofek [15] we shall establish the following

Theorem 2.4. *Let G be an Erdős-Rényi random graph with parameters (n,p) where the "average degree parameter" $d := np$ verifies*

$$c_0 \log(n) \leq d, \tag{2.20}$$

where $c_0 > 0$ is any positive constant. Let A denote the adjacency matrix of G, and $\bar{A}$ its expectation.

Then for every $c > 0$, there exists $c' > 0$ that depends only on c_0 and c, and such that, with probability at least $1 - n^{-c}$,

$$\|A - \bar{A}\| \leq c' \sqrt{d}.$$

As previously mentioned, the proof uses an ϵ-net $\mathcal{N}$ of $\mathcal{B}_n$ and the right-side inequality in (2.18). Let then $x \in \mathcal{N}$ be a vector of the ϵ-net. Distinguish pairs $i < j$ of indices in $[n]$ according to whether $|x_i x_j| \leq \sqrt{d}/n$, in which case (i,j) is termed a **light couple** for x, which we denote $(i,j) \in l(x)$, or whether $|x_i x_j| > \sqrt{d}/n$, in which case (i,j) is termed a **heavy couple** for x, which we denote $(i,j) \in h(x)$.

By Lemma 2.5, inequality (2.18), we have

$$\|A - \bar{A}\| \leq \frac{1}{1-2\epsilon} \left[2S_l + 2S_h\right], \tag{2.21}$$

where

$$\begin{aligned} S_l &:= \sup_{x\in\mathcal{N}} \left| \sum_{(i,j)\in l(x)} x_i (A_{ij} - \frac{d}{n}) x_j \right|, \\ S_h &:= \sup_{x\in\mathcal{N}} \left| \sum_{(i,j)\in h(x)} x_i (A_{ij} - \frac{d}{n}) x_j \right|. \end{aligned}$$

Bounding the contribution S_l of light couples: Chernoff plus union bound arguments .

Fix some $x \in \mathcal{N}$. Let $\lambda \in \mathbb{R}$. Noting $Z_{ij} := A_{ij} - d/n$, write

$$\begin{aligned} \mathbb{E} e^{\lambda \sum_{(i,j)\in l(x)} x_i x_j Z_{ij}} &= \prod_{(i,j)\in l(x)} e^{-\lambda x_i x_j d/n} \left(1 - \tfrac{d}{n} + \tfrac{d}{n} e^{\lambda x_i x_j}\right) \\ &\leq e^{\sum_{(i,j)\in l(x)} -\lambda x_i x_j (d/n) + (d/n)[e^{\lambda x_i x_j} - 1]}, \end{aligned}$$

where we used the inequality $(1+x) \leq e^x$. To proceed we will rely on the inequality

$$\forall x \in \mathbb{R},\ |x| \leq \frac{1}{2} \Rightarrow e^x - 1 - x \leq 2x^2,$$

established by considering the expansion of e^x. Plugged into the above inequality, using the fact that only light couples (i,j) are considered, this ensures that for all λ such that $|\lambda| \leq \frac{n}{2\sqrt{d}}$, one has

$$\mathbb{E} e^{\lambda \sum_{(i,j)\in l(x)} x_i x_j Z_{ij}} \leq e^{\sum_{(i,j)\in l(x)} 2(d/n)\lambda^2 x_i^2 x_j^2}.$$

The Chernoff bounding technique thus yields, taking $\lambda = n/(2\sqrt{d})$:

$$\begin{aligned} \mathbb{P}(\textstyle\sum_{(i,j)\in l(x)} x_i x_j Z_{ij} \geq C\sqrt{d}) &\leq e^{-nC/2} e^{\sum_{(i,j)\in l(x)} (n/2) x_i^2 x_j^2} \\ &\leq e^{-n(C/2 - 1/4)}, \end{aligned}$$

where we exploited the fact that $\|x\| \leq 1$ to obtain

$$\sum_{(i,j)\in l(x)} x_i^2 x_j^2 \leq \frac{1}{2} \sum_{i,j\in[n]} x_i^2 x_j^2 = \frac{1}{2}.$$

Similarly, taking $\lambda = -n/(2\sqrt{d})$ we obtain

$$\mathbb{P}(\sum_{(i,j)\in l(x)} x_i x_j Z_{ij} \leq -C\sqrt{d}) \leq e^{-n(C/2-1/4)}.$$

Taking a union bound, we thus obtain:

$$\mathbb{P}(S_l \geq C\sqrt{d}) \leq 2|\mathcal{N}|e^{-n(C/2-1/4)}.$$

By Lemma 2.6, Inequality (2.19), we thus obtain

$$\mathbb{P}(S_l \geq C\sqrt{d}) \leq 2e^{n[-C/2+1/4+\ln(1+2/\epsilon)]}.$$

By choosing constant C sufficiently large, the contribution of S_l to the upper bound (2.21) can be made less than $2C\sqrt{d}$ with probability exponentially close in n to 1.

Bounding the contribution S_h of heavy couples: the bounded discrepancy property.

We can bound S_h as follows:

$$S_h \leq T_h + U_h, \tag{2.22}$$

where

$$T_h := \sup_{x\in\mathcal{N}} |\sum_{(i,j)\in h(x)} x_i x_j A_{ij}|, \text{ and } U_h := \sup_{x\in\mathcal{N}} |\sum_{(i,j)\in h(x)} x_i x_j (d/n)|. \tag{2.23}$$

For arbitrary $x \in \mathcal{B}_n$, consider two independent random variables I, J uniformly distributed in $[n]$. Write then

$$\begin{aligned} |\sum_{(i,j)\in h(x)} x_i x_j (d/n)| &\leq (d/n) \sum_{i,j\in[n]} |x_i x_j| \mathbf{1}_{|x_i x_j| > \sqrt{d}n} \\ &\leq nd\mathbb{E}[|x_I x_J| \mathbf{1}_{|x_I x_J| > \sqrt{d}/n}] \\ &\leq nd\mathbb{E}[|x_I x_J| \frac{|x_I x_J|}{\sqrt{d}/n}] \\ &\leq n^2\sqrt{d}\mathbb{E}[x_I^2 x_J^2] \\ &\leq \sqrt{d}, \end{aligned}$$

where we used $\mathbb{E}[x_I^2] = (1/n)\|x\|^2 \leq (1/n)$ in the last step. This Markov inequality argument then implies

$$U_h \leq \sqrt{d}. \tag{2.24}$$

The last (and most delicate) step of the argument will leverage global properties of G, namely the **bounded degree property** and the **bounded discrepancy property** to obtain, uniformly in $x \in \mathcal{B}_n$, a bound in $O(\sqrt{d})$ on U_h.

Definition 2.2. *Graph G with vertex set $[n]$ is said to have the bounded degree property with average degree d and tolerance factor $c_1 > 0$ if and only if every vertex $i \in [n]$ of G has degree at most $c_1 d$.*

Graph G is said to have the bounded discrepancy property with tolerance factors $c_2, c_3 > 0$ if and only if the following holds. For any two vertex subsets $A, B \subset [n]$ such that $|B| \geq |A|$, letting $\mu(A,B) := |A| \cdot |B|(d/n)$, and $e(A,B)$ denote the number of edges with one endpoint in A and another endpoint in B, then one of the following two properties holds:

1. $\frac{e(A,B)}{\mu(A,B)} \leq ec_2$;
2. $e(A,B) \ln\left(\frac{e(A,B)}{\mu(A,B)}\right) \leq c_3|B| \ln\left(\frac{n}{|B|}\right)$.

An essential step to bound T_h is the following

Proposition 2.2. *Let graph G satisfy the bounded degree and bounded discrepancy conditions with average degree parameter d and tolerance factors c_1, (c_2, c_3) respectively. Then it holds that*

$$\sup_{x\in\mathcal{B}_n} \sum_{(i,j)\in h(x)} |x_i x_j| A_{ij} \le c_4\sqrt{d}, \tag{2.25}$$

where constant c_4 depends only on c_1, c_2, c_3.

Proof. Consider $x \in \mathcal{B}_n$, and assume without loss of generality that $x_i \ge 0$ for all $i \in [n]$. For each $k \in \mathbb{Z}$, let

$$\gamma_k := 2^k,\ A_k := \{i \in [n] : \frac{\gamma_{k-1}}{\sqrt{n}} \le x_i < \frac{\gamma_k}{\sqrt{n}}\}.$$

Since $x_i \le 1$, necessarily A_k is empty for $k > \lceil \log_2(\sqrt{n})\rceil$. We shall also use the following notations:

$$\begin{array}{ll} a_k := |A_k|, & k \le \lceil \log_2(\sqrt{n})\rceil; \\ \mu_{k\ell} := \mu(A_k, A_\ell) = a_k a_\ell \frac{d}{n}, & k,\ell \le \lceil \log_2(\sqrt{n})\rceil; \\ \lambda_{k\ell} := \frac{e(A_k,A_\ell)}{\mu_{k\ell}}, & k,\ell \le \lceil \log_2(\sqrt{n})\rceil. \end{array}$$

We then have

$$\begin{array}{ll} \sum_{(i,j)\in h(x)} x_i x_j A_{ij} & \le \sum_{k,\ell:\gamma_k\gamma_\ell\ge\sqrt{d}} e(A_k, A_\ell) \frac{\gamma_k}{\sqrt{n}} \frac{\gamma_\ell}{\sqrt{n}} \\ & = \sum_{k,\ell:\gamma_k\gamma_\ell\ge\sqrt{d}} a_k a_\ell \frac{d}{n} \lambda_{k\ell} \frac{\gamma_k}{\sqrt{n}} \frac{\gamma_\ell}{\sqrt{n}} \\ & = \sqrt{d} \sum_{k,\ell:\gamma_k\gamma_\ell\ge\sqrt{d}} \alpha_k \alpha_\ell \sigma_{k\ell}, \end{array}$$

where we introduced the notations

$$\alpha_k := a_k \frac{\gamma_k^2}{n},\ \sigma_{k\ell} := \frac{\lambda_{k\ell}\sqrt{d}}{\gamma_k\gamma_\ell},\ k,\ell \le \lceil \log_2(\sqrt{n})\rceil.$$

We then have

$$\sum_{(i,j)\in h(x)} x_i x_j A_{ij} \le 2\sqrt{d} \sum_{k,\ell} \mathbf{1}_{\gamma_k\gamma_\ell\ge\sqrt{d}} \mathbf{1}_{a_k\le a_\ell} \alpha_k \alpha_\ell \sigma_{k\ell}. \tag{2.26}$$

We shall use repeatedly the following bound, which follows from $x \in \mathcal{B}_n$ and the definition of A_k:

$$\sum_k \alpha_k \le 4 \sum_{i\in[n]} x_i^2 \le 4. \tag{2.27}$$

To upper bound the right-hand side of (2.26) we shall consider the following distinct cases for pairs of indices (k, ℓ) such that $\gamma_k\gamma_\ell \ge \sqrt{d}$ and $a_k \le a_\ell$:

$$\begin{array}{ll} \sigma_{k\ell} \le 1 & \Leftrightarrow (k,\ell) \in C_1, \\ \lambda_{kl} \le ec_2 & \Leftrightarrow (k,\ell) \in C_2, \\ \gamma_k > \sqrt{d}\gamma_\ell & \Leftrightarrow (k,\ell) \in C_3, \\ e(A_k, A_\ell)\log(\lambda_{k\ell}) \le c_3 a_\ell \log(n/a_\ell) & \Leftrightarrow (k,\ell) \in C_4. \end{array}$$

By the bounded discrepancy property, the sets C_2 and C_4 cover all possible pairs k, ℓ.

Write then, by definition of C_1:

$$\sum_{k,\ell} \mathbf{1}_{\gamma_k\gamma_\ell\ge\sqrt{d}} \mathbf{1}_{a_k\le a_\ell} \mathbf{1}_{C_1}(k,\ell) \alpha_k\alpha_\ell\sigma_{k\ell} \le \sum_{k,\ell} \alpha_k\alpha_\ell \le 16,$$

where we used (2.27). Write next, recalling the definition of $\sigma_{k\ell}$:

$$\begin{array}{ll} \sum_{k,\ell} \mathbf{1}_{\gamma_k\gamma_\ell\ge\sqrt{d}} \mathbf{1}_{a_k\le a_\ell} \mathbf{1}_{C_2}(k,\ell) \alpha_k\alpha_\ell\sigma_{k\ell} & \le \sum_{k,\ell} \mathbf{1}_{\gamma_k\gamma_\ell\ge\sqrt{d}} \alpha_k\alpha_\ell \frac{ec_2\sqrt{d}}{\gamma_k\gamma_\ell} \\ & \le \sum_{k,\ell} \alpha_k\alpha_\ell ec_2 \\ & \le 16ec_2. \end{array}$$

By the bounded degree property, one has

$$\lambda_{k\ell} \leq \frac{a_k c_1 d}{a_k a_\ell (d/n)} = \frac{n c_1}{a_\ell}.$$

Thus for fixed k,

$$\begin{aligned}\sum_\ell \mathbf{1}_{C_3}(k,\ell)\alpha_\ell \sigma_{k\ell} &= \sum_\ell \mathbf{1}_{C_3}(k,\ell) a_\ell \tfrac{\gamma_\ell^2}{n} \tfrac{\lambda_{k\ell}\sqrt{d}}{\gamma_k \gamma_\ell} \\ &\leq \sum_\ell \mathbf{1}_{C_3}(k,\ell) a_\ell \tfrac{\gamma_\ell}{n\gamma_k} \tfrac{n c_1 \sqrt{d}}{a_\ell} \\ &= \sum_{\ell:\gamma_k > \gamma_\ell \sqrt{d}} c_1 \tfrac{\gamma_\ell \sqrt{d}}{\gamma_k} \\ &\leq 2c_1,\end{aligned}$$

where we replaced the geometric sum by its value 2. This entails:

$$\sum_{k,\ell} \mathbf{1}_{\gamma_k \gamma_\ell \geq \sqrt{d}} \mathbf{1}_{a_k \leq a_\ell} \mathbf{1}_{C_3}(k,\ell) \alpha_k \alpha_\ell \sigma_{k\ell} \leq 2c_1 \sum_k \alpha_k \leq 8c_1.$$

We shall now consider summations restricted to pairs (k,ℓ) in C_4, the second case of the bounded discrepancy condition. Expressed in terms of $\lambda_{k\ell}$ and α_ℓ this condition reads:

$$\lambda_{k\ell} a_k a_\ell \frac{d}{n} \log(\lambda_{k\ell}) \leq c_3 a_\ell \log\left(\frac{\gamma_\ell^2}{\alpha_\ell}\right),$$

implying

$$\lambda_{k\ell} \sqrt{d} a_k \frac{1}{n} \log(\lambda_{k\ell}) \leq c_3 \frac{1}{\sqrt{d}} \log\left(\frac{\gamma_\ell^2}{\alpha_\ell}\right).$$

This is equivalent by definition of $\sigma_{k\ell}$ and α_k to:

$$\sigma_{k\ell}\alpha_k \log(\lambda_{k\ell}) \leq c_3 \frac{\gamma_k}{\gamma_\ell \sqrt{d}} \left[2\log(\gamma_\ell) + \log(1/\alpha_\ell)\right]. \tag{2.28}$$

We further distinguish three conditions on the couples (k,ℓ), which specify the dominant term from $\lambda_{k\ell}$, γ_ℓ and $1/\alpha_\ell$:

$$\begin{array}{ll} \log(\lambda_{k\ell}) > \frac{1}{4}\left[2\log(\gamma_\ell) + \log(1/\alpha_\ell)\right] & \Leftrightarrow (k,\ell) \in C_1', \\ 2\log(\gamma_\ell) \geq \log(1/\alpha_\ell) & \Leftrightarrow (k,\ell) \in C_2', \\ 2\log(\gamma_\ell) < \log(1/\alpha_\ell) & \Leftrightarrow (k,\ell) \in C_3'. \end{array}$$

For $(k,\ell) \in C_4 \cap C_1'$, (2.28) implies

$$\sigma_{k\ell}\alpha_k \leq 4c_3 \frac{\gamma_k}{\gamma_\ell \sqrt{d}}.$$

Thus

$$\sum_k \mathbf{1}_{C_4 \cap C_1' \setminus C_3}(k,\ell) \sigma_{k\ell}\alpha_k \leq 4c_3 \sum_{k:\gamma_k \leq \gamma_\ell \sqrt{d}} \frac{\gamma_k}{\gamma_\ell \sqrt{d}} \leq 8c_3,$$

which yields

$$\sum_{k,\ell} \mathbf{1}_{\gamma_k \gamma_\ell \geq \sqrt{d}} \mathbf{1}_{a_k \leq a_\ell} \mathbf{1}_{C_4 \cap C_1' \setminus C_3}(k,\ell) \alpha_k \alpha_\ell \sigma_{k\ell} \leq 8c_3 \sum_\ell \alpha_\ell \leq 32 c_3.$$

For $(k,\ell) \in C_2'' := C_2' \setminus C_1'$, one has

$$\log(\lambda_{k\ell}) \leq \frac{1}{4}[2\log(\gamma_\ell) + \log(1/\alpha_\ell)] \leq \log(\gamma_\ell),$$

hence $\lambda_{k\ell} \leq \gamma_\ell$. Assume further $(k,\ell) \notin C_1$. Thus

$$1 \leq \sigma_{k\ell} = \frac{\lambda_{k\ell}\sqrt{d}}{\gamma_k \gamma_\ell} \leq \frac{\sqrt{d}}{\gamma_k}. \tag{2.29}$$

Assume also $(k,\ell) \notin C_2$, so that $\lambda_{k\ell} \geq ec_2$, and without loss of generality, that constant c_2 appearing in the bounded discrepancy property satisfies $c_2 \geq 1$, so that $\log(\lambda_{k\ell}) \geq 1$. Then, if moreover $(k,\ell) \in C_4$, using (2.28) we obtain

$$\sigma_{k\ell}\alpha_k \leq c_3 \frac{\gamma_k}{\gamma_\ell \sqrt{d}} 4 \log(\gamma_\ell).$$

This entails

$$\sum_k \mathbf{1}_{a_k \leq a_\ell} \mathbf{1}_{C_2'' \cap C_4 \setminus (C_1 \cup C_2)}(k,\ell)\alpha_k \sigma_{k\ell} \leq 4c_3 \sum_k \mathbf{1}_{\gamma_k \leq \sqrt{d}} \frac{\gamma_k}{\sqrt{d}} \frac{\log(\gamma_\ell)}{\gamma_\ell},$$

where we used (2.29). Upper-bounding $\log(\gamma_\ell)/\gamma_\ell$ by 1, we get that the above sum is at most $8c_3$. This entails

$$\sum_{k,\ell} \mathbf{1}_{\gamma_k \gamma_\ell \geq \sqrt{d}} \mathbf{1}_{a_k \leq a_\ell} \mathbf{1}_{C_2'' \cap C_4 \setminus (C_1 \cup C_2)}(k,\ell)\alpha_k \alpha_\ell \sigma_{k\ell} \leq 8c_3 \sum_\ell \alpha_\ell \leq 32 c_3.$$

Assume finally that $(k,\ell) \in C_3'' := C_3' \setminus C_1'$. Then $\log(\lambda_{k\ell}) \leq \log(1/\alpha_\ell)$. This implies

$$\sigma_{k\ell} = \frac{\lambda_{k\ell}\sqrt{d}}{\gamma_k \gamma_\ell} \leq \frac{1}{\alpha_\ell} \frac{\sqrt{d}}{\gamma_k \gamma_\ell}.$$

Thus

$$\sum_\ell \mathbf{1}_{\gamma_k \gamma_\ell \geq \sqrt{d}} \mathbf{1}_{a_k \leq a_\ell} \mathbf{1}_{C_3''}(k,\ell)\sigma_{k\ell}\alpha_\ell \leq \sum_{\ell : \gamma_\ell \gamma_k \geq \sqrt{d}} \frac{\sqrt{d}}{\gamma_k \gamma_\ell} \leq 2.$$

This entails

$$\sum_{k,\ell} \mathbf{1}_{\gamma_k \gamma_\ell \geq \sqrt{d}} \mathbf{1}_{a_k \leq a_\ell} \mathbf{1}_{C_3''}(k,\ell)\alpha_k \alpha_\ell \sigma_{k\ell} \leq 8.$$

Collecting all bounds, we arrive at

$$\sum_{(i,j) \in h(x)} x_i x_j A_{ij} \leq 2\sqrt{d}\left[16 + 16ec_2 + 8c_1 + 64c_3 + 8\right].$$

The announced result therefore holds with

$$c_4 := 16[12 + c_1 + 2ec_2 + 8c_3].$$

□

Let us now establish the following

Proposition 2.3. *For an Erdős-Rényi graph $G(n,p)$ with average degree parameter $d = np$ satisfying*

$$c_0 \log(n) \leq d,$$

where $c_0 > 0$ is an arbitrary positive constant, then for any $c > 0$ there exist constants $c_1, c_2, c_3 > 0$ that depend only on c_0 and c and such that the bounded degree and bounded discrepancy properties with corresponding average degree d and tolerance factors c_i hold with probability at least $1 - n^{-c}$.

Proof. **Bounded degree property:** For a Binomial random variable $X \sim \mathrm{Bin}(m,q)$, a Poisson random variable $Y \sim \mathrm{Poi}(mq)$ with the same mean and $\lambda \in \mathbb{R}$, one has

$$\mathbb{E}e^{\lambda X} = (1 - q + qe^\lambda)^m \leq e^{(mq)(e^\lambda - 1)} = \mathbb{E}e^{\lambda Y}.$$

Thus a Binomial random variable admits tighter Chernoff bounds than a Poisson random variable with the same mean.

The degree D_i of any node i in G is distributed as $\mathrm{Bin}(n-1, d/n)$. Thus for any $\kappa > 1$ we have:

$$\mathbb{P}(D_i \geq \kappa d) \leq e^{-dh(\kappa)}$$

where $h(x) := x\ln(x) - x + 1$ is the so-called Cramér transform of a Poi(1) random variable. The probability that there is some $i \in [n]$ with $D_i \geq \kappa d$ is thus no larger than

$$ne^{-dh(\kappa)} \leq n^{1-c_0 h(\kappa)}.$$

It thus suffices to choose $c_1 > 1$ such that: $c_0 h(c_1) > c - 1$ to ensure that the bounded degree property with average degree d and tolerance c_1 holds with probability at least $1 - n^{-c}$.

Bounded discrepancy property: Let us assume that graph G satisfies the bounded degree property with parameter c_1. Let $A, B \subset [n]$ with respective sizes a, b such that $a \leq b$. Let $\mu := \mu(A, B) = ab(d/n)$.

Assume first that $b \geq n/e$. By the bounded degree property, $e(A, B) \leq adc_1$ so that $e(A, B)/\mu \leq ec_1$.

Assume now that $b \leq n/e$. Since $e(A, B)$ is Binomial with mean no larger than μ (the case where it equals μ corresponds to $A \cap B = \emptyset$), by the previous Chernoff bound applied to Binomial random variables, we have for $\kappa > 1$ that

$$\mathbb{P}(e(A, B) \geq \kappa\mu) \leq e^{-\mu h(\kappa)}.$$

To use this in a union bound over sets A, B, we will choose κ such that for all $a \leq b \leq n/e$,

$$e^{-\mu h(\kappa)} \binom{n}{a} \binom{n}{b} \leq n^{-2-c}.$$

A sufficient condition for this to hold, based on the inequality

$$\binom{n}{m} \leq \left(\frac{ne}{m}\right)^m,$$

is given by

$$a\,(1 + \ln(n/a)) + b\,(1 + \ln(n/b)) + (2 + c)\ln(n) \leq \mu h(\kappa).$$

It can be checked (exercise!) that the function $x \to x\ln(n/x)$ is monotonic non-decreasing on $[1, n/e]$. Since $a \leq b$ and $b \leq n/e$, a sufficient condition for the above inequality to hold is then provided by

$$4b\ln(n/b) + (2 + c)\ln(n) \leq \mu h(\kappa).$$

Again by monotonicity, $\ln(n) \leq b\ln(n/b)$, hence a further sufficient condition is

$$(6 + c)\frac{b}{\mu}\ln(n/b) \leq h(\kappa).$$

Remark that for $\kappa \geq 4$, $h(\kappa) \geq \kappa\ln(\kappa)/3$. Let then for all b, $1 \leq b \leq n/e$:

$$\kappa(b) = \inf\{k \in [4, +\infty) : k\ln(k) \geq 3(6 + c)(b/\mu)\ln(n/b)\}.$$

Then by a union bound we have that with probability at least $1 - n^{-c}$, for all $A, B \subset [n]$ with $a := |A| \leq b := |B| \leq n/e$,

$$e(A, B) \leq \kappa(b)\mu(A, B). \tag{2.30}$$

If $\kappa(b) = 4$, then the first condition for bounded discrepancy holds with $ec_2 = 4$. Otherwise, $\kappa(b)\ln(\kappa(b)) = 3(6 + c)(b/\mu)\ln(n/b)$ so that

$$e(A, B) \leq \frac{3(6 + c)b}{\ln(\kappa(b))}\ln(n/b),$$

and in turn

$$e(A, B)\ln(\kappa(b)) \leq 3(6 + c)b\ln(n/b).$$

In view of (2.30), this entails

$$e(A, B)\ln\left(\frac{e(A, B)}{\mu(A, B)}\right) \leq 3(6 + c)|B|\ln\left(\frac{n}{|B|}\right),$$

that is to say the second condition for bounded discrepancy holds, with tolerance $c_3 = 3(6 + c)$. □

Remark (Sharpness of conditions in Theorem 2.4)**.** .

The fact that the spectral radius of $\|A - \bar{A}\|$ is of order at least $\sqrt{d}$, where $d = np$ is the average degree, can readily be obtained by considering the all-ones vector.

This spectral radius can be $\gg \sqrt{d}$ for $d \ll \log(n)$. This is for instance seen by considering $d = \Theta(1)$, for which there are in $\mathcal{G}(n,p)$ isolated stars with $\Theta(\log(n)/\log\log n)$ branches, inducing eigenvalues of A of order $\Theta(\sqrt{\log(n)/\log\log n})$, hence eigenvalues in $A - \bar{A}$ of the same order that is large compared to $\sqrt{d} = O(1)$.

Finer analysis of the exact necessary condition on d for which $\|A - \bar{A}\| = O(\sqrt{d})$: add reference to recent paper by Bordenave et al.

We shall in fact need the following consequence of Theorem 2.4:

Theorem 2.5. *Let A be a random symmetric matrix, with entries A_{ij} independent up to symmetry, $A_{ij} \in [0,1]$, and such that $\mathbb{E}(A_{ij}) \leq d/n$ for some upper bound parameter d. Assume that for some positive constant c_0,*

$$c_0 \log(n) \leq d \leq n(1 - c_0). \tag{2.31}$$

Then for all $c > 0$ there exists $\tilde{c} > 0$ such that with probability at least $1 - 1/n^c$, one has

$$\rho(A - \mathbb{E}(A)) \leq \tilde{c}\sqrt{d}. \tag{2.32}$$

Proof. The centred matrix $A - \mathbb{E}(A)$ has its entries in $[-d/n, 1]$, so that the centred matrix $B := (1 - d/n)(A - \mathbb{E}(A))$ has its entries in $[-d/n, 1 - d/n]$. It is therefore stochastically smaller, for the convex ordering, than the matrix $C - \mathbb{E}(C)$ where C is the adjacency matrix of a $G(n, d/n)$ Erdős-Rényi random graph (maybe more details on this comparison). By Strassen's theorem one can construct the matrices B and C on the same probability space such that

$$\mathbb{E}(C - \mathbb{E}(C)|B) = B.$$

The spectral radius $\rho(\cdot)$ is a convex function. Thus by Jensen's inequality,

$$\rho(B) = \rho(\mathbb{E}(C - \mathbb{E}(C)|B)) \leq \mathbb{E}(\rho(C - \mathbb{E}(C))|B).$$

Let for notational convenience $S = \rho(B)$ and $R = \rho(C - \mathbb{E}(C))$. The previous display then implies $S \leq \mathbb{E}(R|S)$. By Theorem 2.4 we have the existence, for all $c > 1$, of c' such that $R \leq c'\sqrt{d}$ with probability at least $1 - n^{-c}$. Moreover, the spectral radius of an $n \times n$ matrix with entries bounded in absolute value by 1 is at most n.

We thus have, for $t = c'\sqrt{d}$,

$$R \leq n\mathbf{1}_{R>t} + t\mathbf{1}_{R\leq t}.$$

Multiplying by $\mathbf{1}_{S>t+1}$ and taking conditional expectations with respect to S this yields

$$S\mathbf{1}_{S>t+1} \leq n\mathbf{1}_{S>t+1}\mathbb{P}(R > t|S) + t\mathbf{1}_{S>t+1}\mathbb{P}(R \leq t|S),$$

so that

$$\mathbb{P}(R > t|S) \geq \mathbf{1}_{S>t+1}\frac{S-t}{n-t} \geq \mathbf{1}_{S>t+1}\frac{1}{n-t}.$$

In turn, taking expectations this yields

$$\mathbb{P}(S > t+1) \leq n\mathbb{P}(R > t) \leq n^{1-c}.$$

Letting $c' = \phi(c)$ be the constant provided by Theorem 2.4, the previous display gives

$$\mathbb{P}((1 - d/n)\|A - \bar{A}\| \geq \phi(c+1)\sqrt{d} + 1) \leq n^{-c},$$

so that, using the assumption $(1 - d/n) \geq c_0$, and taking n large enough to ensure $1 \leq \sqrt{d}$, Inequality (2.32) indeed holds with probability at least $1 - n^{-c}$ for $\tilde{c} = [\phi(c+1) + 1]/c_0$. □

2.4 Kolchinskii-Giné results for graphon-related matrices

Before stating the main result of this section, we recall without proof the following result.

Theorem 2.6. *(Kato [21]) Let T be a linear, self-adjoint operator acting on a Hilbert space of functions $\mathcal{L}^2(\mathcal{X}, \pi)$, where π is a non-negative measure. Assume that T is compact, i.e. it maps any bounded set to a set whose closure is compact. Then T admits a discrete spectrum $\{\lambda_k\}_{k\geq 1}$, an associated orthonormal basis of eigenfunctions $\{\phi_k\}_{k\geq 1}$, and it can be represented as*

$$Tf(x) = \sum_{k\geq 1} \phi_k(x)\lambda_k \int_{\mathcal{X}} \phi_k(y) f(y)\pi(dy).$$

We now establish the following result, which is a special case of results by von Luxburg et al. [43], themselves variations on earlier results by Koltchinskii and Giné, see [24].

Theorem 2.7. *Fix a compact metric space $\mathcal{X}$, a symmetric, continuous function $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$, and a probability measure π on $\mathcal{X}$. Let $X_1, \ldots, X_n$ be i.i.d., distributed according to π. Let $M^{(n)} = (n^{-1}K(X_i, X_j))_{i,j\in[n]}$. Define the operator T on $\mathcal{L}^2(\mathcal{X}, \pi)$ via*

$$Tf(x) := \int_{\mathcal{X}} K(x, y) f(y)\pi(dy).$$

Consider its spectrum, split into $\{\lambda_i^+\}_{i\geq 1}$, its positive eigenvalues sorted in decreasing order, and $\{\lambda_i^-\}_{i\geq 1}$, its negative eigenvalues sorted in increasing order.

Denote by $\lambda_i^{+,(n)}$ the i-th largest positive eigenvalue of M, and $\lambda_i^{-,(n)}$ its i-th smallest negative eigenvalue. Then for each $i \geq 1$, we have the convergence in probability $\lim_{n\to\infty} \lambda_i^{\pm,(n)} = \lambda_i^{\pm}$.

Let i_0 such that $\lambda_{i_0}^{\pm} \neq 0$ is of multiplicity d, with $\lambda_{i_0}^{\pm} = \cdots = \lambda_{i_0+d-1}^{\pm}$ and $v_i^{\pm,(n)}$ orthonormal eigenvectors of M associated to $\lambda_i^{\pm,(n)}$, $i = i_0, \ldots, i_0 + d - 1$. There exist orthonormal functions $\{\psi_i^{\pm}\}$ for $i = i_0, \ldots, i_0 + d - 1$ of T associated with $\lambda_{i_0}^{\pm}$ such that in probability, $\|v_i^{\pm,(n)} - \{n^{-1/2}\psi_i^{\pm}(X_k)\}_{k\in[n]}\| \to 0$.

Proof. By Example 2.19, p.264 in [21], operator T is compact, and thus admits a discrete spectrum $\{\lambda_i\}_{i\geq 1}$, here considered as sorted by decreasing modulus. Denote by $\{\phi_i\}_{i\geq 1}$ an associated orthonormal basis of eigenfunctions. K is moreover such that

$$K(x, y) = \sum_{i\geq 1} \lambda_i \phi_i(x)\phi_i(y),$$

and

$$\sum_{i\geq 1} \lambda_i^2 = \int_{\mathcal{X}^2} K(x, y)^2 \pi(dx)\pi(dy).$$

Let $\delta_0 > 0$ be defined as

$$\delta_0 := \inf_{j:\lambda_j \neq \lambda_{i_0}^{\pm}} |\lambda_j - \lambda_{i_0}^{\pm}|.$$

Fix some (small) $\epsilon > 0$, and let $j_0 \geq i_0$ be such that

$$\sum_{j>j_0} \lambda_j^2 \leq \epsilon^2,$$

and that $\lambda_{i_0}^{\pm}$ appears d times (its multiplicity) in the sequence $(\lambda_1, \ldots, \lambda_{j_0})$. Let for each $j \in [j_0]$,

$$v_j(k) = \frac{\phi_j(X_k)}{\sqrt{n}},\ k \in [n].$$

For all $i, j \in [i_0]$, one has, by the law of large numbers, almost sure convergence of $\langle v_j, v_j \rangle$ to $\mathbf{1}_{i=j}$. Let $w_1, \ldots, w_{j_0}$ be the n-dimensional vectors obtained by Gram-Schmidt orthonormalization of $(v_1, \ldots, v_{j_0})$. Let $V := [v_1, \ldots, v_{j_0}]$, $W := [w_1, \ldots, w_{j_0}]$, and R the $j_0 \times j_0$ matrix such that

$$VR = W.$$

Thus R converges almost surely to the identity matrix.

Let $w_{j_0+1}, \ldots, w_n$ complement $w_1, \ldots, w_{j_0}$ into an orthonormal basis. Let $\Lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_{i_0})$. Decompose $M^{(n)}$ as

$$\begin{array}{l} M^{(n)} = A + B + C + D, \text{ where :} \\ A = W\Lambda W^\top,\ B = (M^{(n)} - V\Lambda V^\top),\ C = V\Lambda(V^\top - W^\top),\ D = (V - W)\Lambda W^\top. \end{array}$$

The fact that $VR = W$ with R converging almost surely to the $j_0 \times j_0$ identity matrix as $n \to \infty$ ensures that

$$\begin{array}{rl} \|C\|_{op} & \leq \|W\|_{op}\|R^{-1}\|_{op}\|\Lambda\|_{op}\|R^{-1} - I\|_{op}\|W\|_{op} \\ & = \|R^{-1}\|_{op}|\lambda_1|\|R^{-1} - I\|_{op} \\ & \to 0 \text{ almost surely as } n \to \infty. \end{array}$$

A similar evaluation yields the almost sure limit $\lim_{n\to\infty} \|D\|_{op} = 0$.

Note that

$$B_{k\ell} = \frac{1}{n} K'(X_k, X_\ell),$$

with

$$K'(x, y) = K(x, y) - \sum_{j \in [j_0]} \lambda_j \phi_j(x)\phi_j(y) = \sum_{j > j_0} \lambda_j \phi_j(x)\phi_j(y).$$

One then has

$$\|B\|_F^2 = \frac{1}{n^2} \sum_{k,l \in [n]} K'(X_k, X_l)^2.$$

The following

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{k,l\in[n]} K'(X_k, X_l)^2 = \int_{\mathcal{X}^2} K'(x, y)^2 \pi(dx)\pi(dy)$$

holds almost surely, see e.g. Serfling (it can be proven straightforwardly from the usual strong law of large numbers for continuous K' and compact $\mathcal{X}$, but holds more generally). From the expression of K', this yields

$$\lim_{n\to\infty} \|B\|_F^2 = \sum_{j > j_0} \lambda_j^2 \leq \epsilon^2.$$

Thus $M^{(n)} = W\Lambda W^T + E$, where $\|E\|_{op} \leq 2\epsilon$ for large enough n.

We can now apply Weyl's inequality, to deduce that for all fixed $j \geq 1$,

$$\limsup_{n\to\infty} |\lambda_j^{\pm,(n)} - \lambda_j^{\pm}| \leq 2\epsilon.$$

Consider now the eigenvalues $\lambda_j^{\pm,(n)}$ of $M^{(n)}$ for $j = i_0, \ldots, i_0 + d - 1$. Let $\tilde{V} = [\tilde{v}_{i_0}^{\pm}, \ldots, \tilde{v}_{i_0+d-1}^{\pm}]$ where the $\tilde{v}_j^{\pm}$ are an orthonormal system of eigenvectors of $M^{(n)}$ associated with the $\lambda_j^{\pm,(n)}$, and $\tilde{W} = [w_{i_0}^{\pm}, \ldots, w_{i_0+d-1}^{\pm}]$ is constructed from the columns of W whose indices j are such that $\lambda_j = \lambda_{i_0}^{\pm}$. The Davis-Kahan $\sin\Theta$ theorem yields the existence of a $d \times d$-orthogonal matrix Θ such that

$$\|\tilde{V}\Theta - \tilde{W}\|_F \leq \frac{2^{3/2} d^{1/2} \|E\|_{op}}{\delta_0}.$$

The right-hand side is asymptotically no more than $2^{3/2}d^{1/2}2\epsilon/\delta_0$. Let $\hat{V} := [v_{i_0}^{\pm}, \ldots, v_{i_0+d-1}^{\pm}]$ be constructed from the columns v_j of V such that $\lambda_j = \lambda_{i_0}^{\pm}$. Since R converges to the identity, we have

$$\lim_{n\to\infty} \|\tilde{W} - \hat{V}\|_F = 0.$$

This then yields

$$\limsup_{n\to\infty} \|\tilde{V}\Theta - \hat{V}\|_F \leq 2^{3/2}d^{1/2}2\epsilon/\delta_0.$$

Since ϵ can be chosen arbitrarily small, $\tilde{V}$ is asymptotically close in Frobenius norm to $\hat{V}\Theta^\top$. By construction, this latter matrix's columns can be written $(n^{-1/2}\psi_j^{\pm}(X_k))_{k\in[n]}$ for a system of orthonormal eigenfunctions $\psi_j^{\pm}$ of K associated with eigenvalue $\lambda_{i_0}^{\pm}$. □

Chapter 3

Community detection in the strong signal regime

3.1 The Stochastic Block Model and the Graphon Model

For a number of vertices $n \in \mathbb{N}$, a number of *blocks* or *communities* $K > 0$, a probability distribution $\alpha = \{\alpha_k\}_{k\in[K]}$ and a symmetric matrix $P \in [0,1]^{K\times K}$, the Stochastic Block Model $\mathcal{G}(n, \alpha, P)$ is defined as follows.

The types σ_i of nodes $i \in [n]$ are sampled i.i.d. according to α. Conditionally on $\sigma_{[}n] := \{\sigma_i\}_{i\in[n]}$, each (unoriented) edge (i, j) is present with probability P_{σ_i,σ_j}, independently over pairs (i, j), $1 \le i < j \le n$. Accordingly, for a target collection of types $s_{[n]} \in [K]^n$ and an edge set e,

$$\mathbb{P}(\sigma_{[n]} = s_{[n]}, E = e) = \prod_{i\in[n]} \alpha_{s_i} \prod_{i<j} \left(P_{s_i,s_j}\mathbf{1}_{(i,j)\in e} + (1 - P_{s_i,s_j})\mathbf{1}_{(i,j)\notin e}\right).$$

We shall also consider the following extension. Let $\mathcal{X}$ be a compact metric space, endowed with a probability measure π and a symmetric, continuous function $F : \mathcal{X} \times \mathcal{X} \to [0, 1]$. The graphon model $\mathcal{G}(n, \pi, F)$ is then defined as follows. Node types σ_i are sampled i.i.d. from $\mathcal{X}$ according to π. Conditionally on $\sigma_{[n]}$, edge (i, j) is present with probability $F(\sigma_i, \sigma_j)$.

3.2 Spectral community detection for the SBM in the strong signal regime

We define the strong signal regime for the SBM as follows. We assume K, α, and the $K\times K$ symmetric matrix B to be given. We then consider the random graph G distributed as $\mathcal{G}(n, \alpha, (d/n)B)$ where d represents the average degree, or the signal strength in our observed graph G. We say we are in a strong signal regime when d diverges to infinity with n.

We shall first establish the following

Theorem 3.1. *Assume that* $\sqrt{\ln(n)} \ll d \ll n^\delta$ *for some constant* $\delta \in (0, 1)$. *Assume that* $\min_{k\in[K]} \alpha_k > 0$ *and that matrix* $B \in \mathbb{R}_+^{K\times K}$ *has* distinguishable blocks, *i.e. for all* $k \neq \ell \in [K]$, $\exists m \in [K]$ *such that* $B_{km} \neq B_{\ell m}$. *Let* R *be the rank of matrix* B. *The with high probability:*

i) The spectrum of A, *the adjacency matrix of a spectral block model* $G \sim \mathcal{G}(n, \alpha, (d/n)B)$ *consists of* R *eigenvalues* $\lambda_1, \dots, \lambda_R$ *of order* $\Theta(d)$, *and* $n - R$ *eigenvalues of order* $o(d)$.

ii) Spectral embedding based on the orthonormal set of eigenvectors $x_1, \dots, x_R$ *of* A *associated with* $\lambda_1, \dots, \lambda_R$, *associating to vertex* i *the vector* $z_i = \sqrt{n}(x_1(i), \dots, x_R(i)) \in \mathbb{R}^R$ *captures the block structure in the following sense:*

For some $\epsilon > 0$ depending only on B and α, for all vertices $i, j \in [n] \setminus \mathcal{B}$, where $\mathcal{B}$ is a set of "bad" vertices of negligible size $|\mathcal{B}| = o(n)$, one has

$$\|z_i - z_j\| = \begin{cases} o(1) & \text{if } \sigma_i = \sigma_j, \\ \geq \epsilon & \text{otherwise.} \end{cases}$$

Let

$$\overline{A}_{i,j} := \frac{d}{n} B_{\sigma_i \sigma_j} \text{ and } W = A - \overline{A}. \tag{3.1}$$

The proof of Theorem 3.1 proceeds with the following steps. We first etablish the following

Lemma 3.1. *Under the assumptions of Theorem 3.1,*

$$\text{with high probability, } \rho(W) \ll d. \tag{3.2}$$

Proof. Perturbation matrix W is the sum of diagonal matrix $\text{Diag}(\{-(d/n)B_{\sigma_i \sigma_i})$, whose spectral radius is obviously $O(d/n)$ and thus a fortiori $o(d)$, and of matrix $(A_{ij} - \mathbb{E}(A_{ij}))_{i,j \in [n]}$, whose spectral radius can be bounded using the results of the previous chapter. In particular, Theorem 2.5 applies in case $d \leq n^\delta$ with $\delta < 1/5$, with $\bar{b} = O(1)$ and $\omega = \max(d, \log n)$, yielding

$$\rho(W) = O\left(\sqrt{\max(d, \log n)}\right).$$

The assumption $\sqrt{\log n} \ll d$ then implies (3.2).

In the case where $d \geq n^\kappa$ for some $\kappa > 0$, Proposition 2.1 entails that with high probability, $\rho(W) \leq \sqrt{d} n^\xi$ for arbitrary $\xi > 0$. Choosing $\xi < \kappa/2$ implies again (3.2). □

Remark. *Under the assumption $d \gg \log n$, Bernstein's inequality 2.2 applies here (exercise!) with $v = \Theta(d)$ and $L = \Theta(1)$, and can also be used to establish* (3.2).

The next result describes the structure of matrix $\overline{A}$:

Lemma 3.2. *For all $t \in \mathbb{R}^K$, define $\phi(t) \in \mathbb{R}^n$ as $\phi(t)_i := t_{\sigma_i}$. Symmetric matrix $\overline{A} \in \mathbb{R}^{n \times n}$ defined in* (3.1) *verifies*

$$\overline{A}\phi(t) = d\phi(Mt), \quad t \in \mathbb{R}^K, \tag{3.3}$$

where $M = B\text{Diag}(\{\tilde{\alpha}_u\}_{u \in [K]} \in \mathbb{R}^{K \times K}$ and $\tilde{\alpha}_u = n^{-1} \sum_{i \in [n]} \mathbf{1}_{\sigma_i = u}$.

Let $(\mu_u, t_u)_{u \in [R]}$ be pairs of non-zero eigenvalues and associated eigenvectors of M. The spectrum of $\overline{A}$ consists in R non-zero eigenvalues $\lambda_u = d\mu_u$, $u \in [R]$ with associated eigenvectors $\phi(t_u)$, and eigenvalue 0 with multiplicity $n - R$.

Proof. Identity (3.3) follows from the block structure of $\overline{A}$. The image of any vector by $\overline{A}$ is block-constant so that any eigenvector associated with a non-zero eigenvalue is block-constant. Thus any eigenvector associated with a non-zero eigenvalue reads $\phi(t)$ for some $t \in \mathbb{R}^K$. The result follows. □

Corollary 3.1. *By the distinguishability assumption of blocks (for all $k \neq \ell \in [K]$, there exists $m \in [K]$ with $B_{km} \neq B_{\ell m}$) ensures the existence of positive ϵ function of B et $\{\tilde{\alpha}_m\}_{m \in [K]}$ such that for any choice of orthonormal eigenvectors $\bar{x}_1, \ldots, \bar{x}_R$ of $\overline{A}$ associated with its non-zero eigenvalues λ_u, $u \in [R]$, the eigenvectors $\bar{z}_i = \sqrt{n}(\bar{x}_1(i), \ldots, \bar{x}_R(i))^\top \in \mathbb{R}^R$ verify*

$$\forall i, j \in [n], \begin{cases} \sigma_i = \sigma_j \Rightarrow \|\bar{z}_i - \bar{z}_j\| = 0, \\ \sigma_i \neq \sigma_j \Rightarrow \|\bar{z}_i - \bar{z}_j\| > \epsilon. \end{cases} \tag{3.4}$$

Proof. Let $t_u \in \mathbb{R}^K$ be such that $\sqrt{n}\bar{x}_u = \phi(t_u)$. Then noting $\sqrt{\tilde{\alpha}} = \mathrm{Diag}(\{\sqrt{\tilde{\alpha}_u}\}_{u\in[K]})$, the vectors $\{\sqrt{\alpha}t_u\}_{u\in[R]}$ are orthonormal, since the $\bar{x}_u$ are orthonormal. t_u is an eigenvector of M, so that $\sqrt{\alpha}t_u$ is an eigenvector of symmetric matrix $\sqrt{\tilde{\alpha}}B\sqrt{\tilde{\alpha}}$ with associated eigenvalue μ_u. Thus

$$\sqrt{\alpha}B\sqrt{\alpha} = \sum_{u\in[R]} \mu_u \sqrt{\tilde{\alpha}} t_u t_u^T \sqrt{\tilde{\alpha}},$$

hence

$$B = \sum_{u\in[R]} \mu_u t_u t_u^T. \tag{3.5}$$

Clearly for two vertices $i, j \in [n]$ such that $\sigma_i = \sigma_j$, then $\bar{z}_i = \bar{z}_j$. If $\sigma_i = u \neq \sigma_j = v$, as $\bar{z}_i = (t_1(u), \ldots, t_R(u))^\top$ and $\bar{z}_i = (t_1(v), \ldots, t_R(v))^\top$, we conclude from (3.5) that, if $\bar{z}_i = \bar{z}_j$ held, then B would have two identical lines, which contradicts the distinguishability hypothesis. The set of possible collections of eigenvectors $\{t_u\}_{u\in[R]}$ is compact, and the quantities

$$\|\bar{z}_i - \bar{z}_j\| = \sqrt{\sum_{w\in[R]} (t_w(u) - t_w(v))^2}$$

for $u \neq v$ are continuous functions of these vectors, which are strictly positive. They are thus lower-bounded by some $\epsilon > 0$ that is a function of B and $\tilde{\alpha}$ only. □

Proof. (of Theorem 3.1). Weyl's inequalities (1.4) together with (3.2) and Lemma 3.2 ensure the announced property of the eigenvalues of A.

The Davis-Kahane theorem 1.3 ensures that for a system of orthonormal eigenvectors $x_1, \ldots, x_R$ of A associated with its R eigenvalues of magnitude $\Theta(d)$, there exists a corresponding orthonormal system of eigenvectors $\bar{x}_1, \ldots, \bar{x}_R$ of $\overline{A}$ associated with its non-zero eigenvalues and verifying

$$\langle x_i, \bar{x}_i \rangle = 1 - O((\rho/d)^2),$$

where $\rho = \rho(W)$, or equivalently $\|x_i - \bar{x}_i\|^2 = O((\rho/d)^2) = o(1)$. For $i \in [n]$, the R-dimensional vector z_i is defined as $\sqrt{n}(x_1(i), \ldots, x_R(i))^\top$. We define similarly $\bar{z}_1(i) = \sqrt{n}(\bar{x}_1(i), \ldots, \bar{x}_R(i))^\top$. Thus

$$\sum_{i\in[n]} \|z_i - \bar{z}_i\|^2 = n \sum_{j\in[R]} \|x_j - \bar{x}_j\|^2 = n\theta(n),$$

where $\theta(n) = o(1)$. Bienaymé-Tchebitchev inequality then yields:

$$|\{i \in [n] : \|z_i - \bar{z}_i\| \geq \theta(n)^{1/3}\}| \leq n\theta(n)^{1/3} = o(n).$$

Corollary 3.1 ensures that the $\bar{z}_i$ coincide within a block, and differ by at least some $\epsilon > 0$ bounded away from 0 between block, which concludes the proof. □

This result suggests algorithms for the choice of the embedding dimension k used to construct the spectral embedding $z_u := \sqrt{n}(x_1(u), \ldots, x_k(u)) \in \mathbb{R}^k$, for instance letting $\hat{R} = \min\{i \in [n] : |\lambda_i| \geq T|\lambda_{i+1}|\}$, where the eigenvalues λ_i of A are sorted by decreasing absolute value and $T > 1$ is a large constant. With high probability, for large enough T, $\hat{R} = R$ under the assumptions of the theorem.

It can also be used to prove that specific clustering algorithms applied to the R-dimensional spectral embedding will correctly classify all but a vanishing fraction of the vertices.

Consider for instance the following clustering strategy. Having formed ℓ groups $C_1, \ldots, C_\ell$ of vertices in $[n]$, so long as there remain $\epsilon_1 n$ un-grouped vertices, i.e. so long as

$$\sum_{m=1}^{\ell} |C_m| \leq (1 - \epsilon_1)n,$$

choose one un-grouped vertex $i_{\ell+1}$ uniformly at random, and define group $C_{\ell+1}$ as those remaining vertiecs j such that $\|z_{i_{\ell+1}} - z_j\| \leq \epsilon_2$.

Then part ii) of the theorem ensures that, for $\epsilon_1, \epsilon_2 > 0$ sufficiently small, with high probability this procedure produces K groups, and there exists a permutation s of $[K]$ such that for all $\ell \in [K]$:

$$|C_\ell \setminus D_{s(\ell)}| + |D_{s(\ell)} \setminus C_\ell| = o(n), \tag{3.6}$$

where $D_\ell := \{i \in [n] : \sigma_i = \ell\}$. In other words, this procedure correctly reconstructs the K blocks assciated with the underlying vertex classes $\sigma_i \in [K]$, except perhaps for a negligible fraction of vertices, which could be either mis-classified or unclassified.

To see this, assume that property (3.6) is verified at step $\ell < K$. If $\epsilon_1 < \min_{m\in[K]} \alpha_m$, the construction does not stop, and a uniform choice for vertex $i_{\ell+1}$ falls with probability $1 - o(1)$in one of the D_m, $m \in [K] \setminus \{s(1), \ldots, s(\ell)\}$, and does not belong to the set $\mathcal{B}$ of "bad vertices". For $\epsilon_2 > 0$ sufficiently small, ii) ensures that C_ℓ verifies

$$D_m \setminus \mathcal{B} \subset C_\ell \subset D_m \cup \mathcal{B}.$$

Defining $s(\ell + 1) = m$, the announced property (3.6) is then satisfied at step $\ell + 1$. Finally, after the K-th step, the number of remaining vertices is $o(n)$ and thus the process stops.

The schemes we just described involve three parameters T^{-1}, ϵ_1, ϵ_2 and the argument necessitates all three to be sufficiently small with respect to model parameters. Schemes have been proposed for selecting such parameters as functions of the observed graph so as to ensure these schemes succeed with probability $1 - o(1)$ for arbitrary values of the model parameters. Details can be found in [28].

3.3 Inference for graphon models in the strong signal regime

Let $\mathcal{X}$ be a compact metric space, endowed with a probability measure π and a symmetric, continuous function $F : \mathcal{X} \times \mathcal{X} \to [0, 1]$. The graphon model $\mathcal{G}(n, \pi, F)$ is then defined as follows. Node types σ_i are sampled i.i.d. from $\mathcal{X}$ according to π. Conditionally on $\sigma_{[n]}$, edge (i, j) is present with probability $F(\sigma_i, \sigma_j)$.

It can be seen as an extension of the SBM to an infinite collection of blocks. The definition of successful reconstruction in that setup must therefore be adapted.

One possible notion of reconstruction is as follows: infer from the observed graph a symmetric matrix $\hat{K}_{ij}$ providing an accurate estimation of the matrix $F(\sigma_i, \sigma_j)$ of probabilities of edge presence.

We consider the following strong signal regime: assume that the distribution π on metric compact space $\mathcal{X}$ is fixed, together with a continuous symmetric function $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$. Assume then that $F(x, y) = \frac{d}{n} K(x, y)$, for some "signal strength" parameter d. By strong signal regime, we refer to this special setting, with signal strength $d \to +\infty$ with $n \to \infty$.

The following statement will involve the spectral analysis of operator

$$T : f \in \mathcal{L}^2(\pi) \to Tf(y) = \int_{\mathcal{X}} K(x, y) f(y) \pi(dy) \in \mathcal{L}^2(\pi),$$

in terms of its eigenvalues $\lambda_i(T)$, sorted by decreasing absolute value: $|\lambda_1(T)| \geq |\lambda_2(T)| \geq \cdots$, and orthonormal systems of associated eigenfunctions ψ_i:

$$\int_{\mathcal{X}} \psi_i(y)\psi_j(y)\pi(dy) = \mathbf{1}_{i=j}.$$

Theorem 3.2. *Let $A \sim \mathcal{G}(n, \pi, (d/n)K)$. Let $R \geq 1$ be fixed , and such that $|\lambda_{R+1}(T)| < |\lambda_R(T)|$. Let $u_1, \ldots, u_R$ be an orthonormal collection of eigenvectors of A associated with the eigenvalues $\lambda_1(A), \ldots, \lambda_R(A)$ of A. Assume that $\sqrt{\log n} \ll d \ll n^\delta$ for some fixed $\delta \in (0, 1)$. Define for all $i, j \in [n]$,*

$$\hat{K}_{ij} = \sum_{\ell=1}^{R} \lambda_\ell(A) u_\ell(i) u_\ell(j). \tag{3.7}$$

Then with high probability, one has

$$\sum_{i,j\in[n]} \left[\frac{n}{d}\hat{K}_{ij} - K(\sigma_i,\sigma_j)\right]^2 = o(n^2) + O(n^2\epsilon_R^2), \tag{3.8}$$

where $\sigma_i \in \mathcal{X}$ is the type of vertex i, $i \in [n]$, and

$$\epsilon_R^2 = \sum_{\ell>R} \lambda_\ell(K)^2. \tag{3.9}$$

Proof. Let

$$M = \frac{1}{n}\left(K(\sigma_i,\sigma_j)\right)_{i,j\in[n]},$$

and let $\{\lambda_\ell(M)\}_{\ell\in[n]}$ denote its eigenvalues, sorted by decreasing absolute value. Then the argument of Lemma 3.1 applies, showing, with $W = A - dM$, that with high probability $\rho(W) \ll d$. Weyl's inequality (1.4) then ensures that

$$\left|\frac{1}{d}\lambda_\ell(A) - \lambda_\ell(M)\right| = o(1), \ell \in [R]. \tag{3.10}$$

The Davis-Kahane theorem 1.3 also ensures the existence of an orthonormal collection of eigenvectors $v_1,\dots,v_R$ of M associated with $\lambda_1(M),\dots,\lambda_R(M)$ such that

$$\|u_\ell - v_\ell\| = o(1),\ \ell \in [R]. \tag{3.11}$$

Moreover, Theorem 2.7 ensures that the eigenvalues $\lambda_\ell(K)$ of operator T satisfy

$$|\lambda_\ell(M) - \lambda_\ell(T)| = o(1),\ \ell \in [R], \tag{3.12}$$

and guarantees the existence of an orthonormal system of eigenfunctions ψ_ℓ such that

$$\sum_{i\in[n]} \left[v_\ell(i) - \frac{1}{\sqrt{n}}\psi_\ell(\sigma_i)\right]^2 = o(1),\ \ell \in [R]. \tag{3.13}$$

Note that for some constant C_R, one has for each $i,j \in [n]$:

$$\left[\frac{n}{d}\hat{K}_{ij} - K(\sigma_i,\sigma_j)\right]^2 \le C_R\left\{\sum_{\ell\in[R]}[Y_{ij}(\ell,1) + Y_{ij}(\ell,2) + Y_{ij}(\ell,3)] + Y_{ij}(\ell,4)\right\},$$

where:

$$\begin{aligned}
Y_{ij}(\ell,1) &= \left[(\tfrac{n}{d}\lambda_\ell(A) - n\lambda_\ell(T))u_\ell(i)u_\ell(j)\right]^2,\\
Y_{ij}(\ell,2) &= \left[n\lambda_\ell(T)[u_\ell(i)u_\ell(j) - v_\ell(i)v_\ell(j)]\right]^2,\\
Y_{ij}(\ell,3) &= \left[\lambda_\ell(T)[nv_\ell(i)v_\ell(j) - \psi_\ell(\sigma_i)\psi_\ell(\sigma_j)]\right]^2,\\
Y_{ij}(4) &= \left[\textstyle\sum_{\ell\in[R]}\lambda_\ell(T)\psi_\ell(\sigma_i)\psi_\ell(\sigma_j) - K(\sigma_i,\sigma_j)\right]^2.
\end{aligned}$$

Evaluations (3.10), (3.12) together with the fact that the vectors u_ℓ are normed ensures that $\sum_{i,j} Y_{ij}(\ell,1) = o(n^2)$.

Evaluations (3.11) and (3.13) guarantee respectively that $\sum_{i,j} Y_{ij}(\ell,2)$ and $\sum_{i,j} Y_{ij}(\ell,3)$ are both $o(n^2)$.

Finally, the law of large numbers mentioned in the proof of Theorem 2.7 ensures that

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{i,j\in[n]} Y_{ij}(4) = \int_{\mathcal{X}^2}[\sum_{\ell\in[R]}\lambda_\ell(T)\psi_\ell(x)\psi_\ell(y) - K(x,y)]^2\pi(dx)\pi(dy).$$

The conclusion then follows from the fact that the integral in the right-hand side of the previous display equals precisely ϵ_R^2, in view of the spectral decomposition

$$K(x,y) = \sum_{\ell\ge1}\lambda_\ell(T)\psi_\ell(x)\psi_\ell(y).$$

□

Part II

Statistical physics and Belief Propagation

Chapter 4

Graphical models prerequisites

In this chapter we introduce the basic definitions of graphical models, and classical results that will be needed later on.

4.1 Pairwise graphical models and Markov random fields

Definition 4.1. *Let $\mathcal{G} := (\mathcal{V}, \mathcal{E})$ be a simple undirected graph with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$, $\mathcal{X}$ a finite alphabet, and functions $\psi_i : \mathcal{X} \to \mathbb{R}_+$, $i \in \mathcal{V}$, $\psi_e : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$, $e \in \mathcal{E}$. The probability distribution μ on $\mathcal{X}^{\mathcal{V}}$ defined by*

$$\mu(x) := \frac{1}{Z} \prod_{i \in \mathcal{V}} \psi_i(x_i) \prod_{e=(i,j) \in \mathcal{E}} \psi_{i,j}(x_i, x_j)$$

is a pairwise graphical model with underlying graph $\mathcal{G}$. In the above expression, Z is a normalization constant given by

$$Z = \sum_{x \in \mathcal{X}^{\mathcal{V}}} \prod_{i \in \mathcal{V}} \psi_i(x_i) \prod_{e=(i,j) \in \mathcal{E}} \psi_{i,j}(x_i, x_j),$$

and also referred to as the partition function.

Example 4.1. *The Ising model from statistical physics corresponds to the alphabet $\mathcal{X} = \{-1, 1\}$, $\psi_i(x_i) = \exp(h_i x_i)$ and $\psi_{i,j}(x_i, x_j) = \exp(J_{ij} x_i x_j)$. In that case random variable X_i is known as the spin at site (or vertex) i, parameter $h_i \in \mathbb{R}$ is the external field at site i, and J_{ij} is the coupling coefficient between the spins at sites i and j.*

Definition 4.2. *Given an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a finite alphabet $\mathcal{X}$ and a distribution μ on $\mathcal{X}^{\mathcal{V}}$, the pair $(\mathcal{G}, \mu)$ is a Markov random field if there exist functions $\psi_K : \mathcal{X}^K \to \mathbb{R}_+$ indexed by the cliques K of $\mathcal{G}$ and a normalization constant Z such that:*

$$\mu(x) = \frac{1}{Z} \prod_K \psi_K(x_K), \quad , x \in \mathcal{X}^{\mathcal{V}}, \tag{4.1}$$

where we denote $x_K := (x_i)_{i \in K}$, and the product runs over all cliques of $\mathcal{G}$.

It is immediate that a pairwise graphical model is a Markov random field (if the graph has cliques K that contain more than two vertices, take the corresponding ψ_K identically equal to 1).

The factored form of a Markov random field has the following implication. Let A, B, C be subsets of vertices such that C *separates* A from B, that is any path in $\mathcal{G}$ from A to B must traverse C. Let A' denote the set of nodes $i \in V \setminus (A \cup B \cup C)$ such that there is a path in the graph from A to i that does not cross C. Let also $B' = V \setminus (A \cup B \cup C \cup A')$. Then, because C separates A from B, $(A \cup A') \cap (B \cup B') = \emptyset$, and cliques K of G can comprise, besides nodes of C, either nodes of $A \cup A'$, or nodes in $B \cup B'$, but cannot comprise

nodes of both $A \cup A'$ and $B \cup B'$. This implies that $\mu(x)$ can be factored as $F(x_{A\cup A'\cup C})G(x_{B\cup B'\cup C})$. It follows that under μ, conditionally on X_C, X_A and X_B are independent. Indeed this amounts to showing that

$$\mathbb{P}(X_{A\cup C} = x_{A\cup C})\mathbb{P}(X_{B\cup C} = x_{B\cup C}) = \mathbb{P}(X_{A\cup B\cup C} = x_{A\cup B\cup C})\mathbb{P}(X_C = x_C),$$

which follows by writing these probabilities as sums over values y_A, $y_{A'}$, y_B, $y_{B'}$. It constitutes the direct part of the following theorem:

Theorem 4.1. *(Hammersley-Clifford). Given a Markov random field $(\mathcal{G}, \mu)$, for any subsets A, B, C of $\mathcal{V}$ such that C separates A from B, then X_A and X_B are independent conditionally on X_C under μ.*

Conversely, let μ be a distribution on $\mathcal{X}^{\mathcal{V}}$ such that $\mu(x) > 0$ for all $x \in \mathcal{X}^{\mathcal{V}}$ and for all subsets A, B, C of $\mathcal{V}$ such that C separates A from B, then X_A and X_B are independent conditionally on X_C under μ. Then $(\mathcal{G}, \mu)$ is a Markov random field.

Proof. The argument for the direct part has been sketched above. The converse is established as follows.

Fix $x^* \in \mathcal{X}^{\mathcal{V}}$. For any subset $S \subseteq \mathcal{V}$, let

$$\phi_S(x_S) := \prod_{U\subseteq S} \mu(x_U, x^*_{V\setminus U})^{(-1)^{|S\setminus U|}}.$$

We first establish that

$$\mu(x) = \mu(x^*) \prod_{S\subseteq\mathcal{V}, S\neq\emptyset} \phi_S(x_S). \tag{4.2}$$

Recall the identity

$$\sum_{B\subseteq A} (-1)^{|B|} = \sum_{B\subseteq A} (-1)^{|A\setminus B|} = \mathbf{1}_{A=\emptyset}. \tag{4.3}$$

Next note that, since by (4.3), $\sum_{S\subseteq V}(-1)^{|S|} = 0$, one has

$$\prod_{S\subseteq\mathcal{V}, S\neq\emptyset} \phi_S(x_S) = \prod_{S\subseteq\mathcal{V}, S\neq\emptyset} \prod_{U\subseteq S} \mu(x_U, x^*_{V\setminus U})^{(-1)^{|S\setminus U|}} = \prod_{U\subseteq\mathcal{V}} \mu(x_U, x^*_{V\setminus U})^{\kappa_U},$$

where

$$\kappa_U := \sum_{S\neq\emptyset, U\subseteq S\subseteq V} (-1)^{|S\setminus U|}.$$

Using (4.3), one has that $\kappa_\emptyset = -1$, $\kappa_{\mathcal{V}} = 1$ and $\kappa_U = 0$ for $U \neq \emptyset, \mathcal{V}$. This establishes (4.2).

We now show that for each S that is not a clique of $\mathcal{G}$, necessarily $\phi_S(x_S) \equiv 1$.

To that end, first remark that, if for some $j \in S$, $x_j = x^*_j$, necessarily $\phi_S(x_S) = 1$. Indeed:

$$\phi_S(x_S) = \prod_{U\subseteq S\setminus\{j\}} \left(\frac{\mu(x_U, x^*_{\mathcal{V}\setminus U})}{\mu(x_{U+j}, x^*_{\mathcal{V}\setminus(U+j)})} \right)^{(-1)^{|S\setminus U|}},$$

and the ratio in the above equals 1 when $x^*_j = x_j$.

Next assume that S is not a clique, i.e. there are $i, j \in S$ such that (i, j) is not an edge of $\mathcal{G}$. Write then

$$\phi_S(x_S) = \prod_{U\subseteq\mathcal{V}\setminus\{i\}} \left(\frac{\mu(x_U, x^*_{\mathcal{V}\setminus U})}{\mu(x_{U+i}, x^*_{\mathcal{V}\setminus(U+i)})} \right)^{(-1)^{|S\setminus U|}}.$$

Now the above ratio is, by the Markov conditional independence property, independent of the value of x_j. Indeed, for a fixed U, let $K = \mathcal{V} \setminus \{i, j\}$, for $k \in \mathcal{V}$, $y_k = x_k$ if $k \in U$, $y_k = x^*_k$ if $k \in \mathcal{V} \setminus U$. One then has:

$$\frac{\mu(x_U, x^*_{\mathcal{V}\setminus U})}{\mu(x_{U+i}, x^*_{\mathcal{V}\setminus(U+i)})} = \frac{\mathbb{P}(X_i = x^*_i|X_K = y_K)\mathbb{P}(X_j = y_j|X_K = y_K)}{\mathbb{P}(X_i = x_i|X_K = y_K)\mathbb{P}(X_j = y_j|X_K = y_K)} = \frac{\mathbb{P}(X_i = x^*_i|X_K = y_K)}{\mathbb{P}(X_i = x_i|X_K = y_K)}.$$

We may thus replace x_j by x^*_j. Having done so, we can now conclude by the previous result that $\phi_S(x_S) = 1$. □

4.2 Extremal characterization of Gibbs measures

The Markov random field distributions we have just seen are of the form

$$\mu(x) = \frac{1}{Z}\psi(x), \quad x \in \mathcal{X}^{\mathcal{V}},$$

where ψ is given in explicit form, but Z is not. Equivalently, it can be written as the celebrated Boltzmann distribution $Z^{-1}e^{-E(x)/T}$ with energy function $E(x)$ and temperature T by making the identification $\psi(x) = e^{-E(x)/T}$.

We have the following

Definition 4.3. *The Gibbs free energy functional $\mathbb{G}$ is defined on the set $M(\mathcal{X}^{\mathcal{V}})$ of probability measures on $\mathcal{X}^{\mathcal{V}}$ as*

$$\mathbb{G}(\nu) := -H(\nu) - \mathbb{E}_\nu \ln \psi(x), \tag{4.4}$$

where $H(\nu) := \sum_x \nu(x)\ln(1/\nu(x))$ is the Shannon entropy.

We then have the following

Proposition 4.1. *The Gibbs free energy $\mathbb{G}(\nu)$ is strictly convex on $M(\mathcal{X}^{\mathcal{V}})$, is minimal at $\nu = \mu$, the Boltzmann-Gibbs distribution, at which point it equals $-\ln(Z)$.*

Proof. The function $z \to z\ln(z)$ is strictly convex on $\mathbb{R}_+$, hence strict convexity of $\mathbb{G}$. To determine the minimum of $\mathbb{G}$ over $M(\mathcal{X}^{\mathcal{V}})$, introduce the Lagrangian

$$\mathcal{L}(\nu, \lambda) := \mathbb{G}(\nu) + \lambda(1 - \sum_x \nu(x)).$$

Setting $d\mathcal{L}/d\nu(x) = 1 + \ln(\nu(x)) - \ln\psi(x) - \lambda$ to zero yields $\nu(x) = C\psi(x)$ for some constant C. This constant must be Z^{-1} for ν to be a probability measure. Using the expression of $\mu(x)$ it is readily verified that $\mathbb{G}(\mu) = -\ln(Z)$. □

Remark. *Recall that the Kullback-Leibler divergence $D(p\|q)$ between two distributions p, q on the discrete set $\mathcal{X}$ is by definition*

$$D(p\|q) := \sum_{x\in\mathcal{X}} p(x)\ln\left(\frac{p(x)}{q(x)}\right).$$

The Gibbs free energy $\mathbb{G}(\nu)$ can then be expressed as

$$\mathbb{G}(\nu) = -\ln(Z) + D(\nu\|\mu),$$

and the proof of the above proposition gives the classical result that Kullback-Leibler divergences $D(p\|q)$ are non-negative, equal to zero if and only if $p = q$.

4.3 Tree Markov fields and belief propagation

Consider a tree Markov field, that is a Markov random field $(\mathcal{G}, \mu)$ where graph $\mathcal{G}$ is in fact a tree. In that particular case, tasks of interest such as evaluation of the partition function Z, computation of marginal distributions, etc become tractable.

Specifically, assume to be given functions $\psi_i : \mathcal{X} \to \mathbb{R}_+$, $\psi_{ij} : \mathcal{X}^2 \to \mathbb{R}_+$ such that

$$\mu(x) = \frac{1}{Z}\prod_{i\in\mathcal{V}} \psi_i(x_i) \prod_{(ij)\in\mathcal{E}} \psi_{ij}(x_i, x_j), \quad x \in \mathcal{X}^{\mathcal{V}}.$$

For any edge (ij) of $\mathcal{G}$, let $i \to j$ be an arbitrary orientation associated to it. Consider the subtree $\mathcal{T}_{i\to j}$ of $\mathcal{G}$, consisting of nodes $k \in \mathcal{V}$ that remain connected to i in $\mathcal{G}$ if we remove the edge (ij) from $\mathcal{G}$. In other words,

$\mathcal{T}_{i\to j}$ is the connected component of i in the graph $(\mathcal{V}, \mathcal{E}\setminus(ij))$. We let $\mathcal{V}_{i\to j}$, $\mathcal{E}_{i\to j}$ denote the corresponding sets of vertices and edges respectively.

Consider further the Markov random field on $\mathcal{T}_{i\to j}$ corresponding to the functions $\{\psi_k, k\in\mathcal{V}_{i\to j}\}$ and $\{\psi_{k\ell}, (k\ell)\in\mathcal{E}_{i\to j}\}$.

Let $\{b_{i\to j}(x_i)\}_{x_i\in\mathcal{X}}$ denote the corresponding marginal distribution of X_i. We will further let $\mu_i(x_i)$ (respectively, $\mu_{ij}(x_i,x_j)$) denote the marginal distribution of X_i (respectively, (X_i,X_j)) for the original Markov random field.

We then have:

$$b_{i\to j}(x_i) = \frac{1}{Z_{i\to j}} \sum_{x_k, k\in\mathcal{V}_{i\to j}\setminus\{i\}} \prod_{k\in\mathcal{V}_{i\to j}} \psi_k(x_k) \prod_{(k\ell)\in\mathcal{E}_{i\to j}} \psi_{k\ell}(x_k,x_\ell), \tag{4.5}$$

where the normalization constant $Z_{i\to j}$ is given by

$$Z_{i\to j} = \sum_{x_k, k\in\mathcal{V}_{i\to j}} \prod_{k\in\mathcal{V}_{i\to j}} \psi_k(x_k) \prod_{(k\ell)\in\mathcal{E}_{i\to j}} \psi_{k\ell}(x_k,x_\ell). \tag{4.6}$$

For a leaf node i of $\mathcal{G}$, this simplifies to

$$b_{i\to j}(x_i) = \frac{1}{Z_{i\to j}}\psi_i(x_i),\ Z_{i\to j} = \sum_{x_i\in\mathcal{X}} \psi_i(x_i). \tag{4.7}$$

For a non-leaf node i, to obtain a recursive relation on $b_{i\to j}$ we split the product in (4.5) into the factor $\psi_i(x_i)\prod_{k\sim i, k\neq i}\psi_{ki}(x_k,x_i)$ and, for each $k\sim i$, $k\neq i$, the factor

$$\prod_{\ell\in\mathcal{V}_{k\to i}} \psi_\ell(x_\ell) \prod_{(\ell m)\in\mathcal{E}_{k\to i}} \psi_{\ell m}(x_\ell,x_m).$$

This readily implies, by distributing for each $k\sim i$, $k\neq j$ the summation over $x_\ell, \ell\in\mathcal{V}_{k\to i}\setminus\{k\}$ to the corresponding factor, the identity:

$$b_{i\to j}(x_i) = \frac{1}{Z_{i\to j}}\psi_i(x_i) \prod_{k\sim i, k\neq j} \sum_{x_k} \psi_{ki}(x_k,x_i) Z_{k\to i} b_{k\to i}(x_k). \tag{4.8}$$

This identity implies a recursive formula for the normalization constants:

$$Z_{i\to j} = \sum_{x_i\in\mathcal{X}} \psi_i(x_i) \prod_{k\sim i, k\neq j} \sum_{x_k} \psi_{ki}(x_k,x_i) Z_{k\to i} b_{k\to i}(x_k). \tag{4.9}$$

Exactly the same reasoning further gives the following identities:

$$\mu_i(x_i) = \frac{1}{Z}\psi_i(x_i) \prod_{k\sim i} \sum_{x_k} \psi_{ki}(x_k,x_i) Z_{k\to i} b_{k\to i}(x_k), \tag{4.10}$$

and

$$Z = \sum_{x_i\in\mathcal{X}} \psi_i(x_i) \prod_{k\sim i} \sum_{x_k} \psi_{ki}(x_k,x_i) Z_{k\to i} b_{k\to i}(x_k). \tag{4.11}$$

The above equations allow to compute recursively, for each edge of the original tree graph and its two orientations, both the partial normalization constants $Z_{i\to j}$ and the distribution $b_{i\to j}$. They further allow to compute the marginal distributions μ_i and the partition function Z. Moreover, they readily imply that the joint marginal distribution μ_{ij}, for an edge $(ij)\in\mathcal{E}$, is given by

$$\mu_{ij}(x_i,x_j) = \frac{1}{Z}\psi_{ij}(x_i,x_j) Z_{i\to j} b_{i\to j}(x_i) Z_{j\to i} b_{j\to i}(x_j). \tag{4.12}$$

One usually does not keep track of the partial normalization constants $Z_{i\to j}$, and rather expresses the above relations as:

$$b_{i\to j}(x_i) \propto \psi_i(x_i) \prod_{k\sim i, k\neq j} \sum_{x_k} \psi_{ki}(x_k, x_i) b_{k\to i}(x_k), \tag{4.13}$$

$$\mu_i(x_i) \propto \psi_i(x_i) \prod_{k\sim i} \sum_{x_k} \psi_{ki}(x_k, x_i) b_{k\to i}(x_k), \tag{4.14}$$

$$\mu_{ij}(x_i, x_j) \propto \psi_{ij}(x_i, x_j) b_{i\to j}(x_i) b_{j\to i}(x_j). \tag{4.15}$$

The first equation, (4.13), is known in the literature as "belief propagation", and also as the sum-product algorithm. The above derivations establish the following

Theorem 4.2. *For a Markov random field $(\mathcal{G}, \mu)$ whose graph $\mathcal{G}$ is a tree, the belief propagation (or sum-product) algorithm* (4.13)*, initialized on leaf nodes i of $\mathcal{G}$ with* (4.7)*, converges in a finite number of steps. Its limit is such that the marginal distributions μ_i and μ_{ij} for $i \in \mathcal{V}$ and $(ij) \in \mathcal{E}$ are given by* (4.14) *and* (4.15) *respectively.*

We end this section with another remarkable property of tree Markov fields.

Theorem 4.3. *For a tree Markov random field $(\mathcal{G}, \mu)$, the distribution μ can be written as:*

$$\mu(x) = \prod_{i\in\mathcal{V}} \mu_i(x_i) \prod_{(ij)\in\mathcal{E}} \frac{\mu_{ij}(x_i, x_j)}{\mu_i(x_i)\mu_j(x_j)}, \quad x \in \mathcal{X}^{\mathcal{V}}. \tag{4.16}$$

Denoting by d_i the degree of node i in $\mathcal{G}$, it can alternatively be written

$$\mu(x) = \prod_{i\in\mathcal{V}} \mu_i(x_i)^{1-d_i} \prod_{(ij)\in\mathcal{E}} \mu_{ij}(x_i, x_j), \quad x \in \mathcal{X}^{\mathcal{V}}. \tag{4.17}$$

Proof. The proof is by induction on the number of nodes $n = |\mathcal{V}|$. Formula (4.16) obviously holds for $n = 1$. For $n > 1$, let $i \in \mathcal{V}$ be a leaf node of $\mathcal{G}$, and j be the only neighbour of i in $\mathcal{G}$. Node j separates i from $\mathcal{V} \setminus i$. We then have, by the conditional independence property 4.1:

$$\mu(x) = \mu_{ij}(x_i, x_j) \frac{1}{\mu_j(x_j)} \mathbb{P}(X_{\mathcal{V}\setminus i} = x_{\mathcal{V}\setminus i}).$$

Now, $X_{\mathcal{V}\setminus i}$ is a Markov random field on the tree graph obtained from $\mathcal{G}$ by removing node i and the edge (ij). Indeed, this can be shown by summing over $x_i \in \mathcal{X}$ the expression for $\mu(x)$. The induction hypothesis applied to this tree Markov field then gives the result. □

4.4 Chow-Liu trees and maximum likelihood estimation

Given a sample $X(1), \ldots, X(N)$, where each $X(n) \in \mathcal{X}^{\mathcal{V}}$ is a vector with coordinates in $\mathcal{X}$ indexed by $i \in \mathcal{V}$, consider the question of determining a tree Markov field $(\mu, \mathcal{T})$ maximizing the (log-) likelihood

$$L = \sum_{n=1}^{N} \log \mu(X(n)).$$

We shall parameterize the desired tree Markov field by its graph component $\mathcal{T} = (\mathcal{V}, \mathcal{E})$ and by the marginal distributions μ_i, μ_{ij}, $i \in \mathcal{V}$, $(ij) \in \mathcal{E}$.

Introduce for all $i \neq j \in \mathcal{V}$ the empirical distributions

$$\hat{\mu}_i(x) = \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{X_i(n)=x}, \ \hat{\mu}_{ij}(x_i, x_j) = \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{X_i(n)=x_i, X_j(n)=x_j}.$$

The log-likelihood reads, in view of (4.17):

$$\begin{aligned} L &= \sum_{n=1}^{N} \sum_{i\in\mathcal{V}} (1-d_i) \ln[\mu_i(X_i(n))] + \sum_{(i,j)\in\mathcal{E}} \ln\left[\mu_{ij}(X_i(n), X_j(n))\right] \\ &= N \sum_i (1-d_i) \sum_{x_i} \hat{\mu}_i(x_i) \ln \mu_i(x_i) + N \sum_{(i,j)\in\mathcal{E}} \sum_{x_i,x_j} \hat{\mu}_{ij}(x_i,x_j) \ln\left[\mu_{ij}(x_i,x_j)\right] \end{aligned}$$

We denote this formula by $N * \Delta[\hat{\mu}, \mu|\mathcal{T}]$. Let i_0 be a leaf node of $\mathcal{T}$, with edge (i_0, j_0) incident to i_0 in $\mathcal{T}$. Define also $\mathcal{T}'$ to be the tree obtained from $\mathcal{T}$ by removal of i_0 from $\mathcal{V}$ and (i_0, j_0) from $\mathcal{E}$. We then have

$$\Delta[\hat{\mu}, \mu|\mathcal{T}] = \sum_{x,y} \hat{\mu}_{i_0,j_0}(x,y) \ln\left[\frac{\mu_{i_0,j_0}(x,y)}{\mu_{j_0}(y)}\right] + \Delta[\hat{\mu}, \mu|\mathcal{T}']. \tag{4.18}$$

The first term in the right-hand side of (4.18) reads

$$\begin{aligned} \sum_y \hat{\mu}_{j_0}(y) \sum_x \hat{\mu}_{i_0|j_0}(x|y) \ln \mu_{i_0|j_0}(x|y) &= \sum_y \hat{\mu}_{j_0}(y) \sum_x \hat{\mu}_{i_0|j_0}(x|y) \left[\ln\left(\tfrac{\mu_{i_0|j_0}(x|y)}{\hat{\mu}_{i_0|j_0}(x|y)}\right) + \ln\left(\hat{\mu}_{i_0|j_0}(x|y)\right)\right] \\ &= \sum_y \hat{\mu}_{j_0}(y) \left[-D(\hat{\mu}_{i_0|j_0}(\cdot|y) \| \mu_{i_0|j_0}(\cdot|y)) - H(\hat{\mu}_{i_0|j_0}(\cdot|y)\right]. \end{aligned}$$

From the property of Kullback-Leibler divergence discussed in Remark 4.2, we see that we should choose $\mu_{i_0|j_0}(x|y) \equiv \hat{\mu}_{i_0|j_0}(x|y)$ to maximize this first term.

Since the second term $\Delta[\hat{\mu}, \mu|\mathcal{T}']$ in the right-hand side of (4.18) is of similar nature to the left-hand side, with a tree $\mathcal{T}'$ with one node less than $\mathcal{T}$, we obtain by induction on $|\mathcal{V}|$ that for fixed tree $\mathcal{T}$, one should choose $\mu_{ij} = \hat{\mu}_{ij}$ for each edge (i,j) in order to maximize the likelihood of the observations.

Maximum likelihood estimation is then completed by optimizing over tree $\mathcal{T}$. In view of (4.17), this corresponds to the following

$$\max_{\mathcal{T}:\text{ tree on } \mathcal{V}} \sum_{(i,j)\in\mathcal{E}(\mathcal{T})} \sum_{x,y} \hat{\mu}_{ij}(x,y) \ln\left(\frac{\hat{\mu}_{ij}(x,y)}{\hat{\mu}_i(x)\hat{\mu}_j(x)}\right). \tag{4.19}$$

Note that the sum corresponding to edge (i,j) in the above is the mutual information $I(X_i; X_j)$ between variables X_i and X_j with joint distribution $\hat{\mu}_{ij}$.

Thus maximum likelihood estimation is performed by identifying a tree $\mathcal{T}$ on the node set $\mathcal{V}$ of maximal weight, as measured by the sum over its edges (i,j) of the mutual information $I(X_i; X_j)$.

Efficient algorithms for determining a maximum weight tree are available, for instance Kruskal's algorithm iteratively adds edges of maximal weight which do not lead to cycles in the graph being constructed.

4.5 Bethe free energy and belief propagation

For a given distribution ν on $\mathcal{X}^{\mathcal{V}}$, the "energy" term $-\mathbb{E}_\nu \ln(\psi(x))$ in the Gibbs free energy reads, for a pairwise graphical model,

$$\begin{aligned} -\mathbb{E}_\nu \ln(\psi(x)) &= -\sum_{x\in\mathcal{X}^{\mathcal{V}}} \nu(x) \left[\sum_{i\in\mathcal{V}} \ln(\psi_i(x_i)) + \sum_{(ij)\in\mathcal{E}} \ln(\psi_{ij}(x_i,x_j))\right] \\ &= -\sum_{i\in\mathcal{V}} \sum_{x_i\in\mathcal{X}} \nu_i(x_i) \ln(\psi_i(x_i)) - \sum_{(ij)\in\mathcal{E}} \sum_{x_i,x_j\in\mathcal{X}} \nu_{ij}(x_i,x_j) \ln(\psi_{ij}(x_i,x_j)), \end{aligned}$$

where we denoted by ν_i and ν_{ij} the one-point and two-point marginals of distribution ν.

The Bethe free energy provides an approximation of the Gibbs free energy that is expressed only in terms of such marginals ν_i, ν_{ij}. The approximation of the entropy term in terms of such marginals is based on the expression (4.17) of a tree-Markov field distribution, and reads:

$$H_{\text{Bethe}}(\nu) = \sum_{(ij)\in\mathcal{E}} \sum_{x_i,x_j\in\mathcal{X}} \nu_{ij}(x_i,x_j) \ln(1/\nu_{ij}(x_i,x_j)) + \sum_{i\in\mathcal{V}} \sum_{x_i\in\mathcal{X}} (1-d_i)\nu_i(x_i) \ln(1/\nu_i(x_i)).$$

This leads to the expression for the Bethe free energy of a Markov random field in terms of the marginal distributions ν_i, ν_{ij}:

$$\begin{aligned} \mathbb{G}_{\text{Bethe}}(\{\nu_i\},\{\nu_{ij}\}) &= \sum_{(ij)\in\mathcal{E}} \sum_{x_i,x_j\in\mathcal{X}} \nu_{ij}(x_i,x_j) \left[-\ln\psi_{ij}(x_i,x_j) + \ln(\nu_{ij}(x_i,x_j))\right] \\ &\quad + \sum_{i\in\mathcal{V}} \sum_{x_i\in\mathcal{X}} \nu_i(x_i) \left[-\ln\psi_i(x_i) + (1-d_i)\ln(\nu_i(x_i))\right]. \end{aligned} \tag{4.20}$$

Determining μ by minimizing the Gibbs free energy may not be tractable for general graphs $\mathcal{G}$. Instead, one can try to minimize the Bethe free energy (4.20) over sets of distributions ν_i, ν_{ij} that satisfy the natural constraints

$$\sum_{x_i \in \mathcal{X}} \nu_i(x_i) = 1, \ i \in \mathcal{V}, \tag{4.21}$$

$$\sum_{x_i, x_j \in \mathcal{X}} \nu_{ij}(x_i, x_j) = 1, \ (ij) \in \mathcal{E}, \tag{4.22}$$

$$\nu_i(x_i) = \sum_{x_j \in \mathcal{X}} \nu_{ij}(x_i, x_j), \ i \in \mathcal{V}, \ x_i \in \mathcal{X}, \ (ij) \in \mathcal{E}. \tag{4.23}$$

Remark. *1) The number of variables is reduced from $|\mathcal{X}|^{|\mathcal{V}|}$ to $|\mathcal{V}| \cdot |\mathcal{X}| + |\mathcal{E}| \cdot |\mathcal{X}|^2$ when going from Gibbs free energy minimization to Bethe free energy minimization.*

2) In general, for marginal distributions ν_i, ν_{ij} satisfying constraints (4.22), (4.23) *there may not exist any distribution ν on $\mathcal{X}^{\mathcal{V}}$ of which these are the marginals.*

3) The constraints (4.21) *and* (4.22) *are redundant when* (4.23) *holds, but their inclusion facilitates the derivation to follow.*

Minimization of (4.20) under constraints (4.21)–(4.23) can be approached by introducing the Lagrangian

$$\begin{aligned} \mathcal{L}((\nu_i, \nu_{ij}), (\alpha_i, \beta_{ij}, \lambda_{i\to j})) = \ & \mathbb{G}_{\text{Bethe}}(\{\nu_i\}, \{\nu_{ij}\}) + \textstyle\sum_{i \in \mathcal{V}} \alpha_i (\sum_{x_i \in \mathcal{X}} \nu_i(x_i) - 1) \\ & + \textstyle\sum_{(ij) \in \mathcal{E}} \beta_{ij} (\sum_{x_i, x_j \in \mathcal{X}} \nu_{ij}(x_i, x_j) - 1) \\ & + \textstyle\sum_{(i \to j)} \sum_{x_i \in \mathcal{X}} \lambda_{i\to j}(x_i) [\nu_i(x_i) - \sum_{x_j \in \mathcal{X}} \nu_{ij}(x_i, x_j)]. \end{aligned} \tag{4.24}$$

Consider a stationary point of the Lagrangian, i.e. a pair $\{(\nu_i, \nu_{ij}), (\alpha_i, \beta_{ij}, \lambda_{i\to j})\}$ such that partial derivatives of $\mathcal{L}$ with respect to all its variables are zero. The conditions

$$\partial \mathcal{L} / \partial \alpha_i = \partial \mathcal{L} / \partial \beta_{ij} = \partial \mathcal{L} / \partial \lambda_{i\to j}(x_i) = 0$$

state that ν_i and ν_{ij} satisfy the constraints (4.21)–(4.23).

One further has

$$\begin{array}{ll} \frac{\partial \mathcal{L}}{\partial \nu_{ij}(x_i, x_j)} = 0 & \Leftrightarrow -\ln \psi_{ij}(x_i, x_j) + 1 + \ln \nu_{ij}(x_i, x_j) - \lambda_{i\to j}(x_i) - \lambda_{j\to i}(x_j) + \beta_{ij} = 0, \\ \frac{\partial \mathcal{L}}{\partial \nu_i(x_i)} = 0 & \Leftrightarrow -\ln \psi_i(x_i) + (1 - d_i)(1 + \ln \nu_i(x_i)) + \alpha_i + \sum_{j \sim i} \lambda_{i\to j}(x_i) = 0. \end{array}$$

This allows to prove the following

Theorem 4.4. *Assume that $\psi_i(x_i)$ and $\psi_{ij}(x_i, x_j) > 0$ for all $(i, j) \in \mathcal{E}$ and $x_i, x_j \in \mathcal{X}$. There is then a one-to-one correspondence between stationary points of the Lagrangian* (4.24) *associated with Bethe free energy minimization and fixed points $b_{i\to j}(x_i)$ of belief propagation.*

Proof. Consider a stationary point of $\mathcal{L}$. The previous identities obtained by setting to zero the partial derivatives of $\mathcal{L}$ with respect to the primal variables $\nu_i(x_i)$ and $\nu_{ij}(x_i, x_j)$ yield

$$\nu_i(x_i) \propto \exp\left(\tfrac{1}{d_i - 1} [-\ln \psi_i(x_i) + \textstyle\sum_{j \sim i} \lambda_{i\to j}(x_i)] \right), \tag{4.25}$$

$$\nu_{ij}(x_i, x_j) \propto \psi_{ij}(x_i, x_j) \exp(\lambda_{i\to j}(x_i) + \lambda_{j\to i}(x_j)). \tag{4.26}$$

Let us then define

$$b_{i\to j}(x_i) = \exp(\lambda_{i\to j}(x_i)), \quad i \in \mathcal{V}, \ x_i \in \mathcal{X}. \tag{4.27}$$

Constraint (4.23) guarantees that, for all $(ik) \in \mathcal{E}$, as a function of x_i one has

$$\sum_{x_k \in \mathcal{X}} \psi_{ik}(x_i, x_k) b_{k\to i}(x_k) \propto b_{i\to k}(x_i)^{-1} \psi_i(x_i)^{-1/(d_i - 1)} \prod_{\ell \sim i} b_{i\to \ell}(x_i)^{1/(d_i - 1)}.$$

Fix now $j \sim i$, and multiply the above relations for all $k \sim i$, $k \neq j$. This yields

$$\prod_{k\sim i, k\neq j} \sum_{x_k \in \mathcal{X}} \psi_{ik}(x_i, x_k) b_{k\to i}(x_k) \propto \prod_{k\sim i, k\neq j} b_{i\to k}(x_i)^{-1} \psi_i(x_i)^{-1} \prod_{\ell \sim i} b_{i\to \ell}(x_i),$$

which after simplification of the factor $\prod_{k\sim i, k\neq j} b_{i\to k}(x_i)^{-1}$ shows that the corresponding messages $b_{i\to j}(x_i)$ satisfy the belief propagation relation (4.13). It is immediate to verify that the identity (4.26) is equivalent to (4.15) with $\nu_{ij} = \mu_{ij}$. To complete the first implication, note that thanks to the belief propagation relation verified by $b_{i\to j}$, identity (4.26) yields

$$\begin{aligned} \nu_i(x_i) &\propto \psi_i(x_i)^{-1/(d_i-1)} \prod_{j\sim i} \Big(\psi_i(x_i) \prod_{k\sim i, k\neq j} \textstyle\sum_{x_k \in \mathcal{X}} \psi_{ki}(x_k, x_i) \lambda_{k\to i}(x_k) \Big)^{1/(d_i-1)} \\ &= \psi_i(x_i) \prod_{k\sim i} \textstyle\sum_{x_k \in \mathcal{X}} \psi_{ki}(x_k, x_i) \lambda_{k\to i}(x_k), \end{aligned}$$

which is precisely the expression (4.14) when setting $\mu_i = \nu_i$.

Reciprocally, given messages $b_{i\to j}(x_i)$ that satisfy the belief propagation relation (4.13), these must necessarily be strictly positive when functions ψ_i, ψ_{ij} are. We can then define $\lambda_{i\to j}(x_i) := \ln b_{i\to j}(x_i)$. By then letting $\nu_i = \mu_i$, $\nu_{ij} = \mu_{ij}$ with μ_i, μ_{ij} as in (4.14) and (4.15), we again obtain from similar evaluations that necessarily (ν_i, ν_{ij}) constitute a set of primal stationary points of the Lagrangian with associated dual variables $\lambda_{i\to j}$. □

4.6 Notes

The relation between belief propagation and Bethe free energy minimization was identified by Yedidia, Freeman and Weiss [45].

Chapter 5

The tree reconstruction problem

Let $\mathcal{T}$ be a tree graph with root r. We denote by $\mathcal{L}_d$ the set of nodes of generation d, starting with $\mathcal{L}_0 = \{r\}$. we also denote by $\mathcal{T}_d$ the tree down to generation d, by E_d the set of its (unoriented) edges and V_d its vertex set.

Each node i has trait $\sigma_i \in [q]$ inherited from its parent $p(i)$, potentially with errors. Specifically we assume that for all d, conditionally on $\mathcal{T}$ and the spins $\sigma_{V_{d-1}}$ of all nodes of generations $0, \dots, d-1$, nodes i in $\mathcal{L}_d$ have independent spins with distribution

$$\mathbb{P}(\sigma_{\mathcal{L}_d} = s_{\mathcal{L}_d} | \mathcal{T}, \sigma_{V_{d-1}}) = \prod_{i \in \mathcal{L}_d} P_{\sigma_{p(i)} s_i},$$

where P is a stochastic matrix assumed irreducible and with stationary distribution $\{\nu_s\}_{s\in[q]}$. The joint distribution of spins σ_i of all nodes is then fully specified by imposing that σ_r is distributed according to ν. It is easily verified that $\{\sigma_i\}_{i\in V_d}$ is a Markov random field with conditional independence structure prescribed by $\mathcal{T}_d$.

We will use as an illustration the special case where $P_{\tau\tau} = p$, $P_{\tau s} = (1-p)/(q-1)$, $s \neq \tau$, for which ν is uniform. In that particular case, one has for all $s_{V_d} \in [q]^{V_d}$:

$$\mathbb{P}(\sigma_{V_d} = s_{V_d}) \propto \prod_{(i,j)\in E_d} \left[p \mathbf{1}_{s_i = s_j} + \frac{1-p}{q-1} \mathbf{1}_{s_i \neq s_j} \right]. \tag{5.1}$$

This is the so-called symmetric Potts model.

The tree reconstruction problem is then: based on knowledge of $\mathcal{T}_d$, and traits (or spins) σ_i of generation d nodes $i \in \mathcal{L}_d$, can one infer σ_r, the ancestor's trait, non-trivially as $d \to \infty$?

In this chapter we consider Galton-Watson trees conditioned on survival, since in case of extinction the tree reconstruction problem is trivial. An important special case is that of offspring distribution $\mathrm{Poi}(\alpha)$ for fixed $\alpha > 1$. The results in this chapter also admit versions for trees that are not necessarily Galton-Watson, see in particular Evans et al. [14] and Mossel and Peres [35].

5.1 Information theory background

Denote by $H(\nu) := \sum_s \nu_s \ln(1/\nu_s)$ the entropy of a discrete distribution ν, and by $H(X)$ the entropy of the (distribution of a) random variable X.

The **conditional entropy** $H(X|Y)$ of X given Y is by definition

$$\begin{aligned} H(X|Y) &:= \textstyle\sum_{x,y} p_{X,Y}(x,y) \ln\left(\frac{1}{p_{X|Y}(x|y)}\right) \\ &= H(X,Y) - H(Y) \\ &= \textstyle\sum_y p_Y(y) H(\text{Law of } X \text{ conditional on } Y = y). \end{aligned}$$

This last formulation readily extends to the case where the random variable Y is not discrete.

The **mutual information** $I(X;Y)$ between two discrete random variables X, Y is by definition

$$\begin{aligned} I(X;Y) &= H(X) + H(Y) - H(X,Y) \\ &= \sum_{x,y} p_{X,Y}(x,y) \ln\left(\frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)}\right) \\ &= D(p_{X,Y} \| p_X \otimes p_Y). \end{aligned}$$

This last characterization implies, based on properties of Kullback-Leibler divergence, that $I(X;Y)$ is non-negative and equals zero if and only if X and Y are independent.

Given three discrete random variables X, Y, Z, the **conditional mutual information** $I(X;Y|Z)$ is defined as

$$\begin{aligned} I(X;Y|Z) &= H(X|Z) + H(Y|Z) - H(X,Y|Z) \\ &= \sum_z p_Z(z) D(p_{X,Y|Z=z} \| p_{X|Z=z} \otimes p_{Y|Z=z}). \end{aligned}$$

It follows again from properties of the Kullback-Leibler divergence that conditional mutual information $I(X;Y|Z)$ is non-negative, and equals zero if and only if for almost all values z of Z, X and Y are conditionally independent given $Z = z$.

Using this definition, one can obtain by elementary manipulations the following **chain rule** for computing mutual information:

$$I(X;(Y,Z)) = I(X;Z) + I(X;Y|Z).$$

We now recall the following **data processing inequality**:

Lemma 5.1. *Given three discrete random variables X, Y, Z such that X and Z are independent conditionally on Y, it holds that*

$$I(X;Y) \geq I(X;Z).$$

Proof. Compute $I(X;(Y,Z))$ in two ways using the chain rule, to obtain

$$I(X;Y) + I(X;Z|Y) = I(X;Z) + I(X;Y|Z).$$

Conditional independence of X, Z given Y implies that $I(X;Z|Y) = 0$; non-negativity of conditional mutual information $I(X;Y|Z)$ then gives the result. □

5.2 Non-trivial reconstruction

Let $\mathcal{F}_d = \sigma(\mathcal{T}_d, \sigma_{V_d})$, $\mathcal{G}_d = \sigma(\mathcal{T}_d, \sigma_{\mathcal{L}_d})$ and $\hat{\nu}_{s,d} = \mathbb{P}(\sigma_r = s|\mathcal{G}_d)$, $s \in [q]$.

Mutual information between σ_r and $\mathcal{G}_d$ is by definition

$$\begin{aligned} I(\sigma_r;\mathcal{G}_d) &:= H(\sigma_r) - H(\sigma_r|\mathcal{G}_d) \\ &= \mathbb{E}\sum_{s\in[q]} \hat{\nu}_{s,d} \ln(\hat{\nu}_{s,d}/\nu_s) \\ &= \mathbb{E}\sum_{s\in[q]} \mathbb{P}(\sigma_r = s|\mathcal{G}_d) \ln\left(\frac{\mathbb{P}(\sigma_r=s|\mathcal{G}_d)}{\nu_s}\right). \end{aligned}$$

By the data processing inequality, since conditionally on $\mathcal{G}_d$, σ_r and $\mathcal{G}_{d+1}$ are mutually independent (this is true for both Galton-Watson and deterministic trees $\mathcal{T}$), it follows that

$$I(\sigma_r;\mathcal{G}_d) \geq I(\sigma_r;\mathcal{G}_{d+1}).$$

The limit $\lim_{d\to\infty} I(\sigma_r;\mathcal{G}_d)$ therefore exists. We will then say that

Definition 5.1. *Tree reconstruction is feasible if* $\lim_{d\to\infty} I(\sigma_r,\mathcal{G}_d) > 0$.

Recall the following definitions:

Definition 5.2. *Given a filtration $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$, i.e. an increasing family of σ-fields, the family $\{M_d\}_{d\in\mathbb{N}}$ of random variables is a $\mathcal{G}_d$-martingale if for all $d \geq 0$,*

$$\mathbb{E}(M_{d+1}|\mathcal{G}_d) = M_d.$$

Definition 5.3. *A family of random variables X_t indexed by t in some arbitrary set T is uniformly integrable if:*

$$\lim_{A\to+\infty} \sup_{t\in T} \mathbb{E}(|X_t|\mathbf{1}_{|X_t|\geq A}) = 0.$$

Let us now recall two classical results on martingales.

Theorem 5.1. *Given a filtration $\{\mathcal{G}_d\}_{d\in\mathbb{N}}$ and a uniformly integrable $\mathcal{G}_d$-martingale $\{M_d\}_{d\in\mathbb{N}}$, then there exists a $\mathcal{G}_\infty$-measurable and integrable random variable M_∞ such that: $M_d = \mathbb{E}(M_\infty|\mathcal{G}_d)$, and $\lim_{d\to\infty} M_d = M_\infty$, with convergence taking place both almost surely and in $\mathcal{L}^1$.*

Reciprocally, for any integrable, $\mathcal{G}_\infty$-measurable random variable M_∞, the sequence $M_d := \mathbb{E}(M_\infty|\mathcal{G}_d)$ is a uniformly integrable $\mathcal{G}_d$-martingale that converges almost surely and in $\mathcal{L}^1$ to M_∞.

We shall also make use of the following

Theorem 5.2. *Given a decreasing family of σ-fields $\mathcal{H}_d$, for a given integrable random variable X, let $X_d := \mathbb{E}(X|\mathcal{H}_d)$. Then almost surely and in $\mathcal{L}^1$ it holds that $\lim_{d\to\infty} X_d = \mathbb{E}(X|\mathcal{H}_\infty)$.*

Proposition 5.1. *Non-reconstructibilty, i.e. $\lim_{d\to\infty} I(\sigma_r, \mathcal{G}_d) = 0$, is equivalent to: $\hat{\nu}_{s,d} \to \nu_s$ in probability as $d \to \infty$ for all $s \in [q]$.*

Proof. Indeed, let us assume without loss of generality that $\inf_{s\in[q]} \nu_s > 0$ (we can otherwise restrict ourselves to the subset of $\{s \in [q] : \nu_s > 0\}$). On the simplex $M([q])$ of probability distributions μ on $[q]$, the Kullback-Leibler divergence $D(\mu\|\nu)$ is then continuous, non-negative, and equal to zero only at $\mu = \nu$.

Assuming non-reconstructibility, then for all $\epsilon > 0$, the probability that $\{\hat{\nu}_{s,d}\}_{s\in[q]}$ belongs to the set of probability measures μ such that $D(\mu\|\nu) \geq \epsilon)$ must go to zero. The fact that ν is the unique minimizer of $D(\cdot\|\nu)$ on $M([q])$ then implies that $\hat{\nu}_d$ must converge in probability to ν.

Conversely, if $\hat{\nu}_d$ converges in probability to ν as $d \to \infty$, for any $\epsilon > 0$, and d large enough, the probability that $D(\hat{\nu}_d\|\nu) \leq \epsilon$ must be at least $1 - \epsilon$. On the complementary event, one has the deterministic upper bound

$$D(\hat{\nu}_d\|\nu) \leq C := \ln(q) - \ln(\inf_{s\in[q]}(\nu_s)).$$

This gives $\limsup_{d\to\infty} I(\sigma_r; \mathcal{G}_d) \leq \epsilon + C\epsilon$, hence the result. □

Remark. *We can interpret the reconstruction property as follows. Given some loss function $L(s, s')$, when asked to produce an estimate $\hat{s}$ of σ_r based on observation $\mathcal{G}_d$, to minimize average loss $\mathbb{E}L(\sigma_r, \hat{s})$, if $I(\sigma_r, \mathcal{G}_d) \to 0$ as $d \to \infty$, a deterministic rule $\hat{s} \in argmin_{s'\in[q]} \sum_s \nu_s L(s, s')$ cannot be beaten asymptotically. In the converse situation $\lim_{d\to\infty} I(\sigma_r, \mathcal{G}_d) > 0$, for non-trivial loss functions, the rule*

$$\hat{s} \in argmin_{s'\in[q]} \sum_s \hat{\nu}_{s,d} L(s, s')$$

can achieve strictly better performance.

In the particular case $L(s, s') = \mathbf{1}_{s\neq s'}$, the optimal loss is achieved by taking $\hat{s} \in argmax(\hat{\nu}_{s,d})$, and achieves average loss $\mathbb{E}(1 - \sup_{s\in[q]} \hat{\nu}_{s,d})$. Assuming further that ν is uniform on $[q]$, it follows that the limiting loss (which must exist, because $\hat{\nu}_{s,d}$ is a martingale) is strictly less than $(1 - 1/q)$ if and only tree reconstruction is feasible. Indeed, $\sup_{s\in[q]} \hat{\nu}_{s,d} \geq 1/q$, and the asymptotic loss is $1 - 1/q$ if and only if $\sup_s \hat{\nu}_s \to 1/q$ in probability, which is equivalent by Proposition 5.1 to non-reconstructibility.

5.3 Census reconstruction above the Kesten-Stigum threshold

Let $X_{s,d}$ denote the number of individuals at generation d with trait s:

$$X_{s,d} = \sum_{i \in \mathcal{L}_d} \mathbf{1}_{\sigma_i = s}.$$

We also refer to the vector $X_d := \{X_{s,d}\}_{s \in [q]}$ as the *census* of generation d. We will then consider the following

Definition 5.4. *Census reconstruction is feasible if* $\lim_{d\to\infty} I(\sigma_r; X_d) > 0$.

Remark. *The data processing inequality guarantees that the limit* $\lim_{d\to\infty} I(\sigma_r; X_d)$ *exists whenever* $\mathcal{T}$ *is either a deterministic or a Galton-Watson branching tree.*

Still by the data processing inequality, census reconstruction implies tree reconstruction.

When $\mathcal{T}$ is a Galton-Watson tree with offspring distribution $\mathrm{Poi}(\alpha)$, $\{X_d\}_{d\geq 0}$ is a multi-type branching process where a type τ-individual has $\mathrm{Poi}(\alpha P_{\tau s})$ children of type s, independently over $s \in [q]$. This is a special case of the family of multi-type Galton-Watson branching processes studied by Kesten and Stigum in [22].

Let us order the eigenvalues of P in decreasing order of absolute value, and denote them by λ_s, $s \in [q]$: $\lambda_1 = 1 > |\lambda_2| \geq \cdots \geq |\lambda_q|$. We then have the following

Theorem 5.3. *Assume* $\alpha|\lambda_2|^2 > 1$*, and let* $\mathcal{T}$ *be a Galton-Watson branching tree with offspring of mean* $\alpha > 1$ *and bounded second moment. Then census (and hence tree) reconstruction holds.*

Let $\{x_s\}_{s\in[q]}$ *denote a (non-constant) eigenvector associated with eigenvalue* λ_2 *of* P*. Construct from the census* X_d *the random variable*

$$Z_d := (\alpha \cdot \lambda_2)^{-d} \sum_{s \in [q]} X_{s,d} x_s.$$

Then Z_d *contains non-trivial information on* σ_r *in the sense that there exists some* $t \in \mathbb{R}$ *such that* $\lim_{d\to\infty} I(\sigma_r; \mathbf{1}_{Z_d \leq t}) > 0$[1].

Proof. The second result implies the first since Z_d, and hence $\mathbf{1}_{Z_d\leq t}$, are functions of the census X_d, and $\mathcal{G}_d$-measurable, and the data-processing inequality then ensures that $I(\sigma_r; \mathcal{G}_d) \geq I(\sigma_r; X_d) \geq I(\sigma_r; \mathbf{1}_{Z_d\leq t})$.

Now, Z_d is a uniformly integrable $\mathcal{F}_d$-martingale such that $\mathbb{E}(Z_d|\mathcal{F}_0) = x_{\sigma_r}$. Indeed:

$$\begin{aligned}
\mathbb{E}(Z_d|\mathcal{F}_{d-1}) &= (\alpha\cdot\lambda_2)^{-d}\mathbb{E}\left[\textstyle\sum_{i\in\mathcal{L}_{d-1}}\sum_{j:p(j)=i} x_{\sigma_j} \mid \mathcal{F}_{d-1}\right] \\
&= (\alpha\cdot\lambda_2)^{-d}\textstyle\sum_{i\in\mathcal{L}_{d-1}} \alpha \sum_{s\in[q]} P_{\sigma_i s} x_s \\
&= Z_{d-1},
\end{aligned}$$

because x is an eigenvector of P associated to λ_2. To establish uniform integrability, we use the conditional variance formula to show that the second moment of Z_d is bounded:

$$\begin{aligned}
\mathrm{Var}(Z_d) &= \mathrm{Var}(\mathbb{E}(Z_d|\mathcal{F}_{d-1})) + \mathbb{E}(\mathrm{Var}(Z_d|\mathcal{F}_{d-1})) \\
&= \mathrm{Var}(Z_{d-1}) + |\alpha\cdot\lambda_2|^{-2d}\mathbb{E}(\mathrm{Var}(\textstyle\sum_{i\in\mathcal{L}_d} x_{\sigma_i}|\mathcal{F}_{d-1})) \\
&= \mathrm{Var}\, Z_{d-1} + |\alpha\cdot\lambda_2|^{-2d}\mathbb{E}(\textstyle\sum_{i\in\mathcal{L}_{d-1}} \mathrm{Var}(\sum_{j:p(j)=i} x_{\sigma_j}|\mathcal{F}_{d-1})) \\
&\leq \mathrm{Var}\, Z_{d-1} + C(\alpha\cdot\lambda_2)^{-2d}\mathbb{E}|\mathcal{L}_{d-1}|,
\end{aligned}$$

where $C := \sup_{\tau\in[q]} \mathrm{Var}(\sum_{j:p(j)=i} x_{\sigma_j}|\mathcal{F}_{d-1}; \sigma_i = \tau)$ is finite by the assumption that offspring distribution has finite second moment. Since $\mathbb{E}|\mathcal{L}_d| = \alpha^d$, it follows that

$$\begin{aligned}
\mathrm{Var}(Z_d) &\leq \mathrm{Var}\, Z_{d-1} + C(\alpha|\lambda_2|^2)^{-d} \\
&\leq \mathrm{Var}(Z_0) + \textstyle\sum_{k=1}^{d} C(\alpha|\lambda_2|^2)^{-k} \\
&\leq \mathrm{Var}(Z_0) + C\frac{1}{\alpha|\lambda_2|^2-1},
\end{aligned}$$

[1] If λ_2 and x are complex, then this statement should be modified to: there exists some $t \in \mathbb{R}$ such that either $\lim_{d\to\infty} I(\sigma_r; \mathbf{1}_{\Re(Z_d)\leq t}) > 0$, or $\lim_{d\to\infty} I(\sigma_r; \mathbf{1}_{\Im(Z_d)\leq t}) > 0$.

by the assumption that $\alpha|\lambda_2|^2 > 1$.

By the martingale convergence theorem, Z_d converges almost surely and in $\mathcal{L}^1$ to Z_∞ such that $\mathbb{E}(Z_\infty|\mathcal{F}_d) = Z_d$. In particular, $\mathbb{E}(Z_\infty|\mathcal{F}_0] = x_{\sigma_r}$.

Recall that the coordinates $(x_s)_{s\in[q]}$ are not all equal. We could assume this since the constant vector is an eigenvector of P associated with eigenvalue 1.

Assume that for (Lebesgue almost) all $t \in \mathbb{R}$, $\mathbf{1}_{Z_\infty \leq t}$ is independent of σ_r. Thus for Lebesgue almost all $t \in \mathbb{R}$, all $s \in [q]$,

$$\mathbb{P}(Z_\infty \leq t|\sigma_r = s) = \mathbb{P}(Z_\infty \leq t).$$

However, this implies that $\mathbb{E}(Z_\infty|\sigma_r = s)$ is independent of s, which is impossible since it equals x_s, which cannot be constant in s. Thus there exists a set of positive Lebesgue measure of t's such that $\mathbf{1}_{Z_\infty\leq t}$ and σ_r are correlated. Choose one that is a continuity point of all conditional distributions $\mathbb{P}(Z_\infty \in \cdot|\sigma_r = s)$, $s \in [q]$. Then almost surely for all $s \in [q]$,

$$\lim_{d\to\infty} \mathbb{P}(\sigma_r = s|Z_d \leq t) = \mathbb{P}(\sigma_r = s|Z_\infty \leq t).$$

The claim follows, because the latter property implies that $\lim_{d\to\infty} I(\sigma_r; \mathbf{1}_{Z_d\leq t}) = I(\sigma_r; \mathbf{1}_{Z_\infty\leq t}) > 0$. □

Another question of interest is: can one obtain non-trivial information about σ_r if the spins σ_i for $i \in \mathcal{L}_d$ have been slightly distorted? Specifically, assume that each such spin σ_i is probabilistically modified into $\sigma'_i \in [q]$, using some Markov transition kernel $M_{s\tau} = \mathbb{P}(\sigma'_i = \tau|\sigma_i = s)$. We pose the following

Definition 5.5. *Let μ be a distribution on $[q]$. Robust distribution holds with respect to μ if for any $\epsilon > 0$, and Markov transition kernel M given by $M_{s\tau} = \epsilon\mathbf{1}_{s=\tau} + (1-\epsilon)\mu_\tau$, it holds that $\liminf_{d\to\infty} I(\sigma_r; (\mathcal{T}, \sigma'_{\mathcal{L}_d})) > 0$.*

We then have the following

Theorem 5.4. *Let $\mathcal{T}$ be a Galton-Watson tree with offspring distribution of mean $\alpha > 1$ and bounded second moment. Above the Kesten-Stigum threshold, i.e. when $\alpha|\lambda_2|^2 > 1$, robust reconstruction holds, and can be performed from the perturbed census $X'_{s,d} := \sum_{i\in\mathcal{L}_d} \mathbf{1}_{\sigma'_i=s}$. Specifically, let $y = M^{-1}x$, where x is the eigenvector of P associated with λ_2. Let*

$$Z'_d := (\alpha\cdot\lambda_2)^{-d} \sum_{i\in\mathcal{L}_d} y_{\sigma'_i}.$$

Then there is some $t \in \mathbb{R}$ such that

$$\lim_{d\to\infty} I(\sigma_r, \mathbf{1}_{Z'_d\leq t}) > 0.$$

Proof. Vector y is well-defined since M is clearly invertible, for arbitrary $\epsilon > 0$ and distribution μ. By definition of Z'_d and our choice of y, one has

$$\begin{aligned}\mathbb{E}(Z'_d|\mathcal{F}_{d-1}) &= |\alpha\cdot\lambda_2|^{-d} \textstyle\sum_{i\in\mathcal{L}_{d-1}} \alpha \sum_{s\in[q]} \sum_{\tau\in[q]} P_{\sigma_i s} M_{s\tau} y_\tau \\ &= Z_{d-1}.\end{aligned}$$

Moreover, one has:

$$\begin{aligned}\mathrm{Var}(Z'_d|\mathcal{F}_{d-1}) &= |\alpha\cdot\lambda_2|^{-2d} \textstyle\sum_{i\in\mathcal{L}_{d-1}} \mathrm{Var}(\sum_{j:p(j)=i} y_{\sigma'_j}|\mathcal{F}_{d-1}) \\ &\leq C(\alpha|\lambda_2|^2)^{-d}\alpha^{-d}|\mathcal{L}_{d-1}|,\end{aligned}$$

for some suitable constant $C > 0$. However, $\alpha^{-d}|\mathcal{L}_{d-1}|$ is easily seen to be a uniformly integrable martingale when $\alpha > 1$, and thus $\mathbb{E}[(Z'_d - Z_{d-1})^2|\mathcal{F}_{d-1}]$ converges to zero almost surely as $d \to \infty$. Thus the conditional distributions $\mathbb{P}(Z'_d \in \cdot|\sigma_r = s)$ and $\mathbb{P}(Z_d \in \cdot|\sigma_r = s)$ admit the same weak limits for all $s \in [q]$. The result follows. □

5.4 Below the Kesten-Stigum threshold

We now establish the following

Theorem 5.5. *Consider a Galton-Watson branching tree with Poisson offspring distribution of mean $\alpha > 1$. Assume that $\alpha|\lambda_2|^2 < 1$. Then census reconstruction fails, i.e. $X_d := \{X_{s,d}\}_{s\in[q]}$ is such that* $\lim_{d\to\infty} I(\sigma_r; X_d) = 0$.

While we consider here only the case of Poisson offspring, the result admits extensions to other branching trees. In particular, the more delicate case of b-ary trees is handled in Mossel and Peres [35].

Proof. When $\alpha|\lambda_2|^2 < 1$, the results of Kesten and Stigum [22] show that conditionally on $\sigma_r = \tau$ and on survival of the branching process, $\{\alpha^{-d/2}(X_{s,d} - \alpha^d \nu_s)\}_{s\in[q]}$ converges in distribution to a q-dimensional Gaussian vector whose distribution does not depend on the initial value τ of σ_r.

Survival occurs with a probability that does not depend on σ_r in our model.

It then follows that for two initial conditions τ, τ' in $[q]$, we can generate coupled corresponding vectors $X_d^{(\tau)}$, $X_d^{(\tau')}$ such that for all $\epsilon > 0$, $\lim_{d\to\infty} \mathbb{P}(|X_d^{(\tau)} - X_d^{(\tau')}| \geq \epsilon\alpha^{d/2}) = 0$.

The random variables $X_{s,d+1}^{(\tau)}$ and $X_{s,d+1}^{(\tau')}$ admit conditionally on $\mathcal{F}_d$ the distributions

$$X_{s,d+1}^{(t)} \sim \text{Poi}\left(\alpha \sum_{s'\in[q]} X_{s',d}^{(t)} P_{s's}\right),\ t \in \{\tau, \tau'\},$$

with independence for distinct s.

Let $M_s^{(t)} = \mathbb{E}(X_{s,d+1}^{(t)}|\mathcal{F}_d)$, $M_s = \frac{1}{2}(M_s^{(\tau)} + M_s^{(\tau')})$, and

$$\epsilon_s = \frac{1}{2}|M_s^{(\tau)} - M_s^{(\tau')}|M_s^{-1/2}.$$

The coupling result entails the existence of a function α_d going to zero as $d \to \infty$ such that with high probability, $\epsilon_s \leq \alpha_d$. Conditionally on $\mathcal{F}_d$, the variation distance between $X_{d+1}^{(\tau)}$ and $X_{d+1}^{(\tau')}$ is upper-bounded by

$$\sum_{s\in[q]} \sum_{k\geq 0} \left| \frac{e^{-M_s - \epsilon_s\sqrt{M_s}}(M_s + \epsilon_s\sqrt{M_s})^k - e^{-M_s + \epsilon_s\sqrt{M_s}}(M_s - \epsilon_s\sqrt{M_s})^k}{k!} \right|.$$

For fixed s we split the corresponding sum over k into two sums, according to whether $|M_s - k| \leq \omega_d\sqrt{M_s}$ or not, where $\omega_d = 1/\sqrt{\alpha_d}$. The summation over k such that $|M_s - k| > \omega_d\sqrt{M_s}$ is upper bounded by the sum of the two $\pm$ terms

$$\mathbb{P}(|\text{Poi}(M_s \pm \epsilon_s\sqrt{M_s}) - M_s| \geq \omega_d\sqrt{M_s}),$$

which must go to zero on $\{\epsilon_s \leq \alpha_d\}$. The summation over k such that $|M_s - k| \leq \omega_d\sqrt{M_s}$ is upper bounded by

$$\sum_{|k-M_s|\leq\omega_d\sqrt{M_s}} e^{-M_s}\frac{M_s^k}{k!} \left| e^{-\epsilon_s\sqrt{M_s}}(1 + \epsilon_s/\sqrt{M_s})^k - e^{\epsilon_s\sqrt{M_s}}(1 - \epsilon_s/\sqrt{M_s})^k \right|$$

We have the equivalent

$$\begin{aligned} e^{\pm\epsilon_s\sqrt{M_s}}(1 \mp \epsilon_s/\sqrt{M_s})^k &= e^{\pm\epsilon_s\sqrt{M_s} + k(\mp\epsilon_s/\sqrt{M_s} + O(\epsilon_s^2/M_s))} \\ &= e^{O(\epsilon_s\omega_d)} \\ &= 1 + O(\sqrt{\alpha_d}) \end{aligned}$$

on the event that $\epsilon_s \leq \alpha_d$. Thus the second summation is, with high probability, upper bounded by $O(\sqrt{\alpha_d}) = o(1)$. this implies that the variation distance between $X_d^{(\tau)}$ and $X_d^{(\tau')}$ goes to zero as $d \to \infty$.

This implies in turn that $I(\sigma_r; X_d)$ goes to zero. For instance, this can be seen by letting $f_s(x) := \mathbb{P}(X_d = x|\sigma_r = s)/\mathbb{P}(X_d = x)$ for $x \in \mathbb{N}^q$. One then has

$$\begin{aligned} I(\sigma_r; X_d) &= \sum_{s\in[q],x\in\mathbb{N}^q} \nu_s \mathbb{P}(X_d = x|\sigma_r = s) \ln\left(\frac{\mathbb{P}(X_d=x|\sigma_r=s)}{\mathbb{P}(X_d=x)}\right) \\ &= \sum_{x\in\mathbb{N}^q} \mathbb{P}(X_d = x) \sum_{s\in[q]} \nu_s f_s(x) \ln(f_s(x)) \\ &\leq \sum_{x\in\mathbb{N}^q} \mathbb{P}(X_d = x) \sum_{s\in[q]} \nu_s f_s(x)[f_s(x) - 1]. \end{aligned}$$

However $\nu_s f_s(x) \leq 1$ so that the last term is upper bounded by

$$\sum_{s\in[q]} \sum_{x\in\mathbb{N}^q} \mathbb{P}(X_d = x)|f_s(x) - 1|.$$

Note now that, because $\sum_{\tau\in[q]} \nu_\tau f_\tau(x) = 1$:

$$\begin{aligned} \sum_{x\in\mathbb{N}^q} \mathbb{P}(X_d = x)|f_s(x) - 1| &\leq \sum_{\tau\in[q]} \nu_\tau \sum_{x\in\mathbb{N}^q} \mathbb{P}(X_d = x)|f_s(x) - f_\tau(x)| \\ &\leq \sup_{s,\tau\in[q]} d_{var}(\mathbb{P}(X_d \in \cdot|\sigma_r = s), \mathbb{P}(X_d \in \cdot|\sigma_r = \tau)) \end{aligned}$$

to obtain

$$I(\sigma_r; X_d) \leq q \sup_{s,\tau\in[q]} d_{var}(\mathbb{P}(X_d \in \cdot|\sigma_r = s), \mathbb{P}(X_d \in \cdot|\sigma_r = \tau)).$$

□

Remark. *Janson and Mossel [19] showed that below the KS threshold, for bounded degree trees, for any measure μ that is non-degenerate (i.e. $\inf_{s\in[q]} \mu_s > 0$), robust reconstruction with respect to μ is impossible.*

5.5 Sufficient condition for non-reconstruction for two symmetric communities

Consider the case where $q = 2$, for which it will be convenient to denote the two traits by $+1$ and -1, or $+$ and $-$. Assume further a symmetric transition matrix P, with for some fixed $\epsilon \in (0, 1)$, $P_{++} = P_{--} = 1 - \epsilon$, and $P_{+-} = P_{-+} = \epsilon$. In that case, $\lambda_2 = 1 - 2\epsilon$, and we have the following

Theorem 5.6. *For two-type symmetric propagation on a deterministic tree $\mathcal{T}$ such that*

$$\limsup_{d\to\infty} \frac{1}{d} \ln(|\mathcal{L}_d|) \leq \ln(\alpha), \tag{5.2}$$

tree reconstruction fails when $(\lambda_2)^2\alpha < 1$.

To prove Theorem 5.6, we follow [14]. We need the two following lemmas.

Lemma 5.2. *Consider the trees $\mathcal{T}$, $\mathcal{T}'$ depicted on Figure 5.1, where node variables are binary spins, each uniformly distributed with values ± 1, edge weights are in $[0, 1]$ and represent transmission probability, e.g. $\mathbb{P}(\tau_1 = \sigma_r) = (1 + \theta)/2$ or equivalently $\mathbb{E}(\sigma_r\tau_1) = \theta$.*

Then there exists a probability transition matrix $M : \{-1, 1\}^2 \to \{-1, 1\}^2$ such that

$$\mathbb{P}(\sigma_{r'} = s_r, \sigma_{1'} = s_1, \sigma_{2'} = s_2) = \sum_{u_1,u_2=\pm} \mathbb{P}(\sigma_r = s_r, \sigma_1 = u_1, \sigma_2 = u_2) M_{(u_1,u_2),(s_1,s_2)}.$$

Proof. To be completed □

...

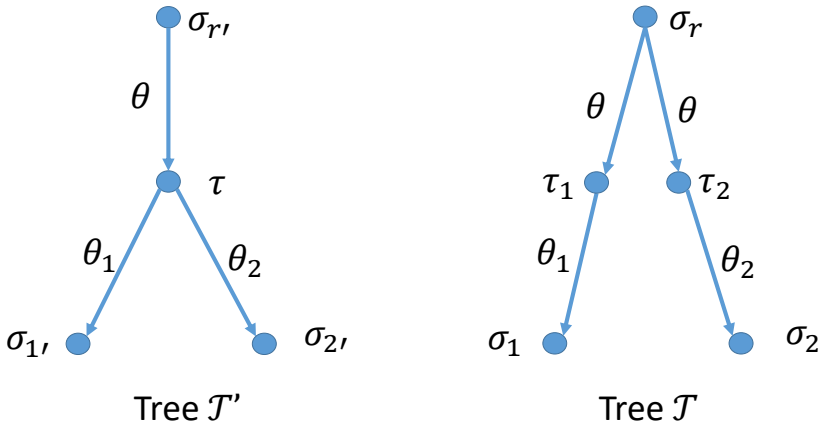

Figure 5.1: Two trees for propagation of binary spins

Lemma 5.3. *Consider two random vectors $U \in \{\pm 1\}^a$, $V \in \{\pm 1\}^b$, that are mutually independent and independent of the spins of the two trees on Figure 5.1. Let $X = \sigma_1 U$, $Y = \sigma_2 V$, $X' = \sigma_{1'} U$, $Y' = \sigma_{2'} V$. Then there is a probability transition matrix M on $\{\pm 1\}^{a+b}$ such that the joint law of $(\sigma_{r'}, (X', Y'))$ is obtained as the joint law of σ_r and the image of (X, Y) by M.*

Proof. Reduce the problem to that of the previous lemma: for each set of 4 states $\{\pm x, \pm y\}$ in $\{\pm 1\}^{a+b}$, use the previous channel... □

Corollary 5.1. *For the binary symmetric transmission of spins on fixed tree $\mathcal{T}$, one has*

$$I(\sigma_r; \sigma_{\mathcal{L}_d}) \leq \sum_{j \in \mathcal{L}_d} I(\sigma_r; \sigma_j).$$

The proof is by induction, and uses the fact that the mutual information $I(X; (Y_1, \ldots, Y_n))$ is, provided the Y_i are independent conditionally on X, upper-bounded by $\sum_{i=1}^n I(X; Y_i)$. It alternates between splitting the tree and applying the previous corollary.

Proof. (of Theorem 5.6). For some node $i \in \mathcal{L}_d$, one has again a symmetric channel between σ_r and σ_i characterized by $\mathbb{E}(\sigma_r \sigma_i) = (1-2\epsilon)^d = \lambda_2^d$ or equivalently, by $\mathbb{P}(\sigma_i = \sigma_r) = [1 + \lambda_2^d]/2$. Thus, letting $\theta_d := \lambda_2^d$,

$$\begin{array}{rl} I(\sigma_r; \sigma_i) & = \sum_{s,t=\pm} \frac{1}{2} \frac{1+st\theta_d}{2} \ln(1 + st\theta_d) \\ & \leq \sum_{s,t=\pm} \frac{1}{2} \frac{1+st\theta_d}{2} (st\theta_d)) \\ & = \theta_d^2. \end{array}$$

Thus by Corollary 5.1, one has

$$I(\sigma_r; \sigma_{\mathcal{L}_d}) \leq |\mathcal{L}_d| \lambda_2^{2d}, \tag{5.3}$$

and under the assumption (5.2), the right-hand side of the above is upper-bounded by $\leq (\alpha \lambda_2^2 + o(1))^d$, hence goes to zero as $d \to \infty$. This concludes the proof. □

Corollary 5.2. *For two-type symmetric propagation on a Galton-Watson tree with mean number of children α, reconstruction fails when $(\lambda_2)^2 \alpha < 1$.*

The corollary in fact follows from the previous bound (5.3) on the mutual information $I(\sigma_r;\sigma_{\mathcal{L}_d})$ for a deterministic tree, which yields here:

$$\begin{aligned} I(\sigma_r;(\sigma_{\mathcal{L}_d},\mathcal{T}_d)) &= I(\sigma_r;\mathcal{T}_d) + I(\sigma_r;\sigma_{\mathcal{L}_d}|\mathcal{T}_d) \\ &\leq 0 + \mathbb{E}((\min(1,|\mathcal{L}_d|\lambda_2^{2d})). \end{aligned}$$

By Markov's inequality, the probability that $|\mathcal{L}_d| \geq \alpha+\delta$ is upper-bounded by $(1+\delta/\alpha)^{-d}$. Picking $\delta>0$ such that $(\alpha+\delta)\lambda_2^2<1$ then implies that $\lim_{d\to\infty} I(\sigma_r;(\sigma_{\mathcal{L}_d},\mathcal{T}_d))=0$.

5.6 Optimal inference and belief propagation

The conditional distribution $\hat{\nu}_{s,d}=\mathbb{P}(\sigma_r=s|\mathcal{G}_d)$ can be computed recursively, following the so-called Belief Propagation (or sum-product) algorithm described in the previous chapter. We rederive it in the context of the tree reconstruction problem. We need the following notation:

$$\nu_s^i := \mathbb{P}(\sigma_i=s|\mathcal{G}_{i,d}),\ i\in V_d,\ s\in[q],$$

where $\mathcal{G}_{i,d}$ is the σ-field generated by $\mathcal{T}_d$ and the spins σ_j of all nodes $j\in\mathcal{L}_d$ that admit i in their ancestry line to the root r. We denote by $\mathcal{L}_{i,d}$ the corresponding set of nodes of $\mathcal{L}_d$.

Let us fix some values $s_{\mathcal{L}_{i,d}}$ for the spins $\sigma_{\mathcal{L}_{i,d}}$.

Write then, introducing the notation $\mathcal{S}(i)$ to denote the set of nodes j such that $p(i)=j$:

$$\begin{aligned} \mathbb{P}(\sigma_i=s|\sigma_{\mathcal{L}_{i,d}}=s_{\mathcal{L}_{i,d}}) &\propto \nu_s \mathbb{P}(\sigma_{\mathcal{L}_{i,d}}=s_{\mathcal{L}_{i,d}}|\sigma_i=s) \\ &= \nu_s \textstyle\sum_{s_j\in[q],j\in\mathcal{S}(i)} \mathbb{P}(\sigma_{\mathcal{L}_{i,d}}=s_{\mathcal{L}_{i,d}},\sigma_{\mathcal{S}(i)}=s_{\mathcal{S}(i)}|\sigma_i=s) \\ &= \nu_s \textstyle\sum_{s_j\in[q],j\in\mathcal{S}(i)} \prod_{j\in\mathcal{S}(i)} \mathbb{P}(\sigma_{\mathcal{L}_{j,d}}=s_{\mathcal{L}_{j,d}},\sigma_j=s_j|\sigma_i=s) \\ &= \nu_s \textstyle\prod_{j\in\mathcal{S}(i)} \sum_{s_j\in[q]} \mathbb{P}(\sigma_{\mathcal{L}_{j,d}}=s_{\mathcal{L}_{j,d}}|\sigma_j=s_j)\mathbb{P}(\sigma_j=s_j|\sigma_i=s) \\ &\propto \nu_s \textstyle\prod_{j\in\mathcal{S}(i)} \sum_{s_j\in[q]} \frac{\mathbb{P}(\sigma_j=s_j|\sigma_{\mathcal{L}_{j,d}}=s_{\mathcal{L}_{j,d}})}{\nu_{s_j}} P_{ss_j}, \end{aligned}$$

where we used Bayes' formula and conditional independence. This computation proves the following recursive formula for ν^i:

$$\nu_s^i = \frac{1}{Z^i}\nu_s \prod_{j\in\mathcal{S}(i)} \sum_{s_j\in[q]} \frac{\nu_{s_j}^j}{\nu_{s_j}} P_{ss_j}, \tag{5.4}$$

where the normalization constant Z^i is such that $\sum_{s\in[q]}\nu_s^i=1$.

Remark that when $\nu_s^j\equiv\nu_s$, $j\in\mathcal{S}(i)$, $s\in[q]$, one obtains $\nu_s^i=\nu_s$. In other words, the distribution ν is a trivial fixed point of the Belief Propagation recursion.

Besides providing an algorithm for determining the conditional distribution of a node's spin, it also provides a basis for analyzing the feasibility of tree reconstruction.

Let us denote by p_k the probability that a node has k children, our primary example corresponding to Poisson offspring, $p_k=e^{-\alpha}\alpha^k/k!$. Let us also define the mapping

$$\begin{aligned} F_k:\ M([q])^k &\to M([q]) \\ \eta_1,\dots,\eta_k &\to \left\{ \tfrac{1}{Z_k(\eta_1,\dots,\eta_k)}\nu_s \textstyle\prod_{j=1}^k \sum_{s_j\in[q]} \frac{\eta_j(s_j)}{\nu_{s_j}} P_{ss_j} \right\}_{s\in[q]}, \end{aligned} \tag{5.5}$$

where $Z_k(\eta_1,\dots,\eta_k)$ is the normalization constant.

Denote by $Q_{\tau,d}$ the probability distribution of the random vector $\{\mathbb{P}(\sigma_r=s|\mathcal{G}_d)\}_{s\in[q]}$ in the simplex $M([q])$ of probability distributions on $[q]$, conditionally on $\sigma_r=\tau$. (5.4) then implies the following recursive characterization of $Q_{\tau,d}$, where ϕ is any continuous function on $M([q])$:

$$\int_{M([q])} \phi(\eta) Q_{\tau,d+1}(d\eta) = \sum_{k\geq 0} p_k \int_{M([q])^k} \phi(F_k(\eta_1,\dots,\eta_k)) \prod_{\ell=1}^k \sum_{s_\ell\in[q]} P_{\tau s_\ell} Q_{s_\ell,d}(d\eta_\ell). \tag{5.6}$$

This stochastic recursive equation is known as the **density evolution** equation (see Mézard and Montanari [37]).

Consider now the law $\hat{Q}_d$ on $M([q])$ of the random distribution $\mathbb{P}(\sigma_r \in \cdot | \mathcal{G}_d)$, no longer conditioned on the exact value of σ_r. Writing this random, $\mathcal{G}_d$-measurable distribution as $\hat{\nu}_d$, observe that

$$\begin{aligned}\textstyle\int_{M([q])} \phi(\eta) Q_{\tau,d}(d\eta) &= \mathbb{E}[\phi(\hat{\nu}_d)|\sigma_r = \tau] \\ &= \tfrac{1}{\nu_\tau}\mathbb{E}[\phi(\hat{\nu}_d)\mathbf{1}_{\sigma_r=\tau}] \\ &= \tfrac{1}{\nu_\tau}\mathbb{E}[\phi(\hat{\nu}_d)\hat{\nu}_d(\tau)] \\ &= \textstyle\int_{M([q])} \phi(\eta)\frac{\eta(\tau)}{\nu_\tau}\hat{Q}(d\eta),\end{aligned}$$

or put differently,

$$Q_{\tau,d}(d\eta) = \frac{\eta(\tau)}{\nu_\tau}\hat{Q}_d(d\eta). \tag{5.7}$$

This allows to recover the simple identity

$$\hat{Q}_d = \sum_{\tau\in[q]} \nu_\tau Q_{\tau,d}.$$

Using these we may re-express (5.6) in terms of the unconditional distributions $\hat{Q}_d$ only, as:

$$\int_{M([q])} \phi(\eta)\hat{Q}_{d+1}(d\eta) = \sum_{\tau\in[q]}\nu_\tau \sum_{k\geq 0} p_k \int_{M([q])^k} \phi(F_k(\eta_1,\dots,\eta_k)) \prod_{\ell=1}^k \sum_{s_\ell\in[q]} P_{\tau s_\ell}\frac{\eta_\ell(s_\ell)}{\nu_{s_\ell}}\hat{Q}_d(d\eta_\ell). \tag{5.8}$$

Note that necessarily, for all d, $\int_{M([q])}\eta\hat{Q}_d(d\eta) = \nu$. We shall call fixed points $\hat{Q}$ of (5.8) that verify this condition **consistent fixed points**. We then have the following

Lemma 5.4. *Any fixed point distribution $\hat{Q}$ of* (5.8) *is consistent.*

Proof. Given a fixed point $\hat{Q}$ of (5.8), let for any $s\in[q]$, $m_s := \int_{M([q])}\eta(s)\hat{Q}(d\eta)$. Taking $\phi(\eta)=\eta(s)$ in (5.8) gives:

$$\begin{aligned} m_s &= \sum_{\tau\in[q]}\nu_\tau\sum_{k\geq 0}p_k\int_{M([q])^k}\frac{\nu_s\prod_{j=1}^k\sum_{t_j\in[q]}\frac{\eta_j(t_j)}{\nu_{t_j}}P_{st_j}}{Z_k(\eta_1,\dots,\eta_k)}\prod_{\ell=1}^k\sum_{s_\ell\in[q]}P_{\tau s_\ell}\frac{\eta_\ell(s_\ell)}{\nu_{s_\ell}}\hat{Q}(d\eta_\ell) \\ &= \sum_{k\geq 0}p_k\int_{M([q])^k}\nu_s\prod_{j=1}^k\sum_{t_j\in[q]}\frac{\eta_j(t_j)}{\nu_{t_j}}P_{st_j}\hat{Q}(d\eta_j) \\ &= \nu_s\sum_{k\geq 0}p_k\Big(\sum_{t\in[q]}\frac{m_t}{\nu_t}P_{st}\Big)^k. \end{aligned} \tag{5.9}$$

Summing this identity over $s\in[q]$ gives

$$1 = \sum_{s\in[q]}\nu_s\sum_{k\geq 0}p_k\Big(\sum_{t\in[q]}\frac{m_t}{\nu_t}P_{st}\Big)^k.$$

By convexity of the function $x\to x^k$ on $\mathbb{R}_+$ for all $k\in\mathbb{N}$, it follows from Jensen's inequality that

$$\begin{aligned}\textstyle\sum_{s\in[q]}\nu_s\big(\sum_{t\in[q]}\frac{m_t}{\nu_t}P_{st}\big)^k &\geq \big(\textstyle\sum_{s\in[q]}\nu_s\sum_{t\in[q]}\frac{m_t}{\nu_t}P_{st}\big)^k \\ &= \big(\textstyle\sum_{t\in[q]}m_t\sum_{s\in[q]}\nu_s P_{st}\frac{1}{\nu_t}\big)^k \\ &= 1,\end{aligned}$$

where in the last step we used the fact that ν is invariant for P. We must therefore have equality in Jensen's inequality for all $k\geq 0$ such that $p_k>0$. Since for some $k>1$, we must have $p_k>0$ (otherwise tree reconstruction is uninteresting), and the corresponding function $k\to x^k$ is then strictly convex, this implies that $s\to\sum_{t\in[q]}\frac{m_t}{\nu_t}P_{st}$ must be identically 1. In view of (5.9), this implies that $m_s=\nu_s$ for all $s\in[q]$. □

We then have the following

Theorem 5.7. *The distributional recursion* (5.8) *admits a unique fixed point if and only if the tree reconstruction is infeasible.*

Proof. Let δ_τ denote the Dirac distribution on τ. Initialized with $\hat{Q}_0 = \sum_{\tau\in[q]} \nu_\tau \delta_\tau$, recursion (5.8) characterizes the distribution $\hat{Q}_d$ of the distribution $\hat{\nu}_d$ of the root spin σ_r conditionally on $\mathcal{G}_d$.

Define $\mathcal{H}_d = \sigma(\mathcal{G}_d, \mathcal{T}, \{\sigma_{\mathcal{L}_k}\}_{k>d})$. Conditionally on $\mathcal{G}_d$, σ_r is independent of $\mathcal{T}$ and $\{\sigma_{\mathcal{L}_k}\}_{k>d}$. Thus $\hat{\nu}_d$ is also the distribution of σ_r conditionally on $\mathcal{H}_d$. The sigma-fields $\mathcal{H}_d$ decrease with d. Thus by the backward martingale convergence theorem 5.2, $\hat{\nu}_d$ converges almost surely to the limit $\hat{\nu}_\infty$, which is the conditional distribution of σ_r given $\mathcal{H}_\infty$. It follows that $\hat{Q}_d$ converges weakly to $\hat{Q}_\infty$, the distribution of $\hat{\nu}_\infty$. By continuity (details to be added), then necessarily $\hat{Q}_\infty$ is a fixed point of (5.8).

The Dirac mass on distribution ν is a fixed point of (5.8). Under the assumption that there is a unique fixed point, $\hat{\nu}_d$ then converges in probability to ν, and Proposition 5.1 implies that reconstruction is infeasible.

Conversely, assume that there is a non-trivial fixed point $\hat{Q}$ of (5.8). Let then, for $\tau \in [q]$,

$$Q_\tau(d\eta) = \frac{\eta(\tau)}{\nu_\tau}\hat{Q}(d\eta).$$

By Lemma 5.4, $\hat{Q}$ is consistent, and this thus defines a probability distribution on $M([q])$. The fixed point relation (5.8) of $\hat{Q}$ implies that the Q_τ are fixed points of the conditional recursive distribution equation (5.6). Given observations σ_i, $i \in \mathcal{L}_d$, construct the distributions $\eta_i = Q_{\sigma_i}$, and propagate them towards the root r of the tree using belief propagation (5.4).

The fixed point (5.6) verified by the Q_τ implies that the obtained distribution at the root is distributed according to Q_{σ_r}. Determine at random an estimate $\hat{\sigma}_r$ of σ_r by setting $\hat{\sigma}_r = s$ with probability $\eta(s)$, where η is the belief thus obtained at the root. Write then

$$\begin{aligned} I(\sigma_r;\hat{\sigma}_r) &= \sum_{s,\tau\in[q]} \nu_s \int_{M([q])} \eta(\tau) Q_s(d\eta) \ln\left(\frac{\nu_s \int_{M([q])} \eta(\tau) Q_s(d\eta)}{\nu_s \nu_\tau}\right) \\ &= \sum_{s,\tau\in[q]} \int_{M([q])} \eta(s)\eta(\tau)\hat{Q}(d\eta) \ln\left(\frac{\int_{M([q])} \eta(s)\eta(\tau)\hat{Q}(d\eta)}{\nu_s\nu_\tau}\right) \\ &= D(m\|\nu\otimes\nu), \end{aligned}$$

where m is the distribution on $[q]\times[q]$ specified by

$$m_{s\tau} = \int_{M([q])} \eta(s)\eta(\tau)\hat{Q}(d\eta).$$

Assume that $I(\sigma_r;\hat{\sigma}_r) = 0$. Then necessarily, $m = \nu\otimes\nu$. This implies in particular that for all $s \in [q]$,

$$\int_{M([q])} \eta(s)^2 \hat{Q}(d\eta) = \left(\int_{M([q])} \eta(s)\hat{Q}(d\eta)\right)^2.$$

Then necessarily, $\eta(s) = \nu_s$ $\hat{Q}$-almost surely for all $s \in [q]$, i.e. $\hat{Q} = \delta_\nu$, a contradiction. □

5.7 Notes

The proof of Theorem 5.7 builds on that of Proposition 1 in Mézard and Montanari [29], that it extends by relaxing the assumption that ν is uniform.

Sly [40] used the density evolution equation to prove that for b-ary trees and the symmetric Potts model (5.1), when $q \geq 4$, reconstruction is feasible below the Kesten-Stigum threshold. Such results had earlier been conjectured in Mézard and Montanari [29].

Chapter 6

Community Reconstruction

We now assume that P is reversible for ν, i.e. $\nu_s P_{st} = \nu_t P_{ts}$, $s, t \in [q]$. We consider the following problem. Given n nodes $i \in [n]$, assign to each a spin $\sigma_i \in [q]$, drawn i.i.d. according to ν. Conditionally on node spins, independently for any pair of nodes (i, j) in $[n]$, with respective spins $\sigma_i = s$, $\sigma_j = t$, create an edge between them with probability R_{st}/n, where

$$R_{st} = \alpha P_{st} \frac{1}{\nu_t}.$$

The reversibility condition on P implies that R is symmetric. Moreover conditionally on node spins, the average number of neighbors of any node i is asymptotic to α as $n \to \infty$, irrespective of the value $\sigma_i \in [q]$.

The resulting random graph is known as the *stochastic block model*; it generalizes the Erdős-Rényi random graph $\mathcal{G}(n, \alpha/n)$, which would correspond to the case where $R_{st} \equiv \alpha$. We will also make use of the *mean progeny matrix* $(M_{st})_{s,t\in[q]}$ describing the average number of spin t-neighbors of a node with spin s. Is is readily verified that $M = \alpha P$.

6.1 Inference problems

Reconstruction is the problem of providing estimates of the underlying block structure, possibly through spin estimates $\hat{\sigma}_i$ determined from the observed graph.

We have the following definition

Definition 6.1. *The overlap of spin estimates $\hat{\sigma}_i$ is by definition*

$$\max_{\pi} \frac{1}{n} \sum_{i\in[n]} \mathbf{1}_{\pi(\sigma_i)=\hat{\sigma}_i} - \sup_{s\in[q]} \nu_s, \tag{6.1}$$

where π runs over permutations of $[q]$.

We say that weak reconstruction is feasible (respectively, polynomial-time feasible) if there exist estimates $\hat{\sigma}_i$ computed from G (respectively, computed in polynomial time from G), that achieve with high probability overlap at least ϵ for some $\epsilon > 0$.

Remark. *Zero overlap can always be achieved by taking $\hat{\sigma}_i \equiv 1$. In the case where $\nu_s \equiv 1/q$, it can also be achieved by assigning the $\hat{\sigma}_i$ independently of G, in an i.i.d. manner.*

An alternative definition of weak reconstruction consists in requiring that there exist estimates $\hat{\sigma}_i$ such that the two empirical distributions of the σ_i and $\hat{\sigma}_i$ be asymptotically correlated, i.e. for some $\epsilon > 0$, that with high probability,

$$\liminf_{n\to\infty} \sum_{s,t\in[q]} p_n(s,t) \ln \left(\frac{p_n(s,t)}{\nu_s q_n(t)} \right) \geq \epsilon, \tag{6.2}$$

where $p_n(s,t) = \frac{1}{n}\sum_{i\in[n]} \mathbf{1}_{\sigma_i=s,\hat{\sigma}_i=t}$ *and* $q_n(t) = \sum_{s\in[q]} p_n(s,t)$.

The first notion always implies the second. By the argument of [9] in the proof of Theorem 5, it holds that when ν *is uniform on* $[q]$*, the second property implies that a possibly different estimate* $\hat{\sigma}_i'$ *can be constructed that achieves strictly positive overlap and thus in that case the two notions coincide. When* ν *is not uniform however, the two notions may differ.*

6.2 Weak community reconstruction implies tree reconstruction

We have the following

Lemma 6.1. *Let* $d \leq c\ln(n)$ *for* $c > 0$ *small enough, and let* $i \in [n]$*. Then conditional on* $\sigma_i = s$*, the law of the d-neighborhood of i in G, constituted of graph edges and node spins, converges in variation to that of the Poisson Galton-Watson branching process studied in the previous section.*

We just sketch the idea and refer the reader to [9] for details. The Stein-Chen method entails that $|\mathrm{Poi}(\lambda) - \mathrm{Bin}(n,\lambda/n)|_{\mathrm{Var}} \leq 2\lambda/n$. We may further show that with high probability, the d-neighborhood in G of node i is cycle-free, and that its size is $o(n^{c'})$ for some constant c' less than $1/2$ for small enough $c > 0$. The proof then consists in building the d-neighborhood of i by successively considering nodes j already added to this neighborhood, and adding their yet undiscovered neighbors of each type s. At each step, the probability that the number of newly added type s neighbors differs from $\mathrm{Poi}(\alpha P_{\sigma_j s})$ is thus at most $O(n^{c'-1}) + O(1/n)$. The overall probability of failure is at most $O(n^{2c'-1}) = o(1)$.

We need another Lemma, whose idea appeared in Mossel, Neeman and Sly [34], and which was extended in Gulikers, Lelarge and Massoulié [17]. In both references, only the case of symmetric SBM on two communities was treated, but the proof generalizes directly to give:

Lemma 6.2. *Let* i *be chosen uniformly at random in* $[n]$*,* $d \leq c\ln(n)$ *for* $c > 0$ *small enough,* U *be the d-neighborhood of* i *in* G*,* V *the set of nodes at distance* $d+1$ *from* i *in* G *and* $W = [n] \setminus (U \cup V)$*. Then*

$$\forall \epsilon > 0,\ \lim_{n\to\infty} \mathbb{P}(|\mathbb{P}(\sigma_i = s|\sigma_{V\cup W}, G) - \mathbb{P}(\sigma_i = s|\sigma_V, G|_{U\cup V})| \geq \epsilon) = 0.$$

Essentially the Lemma states that the spins σ_i on the graph G follow an approximate version of the Markov random field conditional independence property.

Combined with the previous lemma it entails that, for randomly selected i and j, provided $d \to \infty$, assuming the tree reconstruction problem is infeasible, with high probability

$$\mathbb{P}(\sigma_i = s|G, \sigma_W, \sigma_j = t) \to \nu_s.$$

One has a fortiori

$$\mathbb{P}(\sigma_i = s|G, \sigma_j = t) \to \nu_s.$$

This implies that for any estimates $\hat{\sigma}_i = f_i(G)$, with high probability (using Bienaymé-Tchebitchev inequality), for all s, t,

$$p_n(s,t) = \nu_s q_n(t) + o(1).$$

In turn this readily implies the following

Theorem 6.1. *Reconstruction in the SBM is impossible according to the second, weaker definition* (6.2) *proposed in the Remark 6.1 whenever the associated tree reconstruction is impossible.*

The inverse implication is not expected to hold true in general, see for instance Moore [33] p.29, which suggests the existence for a particular example of 5-community symmetric block model of parameter ranges for which tree reconstruction is possible while the corresponding weak community reconstruction is not.

This begs the question: can one identify the exact condition on model parameters for which weak community reconstruction is feasible?

6.3 Conjectured condition for community reconstruction

Such a condition has been rigorously identified in Coja-Oghlan et al. [11] for symmetric, disassortative Stochastic Block Models. The corresponding results in [11] are in fact a confirmation of a more general prediction from statistical physics that might be called the Bethe Ansatz, after the Bethe free energy introduced in Chapter 4 for pairwise graphical models. We now try to convey its general idea.

Consider the density evolution equation (5.8) of the tree reconstruction problem associated with the community reconstruction problem at hand. Let $\hat{Q}$ denote any fixed point solution of (5.8).

The premise for the argument below is that this fixed point characterizes the distribution of messages that are stationary points of belief propagation in a large graph limit, and hence candidates for distributions achieving the minimum of the Bethe free energy.

Now, if the graph we are considering was a tree, for an arbitrary node i with degree D_i, and i.i.d. samples $\eta_1, \ldots, \eta_{D_i}$, according to formula (4.14), the conditional distribution μ_i of node i's community given the rest of the graph would read

$$\mu_i(s) \propto \nu_s \prod_{j=1}^{D_i} \Big(\sum_{s_j \in [q]} P_{s s_j} \frac{\eta_j(s_j)}{\nu_{s_j}} \Big). \tag{6.3}$$

In our particular setup, where D_i admits a Poisson distribution, it turns out that μ_i is in fact distributed according to $\hat{Q}$.

Similarly, for an arbitrary edge (i, j), given i.i.d. samples η, η' drawn according to $\hat{Q}$, according to (4.15), the conditional joint distribution μ_{ij} of the types of nodes i and j would read

$$\mu_{ij}(s, s') \propto \frac{P_{ss'}}{\nu_{s'}} \eta(s) \eta'(s'). \tag{6.4}$$

These can in turn be used to characterize the Bethe free energy of the corresponding stationary set of messages, by applying formula (4.20), as:

$$\begin{aligned} \mathbb{G}_{\text{Bethe}}(\{\mu_i\}, \{\mu_{ij}\}) &= \textstyle\sum_{(ij) \in \mathcal{E}} \sum_{s_i, s_j \in [q]} \mu_{ij}(s_i, s_j) \left[-\ln(\frac{P_{s_i, s_j}}{\nu_{s_j}}) + \ln(\mu_{ij}(s_i, s_j)) \right] \\ &\quad + \textstyle\sum_{i \in \mathcal{V}} \sum_{s_i \in [q]} \mu_i(s_i) \left[-\ln \nu_{s_i} + (1 - D_i) \ln(\mu_i(s_i)) \right] \\ &\approx n(-E(\hat{Q}) - H(\hat{Q})), \end{aligned}$$

for functionals E and H defined by

$$-E(\hat{Q}) := \int \hat{Q}(d\eta) \sum_s \eta(s) [-\ln(\nu_s)] + \frac{\alpha}{2} \int \int \hat{Q}(d\eta) \hat{Q}(d\eta') \left(\sum_{s, s'} \mu_{ij}(s, s') [-\ln(\frac{P_{s,s'}}{\nu_{s'}})] \right),$$

and

$$\begin{aligned} -H(\hat{Q}) &:= \tfrac{\alpha}{2} \textstyle\int \int \hat{Q}(d\eta) \hat{Q}(d\eta') \sum_{s, s' \in [q]} \mu_{ij}(s, s') \ln(\mu_{ij}(s, s')) \\ &\quad + \textstyle\sum_{d \geq 0} \mathbb{P}(\text{Poi}(\alpha) = d) \int \cdots \int \hat{Q}(d\eta_1) \ldots \hat{Q}(d\eta_d) \sum_{s \in [q]} \mu_i(s)(1 - d) \ln(\mu_i(s)). \end{aligned}$$

In the above formula, μ_{ij} stands for the distribution in (6.4), and μ_i stands for the distribution in (6.3) with $D_i = d$.

The ansatz then states that when the graph on which we consider a Markov random field is locally tree-like with long-range correlations that are "weak" in some suitable sense, the posterior distribution of spins is well approximated by the minimizer Q^* of the Bethe free energy functional we have just identified over fixed points $\hat{Q}$ of the recursive equation.

In turn, it predicts that the mutual information $I(\sigma_{[n]}; G)$ between the observed graph G and the spin vector $\sigma_{[n]}$ is asymptotically equivalent to $n[\sum_s \nu_s \ln(1/\nu_s) - H(Q^*)]$. Thus reconstruction is feasible if and only if $H(Q^*) > \sum_s \nu_s \ln(1/\nu_s)$.

[11] establishes correctness of this formula and adaptations thereof to various models (including symmetric disassortative SBM, random graph coloring, and random linear binary codes). See also [13], which addresses the related problem of establishing a law of large numbers for the logarithm of the partition function of Gibbs measures on a sparse random graph.

6.4 Failure of classical spectral methods in sparse case

Spectrum pollution: Combination of first and second moment methods imply that in each community s, there exist nodes i with: degree $d = \Theta(\ln(n)/\ln\ln n)$, and all neighbors of degree at most $\delta \ll d$. Moreover this still holds for some i in the complement of the giant component. Hence there are eigenvectors with associated eigenvalue $\Theta(\sqrt{\ln(n)/\ln\ln n})$ and support in some non-giant connected component. These would not provide information on the spins. With some additional work one can show that $\sqrt{\ln(n)/\ln\ln n}$ is the correct order of magnitude for the Perron-Frobenius eigenvalue of such graphs.

Another argument for showing that information is not present in the eigenvectors related to large eigenvalues is as follows. Assume that for some eigenvalue μ, the corresponding normed eigenvector x is non-localized, i.e. for some $\epsilon > 0$, the set $\mathcal{I}$ of nodes i such that $|x_i| \geq \epsilon/\sqrt{n}$ has macroscopic size, i.e.

$$|\mathcal{I}| \geq \epsilon n.$$

Write then by the relation $\mu x = Ax$,

$$i \in \mathcal{I} \Rightarrow \mu \frac{\epsilon}{\sqrt{n}} \leq \sum_{j \sim i} |x_j|.$$

Summed over $i \in \mathcal{I}$ this yields

$$\mu \epsilon^2 \sqrt{n} \leq \sum_{j \in [n]} |x_j| \times |\{i \in \mathcal{I} : i \sim j\}| \leq \sqrt{\sum_{j \in [n]} d_j^2}$$

by Cauchy-Schwarz, where d_j is the degree of node j. Since the sum of squares of degrees concentrates around $n(\alpha^2 + \alpha)$, we get with high probability:

$$\mu \leq \frac{\sqrt{\alpha^2 + \alpha}}{\epsilon^2}.$$

Thus any eigenvalue μ that is large compared to λ cannot have an eigenvector that is delocalized. While not a proof, this gives a plausibility argument that standard spectral clustering based on embedding nodes from eigenvectors of leading eigenvalues of G's adjacency matrix can't lead to successful weak reconstruction, even above the Kesten-Stigum threshold.

6.5 Spectral redemption

The Belief Propagation equations can be written in terms of distributions $\psi^{i \to j}$ that are passed along neighbor nodes $i \sim j$, and are updated as follows:

$$\psi_s^{i \to j} = \frac{\nu_s \prod_{k \sim i, k \neq j} \sum_{s_k \in [q]} \psi_{s_k}^{k \to i} R_{s s_k}}{\sum_{t \in [q]} \nu_t \prod_{k \sim i, k \neq j} \sum_{s_k \in [q]} \psi_{s_k}^{k \to i} R_{t s_k}}.$$

The following conjecture was made by Decelle et al. [12]:

Conjecture 6.1. *When above the Kesten-Stigum threshold, BP initialized with random starting points converges to some limit messages $\psi_s^{i \to j}$ such that positive overlap is achieved by $\hat{\sigma}_i \in argmax_{s \in [q]} \psi_s^i$, where*

$$\psi_s^i \propto \nu_s \prod_{j \sim i} \sum_{s_j \in [q]} \psi_{s_j}^{j \to i} R_{s s_j}.$$

Despite convincing numerical evidence, this conjecture is still open, and an analysis of the convergence properties of regular BP message passing on G drawn from the Stochastic Block Model is still missing. The challenge of analyzing BP therefore led Krzakala et al. [25] to consider its linearized version instead. Writing

$\psi_s^{i\to j} = \nu_s(1+\epsilon_s^{i\to j})$, the first order linearization of BP around the trivial fixed point distribution ν reads, using the relation between R_{st} and P_{st} and reversibility of P with respect to ν:

$$\epsilon_s^{i\to j} = \sum_{k\sim i, k\neq j} \sum_{s_k\in[q]} \epsilon_{s_k}^{k\to i}[P_{ss_k} - \nu_{s_k}].$$

We can further drop the term ν_{s_k} in the above since the normalization constraint implies that $\sum_s \epsilon_s^{k\to i}\nu_s = 0$.

Introduce the non-backtracking matrix B, whose columns and rows are indexed by oriented edges $(i\to j)$ in graph G, hence it is of size $2m$ where m is the number of non-oriented edges of G, defined by

$$B_{i\to j,k\to\ell} := \mathbf{1}_{j=k}\mathbf{1}_{\ell\neq i}.$$

It is then readily seen that the linearized BP equations read

$$\epsilon = (B^\top \otimes P)\epsilon.$$

Krzakala et al. [25] thus conjectured that positive overlap can be achieved above KS by constructing estimates $\hat{\sigma}_i$ from the leading eigenvectors of non-backtracking matrix B. One motivation for this conjecture is the separation between parameters R_{st}, ν_s and the non-backtracking nature of BP through the tensor product between B and P. Thus a single spectral clustering algorithm based solely on B could extract signal from G drawn from any SBM.

Characterization of spectrum of B:

Theorem 4 of [9] gives a characterization of the spectrum of B. To state it, we need the following notation. Let $\lambda_i(M)$ denote the eigenvalues of the mean progeny matrix αP sorted by decreasing absolute value, hence $\lambda_i(M) = \alpha\lambda_i(P)$. Let $\lambda_i(B)$ denote the eigenvalues of the non-backtracking matrix B sorted by decreasing modulus.

Let $x_i \in \mathbb{R}^q$ denote an eigenvector of M associated with eigenvalue $\lambda_i(M)$. Let $y_i \in \mathbb{R}^{2m}$ be defined for any oriented edge $e = (u,v)$ by $y_i(e) = x_i(\sigma_u)$. Finally let $z_i = B^\ell B^{\top\ell} y_i$, where $\ell = c\ln(n)$ for positive sufficiently small constant $c > 0$. Then we have the following

Theorem 6.2. *Let $r_0 = \sup\{i\in[q] : |\lambda_i(M)|^2 > \lambda_1(M)\}$. Then for all $i \in [r_0]$, $|\lambda_i(B) - \lambda_i(M)|$ converges to zero in probability. For $i > r_0$, $|\lambda_i(B)| \le \sqrt{\lambda_1(M)} + o(1)$.*

Moreover for $i \in [r_0]$ such that eigenvalue $\lambda_i(M)$ is simple, B admits an eigenvector ξ_i associated with $\lambda_i(B)$ that is asymptotically parallel to z_i, i.e.

$$\lim_{n\to\infty} \frac{\langle z_i, \xi_i\rangle}{\|z_i\|\cdot\|\xi_i\|} = 1.$$

Moreover for such $i \in [r_0]$ with $i > 1$, the vector $\phi \in \mathbb{R}^n$ defined by $\phi_u = \sqrt{n}\sum_{v\sim u}\xi_i(uv)$ is such that the empirical distribution $\frac{1}{n}\sum_{u\in[n]}\delta_{(\sigma_u,\phi_u)}$ converges in probability to the distribution of a pair (σ,ϕ) such that $I(\sigma,\phi) > 0$.

The intuition for why iterates of B applied to vectors y_i might be good candidates for eigenvectors of B is as follows. Fix some oriented edge $u\to v$. By definition of y_i,

$$B^{\top\ell} y_i(u\to v) = \sum_{(e,f)} x_i(\sigma_e)|\{\text{number of length } \ell \text{ nonbacktracking walks from } (e,f) \text{ to } (u,v)\}|.$$

By coupling the neighborhood of node v with a GW branching process, the latter sum corresponds to the martingale we studied earlier, up to factor $(\alpha\lambda_i(P))^\ell$, whose mean is $x_i(\sigma_u)$. It is uniformly integrable, hence converges to a limit, with distribution that depends on σ_u, that we denote $\Delta_{u,v}$. Thus heuristically,

$$B^\top(B^{\top\ell} y_i)(u\to v) \approx (\lambda_i(M))^{\ell+1}\Delta_{u,v} \approx \lambda_i(M) y_i(u\to v).$$

Still heuristically, this gives some indication for why the entries of vector ϕ, obtained by projecting eigenvector ξ_i down to $\mathbb{R}^n$, might be correlated with the spin vector, just as the census in the tree reconstruction problem was correlated with the spin at the root node.

Remark. *In [41], the structure of the eigenvectors ξ_i of B when $\lambda_i(M)$ has multiplicity larger than 1 is elucidated. [41] also establishes the following result. Let $\xi(v) = \sqrt{n}\sum_{u\sim v}\xi_2(u,v)$, where ξ_2 is a normed eigenvector of B associated with $\lambda_2(B)$. Partition then the nodes $v \in V$ into two sets I^+, I^- in a probabilistic manner by letting*

$$\mathbb{P}(v \in I^+|\xi_2) = \frac{1}{2} + \frac{1}{2K}\xi(v)\mathbf{1}_{|\xi(v)|\leq K}.$$

Assign label 1 to each node in I^+ and label 2 to each node in I^-. Then, for equal-sized communities ($\nu_r \equiv 1/r$) and when above the Kesten-Stigum threshold, this procedure achieves strictly positive overlap.

6.6 Existence of hard phase

Argument in Banks et al. [5]: consider the symmetric SBM on q blocks, with edge probabilities c_{in}/n, c_{out}/n. The average degree d is $[c_{in} + (q-1)c_{out}]/q$. The mean progeny matrix admits d as its Perron-Frobenius eigenvalue, and $\lambda_2 = (c_{in} - c_{out})/q$ is its only other eigenvalue (with multiplicity $q-1$). Let $\lambda := \lambda_2/d$.

Say that a partition of node set $[n]$ is **good** if it splits it into q equal-sized sets and its number of within-group edges m_{in} and its number of accross-group edges m_{out} verify

$$|m_{in} - \overline{m}_{in}| \leq n^{2/3}, \quad |m_{out} - \overline{m}_{out}| \leq n^{2/3},$$

where $\overline{m}_{in} = (nc_{in})/(2q)$, $\overline{m}_{out} = [n(q-1)c_{out}]/(2q)$. It is shown in [5] that, provided

$$d > d^{upper} := \frac{2q\ln(q)}{[1+(q-1)\lambda]\ln(1+(q-1)\lambda) + (q-1)(1-\lambda)\ln(1-\lambda)},$$

then with high probability, any good partition necessarily has overlap $\beta(d,q,\lambda) > 0$ with the true partition, where $\beta(d,q,\lambda)$ is defined as the solution of a fixed-point equation. Thus a brute-force search succeeds to achieve positive overlap with high probability. The proof follows the first moment method: it shows that the expected number of good partitions with overlap below the announced one goes to zero as $n \to \infty$.

6.7 Nature of hard phase

The intuition for the nature of the hard phase is as follows. Feasibility corresponds to the fact that the Bayes posterior distribution $\mathbb{P}(\sigma = \cdot|G)$ puts with high probability mass $\Omega(1)$ on a set of spin vectors whose overlap is bounded away from zero. Hardness corresponds to the fact that for an initialization of belief vectors that is independent of σ, with high probability BP iterations will converge to the uninformative, trivial fixed point. The hard phase therefore corresponds to a case where the trivial fixed point for BP is attractive while there exists a set of "good configurations" which captures sizeable mass in the posterior distribution of the spin vector. Since the energy of the good configurations is lower than that of uncorrelated ones, there must therefore exist an entropy barrier between the two stable points of BP (add plot).

6.8 Non-backtracking matrices and Ramanujan graphs

Ramanujan graphs, introduced by Lubotzky, Phillips and Sarnak [27] are by definition d-regular graphs whose adjacency matrix A is such that

$$\sup_{\lambda\in \mathrm{Sp}(A), |\lambda|\neq d} \{|\lambda|\} \leq 2\sqrt{d-1};$$

see also Lubotzky [27] for more background. We now relate this property to that of the spectrum of the nonbacktracking matrix B of the graph. Let thus $G = (V,E)$ be a graph on $n := |V|$ nodes with $m := |E|$

edges, and define B to be the $2m \times 2m$ matrix indexed by pairs (e, f) of oriented edges on the edge set E of G such that

$$B_{ef} = \mathbf{1}_{e_2=f_1} \mathbf{1}_{e_1 \neq f_2}$$

where $e = (e_1, e_2)$ and $f = (f_1, f_2)$. We shall use the notation $e^{-1} = (e_2, e_1)$, i.e. e^{-1} is the reversal of edge e.

We assume that the oriented edges are ordered as $e(1), \ldots, e(2m)$ so that for $i \leq m$, $e(i+m) = e(i)^{-1}$.

Following Horton, Stark and Terras [18] set $J = \begin{pmatrix} 0 & I_m \\ I_m & 0 \end{pmatrix}$. Then define the $n \times 2m$ start matrix S and the $n \times 2m$ terminal matrix T by setting $S_{ve} = \mathbf{1}_{v=e_1}$, $v \in V$, $e \in E$, and $T_{ve} = \mathbf{1}_{v=e_2}$, $v \in V$, $e \in E$.

Further define A to be the graph's adjacency matrix, D to be the diagonal matrix of node degrees and $Q = D - I_n$.

We then have (Proposition 1 p. 14 in [18]):

$$\begin{array}{l} SJ = T, \quad TJ = S; \\ A = ST', \quad SS' = TT' = D = Q + I_n; \\ B + J = T'S. \end{array} \tag{6.5}$$

We finally reproduce from [18] the identity between $(n+2m) \times (n+2m)$ matrices

$$\begin{pmatrix} I_n & 0 \\ T' & I_{2m} \end{pmatrix} \begin{pmatrix} (1-u^2)I_n & uS \\ 0 & I_{2m} - uB \end{pmatrix} = \begin{pmatrix} I_n - uA + u^2 Q & uS \\ 0 & I_{2m} + uJ \end{pmatrix} \begin{pmatrix} I_n & 0 \\ T' - uS' & I_{2m} \end{pmatrix}, \tag{6.6}$$

where $u \in \mathbb{C}$ is an arbitrary complex scalar, and which directly follows from (6.5).

Taking the determinant of the above gives the celebrated Ihara-Bass formula, that we now reproduce:

$$(1-u^2)^{n-m} \mathrm{Det}(I - uB) = \mathrm{Det}(I - uA + u^2 Q). \tag{6.7}$$

The previous identities (6.5,6.6) can further be used to establish basic correspondences between eigenvalue-eigenvector pairs (λ, y) of B and and vectors $x \in \mathbb{C}^n$ in the kernel of the $n \times n$ symmetric matrix $\lambda^2 I_n - \lambda A + Q$. We have the following

Proposition 6.1. *Let (λ, y) be an eigen-pair of B such that $\lambda \notin \{-1, 0, 1\}$. Then the vector $x := Sy$ is in the kernel of $\lambda^2 I_n - \lambda A + Q$.*

Conversely, let $x \in \mathbb{C}^n$ be a non-zero vector in the kernel of $\lambda^2 I_n - \lambda A + Q$, for some $\lambda \notin \{-1, 0, 1\}$. Then the vector $y := (\lambda J - I_{2m})S'x$ is non-zero, and in the kernel of $\lambda I_{2m} - B$.

Proof. Let $u = \lambda^{-1}$ and $z = -\frac{u}{1-u^2} Sy$, where (λ, y) is an eigenpair of B such that $\lambda \notin \{-1, 0, 1\}$. Write

$$\begin{aligned} (I - uA + u^2 Q)Sy &= Sy - uST'Sy + u^2(SS' - I_n)Sy \\ &= Sy - uS(B+J)y + u^2(SJT'S - S)y \\ &= Sy - uSu^{-1}y - uSJy + u^2 SJ(B+J)y - u^2 Sy \\ &= -uSJy + u^2 SJu^{-1}y + u^2 Sy - u^2 Sy \\ &= 0, \end{aligned}$$

where we used identities (6.5), the fact that y is an eigenvector of B with eigenvalue u^{-1}, symmetry of J and $J^2 = I_{2m}$.

For the converse statement, setting again $u = \lambda^{-1}$, write

$$\begin{aligned} (I_{2m} - uB)uy &= (I_{2m} - uB)(J - uI_{2m})S'x \\ &= (J - uI_{2m} - uBJ + u^2 B)S'x \\ &= (JS' - uS' - u(T'S - J)JS' + u^2(T'S - J)S')x \\ &= (T' - uT'SJS' + u^2 T'SS' - u^2 T')x \\ &= T'(I_n - uST' + u^2(Q + I_n) - u^2 I_n)x \\ &= T'(I_n - uA + u^2 Q)x \\ &= 0, \end{aligned}$$

where we used repeatedly the identities (6.5) and finished with the identity $(I_n - uA + u^2 Q)x = 0$, holding by assumption. To show that $y \neq 0$, compute $Sy = (\lambda^2 - 1)x$ and use $\lambda \notin \{-1, 1\}$, $x \neq 0$. □

We then have the

Corollary 6.1. *A d-regular graph is Ramanujan if and only if the spectrum of its non-backtracking matrix B consists of eigenvalues with modulus $d-1$ or at most $\sqrt{d-1}$.*

Proof. By Proposition 6.1, $\lambda \in \mathrm{Sp}(B)$ if and only if $\lambda \in \{-1, 0, 1\}$ or $\lambda + (d-1)/\lambda \in \mathrm{Sp}(A)$. Equivalently, $\lambda \in \{-1, 0, 1\}$ or $\lambda = \frac{\mu \pm \sqrt{\mu^2 - 4(d-1)}}{2}$ for some $\mu \in \mathrm{Sp}(A)$. For $|\mu| = d$, this gives $\lambda \in \{d-1, 1\}$. Thus the graph is Ramanujan if and only if the spectrum of B consists of eigenvalues in $\{-1, 0, 1, d-1\}$, and equal to $\frac{\mu \pm \sqrt{\mu^2 - 4(d-1)}}{2}$ $|\mu| < d$for some μ with $|\mu| \leq 2\sqrt{d-1}$. The corresponding eigenvalue has modulus $\sqrt{d-1}$, hence the result. □

6.9 From Kesten-Stigum thresholds to the Baik-Ben Arous-Péché phase transition

The Baik-Ben Arous-Péché transition in random matrix theory implies in particular the following (see Benaych-Georges and Nadakuditi [7]). Consider a symmetric $n \times n$ noise matrix X_n, with independent, zero mean, Gaussian entries with variance σ^2/n off-diagonal and $2\sigma^2/n$ on the diagonal. The Wigner semi-circle law entails that its spectral measure converges almost surely to the distribution with density

$$\frac{\sqrt{4\sigma^2 - x^2}}{2\sigma^2 \pi} \mathbf{1}_{|x| \leq 2\sigma}.$$

Then for P_n with rank r and non-zero eigenvalues θ_i, $i \in [r]$, such that $|\theta_1| \geq \cdots \geq |\theta_r|$ one has the following for all $i \in [r]$ as $n \to \infty$:

$$\lambda_i(X_n + P_n) \overset{\text{a.s.}}{\to} \begin{cases} \theta_i + \frac{\sigma^2}{\theta_i} & \text{if } |\theta_i| > \sigma, \\ \pm 2\sigma & \text{otherwise.} \end{cases} \tag{6.8}$$

Note that reconstruction of the θ_i and the corresponding eigenvectors is again a planted structure reconstruction problem, this time under a different assumption on the background noise in which the structure has been planted.

A parallel can be drawn with Theorem 6.2. The adjacency matrix A corresponds, conditionally on the node spins, to a noise matrix X_n with variances $p_{uv}(1 - p_{uv})$ where $p_{uv} = \frac{R_{\sigma_u \sigma_v}}{n}$. The sum of variances on row u is thus asymptotic to

$$\sum_{s \in [q]} \nu_s R_{\sigma_u s} = \alpha = \lambda_1(M).$$

We thus have the correspondance $\sigma^2 \leftrightarrow \lambda_1(M)$. Moreover, the expectation of matrix A conditional on the spins has rank at most q, and eigenvalues corresponding precisely to the spectrum of the mean progeny matrix M. Thus the eigenvalues $\lambda_i(M)$ are visible in the spectrum of B if and only if they satisfy the Kesten-Stigum condition $\lambda_i(M)^2 > \lambda_1(M)$.

This corresponds precisely to the condition $|\theta_i| > \sigma$. Let us show that we can push the correspondence further, at least when the average degree α is large. Assume that for some fixed $u > 0$,

$$\lambda_i(M)^2 = (1 + u)\alpha,$$

and that the degrees d_i of nodes are such that $\sup_{i \in [n]} |d_i - 1 - \alpha| = o(\alpha)$. In full rigour Theorem 6.2 is proven for $\alpha = O(1)$; if it was extended to this range where $\alpha \gg 1$, we could justify the following heuristic argument.

There exists $\lambda = \lambda_i(B) = \pm\sqrt{(1 + u + o(1))\alpha}$. Thus,

$$\mathrm{Det}[\lambda^2 I - \lambda A + \alpha I + \alpha \epsilon] = 0,$$

where ϵ is a diagonal matrix with entries $\alpha^{-1}(d_i - 1 - \alpha) = o(1)$. Thus

$$\mathrm{Det}[(\lambda/\alpha)A - (\lambda^2/\alpha + 1)I - \epsilon] = 0$$

The Bauer-Fike theorem implies that the spectrum of $(\lambda/\alpha)A - (\lambda^2/\alpha + 1)I$ contains an eigenvalue that is $o(1)$. Thus A admits an eigenvalue μ that satisfies (using $\lambda \ll \alpha$)

$$\mu = \frac{\lambda^2/\alpha + 1 + o(1)}{\lambda/\alpha} = \lambda + \frac{\alpha}{\lambda} + o(1),$$

which corresponds precisely to (6.8).

6.10 Conclusion

Many exciting areas not covered here, e.g. Approximate Message Passing (see e.g. [20]), a method well suited to analyze models with average degree $\alpha \gg 1$, a regime for which averaging phenomena induce simplifications compared to the sparse regime. A number of exciting open problems too: nature of hard phase, in terms of landscape for Belief Propagation dynamics, not sufficiently understood; analysis of alternative planted structure reconstruction problems. More subtle phenomena can also be considered, such as the *spinodal transition* (see e.g. [39]), in some models displaying a phase where Belief Propagation successfully performs non-trivial reconstruction, but nevertheless achieves suboptimal performance.

Finally, this domain sheds new light on the problem of computational complexity, bringing to bear tools of statistical physics and information theory, and revealing a more refined picture than worst case approaches popular in Computer Science do.

Chapter 7

Detection problems

In the previous chapter we considered the problem of reconstructing some structure present in an observed graph, namely the partition into blocks of nodes underlying the Stochastic Block Model distribution of the graph.

We now consider the following hypothesis testing problem: given some graph, does it present a specific underlying structure or not? More precisely, the null hypothesis H_0 corresponds to the case where the observed graph has no specific structure, e.g. it is drawn from the Erdős-Rényi distribution $\mathcal{G}(n, \alpha/n)$, while under H_1 it displays some structure, e.g. a block structure for a graph drawn from the stocahstic block model $\mathcal{G}(n, \nu, (\alpha/n)P)$ for some irreducible, reversible transition matrix P on $[q]$, with stationary distribution ν. Below we will denote by $\mathbb{P}_n$ the distribution of the observations under the null hypothesis and by $\mathbb{Q}_n$ the corresponding distribution under the alternative hypothesis, omitting the subscript n when convenient.

The likelihood ratio $Y = \frac{d\mathbb{P}}{d\mathbb{Q}}$ plays a key role in this hypothesis testing problem. Indeed the Neyman-Pearson lemma states the following. A test $T(G) \in \{0, 1\}$ that maximizes the probability of correct detection $\mathbb{P}(T = 1)$ while guaranteeing a probability of false detection $\mathbb{Q}(T = 1)$ to be below a given threshold $\epsilon > 0$ can be constructed by setting $T = 1$ if $Y \geq t$, $T = 0$ if $Y < t$, and issuing a random value for T when $Y = t$, for some choice of threshold t.

We shall adopt the following definition of detection:

Definition 7.1. *We say that detection between $\{\mathbb{P}_n\}$ and $\{\mathbb{Q}_n\}$ is feasible if there is a series of tests $\{T_n\}$ such that*

$$\lim_{n\to\infty} (\mathbb{P}_n(T_n = 0) + \mathbb{Q}_n(T_n = 1)) = 0.$$

Another relevant notion from hypothesis testing theory is the following

Definition 7.2. *The sequence $\{\mathbb{P}_n\}_{n>0}$ of distributions is* contiguous *with respect to $\{\mathbb{Q}_n\}_{n>0}$ if for any sequence of events E_n, we have the following implication:*

$$\lim_{n\to\infty} \mathbb{Q}_n(E_n) = 0 \Rightarrow \lim_{n\to\infty} \mathbb{P}_n(E_n) = 0.$$

The following result is easily shown:

Lemma 7.1. *Assume that there is a finite constant $C > 0$ such that $\sup_{n>0} \mathbb{E}_{\mathbb{Q}_n} Y_n^2 \leq C$, where $Y_n = \frac{d\mathbb{P}_n}{d\mathbb{Q}_n}$. Then the sequence $\{\mathbb{P}_n\}_{n>0}$ is contiguous with respect to $\{\mathbb{Q}_n\}_{n>0}$.*

Proof. For events E_n, write

$$\mathbb{P}_n(E_n) = \mathbb{E}_{\mathbb{Q}_n}(\mathbf{1}_{E_n} Y_n) \leq \sqrt{\mathbb{Q}_n(E_n)\mathbb{E}_{\mathbb{Q}_n} Y_n^2} \leq \sqrt{\mathbb{Q}_n(E_n)C}.$$

The implication directly follows. □

We will be using the following consequence:

Proposition 7.1. *If $\{\mathbb{P}_n\}$ is contiguous with respect to $\{\mathbb{Q}_n\}$, and thus in particular when the assumption of the previous lemma holds, detection between $\{\mathbb{P}_n\}$ and $\{\mathbb{Q}_n\}$ is not feasible.*

Proof. Let $E_n = \{T_n = 1\}$, where T_n is a test assumed to succeed at detection between the two sequences. Thus necessarily, $\lim_{n\to\infty} \mathbb{Q}_n(E_n) = 0$. By contiguity, it follows that $\lim_{n\to\infty} \mathbb{P}_n(E_n) = 0$, and thus $\lim_{n\to\infty} \mathbb{P}_n(T_n = 1) = 0$. This contradicts successful detection, which requires instead $\mathbb{P}_n(T_n = 0) = 1 - \mathbb{P}_n(T_n = 1)$ to tend to zero. □

Lemma 7.2. *The variation distance $|\mathbb{P}_n - \mathbb{Q}_n|_{var}$ is upper-bounded by $2\sqrt{\mathbb{E}_{\mathbb{Q}_n}(Y_n^2) - 1}$. As a consequence, if $\lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}_n} Y_n^2 = 1$, then $\lim_{n\to\infty} |\mathbb{P}_n - \mathbb{Q}_n|_{var} = 0$, and the two sequences $\{\mathbb{P}_n\}$, $\mathbb{Q}_n\}$ are mutually contiguous.*

Proof. By definition,

$$\begin{aligned} |\mathbb{P}_n - \mathbb{Q}_n|_{var} = 2\sup_A |\mathbb{P}_n(A) - \mathbb{Q}_n(A)| &= 2\sup_A |\mathbb{E}_{\mathbb{Q}_n} \mathbf{1}_A(1 - Y_n)| \\ &\leq 2\sup_A \sqrt{\mathbb{Q}_n(A)\mathbb{E}_{\mathbb{Q}_n}(Y_n - 1)^2} \\ &\leq 2\sqrt{\mathbb{E}_{\mathbb{Q}_n}(Y_n - 1)^2} \\ &= 2\sqrt{\mathbb{E}_{\mathbb{Q}_n} Y_n^2 - 1}, \end{aligned}$$

where we used the fact that $\mathbb{E}_{\mathbb{Q}_n}(Y) = 1$. The fact that mutual contiguity holds when the variation distance $|\mathbb{P}_n - \mathbb{Q}_n|_{var}$ tends to zero follows from $|\mathbb{P}_n(E_n) - \mathbb{Q}_n(E_n)| \leq |\mathbb{P}_n - \mathbb{Q}_n|_{var}$. □

We will discuss first detection in the binary symmetric block model, and then consider the planted clique detection model, a variant with a different scaling in terms of edge presence probabilities and block sizes.

7.1 Detection for the binary symmetric block model

The spins σ_i, $i \in [n]$ are chosen uniformly, i.i.d. in $\{\pm\}$, and $\mathbb{P}((uv) \in E(G)|\sigma_{[n]})$ equals a/n if $\sigma_i = \sigma_j$, and b/n otherwise. This corresponds to $\nu = (1/2, 1/2)$, $\alpha = (a + b)/2$, and

$$P = \begin{pmatrix} \frac{a}{a+b} & \frac{b}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix}.$$

We know from the previous chapter that block reconstruction is possible in polynomial time if $\tau := \frac{(a-b)^2}{2(a+b)} > 1$, and impossible (irrespective of computational resources) if $\tau < 1$. The same transition point determines feasibility of detection:

Theorem 7.1. *Let $\mathbb{P}$ denote the law of the binary stochastic block model with parameters (n, a, b), and $\mathbb{Q}_n$ the law of the Erdős-Rényi graph $\mathbb{G}(n, \alpha/n)$, where $\alpha = (a+b)/2$. Let $\tau := \frac{(a-b)^2}{2(a+b)}$. Then detection is feasible when $\tau > 1$, and infeasible when $\tau < 1$.*

Proof. We rely for the direct part on Theorem 6.2, which guarantees that, when $\tau > 1$, the two eigenvalues $\lambda_1(B)$ and $\lambda_2(B)$ of the non-backtracking matrix of the symmetric Stochastic Block Model converge in probability as $n \to \infty$ to $\alpha = (a + b)/2$, and $(a - b)/2$ respectively. On the other hand, Theorem 6.2 implies that under distribution $\mathbb{Q}_n$, we have again convergence in probability of $\lambda_1(B)$ to α, but with high probability, $|\lambda_2(B)| \leq \sqrt{\alpha} + o(1)$.

Thus the test $T_n = 1$ if and only if $|\lambda_2(B)| \geq (1 + \epsilon)\sqrt{\alpha}$ succeeds whenever $(1 + \epsilon)\sqrt{\alpha} < |a - b|/2$.

To show the converse when $\tau < 1$, we shall rely on Proposition 7.1, and establish that the likelihood ratio Y_n has bounded second moment under $\mathbb{Q}_n$. To ease notation, we drop indices n. Write

$$Y_n = 2^{-n} \sum_{s\in\{\pm\}^n} \frac{\mathbb{P}(G|\sigma_{[n]} = s)}{\mathbb{Q}(G)} = 2^{-n} \sum_{s\in\{\pm\}^n} \prod_{(u,v)} W_{uv}(s), \tag{7.1}$$

where the product is over unordered pairs (u,v) of nodes in $[n]$, and

$$W_{uv}(s) := \begin{cases} \frac{2a}{a+b} & \text{if } s_u = s_v \text{ and } (uv) \in E(G), \\ \frac{2b}{a+b} & \text{if } s_u \neq s_v \text{ and } (uv) \in E(G), \\ \frac{1-a/n}{1-(a+b)/(2n)} & \text{if } s_u = s_v \text{ and } (uv) \notin E(G), \\ \frac{1-b/n}{1-(a+b)/(2n)} & \text{if } s_u \neq s_v \text{ and } (uv) \notin E(G). \end{cases}$$

For fixed $s,t \in \{\pm\}^n$, we let $W_{uv} = W_{uv}(s)$ and $V_{uv} = W_{uv}(t)$. It can be checked directly that the following identities hold:

$$\mathbb{E}_{\mathbb{Q}} W_{uv} = 1, \tag{7.2}$$

$$s_u s_v t_u t_v = + \Rightarrow \mathbb{E}_{\mathbb{Q}} W_{uv} V_{uv} = 1 + \frac{1}{n}\frac{(a-b)^2}{2(a+b)} + \frac{(a-b)^2}{4n^2} + O(n^{-3}), \tag{7.3}$$

$$s_u s_v t_u t_v = - \Rightarrow \mathbb{E}_{\mathbb{Q}} W_{uv} V_{uv} = 1 - \frac{1}{n}\frac{(a-b)^2}{2(a+b)} - \frac{(a-b)^2}{4n^2} + O(n^{-3}), \tag{7.4}$$

$$\tag{7.5}$$

Let now $\rho = \rho(s,t) := \frac{1}{n}\sum_{i\in[n]} s_i t_i$, and $S_\pm := |\{(u,v) : s_u s_v t_u t_v = \pm\}|$. One then has:

$$\rho^2 = \frac{1}{n} + \frac{2}{n^2}\sum_{u\neq v} s_u s_v t_u t_v = \frac{1}{n} + \frac{2}{n^2}(S_+ - S_-).$$

Also, $2n^{-2}(S_+ + S_-) = 1 - n^{-1}$. Thes two equations give

$$S_+ = (1+\rho^2)\frac{n^2}{4} - \frac{n}{2}, \quad S_- = (1-\rho^2)\frac{n^2}{4}. \tag{7.6}$$

Put together, these relations will imply that, when $\tau < 1$, one has

$$\mathbb{E}_{\mathbb{Q}} Y_n^2 = (1+o(1))\frac{e^{-\tau/2-\tau^2/4}}{\sqrt{1-\tau}}, \tag{7.7}$$

which will then imply contiguity of $\mathbb{P}_n$ with respect to Q_n, and hence infeasibility of detection. We now establish (7.7). Letting $\gamma = \frac{\tau}{n} + \frac{(a-b)^2}{4n^2}$, Write

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} Y_n^2 &= 2^{-2n}\textstyle\sum_{s,t\in\{\pm\}^n}\prod_{(u,v)} \mathbb{E}_{\mathbb{Q}} W_{uv} V_{uv} \\ &= 2^{-2n}\textstyle\sum_{s,t\in\{\pm\}^n}(1+\gamma+O(n^{-3}))^{S_+}(1-\gamma+O(n^{-3}))^{S_-}, \end{aligned}$$

where we used the representation (7.1) of Y_n together with (7.3) and (7.4). Using the relation $(1+x/n)^{n^2} = (1+o(1))e^{nx-x^2/2}$, valid for fixed x as $n\to\infty$, replacing in the last expression S_+ and S_- by their expressions (7.6), we obtain

$$\mathbb{E}_{\mathbb{Q}} Y_n^2 = (1+o(1))e^{-\tau/2-\tau^2/4}2^{-2n}\sum_{s,t\in\{\pm\}^n}\exp\left(\frac{\rho^2}{2}[\frac{(a-b)^2}{4}+n\tau]\right).$$

The summation together with its prefactor 2^{-2n} can be interpreted as the expectation $\mathbb{E}\exp(Z_n^2/2[\tau + (a-b)^2/(4n)])$ where $Z_n = n^{-1/2}\sum_{i=1}^n \xi_i$, and the ξ_i are i.i.d. uniform on $\{\pm\}$.

By the central limit theorem, and continuity of the function $z \to \exp(\tau z^2/2)$, the random variable $\exp(\tau Z_n^2/2)$ converges in distribution as $n\to\infty$ to $\exp(\tau Z^2/2)$, where $Z \sim \mathcal{N}(0,1)$. To conclude, it remains to show that $\lim_{n\to\infty}\mathbb{E}\exp(Z_n^2/2[\tau+(a-b)^2/(4n)]) = (1-\tau)^{-1/2}$. This will follow form the fact that $\mathbb{E}\exp(\tau Z^2/2) = (1-\tau)^{-1/2}$ (which is readily verified), and uniform integrability of the random variables $\exp(\tau Z_n^2/2)$. To establish this uniform integrability, write

$$\mathbb{P}(e^{\tau Z_n^2/2} \geq M) = \mathbb{P}(|Z_n| \geq \sqrt{2\ln(M)/\tau}) \leq 2e^{-2\ln(M)/(2\tau)},$$

where the last inequality follows from Hoeffding's inequality. This last term reads $2M^{-1/\tau}$, and is integrable in M for M larger than 1, which guarantees uniform integrability, and the announced convergence. □

Remark. *The above proof is taken from Mossel, Neeman and Sly [34]. For the direct part of the proof, they construct a test based on counts of short cycles in the graph rather than on the second largest eigenvalue of the non-backtracking matrix B as we do here.*

The authors of [34] further show, for $\tau < 1$, contiguity of $\mathbb{Q}_n$ with respect to $\mathbb{P}_n$. This implies that consistent estimation of model parameters a, b is impossible. Indeed assume that there is an estimator $\hat{a}_n$ that converges in probability under $\mathbb{P}_n$ to a. Then for arbitrary fixed $\epsilon > 0$, let $E_n = \{|\hat{a} - a| \geq \epsilon\}$, $\mathbb{P}_n(E_n) \to 0$. However the Erdős-Rényi graph $\mathcal{G}(n, \alpha)$ is also a binary symmetric block model with parameters identical to α. The estimator $\hat{a}$ should thus converge in probability to α under Q_n, a contradiction when $a \neq \alpha$.

7.2 Planted clique detection: informational threshold

We consider as a null model (hypothesis H_0) the Erdős-Rényi graph $\mathcal{G}(n, 1/2)$, and for some integer $k > 0$, as the alternative hypothesis H_1, an ErdHos-Rényi graph distributed as under H_0, to which we added all the edges connecting nodes in a set K of siz k, chosen uniformly at random from k-subsets of $[n]$. We shall denote by $\binom{[n]}{k}$ the collection of such subsets.

We then have the following

Theorem 7.2. *Let $\epsilon > 0$ be fixed. Assume that $k = (1-\epsilon)2\log_2(n)$. Then the variation distance $|\mathbb{Q}_n - \mathbb{P}_n|_{var}$ tends to zero as $n \to \infty$. The two sequences are thus contiguous, and detection is infeasible.*

On the other hand, when $k \geq (1+\epsilon)\log_2(n)$, detection is feasible based on the following test: select H_1 if and only if the observed graph contains a clique of size $k_0 = (1+\epsilon)\log_2(n)$.

Proof. For the first part, we shall rely on Lemma 7.2 and show that under $\mathbb{Q}_n$, the second moment of the likelihood ratio Y_n goes to 1 as $n \to \infty$. Notice first that

$$\begin{aligned} Y_n(g) &= \frac{1}{\binom{n}{k}} \sum_{C \in \binom{[n]}{k}} \frac{\mathbb{P}_n(G=g|K=C)}{\mathbb{Q}_n(g)} \\ &= \frac{1}{\binom{n}{k}} \sum_{C \in \binom{[n]}{k}} 2^{\binom{k}{2}} \mathbf{1}_{C \text{ clique of } g}. \end{aligned}$$

Then

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} Y_n^2 &= \left(\frac{1}{\binom{n}{k}}\right)^2 \sum_{C,C' \in \binom{[n]}{k}} 2^{2\binom{k}{2}} \mathbb{Q}(C, C' \text{ cliques of } G) \\ &= \frac{1}{\binom{n}{k}} \sum_{C \in \binom{[n]}{k}} 2^{2\binom{k}{2}} \mathbb{Q}(C, [k] \text{ cliques of } G) \\ &= \frac{1}{\binom{n}{k}} \sum_{\ell=0}^{k} 2^{2\binom{k}{2}} \binom{k}{\ell} \binom{n-k}{k-\ell} \left(\frac{1}{2}\right)^{2\binom{k}{2} - \binom{\ell}{2}} \\ &\leq \frac{1}{\binom{n}{k}} \sum_{\ell=0}^{k} \binom{k}{\ell} \binom{n-k}{k-\ell} 2^{k\ell/2}, \end{aligned}$$

where we used symmetry with respect to the set C' to fix it to $C' = [k]$, introduced the size ℓ of $C' \cap C$, and finally bounded $\binom{\ell}{2}$ by $k\ell/2$. By replacing k by its upper bound $(1-\epsilon)2\log_2(n)$, the last term in the above display is upper-bounded by $n^{(1-\epsilon)\ell}$. We thus have

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} Y_n^2 &\leq (1+o(1)) \frac{k!}{n^k} \sum_{\ell=0}^{k} \binom{k}{\ell} \frac{n^{k-\ell}}{(k-\ell)!} n^{(1-\epsilon)\ell} \\ &\leq \sum_{\ell=0}^{k} \binom{k}{\ell} \frac{k!}{(k-\ell)!} n^{-\epsilon\ell}. \end{aligned}$$

Computing the ratio of consecutive terms in the above sum, it follows that these are decreasing with ℓ, so that

$$\mathbb{E}_{\mathbb{Q}} Y_n^2 \leq 1 + k\binom{k}{1} \frac{k!}{(k-1)!} n^{-\epsilon} = 1 + k^3 n^{-\epsilon} = 1 + o(1).$$

This concludes the first part of the proof.

We now show that with high probability under $\mathbb{Q}_n$, G contains no clique of size $k_0 = (1+\epsilon)\log_2(n)$. To that end we use the union bound:

$$\begin{aligned}\mathbb{Q}_n(G \text{ contains a clique of size } k_0) &\leq \tbinom{n}{k_0}\mathbb{Q}_n([k_0] \text{ clique of } G)\\ &= \tbinom{n}{k_0}2^{-\binom{k_0}{2}}\\ &\leq n^{k_0}n^{-(1+\epsilon)(k_0-1)}\\ &\leq n^{-\epsilon k_0+1+\epsilon},\end{aligned}$$

and for fixed $\epsilon > 0$, this last term goes to zero as $n \to \infty$. Thus the probability of deciding H_1 under $\mathbb{Q}_n$ goes to zero as $n \to \infty$, while by design, the test always decides H_1 when H_1 holds. □

7.3 Planted clique detection: computational threshold

In this section we shall establish that, provided the size parameter k verifies $k = \Omega(\sqrt{n})$, detection and reconstruction of a planted clique of size k are both feasible in polynomial time.

We shall consider spectral methods. Their analysis will rely on the following result, that follows from Theorem 2.1.22 in [3]. It applies to so-called Wigner matrices, that is $n \times n$ symmetric matrices M with i.i.d. entries $M_{ii} = Y_i$ on the diagonal, i.i.d. entries $Z_{ij} = M_{ij}$ off diagonal, $i \in [n]$, $j \in [n]$, $i < j$, such that $\mathbb{E}(Y_1) = \mathbb{E}(Z_{12}) = 0$, $\mathrm{Var}(Z_{12}) = 1$:

Theorem 7.3. *Consider an $n \times n$ Wigner matrix M as above. Assume moreover that, for some fixed constant $c > 0$, it holds that*

$$\forall k \geq 1,\ r_k := \max(\mathbb{E}(|Y_1|^k, \mathbb{E}(|Z_{12}|^k)) \leq k^{ck}. \tag{7.8}$$

Then the largest eigenvalue $\lambda_1(M)$, multiplied by $1/\sqrt{n}$, converges in probability to 2 as $n \to \infty$.

Moreover, for all fixed $\delta > 0$ and $c' > 0$, one has

$$\mathbb{P}(\lambda_1(M) \geq \sqrt{n}(2+\delta)) = o(n^{-c'}). \tag{7.9}$$

Remark. *The second property is obtained by a careful application of Füredi and Komlos' trace method, of which we gave a crude illustration in the first chapters. This yields the upper bound*

$$\mathbb{P}(\lambda_1(M) \geq \sqrt{n}(2+\delta)) \leq n(1+\delta/2)^{-2k}\frac{1}{1-k^c/n},$$

where c is the constant appearing in (7.8), and k is any integer such that $k^c < n$. The result then follows by choosing $k = n^{1/2c}$.

Let G_0 be a $\mathcal{G}(n, 1/2)$ Erdős-Rényi graph, and K a random subset of $[n]$ of size k. To determine whether we observe G_0 or G_1, that is G_0 plus the added clique K, and in the latter case estimate set K, we proceed as follows.

For some fixed $\epsilon > 0$, first remove uniformly at random, and independently of the edges in the observed graph G, ϵn nodes from $[n]$, and consider the induced graph G' on the remaining $n' := (1-\epsilon)n$ nodes. Let K' be the nodes of the planted clique, if any, retained in G', and $k' := |K'|$ its size.

Consider then the following $n' \times n'$ matrices associated with G'_0 and G'_1, the induced graphs without and with planted clique:

$$S^{0,1}_{ij} = \begin{cases} +1 & \text{if } (i,j) \in E(G'_{0,1}),\\ -1 & \text{otherwise.}\end{cases}$$

Then S^0 is a Wigner matrix, to which Theorem 7.3 applies, yielding for all fixed $\delta > 0$, $c' > 0$:

$$\mathbb{P}(\lambda_1(S^0) \geq \sqrt{n'}(2+\delta)) = o(n^{-c'}). \tag{7.10}$$

The difference $\Delta := S^1 - S^0$ is given by

$$\Delta_{ij} = \begin{cases} |S^0_{ij}| - S^0_{ij} & \text{if } i,j \in K', \\ 0 & \text{otherwise.} \end{cases}$$

Note $\bar{\Delta} := \mathbb{E}(\Delta|K')$. This is a block matrix, with eigenvector $(1/\sqrt{k'})\mathbf{1}_{K'}$, associated with eigenvalue $k'-1$. Moreover, matrix $\Delta - \bar{\Delta}$ is, conditionally on K', a Wigner matrix on its $K' \times K'$ block, and zero elsewhere. Theorem 7.3 applies to give for all $\delta, c' > 0$:

$$\mathbb{P}(\lambda_1(\Delta - \bar{\Delta}) > \sqrt{k'}(2+\delta)|K') = o(k'^{-c'}).$$

Write then

$$S^1 = \bar{\Delta} + W,$$

where $W = W^0 + (\Delta - \bar{\Delta})$.

We then have the following

Theorem 7.4. *Assume that $k = \theta\sqrt{n}$, where $\theta > 4$ and $\epsilon > 0$ is such that $\theta(1-\epsilon) > 4$. Fix $r > 0$ such that $\theta(1-\epsilon) - 2 > r > 2$.*

Let S be the signed matrix constructed from G', that will coincide with S^i under H_i, $i \in \{0,1\}$. Then the test that decides H_1 if $\lambda_1(S) \geq r\sqrt{n'}$ and H_0 otherwise succeeds with probability $1 - o(n^{-c'})$ for all $c' > 0$.

Proof. Under H_0, then $\lambda_1(S) = \lambda_1(S^0) \leq (2+\delta)\sqrt{n'}$ with probability $1 - o(n^{-c'})$ for all fixed $\delta > 0$. The announced guarantee holds because $\delta = r - 2 > 0$ by assumption.

Under H_1, using Weyl's inequality,

$$\lambda_1(S) \geq \lambda_1(\bar{\Delta}) - \|W^0\|_{op} - \|\Delta - \bar{\Delta}\|_{op}.$$

Thus for all $\delta, c' > 0$ fixed, conditional on K', with probability $1 - o(n^{-c'})$, $\lambda_1(S) \geq k' - 1 - (2+\delta)\sqrt{n'} - (2+\delta)\sqrt{k'}$. Since, for any $\epsilon' > 0$, with probability $1 - o(n^{-c'})$ one has $k' \geq (1-\epsilon-\epsilon')k \geq \theta(1-\epsilon-\epsilon')\sqrt{n'}$, the result follows by our choice of r. □

Remark. *The setting aside of a set J of ϵn nodes is in fact not necessary for detection, as can be seen from the previous proof. Its use will appear in the estimation of the planted clique K.*

We now turn to the estimation of K, assuming that we are under H_1. Let x be a normed eigenvector of S^1 associated with $\lambda_1(S^1)$. Then, by the results on perturbations of eigenvectors, provided that $2\|W\|_{op} < k' - 1$, we have

$$|\langle x, 1/\sqrt{k'}\mathbf{1}_{K'}\rangle| \geq \sqrt{1 - \frac{\|W\|^2_{op}}{[k' - 1 - \|W\|_{op}]^2}},$$

and the right-hand side is $\beta = \Omega(1)$ with probability $1 - o(n^{-c'})$, provided $\theta(1-\epsilon) > 4$. Thus we have for some sign s:

$$s\sum_{i \in K'} x_i \geq \beta\sqrt{k'},$$

and a fortiori:

$$\sum_{i \in K'} |x_i| \geq \beta\sqrt{k'}.$$

Let

$$C = \{i \in K' : |x_i| > \alpha/\sqrt{k'}\},\ D = \{i \notin K' : |x_i| > \alpha/\sqrt{k'}\},$$

where $\alpha > 0$ is a constant chosen strictly less than β. We then have:

$$\beta\sqrt{k'} \leq \sum_{i \in C} |x_i| + k'\alpha/\sqrt{k'}.$$

Also, by Cauchy-Schwartz inequality,

$$\sum_{i \in C} |x_i| \leq \sqrt{|C|}.$$

Thus:

$$\sqrt{|C|} \geq (\beta - \alpha)\sqrt{k'}.$$

Also, we have

$$\frac{\alpha^2}{k'}(|C| + |D|) \leq \sum_{C \cup D} x_i^2 \leq 1.$$

Thus, among the nodes $i \in C \cup D$, at least a fraction $(\beta - \alpha)^2\alpha^2$ belongs to C.

We now use the set J of ϵn nodes initially set aside. For each $j \in J$, if it does not belong to K, it has $\mathrm{Bin}(|C| + |D|, 1/2)$ neighbors in $C \cup D$. If it does belong to K on the other hand, it has $|C| + \mathrm{Bin}(|D|, 1/2)$ neighbors within $C \cup D$. Thus, selecting some suitable $\epsilon' > 0$, and letting

$$J^* := \{j \in J : \sum_{i \in C \cup D} \mathbf{1}_{j \sim i} \geq (|C| + |D|)(1/2 + \epsilon')\}$$

will, for each constant $c' > 0$, with probability $1 - o(n^{-c'})$, contain all nodes in $J \cap K$ and no other.

The estimation procedure then consists in returning, as an estimate of K, those nodes in J^* together with the nodes i of G' that are neighbors of all nodes $j \in J^*$. By the previous analysis, for $k = \theta\sqrt{n}$ with $\theta(1 - \epsilon) > 4$, this method succeeds with high probability. Indeed, on the event that $J^* = J \cap K$, all nodes in K' are automatically included in the estimated set $\hat{K}$, while each node of G' not in K' is connected to all nodes in J^* with probability $2^{-|J^*|}$. Thus for all $c' > 0$, with probability $1 - o(n^{-c'})$, $\hat{K} = K$.

It now remains to address the case when $k = \theta\sqrt{n}$ with $\theta = \Omega(1)$ and $\theta \leq 4$. In that situation, we can rely on the following technique due to Alon et al. [1]. Let $\ell > 0$ be an integer such that $\theta > 4\sqrt{2^{-\ell}}$. Consider detection first.

Pick sequentially all subsets of ℓ vertices. If they form a clique, consider the collection of other nodes that are neighbors of each of them. Under H_0, these remaining nodes are connected according to an Erdős-Rényi graph, and thus the test based on the leading eigenvalue of the associated S matrix will fail, and this with high probability for the polynomially many such collections of ℓ vertices, in view of our guarantee in $1 - o(n^{-c'})$ on the success probability of each such test.

Under H_1, for some collection of ℓ vertices all belonging to the planted clique K', the derived graph will consist of $K' - \ell$ nodes in the clique, plus $n'' = \mathrm{Bin}(n' - K', 2^{-\ell})$ vertices. Besides clique edges, the remaining edges in the derived graph are present with probability $1/2$. This is thus a graph with $n'' \approx 2^{-\ell}n'$ nodes and a planted clique of size $\theta\sqrt{n}$. The ratio $\theta\sqrt{n}/\sqrt{n''}$ is thus strictly larger than 4 by our choice of ℓ.

By our previous analysis, the test must then succeed with high probability on such an induced graph. Again by the previous analysis, with high probability, on such an induced graph we can apply our detection procedure based on the set J of ϵn nodes initially set aside, and this will allow reconstruction of the planted clique.

We have thus shown the following

Theorem 7.5. *For $k = \theta\sqrt{n}$, and $\theta = \Omega(1)$, there exist polynomial time procedures which succeed with high probability for (i) detecting whether a size-k clique has been planted in a $\mathcal{G}(n, 1/2)$ Erdős-Rényi graph, and (ii) identifying all nodes in the planted clique.*

Chapter 8

Semi-definite programming approaches

A semi-definite program is an optimization program of the following form:

$$\begin{array}{ll} \text{Minimize} & \langle C, X\rangle \\ \text{Over} & X \in \mathcal{S}_n^+ \\ \text{Such that} & \langle A_i, X\rangle = b_i,\ i = 1, \dots, m, \end{array}$$

where $\mathcal{S}_n^+$ is the cone of semi-definite positive symmetric $n \times n$ matrices, $\langle A, B\rangle = \mathrm{Tr}(AB^\top)$ is the Frobenius scalar product between matrices, $C, A_1, \dots, A_m$ are symmetric $n \times n$ matrices, and $b_1, \dots, b_n \in \mathbb{R}$.

The key property of such a program is that it is a convex minimization problem. Indeed, $\mathcal{S}_n^+$ is a convex cone, as can be readily verified from the definition of positive semi-definiteness. It can be solved in polynomial time, e.g. using the ellipsoid method initially developed by Kamarkar to solve linear problems. More precisely, a solution within an additive error of ϵ can be found in time polynomial in n, m and $\log(1/\epsilon)$.

Several NP-complete combinatorial optimization problems admit convex relaxations into semi-definite programs, whose solution (accessible in polynomial time) can be post-processed to yield an approximate solution of the original combinatorial optimization problem with bounded sub-optimality.

This chapter will introduce the Goemans-Williamson approximate solution to the max-cut graph partitioning algorithm. It will then establish the Grothendieck inequality, and how it can lead to approximation of the cut-norm of matrices. Finally it will illustrate an approach to obtain non-trivial block reconstruction in stochastic block model well above the Kesten-Stigum threshold, based on the solution of an adequate semi-definite program.

8.1 Max-cut and the Goemans-Williamson algorithm

Given a graph $G = (V, E)$, the max-cut problem amounts to finding a partition of V into two sets V_-, V_+ such that the number $|E(V_+, V_-)|$ of edges across this partition is maximal. It is NP-complete, contrarily to the min-cut problem which, thanks to the max-flow min-cut theorem, can be solved in polynomial time.

Letting A denote the adjacency matrix of G, the max-cut optimization problem can be written as

$$\begin{array}{ll} \text{Maximize} & \sum_{u,v\in[n]} A_{uv}[1 - Y_{uv}] \\ \text{Over} & Y \in \mathcal{S}_n^+ \\ \text{Such that} & Y_{uv} = \sigma_u\sigma_v,\ u, v \in [n] \\ \text{where} & \sigma_u \in \{+1, -1\},\ u \in [n]. \end{array}$$

We let $MC(G)$ denote the optimal value of this problem. Equivalently, the last constraint requires that Y is of rank 1, and has all its entries in $\{-1, 1\}$. If we suppress this last constraint, and only impose that $Y_{uu} = 1$

for all $u \in [n]$, we obtain the following semi-definite relaxation:

$$\begin{array}{ll} \text{Maximize} & \sum_{u,v\in[n]} A_{uv}[1 - Y_{uv}] \\ \text{Over} & Y \in \mathcal{S}_n^+ \\ \text{Such that} & Y_{uu} = 1,\ u \in [n]. \end{array}$$

We let $GW(G)$ denote the optimal value of this problem. By construction, $GW(G) \geq MC(G)$.

Assume that we have obtained an optimal solution Y^*. Taking a symmetric square root Z of matrix Y^*, and noting z_u the u-th column of Z, one has $Y^*_{uv} = \langle z_u, z_v\rangle$, and the constraint on the diagonal of Y^* guarantees that the vectors z_u lie on the unit sphere $\mathcal{S}^{n-1}$ of $\mathbb{R}^n$.

The z_u vectors can then be used in a randomized algorithm as follows.

1. Pick a uniform vector z on $\mathcal{S}^{n-1}$. Construct the sign vector σ by letting $\sigma_u = \text{sign}(\langle z, z_u\rangle)$. Let $C(\sigma)$ denote the corresponding cut-size.

2. Repeat N times.

3. Output the largest cut obtained over the N iterations.

One then has the following result:

Theorem 8.1. *The expected value $\mathbb{E}(C(\sigma))$ of the cut resulting from the above procedure is larger than $0.878MC(G)$.*

Proof. For two indices u, v, one has

$$\mathbb{P}(\sigma_u = +, \sigma_v = -) = \frac{\arccos\langle z_u, z_v\rangle}{2\pi}.$$

Therefore

$$\begin{array}{rl} \mathbb{E}C(\sigma) & = \sum_{u,v\in[n]} A_{uv} 2\frac{\arccos\langle z_u,z_v\rangle}{2\pi} \\ & = \frac{1}{\pi}\sum_{u,v\in[n]} A_{uv} \arccos\langle z_u, z_v\rangle. \end{array}$$

It can be verified by calculus that for all $x \in [-1, 1]$,

$$\frac{1}{\pi}\arccos(x) \geq \beta\frac{1}{2}(1 - x),$$

where $\beta = 0.87856$. This implies that $\mathbb{E}C(\sigma) \geq \beta GW(G)$, and the result follows from $GW(G) \geq MC(G)$. □

Corollary 8.1. *For $\epsilon > 0$, the above randomized algorithm outputs a cut that is less than $(1 - \epsilon)\beta MC(G)$ with probability at most $[1 + \epsilon\bar{C}/(n^2/4 - \bar{C})]^{-N}$, where $\bar{C} := \mathbb{E}C(\sigma)$.*

The corollary thus shows that for large enough N, the randomized algorithm produces with high probability a cut that is within $\beta(1-\epsilon)$ of $MC(G)$. It simply relies on the fact that, letting $p = \mathbb{P}(C(\sigma) \leq (1-\epsilon)\bar{C})$, since $C(\sigma) \in [0, n^2/4]$, one has

$$p(1 - \epsilon)\bar{C} + (1 - p)n^2/4 \geq \bar{C},$$

so that $p \leq [n^2/4 - \bar{C}]/[n^2/4 - \bar{C} + \epsilon\bar{C}]$.

Remark. *A more detailed presentation can be found in Mohar [31], which describes in particular a derandomization scheme due to Goemans and Williamson, allowing to find a cut of size at least $\beta MC(G)$ with probability 1 in polynomial time.*

8.2 The Grothendieck inequality

For a rectangular matrix $M \in \mathbb{R}^{n\times m}$, consider the following norm:

$$\|M\|_{\infty\to 1} := \sup_{x_i,y_j\in\{-1,1\}} \sum_{i\in[n],j\in[m]} M_{ij}x_iy_j.$$

A semi-definite relaxation of the combinatorial optimization defining $\|M\|_{\infty\to 1}$ is as follows:

$$\begin{array}{ll} \text{Maximize} & \sum_{i\in[n],j\in[m]} M_{ij}\langle u_i, v_j\rangle \\ \text{Over} & u_i, v_j \in \mathbb{R}^{n+m} \\ \text{Such that} & \|u_i\| = 1,\ \|v_j\| = 1. \end{array} \tag{8.1}$$

To see that this is indeed a semi-definite program, one may rewrite the objective function as $\langle \hat{M}, Y\rangle$ where $\hat{M}$ and Y are square symmetric matrices of size $n+m$, $\hat{M}$ is given by

$$\hat{M} = \frac{1}{2}\begin{pmatrix} 0 & M^\top \\ M & 0 \end{pmatrix}.$$

The matrix Y is assumed semi-definite positive, with ones on the diagonal. Then considering a square root of matrix Y, its first m columns as vectors v_j, and its last n columns as vectors u_i, it follows that $Y_{m+i,j} = \langle u_i, v_j\rangle$, and the maximization of $\langle \hat{M}, Y\rangle$ over semi-definite positive Y with ones on the diagonal is indeed equivalent to the above optimization.

Let $f(M)$ denote the value of the semi-definite program (8.1). Clearly, $f(M) \geq \|M\|_{\infty\to 1}$. The Grothendieck inequality is the following

Theorem 8.2. *For any $n \times m$ real matrix M, the optimal value $f(M)$ of (8.1) satisfies*

$$f(M) \leq K_G\|M\|_{\infty\to 1}, \tag{8.2}$$

for some universal constant K_G, with $K_G \leq \frac{\pi}{2\ln(1+\sqrt{2})}$.

Proof. □

see Alon and Naor [2]). Our in

8.3 Application: semi-definite programming for block reconstruction in the SBM

Part III

Spectral methods for the sparse SBM

Chapter 9

Local convergence of sparse SBMs

We begin the systematic study of the sparse SBM by studying the local structure around an arbitrary vertex. Throughout this chapter, G is a random graph generated according to $\mathrm{SBM}(n, \pi, P/n)$, with average degree d.

Preliminaries

Labeled rooted graphs A labeled rooted graph is a triplet $g_* = (g, o, \sigma)$ consisting of a graph $g = (V, E)$, a *root* $o \in V$, and a labeling function $\sigma : V \to \mathbb{N}$. The set of such graphs with $V \subseteq \mathbb{N}$ will be denoted by $\mathcal{G}_*$. The notions of distance, induced subgraphs and neighbourhoods extend naturally to this setting.

For a graph $g = (V, E)$, $x \in V$ for $t \geq 0$, we denote by $(g, x)_t$ the subgraph of G induced by the vertices at distance at most t from x, rooted at x. If g has a label function σ, we can consider $(g, x)_t$ as an element of $\mathcal{G}_*$ as well. The boundary of this set, or equivalently the set of vertices at a distance exactly t from x, is written as $\partial(g, x)_t$.

When the root x is chosen uniformly at random in V, this turns g into a random element of $\mathcal{G}_*$, denoted $U(g)$.

Notions of weak convergence For two rooted (possibly labeled) graphs $(g, o), (g', o')$, define

$$k = \sup\{t \in \mathbb{N} : (g, o)_t \simeq (g', o')_t\},$$

where the isomorphism between $(g, o)_t$ and $(g', o')_t$ also preserves labels. Then, we can define the *distance* between (g, o) and (g', o') as

$$d_*((g, o), (g', o')) = 2^{-k}, \tag{9.1}$$

with the convention that $2^{-\infty} = 0$. This turns $\mathcal{G}_*$ into a compact metric space and hence we can define a weak convergence of measures:

$$\mu_n \xrightarrow{\mathcal{D}} \mu \quad \iff \quad \lim_{n\to\infty} \mathbb{E}_{\mu_n} f = \mathbb{E}_\mu f \tag{9.2}$$

for any function $f : \mathcal{G}_* \to \mathbb{R}$ that is continuous with respect to the metric d_*. Since $(\mathcal{G}_*, d_*)$ is a totally disconnected metric space, the following equivalence also holds: the sequence $(\mu_n)_{n\in\mathbb{N}}$ converges weakly to μ if and only if for all finite rooted graphs γ and depth $t \geq 0$,

$$\lim_{n\to\infty} \mathbb{P}_{\mu_n}((g, o)_t \simeq \gamma) = \mathbb{P}_\mu((g, o)_t \simeq \gamma). \tag{9.3}$$

Any of those equivalent definitions corresponds to the so-called Benjamini-Schramm convergence, from the seminal work [8].

In this chapter, we show in essence that the random graph $U(G)$ converges in distribution to the Galton-Watson tree (T, ρ) defined in [TODO]. We will consider both versions of the weak convergence, and show slightly stronger results in both cases: we show (9.3) for a depth t that is logarithmic in n, and we provide rigorous convergence speed bounds for specific functionals in (9.2).

9.1 Neighbourhood growth rates

In expectation, the size of $\partial(G,x)_t$ (resp. $\partial(T,\rho)_t$) grows at most like d^t. We provide quantitative versions of this bound.

Lemma 9.1. *There exists absolute constants $c_0, c_1 > 0$ such that for any $s > 0$ and $x \in V$,*

$$\mathbb{P}\big(\forall t \geq 0,\ |\partial(G,x)_t| \leq sd^t\big) \geq 1 - c_0 e^{-c_1 s}, \tag{9.4}$$

and the same holds for (T,ρ).

Proof. We begin by showing that this result holds for (T,ρ). For $k \geq 1$, define

$$\varepsilon_t = d^{-t/2}\sqrt{t} \quad \text{and} \quad f_t = \prod_{t'=1}^{t}(1+\varepsilon_{t'}).$$

Since the series $\sum \varepsilon_t$ is bounded, so is f_t, and hence there exists a constant c_1 so that for all $t \geq 1$,

$$\varepsilon_t \leq c_1 \quad \text{and} \quad 1 \leq f_t \leq c_1. \tag{9.5}$$

Let $S_t = |\partial(T,\rho)_t|$; we have $S_0 = 1$, and for $t \geq 0$

$$S_{t+1} = \sum_{k=1}^{S_t} Y_k,$$

where the Y_k are i.i.d Poi(d) random variables. By a Chernoff bound, for any integer $\ell \geq 1$ and $s \leq 1$,

$$\mathbb{P}\left(\sum_{k=1}^{\ell} Y_k \geq \ell d s\right) \leq e^{-\ell\mu_1\gamma(s)}, \tag{9.6}$$

where we defined $\gamma(s) = s\log(s) - s + 1$. In particular, we have

$$\mathbb{P}\big(S_{t+1} \geq sf_{t+1}d^{t+1} \,\big|\, S_t \geq sf_t d^t\big) \leq e^{-sd^{t+1}f_t\gamma(1+\varepsilon_t)} \leq e^{-\theta s d^{t+1}\varepsilon_{t+1}^2} = e^{-\theta s(t+1)},$$

using the existence of some $\theta > 0$ such that $\gamma(1+x) \geq \theta x^2$ for $x \in [0, c_1]$.

Finally, using the bounds in (9.5), for any $s \geq \max(1/\theta, 1/c_1)$,

$$\mathbb{P}(\exists t : S_t \geq sc_1 d^{t+1}) \leq \sum_{t=1}^{\infty} e^{-\theta s t} = \frac{e^{-\theta s}}{1-e^{-\theta s}},$$

from which the statement of Lemma 9.1 ensues for suitably redefined constants c_0, c_1.

Now, consider the random graph x; conditioned on $(G,x)_t$, if $y \in \partial(G,x)_t$ has type i, the number of undiscovered neighbors of y is upper bounded stochastically by

$$V_i = \sum_{j=1}^{r} V_{ij},$$

where the V_{ij} are independent random variables with $V_{ij} \sim \mathrm{Bin}(n_i, W_{ij}/n)$. Hence, for any $\theta \geq 0$, using that $1 + x \leq e^x$ we get that

$$\begin{aligned}
\mathbb{E}\big[e^{\theta V_i}\big] &= \prod_{j=1}^{r}\left(1 - \frac{\pi_i W_{ij}}{n_i} + \frac{\pi_i W_{ij}}{n_i}e^{\theta}\right)^{n_i} \\
&\leq \exp\left((e^{\theta}-1)\sum_{j=1}^{r}\pi_i W_{ij}\right) \\
&= e^{-d(e^{\theta}-1)},
\end{aligned}$$

which is the charactersitic function of a Poi(d) random variable. Hence, the inequality (9.6) also holds for G, and the same proof applies. □

This tail bound can then be turned into moment bounds via classical arguments

Lemma 9.2. *There exists a universal constant c such that for every $p \geq 1$:*

$$\mathbb{E}\left[\max_{t\geq 0}\left(\frac{|\partial(G,x)_t|}{d^t}\right)^p\right]^{\frac{1}{p}} \leq cp \quad \text{and} \quad \mathbb{E}\left[\max_{t\geq 0}\max_{x\in[n]}\left(\frac{|\partial(G,x)_t|}{d^t}\right)^p\right]^{\frac{1}{p}} \leq c(\ln(n)+p)$$

As a result, for any $t \geq 0, p \geq 1$, we have

$$\mathbb{E}[|(G,x)_t|^p]^{\frac{1}{p}} \leq c'pd^t \quad \text{and} \quad \mathbb{E}\left[\max_{x\in[n]}|(G,x)_t|^p\right]^{\frac{1}{p}} \leq c'(\ln(n)+p)d^t.$$

The same inequalities hold for (T,ρ) (resp. n independent copies of (T,ρ)).

Proof. For any $p \geq 0$

$$\begin{aligned}\mathbb{E}\left[\max_{t\geq 0}\left(\frac{|\partial(G,x)_t|}{d^t}\right)^p\right] &= p\int_0^\infty s^{p-1}\mathbb{P}\left(\max_{t\geq 0}\frac{|\partial(G,x)_t|}{d^t}\geq s\right)ds\\ &\leq c_0 p\int_0^\infty s^{p-1}e^{-c_1 s}ds\\ &\leq c^p p!\end{aligned}$$

via the Gamma function expression for the factorial. This implies the first inequality; the second one is done identically, using

$$\mathbb{P}\left(\max_{t\geq 0}\max_{x\in[n]}\frac{|\partial(G,x)_t|}{d^t}\geq s\right) \leq \max\left(1, \sum_{x\in[n]}\mathbb{P}\left(\max_{t\geq 0}\frac{|\partial(G,x)_t|}{d^t}\geq s\right)\right).$$

Now, for any $t \geq 0$,

$$|(G,x)_t| = \sum_{t'=0}^{t} d^{t'}\frac{|\partial(G,x)_{t'}|}{d^{t'}} \leq c'd^t\max_{t'\geq 0}\frac{|\partial(G,x)_{t'}|}{d^{t'}}$$

We can then take the p-th power and the expectation on both sides to get the desired result. □

Taking $p = \ln(n)$, and using the Markov inequality, the following easy corollary ensues:

Corollary 9.1. *There exists an absolute constant $c > 0$ such that with probability at least $1 - 1/n$, for any $x \in [n]$ and $t \geq 0$,*

$$|(G,x)_t| \leq c\ln(n)d^t$$

9.2 Weak convergence of sparse SBMs

We now consider the convergence of local subgraphs of G, by showing the existence of a coupling between G and (T,ρ) up to logarithmic depth. The main ingredient of this section is the following exploration process of the neighbourhood of $x \in [n]$:

Definition 9.1. *Start with $A_0 = \{x\}$, and at stage $t \geq 0$, if A_t is not empty, take a vertex x_t in A_t at minimal distance from v, reveal its neighbourhood N_t in $[n]\setminus A_t$, and update $A_{t+1} = (A_t \cup N_t)\setminus\{x_t\}$. Denote $\mathcal{F}_t$ the filtration generated by $(A_0,\dots,A_t)$, and $D_t = \bigcup_{0\leq s\leq t} A_s$ the set of discovered vertices at time t.*

We first show that G is locally tree-like, i.e. it contains only a few cycles. We say that a graph g is h-tangle-free is for any vertex $x \in g$, the h-neighbourhood of x contains at most one cycle.

Lemma 9.3. *Let $h \leq \kappa\log_d(n)$ with $0 \leq \kappa \leq 1/2$. Then:*

1. *the random graph G is h-tangle-free with probability at least $1 - cn^{2\kappa-1}$,*

2. *the probability that a fixed vertex x has a cycle in its h-neighbourhood is at most $cn^{\kappa-1}$.*

Proof. We first prove the second statement. Let τ be the first time at which every vertex in $(G,x)_h$ has been revealed. It is clearly a stopping time for the filtration $(\mathcal{F}_t)_{t\geq 0}$; and by construction, given $\mathcal{F}_\tau$, the set of discovered edges forms a spanning tree of $(G,x)_h$. Hence, conditioned on $\mathcal{F}_\tau$, the number of undiscovered edges in $(G,x)_h$ is stochastically upper bounded by $\mathrm{Bin}(m, a/n)$ where $m = |(G,x)_h|$ and $a = \max_{i,j} W_{ij}$. Using Lemma 9.2, we immediately get

$$\mathbb{P}((G,x)_h \text{ is not a tree}) \leq \frac{a\mathbb{E}[|(G,x)_h|]}{n} \leq cn^{\kappa-1}.$$

We now turn to the first statement. If G is not h-tangle-free, it means that for $x \in [n]$, there are two undiscovered edges in the exploration process of $(G,x)_h$. Using the fact that

$$\mathbb{P}(\mathrm{Bin}(m,q) \geq 2) \leq q^2 m(m-1) \leq q^2 m^2,$$

we find

$$\mathbb{P}(G \text{ is } h\text{-tangled}) \leq \sum_{x\in[n]} \frac{a^2\mathbb{E}\big[|(G,x)_h|^2\big]}{n^2} \leq \frac{cd^{2h}}{n} \leq cn^{2\kappa-1},$$

as required. □

Before proceeding to the main proof of this section, we recall a few facts about distance in measure space. Recall that the total variation distance $\mathrm{d_{TV}}$ between two measures $\mathbb{P}_1, \mathbb{P}_2$ on the same probability space $(\Omega, \mathcal{F})$ is defined as

$$\mathrm{d_{TV}}(\mathbb{P}_1, \mathbb{P}_2) = \sup_{\mathcal{A}\in\mathcal{F}} |\mathbb{P}_1(A) - \mathbb{P}_2(A)|$$

It is also defined by the following characterization:

$$\mathrm{d_{TV}}(\mathbb{P}_1, \mathbb{P}_2) = \inf_{\substack{X_1\sim\mathbb{P}_1 \\ X_2\sim\mathbb{P}_2}} \mathbb{P}(X_1 \neq X_2)$$

where the supremum is over all *couplings* of $\mathbb{P}_1, \mathbb{P}_2$, i.e. probability measures $\mathbb{P}$ on $(\Omega^2, \mathcal{F}\otimes\mathcal{F})$ such that the marginal distributions of $\mathbb{P}$ are $\mathbb{P}_1$ and $\mathbb{P}_2$.

Proposition 9.1. *Let $h \sim \kappa \log_d(n)$ for $\kappa < 1/2$. Then, for any $x \in [n]$, if the random tree (T,ρ) is started from $\sigma(\rho) = \sigma(x)$, there exists a universal constant $c > 0$ such that*

$$\mathrm{d_{TV}}\left(\mathcal{L}((G,x)_h), \mathcal{L}((T,\rho)_h)\right) \leq c\ln(n)^2 n^{2\kappa-1} \tag{9.7}$$

Proof. Consider again the exploration process of Definition 9.1, and let τ be the time at which all of $(G,x)_h$ has been revealed. We perform the same exploration process, denoted (A'_t, N'_t) in parallel on (T,ρ), which corresponds to a breadth-first traversal of the tree. At each timestep t, denote by $Y_t = (Y_t(1), \ldots, Y_t(r))$ the distribution of spins in N_t (resp. Y'_t that of N'_t). Let $\mathbb{P}_t$ be the law of Y_t given $\mathcal{F}_t$, and by $\mathbb{Q}_t$ the law of Y'_t (no conditioning needed).

From Corollary 9.1 and Lemma 9.3, with probability at least $1 - cn^{\kappa-1}$, we have $\tau \leq c\ln(n)n^\kappa$ and $(G,x)_h$ is a tree. Hence, by recursion, it suffices to show that for each timestep t, if the coupling holds until t, then

$$\mathrm{d_{TV}}(\mathbb{P}_t, \mathbb{Q}_t) \leq c\ln(n)d^{\kappa-1}. \tag{9.8}$$

Assume that $\sigma(x_t) = i$. Letting $n_t(j)$ be the number of vertices with spin j in $[n] \setminus D_t$, then given $\mathcal{F}_t$ the random variables $(Y_t(j))_{j\in[r]}$ are independent and $Y_t(j)$ has distribution $\mathrm{Bin}(n_t(j), Wij/n)$. The random

variables $(Y_t'(j))_{j\in[r]}$ are also independent, and $Y_t'(j)$ has distribution $\mathrm{Poi}(\pi_i W_{ij})$. We shall couple those processes using the following classical bounds (see e.g. [6]):

$$\mathrm{d_{TV}}\left(\mathrm{Bin}\left(m,\frac{\lambda}{m}\right),\mathrm{Poi}(\lambda)\right)\le\frac{\lambda}{m}\quad\text{and}\quad\mathrm{d_{TV}}(\mathrm{Poi}(\lambda),\mathrm{Poi}(\lambda'))\le|\lambda-\lambda'|$$

Using the triangle inequality, we have

$$\begin{aligned}\mathrm{d_{TV}}(\mathbb{P}_t,\mathbb{Q}_t)&\le\mathrm{d_{TV}}\left(\mathbb{P}_t,\bigotimes_{j\in[r]}\mathrm{Poi}\left(\frac{n_t(j)W_{ij}}{n}\right)\right)+\mathrm{d_{TV}}\left(\bigotimes_{j\in[r]}\mathrm{Poi}\left(\frac{n_t(j)W_{ij}}{n}\right),\mathbb{Q}_t\right)\\&\le\sum_{j\in[r]}\left(\frac{W_{ij}}{n}+W_{ij}\left|\frac{n_t(i)}{n}-\pi_i\right|\right)\\&\le c\frac{|(G,x)_h|}{n},\end{aligned}$$

which proves (9.8). □

9.3 Weak law of large numbers for graph functionals

We now leverage the results of the previous section to show concentration properties for local functionals. We say that a function $f:\mathcal{G}_*\to\mathbb{R}$ is h-local if $f(g,o)$ is only a function of $(g,o)_h$.

Lemma 9.4. *Let $f,\psi:\mathcal{G}_*\to\mathbb{R}$ be two t-local functions such that $|f(g,o)|\le\psi(g,o)$ and ψ is non-decreasing by the addition of edges. Then there exists $c>0$ such that for all $p\ge2$,*

$$\mathbb{E}\left[\left|\sum_{o\in[n]}f(G,o)-\mathbb{E}\sum_{o\in[n]}f(G,o)\right|^p\right]^{1/p}\le c\sqrt{n}p^{3/2}d^t\mathbb{E}\left[\max_{o\in[n]}\psi(G,o)^{2p}\right]^{\frac{1}{2p}}$$

Proof. For $x\in[n]$, define E_x the set of edges of the form $\{x,y\}$ with $x\le y$. Then $(E_x)_{x\in[n]}$ is an independent vector, and since the union of E_x is the whole graph G we can write

$$Y:=\sum_{x\in[n]}f(G,o)=F(E_1,\dots,E_n)$$

for some measurable function F. Define G_x the graph with vertex set V and edge set $E\setminus E_x$, and

$$Y_x=\sum_{o\in[n]}f(G_x,o).$$

The random variable Y_x is $\bigcup_{y\neq x}E_y$-measurable, and hence we can use the Efron-Stein inequality (see [10], Theorem 15.5): for any $p\ge2$,

$$\mathbb{E}[|Y-\mathbb{E}[Y]|^p]\le(c\sqrt{p})^p\mathbb{E}\left[\left(\sum_{x\in[n]}(Y-Y_x)^2\right)^{p/2}\right].\tag{9.9}$$

Since f is t-local, for a given $x\in[n]$, the difference $f(G,o)-f(G_x,o)$ is nonzero only if $x\in(G,o)_t$. Consequently,

$$\begin{aligned}|Y-Y_x|&\le\sum_{o\in(G,x)_t}|f(G,o)-f(G_x,o)|\\&\le\sum_{o\in(G,x)_t}\psi(G,o)+\psi(G_x,o)\\&\le2|(G,x)_t|\cdot\max_{o\in[n]}\psi(G,o).\end{aligned}$$

Using Hölder's inequality, we find

$$\mathbb{E}\left[\left(\sum_{x\in[n]}(Y-Y_x)^2\right)^{p/2}\right] \le n^{p/2-1}2^p\mathbb{E}\left[\sum_{x\in[n]}\left(|(G,x)_t|\cdot\max_{o\in[n]}\psi(G,o)\right)^p\right]$$
$$\le n^{p/2}\sqrt{\mathbb{E}[|(G,x)_t|^{2p}]\,\mathbb{E}\left[\max_{o\in[n]}\psi(G,o)^{2p}\right]}$$

The desired bound then follows from (9.9) and Lemma 9.2. □

The second bound relates the expectation of graph functionals with their equivalents on trees. It can be viewed as a generalization of (9.2), with guarantees on the convergence speed.

Lemma 9.5. *Let $h \sim \kappa \log_d(n)$ with $0 < \kappa < 1/2$, and (T,ρ) be a Galton-Watson tree with random root spin $\sigma(\rho) \sim \pi$. There exists $c > 0$ such that for any h-local function $f : \mathcal{G}_* \to \mathbb{R}$,*

$$\left|\frac{1}{n}\sum_{x\in[n]}\mathbb{E}f(G,x) - \mathbb{E}f(T,\rho)\right| \le c\ln(n)n^{\kappa-1/2}\sqrt{\max_{x\in[n]}\mathbb{E}[f(G,x)^2]} \vee \sqrt{\max_{i:\sigma(\rho_i)=i}\mathbb{E}[f(T,\rho_i)^2]}$$

Proof. Notice that by definition of (T,ρ), we have

$$\mathbb{E}f(T,\rho) = \frac{1}{n}\sum_{x\in[n]}\mathbb{E}f(T_x,\rho_x),$$

where $\sigma(\rho_x) = \sigma(x)$. Let $\mathcal{E}_x$ be the event "the coupling between $(G,x)_h$ and $(T_x,\rho_x)_h$ fails"; since f is h-local, we have $f(G,x) = f(T_x,\rho_x)$ on $\mathcal{E}_x^c$. From the Cauchy-Schwarz inequality, for any $x \in [n]$,

$$|\mathbb{E}f(G,x) - \mathbb{E}f(T_x,\rho_x)| \le \mathbb{E}[|f(G,x) - f(T_x,\rho_x)|\mathbf{1}_{\mathcal{E}_x}]$$
$$\le \sqrt{\mathbb{P}(\mathcal{E}_x)}\sqrt{\mathbb{E}[(f(G,x) - f(T_x,x))^2]}$$
$$\le \sqrt{c\ln(n)^2n^{2\kappa-1}}\left(\sqrt{\mathbb{E}[f(G,x)^2]} + \sqrt{\mathbb{E}[f(T_x,x)^2]}\right),$$

where the last inequality follows from Proposition 9.1. □

Combining the two previous results, we show a proposition that both encapsulates the weak local convergence and the concentration properties of G:

Proposition 9.2. *Let $h \sim \kappa\log_d(n)$ with $0 < \kappa < 1/2$. Let $f : \mathcal{G}_* \to \mathbb{R}$ be a h-local function such that $|f(g,o)| \le \alpha|(g,o)_h|^\beta$ for some $\alpha,\beta > 0$. Then, there exists a universal constant $c > 0$ such that for any $s \ge 0$, with probability at least $1 - n^{-s}$,*

$$\left|\frac{1}{n}\sum_{x\in[n]}f(G,x) - \mathbb{E}f(T,\rho)\right| \le c\alpha\log(n)^{5/2+\beta}n^{\kappa(1+\beta)-1/2}$$

Proof. We use the following version of the Markov inequality: for a random variable X and $p \ge 1$,

$$\mathbb{P}\Big(|X| > a\mathbb{E}\big[X^{2p}\big]^{\frac{1}{2p}}\Big) \le a^{-2p}. \tag{9.10}$$

Let $\psi(g,o) = \alpha|(g,o)_h|^\beta$; it is easily checked that ψ satisfies the conditions of Lemma 9.4. Further, by Lemma 9.2, we have

$$\mathbb{E}\left[\max_{o\in[n]}\psi(G,o)^{2p}\right]^{\frac{1}{2p}} \le c\alpha(\log(n)\vee p)^\beta d^{\beta h}.$$

We apply (9.10) with $a = e$ and $2p \geq s\ln(n)$: with probability at least n^{-s},

$$\left| \frac{1}{n} \sum_{x \in [n]} f(G,x) - \mathbb{E}\left[\frac{1}{n} \sum_{x \in [n]} f(G,x) \right] \right| \leq c\alpha \ln(n)^{5/2+\beta} n^{\kappa(1+\beta)-1/2}.$$

The exponent in n above is more than the one in Lemma 9.5, hence by the triangle inequality, upon adjusting the constant c,

$$\left| \frac{1}{n} \sum_{x \in [n]} f(G,x) - \mathbb{E} f(T,\rho) \right| \leq c\alpha \ln(n)^{5/2+\beta} n^{\kappa(1+\beta)-1/2}.$$

□

Chapter 10

Tree functionals and pseudo-eigenvectors

We have shown in the previous chapter that all local functionals of G concentrate around their tree equivalents. We will leverage this fact to build pseudo-eigenvectors for the non-backtracking matrix B, stemming from local functionals.

10.1 Tree martingales

For a vector $\xi \in \mathbb{R}^r$ and $t \geq 0$, we define the functional

$$f_{\xi,t}(g,o) = \sum_{x_t \in \partial(g,o)_t} \xi_{\sigma(x_t)} \tag{10.1}$$

When ξ is an eigenvector of M, these correspond to the classical tree processes defined by Kesten and Stigum [23, 22]. Let (T,ρ) be a Galton-Watson tree, and define $\mathcal{F}_t$ the σ-algebra generated by $(T,\rho)_t$. Throughout this section, we denote by $\mathbb{E}_t$ the conditional expectation $\mathbb{E}[\,\cdot \mid \mathcal{F}_t]$.

We begin with a small lemma, that will prove very useful:

Lemma 10.1. *Let $x_t \in \partial(T,\rho)_t$ for some $t \geq 0$, and $\xi \in \mathbb{R}^r$ any vector. Then*

$$\mathbb{E}_t \left[\sum_{x_t \to x_{t+1}} \xi_{\sigma(x_{t+1})} \right] = [M\xi](\sigma(x_t)),$$

where the sum ranges over all children x_{t+1} of x_t in T.

Proof. Let $\sigma(x_t) = i$. Conditioned on $\mathcal{F}_t$, we have

$$\sum_{x_t \to x_{t+1}} \xi_{\sigma(x_{t+1})} = \sum_{j \in [r]} Y_j \xi_j,$$

where $Y_j \sim \mathrm{Poi}(M_{ij})$. Thus,

$$\mathbb{E}_t \sum_{x_t \to x_{t+1}} \xi_{\sigma(x_{t+1})} = \sum_{j \in [r]} M_{ij}\xi_j = [M\xi](i).$$

□

Proposition 10.1. *Let (μ,ϕ) be an eigenpair of M. Then the random process*

$$Z_t = \mu^{-t} f_{\phi,t}(T,\rho) \tag{10.2}$$

is a $(\mathcal{F}_t)_t$-martingale, with common expectation $\phi_{\sigma(\rho)}$.

Proof. Let $t \geq 0$ be fixed. We have

$$Z_{t+1} - Z_t = \mu^{-(t+1)} \sum_{x_t \in \partial(T,\rho)_t} \left(\sum_{x_t \to x_{t+1}} \phi_{\sigma(t+1)} - \mu \phi_{\sigma(x_t)} \right).$$

By Lemma 10.1, we have for any $x_t \in \partial(T,\rho)_t$

$$\mathbb{E}_t \left[\sum_{x_t \to x_{t+1}} \phi_{\sigma(x_{t+1})} \right] = [M\phi](\sigma(x_t)) = \mu \phi_{\sigma(x_t)},$$

which concludes the proof. □

Our tools also allow to compute the correlation of those two martingales, which reduces to the study of their increments:

Lemma 10.2. *Let (μ, ϕ) and (μ', ϕ') be two (not necessarily distinct) eigenpairs of M, and consider the martingales Z_t and Z'_t as in* (10.2)*. Then*

$$\mathbb{E}[(Z_{t+1} - Z_t)(Z'_{t+1} - Z'_t)] = [\mu\mu']^{-(t+1)} [M^{t+1}(\phi \circ \phi')](\sigma(\rho))$$

As a result, the martingale Z_t converges almost surely and in $\mathcal{L}^2$ whenever $\mu^2 > d$, and for any $0 \leq t \leq t'$,

$$\mathbb{E}[Z_t Z'_{t'}] = \sum_{s=0}^{t} \frac{[M^s(\phi \circ \phi')](\sigma(\rho))}{(\mu\mu')^s}.$$

Proof. Denote by $\Delta_t = Z_{t+1} - Z_t$ (resp. $\Delta'_t = Z'_{t+1} - Z'_t$) the martingale increments. We first compute the conditional expectation $\mathbb{E}_t[\Delta_t \Delta'_t]$:

$$\mathbb{E}_t[\Delta_t \Delta'_t] = [\mu\mu']^{-(t+1)} \sum_{x_t, x'_t \in \partial(T,\rho)_t} E(x_t, x'_t),$$

with the correlation term $E(x_t, x'_t)$ given by

$$E(x_t, x'_t) = \mathbb{E}_t \left[\left(\sum_{x_t \to x_{t+1}} \phi_{\sigma(x_{t+1})} - \mu \phi_{\sigma(x_t)} \right) \left(\sum_{x'_t \to x'_{t+1}} \phi'_{\sigma(x'_{t+1})} - \mu' \phi'_{\sigma(x'_t)} \right) \right]$$

Since disjoint subsets of a Galton-Watson tree are independent, $E(x_t, x'_t)$ is zero except when $x_t = x'_t$. Letting $\sigma(x_t) = i$, we have as before

$$\begin{aligned}
E(x_t, x_t) &= \mathbb{E} \left[\left(\sum_{j \in [r]} Y_j \phi_j - \mu \phi_i \right) \left(\sum_{j \in [r]} Y_j \phi'_j - \mu' \phi'_i \right) \right] \\
&= \mathbb{E} \left[\left(\sum_{j \in [r]} (Y_j - M_{ij}) \phi_j \right) \left(\sum_{j \in [r]} (Y_j - M_{ij}) \phi'_j \right) \right] \\
&= \mathbb{E} \left[\sum_{j \in [r]} M_{ij} \phi_j \phi'_j \right] \\
&= [M(\phi \circ \phi')](i),
\end{aligned}$$

where we used the fact that $\mathbb{E}[Y_j] = \mathrm{Var}(Y_j) = M_{ij}$.

Now, we are in position to repeatedly apply Lemma 10.1:

$$\begin{aligned}\mathbb{E}[\Delta_t \Delta_t'] &= [\mu\mu']^{-(t+1)} \sum_{x_t \in \partial(T,\rho)_t} [M(\phi \circ \phi')](\sigma(x_t)) \\ &= [\mu\mu']^{-(t+1)} [M^{t+1}(\phi \circ \phi')](\sigma(\rho)),\end{aligned}$$

which completes the first part of the proof. Now, for $0 \leq t \leq t'$, since the increments are centered,

$$\begin{aligned}\mathbb{E}[Z_t Z_{t'}'] &= \mathbb{E}[Z_0 Z_0'] + \sum_{s=0}^{t-1} \mathbb{E}[\Delta_t \Delta_t'] \\ &= \phi_{\sigma(\rho)} \phi_{\sigma(\rho)}' + \sum_{s=0}^{t-1} [\mu\mu']^{-(t+1)} [M^{t+1}(\phi \circ \phi')](\sigma(\rho)),\end{aligned}$$

which corresponds to the second part of the proof. Finally, since the spectral radius of M is d, $\mathbb{E}[Z_t^2]$ is bounded whenever $\mu^2 > d$. By the Doob martingale convergence theorem, this implies that Z_t converges both almost surely and in $\mathcal{L}^2$. □

10.2 A top-down approach

Although the proofs in the previous section are necessary to obtain the martingale convergence properties, they can seem unnecessarily complex. We present here another "top-down" point of view, that recovers essentially the same results.

From [TODO], it is clear that the law of (T, ρ) only depends on the distribution of $\sigma(\rho)$. Thus, for a functional $f : \mathcal{G}_* \to \mathbb{R}$, we can define its Galton-Watson transform $\overline{f} \in \mathbb{R}^r$ as

$$\overline{f}(i) = \mathbb{E}_{\sigma(\rho)=i} f(T, \rho). \tag{10.3}$$

Although this transformation obviously loses a lot of the strucure of f, Proposition 9.2 indicates that $\overline{f}$ encapsulates all the necessary information about the action of f on G. In this framework, the results of the previous section are reformulated as follows:

Proposition 10.2. *Let (μ, ϕ) and (μ', ϕ') be two eigenpairs of M, and define $f_{\xi,t}$ as in* (10.1)*. Then*

$$\overline{f_{\phi,t}} = \mu^t \phi, \tag{10.4}$$

$$\overline{f_{\phi,t} f_{\phi',t}} = \sum_{s=0}^{t} [\mu\mu']^{t-s} M^s(\phi \circ \phi'), \tag{10.5}$$

and if $F_{\phi,t} = (f_{\phi,t+1} - \mu f_{\phi,t})^2$, then

$$\overline{F_{\phi,t}} = M^{t+1}(\phi \circ \phi') \tag{10.6}$$

The top-down transformation All tree functionals we have considered so far can be understood in terms of the following transformation ∂, defined for any functional $f : \mathcal{G}_* \to \mathbb{R}$:

$$\partial f(T, \rho) = \sum_{\rho' \sim \rho} f(T', \rho'), \tag{10.7}$$

where T' is the subtree of T rooted at ρ'. An important property is that all subtrees (T', ρ') are independent with the same distribution (depending on the spin of ρ'). This representation makes the transformation ∂ very convenient; in particular, Lemma 10.1 implies that for any functional $f : \mathcal{G}_* \to \mathbb{R}$,

$$\overline{\partial f} = M\overline{f}. \tag{10.8}$$

It is also easy to check that the functional $f_{\xi,t}$ defined in (10.1) satisfies the following recurrence relation:

$$f_{\xi,t+1}(g,o) = \partial f_{\xi,t}(g,o). \tag{10.9}$$

This allows to recover virtually all results of the previous section, without the martingale property. The correlations can be obtained through the formula (similarly easy to prove)

$$\overline{\partial f_1 \circ \partial f_2} = M\overline{f_1 \circ f_2} + \overline{\partial f_1} \circ \overline{\partial f_2}. \tag{10.10}$$

10.3 Pseudo-eigenvectors of B

We now explain how those estimates on trees can be translated to eigenvector equations on B. For the remainder of this section, we set

$$\ell = \lfloor \kappa \log_d(n) \rfloor, \tag{10.11}$$

for some constant κ that will be fixed later. For $i \in [r_0]$, define the pseudo-right (resp. left) eigenvectors u_i, v_i as

$$u_i = \frac{B^\ell \check{\chi}_i}{\sqrt{n}\mu_i^{\ell+1}} \quad \text{and} \quad v_i = \frac{(B^*)^\ell \chi_i}{\sqrt{n}\mu_i^\ell}. \tag{10.12}$$

We also let U (resp. V) be the $m \times r_0$ matrix whose columns are the u_i (resp. v_i). The matrix of pseudo-eigenvalues is simply

$$\Sigma = \mathrm{diag}(\mu_1, \dots, \mu_{r_0}). \tag{10.13}$$

Finally, we introduce the quantities $\gamma_i^{(t)}$ defined as

$$\gamma_i^{(t)} = \frac{1-\tau_i^{t+1}}{1-\tau_i} = \gamma_i + O(\tau_i^t), \tag{10.14}$$

where the limiting value γ_i is simply

$$\gamma_i = \frac{1}{1-\tau_i}. \tag{10.15}$$

We shall show the following proposition:

Proposition 10.3. *There exists a universal constant c and a choice of κ in* (10.11) *such that with probability at least $1 - cn^{-1/4}$ the following holds:*

$$\|U^*U - \mathrm{diag}(\{\tau_i \gamma_i^{(\ell)}\}_{i\in[r_0]})\| \le cn^{-1/4} \tag{10.16}$$

$$\|V^*V - \mathrm{diag}(\{d\gamma_i^{(\ell)}\}_{i\in[r_0]})\| \le cn^{-1/4} \tag{10.17}$$

$$\|U^*V - I_{r_0}\| \le cn^{-1/4} \tag{10.18}$$

$$\|V^*B^\ell U - \Sigma^\ell\| \le cn^{-1/4} \tag{10.19}$$

At first glance, it is not clear how those results relate to the eigenvalue decomposition of B (or B^ℓ). However, if we assume that U and V are orthogonal (i.e. that the RHS of (10.18) is zero), and define

$$S = U^*\Sigma^\ell V,$$

then:

- the eigenvalues of S are exactly the μ_i^ℓ,
- S satisfies exactly Equation (10.19).

In fact, Proposition 10.3 already implies that the eigenvalues of $P_{\mathrm{im}(U)}B^\ell P_{\mathrm{im}(V)}$ (where P_H denotes the orthogonal projection on a subspace H) are indeed close to the μ_i^ℓ. Studying the eigenvalues of B on the orthogonal of these subspaces requires completely different methods, and will be the focus of the next chapter.

We first give an overview of the arguments used. Let g be any graph, and $e = (e_1, e_2)$ an oriented edge in g. We define the graph functional

$$h_{\phi,t}(g,e) = \sum_{x_0=e_2,\dots,x_t} \phi_{\sigma(x_t)}, \tag{10.20}$$

where the sum ranges over all paths $\gamma = (x_0, \dots, x_t)$ such that (e_1, γ) is non-backtracking. The following facts are then immediate:

- by definition of B and $\check\chi_i$, for any $e \in \vec E$,
$$h_{\varphi_i,t}(G,e) = [B^t\check\chi_i](e) \tag{10.21}$$
- if $(g,e_1)_{t+1}$ is tangle-free, then there is at most two non-backtracking paths from e to any vertex x, hence
$$|h_{\varphi_i,t}(g,e)| \le |(g,e_1)_{t+1}| \tag{10.22}$$
- finally, consider a Galton-Watson tree (T,ρ) and $e = (\rho,\rho')$. Then
$$h_{\phi,t}(T,e) = f_{\phi,t}(T',\rho'), \tag{10.23}$$
where T' is the subtree of T whose root is ρ'.

We therefore have an explicit link between the graph functionals studied before and quantities of the form $B^t\check\chi_i$, which we shall use to show Proposition 10.3. Since all matrices involved are of size $r_0 \times r_0$, it is enough to show that the inequalities hold element by element. We show slightly stronger statements, and leave to the reader to check that these lemmas imply the desired inequalities.

Lemma 10.3. *Let ℓ as in* (10.11)*, with $\kappa < 1/24$. Then, with probability $1 - cn^{-1/4}$, for any $0 \le t \le 3\ell$ and $i,j \in [r]$,*

$$\left|\langle \chi_i, B^t\check\chi_j\rangle - n\mu_j^{t+1}\delta_{ij}\right| \le cn^{3/4}$$

Proof. Consider the graph functional

$$f(g,o) = \mathbf{1}_{(g,o)_{t+1}\text{ is tangle-free}}\,\varphi_j(\sigma(o)) \sum_{e:e_1=o} h_{\varphi_i,t}(g,e).$$

On the one hand, assuming that G is 3ℓ-tangle-free (which happens with probability $1 - n^{-1/4}$ as soon as $\kappa < 1/8$), we have

$$\sum_{x\in[n]} f(G,x) = \sum_{e\in\vec E} \varphi_i(e_1)[B^t\check\chi_j](e) = \langle \chi_i, B^t\check\chi_j\rangle.$$

On the other hand, we have

$$f(T,\rho) = \varphi_i(\sigma(\rho)) \sum_{\rho'\sim\rho} f_{\varphi_j,t}(T',\rho') = \varphi_i(\sigma(\rho)) f_{\varphi_j,t+1}(T,\rho).$$

Hence, for any $k \in [r]$,

$$\overline{f}(k) = \mu_j^{t+1}\varphi_i(k)\varphi_j(k).$$

Since f is obviously $t+1$-local, the bound (10.22) allows us to use Proposition 9.2: with probability at least $1 - n^{-1/4}$,

$$\left|\sum_{x\in[n]} f(G,x) - n\sum_{k\in[r]} \pi_k\overline{f}(k)\right| \le c\log(n)^{7/2}n^{6\kappa+1/2}.$$

It then remains to notice that by Equation [TODO],

$$\sum_{k\in[r]} \pi_k \varphi_i(k)\varphi_j(k) = \delta_{ij}.$$

□

This lemma implies the bound (10.18) (resp. (10.19)) by taking $t = 2\ell$ (resp. $t = 3\ell$). For the two other bounds, the proof is slightly more complicated, but follows the same principle:

Lemma 10.4. *Let ℓ as in* (10.11)*, with $\kappa < 1/12$. Then, with probability $1 - cn^{-1/4}$, for any $0 \le t \le \ell$ and $i, j \in [r]$,*

$$\left|\langle B^t \check{\chi}_i, B^t \check{\chi}_j\rangle - nd\mu_i^{2t}\gamma_i^{(t)}\right| \le cn^{3/4}.$$
$$\left|\langle (B^*)^t \chi_i, (B^*)^t \chi_j\rangle - nd\mu_i^{2t}\gamma_i^{(t)}\right| \le cn^{3/4}.$$

Proof. We first show how the first inequality implies the second: since $P^2 = I$ and $P = P^*$, we have

$$\begin{aligned}\langle (B^*)^t \chi_i, (B^*)^t \chi_j\rangle &= \langle P(B^*)^t \chi_i, P(B^*)^t \chi_j\rangle \\ &= \langle B^t P\chi_i, B^t P\chi_j\rangle \\ &= \langle B^t \check{\chi}_i, B^t \check{\chi}_j\rangle,\end{aligned}$$

where we used the parity-time symmetry of P. Hence, it suffices to prove the first inequality. Define this time the functional

$$f(g,o) = \mathbf{1}_{(g,o)_{t+1} \text{ is tangle-free}} \sum_{e:e_1=o} h_{\varphi_i,t}(g,e)h_{\varphi_j,t}(g,e)$$

It is easy to check that both

$$\sum_{x\in[n]} f(G,x) = \langle B^t \check{\chi}_i, B^t \check{\chi}_j\rangle$$

and

$$f(T,\rho) = \partial[f_{\varphi_i,t} f_{\varphi_j,t}](T,\rho),$$

where ∂ is the operator defined in (10.7). From Lemma 10.8 and Proposition 10.2, we deduce

$$\overline{f} = \sum_{s=0}^{t} [\mu_i\mu_j]^{t-s} M^{s+1}(\varphi_i \circ \varphi_j).$$

In order to apply Proposition 9.2 (with $\beta = 2$), we need to compute $\langle \pi, \overline{f}\rangle$. Since π is a left eigenvector of M with eigenvalue d, we have

$$\langle \pi, \overline{f}\rangle = \langle \pi, \varphi_i \circ \varphi_j\rangle \sum_{s=0}^{t} [\mu_i\mu_j]^{t-s} d^{s+1}.$$

The first term is equal to δ_{ij} as above, and the second is a geometric sum, which when $i = j$ is equal to $d\mu_i^{2t}\gamma_i^{(t)}$. This completes the proof. □

Chapter 11

Tangle-free decomposition and the trace method

The methods we have developed so far gave us a set of approximate right (resp. left) eigenvectors of B^ℓ, the u_i (resp. v_i) for $i \in r_0$. It remains to bound all of the $|\vec{E}| - r_0$ eigenvalues to apply perturbation bounds. We will do so using the trace method already encountered in Chapter 2. This is an intuitive choice, since B^ℓ_{ef} already counts the number of non-backtracking paths between e and f. However, several problems arise:

- since B is of size $m \times m$, which is random, finding a suitable centered version of B is non-trivial;
- we need to bound only the eigenvalues in $\mathrm{im}(U)^\perp$, which itself depends on B^ℓ;
- the path counting estimates of Proposition 2.1 are too coarse for the bounds we need.

Hence, we need a preprocessing step on B^ℓ before applying the trace method: this is known as the *tangle-free decomposition*.

11.1 Tangle-free decomposition of B^ℓ

Define the set of oriented edges of the complete graph $\vec{E}(V)$ as

$$\vec{E}(V) = \{(x, y) \mid x, y \in V, x \neq y\}.$$

We can consider B as a matrix in $\vec{E}(V)$-space by setting

$$B_{ef} = A_e A_f \mathbf{1}_{e \to f}.$$

This embedding preserves the spectral properties of B; by abuse of notation, we use the same notation for B when viewed as a matrix in $\vec{E}$ and $\vec{E}(V)$. We can extend the vectors $\chi_i, \check{\chi}_i$ accordingly:

$$\chi_i(e) = \varphi_i(\sigma(e_1)) \quad \text{and} \quad \check{\chi}_i(e) = \varphi_i(\sigma(e_2)).$$

For $k \in \mathbb{N}$, $e, f \in \vec{E}(V)$, we let $\mathcal{F}^k_{ef}$ be the set of tangle-free paths of length k from e to f. Then, when G is ℓ-tangle-free, for any $k \leq \ell$ we have $B^k = B^{(k)}$, where

$$B^{(k)}_{ef} = \sum_{x \in \mathcal{F}^{k+1}_{ef}} \prod_{t=0}^{k} A_{x_t x_{t+1}} \tag{11.1}$$

In order to get the zero-mean property needed for the trace method, we define the centered version of A:

$$\underline{A}_{xy} = A_{xy} - \frac{W_{\sigma(x)\sigma(y)}}{n},$$

from where we can define $\Delta^{(k)}$ as

$$\Delta^{(k)}_{ef} = \sum_{x \in \mathcal{F}^{k+1}_{ef}} \prod_{t=0}^{k} \underline{A}_{x_t x_{t+1}} \tag{11.2}$$

We link $B^{(\ell)}$ and $\Delta^{(\ell)}$ using the following identity: for any $a, b \in \mathbb{R}^\ell$,

$$\prod_{t=0}^{\ell} a_t = \prod_{t=0}^{\ell} b_t + \sum_{k=0}^{\ell} \prod_{t=0}^{k-1} b_t \, (a_k - b_k) \prod_{t=k+1}^{\ell} a_t,$$

with the convention that an empty product is equal to 1. Applying this identity to the products in (11.1) and (11.2),

$$B^{(\ell)}_{ef} = \Delta^{(\ell)}_{ef} + \sum_{k=0}^{\ell} \sum_{x \in \mathcal{F}^{k+1}_{ef}} \prod_{t=0}^{k-1} \underline{A}_{x_t x_{t+1}} \left(\frac{W_{\sigma(x_k)\sigma(x_{k+1})}}{n} \right) \prod_{t=k+1}^{\ell} A_{x_t x_{t+1}} \tag{11.3}$$

Define the matrix $K^{(2)}$ as

$$K^{(2)}_{ef} = \mathbf{1}_{e \overset{(2)}{\to} f} W_{\sigma(e_2)\sigma(f_1)},$$

where $e \overset{(2)}{\to} f$ means that (e_1, e_2, f_1, f_2) is non-backtracking. Then, the middle sum of (11.3) is very similar to $\Delta^{(k-1)} K^{(2)} B^{(\ell-k-1)}$, with the following caveat: the concatenation of two tangle-free paths is not necessarily tangle-free! To remediate this, we let $\mathcal{F}^{\ell+1}_{k,ef}$ be the set of non-backtracking paths $(x_0, \dots, x_{\ell+1})$ such that:

- $(x_0, \dots, x_k)$ and $(x_{k+1}, \dots, x_{\ell+1})$ are both tangle-free
- $(x_0, \dots, x_{\ell+1})$ is *not* tangle-free.

Then, for $1 \le k \le \ell - 1$, we have

$$\sum_{x \in \mathcal{F}^{\ell+1}_{ef}} \prod_{t=0}^{k-1} \underline{A}_{x_t x_{t+1}} W_{\sigma(x_k)\sigma(x_{k+1})} \prod_{s=tk+1}^{\ell} A_{x_t x_{t+1}} = [\Delta^{(k-1)} K^{(2)} B^{(\ell-k-1)}]_{ef} - [R^{(\ell)}_k]_{ef},$$

where $R^{(\ell)}_k$ is simply

$$R^{(\ell)}_k = \sum_{x \in \mathcal{F}^{\ell+1}_{k,ef}} \prod_{t=0}^{k-1} \underline{A}_{x_t x_{t+1}} W_{\sigma(x_k)\sigma(x_{k+1})} \prod_{t=k+1}^{\ell} A_{x_t x_{t+1}}.$$

Finally, to deal with the case $k = 0$ and $k = \ell$, we introduce

$$K_{ef} = \mathbf{1}_{e \to f} W_{\sigma(e_1)\sigma(e_2)} \quad \text{and} \quad K'_{ef} = \mathbf{1}_{e \to f} W_{\sigma(f_1)\sigma(f_2)}$$

We can then write

$$B^{(\ell)} = \Delta^{(\ell)} + \frac{1}{n} K B^{(\ell-1)} + \frac{1}{n} \sum_{k=1}^{\ell-1} \Delta^{(k-1)} K^{(2)} B^{(\ell-k-1)} + \frac{1}{n} \Delta^{(\ell-1)} K' - \frac{1}{n} \sum_{k=1}^{\ell-1} R^{(\ell)}_k. \tag{11.4}$$

The last step is to understand the matrix $K^{(2)}$. Let

$$\bar{W} = \sum_{i=1}^{r} \mu_i \check{\chi}_i \chi_i^* \quad \text{and} \quad L = K^{(2)} - \bar{W},$$

then it is easy to check that

$$\bar{W}_{ef} = W_{\sigma(e_2)\sigma(f_1)},$$

and hence L_{ef} is zero except if $e_1 = f_1$, $e_2 = f_1$ or $e_2 = f_2$. We thus define the (hopefully small) matrices

$$S_k^{(\ell)} = \Delta^{(k-1)} L B^{(\ell-k-1)}.$$

All of the discussion of this section can be summarized in the following proposition:

Proposition 11.1. *Let $\ell \geq 0$, and assume that the graph G is ℓ-tangle-free. Then, for any x in $\mathbb{R}^{\vec{E}(V)}$ such that $\|x\| = 1$, we have*

$$\begin{aligned}\left\|B^\ell x\right\| \leq \left\|\Delta^{(\ell)}\right\| + \frac{1}{n}\left\|KB^{(\ell-1)}\right\| + \frac{d}{n}\sum_{k=1}^{\ell-1}\sum_{i=1}^{r}\left\|\Delta^{(k-1)}\check{\chi}_i\right\| \left|\langle \chi_i, B^{\ell-k-1}x\rangle\right| \\ + \frac{1}{n}\sum_{k=1}^{\ell-1}\|S_t^{(\ell)}\| + \frac{1}{n}\left\|\Delta^{(\ell-1)}K'\right\| + \frac{1}{n}\sum_{k=1}^{\ell-1}\left\|R_t^{(\ell)}\right\|.\end{aligned} \tag{11.5}$$

Before moving on, we expand on why we expect each term in (11.5) to be small:

1. $\|\Delta^{(k)}\|$, $\|\Delta^{(k)}\chi_i\|$ and $\|R_k^{(\ell)}\|$ involve matrices with zero-mean random elements, so their norm is bounded by the trace method;

2. since the matrices K and L are very sparse with bounded entries, the norms of $S_k^{(\ell)}$ and $KB^{(\ell-1)}$ are also bounded;

3. finally, if $x \in \mathrm{im}(V)^\perp$, then
$$\langle (B^*)^\ell \chi_i, x\rangle = 0.$$
Since $(B^*)^\ell\chi_i$ is an approximate eigenvector of B^*, $\langle (B^*)^{\ell-t-1}\chi_i, x\rangle$ remains small enough.

Remark. *We shall only prove Equation TODO in Proposition TODO, bounding $B^\ell P_{\mathrm{im}(V)^\perp}$. The bound for $\mathrm{im}(U)^\perp$ is almost identical; reversing the role of A and $\underline{A}$ in* (11.3) *and multiplying par x on the left yields an expression almost identical to* (11.5)*, but depending on $\langle x, B^{\ell-t-1}\check{\chi}_i\rangle$ instead. The rest of the proof proceeds identically.*

11.2 Non-backtracking trace method

The goal of this section if to prove the following proposition:

Proposition 11.2. *Let $\ell \sim \kappa \log_d(n)$ with $\kappa > 0$. There exist constants c, s such that for n large enough, with probability at least $1 - n^{-s}$, the following bounds hold for $0 \leq k \leq \ell$ and $i \in [r]$:*

$$\left\|\Delta^{(k)}\right\| \leq \log(n)^c d^{k/2} \tag{11.6}$$

$$\left\|\Delta^{(k)}\check{\chi}_i\right\| \leq \log(n)^c \sqrt{n} d^{k/2} \tag{11.7}$$

$$\left\|R_k^{(\ell)}\right\| \leq \log(n)^c d^{\ell-k/2} \tag{11.8}$$

$$\|S_k^{(\ell)}\| \leq \sqrt{n}\log(n)^c d^{\ell-k/2} \tag{11.9}$$

$$\left\|B^{(k)}\right\| \leq \log(n)^c d^k \tag{11.10}$$

$$\left\|KB^{(k)}\right\| \leq \log(n)^c \sqrt{n} d^k \tag{11.11}$$

We shall only prove two of these bounds, namely (11.6) and (11.10). These will give some good examples of a more involved trace method for both zero-mean and nonzero-mean matrices. Note that the bounds (11.7) and (11.11) are significant improvements on the naive bounds of the form

$$\left\|\Delta^{(k)}\check{\chi}_i\right\| \le \|\Delta^{(k)}\|\|\check{\chi}_i\| \quad \text{and} \quad \left\|KB^{(k)}\right\| \le \|K\|\|B^{(k)}\|$$

which give a linear dependence in n in both cases.

11.2.1 Basics of the trace method

We start with $\Delta^{(k)}$. Since this matrix is not symmetric, the corresponding version of (2.1) is the following: for any $m \ge 0$,

$$\|\Delta^{(k)}\|^{2m} \le \operatorname{Tr}\left[\left(\Delta^{(k)}\Delta^{(k)*}\right)^m\right] \tag{11.12}$$

We can now expand the right-hand side, using the convention $e_{2m+1} = e_1$ and the symmetry condition $\left(\Delta^{(k)}\right)_{ef} = \left(\Delta^{(k)}\right)_{f^{-1}e^{-1}}$:

$$\begin{aligned}
\|\Delta^{(k)}\|^{2m} &\le \sum_{e_1,\dots,e_{2m}} \prod_{i=1}^{m} \left(\Delta^{(k)}\right)_{e_{2i-1},e_{2i}} \left(\Delta^{(k)}\right)_{e_{2i+1},e_{2i}} \\
&= \sum_{e_1,\dots,e_{2m}} \prod_{i=1}^{m} \left(\Delta^{(k)}\right)_{e_{2i-1},e_{2i}} \left(\Delta^{(k)}\right)_{e_{2i}^{-1},e_{2i+1}^{-1}} \\
&= \sum_{\gamma\in W_{k,m}} \prod_{i=1}^{2m}\prod_{t=1}^{k} \underline{A}_{\gamma_{i,t-1}\gamma_{i,t}},
\end{aligned} \tag{11.13}$$

where $W_{k,m}$ is the set of sequences of paths $\gamma = (\gamma_1, \dots, \gamma_{2m})$ such that, for all $i \in [2m]$, $\gamma_i = (\gamma_{i,0}, \dots, \gamma_{i,k}) \in [n]^{k+1}$ is non-backtracking tangle-free, and we have

$$(\gamma_{i,k-1}, \gamma_{i,k}) = (\gamma_{i+1,1}, \gamma_{i+1,0}), \tag{11.14}$$

again with the convention that $\gamma_{2m+1} = \gamma_1$. Note that the same inequality also holds for B^k, becoming

$$\|B^k\|^{2m} \le \sum_{\gamma\in W_{k,m}} \prod_{i=1}^{2m}\prod_{t=1}^{k} A_{\gamma_{i,t-1}\gamma_{i,t}} \tag{11.15}$$

To any path sequence $\gamma \in W_{k,m}$, we associate the undirected graph $G_\gamma = (V_\gamma, E_\gamma)$ with

$$V_\gamma = \{\gamma_{i,t} \mid 1 \le i \le 2m, 0 \le t \le k\} \quad \text{and} \quad E_\gamma = \{\{\gamma_{i,t-1}, \gamma_{i,t}\} \mid 1 \le i \le 2m, 1 \le t \le k\}$$

We now group the sequences in $W_{k,m}$ according to their topological properties. We say that two sequences γ and γ' are equivalent, and write $\gamma \sim \gamma'$, if there exists a permutation $\mathfrak{s} \in \mathfrak{S}_n$ such that $\gamma' = \mathfrak{s} \circ \gamma$, where

$$(\mathfrak{s} \circ \gamma)_{i,t} = \mathfrak{s}(\gamma_{i,t}).$$

It is straightforward to check that $\sim$ is an equivalence relation, which partitions $W_{k,m}$ in disjoint classes, and that $|V_\gamma|$ and $|E_\gamma|$ are constant on each class. For $s, a \ge 0$, we thus define $\mathcal{W}_{k,m}(s,a)$ the set of equivalence classes $[\gamma] \in W_{k,m}/\sim$ such that $|V_\gamma| = s$ and $|E_\gamma| = a$. The graphs G_γ are necessarily connected by the boundary condition (11.14), hence $\mathcal{W}_{k,m}(s,a)$ is nonempty only if $a - s + 1 \ge 0$. In order to bound the sum in (11.13), we therefore need the following steps:

- bound the size of $\mathcal{W}_{k,m}(s,a)$,
- bound the contribution of a single equivalence class $[\gamma] \in \mathcal{W}_{k,m}(s,a)$.

11.2.2 Bounding $|\mathcal{W}_{k,m}(s,a)|$

We show the following proposition:

Proposition 11.3. *Let $s, a \geq 0$ with $a - s + 1 \geq 0$. Then*

$$|\mathcal{W}_{k,m}(s,a)| \leq (2km)^{6m(a-s+1)+2m}.$$

Since the proof of this result is quite involved, we break it into several steps.

Preliminaries Throughout this section, we will consider a path sequence γ as a single continuous path indexed $T = \{(i,t) \mid 1 \leq i \leq 2m, 0 \leq t \leq k\}$ with the lexicographic ordering. We define the *canonical* representative of a class $[\gamma'] \in \mathcal{W}_{k,m}(s,a)$ as the only path $\gamma \in [\gamma']$ such that $V_\gamma = \{1, \dots, s\}$ and the vertices of V_γ are visited in order. We leave to the reader to check that each class indeed contains exactly one such path.Therefore, bounding $|\mathcal{W}_{k,m}(s,a)|$ is equivalent to bounding the number of canonical paths. In turn, such an upper bound is equivalent to finding an injective encoding of the set of canonical paths.

For $(i,t) \in T$, we let $e_{i,t} = \gamma_{i,t}, \gamma_{i,t+1}$. If $\gamma_{i,t+1}$ is visited for the first time, we say that $e_{i,t}$ is a *tree edge*. It is easy to check that the set of tree edges indeed forms a tree $\mathcal{T}$; any edge not belonging to $\mathcal{T}$ is an *excess edge*. The number of such edges is $g = a - s + 1$.

A first encoding Each path γ_i can be subdivided into sections with the following structure:

1. a (possibly empty) sequence of *tree times*, where $e_{i,t}$ is an already discovered edge of $\mathcal{T}$,
2. a (possibly empty) sequence of *first times*, where $\gamma_{i,t+1}$ is an undiscovered vertex at time (i,t),
3. and finally a *cycling time*, where $e_{i,t}$ is an excess edge.

We encode each of this section with the following information: the first and last vertices (x_0, x_1) of the sequence of tree times, and the cycling time τ (if the last section does not have an excess edge, we let $\tau = k$). This information allows us to recover the entire section, for the following reasons:

- there is a unique non-backtracking path between vertices x_0 and x_1 on $\mathcal{T}$, so the sequence of tree times is uniquely determined,
- the number of first times is $\tau - (\tau' + 1 + d_{\mathcal{T}}(x_0, x_1))$, where τ' is the previous cycling time, and the vertices are visited in order,
- and finally the excess edge is (x_2, x_0'), where x_2 is the end of the sequence of first times and x_0' is the first mark of the next cycling time.

This yields a first encoding of canonical paths, using $s^2 k$ possibilities for each cycling time.

Refining the encoding The issue with this first encoding is that the number of cycling times is potentially very high. However, we haven't used the fact that each path γ_i is tangle-free. First consider the case where γ_i contains no cycles; then each excess edge is used at most once si doing otherwise would create a cycle; hence the number of cycling times is at most g.

Otherwise, let t_1 be the first time t where $\gamma_{i,t} \in \{\gamma_{i,0}, \dots \gamma_{i,t-1}\}$, and $t_0 \leq t_1$ such that $\gamma_{i,t_1} = \gamma_{i,t_0}$. Then the cycle $(\gamma_{i,t_0}, \dots, \gamma_{i,t_1})$ is the only cycle formed by γ_i, call it $\mathcal{C}$. We add the *exit* mark t_2 to γ_i, which is the first time $t \geq t_1$ such that $e_{i,t}$ is not an edge of $\mathcal{C}$. The cycling times t with $0 \leq t \leq t_1$ or $t_2 \leq t \leq k$ are called *important imes*; other times are called *superfluous times*. Then:

- no two important times can share an edge, since otherwise it would create a second cycle in γ_i,
- we do not need to mark the superfluous times, since the path γ_i is simply cycling around $\mathcal{C}$ between t_1 and t_2.

As a result, the number of marks is at most g per path, with s^2k choices for them, plus a possible additional cycling mark with s choices. This yields the bound

$$|\mathcal{W}_{k,m}(s,a)| \leq s^{2m}(s^2k)^{2mg}, \tag{11.16}$$

which is much lower than the one in Proposition 11.3.

11.2.3 Bounding the contribution of a path class

We now fix a path class $[\gamma] \in \mathcal{W}_{k,m}(s,a)$, and define its *contribution* $c_{[\gamma]}$ as

$$c_{[\gamma]} = \frac{1}{n^s} \sum_{\mathfrak{s}} \mathbb{E}\left[\prod_{i=1}^{2m} \prod_{t=1}^{k} \underline{A}_{\mathfrak{s}(\gamma_{i,s-1})\mathfrak{s}(\gamma_{i,s})}\right],$$

where $\mathfrak{s} : [s] \to [n]$ ranges over all possible injections from V_γ to $[n]$. The goal is to show the following:

Proposition 11.4. *For any $[\gamma] \in \mathcal{W}_{k,m}(s,a)$, we have*

$$c_{[\gamma]} \leq \left(\frac{d}{n}\right)^{s-1} \left(\frac{p_{\max}}{n}\right)^{a-s+1}, \tag{11.17}$$

where $p_{\max}$ is an upper bound on the entries of P.

Proof. We first rewrite this sum as

$$c_{[\gamma]} = \frac{1}{n^s} \sum_{\mathfrak{s}} \mathbb{E}\left[\prod_{e \in E_\gamma} \underline{A}^{m_e}_{\mathfrak{s}(e_1)\mathfrak{s}(e_2)}\right],$$

where m_e is the multiplicity of an edge e in G_γ. In particular, note that $c_{[\gamma]}$ is zero except when $m_e > 1$ for all edges in E_γ. Using the inequality

$$\mathbb{E}\big[\underline{A}^m_{xy}\big] \leq \frac{P_{\sigma(x)\sigma(y)}}{n} \tag{11.18}$$

for all x, y and the independence of the entries of $\underline{A}$, we have

$$c_{[\gamma]} = \frac{1}{n^s} \sum_{\mathfrak{s}} \prod_{e \in E_\gamma} \frac{P_{\sigma\circ\mathfrak{s}(e_1),\sigma\circ\mathfrak{s}(e_2)}}{n}.$$

Now, fix a spanning tree $\mathcal{T}$ of G_γ. For $e \notin \mathcal{T}$, we simply use the bound $P_{ij} \leq p_{\max}$. Let x be a leaf of $\mathcal{T}$, and y the only vertex connected to x. For any injection $\mathfrak{s}$, define $\mathfrak{s}_{-x}$ the restriction of s to $[k] \setminus x$. Then, we can decompose

$$\sum_{\mathfrak{s}} \prod_{e \in E_\gamma} \frac{P_{\sigma\circ\mathfrak{s}(e_1),\sigma\circ\mathfrak{s}(e_2)}}{n} = \sum_{\mathfrak{s}_{-x}} \prod_{e \in \mathcal{T}\setminus\{x,y\}} \frac{P_{\sigma\circ\mathfrak{s}_{-x}(e_1),\sigma\circ\mathfrak{s}_{-x}(e_2)}}{n} \cdot \sum_{\mathfrak{s}(x)} \frac{P_{\sigma\circ\mathfrak{s}_{-x}(y),\sigma\circ\mathfrak{s}(x)}}{n}, \tag{11.19}$$

where the second sum ranges over $\mathfrak{s}(x) \notin \mathrm{im}(\mathfrak{s}_{-x})$. Letting $i = \sigma \circ \mathfrak{s}_{-x}(y)$,

$$\sum_{\mathfrak{s}(x)} \frac{P_{\sigma\circ\mathfrak{s}_{-x}(y),\sigma\circ\mathfrak{s}(x)}}{n} \leq \sum_{\mathfrak{s}(x)\in[n]} \frac{P_{\sigma\circ\mathfrak{s}_{-x}(y),\sigma\circ\mathfrak{s}(x)}}{n} \leq \sum_{j\in[r]} \frac{P_{ij}n_j}{n} = d.$$

Note that the sum

$$\sum_{\mathfrak{s}_{-x}} \prod_{e \in \mathcal{T}\setminus\{x,y\}} \frac{P_{\sigma\circ\mathfrak{s}_{-x}(e_1),\sigma\circ\mathfrak{s}_{-x}(e_2)}}{n}$$

is the exact equivalent of the LHS of (11.19) for the tree $T \setminus x$. As a result, an immediate recursion yields

$$\sum_{\mathfrak{s}(x)} \frac{P_{\sigma\circ\mathfrak{s}_{-x}(y),\sigma\circ\mathfrak{s}(x)}}{n} \leq nd^{s-1}, \tag{11.20}$$

where the factor of n corresponds to the image of the tree root. Proposition 11.4 then ensues from dividing by n^s on both sides. □

Remark. *Since Eq.* (11.18) *also holds for the matrix A, the contribution for a path class $[\gamma]$ to the bound on $\|B^k\|^{2m}$ also satisfies Proposition 11.4.*

11.2.4 Wrapping up the trace method

Now that we have both of the needed ingredients, we are ready to bound $\mathbb{E}[\|\Delta^{(k)}\|^{2m}]$. We write

$$\begin{aligned}\mathbb{E}[\|\Delta^{(k)}\|^{2m}] &\leq \sum_{\gamma\in W_{k,m}} \mathbb{E}\left[\prod_{i=1}^{2m}\prod_{t=1}^{k} \underline{A}_{\gamma_{i,t-1}\gamma_{i,t}}\right] \\ &\leq \sum_{s,a} \sum_{[\gamma]\in\mathcal{W}_{k,m}(s,a)} n^s c_{[\gamma]} \\ &\leq \sum_{s,a} (2km)^{6m(a-s+1)+2m} \left(\frac{d}{n}\right)^{s-1} \left(\frac{p_{\max}}{n}\right)^{s-a+1} n^s \end{aligned} \tag{11.21}$$

We just need to determine the bounds on s, a. The crucial remark is that since each edge of G_γ has to be visited twice for c_γ to be nonzero, we have $s-1 \leq a \leq km$, and hence $0 \leq a-s+1 \leq km$. This gives

$$\begin{aligned}\mathbb{E}[\|\Delta^{(k)}\|^{2m}] &\leq n(2km)^{2m} \sum_{s=2}^{km+1} d^{s-1} \sum_{g=0}^{km} \left(\frac{(2km)^{6m}}{n}\right)^g \\ &\leq n(2km)^{2m} kmd^{km} \sum_{g=0}^{\infty} \left(\frac{(2km)^{6m}}{n}\right)^g \end{aligned}$$

We now choose

$$m = \frac{c\log(n)}{\log(\log(n))};$$

for a sufficiently small choice of c, it is easy to check that

$$n^{1/(2m)} \leq \log(n)^{1/(2c)}, \quad 2km \leq \log(n)^2 \quad \text{and} \quad (2km)^{6m} = o(n).$$

In particular, the above geometric series converges, and (11.6) ensues via a Markov bound.

Remark. *The bound on B is almost identical: Eq.* (11.21) *still holds, but for different bounds on s, a. Indeed, each edge does not need to be visited twice, so the weaker $s-1 \leq a \leq 2km$ holds instead. One can then easily check that this bound on s implies* (11.10)*.*

11.3 An approximate eigenvector equation

The goal of this section is to deal with the term $\langle \chi_i, B^{\ell-k-1}x\rangle$ in Proposition 11.1. In view of what has already been proven, the following bound will be sufficient:

Proposition 11.5. *Let ℓ as in* (10.11)*, with $\kappa < 1/12$. With probability $1-n^{-\varepsilon}$, we have for any $0 \leq t \leq \ell$, $i \in [r]$ and $x \in \mathrm{im}(V)^\perp$:*

$$\langle \chi_i, B^t x\rangle \leq c\log(n)^c \sqrt{n}\, d^{t/2} \tag{11.22}$$

The case $i > r_0$ is immediate, and can be treated directly: by Lemma 10.4 and the triangular inequality,

$$\begin{aligned}\langle \chi_i, B^t x\rangle &\le \|(B^*)^t \chi_i\| \\ &\le \sqrt{n\mu_i^{2(t)}\gamma_i^{(t)} + cn^{3/4}} \\ &\le c\sqrt{n}\, d^{t/2},\end{aligned}$$

having used the fact that $\mu_i^{2t}\gamma_i^{(t)} \le d^t$ whenever $\mu_i < \sqrt{d}$.

For $i \in [r_0]$, the situation is different since $\mu_i^{2t} \gg d^t$. However, we can write the telescopic sum

$$\begin{aligned}\mu_i^{-t}\langle \chi_i, B^t x\rangle &= \mu_i^{-t}\langle \chi_i, B^t x\rangle - \mu_i^{-\ell}\langle \chi_i, B^\ell x\rangle \\ &= \sum_{s=t}^{\ell-1} \mu_i^{-s}\langle \chi_i, B^s x\rangle - \mu_i^{-(s+1)}\langle \chi_i, B^{s+1} x\rangle \\ &\le \sum_{s=t}^{\ell-1} |\mu_i|^{-(s+1)} \cdot \|B^{s+1}\check{\chi}_i - \mu_i B^s \check{\chi}_i\|,\end{aligned}$$

where in the first line we used that $\langle \chi_i, B^\ell x\rangle = 0$. Our aim is thus to show an appxroximate eigenvector equation on $\check{\xi}$. Specifically, we show the following bound:

Lemma 11.1. *Let ℓ as in (10.11), with $\kappa < 1/12$. With probability at least $1 - n^{-\epsilon}$, for any $i \in [r_0]$ and $0 \le t \le \ell$,*

$$\left\|B^{s+1}\check{\chi}_i - \mu_i B^s \check{\chi}_i\right\| \le TODO$$

Proof. Define the functional

$$f(g, o) = \sum_{e : e_1 = o} \left(h_{\phi_i, t+1}(g, e) - \mu_i h_{\phi_i, t}(g, e)\right)^2, \tag{11.23}$$

where we recall that $h_{\phi_i,t}$ is defined in (10.20). □

Bibliography

[1] N. Alon, M. Krivelevich, and B. Sudakov. Finding a large hidden clique in a random graph. *Random Structures and Algorithms*, 13(3-4):457–466, 1998.

[2] N. Alon and A. Naor. Approximating the cut-norm via grothendieck's inequality. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, pages 72–80, 2004.

[3] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[4] Z. Bai and J. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*. Springer, 2010.

[5] J. Banks, C. Moore, J. Neeman, and P. Netrapalli. Information-theoretic thresholds for community detection in sparse networks. In *Proceedings of the 29th Conference on Learning Theory, COLT 2016, New York, USA, June 23-26, 2016*, pages 383–416, 2016.

[6] A. D. Barbour and L. H. Y. Chen, editors. *An introduction to Stein's method*, volume 4 of *Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore*. Singapore University Press, Singapore, 2005.

[7] F. Benaych-Georges and R. R. Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 227(1):494 – 521, 2011.

[8] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electronic J. Probab.*, 6:–, 2001.

[9] C. Bordenave, M. Lelarge, and L. Massoulié. Nonbacktracking spectrum of random graphs: Community detection and nonregular ramanujan graphs. *Ann. Probab.*, 46(1):1–71, 01 2018.

[10] S. Boucheron, G. Lugosi, and P. Massart. On concentration of self-bounding functions. *Electronic Journal of Probability*, (14):1884–1899, 2009.

[11] A. Coja-Oghlan, F. Krzakala, W. Perkins, and L. Zdeborova. Information-theoretic thresholds from the cavity method. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 146–157, New York, NY, USA, 2017. ACM.

[12] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová. Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Phys. Rev. E*, 84:066106 (1–19), Dec 2011.

[13] A. Dembo and A. Montanari. Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.*, 24(2):137–211, 07 2010.

[14] W. Evans, C. Kenyon, Y. Peres, and L. J. Schulman. Broadcasting on trees and the ising model. *The Annals of Applied Probability*, 10(2):410–433, 2000.

[15] U. Feige and E. Ofek. Spectral techniques applied to sparse random graphs. *Random Struct. Algorithms*, 27(2):251–275, Sept. 2005.

[16] Z. Füredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, (1(3)):233–241, 1981.

[17] L. Gulikers, M. Lelarge, and L. Massoulié. An impossibility result for reconstruction in the degree-corrected stochastic block model. *Ann. Appl. Probab.*, 28(5):3002–3027, 10 2018.

[18] M. D. Horton, H. M. Stark, and A. A. Terras. What are zeta functions of graphs and what are they good for? *Contemporary Mathematics, Quantum Graphs and Their Applications; Edited by Gregory Berkolaiko, Robert Carlson, Stephen A. Fulling, and Peter Kuchment*, 415:173–190, 2006.

[19] S. Janson and E. Mossel. Robust reconstruction on trees is determined by the second eigenvalue. *The Annals of Probability*, 32(3B):2630–2649, 2004.

[20] A. Javanmard and A. Montanari. State evolution for general approximate message passing algorithms, with applications to spatial coupling. *CoRR*, abs/1211.5164, 2012.

[21] T. Kato. *Perturbation Theory for Linear Operators*. Springer, 1966.

[22] H. Kesten and B. P. Stigum. Additional limit theorems for indecomposable multidimensional galton-watson processes. *The Annals of Mathematical Statistics*, pages 1463–1481, 1966.

[23] H. Kesten and B. P. Stigum. Additional limit theorems for indecomposable multidimensional Galton-Watson processes. *Ann. Math. Statist.*, (37):1463–1481, 1966.

[24] V. I. Koltchinskii. Asymptotics of spectral projections of some random matrices approximating integral operators. In E. Eberlein, M. Hahn, and M. Talagrand, editors, *High Dimensional Probability*, pages 191–227, Basel, 1998. Birkhäuser Basel.

[25] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, and P. Zhang. Spectral redemption in clustering sparse networks. *Proceedings of the National Academy of Sciences*, 110(52):20935–20940, 2013.

[26] T. C. Kwok, L. C. Lau, Y. T. Lee, S. O. Gharan, and L. Trevisan. Improved cheeger's inequality: analysis of spectral partitioning algorithms through higher order spectral gap. In *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 11–20, 2013.

[27] A. Lubotzky. Cayley graphs: eigenvalues, expanders and random walks. *Surveys in Combinatorics, London Math. Soc. Lecture Notes*, 218:155–189, 1995.

[28] L. Massoulié and D. Tomozei. Distributed user profiling via spectral methods. *Stochastic Systems*, 4:1–43, 2014.

[29] M. Mézard and A. Montanari. Reconstruction on trees and spin glass transition. *Journal of Statistical Physics*, 2006.

[30] B. Mohar. Some applications of laplace eigenvalues of graphs. In *Graph symmetry*, pages 225–275. Springer, 1997.

[31] B. Mohar. Some applications of laplace eigenvalues of graphs. In *GRAPH SYMMETRY: ALGEBRAIC METHODS AND APPLICATIONS, VOLUME 497 OF NATO ASI SERIES C*, pages 227–275. Kluwer, 1997.

[32] B. Mohar. A strengthening and a multipartite generalization of the alon-boppana-serre theorem. *Proc. Amer. Math. Soc.*, 138:3899–3909, 2010.

[33] C. Moore. The computer science and physics of community detection: Landscapes, phase transitions, and hardness. *Bulletin of the EATCS*, 121, 2017.

[34] E. Mossel, J. Neeman, and A. Sly. Reconstruction and estimation in the planted partition model. *Probability Theory and Related Fields*, 162(3-4):431–461, 2015.

[35] E. Mossel and Y. Peres. Information flow on trees. *The Annals of Probability*, 13(3):817–844, 2003.

[36] R. Murty. Ramanujan graphs. *J. Ramanujan Math. Soc.*, 18(1):1–20, 2003.

[37] M. Mézard and A. Montanari. *Information, Physics, and Computation*. Oxford University Press, Inc., New York, NY, USA, 2009.

[38] S. Péché and A. Soshnikov. Wigner random matrices with non symmetrically distributed entries. *J. Stat. Phys.*, 129:857–883, 2007.

[39] F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová. Typology of phase transitions in bayesian inference problems. *CoRR*, abs/1806.11013, 2018.

[40] A. Sly. Reconstruction for the potts model. *Ann. Probab.*, 39(4):1365–1406, 07 2011.

[41] L. Stephan and L. Massoulié. Robustness of spectral methods for community detection. In *Conference on Learning Theory, COLT 2019, 25-28 June 2019, Phoenix, AZ, USA*, pages 2831–2860, 2019.

[42] J. A. Tropp. An introduction to matrix concentration inequalities. *Found. Trends Mach. Learn.*, 8(1-2):1–230, May 2015.

[43] U. von Luxburg, M. Belkin, and O. Bousquet. Consistency of spectral clustering. *Ann. Statist.*, 36(2):555–586, 04 2008.

[44] V. H. Vu. Spectral norm of random matrices. *Combinatorica*, 27(6):721–736, 2007.

[45] J. S. Yedidia, W. T. Freeman, and Y. Weiss. Exploring artificial intelligence in the new millennium. chapter Understanding Belief Propagation and Its Generalizations, pages 239–269. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2003.

[46] Y. Yu, T. Wang, and R. J. Samworth. A useful variant of the davis–kahan theorem for statisticians. *Biometrika*, 102(2):315–323, 2015.