

SPECTRAL PROPERTIES FOR THE LAPLACIAN OF A GENERALIZED WIGNER MATRIX

ANIRBAN CHATTERJEE AND RAJAT SUBHRA HAZRA

ABSTRACT. In this article, we consider the spectrum of a Laplacian matrix, also known as Markov matrices where the entries of the matrix are independent but have a variance profile. Motivated by recent works on generalized Wigner matrices we assume that the variance profile gives rise to a sequence of graphons. Under the assumption that these graphons converge, we show that the limiting spectral distribution converges. We give an expression for the moments of the limiting measure in terms of graph homomorphisms. In some special cases, we identify the limit explicitly. We also study the spectral norm and derive the order of the maximum eigenvalue. We show that our results cover Laplacians of various random graphs including inhomogeneous Erdős-Rényi random graphs, sparse W-random graphs, stochastic block matrices and constrained random graphs.

1. INTRODUCTION

The Laplacian of a graph is used in various areas of combinatorics, statistical physics, and probability. Given a graph G on N vertices, the Laplacian is given by $A_N - D_N$, where A_N is the adjacency matrix and D_N , is the degree matrix, that is, a diagonal matrix with i -th diagonal entry being the degree of the graph. If G is a simple graph then the entries of A_N are either 0 or 1. When the entries of A_N are no longer restricted to 0 and 1, the Laplacian is referred to as the Markov matrix. In this article, we study the behavior of the eigenvalues of the Laplacian when A_N is a (generalized) Wigner matrix where the entries are independent but have a variance profile. These setups come up in random graph models when edge weights are independent but not identically distributed and the variance depends on the size of the graph. For example, we can consider the case of inhomogeneous Erdős-Rényi graphs where the vertex set is $[N] = \{1, \dots, N\}$ and any two vertices i and j are connected independently with probability $p_{i,j}$. Other examples include the adjacency matrix of a configuration model where the edges are no longer independent. This article aims to study the behavior of the empirical spectral distribution of the Laplacian matrix under such a variance profile and analyze the behavior of the spectral norm.

A Wigner matrix is a Hermitian random matrix whose entries are i.i.d random variables up to the symmetry constraint, and have zero expectation and variance 1. It is well known that for Wigner matrices the empirical spectral distribution (ESD) converges weakly almost surely to the semicircle law. The constant variance condition and the i.i.d requirement has subsequently been relaxed in Erdős et al. [19], Erdős et al. [20] to show the convergence of ESD to the semicircle law under the setup where entries can have different variances and each column of the variance profile is stochastic. Wigner matrices with a variance profile has also been considered in Ajanki et al. [2], Anderson and Zeitouni [3], Chakrabarty [10], Hachem et al. [21], Shlyakhtenko [33]. Non-symmetric random matrices with variance profiles were considered in Cook et al. [16]. Zhu [41] introduce a graphon approach to finding the moments of the limiting distribution of ESD of a Wigner-type random matrix where the entries satisfy a Lindeberg type assumption and we shall follow the setup of that article. In Wigner matrices with variance profile, generally, the variance matrix is assumed to have some structure and in particular, it was assumed in Zhu [41] that it gives rise to an empirical graphon which converges to a graphon. In that case, the limiting spectral distribution can be described in terms of this limiting graphon. In many important cases, the limit is not Wigner's semicircle law. The importance of assuming a variance profile lies in the fact that it can be used to model various stochastic block matrices (Abbe [1]). Under some non-sparsity assumption (average degree goes to infinity) it is known that the ESD of adjacency of homogeneous

2000 *Mathematics Subject Classification.* 60B20, 05C80, 46L54.

Key words and phrases. Adjacency matrices, Inhomogeneous Erdős-Rényi random graph, Largest eigenvalue, Scaling limit, Stochastic block model.

Erdős–Rényi converges to the semicircle law (Tran et al. [36]). The inhomogeneous extension was done subsequently in Chakrabarty et al. [11], which falls in the setup of the Wigner matrix with a variance profile.

The graph Laplacian is a counterpart of the continuous Laplacian which is well-known in the theory of diffusions and also related to a flow in the network. The spectral graph theory is the study of the properties of a graph in relationship to the characteristic polynomial, eigenvalues, and eigenvectors of its adjacency or Laplacian matrix. Laplacian matrix has relations with the number of spanning forest of the graph (through Kirchoff’s theorem), the algebraic connectivity, number of connected components (through the multiplicity of zeroes). We refer the readers to the monograph Chung [15] for applications of spectral analysis to graph theory. In recent statistical and machine learning applications, it has found good use in the spectral clustering techniques (Couillet and Benaych-Georges [17], von Luxburg et al. [38], Zhou and Amini [40]) and community detection algorithms (Chen et al. [13]).

Fundamental work on random Laplacian matrices was done in Bryc et al. [9] and the convergence of ESD of Laplacian matrices under the i.i.d. setup was determined. The limiting law turns out to be free convolution between the semicircle law and the standard Gaussian distribution. The ESD of Laplacian of sparse Erdős–Rényi is considered in Jiang [23]. The normalized Laplacian in the non-sparse setting was considered in Chi [14]. The local laws Laplacian of the Erdős–Rényi graph was considered in Huang and Landon [22]. They showed that the Stieltjes transform of the empirical eigenvalue distribution is well-approximated by the Stieltjes transform of free convolution of the semicircle law and a standard Gaussian down to the scale N^{-1} . They also show that the gap statistics and averaged correlation functions coincide with the Gaussian Orthogonal Ensemble in the bulk. Ding and Jiang [18] discusses the convergence of ESD of adjacency and Laplacian of random graphs, under the assumption that the variance of entries of $N \times N$ adjacency matrix is constant, only depending on N . There have been few studies on the spectral norm of the Laplacian matrix. Bryc et al. [9] show that for the mean-centered case the order becomes $O(\sqrt{N \log N})$ whereas it changes to $O(N)$ when mean centering is not considered. The law of large numbers for the spectral norm and the largest eigenvalues under the assumption of independence was discussed in Ding and Jiang [18]. They restricted the setup such that the entries satisfy symmetric constraint and the mean and variance depends only on N and not on the index of the entries and showed that the order remains the same for the mean-centered version. Bordenave et al. [6] considered the asymmetric setup and showed that the order of growth of the largest singular value remains $O(\sqrt{N \log N})$ tallying with the previous two works. The same rate of growth of the spectral norm was studied in Ding and Jiang [18], Jiang [24]. We are not aware of any literature which deals with the fluctuations of the spectral norm in these settings.

Main contribution of the article: As mentioned above, we take A_N to be a matrix with independent entries but having a variance profile. Our main assumption is similar to that of Zhu [41], that is, the variance profile matrix gives rise to a graphon W_N which converges in the cut-metric to a limiting graphon W . In Zhu [41] it was shown that the limiting spectral distribution of scaled A_N can be identified through its moments. It is well-known that if C_k is the k -th Catalan number (or $2k$ -th moment of limiting spectral distribution of scaled Wigner matrix) then C_k also counts the number of planer trees on $k + 1$ vertices. In the homogeneous setting when all the variances are the same, each planar tree contributes 1. In the inhomogeneous setting, each planer tree T has a non-negligible contribution, namely, it contributes $t(T, W)$, which indicates the number of copies of a planar tree T in graphon W (more explicitly, see (2.3.2) in the next section).

In the case of Laplacian, the identification of the moments and limit becomes a significantly difficult problem. One can show that ESD of scaled $A_N - D_N$ is the same as ESD of $A_N - \hat{D}_N$ where \hat{D}_N is independent of A_N and same in distribution as D_N . Since A_N is turning out to be a Wigner matrix with a variance profile, so it is not immediate that scaled A_N and \hat{D}_N are asymptotically freely independent as in the i.i.d. setting. It can be shown that when the variance profile or the limiting graphon is multiplicative, then free independence helps us to characterize the limit. We explore the combinatorial expression of the moments in terms of graph homomorphisms. We show that the moments can be expressed in terms of a mixture of Gaussian moments and $t(\tilde{T}, W)$ where \tilde{T} will a modification of the planar tree and the expression $t(\tilde{T}, W)$ indicate the number of copies of this modified tree in the limiting graphon. We

are not aware of such existing expressions for the limits of the moments of random Laplacian matrix. It is well known that the limit in the case of adjacency matrix is known as operator valued semi-circular law and it has connections to freeness over amalgamation (Mingo and Speicher [28], Nica et al. [30]). We strongly believe that this connection extends to Laplacian case too but we don't explore this aspect in the present article.

We derive various interesting examples, especially in random graphs which fall in our setting, for example, inhomogeneous Erdős–Rényi, Sparse W -random graphs and constrained random graphs. The limit is explicitly identified in some special cases when the entries have constant variance and the limiting graphon has a multiplicative structure. We derive the order of the spectral norm when the entries satisfy a bit more restrictive condition. Inspired by the methodology of Bryc et al. [9], we use strong Gaussian approximation which imposes some restrictions on the entries of A_N . We show that their methodology can be extended to a large extent to cover the inhomogeneous setting. It would be interesting to derive the fluctuations of the spectral norm in the above setting. We leave this aspect of analysis for future work.

Outline of the article: The article is arranged in the following way. In Section 2 we introduce the graphon setting briefly and state the main results about the empirical spectral distribution (Theorem 2.1 and Corollary 2.1). We identify the limiting spectral distribution in the multiplicative setting in Theorem 2.2. Subsection 2.5 is dedicated to the description of the moments of the limiting distribution and we compute some lower-order moments to give an idea of how the expression can be used. In Theorem 2.3, Theorem 2.4 we derive almost sure bounds on the spectral norm in the centered case. We discuss some examples which satisfy our assumptions in Section 3. In Section 4 we show some simulations on how the LSD looks like for different graphons. Section 5 and Section 6 are dedicated to the proof of the results on ESD and spectral norm respectively.

Acknowledgment. The research of RSH was supported by the MATRICS grant of SERB. A part of the work was done by AC for the master dissertation in Indian Statistical Institute, Kolkata.

2. THE SETUP AND THE RESULTS

2.1. Laplacian Matrix. For any symmetric $N \times N$ matrix A with eigenvalues $\lambda_1, \dots, \lambda_N$, the *empirical spectral distribution* (ESD) of A is defined by the probability measure

$$\text{ESD}(A) = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i}$$

In this paper we would study a particular class of structured random matrices called the *generalised Wigner matrices*. A generalised Wigner matrix is a random matrix $A_N = ((X_{i \wedge j, i \vee j}))_{N \times N}$ satisfying

- $\{X_{i,j} : 1 \leq i \leq j \leq N\}$ are independent real valued random variables;
- $\mathbb{E}[X_{i,j}] < \infty, \quad \forall 1 \leq i \leq j \leq N;$
- $\mathbb{E}[X_{i,j}^2] < \infty, \quad \forall 1 \leq i \leq j \leq N.$

The random variables also depend on N , but for notational simplification we remove the dependency. It is easy to draw parallel between such matrix and the adjacency matrix of a graph on N vertices having edge weight $X_{i,j}$ on the edge between the vertices i and j . As a result we would sometimes use the term adjacency matrix to denote the generalised Wigner matrices. Correspondingly we can define the Laplacian of A_N as

$$\Delta_N(i, j) = \begin{cases} A_N(i, j) & \text{if } i \neq j \\ - \sum_{k=1, k \neq i}^N A_N(i, k) & \text{if } i = j. \end{cases} \quad (2.1.1)$$

Since the row sum of Δ_N is zero and the infinitesimal generators of continuous-time Markov processes on finite state spaces are given by matrices with row-sums zero. Such matrices are also referred to as

Markov matrix in literature (see Bryc et al. [9]). This paper is mainly concerned with the mean centered version of the above matrices, which we denote by

$$A_N^0 = \frac{1}{\sqrt{N}}(A_N - \mathbb{E}(A_N)) \quad (2.1.2)$$

$$\Delta_N^0 = \frac{1}{\sqrt{N}}(\Delta_N - \mathbb{E}(\Delta_N)). \quad (2.1.3)$$

Define the variance profile matrix corresponding to A_N by $\Sigma_N = ((\sigma_{i,j}^2))_{N \times N}$, where $\sigma_{i,j}^2 = \mathbb{E}[(X_{i,j} - \mathbb{E}(X_{i,j}))^2] > 0$ for all $1 \leq i, j \leq N$. Some kind of convergence assumption on Σ_N is necessary for getting a limit result for the above matrices. We shall assume that the variance profile gives rise to a graphon in the limit.

2.2. Graphons and Convergence of Graph Sequences. Understanding large networks is a fundamental problem in modern graph theory and to properly define a limit object, an important issue is to have a good definition of convergence for graph sequences. The theory of graphons (Lovász and Szegedy [27]) as limits of dense graph sequences aims to provide a solution to this problem.

In our approach, we would define the variance profile matrix Σ_N as a graphon sequence. The convergence of empirical spectral distribution is connected to the convergence of this graphon sequence associated with Σ_N . We provide a brief introduction to graphon theory and for more details refer to Lovász [26].

A *graphon* is a measurable function $W : [0, 1]^2 \rightarrow [0, 1]$ such that $W(x, y) = W(y, x)$ for all $x, y \in [0, 1]$. Let \mathcal{W} to be the space of all graphons. To define the cut-metric on \mathcal{W} , let Φ denote the set of all bijective, Lebesgue measure preserving $\sigma : [0, 1] \rightarrow [0, 1]$. For two graphons W_1 and W_2 , the cut-distance is defined as

$$d_{\square}(W_1, W_2) = \sup_{S, T \subseteq [0, 1]} \left| \int_{S \times T} (W_1(x, y) - W_2(x, y)) \, dx dy \right|,$$

where S and T ranges over all measurable subsets of $[0, 1]$.

Then the cut metric is given by

$$\delta_{\square}(W_1, W_2) = \inf_{\sigma \in \Phi} d_{\square}(W_1, W_2^{\sigma})$$

where $W_2^{\sigma} = W_2(\sigma(x), \sigma(y))$. This forms a pseudo-metric and hence one says $W_1 \sim W_2$ if $\delta_{\square}(W_1, W_2) = 0$. Let $\widetilde{\mathcal{W}}$ be the space of all equivalence classes. It is known that $(\widetilde{\mathcal{W}}, \delta_{\square})$ is a compact metric space. Every weighted graph G can be associated with a graphon.

Definition 2.1. Consider a weighted graph $G = (V, E, (w_e)_{e \in E})$ and for $j \in \{1, \dots, |V| - 1\}$ define

$$I_j = \left[\frac{j-1}{|V|}, \frac{j}{|V|} \right) \text{ and } I_{|V|} = \left[1 - \frac{1}{|V|}, 1 \right].$$

Then we define the empirical graphon of G as

$$W^G(x, y) = \begin{cases} w_e & \text{if } e = (i, j) \in E(G), \quad (x, y) \in I_i \times I_j \\ 0 & \text{otherwise.} \end{cases}$$

Observe that any empirical graphon $W^G \in \mathcal{W}$, if the weights lie in $[0, 1]$. Using the cut metric, we are able to compare two graphs with different sizes and measure their similarity, which defines a type of convergence of graph sequences whose limiting object is the graphon. Another way of defining convergence of graphs is to consider the graph homomorphisms

Definition 2.2. For any graphon W and a finite simple graph $F = (V, E)$ (without loops), define the homomorphism density from F to W as

$$t(F, W) = \int_{[0, 1]^{|V|}} \prod_{\{i, j\} \in E} W(x_i, x_j) \prod_{i \in V} dx_i. \quad (2.2.1)$$

It is natural to think two graphon W_1 and W_2 are similar if they have similar homomorphism densities from any finite graph G . Let $\{W_n\}_{n \in \mathbb{N}}$ be a sequence of graphons. We say $\{W_n\}_{n \in \mathbb{N}}$ is *convergent from the left* if $t(F, W_n)$ converges for any finite simple (no loops, no multi-edges, no directions) graph F .

The homomorphism density characterises convergence under the cut metric. Lovász [26, Theorem 11.5] gives a characterisation of convergence in the space \mathcal{W} . Let $\{W_n\}_{n \in \mathbb{N}}$ be a sequence of graphons in \mathcal{W}_0 and let $W \in \mathcal{W}_0$. Then $t(F, W_n) \rightarrow t(F, W)$ for all finite simple graphs F if and only if $\delta_{\square}(W_n, W) \rightarrow 0$. We now describe the assumptions needed for our results. They are very similar to the ones mentioned in Zhu [41].

2.3. Limiting spectral distribution of Laplacian. Let A_N be a $N \times N$ *generalised Wigner matrix* with a variance profile matrix Σ_N satisfying the following conditions:

L.1 (Bounded variance) There exists a constant $C > 0$ such that

$$\text{Var}(X_{i,j}) \leq C, \quad \forall 1 \leq i, j \leq N, N \geq 1.$$

Without loss of generality we assume $C \leq 1$.

L.2 (Lindeberg's Condition) for any constant $\eta > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{N^2} \sum_{1 \leq i, j \leq N} \mathbb{E} \left[|X_{i,j} - \mathbb{E}[X_{i,j}]|^2 \mathbf{1} \left(|X_{i,j} - \mathbb{E}[X_{i,j}]| \geq \eta \sqrt{N} \right) \right] = 0. \quad (2.3.1)$$

L.3 (Graphon convergence) Consider the graph

$$G^{\Sigma_N} = \left([N], \{(i, j) : 1 \leq i \leq j\}, (\sigma_{i,j}^2)_{1 \leq i \leq j \leq N} \right)$$

and the corresponding empirical graphon W_N . We assume there exists a graphon $W \in \mathcal{W}$ such that

$$\delta_{\square}(W_N, W) \rightarrow 0.$$

Remark 2.1. *There are multiple examples of random matrices where the above assumptions are satisfied. We deal later with some examples arising out of random graphs like inhomogeneous Erdős-Rényi, W -sparse random graphs, constrained random graphs and stochastic block model. In some cases the assumptions were already verified in Zhu [41].*

Before stating the result on convergence of ESD of the centered Laplacian matrix (2.1.3), for the sake of completeness let us take a look at the result on the matrix A_N . To describe the limiting moments we will need the definition of rooted planar tree and this play a crucial role also in the description of moments of the Laplacian.

The *rooted planar tree* is a planar graph with no cycles, with one distinguished vertex as a root, and with a choice of ordering at each vertex. The ordering defines a way to explore the tree starting at the root. One of the algorithms used for traversing the rooted planar trees is *depth-first search*. An enumeration of the vertices of a tree is said to have depth-first search order if it is the output of the depth-first search.

It was shown in Zhu [41, Theorem 3.2] that under the assumptions **L.1–L.3**,

$$\lim_{N \rightarrow \infty} \text{ESD}(A_N^0) = \mu \text{ weakly almost surely,}$$

where μ denotes the unique probability measure identified by the following moments

$$\int x^{2k} d\mu = \sum_{j=1}^{C_k} t(T_j^{k+1}, W), \quad \int x^{2k+1} d\mu = 0, \quad k \geq 0, \quad (2.3.2)$$

where T_j^{k+1} is the j^{th} rooted planar tree with $k+1$ vertices and C_k is the k^{th} Catalan number.

Then for the centered Laplacian defined in (2.1.3) we have the following result

Theorem 2.1. *Under assumptions **L.1–L.3**,*

$$\lim_{N \rightarrow \infty} \text{ESD}(\Delta_N^0) = \nu \text{ weakly in probability}$$

where ν is the unique symmetric probability measure on \mathbb{R} . Further if there exists an open set $U \subseteq [0, 1]^2$ such that $W > 0$ on U , then ν has unbounded support.

2.4. Identification of Limiting Spectral Distribution. The limiting spectral distribution can be identified with standard distributions under certain simplifications. Along with the assumption of boundedness (L.1) we assume that

L.4 $\mathbb{E}X_{i,j} = \mu_N$ and $\sigma_{i,j}^2 = \sigma_N^2 > 0$ for all $1 \leq i, j \leq N, N \geq 1$.

L.5 There exists $\delta > 0$ such that

$$\sup_{1 \leq i, j \leq N, N \geq 1} \mathbb{E} \left[\left| \frac{X_{i,j} - \mu_N}{\sigma_N} \right|^{2+\delta} \right] < \infty.$$

Defining $\{\lambda_i(A) : 1 \leq i \leq N\}$ as the eigenvalues of a $N \times N$ matrix A , we have the following result.

Corollary 2.1. Suppose L.1, L.4 and L.5 holds. Set

$$\tilde{F}_N(x) = \frac{1}{N} \sum_{i=1}^N I \left\{ \frac{\lambda_i(\Delta_N) - N\mu_N}{\sqrt{N}\sigma_N} \leq x \right\}, \quad \forall x \in \mathbb{R}.$$

Then, as $N \rightarrow \infty$, \tilde{F}_N converges to a distribution function F in probability where F denotes the distribution of the free additive convolution γ_M of the semicircle law and the standard normal distribution.

The above result identifies the limiting measure in terms of additive free convolution of two measures. The identification of the limiting measure can be achieved in the case when W has a *multiplicative structure*, that is

$$W(x, y) = r(x)r(y), \quad x, y \in [0, 1] \quad (2.4.1)$$

for some continuous function $r : [0, 1] \rightarrow [0, 1]$. The statement is based on the theory of self-adjoint operators affiliated with a W^* -probability space. It is important to note that we do not assume the initial variances to have a multiplicative structure but the multiplicative structure only appears in the limit. We recall some definitions mentioned in Chakrabarty et al. [11]. For details of free probability of unbounded operators we refer to Anderson et al. [4].

Recall (\mathcal{A}, τ) is called a W^* probability space if \mathcal{A} is a closed (in weak operator topology) C^* algebra of bounded operators on some Hilbert space and τ is a state. A densely defined operator T is said to be affiliated to \mathcal{A} if for every bounded measurable function h , $h(T) \in \mathcal{A}$. For an affiliated operator T , its law $\mathcal{L}(T)$ is the unique probability measure on \mathbb{R} satisfying

$$\tau(h(T)) = \int_{\mathbb{R}} h(x)(\mathcal{L}(T))dx.$$

For two or more self adjoint operators T_1, \dots, T_n , a description of their joint distribution is a specification of

$$\tau(h_1(T_{i_1}), \dots, h_k(T_{i_k})),$$

for all $k \geq 1$, all $i_1, \dots, i_k \in \{1, 2, \dots, n\}$, and all bounded measurable functions h_1, \dots, h_k from \mathbb{R} to itself.

Definition 2.3. Let (\mathcal{A}, τ) be a W^* -probability space and $a_1, a_2 \in \mathcal{A}$. Then a_1 and a_2 are freely independent if

$$\tau(p_1(a_{i_1}), \dots, p_n(a_{i_n})) = 0$$

for all $n \geq 1$, all $i_1, \dots, i_n \in \{1, 2\}$ with $i_j \neq i_{j+1}, j = 1, \dots, n-1$, and all polynomials p_1, \dots, p_n in one variable satisfying

$$\tau(p_j(a_{i_j})) = 0, \quad j = 1, \dots, n$$

For operators a_1, \dots, a_k and b_1, \dots, b_m affiliated with \mathcal{A} , the collections (a_1, \dots, a_k) and (b_1, \dots, b_m) are freely independent if and only if

$$p(h_1(a_1), \dots, h_k(a_k)) \text{ and } q(g_1(b_1), \dots, g_m(b_m))$$

are freely independent for all bounded measurable h_1, \dots, h_k and g_1, \dots, g_m , and all polynomials p and q in k and m non-commutative variables, respectively. It is immediate that the two operators in the above display are bounded, and hence belong to \mathcal{A} .

Now we are ready to state our theorem.

Theorem 2.2. *If W is as in (2.4.1), then under L.1-L.3 with the cut-metric convergence in L.3 replaced by $W_N \xrightarrow{L_1} W$ the limiting measure ν can be identified as,*

$$\nu = \mathcal{L} \left(r^{1/2}(T_u) T_s r^{1/2}(T_u) + \alpha r^{1/4}(T_u) T_g r^{1/4}(T_u) \right) \quad (2.4.2)$$

where

$$\alpha = \left(\int_0^1 r(x) dx \right)^{1/2}$$

Here, T_g and T_u are commuting self-adjoint operators affiliated with a W^* -probability space (\mathcal{A}, τ) such that, for bounded measurable functions h_1, h_2 from \mathbb{R} to itself,

$$\tau (h_1(T_g) h_2(T_u)) = \left(\int_{-\infty}^{\infty} h_1(x) \phi(x) dx \right) \left(\int_0^1 h_2(u) du \right) \quad (2.4.3)$$

with ϕ the standard normal density. Furthermore, T_s has a standard semicircle law and is freely independent of (T_g, T_u) .

Such identification was first achieved in the results of Chakrabarty et al. [11] in the case of inhomogeneous Erdős-Rényi random graphs. We show that similar limit appears in the general case too.

Remark 2.2. *Under such multiplicative structure the limiting spectral distribution of the matrix A_N can also be identified. If we keep the assumption (2.4.1) on the limiting graphon W and the empirical graphon W_N corresponding to the variance profile as in Theorem 2.2, then the limiting spectral distribution of the scaled and centered matrix A_N^0 is given by*

$$\mu = \mathcal{L} \left(r^{1/2}(T_u) T_s r^{1/2}(T_u) \right) \quad (2.4.4)$$

where T_u and T_s are as defined in Theorem 2.2. One should note that the measure μ in 2.4.4 is the same as the free multiplicative convolution of the standard semicircle law and the law of $r(U)$, where U is a standard uniform random variable.

2.5. Description of the moments. In this subsection, we briefly describe the moments of the limiting measure ν . It turns out to be an interesting combinatorial problem to come up with an expression for the moments. To describe the limiting moments, we need to introduce some notions. Fix k in $2\mathbb{N}$. Consider the multiset $\{m_1, m_2, \dots, m_k, n_1, n_2, \dots, n_k\} \in \mathbb{N}$ such that $\sum_{p=1}^k (m_p + n_p) = k$ and $m = \sum_{p=1}^k m_p$ is even. Then consider a rooted planar tree T on $(\frac{m}{2} + 1)$ vertices and take $\tilde{i} = i_1 i_2 \dots i_{m+1}$ to be the depth first search on T . Then we must have $i_{m+1} = i_1$. So for notational simplicity we identify 1 by $m + 1 = 1 + \sum_{p=1}^k m_p$. Then consider $\{s_1, \dots, s_{\eta_s}\} \subseteq \{1, 2, \dots, k\}$ such that the indices in \tilde{i} corresponding to the set $\{1 + \sum_1^{s_1} m_p, \dots, 1 + \sum_1^{s_{\eta_s}} m_p\}$, that is $\{i_{1+\sum_1^{s_1} m_p}, \dots, i_{1+\sum_1^{s_{\eta_s}} m_p}\}$ represents the s^{th} vertex of T . Observe that η_s can be equal 0. Now suppose that $\sum_{j=1}^{\eta_s} n_{s_j}$ is even for all vertex s in T . Then define $\hat{n}_s = \frac{1}{2} \sum_{j=1}^{\eta_s} n_{s_j}$ (if η_s is 0, then the sum is also 0). Construct a new graph \tilde{T} by attaching \hat{n}_s many vertices to the vertex s of T , for all $s \in \{1, 2, \dots, \frac{m}{2} + 1\}$. Only modification that we are doing to T is by adding leaf nodes. Hence the modified graph is still a tree. An example of such modification at the vertex s is shown in Figure 1.

For $t \in \mathbb{N}$, define

$$\mathbf{m}_t = \mathbb{E} [Z^t], \quad Z \sim N(0, 1).$$

Then define the function

$$f(T) = \begin{cases} \prod_{s=1}^{\frac{m}{2}+1} \mathbf{m}_{2\hat{n}_s} & \text{if } 2\hat{n}_s \text{ is even for all } s, \\ 0 & \text{otherwise.} \end{cases} \quad (2.5.1)$$

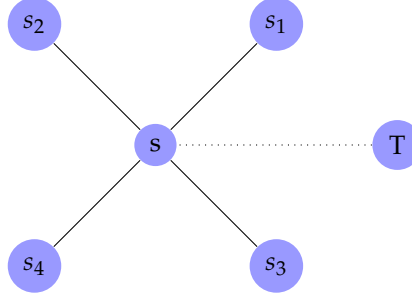


FIGURE 1. The modification of graph T at vertex s with $\hat{n}_s = 4$

Example 2.1. We provide an example of such an modification. Consider $k = 12$, $\tilde{m} = \{m_i\}_{i=1}^{12}$ such that $m_i = 2$ for $i = 1, 2, 4$ and 0 otherwise and finally $\tilde{n} = \{n_i\}_{i=1}^{12}$ such that $n_i = 2$ for $i = 1, 2, 5$ and 0 otherwise. Then $m = \sum_{i=1}^k m_i = 6$. Hence we consider the rooted planar tree T_4 as in Figure 2a.

Then consider the depth first search $1 \rightarrow 2 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 2 \rightarrow 1 = i_1 \rightarrow i_2 \rightarrow i_3 \rightarrow i_4 \rightarrow i_5 \rightarrow i_6 \rightarrow i_7$. Now observe that

$$1 + \sum_{i=1}^j m_i = \begin{cases} 3 & j = 1 \\ 5 & j = 2, 3 \\ 7 & \text{otherwise} \end{cases}$$

Observe that by definition $i_1 = i_7$ corresponds to the vertex 1. Now let us look at 1^{st} vertex, that is vertex 1. Observe that $i_{1+\sum_{l=1}^j m_l}$ falls on vertex 1 if $4 \leq j \leq 12$. Then by definition $\eta_1 = 9$ and the set $\{s_1, \dots, s_{\eta_s}\}$ for $s = 1$ is $\{4, 5, \dots, 12\}$. Then $\hat{n}_1 = \frac{1}{2} \sum_{j=1}^9 n_{s_j} = 1$. Then we look at vertex 2. Observe that i_2, i_4, i_6 corresponds to vertex 2. But $2, 4, 6 \notin \{1 + \sum_{l=1}^j m_l : \forall 1 \leq j \leq 12\}$. Hence $\eta_2 = 0$. Lets look at 4^{rd} vertex. Observe that only $i_{1+\sum_{l=1}^j m_l}$ for $j = 2, 3$ falls on vertex 4. Then by definition $\eta_4 = 2$ and the set $\{s_1, \dots, s_{\eta_s}\}$ for $s = 4$ is $\{2, 3\}$. Then $\hat{n}_4 = 1$. Similarly we can show that $\hat{n}_3 = 1$. The modified graph is as in Figure 2b

One thing to note is that the modification is independent of the labeling of the vertices. It only depends upon the number of overlap certain indices have with the vertices.

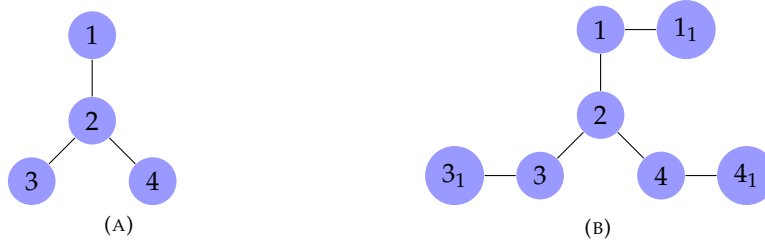


FIGURE 2. (a) Rooted planar tree T_4 , (b) Modified tree \tilde{T}_4

Let \mathcal{P}_{2k} be the multiset of all numbers $(m_1, \dots, m_{2k}, n_1, \dots, n_{2k})$ which appear when we expand $(a+b)^{2k}$ for two non-commutative variables a and b and write it as

$$(a+b)^{2k} = \sum_{(m_1, \dots, m_{2k}, n_1, \dots, n_{2k}) \in \mathcal{P}_{2k}} \prod_{i=1}^{2k} a^{m_i} b^{n_i}.$$

Observe that $\sum_{j=1}^{2k} (m_j + n_j) = 2k$. We identify the moments of the limiting measure ν through the following formula.

$$\int x^{2k} d\nu = \sum_{\mathcal{P}_{2k}} \sum_{r=1}^{\lfloor \frac{m}{2} \rfloor} t\left(\tilde{T}_r^{\frac{m}{2}+1}, W\right) f\left(T_r^{\frac{m}{2}+1}\right) \mathbf{1}_{\{m \in 2\mathbb{N} \cup \{0\}\}}, \int x^{2k+1} d\nu = 0, \forall k \geq 0 \quad (2.5.2)$$

where T_r^q denote the r^{th} rooted planar tree on q many vertices with \tilde{T}_r^q the corresponding modification as stated in Section 2.5.

Remark 2.3. Since the expression of moments in (2.5.2) look complicated, we elucidate by calculating the second and the fourth moments.

First, let us look at the second moment. The possible choices for the vector (m_1, m_2, n_1, n_2) are

$$(2, 0, 0, 0), (1, 0, 1, 0), (0, 1, 1, 0) \text{ and } (0, 0, 2, 0).$$

Observe that for the second and third choice, there wont be any contributions since $m = m_1 + m_2$ is odd. For the first choice number of trees is $C_{m/2} = C_1 = 1$, that is a single edge (T_1^2). Since here $n_1 = n_2 = 0$, then we don't need any modification of the tree. Thus

$$t\left(\tilde{T}_1^2, W\right) = \int W dx dy \text{ and } f(T_1) = 1$$

Now for the choice $(0, 0, 2, 0)$, $m = 0$ and hence number of trees are $C_0 = 1$. The tree is basically a single vertex T_1^1 . Then the depth first search would yield just $i_1 = 1$. Observe that $1 + m_1 = 1 + m_1 + m_2 = 1$. So $\hat{n}_1 = \frac{1}{2}(n_1 + n_2) = 1$. Hence the modified tree is same as T_1^2 . Thus the contribution becomes

$$t(\tilde{T}_1^1, W) = t(T_1^2, W) = \int W(x, y) dx dy \text{ and } f(T_1^1) = 1$$

Hence the second moment is

$$2 \int W(x, y) dx dy$$

Going as in the second moment we can determine the contributory terms in the expression of the fourth moment and their corresponding contributions. We provide explicit calculations for a few terms, the rest follows similarly. Let us look at the term corresponding to $(2, 0, 0, 0, 2, 0, 0, 0)$. Here $m = 2$ and $C_{m/2} = C_1 = 1$. So the contributory rooted planar tree is a single edge (T_1^2). The depth first search would give the closed walk, $i_1 \rightarrow i_2 \rightarrow i_3 = i_1$. Now observe that

$$1 + \sum_{l=1}^j m_l = 3, \quad j = 1, 2, 3, 4.$$

Then the number of overlaps of the form $i_{1+\sum_{l=1}^j m_l}$ with the first vertex is $\eta_1 = 4$ and the values of such j are $j = 1, 2, 3, 4$. Hence

$$\sum_{j=1}^4 n_j = 2 \implies \hat{n}_1 = 1$$

Then the modification of the first vertex is given by adding a single edge to the vertex 1. Observe that no such overlap is possible for the second vertex. Hence there is no modification for the second vertex. So for this case the modified tree \tilde{T}_1^2 is given as in Figure 3a and the contribution from this tree would be

$$t\left(\tilde{T}_1^2, W\right) \mathbf{m}_2$$



FIGURE 3. Modified trees for the fourth moment

Next let us look at the term corresponding to $(1,1,0,0,2,0,0,0)$. Here $m = 2$ and $C_{m/2} = 1$. So again the contributory rooted planar tree is T_1^2 . The depth first search is same as the previous case. Observe that here

$$1 + \sum_{l=1}^j m_l = \begin{cases} 2 & j = 1 \\ 3 & j = 2, 3, 4 \end{cases}$$

Then the number of overlaps of the form $i_{1+\sum_{l=1}^j m_l}$ with the first vertex is $\eta_1 = 3$ and the values of such j are $j = 2, 3, 4$. But then it is easy to observe that $\hat{n}_1 = 0$. Hence there is no modification of the first vertex. There is a single overlap with the second vertex given by i_{1+m_1} . Then $\hat{n}_2 = 1$ and hence we attach a single leaf node to the second vertex. So here again the modified tree is the same as the previous case, but with different labeling, given in Figure 3b. Going similarly for the other terms, the fourth moment is given by

$$(2 + 4\mathbf{m}_2 + \mathbf{m}_4)t(T_1^3, W)$$

where T_1^3 is the rooted planar tree on 3 vertices. (Note that there are 2 rooted planar trees on 3 vertices, but the homomorphism density corresponding to both are equal, hence we can take any one of the trees).

2.6. Spectral Norm of Laplacian. In this subsection we discuss the asymptotics of the spectral norm of the Laplacian matrix Δ_N . In the case when entries are taken to be i.i.d. some results are known. The first order asymptotics, that is, the order of growth of the Laplacian was described in Bryc et al. [9].

For a $N \times N$ matrix \mathbf{M} , the spectral norm of \mathbf{M} is defined as

$$\|\mathbf{M}\| = \sup_{\substack{\mathbf{x} \in \mathbb{R}^N \\ \|\mathbf{x}\|_2 = 1}} \|\mathbf{M}\mathbf{x}\|_2$$

where $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + \dots + x_N^2}$ for $\mathbf{x} = (x_1, x_2, \dots, x_N)^T \in \mathbb{R}^N$.

Observe that using symmetry of A_N and Δ_N we have

$$\|A_N\| = \max\{\lambda_{\max}(-A_N), \lambda_{\max}(A_N)\}, \quad \|\Delta_N\| = \max\{\lambda_{\max}(-\Delta_N), \lambda_{\max}(\Delta_N)\} \quad (2.6.1)$$

In order to proceed with our results we consider the following assumptions on the entries of A_N in line with the assumptions in Section 2.3.

S.1 (Bounded variance) Let $\sigma_{i,j}^2 = \text{Var}(X_{ij})$ and let $A_1 = \inf_{i,j \geq 1} \sigma_{i,j}^2$ and $A_2 = \sup_{i,j \geq 1} \sigma_{i,j}^2$. Suppose $A_1 > 0$ and $A_2 < \infty$. Without loss of generality we can take $A_2 \leq 1$.

S.2 (Higher moments) There exists $\delta > 0$ and $0 < K < \infty$ such that

$$\mathbb{E} \left[|X_{i,j} - \mathbb{E}X_{i,j}|^{6+\delta} \right] \leq K, \quad \text{for all } i, j \in \mathbb{N}.$$

S.3 (Graphon convergence) Consider the graph

$$G^{\Sigma_N} = \left([N], \{(i, j) : 1 \leq i \leq j\}, (\sigma_{i,j}^2)_{1 \leq i \leq j \leq N} \right)$$

and the corresponding empirical graphon W_N . Suppose there exists a graphon $W \in \mathcal{W}_0$ such that

$$\delta_{\square}(W_N, W) \rightarrow 0.$$

The assumption [L.3](#) is restated here as assumption [S.3](#) for convenience. It is needed for the convergence of the ESD of adjacency matrix A_N to a compactly supported probability measure.

Remark 2.4. *It can be easily seen that assumption [S.2](#) implies assumption [L.2](#) (Lindeberg's Condition) for any constant $\eta > 0$, that is*

$$\lim_{n \rightarrow \infty} \frac{1}{N^2} \sum_{1 \leq i, j \leq N} \mathbb{E} \left[|X_{i,j} - \mathbb{E}X_{i,j}|^2 \mathbf{1} \left(|X_{i,j}| \geq \eta \sqrt{N} \right) \right] = 0, \quad \forall \eta > 0 \quad (2.6.2)$$

Under the above set of assumptions we have the following result on spectral norm of the Laplacian matrix Δ_N .

Theorem 2.3. *Suppose that $A_1 = \inf_{i,j \geq 1} \sigma_{i,j}^2$ and $A_2 = \sup_{i,j \geq 1} \sigma_{i,j}^2$, and $\mathbb{E}X_{i,j} = 0$ for all $1 \leq i \leq j$. Then under the above set of assumptions we have*

$$\liminf_{N \rightarrow \infty} \frac{\|\Delta_N\|}{\sqrt{2N \log N}} \geq A_1^{\frac{1}{2}} \text{ a.s.}, \text{ and } \limsup_{N \rightarrow \infty} \frac{\|\Delta_N\|}{\sqrt{2N \log N}} \leq (2A_2)^{\frac{1}{2}} \text{ a.s.} \quad (2.6.3)$$

We are able to identify the in probability limits under an additional assumption, which we state below

S.4 (Uniform row stochasticity) For some $0 < r \leq 1$,

$$\lim_{N \rightarrow \infty} \sup_{i \in \mathbb{N}} \left| \frac{1}{N} \sum_{j=1}^n \sigma_{i,j}^2 - r \right| = 0.$$

Recall that by assumption [S.1](#) we have taken $A_2 \leq 1$, hence we can take $0 < r \leq 1$. Then we have the following theorem

Theorem 2.4. *Under the assumptions [S.1-S.4](#) we have*

$$\lim_{N \rightarrow \infty} \frac{\|\Delta_N\|}{\sqrt{2N \log N}} = \sqrt{r} \text{ in probability}$$

where $\|\cdot\|$ is the spectral norm.

The proof of [Theorem 2.4](#) is similar to the proof of [Theorem 2.3](#) and hence it will be omitted. We also look at the situation where the entries of the adjacency matrix have non-zero mean, that is $\exists N \geq 1$ and $1 \leq i \leq j \leq N$ such that $\mathbb{E}A_N(i, j) \neq 0$. The limit of the spectral norm can be identified under certain assumptions which we outline below. It is an immediate corollary of [Theorem 2.3](#).

S.5 There exists a constant $m \geq 0$ such that

$$\lim_{N \rightarrow \infty} \frac{\|\mathbb{E}\Delta_N\|}{N} = m$$

Under the above additional assumption we have the following [corollary](#),

Corollary 2.2. *Under the set of assumptions [S.1-S.3](#), along with the additional assumption [S.5](#), we have*

$$\lim_{N \rightarrow \infty} \frac{\|\Delta_N\|}{N} = m, \quad \text{a.s.}$$

3. SOME EXAMPLES

This section is devoted to examples and we show various models, mainly related to random networks which fall within the purview of the assumptions mentioned in the last section.

3.1. Inhomogenous Erdős-Rényi graphs. We define the setting of inhomogenous Erdős-Rényi graph following Chakrabarty et al. [11]. Let $f : [0, 1]^2 \rightarrow [0, 1]$ be a continuous function, satisfying

$$f(x, y) = f(y, x) \quad \forall x, y \in [0, 1]. \quad (3.1.1)$$

Consider $(\varepsilon_N : N \geq 1)$ to be a sequence of fixed positive real numbers satisfying

$$\lim_{N \rightarrow \infty} \varepsilon_N = 0, \quad \lim_{N \rightarrow \infty} N\varepsilon_N = \infty \quad (3.1.2)$$

Consider the random graph \mathbb{G}_N on vertices $\{1, \dots, N\}$ where, for each (i, j) with $1 \leq i < j \leq N$, an edge is present between vertices i and j with probability $\varepsilon_N f\left(\frac{i}{N}, \frac{j}{N}\right)$, independently of other pair of vertices. In particular, \mathbb{G}_N is an undirected graph with no self edges and no multiple edges. Boundedness of f ensures that $\varepsilon_N f\left(\frac{i}{N}, \frac{j}{N}\right) \leq 1$ for all $1 \leq i < j \leq N$ when N is large enough. Without loss of generality we assume that $\varepsilon_N < 1$ for all $N \geq 1$. If $f \equiv c$ with c a constant, then \mathbb{G}_N is the (homogeneous) Erdős-Rényi graph with edge retention probability $\varepsilon_N c$. The adjacency matrix is denoted by A_N . Clearly, A_N is a symmetric random matrix whose diagonal entries are zero and whose upper triangular entries are independent Bernoulli random variables, i.e.

$$A_N(i, j) \sim \text{Ber}\left(\varepsilon_N f\left(\frac{i}{N}, \frac{j}{N}\right)\right).$$

Define Δ_N to be the Laplacian of the graph \mathbb{G}_N corresponding to A_N as in definition 2.1.1. The following theorem, which is a restatement of Theorem 1.2 in Chakrabarty et al. [11] states the existence of limiting spectral distribution of Δ_N under suitable scaling.

Proposition 3.1. *Chakrabarty et al. [11, Theorem 1.2] There exists a unique, symmetric probability measure ν on \mathbb{R} such that*

$$\lim_{N \rightarrow \infty} \text{ESD}\left(\frac{1}{\sqrt{N\varepsilon_N}}(\Delta_N - D_N)\right) = \nu \text{ weakly in probability}$$

where

$$D_N = \text{diag}(\mathbb{E}\Delta_N(1, 1), \dots, \mathbb{E}\Delta_N(N, N))$$

Furthermore, if $f \neq 0$, then the support of ν is unbounded.

Proof. Observe that using Lemma 2.1 in Chakrabarty et al. [11] we have

$$\lim_{N \rightarrow \infty} L\left(\frac{1}{\sqrt{N\varepsilon_N}}\text{ESD}(\Delta_N^0), \frac{1}{\sqrt{N\varepsilon_N}}\text{ESD}(\Delta_N - D_N)\right) = 0 \text{ in probability,}$$

where L is the Levy metric¹ and

$$\Delta_N^0 = \Delta_N - \mathbb{E}\Delta_N.$$

Hence it is enough to look at the limiting spectral distribution of $\frac{1}{\sqrt{N\varepsilon_N}}\text{ESD}(\Delta_N^0)$. Define $B_N = \frac{1}{\sqrt{\varepsilon_N}}(A_N - \mathbb{E}A_N)$. Then $\frac{1}{\sqrt{\varepsilon_N}}\Delta_N^0$ is the centered Laplacian corresponding to B_N .

Observe that $\mathbb{E}B_{ij}^{(N)} = 0$ and further it is easy to observe that due to $N\varepsilon_N \rightarrow \infty$ and $|B_{ij}^{(N)}| \leq \frac{2}{\sqrt{\varepsilon_N}}$, we must have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{1 \leq i, j \leq N} \mathbb{E} \left[|B_{ij}^{(N)}|^2 \mathbf{1}_{\{|B_{ij}^{(N)}| \geq \eta\sqrt{N}\}} \right] = 0, \quad \forall \eta > 0$$

¹The Lévy-Prokhorov metric $L : \mathcal{P}(\mathbb{R})^2 \rightarrow [0, +\infty)$ between two probability measures μ and ν is given by

$$L(\mu, \nu) = \inf\{\varepsilon > 0 \mid \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon \quad \forall A \in \mathcal{B}(\mathbb{R})\} \quad (3.1.3)$$

where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ -algebra on \mathbb{R} and A^ε is the ε -neighbourhood of A .

Hence assumption [L.2](#) is satisfied. Now consider $\Sigma^{(N)}$ to be the variance profile matrix of B_N . Then for $i \neq j$

$$\Sigma_{ij}^{(N)} = \frac{\varepsilon_N f\left(\frac{i}{N}, \frac{j}{N}\right) \left(1 - \varepsilon_N f\left(\frac{i}{N}, \frac{j}{N}\right)\right)}{\varepsilon_N} = f\left(\frac{i}{N}, \frac{j}{N}\right) - \varepsilon_N f\left(\frac{i}{N}, \frac{j}{N}\right)^2$$

and $\Sigma_{ii}^{(N)} = 0$, $1 \leq i \leq N$. Then observe that using the fact that $\sup_N \varepsilon_N < 1$ and $|f| \leq 1$, we have $\Sigma_{ij}^{(N)} < C$ for some constant C for $1 \leq i, j \leq N$ and for all $N \geq 1$. Let W_N be the empirical graphon corresponding to $\Sigma^{(N)}$. Let

$$G^N = \left([N], \{(i, j) : 1 \leq i \leq j\}, \left(f\left(\frac{i}{N}, \frac{j}{N}\right) \right)_{1 \leq i \leq j \leq N} \right)$$

be the weighted graph on $[N]$ vertices with edge weights $\beta_{ij} = f\left(\frac{i}{N}, \frac{j}{N}\right)$ for edge $i \neq j, 1 \leq i, j \leq N$ and $\beta_{ii} = 0, 1 \leq i \leq N$. Let \widehat{W}_N be the empirical graphon corresponding to G^N . Then we have

$$W_N(x, y) = \widehat{W}_N(x, y) + o(1) \tag{3.1.4}$$

where $o(1)$ is uniformly over all $(x, y) \in [0, 1]^2$. Now observe that by definition f is also a graphon and is uniformly continuous in $[0, 1]^2$, hence given $\eta > 0 \exists N_\eta \in \mathbb{N}$ such that $\forall N \geq N_\eta$,

$$\left| f\left(\frac{i}{N}, \frac{j}{N}\right) - f(x, y) \right| < \eta, \quad \forall (x, y) \in I_i \times I_j \quad \forall 1 \leq i, j \leq N$$

Then observe that

$$\begin{aligned} \|\widehat{W}_N - f\|_{\square} &= \sup_{S, T \subseteq [0, 1]} \left| \int_{S \times T} \widehat{W}_N(x, y) - f(x, y) dx dy \right| \\ &\leq \int_{[0, 1]^2} \left| \widehat{W}_N(x, y) - f(x, y) \right| dx dy \\ &= \sum_{i, j=1}^N \int_{I_i \times I_j} \left| \widehat{W}_N(x, y) - f(x, y) \right| dx dy \\ &= \sum_{i, j=1}^N \int_{I_i \times I_j} \left| f\left(\frac{i}{N}, \frac{j}{N}\right) - f(x, y) \right| dx dy \leq \eta, \quad \forall N \geq N_\eta \end{aligned} \tag{3.1.5}$$

Hence we have $\delta_{\square}(\widehat{W}_N, f) \rightarrow 0$. Then using [\(3.1.4\)](#) we have $\delta_{\square}(W_N, f) \rightarrow 0$. Thus assumptions [L.1-L.3](#) are satisfied. Further when $f \not\equiv 0$ there exists an open set $U \subseteq [0, 1]^2$ such that $f > 0$ on U . The result follows from [Theorem 2.1](#). \square

Under a different situation, we are well equipped to look at the spectral norm of inhomogenous Erdős-Rényi graph. Instead of the sparse setup of [\(3.1.2\)](#), we consider the dense situation

$$\lim_{N \rightarrow \infty} \varepsilon_N = \varepsilon_\infty \tag{3.1.6}$$

for some $\varepsilon_\infty \in (0, 1)$. Further we assume that there exists a $\lambda > 0$ such that

$$\inf_{x, y \in [0, 1]^2} f(x, y) > \lambda. \tag{3.1.7}$$

Once again we consider the matrix $B_N = ((B_{ij}^{(N)}))_{i, j}$ defined in [Proposition 3.1](#). Using the fact that $|B_{ij}^{(N)}| \leq \frac{2}{\sqrt{\varepsilon_N}}$ and [\(3.1.6\)](#) we must have

$$\mathbb{E} \left[\left| B_{ij}^{(N)} \right|^7 \right] \leq \frac{2^7}{\varepsilon_N^{7/2}} \leq K$$

for some $K > 0$. Thus assumption [S.2](#) is satisfied. Recall the variance profile matrix of B_N given by

$$\Sigma_{ij}^{(N)} = f\left(\frac{i}{N}, \frac{j}{N}\right) - \varepsilon_N f\left(\frac{i}{N}, \frac{j}{N}\right)^2$$

Then using the fact that f is a continuous function from a compact set, lower bounded by $\lambda > 0$, we can easily see that $\Sigma_{ij}^{(N)}$ satisfies assumption [S.1](#). Further continuity along with compact support shows that the function f^2 is uniformly continuous. Using an argument similar to [\(3.1.5\)](#), it can be shown that the empirical graphon corresponding to $\Sigma_{ij}^{(N)}$ converges to

$$W(x, y) = f(x, y) - \varepsilon_\infty f(x, y)^2$$

in the cut metric. By definition, W is a well defined graphon, and hence assumption [S.3](#) is satisfied. Thus we have the following result using [Theorem 2.3](#).

Proposition 3.2. *Under the additional assumptions [\(3.1.6\)](#) and [\(3.1.7\)](#), we have*

$$\|\Delta_N\| = \Theta(\sqrt{N \log N}) \text{ a.s.}$$

where $h(n) = \Theta(g(n))$ implies that for all large enough n , $c_1 g(n) \leq h(n) \leq c_2 g(n)$ for some positive constants c_1 and c_2 .

3.2. Sparse W -random graphs. Given a graphon $W : [0, 1]^2 \rightarrow [0, 1]$, as stated in the definitions of Borgs et al. [\[8\]](#) a sequence of sparse random graphs G_N can be generated in the following way. We choose ε_N to be a sparsity parameter such that $\sup_N \varepsilon_N < 1$ and $\varepsilon_N \rightarrow 0$ and $N\varepsilon_N \rightarrow \infty$. Let $\{X_i : i \geq 1\}$ be i.i.d. $U[0, 1]$ consider the random graph G_N , where i and j are connected with probability $\varepsilon_N W(X_i, X_j)$ independently for all $i \neq j$. The graph G_N is called the *sparse W -random graph* and is denoted by $\mathcal{G}(N, W, \varepsilon_N)$. The adjacency matrix of such random graphs were studied in Zhu [\[41\]](#). In the following theorem we show the existence of an unique limiting spectral distribution of the Laplacian matrix. The proof essentially follows the ideas from Zhu [\[41\]](#) and we present the ideas for completeness.

Proposition 3.3. *Let $\mathcal{G}(N, W, \varepsilon_n)$ be a sequence of sparse W -random graphs with adjacency matrix A_N and the corresponding Laplacian matrix Δ_N . Further assume that W is regular, that is*

$$\int_0^1 W(x, y) dy = d_W, \quad \forall x \in [0, 1]$$

where d_W is a constant only depending on the graphon W . Then the empirical spectral distribution of $\frac{\Delta_N - \mathbb{E}\Delta_N}{\sqrt{N\varepsilon_N}}$ converges in probability to an unique symmetric probability measure ν . Further if there exists an open set $U \subseteq [0, 1]^2$ such that $W > 0$ on U , then ν have unbounded support.

Proof. The proof will follow from the proof of Zhu [\[41, Theorem 5.1\]](#) as long as we can show the following,

$$L\left(\frac{\Delta_N - \mathbb{E}\Delta_N}{\sqrt{N\varepsilon_N}}, \frac{\Delta_N - \mathbb{E}[\Delta_N | X_1, \dots, X_N]}{\sqrt{N\varepsilon_N}}\right) \xrightarrow{p} 0 \quad (3.2.1)$$

Observe that,

$$\mathbb{E}W(X_1, X_2) = \int \int W(x, y) dx dy = d_W \text{ and } \mathbb{E}W(X_1, X_2)W(X_1, X_3) = d_W^2 \quad (3.2.2)$$

By definition,

$$\begin{aligned}
& \frac{1}{N} \operatorname{Tr} \left(\frac{\mathbb{E} [\Delta_N | X_1, X_2, \dots, X_N] - \mathbb{E} \Delta_N}{\sqrt{N \varepsilon_N}} \right)^2 \\
&= \frac{\varepsilon_N}{N^2} \left[\sum_{i=1}^N \left(\sum_{j=1, j \neq i}^N (W(X_i, X_j) - d_W) \right)^2 + \sum_{i=1}^N \sum_{j=1, j \neq i}^N (W(X_i, X_j) - d_W)^2 \right] \\
&= \frac{\varepsilon_N}{N^2} \sum_{i=1}^N \left(\sum_{j=1, j \neq i}^N W(X_i, X_j) - d_W \right)^2 + O(\varepsilon_N) \\
&= \frac{\varepsilon_N}{N^2} \sum_{i=1}^N \sum_{j \neq k, j, k \neq i} (W(X_i, X_k) - d_W) (W(X_i, X_j) - d_W) + O(\varepsilon_N)
\end{aligned}$$

Then by (3.2.2),

$$\begin{aligned}
& \mathbb{E} \frac{1}{N} \operatorname{Tr} \left(\frac{\mathbb{E} [\Delta_N | X_1, X_2, \dots, X_N] - \mathbb{E} \Delta_N}{\sqrt{N \varepsilon_N}} \right)^2 \\
&= \frac{\varepsilon_N}{N^2} \mathbb{E} \left[\sum_{i=1}^N \sum_{j \neq k, j, k \neq i} (W(X_i, X_k) - d_W) (W(X_i, X_j) - d_W) \right] + O(\varepsilon_N) \\
&= \frac{\varepsilon_N}{N^2} \sum_{i=1}^N \sum_{j \neq k, j, k \neq i} \left(\mathbb{E} W(X_i, X_k) W(X_i, X_j) - d_W^2 \right) + O(\varepsilon_N) = O(\varepsilon_N)
\end{aligned}$$

Finally an application of [5, Corollary A.41] proves (3.2.1). \square

3.3. Constrained Random Graph. Constraints in the random graph are very important especially in the context of Gibbs measure related to the graph. The constraints impose certain dependence in the graph. Let S_N be the set of all simple graphs on N vertices. One of the important classes is the so-called *canonical ensemble* which puts a probability distribution on the set of simple graphs in a way that entropy is maximized and the average degree is equal to a prescribed value. There has been a recent interest in the breaking of ensemble equivalence study where the canonical ensembles play a crucial role. We refer to Squartini et al. [34], Touchette [35] for further details. The spectrum of the adjacency matrix for constrained random graph was derived in Chakrabarty et al. [11] and here we derive similar results for the Laplacian matrix.

The canonical ensemble measure P_N is the unique probability distribution on S_N with the following two properties:

- (I) The *average degree* of vertex i , defined by $\sum_{G \in S_N} k_i(G) P_N(G)$, equals k_i^* for all $1 \leq i \leq N$, where $k^* = (k_i^*) : 1 \leq i \leq N$ is a fixed sequence of positive integers of which we only require to be graphical.
- (II) The *entropy* of P_N , defined by $-\sum_{G \in S_N} P_N(G) \log P_N(G)$, is maximal.

It is known that because of property (II), P_N takes the form

$$P_N(G) = \frac{1}{Z_N(\theta^*)} \exp \left(- \sum_{i=1}^N \theta_i^* k_i(G) \right), \quad G \in S_N,$$

where $\theta^* = (\theta_i^* : 1 \leq i \leq N)$ is a sequence of real-valued Lagrange multipliers that must be chosen in such a way that property (I) is satisfied. The normalisation constant $Z_N(\theta^*)$, which depends on θ^* , is called the partition functions in Gibbs theory. The matching of property (I) uniquely fixes θ^* , namely, it turns out that (Squartini et al. [34])

$$P_N(G) = \prod_{1 \leq i < j \leq N} (p_{ij}^*)^{A_N[G](i,j)} (1 - p_{ij}^*)^{1 - A_N[G](i,j)}, \quad G \in S_N,$$

where $A_N[G]$ is the adjacency matrix of G , and p_{ij}^* represent a *reparametrisation* of the Lagrange multipliers, namely,

$$p_{ij}^* = \frac{x_i^* x_j^*}{1 + x_i^* x_j^*}, \quad 1 \leq i \neq j \leq N, \quad (3.3.1)$$

with $x_i^* = e^{-\theta_i^*}$. Thus, we see that P_N is nothing other than an inhomogenous Erdős-Rényi random graph where the probability that vertices i and j are connected by an edge equals p_{ij}^* . In order to match property (I), these probabilities must satisfy

$$k_i^* = \sum_{j \neq i, j=1}^N p_{ij}^*, \quad 1 \leq i \leq N, \quad (3.3.2)$$

which constitutes a set of N equations for the N unknowns x_1^*, \dots, x_N^* .

In order to state the next result, we need to make some assumptions on the sequence $(k_{Ni}^* : 1 \leq i \leq N)$. For the sake of simplification the dependence on N would be suppressed from notation.

Proposition 3.4. *Let $(k_i^* : 1 \leq i \leq N)$ be a graphical sequence of positive integers. Define*

$$m_N = \max_{1 \leq i \leq N} k_i^*$$

Assume that

$$\lim_{N \rightarrow \infty} m_N = \infty, \quad \lim_{N \rightarrow \infty} m_N / \sqrt{N} = 0, \quad (3.3.3)$$

and consider the graph $G^{k_N} = \left([N], \{(i, j) : 1 \leq i < j\}, (k_i^* k_j^* / m_N)_{1 \leq i < j \leq N} \right)$ and the corresponding empirical graphon W_N^k . Further assume there exists a graphon $W \in \mathcal{W}_0$ such that

$$\delta_{\square}(W_N^k, W) \rightarrow 0.$$

Let x_i^* and p_{ij}^* be determined by (3.3.1) and (3.3.2). Let Δ_N be the Laplacian matrix of an inhomogenous Erdős-Rényi random graph on N vertices, with p_{ij}^* the probability of an edge being present between vertices i and j for $1 \leq i \neq j \leq N$. Then there exists a unique, symmetric probability measure ν_k on \mathbb{R} such that

$$\lim_{N \rightarrow \infty} \text{ESD} \left(\frac{1}{\sqrt{N} \varepsilon_N} (\Delta_N - \mathbb{E}(\Delta_N)) \right) = \nu_k, \text{ weakly in probability} \quad (3.3.4)$$

where $\varepsilon_N = \frac{m_N^2}{\sum_{1 \leq l \leq N} k_l^*}$. Further if W is positive on an open set in $[0, 1]$, then support of ν_k is unbounded.

Proof. Consider A_N to be adjacency matrix of the inhomogenous Erdős-Rényi graph defined in the proposition. Define

$$\sigma_N = \sum_{1 \leq l \leq N} k_l^*,$$

It is known that (Squartini et al. [34])

$$\max_{1 \leq l \leq N} x_l^* = o(1),$$

in which case (3.3.1) and (3.3.2) gives

$$x_i^* = [1 + o(1)] \frac{k_i^*}{\sqrt{\sigma_N}}, \quad p_{ij}^* = [1 + o(1)] \frac{k_i^* k_j^*}{\sigma_N}, \text{ as } N \rightarrow \infty \quad (3.3.5)$$

with the error term *uniform* in $1 \leq i \neq j \leq N$. Then by definition

$$\varepsilon_N = \frac{m_N^2}{\sigma_N} \quad (3.3.6)$$

It follows from (3.3.3) that

$$\lim_{N \rightarrow \infty} \varepsilon_N = 0, \quad \lim_{N \rightarrow \infty} N\varepsilon_N = \infty$$

Define $B_N = \frac{A_N}{\sqrt{\varepsilon_N}}$, it is easy to observe that

$$|B_N(i, j) - \mathbb{E}B_N(i, j)| \leq \frac{2}{\sqrt{\varepsilon_N}}$$

The bound above immediately implies B_N satisfies the Lindeberg condition. Remember that by definition the empirical graphon W_N^k is such that

$$W_N(x, y) = \begin{cases} \frac{k_i^* k_j^*}{m_N^2} & \text{if } (x, y) \in I_i \times I_j, 1 \leq i \neq j \leq N \\ 0 & \text{if } (x, y) \in I_i \times I_i, 1 \leq i \leq N \end{cases}$$

For $1 \leq i \neq j \leq N$,

$$\begin{aligned} \text{Var}(B_N(i, j)) &= \frac{1}{\varepsilon_N} p_{ij}^* (1 - p_{ij}^*) \\ &= \frac{k_i^* k_j^*}{m_N^2} \left(1 - \varepsilon_N \frac{k_i^* k_j^*}{m_N^2} \right) + o(1). \end{aligned}$$

Define an empirical graphon \tilde{W}_N^k as

$$\tilde{W}_N^k(x, y) = \begin{cases} \text{Var}(B_N(i, j)) & \text{if } (x, y) \in I_i \times I_j, 1 \leq i \neq j \leq N \\ 0 & \text{if } (x, y) \in I_i \times I_i, 1 \leq i \leq N \end{cases}$$

Then by definition, for all $1 \leq i \neq j \leq N$,

$$\left| W_N^k(x, y) - \tilde{W}_N^k(x, y) \right| \leq \left| \varepsilon_N \left(\frac{k_i^* k_j^*}{m_N^2} \right)^2 \right| + o(1) \xrightarrow{N \rightarrow \infty} 0, \quad \forall (x, y) \in I_i \times I_j \quad (3.3.7)$$

DCT combined with (3.3.7) gives

$$\delta_{\square}(W_N^k, \tilde{W}_N^k) \leq \int_{[0,1]^2} \left| W_N^k(x, y) - \tilde{W}_N^k(x, y) \right| dx dy \rightarrow 0.$$

Thus we have $\delta_{\square}(\tilde{W}_N^k, W) \rightarrow 0$. Hence by an appeal to Theorem 2.1 the proof is completed. \square

Remark 3.1. A concrete example of a graphical sequence $(k_i^* : 1 \leq i \leq N)$ satisfying (3.3.3)-(3.3.4) can be constructed from the example considered in Chakrabarty et al. [11, Remark 5.1]. For $N \geq 1$, let

$$k_i^* = \lfloor i^{1/3} \rfloor, \quad 1 \leq i \leq N$$

Then Theorem 7.12 from van der Hofstad [37] implies that $(k_i^* : 1 \leq i \leq N)$ is graphical for N large enough. Since $m_N = \lfloor N^{1/3} \rfloor$ it is immediate that (3.3.3) holds. Define

$$W(x, y) = (xy)^{1/3}, \quad (x, y) \in [0, 1]^2$$

Then using the L^1 bound on $\delta_{\square}(W_N^k, W)$, where W_N^k is as defined in Proposition 3.4, we can show that

$$\delta_{\square}(W_N^k, W) \rightarrow 0$$

Among other examples one can consider also the random block matrices and sparse stochastic block models considered in Zhu [41]. The results for the Laplacian matrices hold under the assumptions stated in Section 6 and Section 7 of Zhu [41]. To avoid repetitions we skip the results.

4. SIMULATIONS

This section is devoted to simulation study for a clearer picture of the above results. We consider two situations, firstly we consider the inhomogenous Erdős-Rényi graph and in the second case, we generate the elements of adjacency matrices independently having a Gaussian distribution while respecting the symmetry constant.

For the inhomogenous Erdős-Rényi graph we choose $f(x, y) = \sqrt{xy}$ i.e. f is a product of two same functions $r(t) = \sqrt{t}$ in $L^2[0, 1]$. Then, in the Figure 4(A) the eigenvalues of the centered Laplacian matrix are plotted under the scaling $\sqrt{N\varepsilon_N}$ and hence it is clear that the ESD converges to a symmetric distribution. In the simulation N and ε_N are chosen to be 1000 and 0.25 respectively. In Figure 4 we show a comparison with the limiting spectral distribution under usual Erdős-Rényi graph with edge retention probability 0.25.

For the Gaussian case, we consider two graphons W as $W(x, y) = \sqrt{xy}$ and $W(x, y) = \frac{1}{2}(x(1 - y) +$

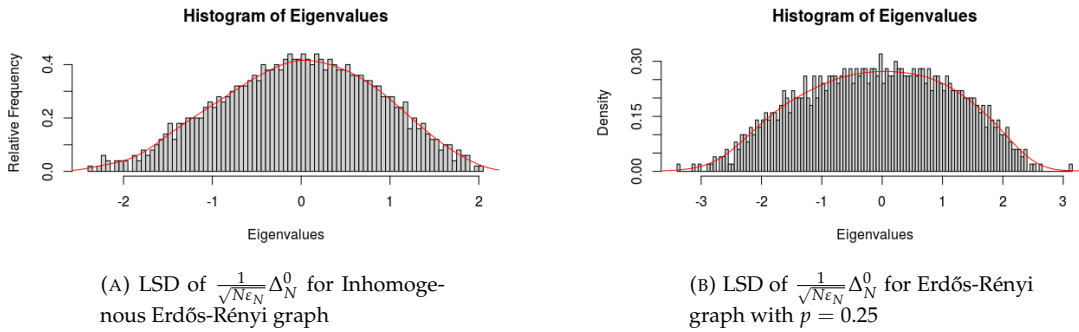


FIGURE 4. ESD in both Inhomogenous and Homogenous Erdős-Rényi graph

$y(1 - x)$). Then the elements of the adjacency matrix are generated as $X_{ij}^{(n)} \sim N\left(0, W\left(\frac{i}{n}, \frac{j}{n}\right)\right)$. The graphons considered here are uniformly continuous and hence by (3.1.5) the empirical graphon constructed using above variance profile converges in the cut norm to W . We consider $n = 1000$ so that the adjacency matrix is a 1000×1000 matrix. In Figure 5 the eigenvalues of the Laplacian matrix under the scaling \sqrt{n} and hence we can observe that ESD converges to a symmetric distribution.

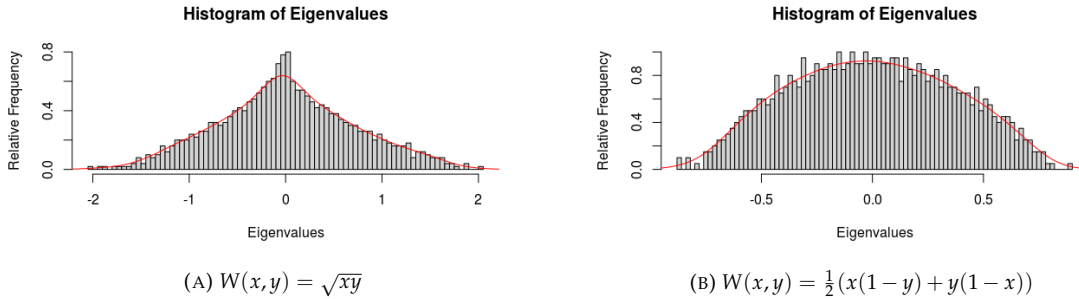


FIGURE 5. ESD of $\frac{1}{\sqrt{n}}\Delta_n$ when edge weights are Gaussian

5. PROOF OF THEOREMS 2.1, 2.2 AND COROLLARY 2.1

5.1. Preparatory Lemmas for Theorem 2.1. The proof of Theorem 2.1 rely on several preparatory lemmas which are organised in this section. One of the crucial steps in studying the properties of ESD is

to replace each entry by a Gaussian random variable, which we outline in the following lemma. Let $\{G_{i,j}, 1 \leq i \leq j\}$ be a family of i.i.d. standard Gaussian random variables. Define $N \times N$ matrices A_N^g and Δ_N^g by

$$A_N^g(i, j) = \frac{\sigma_{i \wedge j, i \vee j}}{\sqrt{N}} G_{i \wedge j, i \vee j} \quad \forall 1 \leq i, j \leq N \quad (5.1.1)$$

$$\Delta_N^g(i, j) = \begin{cases} A_N^g(i, j) & \text{if } i \neq j \\ -\sum_{k \neq i, k=1}^N A_N^g(i, k) & \text{if } i = j \end{cases} \quad (5.1.2)$$

Consider a three times continuously differentiable function $h : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\max_{0 \leq j \leq 3} \sup_{x \in \mathbb{R}} |h^{(j)}(x)| < \infty$$

where $h^{(j)}$ denotes the j -th derivative. For a $N \times N$ real symmetric matrix M define the Stieltjes transform of the ESD of M as

$$H_N(M) = \frac{1}{N} \text{Tr} \left((M - zI_N)^{-1} \right), \quad z \in \mathbb{C} \setminus \mathbb{R}.$$

The next result shows that real and imaginary part of Stieltjes transform of A_N^0 and Δ_N^0 are close to the Gaussian counterparts. Since one knows that convergence of ESD is equivalent to showing the convergence of the corresponding Stieltjes transform, one can work with Gaussian random variables.

Lemma 5.1. (Gaussianisation) *Let entries of A_N^0 satisfy the assumptions [L.1](#) and [L.2](#) and A_N^g and Δ_N^g be defined as in [5.1.1](#) and [5.1.2](#) then*

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[h \left(\mathcal{R}H_N \left(\Delta_N^g \right) \right) - h \left(\mathcal{R}H_N \left(\Delta_N^0 \right) \right) \right] = 0 \quad (5.1.3)$$

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[h \left(\mathcal{I}H_N \left(\Delta_N^g \right) \right) - h \left(\mathcal{I}H_N \left(\Delta_N^0 \right) \right) \right] = 0 \quad (5.1.4)$$

where \mathcal{R} and \mathcal{I} denotes the real and imaginary parts respectively. Similar statement holds true for A_N^0 and A_N^g .

The proof of Lemma 5.1 is routine and hence is skipped here and presented in the Appendix. The next lemma allows for minor tweaks in the diagonal entries of Δ_N^g .

Lemma 5.2. *Define a $N \times N$ matrix by*

$$\bar{A}_N(i, j) = \frac{\sigma_{i, j}}{\sqrt{N}} G_{i \wedge j, i \vee j}, \quad 1 \leq i, j \leq N \quad (5.1.5)$$

and let

$$\bar{\Delta}_N = \bar{A}_N - X_N \quad (5.1.6)$$

where X_N is a diagonal matrix of order N defined by

$$X_N(i, i) = \sum_{k \neq i} \bar{A}_N(i, k), \quad 1 \leq i \leq N.$$

Then

$$\lim_{N \rightarrow \infty} L \left(\text{ESD} \left(\Delta_N^g \right), \text{ESD} \left(\bar{\Delta}_N \right) \right) = 0 \text{ in probability.} \quad (5.1.7)$$

Proof. Observe that by Bai and Silverstein [[5](#), Corollary A.41]

$$\mathbb{E} \left[L^3 \left(\text{ESD} \left(\Delta_N^g \right), \text{ESD} \left(\bar{\Delta}_N \right) \right) \right] \leq \frac{1}{N} \mathbb{E} \left[\text{Tr} \left[\Delta_N^g - \bar{\Delta}_N \right]^2 \right] \quad (5.1.8)$$

Since $\bar{\Delta}_N$ and Δ_N^g differs only in the diagonal entries, then we have $\left(\bar{\Delta}_N - \Delta_N^g \right)_{i, i} = A_N^g(i, i)$, implying

$$\frac{1}{N} \mathbb{E} \left[\text{Tr} \left[\Delta_N^g - \bar{\Delta}_N \right]^2 \right] = \frac{1}{N} \mathbb{E} \left[\sum_{i=1}^N A_N^g(i, i)^2 \right] = \frac{1}{N^2} \sum_{i=1}^N \sigma_{i, i}^2 = O \left(\frac{1}{N} \right) \rightarrow 0$$

The last order coming from **L.1**. Hence $\mathbb{E} \left[L^3 \left(\text{ESD} \left(\Delta_N^g \right), \text{ESD} \left(\bar{\Delta}_N \right) \right) \right] \rightarrow 0$ which show (5.1.8). \square

The (diagonal) entries of X_N are nothing but the row sums of A_N^g . However the correlation between an entry of A_N^g and that of X_N is small. The following decoupling lemma, shows that it does not hurt when the entries of X_N are replaced by a mean-zero Gaussian random variable of the same variance that is independent of A_N^g .

Lemma 5.3. *Chakrabarty et al. [11, Lemma 2.4] Let $(Z_i : i \geq 1)$ be a family of i.i.d. standard normal random variables, independent of $(G_{i,j} : 1 \leq i \leq j)$. Define a diagonal matrix Y_N of order N by*

$$Y_N(i, i) = Z_i \sqrt{\frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^N \sigma_{i,j}^2}, \quad 1 \leq i \leq N$$

and let

$$\tilde{\Delta}_N = \bar{A}_N + Y_N \quad (5.1.9)$$

Then for every $k \in \mathbb{N}$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \left(\text{Tr} \left[(\tilde{\Delta}_N)^{2k} - (\bar{\Delta}_N)^{2k} \right] \right) = 0 \quad (5.1.10)$$

and

$$\lim_{N \rightarrow \infty} \frac{1}{N^2} \mathbb{E} \left(\text{Tr}^2 \left[(\tilde{\Delta}_N)^k \right] - \text{Tr}^2 \left[(\bar{\Delta}_N)^k \right] \right) = 0 \quad (5.1.11)$$

We skip the proof of the above lemma since it is verbatim same as Lemma 2.4 of Chakrabarty et al. [11]. In the next lemma we show the convergence of ESD of the above defined diagonal matrix Y_N .

Lemma 5.4. *Under the assumptions **L.1-L.3**, there exists a unique probability distribution ζ on \mathbb{R} such that*

$$\lim_{N \rightarrow \infty} \text{ESD} (Y_N) = \zeta \text{ weakly in probability}$$

Further if there exists an open set $U \subseteq [0, 1]^2$ such that $W > 0$ on U , then ζ have unbounded support.

Proof. We would be using method of moments to prove our result. After showing moment convergence we would show that the limits uniquely determine the distribution. And finally we would show the unbounded support.

Part 1: Convergence of Moments

Define m_k^N to be the k^{th} moment of ESD of Y_N ,

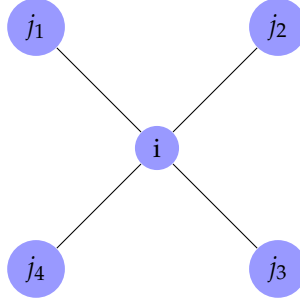
$$m_k^N = \frac{1}{N} \text{Tr} \left(Y_N^k \right) = \frac{1}{N} \sum_{i=1}^N \left(\sum_{\substack{j=1 \\ j \neq i}}^N \frac{\sigma_{i,j}^2}{N} \right)^{\frac{k}{2}} Z_i^k$$

We first show that variance of m_k^N goes to zero as $N \rightarrow \infty$. Note that using $\{Z_i\}$ is a collection of iid standard Gaussian random variables we have

$$\text{Var} \left(m_k^N \right) = \text{Var} \left(Z_1^k \right) \frac{1}{N^2} \sum_{i=1}^N \left(\sum_{\substack{j=1 \\ j \neq i}}^N \frac{\sigma_{i,j}^2}{N} \right)^k.$$

Now define

$$S_1^k := \frac{1}{N} \sum_{i=1}^N \left(\sum_{\substack{j=1 \\ j \neq i}}^N \frac{\sigma_{i,j}^2}{N} \right)^k = \frac{1}{N^{k+1}} \sum_{i=1}^N \sum_{j_1 \dots j_k \neq i} \sigma_{i,j_1}^2 \dots \sigma_{i,j_k}^2$$

FIGURE 6. Star graph F_4

and

$$\tilde{S}_1^k := \frac{1}{N^{k+1}} \sum_{i=1}^N \sum_{j_1 \dots j_k} \sigma_{i,j_1}^2 \dots \sigma_{i,j_k}^2. \quad (5.1.12)$$

It immediately follows that

$$\left| S_1^k - \tilde{S}_1^k \right| \leq \frac{1}{N^{k+1}} O(N^k) = O\left(\frac{1}{N}\right) \xrightarrow{N \rightarrow \infty} 0 \quad (5.1.13)$$

Consider the star graph $F_k = (V_k, E_k)$ with $k+1$ vertices and k edges, with the internal node labelled as i and the leaves labelled as j_1, \dots, j_k . An example of such a star graph is shown in Figure 6. Then we have (recall the definition of homomorphism density (2.2.1))

$$\begin{aligned} t(F_k, W_N) &= \int_{[0,1]^{k+1}} \prod_{(u,v) \in E_k} W_N(x_u, x_v) \prod_{u \in V_k} dx_u \\ &= \sum_{i, j_1 \dots j_k=1}^N \int_{I_i \otimes_{l=1}^k I_{j_l}} \sigma_{i,j_1}^2 \dots \sigma_{i,j_k}^2 dx_i dx_{j_1} \dots dx_{j_k} \\ &= \frac{1}{N^{k+1}} \sum_{i=1}^N \sum_{j_1 \dots j_k} \sigma_{i,j_1}^2 \dots \sigma_{i,j_k}^2 =: \tilde{S}_1^k \end{aligned}$$

Using [L.3](#) we have $t(F_k, W_N) \rightarrow t(F_k, W) < \infty$, implying

$$S_1^k \xrightarrow{N \rightarrow \infty} t(F_k, W). \quad (5.1.14)$$

Hence,

$$\text{Var}(m_k^N) = \frac{\text{Var}\left(Z_1^k\right)}{N} S_1^k \xrightarrow{N \rightarrow \infty} 0. \quad (5.1.15)$$

Now we show that $\mathbb{E}(m_k^N)$ converges. Observe that

$$\mathbb{E}(m_k^N) = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ \mu_k \frac{1}{N} \sum_{i=1}^N \left(\sum_{j \neq i} \frac{\sigma_{i,j}^2}{N} \right)^{\frac{k}{2}} & \text{if } k \text{ is even,} \end{cases}$$

where μ_k is the k^{th} moment of standard Gaussian distribution. If k is even, then we have $\mathbb{E}(m_k) = \mu_k S_1^{\frac{k}{2}}$. So combining (5.1.14) and (5.1.15), we infer

$$m_k^N \xrightarrow{L_2} \begin{cases} 0 & \text{if } k \text{ is odd} \\ \mu_k t\left(F_{\frac{k}{2}}, W\right) & \text{if } k \text{ is even} \end{cases} \quad (5.1.16)$$

Part 2: Uniqueness of the Limiting distribution

Before proceeding let us define

$$\eta_k = \begin{cases} 0 & \text{if } k \text{ is odd} \\ \mu_k t\left(F_{\frac{k}{2}}, W\right) & \text{if } k \text{ is even} \end{cases} \quad (5.1.17)$$

then by Lemma B.1 from Bai and Silverstein [5] we can easily see that there exists a probability measure ζ identified by the moments η_k . To deal with the uniqueness of ζ observe that due to L.1, for all k even we can find $C > 0$ such that $t(F_{k/2}, W) \leq C^{k/2}$. Then for $k \in \mathbb{N}$

$$\frac{1}{2k} \eta_{2k}^{\frac{1}{2k}} \preceq \frac{1}{2k} \left(\frac{2k!}{k!} \right)^{\frac{1}{2k}}. \quad (5.1.18)$$

Using the Sterling's approximation it is immediate that $\limsup_{k \rightarrow \infty} \frac{1}{2k} \eta_{2k}^{\frac{1}{2k}} < \infty$. By applying Theorem 1 of Lin [25], the probability measure ζ is uniquely identified by the moment sequence η_k .

Part 3: Unbounded support of μ

Suppose there exists an open $U \subseteq [0, 1]^2$ such that $W > 0$ on it. For $k \in \mathbb{N}$ define

$$\beta_k = \begin{cases} 0 & \text{if } k \text{ is odd} \\ t\left(F_{\frac{k}{2}}, W\right) & \text{if } k \text{ is even} \end{cases}$$

Since $W > 0$ on U , then by definition $t\left(F_{\frac{k}{2}}, W\right) > 0$. Then again using Lemma B.1 from Bai and Silverstein [5] there exists a probability measure κ having the above moment sequence. By Theorem 1 of Lin [25], we can say that the moment sequence β_k uniquely identifies the probability measure κ . Now consider $X \sim \kappa$ and consider $Z \sim N(0, 1)$ independently of X . Then it is easy to observe that XZ have the moment sequence η_k and hence $XZ \sim \zeta$. It easily follows from here that if we take any $M > 0$ then $P(XZ > M) > 0$ and

$$\sup(\text{supp}(XZ)) = \sup(\text{supp}(\zeta)) = \infty$$

This completes the proof of Lemma 5.4. □

Notations: We now recall some notations from Chakrabarty et al. [11] and we refer the reader to Nica and Speicher [29] for the combinatorial properties of non-crossing pair partition and Kreweras complement.

For $k \in \mathbb{N}$ and Π a partition of $\{1, \dots, 2k\}$, define

$$\Psi(\Pi, N) = \left\{ i \in \{1, \dots, N\}^{2k} : i_u = i_v \iff u, v \text{ belong to the same block of } \Pi \right\}$$

For an even positive integer k , $NC_2(k)$ is the set of non-crossing pair partitions of $\{1, 2, \dots, k\}$. For $\sigma \in NC_2(k)$, its Kreweras complement $K(\sigma)$ is the maximal non-crossing partition $\bar{\sigma}$ of $\{\bar{1}, \dots, \bar{k}\}$, such that $\sigma \cup \bar{\sigma}$ is a non-crossing partition of $\{1, \bar{1}, \dots, k, \bar{k}\}$. For $\sigma \in NC_2(k)$ and $N \geq 1$, define

$$S(\sigma, N) = \{i \in \{1, \dots, N\}^k : i_u = i_v \iff u, v \text{ belong to the same block of } K(\sigma)\} \quad (5.1.19)$$

and

$$C(k, N) = \{1, \dots, N\}^k \setminus \left(\bigcup_{\sigma \in NC_2(k)} S(\sigma, N) \right) \quad (5.1.20)$$

In other words, $S(\sigma, N)$ is the same as $\Psi(K(\sigma), N)$. We will use this fact in the upcoming proof.

5.2. Proof of Theorem 2.1. Combining Lemmas 5.1 and 5.2 we find that $\text{ESD}(\Delta_N^0)$ and $\text{ESD}(\bar{\Delta}_N)$ have the same in probability limit. Next we use method of moments on $\text{ESD}(\bar{\Delta}_N)$. By Lemma 5.3 it is enough to look at the moments of $\text{ESD}(\tilde{\Delta}_N)$ where $\tilde{\Delta}_N$ is as defined in (5.1.9). The rest of the proof is organised as follows. First we show the L_2 convergence of moments of $\text{ESD}(\tilde{\Delta}_N)$. We work separately for even and odd moment. Convergence of the even moment is more involved, and the odd moment convergence follows along similar lines. In case of even moments we appeal to the fact that if X is a random variable and $\mathbb{E}X$ converges to α for some $\alpha \in \mathbb{R}$, and $\mathbb{E}X^2$ converges to α^2 , then X converges to α is L_2 . We will follow the combinatorial ideas from Zhu [41] and express the expected value of moments in terms of graph homomorphism and use graphon convergence assumption (L.3) to find the limit. Finally an appeal to Theorem 1 from Lin [25] would show the uniqueness of the distribution.

5.2.1. Convergence of Moments. Fix $k \in \mathbb{N} \cup \{0\}$. We are interested to look at the L_2 convergence of k^{th} moment of $\text{ESD}(\tilde{\Delta}_N)$ given by

$$\frac{1}{N} \text{Tr} \left[(\tilde{\Delta}_N)^k \right].$$

We would deal with the odd and even moments separately. We shall show that the odd moments converge to 0 and the even moments converge to $\sum_{\mathcal{P}_k} \sum_{\sigma \in \text{NC}_2(\sum m_p)} \beta(\sigma) \mathcal{E}(\sigma)$ where \mathcal{P}_k is defined in Section 2.5, $\beta(\sigma) \geq 0$ and $\mathcal{E}(\sigma)$ will be defined in (5.2.20) and (5.2.21) respectively.

Case 1: k is even

By definition we have

$$\frac{1}{N} \text{Tr} \left(\tilde{\Delta}_N^k \right) = \frac{1}{N} \sum_{\substack{m_1, m_2, \dots, m_k \\ n_1, n_2, \dots, n_k}} \text{Tr} \left(\bar{A}_N^{m_1} Y_N^{n_1} \dots \bar{A}_N^{m_k} Y_N^{n_k} \right) \quad (5.2.1)$$

where the sum is over all the terms in the expansion of $(\bar{A}_N + Y_N)^k$, i.e. we have 2^k many terms and for every choice of $m_1, n_1, \dots, m_k, n_k$ we have $\sum_{i=1}^k m_i + n_i = k$. We can take k many expressions in each term of the expansion, since we allow the exponents to be 0. So enough to look at L_2 convergence of $\frac{1}{N} \text{Tr} \left(\bar{A}_N^{m_1} Y_N^{n_1} \dots \bar{A}_N^{m_k} Y_N^{n_k} \right)$. Let

$$M_j = \sum_{p=1}^j m_j + n_j, \quad j = 1, \dots, k.$$

Observe that $M_k = k$. Take \tilde{i} such that $\tilde{i} = (\tilde{i}_1, \dots, \tilde{i}_{M_k+1}) \in \{1, \dots, N\}^{M_k+1}$ and $\tilde{i}_{1+M_k} = \tilde{i}_1$. Then

$$\begin{aligned} & \frac{1}{N} \text{Tr} \left(\bar{A}_N^{m_1} Y_N^{n_1} \dots \bar{A}_N^{m_k} Y_N^{n_k} \right) \\ &= \frac{1}{N} \sum_{\tilde{i}} \prod_{j=1}^{m_1} \bar{A}_N(i_j, i_{j+1}) \prod_{j=m_1+1}^{M_1} Y_N(i_j, i_{j+1}) \dots \prod_{j=1+M_{k-1}}^{m_k+M_{k-1}} \bar{A}_N(i_j, i_{j+1}) \prod_{j=1+M_k+M_{k-1}}^{M_k} Y_N(i_j, i_{j+1}) \\ &= \frac{1}{N} \sum_i \prod_{j=1}^{m_1} \bar{A}_N(i_j, i_{j+1}) Y_N^{n_1}(i_{m_1+1}, i_{m_1+1}) \dots \prod_{j=1+\sum_{p=1}^{k-1} m_p}^{\sum_{p=1}^k m_p} \bar{A}_N(i_j, i_{j+1}) Y_N^{n_k}(i_1, i_1) \\ &= \frac{1}{N} \sum_{i_1, i_2, \dots, i_{\sum m_p}=1}^N \prod_{j=1}^{\sum m_p} \bar{A}_N(i_j, i_{j+1}) \prod_{j=1}^k Y_N^{n_j} \left(i_{1+\sum_{p=1}^j m_p}, i_{1+\sum_{p=1}^j m_p} \right) \\ &= \frac{1}{N} \sum_{i_1 \dots i_{\sum m_p}} \prod_{j=1}^{\sum m_p} G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}} \prod_{j=1}^{\sum m_p} \frac{\sigma_{i_j, i_{j+1}}}{\sqrt{N}} \prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i_{1+\sum_{p=1}^j m_p}} \sigma_{i_{1+\sum_{p=1}^j m_p}, t}^2 \right)^{\frac{n_j}{2}} \prod_{j=1}^k Z_{i_{1+\sum_{p=1}^j m_p}}^{n_j}. \quad (5.2.2) \end{aligned}$$

Here recall that $(Z_i : i \geq 1)$ is a family of i.i.d. standard Normal random variables, independent of $(G_{i,j} : 1 \leq i \leq j)$ which are also independent standard Normal random variables. Also above $i =$

$(i_1, i_2, \dots, i_{\sum m_p})$ with $i_{1+\sum_{p=1}^k m_p} = i_1$. Using definition of Kreweras complement we have the following decomposition of (5.2.2)

$$\sum_{i_1, i_2, \dots, i_{\sum m_p}} = \sum_{i \in C(\sum m_p, N)} + \sum_{\sigma \in NC_2(\sum m_p)} \sum_{i \in S(\sigma, N)} \quad (5.2.3)$$

Now $i \in S(\sigma, N)$ is same as saying $i \in \Psi(K(\sigma), N)$. Observe that

$$\begin{aligned} \mathbb{E} \left[\frac{1}{N} \text{Tr} (\bar{A}_N^{m_1} Y_N^{n_1} \dots \bar{A}_N^{m_k} Y_N^{n_k}) \right] &= \frac{1}{N} \sum_{i_1, \dots, i_{\sum m_p}} \mathbb{E} \left[\prod_{j=1}^{\sum m_p} G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}} \right] \prod_{j=1}^{\sum m_p} \frac{\sigma_{i_j, i_{j+1}}}{\sqrt{N}} \times \\ &\prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i} \sum_{1+\sum_{p=1}^j m_p} \sigma_{i, t}^2 \right)^{\frac{n_j}{2}} \mathbb{E} \left[\prod_{j=1}^k Z_{i_{1+\sum_{p=1}^j m_p}}^{n_j} \right]. \end{aligned} \quad (5.2.4)$$

Using (5.2.3) let us first look at the sum over $S(\sigma, N)$. Consider a partition Π of $\{1, 2, \dots, \sum m_p\}$, and take $i \in \Pi$, then observe that $\mathbb{E} \left[\prod_{j=1}^{\sum m_p} G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}} \right]$ does not depend on i , but on the partition Π . Define

$$\Phi(\Pi) := \mathbb{E} \left[\prod_{j=1}^{\sum m_p} G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}} \right].$$

First we focus on the factor

$$\mathbb{E} \left[\prod_{j=1}^k Z_{i_{1+\sum_{p=1}^j m_p}}^{n_j} \right].$$

from (5.2.4). One easy observation is that it only depends upon the partition Π where i belongs. Now consider a partition Π and suppose $i \in \Pi$. For notational simplicity we identify 1 by $1 + \sum_{p=1}^k m_p$. Then consider the blocks where the indices $\{m_1 + 1, m_1 + m_2 + 1, \dots, 1 + \sum_{p=1}^k m_p\}$ belongs. Then we have

$$\mathbb{E} \left[\prod_{j=1}^k Z_{i_{1+\sum_{p=1}^j m_p}}^{n_j} \right] = \mathbb{E} \left[\prod_{u \in \Pi} \prod_{\substack{j=1 \\ 1+\sum_{p=1}^j m_p \in u}}^k Z_{i_{1+\sum_{p=1}^j m_p}}^{n_j} \right] = \prod_{u \in \Pi} \mathbb{E} \left[\prod_{\substack{j=1 \\ 1+\sum_{p=1}^j m_p \in u}}^k Z_{l_u}^{n_j} \right] \quad (5.2.5)$$

Where u denotes a block in Π and l_u denotes the corresponding representative element. So now if for some block $u \in \Pi$,

$$\sum_{\substack{j=1 \\ 1+\sum_{p=1}^j m_p \in u}}^k n_j \equiv 1 \pmod{2} \quad (5.2.6)$$

then the expectation in (5.2.5) is 0, and hence in turn the whole expression is 0. So while looking at $\sum_{i \in S(\sigma, N)}$ in (5.2.3) we would only be looking at the case where for all $u \in K(\sigma)$

$$\sum_{\substack{j=1 \\ 1+\sum_{p=1}^j m_p \in u}}^k n_j \equiv 0 \pmod{2} \quad (5.2.7)$$

holds. Now remember the expectation in (5.2.5) does not depend on the choice of $i \in S(\sigma, N)$, hence can go out of the sum. So we would be focusing on

$$\frac{1}{N} \sum_{i \in S(\sigma, N)} \Phi(K(\sigma)) \prod_{j=1}^{\sum m_p} \frac{\sigma_{i_j, i_{j+1}}}{\sqrt{N}} \prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i} \sum_{1+\sum_{p=1}^j m_p} \sigma_{i_{1+\sum_{p=1}^j m_p}, t}^2 \right)^{\frac{n_j}{2}} \quad (5.2.8)$$

Define $m = \sum_{p=1}^k m_p$. We now use a combinatorial identity from proof of Theorem 1.2 in Chakrabarty et al. [11].

$$\lim_{N \rightarrow \infty} N^{-(\frac{m}{2}+1)} \Phi(\Pi) \# \Psi(\Pi, N) = \begin{cases} 1 & \text{if } m \text{ is even and } \Pi = K(\sigma) \text{ for some } \sigma \in NC_2(m) \\ 0 & \text{otherwise} \end{cases} \quad (5.2.9)$$

The above follows from standard arguments leading to proof of Wigner's semicircle law using method of moments. Observe that if

$$m \equiv 1 \pmod{2}, \text{ then } NC_2(m) = \emptyset \quad (5.2.10)$$

where \emptyset denotes the null set. Hence we can safely ignore this case. So when considering the sum over $NC_2(m)$ we assume that m is even. Since we are having $\Pi = K(\sigma)$, and recalling that $\Psi(K(\sigma), N) = S(\sigma, N)$, then using (5.2.9) we have

$$\begin{aligned} & \frac{1}{\# \Psi(K(\sigma), N)} \sum_{i \in \Psi(K(\sigma), N)} \prod_{j=1}^m \sigma_{i_j, i_{j+1}} \prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i} \sum_{1+\sum_{p=1}^j m_p} \sigma_{i_{1+\sum_{p=1}^j m_p}, t}^2 \right)^{\frac{n_j}{2}} \\ & \approx \frac{1}{N^{\frac{m}{2}+1}} \sum_{i \in \Psi(K(\sigma), N)} \prod_{j=1}^m \sigma_{i_j, i_{j+1}} \prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i} \sum_{1+\sum_{p=1}^j m_p} \sigma_{i_{1+\sum_{p=1}^j m_p}, t}^2 \right)^{\frac{n_j}{2}} \end{aligned} \quad (5.2.11)$$

The above follows from $\# \Psi(K(\sigma), N) = O(N^{\frac{m}{2}+1})$.²

Now let us consider the product $\prod_{j=1}^m \sigma_{i_j, i_{j+1}}$. Observe that if all coordinates of i were distinct then $i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_m \rightarrow i_1$ forms a closed walk (G_W) on m vertices. Now we have $i \in K(\sigma)$, and $K(\sigma)$ can have only $\frac{m}{2} + 1$ many distinct block implies there are only $\frac{m}{2} + 1$ many distinct values in i . Hence we have the following modification of G_W . Glue together the vertices i_a and i_b which appear in the same block. Since the previous graph G_W was a closed walk, which is a connected graph, then the new graph (denoted by $G = (V, E)$) will be connected. Observe V is the blocks in i , that is, the blocks in Π and E is the edges between them (without repetition). Observe that G only depends upon the positions in i which are equal and which are not, hence the graph is independent of choice of $i \in S(\sigma, N)$. G only depends upon the blocks of $K(\sigma)$.

In the product $\prod_{j=1}^m \sigma_{i_j, i_{j+1}}$ the number of times the unordered pair (i_j, i_{j+1}) would appear is same as the number of times the edge between i_j and i_{j+1} is traversed in the graph G while following the previous closed walk. If i_j and i_{j+1} belong in the same block, then the edge between them is basically a self loop. Since we are looking undirected graph, then the total number of repetition of the edge (i_j, i_{j+1}) is the same as the total number of appearance of (i_j, i_{j+1}) and (i_{j+1}, i_j) in the product. This takes care of the symmetry constraint. Then we have

$$\prod_{j=1}^m \sigma_{i_j, i_{j+1}} = \prod_{e \in E} \sigma_e^{t_e} \quad (5.2.12)$$

² \approx implies they are same in the limit.

where $e = (a, b)$ denotes the edge between vertex a and b and t_e is the number of times the edge is repeated in the closed walk. Independence of the graph from i gives

$$\Phi(K(\sigma)) = \mathbb{E} \left[\prod_{e \in E} G_e^{t_e} \right]$$

We consider three exhaustive cases

- C.1** $\forall e \in E, t_e = 2$;
- C.2** $\exists e \in E$ such that $t_e = 1$;
- C.3** $\forall e \in E, t_e \geq 2$ and $\exists e \in E$ such that $t_e > 2$.

Let us first deal with **C.2**. Suppose e_1 is the edge appearing only once. Then

$$\Phi(K(\sigma)) = \mathbb{E}(G_{e_1}) \mathbb{E} \left(\prod_{e \in E \setminus \{e_1\}} G_e^{t_e} \right) = 0 \quad \text{since } G_e \sim N(0, 1)$$

But this would contradict (5.2.9). Hence this case does not happen.

Now let us look at **C.3**. Using (5.2.12) we have $\sum_{e \in E} t_e = m$. By **C.3** we have $\sum_{e \in E} t_e > 2|E|$, where $|E|$ denotes the number of edges in G . Since G is a connected graph, then $|E| \geq |V| - 1 = \frac{m}{2}$. Then $\sum_{e \in E} t_e > m$. This is a contradiction again and hence this case is also not possible.

In order to deal with **C.1** we break it into two sub-cases.

Sub-Case 1: The graph G has a self edge. Then we have a connected graph G having a self loop. Even if we remove the self loop still the graph remains connected. Say, e_1 is the self loop. Then consider the spanning tree of G which we denote by S , then $|E(S)| = m/2$, then observe

$$t_{e_1} + \sum_{e \in E(S)} t_e \leq \sum_{e \in E} t_e$$

which implies $2 + \frac{m}{2} \times 2 \leq m$ since here $t_e = 2$ for all $e \in E$. This is a contradiction and it shows that the graph cannot have a self edge.

Sub-Case 2: We have a connected graph G having no self loop and $\frac{m}{2} + 1$ vertices and $t_e = 2, \forall e \in E$. Then using $\sum_{e \in E} t_e = m$ we have $|E| = \frac{m}{2}$. Hence G is a tree.

Since the graph G depends only on $\sigma \in NC_2(m)$, then we redefine it as $G \equiv T(\sigma) = (V(\sigma), E(\sigma))$, where $V(= V(\sigma))$ and $E(= E(\sigma))$ denotes the vertices and edges respectively.

Consider the s^{th} block of $K(\sigma)$ and say the representative element is l_s and define

$$\gamma_s = \# \left\{ 1 \leq j \leq k : 1 + \sum_{i=1}^j m_i \in \{s^{th} \text{ block}\} \right\}$$

and

$$\{s_1, s_2, \dots, s_{\gamma_s}\} = \left\{ 1 \leq j \leq k : 1 + \sum_{i=1}^j m_i \in \{s^{th} \text{ block}\} \right\}$$

Then, since $i_a = i_b$ if a, b are in the same block, it is easy to see that

$$\prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i} \sum_{1 + \sum_{p=1}^j m_p = t} \sigma_{i \rightarrow j}^2 m_p^t \right)^{\frac{n_j}{2}} = \prod_{s=1}^{\frac{m}{2} + 1} \left(\frac{1}{N} \sum_{t \neq l_s} \sigma_{l_s, t}^2 \right)^{\sum_{j=1}^{\gamma_s} n_{s_j} / 2} \quad (5.2.13)$$

Observe that from (5.2.7) we have

$$\sum_{j=1}^{\gamma_s} n_{s_j} = \sum_{\substack{j=1 \\ 1+\sum_{p=1}^j m_p \in s}}^k n_j \equiv 0 \pmod{2}.$$

So define $\tilde{n}_s = \sum_{j=1}^{\gamma_s} n_{s_j}/2$ and then $\sum_s \tilde{n}_s = \frac{1}{2} \sum_{j=1}^k n_j$, where \sum_s denotes sum over the blocks of $K(\sigma)$. Then

$$\prod_{s=1}^{\frac{m}{2}+1} \left(\frac{1}{N} \sum_{t \neq l_s} \sigma_{l_s, t}^2 \right)^{\sum_{j=1}^{\gamma_s} n_{s_j}/2} = \frac{1}{N^{\sum_s \tilde{n}_s}} \prod_{s=1}^{\frac{m}{2}+1} \left(\sum_{p_1 \dots p_{\tilde{n}_s} \neq l_s} \sigma_{l_s, p_1}^2 \dots \sigma_{l_s, p_{\tilde{n}_s}}^2 \right) \quad (5.2.14)$$

Then combining (5.2.13) and (5.2.14), (5.2.11) becomes

$$\begin{aligned} & \frac{1}{N^{\frac{m}{2}+1}} \sum_{i \in \Psi(K(\sigma), N)} \prod_{j=1}^m \sigma_{i_j, i_{j+1}} \prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i} \sum_{1+\sum_{p=1}^j m_p \in t} \sigma_{i, t}^2 \right)^{\frac{n_j}{2}} \\ &= \frac{1}{N^{\frac{m}{2}+1}} \sum_{l_1 \neq \dots \neq l_{\frac{m}{2}+1}} \prod_{(u, v) \in E} \sigma_{l_u, l_v}^2 \prod_{s=1}^{\frac{m}{2}+1} \left(\frac{1}{N^{\tilde{n}_s}} \sum_{p_1, \dots, p_{\tilde{n}_s} \neq l_s} \sigma_{l_s, p_1}^2 \dots \sigma_{l_s, p_{\tilde{n}_s}}^2 \right) \end{aligned} \quad (5.2.15)$$

$$= \frac{1}{N^{1+\frac{\sum m_p + n_p}{2}}} \sum_{\substack{l_1 \neq \dots \neq l_{\frac{m}{2}+1} \\ p_{(s,1)} \dots p_{(s, \tilde{n}_s)} \neq l_s \\ \forall s=1 \dots \frac{m}{2}+1}} \prod_{(u, v) \in E} \sigma_{l_u, l_v}^2 \prod_{s=1}^{\frac{m}{2}+1} \prod_{t=1}^{\tilde{n}_s} \sigma_{l_s, p_{(s,t)}}^2 \quad (5.2.16)$$

where for a fixed $s \in \{1, \dots, 1 + m/2\}$, $p_{(s,i)}$ denotes the index p_i in (5.2.15). We modify the graph $T(\sigma)$ as follows. We take vertex s from $T(\sigma)$ and join \tilde{n}_s many vertices to it, denote those by $\{(s, 1), (s, 2), \dots, (s, \tilde{n}_s)\}$, so that s becomes the internal node of a star graph with the vertices $\{(s, 1), (s, 2), \dots, (s, \tilde{n}_s)\}$ forming the leaves. An example of the modification is shown in Figure 7. Consider the new graph to be

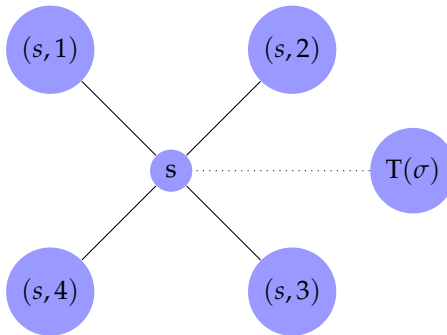


FIGURE 7. Star graph around vertex s , with $\tilde{n}_s = 4$

$$\tilde{T}(\sigma) = (\tilde{V}(\sigma), \tilde{E}(\sigma)), \text{ where } |\tilde{V}(\sigma)| = \frac{1}{2} \sum_{p=1}^k m_p + \frac{1}{2} \sum_s \tilde{n}_s + 1 = \frac{k}{2} + 1 \text{ and } |\tilde{E}(\sigma)| = \frac{m}{2} + \sum_s \tilde{n}_s = \frac{k}{2}.$$

Then observe that by construction $\tilde{T}(\sigma)$ is a tree and the homomorphism density $t(\tilde{T}(\sigma), W_N)$ becomes

$$\begin{aligned}
t(\tilde{T}(\sigma), W_N) &= \int_{[0,1]^{1+\sum \frac{m_p+n_p}{2}}} \prod_{(u,v) \in \tilde{E}} W_N(x_u, x_v) \prod_u dx_u \\
&= \sum_{\substack{l_1 \dots l_{m/2+1} \\ p_{(s,1)} \dots p_{(s,\tilde{n}_s)} \\ \forall s=1 \dots \frac{m}{2}+1}} \int_{\otimes_{j=1}^{\frac{m}{2}+1} I_{l_j} \otimes_{s=1}^{\frac{m}{2}+1} \otimes_{t=1}^{\tilde{n}_s} I_{p_{(s,t)}}} \prod_{(u,v) \in \tilde{E}} W_N(x_u, x_v) \prod_u dx_u \\
&= \frac{1}{N^{1+\frac{\sum m_p+n_p}{2}}} \sum_{\substack{l_1 \dots l_{m/2+1} \\ p_{(s,1)} \dots p_{(s,\tilde{n}_s)} \\ \forall s=1 \dots \frac{m}{2}+1}} \prod_{(u,v) \in E} \sigma_{l_u, l_v}^2 \prod_{s=1}^{\frac{m}{2}+1} \prod_{t=1}^{\tilde{n}_s} \sigma_{l_s, p_{(s,t)}}^2
\end{aligned} \tag{5.2.17}$$

$$\stackrel{N \rightarrow \infty}{\longrightarrow} t(\tilde{T}(\sigma), W) := \beta(\sigma) \tag{5.2.18}$$

where the last limit follows from assumption **L.1**. The expression from (5.2.16) is redefined as

$$S_1 = \frac{1}{N^{1+\frac{\sum m_p+n_p}{2}}} \sum_{\substack{l_1 \neq \dots \neq l_{m/2+1} \\ p_{(s,1)} \dots p_{(s,\tilde{n}_s)} \neq l_s \\ \forall s=1 \dots \frac{m}{2}+1}} \prod_{(u,v) \in E} \sigma_{l_u, l_v}^2 \prod_{s=1}^{\frac{m}{2}+1} \prod_{t=1}^{\tilde{n}_s} \sigma_{l_s, p_{(s,t)}}^2$$

Then using (5.2.17) and a counting argument it follows that

$$\left| S_1 - t(\tilde{T}(\sigma), W_N) \right| = O\left(N^{-1-\frac{\sum m_p+n_p}{2}} N^{\frac{\sum m_p+n_p}{2}} \right) = O\left(\frac{1}{N} \right) \stackrel{N \rightarrow \infty}{\longrightarrow} 0$$

Hence using homomorphism density convergence from (5.2.18) we conclude

$$S_1 = \frac{1}{N^{1+\frac{\sum m_p+n_p}{2}}} \sum_{\substack{l_1 \neq \dots \neq l_{m/2+1} \\ p_{(s,1)} \dots p_{(s,\tilde{n}_s)} \neq l_s \\ \forall s=1 \dots \frac{m}{2}+1}} \prod_{(u,v) \in E} \sigma_{l_u, l_v}^2 \prod_{s=1}^{\frac{m}{2}+1} \prod_{t=1}^{\tilde{n}_s} \sigma_{l_s, p_{(s,t)}}^2 \stackrel{N \rightarrow \infty}{\longrightarrow} t(\tilde{T}(\sigma), W) \tag{5.2.19}$$

Recall that by construction $\tilde{T}(\sigma)$ depends only on σ for fixed values of m_i, n_i for all $1 \leq i \leq k$. Remember that we took (5.2.7) holding for all $u \in K(\sigma)$. Hence redefine

$$\beta(\sigma) = \begin{cases} t(\tilde{T}(\sigma), W) & \text{if (5.2.7) holds for all } u \in K(\sigma) \\ 0 & \text{otherwise} \end{cases} \tag{5.2.20}$$

Also recall that $\mathbb{E} \left[\prod_{j=1}^k Z_i^{n_j} \right]_{1+\sum_{p=1}^j m_p}$ does not depend on choice of i , rather it only depends on σ . Define

$$\mathcal{E}(\sigma) = \mathbb{E} \left[\prod_{j=1}^k Z_i^{n_j} \right]_{1+\sum_{p=1}^j m_p}. \tag{5.2.21}$$

Note that by the discussion following (5.2.6), $\mathcal{E}(\sigma)$ becomes 0 if (5.2.7) does not hold for all $u \in K(\sigma)$. Then expectation of (5.2.2) becomes

$$\begin{aligned} \frac{1}{N} \sum_{i \in S(\sigma, N)} \mathbb{E} \left[\prod_{j=1}^{\sum m_p} G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}} \right] \prod_{j=1}^{\sum m_p} \frac{\sigma_{i_j, i_{j+1}}}{\sqrt{N}} \\ \prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i} \sum_{1+\sum_{p=1}^j m_p} \sigma_{i_{1+\sum_{p=1}^j m_p}, t}^2 \right)^{\frac{n_j}{2}} \mathbb{E} \left[\prod_{j=1}^k Z_{i_{1+\sum_{p=1}^j m_p}}^{n_j} \right] \\ \xrightarrow{N \rightarrow \infty} \beta(\sigma) \mathcal{E}(\sigma) \end{aligned} \quad (5.2.22)$$

where $\beta(\sigma)$ and $\mathcal{E}(\sigma)$ depends only on the choice of m_i, n_i for all $i = 1, 2, \dots, k$ and the corresponding partition σ . (Using (5.2.10) we can safely replace the above expression by 0 when $\sum_1^k m_p$ is odd).

Now observe that we have shown convergence of expectation. But our objective was to show L_2 convergence. For any $i \in \{1, 2, \dots, N\}^m$ define:

$$P_i = \prod_{j=1}^m G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}} \prod_{j=1}^m \frac{\sigma_{i_j, i_{j+1}}}{\sqrt{N}} \prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i} \sum_{1+\sum_{p=1}^j m_p} \sigma_{i_{1+\sum_{p=1}^j m_p}, t}^2 \right)^{\frac{n_j}{2}} \prod_{j=1}^k Z_{i_{1+\sum_{p=1}^j m_p}}^{n_j}$$

and by (5.2.22) we have $\mathbb{E} \left[\frac{1}{N} \sum_{i \in S(\sigma, N)} P_i \right] \rightarrow \beta(\sigma) \mathcal{E}(\sigma)$. To show that

$$\frac{1}{N} \sum_{i \in S(\sigma, N)} P_i \xrightarrow{L_2} \beta(\sigma) \mathcal{E}(\sigma),$$

it is enough to show that

$$\mathbb{E} \left[\left(\frac{1}{N} \sum_{i \in S(\sigma, N)} P_i \right)^2 \right] \rightarrow \beta^2(\sigma) \mathcal{E}^2(\sigma).$$

With that goal observe that

$$\mathbb{E} \left[\frac{1}{N} \sum_{i \in S(\sigma, N)} P_i \right]^2 \rightarrow \beta^2(\sigma) \mathcal{E}^2(\sigma).$$

Let us call $i, j \in \mathbb{N}^m$ to be disjoint if no co-ordinate of i matches any co-ordinates of j i.e. $\min_{1 \leq u, v \leq m} |i_u - j_v| \geq 1$. Now observe

$$\begin{aligned} \left[\mathbb{E} \left[\frac{1}{N} \sum_{i \in S(\sigma, N)} P_i \right] \right]^2 &= \frac{1}{N^2} \sum_{i, j \in S(\sigma, N)} \mathbb{E}(P_i) \mathbb{E}(P_j) \\ &= \frac{1}{N^2} \sum_{\substack{i, j \in S(\sigma, N) \\ \text{are disjoint}}} \mathbb{E}(P_i) \mathbb{E}(P_j) + \frac{1}{N^2} \sum_{\substack{i, j \in S(\sigma, N) \\ \text{are not disjoint}}} \mathbb{E}(P_i) \mathbb{E}(P_j) \end{aligned} \quad (5.2.23)$$

If i and j are not disjoint then there is at least one common coordinate. Which implies there is a block in i having same values as some block in j . Then we have

$$\{i, j \in S(\sigma, N) \mid i, j \text{ are not distinct}\} = \bigsqcup_{l=1}^{\frac{m}{2}+1} \{i, j \in S(\sigma, N) \mid i, j \text{ have exactly } l \text{ blocks common}\} \quad (5.2.24)$$

where \bigsqcup denotes disjoint union and define for $l = 1, 2, \dots, (\frac{m}{2} + 1)$,

$$H_l = \{i, j \in S(\sigma, N) : i, j \text{ have exactly } l \text{ blocks common}\}$$

By (5.2.8) we have

$$\begin{aligned} \mathbb{E}[P_i] &= N^{-\frac{m}{2}} \Phi(K(\sigma)) \prod_{j=1}^m \sigma_{i_j, i_{j+1}} \prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i} \sigma_{i_{1+\sum_{p=1}^j m_p}, t}^2 \right)^{\frac{n_j}{2}} \mathbb{E} \left[\prod_{j=1}^k Z_{i_{1+\sum_{p=1}^j m_p}}^{n_j} \right] \\ &\leq MN^{-\frac{m}{2}} \end{aligned} \quad (5.2.25)$$

for some $0 < M < \infty$, which depends only on σ . Observe that the above bound follows from L.1. Then using (5.2.25), we have for all $l \in \{1, 2, \dots, (\frac{m}{2} + 1)\}$,

$$\frac{1}{N^2} \sum_{i, j \in H_l} \mathbb{E}(P_i) \mathbb{E}(P_j) \leq \frac{1}{N^{m+2}} \sum_{i, j \in H_l} 1 \leq \frac{1}{N^{m+2}} O(N^{m+2-k}) \rightarrow 0 \text{ as } N \rightarrow \infty \quad (5.2.26)$$

Using (5.2.26) we conclude

$$\lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i, j \in H_l} \mathbb{E}(P_i) \mathbb{E}(P_j) = 0$$

Then using (5.2.24) we have

$$\lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{\substack{i, j \in S(\sigma, N) \\ \text{are not disjoint}}} \mathbb{E}(P_i) \mathbb{E}(P_j) = \lim_{N \rightarrow \infty} \sum_{l=1}^{\frac{m}{2}+1} \frac{1}{N^2} \sum_{i, j \in H_l} \mathbb{E}(P_i) \mathbb{E}(P_j) = 0 \quad (5.2.27)$$

Combining (5.2.27) and (5.2.23) we have

$$\lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{\substack{i, j \in S(\sigma, N) \\ \text{are disjoint}}} \mathbb{E}(P_i) \mathbb{E}(P_j) = \beta^2(\sigma) \mathcal{E}^2(\sigma) \quad (5.2.28)$$

We have

$$\begin{aligned} \mathbb{E} \left[\left(\frac{1}{N} \sum_{i \in S(\sigma, N)} P_i \right)^2 \right] &= \frac{1}{N^2} \sum_{i, j \in S(\sigma, N)} \mathbb{E}(P_i P_j) \\ &= \frac{1}{N^2} \sum_{\substack{i, j \in S(\sigma, N) \\ \text{are disjoint}}} \mathbb{E}(P_i) \mathbb{E}(P_j) + \frac{1}{N^2} \sum_{\substack{i, j \in S(\sigma, N) \\ \text{are not disjoint}}} \mathbb{E}(P_i P_j) \end{aligned} \quad (5.2.29)$$

Then for $i, j \in S(\sigma, N)$ such that they have at least one common coordinate, we have

$$\begin{aligned} \mathbb{E}(P_i P_j) &= N^{-m} \tilde{\Phi}_{i, j} \prod_{l=1}^m \sigma_{i_l, i_{l+1}}^2 \prod_{l=1}^k \left[\frac{1}{N} \sum_{t \neq i_{1+\sum_{p=1}^l m_p}} \sigma_{i_{1+\sum_{p=1}^l m_p}, t} \right]^{\frac{n_l}{2}} \\ &\quad \prod_{q=1}^m \sigma_{j_q, j_{q+1}}^2 \prod_{q=1}^k \left[\frac{1}{N} \sum_{t \neq j_{1+\sum_{p=1}^q m_p}} \sigma_{j_{1+\sum_{p=1}^q m_p}, t} \right]^{\frac{n_q}{2}} \mathbb{E} \left[\prod_{l=1}^k Z_{i_{1+\sum_{p=1}^l m_p}}^{n_l} Z_{j_{1+\sum_{p=1}^l m_p}}^{n_l} \right] \end{aligned} \quad (5.2.30)$$

where

$$\tilde{\Phi}_{i, j} = \mathbb{E} \left[\prod_{l=1}^m G_{i_l \wedge i_{l+1}, i_l \vee i_{l+1}} G_{j_l \wedge j_{l+1}, j_l \vee j_{l+1}} \right]$$

It is easy to observe that right hand side of (5.2.30) is bounded by some constant depending on $\sigma \in NC_2(m)$. Then as before we have $\mathbb{E}(P_i P_j) \leq \tilde{M} N^{-m}$ for some $\tilde{M} > 0$, depending on σ , this again follows

similarly as above from L.1. Then second term in (5.2.29) goes to 0 by similar counting argument as in (5.2.26). Then combining (5.2.28) and (5.2.29) we have

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\left(\frac{1}{N} \sum_{i \in S(\sigma, N)} P_i \right)^2 \right] = \beta^2(\sigma) \mathcal{E}^2(\sigma)$$

Hence we have shown that

$$\frac{1}{N} \sum_{i \in S(\sigma, N)} P_i \xrightarrow{L_2} \beta(\sigma) \mathcal{E}(\sigma)$$

Now we want to show that $\frac{1}{N} \sum_{i \in C(m, N)} P_i \rightarrow 0$ in L^2 . Observe here m can be odd or even. Define $D = \{\Pi \in \text{all partitions of } [m] \text{ such that } \Pi \neq K(\sigma), \forall \sigma \in NC_2(m)\}$. Observe that

$$\frac{1}{N} \sum_{i \in C(m, N)} P_i = \frac{1}{N} \sum_{\Pi \in D} \sum_{i \in \Psi(\Pi, N)} P_i$$

So enough to show

$$\frac{1}{N} \sum_{i \in \Psi(\Pi, N)} P_i \xrightarrow{L^2} 0 \text{ for all } \Pi \in D.$$

With that objective fix $\Pi \in D$,

$$\begin{aligned} \mathbb{E} \left[\left(\frac{1}{N} \sum_{i \in \Psi(\Pi, N)} P_i \right)^2 \right] &= \frac{1}{N^2} \sum_{i, j \in \Psi(\Pi, N)} \mathbb{E}(P_i P_j) \\ &= \frac{1}{N^2} \sum_{\substack{i, j \in \Psi(\Pi, N) \\ \text{are disjoint}}} \mathbb{E}(P_i) \mathbb{E}(P_j) + \frac{1}{N^2} \sum_{\substack{i, j \in \Psi(\Pi, N) \\ \text{are not disjoint}}} \mathbb{E}(P_i P_j) \end{aligned}$$

Define

$$\begin{aligned} E(\Pi) &= \{i, j \in \Psi(\Pi, N) : i, j \text{ are disjoint}\} \\ H_l(\Pi) &= \{i, j \in \Psi(\Pi, N) : i, j \text{ have exactly } l \text{ blocks common}\} \end{aligned}$$

Suppose number of blocks in Π is $b(\Pi)$. Then observe

$$\{i, j \in \Psi(\Pi, N) : i, j \text{ are not distinct}\} = \bigsqcup_{k=1}^{b(\Pi)} H_k(\Pi)$$

Recall that $\Phi(\Pi)$ is independent of choice of $i, j \in \Pi$, it only depends upon Π . Then

$$0 \leq \frac{1}{N^2} \sum_{i, j \in E(\Pi)} E(P_i) E(P_j) \leq \frac{1}{N^{m+2}} \Phi(\Pi)^2 |E(\Pi)| \frac{1}{|E(\Pi)|} \sum_{i, j \in E(\Pi)} 1 \quad (5.2.31)$$

$$\leq \left(\frac{1}{N^{\frac{m}{2}+1}} \Phi(\Pi) \#\Psi(\Pi, N) \right)^2 \rightarrow 0 \text{ as } N \rightarrow \infty. \quad (5.2.32)$$

The last limit follows from (5.2.9). Using above definitions we have

$$\frac{1}{N^2} \sum_{\substack{i, j \in \Psi(\Pi, N) \\ \text{are not disjoint}}} \mathbb{E}(P_i P_j) = \frac{1}{N^2} \sum_{k=1}^{b(\Pi)} \sum_{i, j \in H_k(\Pi)} \mathbb{E}(P_i P_j) \quad (5.2.33)$$

Let us look at the case when we have exactly one block common in i and j . Then $|H_1(\Pi)| = O(N^{2b(\Pi)-1})$. It is easy to see that there exists $h > 0$, depending on Π such that

$$\mathbb{E}(P_i P_j) \leq h N^{-m}, \quad \forall i, j \in \Psi(\Pi, N).$$

The above bound follows from the discussion after (5.2.30), replacing $K(\sigma)$ by Π . Now consider two cases depending upon the values of $b(\Pi)$.

Case 1: Let $b(\Pi) \leq \frac{m}{2} + 1$. Then we have

$$\frac{1}{N^2} \sum_{i,j \in H_1(\Pi)} \mathbb{E}(P_i P_j) \leq O\left(N^{-m-2} N^{m+1}\right) = O(N^{-1}) \xrightarrow{N \rightarrow \infty} 0. \quad (5.2.34)$$

Case 2: Let $b(\Pi) > \frac{m}{2} + 1$. Observe

$$\mathbb{E}(P_i P_j) \leq \frac{C}{N^m} \mathbb{E} \left[\prod_{l=1}^m G_{i_l \wedge i_{l+1}, i_l \vee i_{l+1}} \prod_{l=1}^m G_{j_l \wedge j_{l+1}, j_l \vee j_{l+1}} \right],$$

for some constant $C > 0$. Now consider the two closed walks (remember we had $i_{m+1} = i_1$ and $j_{m+1} = j_1$)

$$i_1 \rightarrow i_2 \rightarrow i_3 \rightarrow \dots \rightarrow i_m \rightarrow i_{m+1} \text{ and } j_1 \rightarrow j_2 \rightarrow j_3 \rightarrow \dots \rightarrow j_m \rightarrow j_{m+1}.$$

Now join them at the common coordinates that is where $i_p = j_q$ and for some $p, q \in \{1, \dots, m\}$ and glue together the coordinates of i which are in the same blocks of Π , and do the same for j . Then consider the graph $\widehat{G} = (\widehat{V}, \widehat{E})$ and then we have

$$\prod_{l=1}^m G_{i_l \wedge i_{l+1}, i_l \vee i_{l+1}} \prod_{l=1}^m G_{j_l \wedge j_{l+1}, j_l \vee j_{l+1}} = \prod_{e \in \widehat{E}} G_e^{t_e} \quad (5.2.35)$$

and hence

$$\mathbb{E}(P_i P_j) \leq \frac{C}{N^m} \mathbb{E} \left[\prod_{e \in \widehat{E}} G_e^{t_e} \right] \quad (5.2.36)$$

where t_e is the number of times the edge e is repeated in the graph \widehat{G} . Observe that each edge is repeated at least once, since the two closed walks were connected graph and in the resultant graph we did not remove any edge. So if $\mathbb{E}(P_i P_j) \neq 0$, then $\forall e \in \widehat{E}, t_e \geq 2$. Since $i, j \in H_1(\Pi)$, then total number of distinct block in the combined vector (i, j) would be $2b(\Pi) - 1$. Now by construction the blocks form the vertices of \widehat{G} , hence as the graph is connected then $|\widehat{E}| \geq 2b(\Pi) - 2$. Then

$$\sum_{e \in \widehat{E}} t_e \geq 2(2b(\Pi) - 2) > 2 \left(2 \left(\frac{m}{2} + 1 \right) - 2 \right) = 2m,$$

but by (5.2.35) we have $\sum_{e \in \widehat{E}} t_e = 2m$. Hence a contradiction which implies that in this case $\mathbb{E}(P_i P_j) = 0, \forall i, j \in H_1(\Pi)$.

So we have shown that

$$\frac{1}{N^2} \sum_{i,j \in H_1(\Pi)} \mathbb{E}(P_i P_j) \xrightarrow{N \rightarrow \infty} 0 \quad (5.2.37)$$

Similarly we can draw the same conclusion for any $l \in \{1, 2, \dots, b(\Pi)\}$. Hence we can conclude by (5.2.33) that

$$\frac{1}{N^2} \sum_{\substack{i,j \in \Psi(\Pi, N) \\ \text{are not distinct}}} \mathbb{E}(P_i P_j) \xrightarrow{N \rightarrow \infty} 0 \quad (5.2.38)$$

Then combining (5.2.32) and (5.2.38) we have proved the L_2 convergence for $C(m, N)$. Hence using (5.2.1) we have

$$\frac{1}{N} \text{Tr} \left(\widetilde{\Delta}_N^k \right) \xrightarrow{L_2} \sum_{\mathcal{P}_k} \sum_{\sigma \in NC_2(\Sigma_{m_p})} \beta(\sigma) \mathcal{E}(\sigma) \quad (5.2.39)$$

where \mathcal{P}_k is as defined in Section 2.5.

Case 2: k is odd

Recall the expansion of moment expression as in (5.2.1) in the beginning of the previous case.

$$\frac{1}{N} \operatorname{Tr} \left(\tilde{\Delta}_N^k \right) = \frac{1}{N} \sum_{\substack{m_1, m_2, \dots, m_k \\ n_1, n_2, \dots, n_k}} \operatorname{Tr} \left(\bar{A}_N^{m_1} Y_N^{n_1} \dots \bar{A}_N^{m_k} Y_N^{n_k} \right) \quad (5.2.40)$$

where the sum is over all the terms in the expansion of $(\bar{A}_N + Y_N)^k$, that is we have 2^k many terms and for every choice of $m_1, n_1, \dots, m_k, n_k$ we have $\sum_{p=1}^k m_p + n_p = k$. So in order to show $\frac{1}{N} \operatorname{Tr} \left(\tilde{\Delta}_N^k \right) \xrightarrow{L_2} 0$, it is enough to show that

$$\frac{1}{N} \operatorname{Tr} \left(\bar{A}_N^{m_1} Y_N^{n_1} \dots \bar{A}_N^{m_k} Y_N^{n_k} \right) \xrightarrow{L_2} 0$$

Then due to L.1 we have

$$\left| \frac{1}{N} \operatorname{Tr} \left(\bar{A}_N^{m_1} Y_N^{n_1} \dots \bar{A}_N^{m_k} Y_N^{n_k} \right) \right| \leq \frac{1}{N} \sum_{i_1 \dots i_{\sum m_p}} \left| \prod_{j=1}^{\sum m_p} G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}} \right| \left| \prod_{j=1}^{\sum m_p} \frac{1}{\sqrt{N}} \right| \left| \prod_{j=1}^k Z_{i_{1+\sum_{p=1}^j m_p}}^{n_j} \right| \quad (5.2.41)$$

Sub-Case 1: $\sum_{p=1}^k m_p$ is odd

Observe that from (5.2.41),

$$\begin{aligned} \mathbb{E} \left[\frac{1}{N^2} \operatorname{Tr}^2 \left(\bar{A}_N^{m_1} Y_N^{n_1} \dots \bar{A}_N^{m_k} Y_N^{n_k} \right) \right] &\leq \mathbb{E} \left[\frac{1}{N^{2+\sum m_p}} \sum_{i,j} \prod_{l=1}^{\sum m_p} G_{i_l \wedge i_{l+1}, i_l \vee i_{l+1}} \prod_{l=1}^{\sum m_p} G_{j_l \wedge j_{l+1}, j_l \vee j_{l+1}} \right] \\ &\quad \mathbb{E} \left[\prod_{l=1}^k Z_{i_{1+\sum_{p=1}^l m_p}}^{n_l} Z_{j_{1+\sum_{p=1}^l m_p}}^{n_l} \right] \end{aligned}$$

where i, j are the m dimensional vectors defined in (5.2.2). The last term is uniformly bounded over i, j . (Since there are finitely such values depending on the partitions of $\{1, \dots, m\}$, where $m = \sum_{j=1}^k m_j$). Then

$$\begin{aligned} \mathbb{E} \left[\frac{1}{N^2} \operatorname{Tr}^2 \left(\bar{A}_N^{m_1} Y_N^{n_1} \dots \bar{A}_N^{m_k} Y_N^{n_k} \right) \right] &\leq \mathbb{E} \left[\frac{1}{N^{2+\sum m_p}} \sum_{i,j} \prod_{l=1}^{\sum m_p} G_{i_l \wedge i_{l+1}, i_l \vee i_{l+1}} \prod_{l=1}^{\sum m_p} G_{j_l \wedge j_{l+1}, j_l \vee j_{l+1}} \right] \\ &\leq \mathbb{E} \left[\left(\frac{1}{N} \operatorname{Tr} \left(U_N^{\sum m_p} \right) \right)^2 \right] \xrightarrow{N \rightarrow \infty} 0 \end{aligned}$$

where U_N is a $N \times N$ Wigner matrix having all entries i.i.d. from $N(0, 1/N)$ and satisfying the symmetry constraint. Observe that $\frac{1}{N} \operatorname{Tr} \left(U_N^{\sum m_p} \right)$ is the $(\sum m_p)^{th}$ moment of $\operatorname{ESD}(Z_N)$. The limit follows by the classical estimates for Wigner matrices with i.i.d. entries (see for example Anderson et al. [4]).

Sub-Case 2: $\sum_{p=1}^k m_p$ is even

Define

$$\tilde{P}_i = \frac{1}{N^{\sum m_p/2}} \prod_{l=1}^{\sum m_p} G_{i_l \wedge i_{l+1}, i_l \vee i_{l+1}}$$

So we have to show

$$\frac{1}{N} \sum_i \tilde{P}_i \prod_{l=1}^k Z_{i_{1+\sum_{p=1}^l m_p}}^{n_l} \xrightarrow{L_2} 0$$

Now take Π to be a partition of $\{1, \dots, m\}$, then due to a decomposition similar to (5.2.3) enough to show that

$$\frac{1}{N} \sum_{i \in \Psi(\Pi, N)} \tilde{P}_i \prod_{l=1}^k Z_{i_{1+\Sigma_{p=1}^l}^{m_p}}^{n_l} \xrightarrow{L_2} 0$$

Then observe

$$\begin{aligned} \mathbb{E} \left[\left[\frac{1}{N} \sum_{i \in \Psi(\Pi, N)} \tilde{P}_i \prod_{l=1}^k Z_{i_{1+\Sigma_{p=1}^l}^{m_p}}^{n_l} \right]^2 \right] &= \frac{1}{N^2} \sum_{E(\Pi)} \mathbb{E}(\tilde{P}_i \tilde{P}_j) \mathbb{E} \left(\prod_{l=1}^k Z_{i_{1+\Sigma_{p=1}^l}^{m_p}}^{n_l} \right) \mathbb{E} \left(\prod_{l=1}^k Z_{j_{1+\Sigma_{p=1}^l}^{m_p}}^{n_l} \right) \\ &+ \frac{1}{N^2} \sum_{\substack{i \text{ and } j \\ \text{are not distinct}}} \mathbb{E} \left(\tilde{P}_i \tilde{P}_j \prod_{l=1}^k Z_{i_{1+\Sigma_{p=1}^l}^{m_p}}^{n_l} \prod_{l=1}^k Z_{j_{1+\Sigma_{p=1}^l}^{m_p}}^{n_l} \right) \end{aligned}$$

Now since k is odd, then $\sum m_p + n_p = k$ is odd, and hence $\sum n_p$ is odd. Hence it's easy to see that the first term is 0. So we focus on the second term only. Using independence and the uniform bound of

$$\mathbb{E} \left(\prod_{l=1}^k Z_{i_{1+\Sigma_{p=1}^l}^{m_p}}^{n_l} \prod_{l=1}^k Z_{j_{1+\Sigma_{p=1}^l}^{m_p}}^{n_l} \right)$$

we only focus on $\mathbb{E}(\tilde{P}_i \tilde{P}_j)$. Similar to (5.2.33) it is enough to show that

$$\frac{1}{N^2} \sum_{H_k(\Pi)} \mathbb{E}(\tilde{P}_i \tilde{P}_j) \rightarrow 0$$

Consider $b(\Pi)$ to be number of blocks in Π and remember $m = \sum m_p$. Let us look at the case when we have exactly one block common in i and j . Then $|H_1(\Pi)| = O(N^{2b(\Pi)-1})$. It is easy to see that there exists $h > 0$, depending on Π such that $\mathbb{E}(\tilde{P}_i \tilde{P}_j) \leq hN^{-m}$ for all $i, j \in \Psi(\Pi, N)$. This essentially follows as above in previous case. Now consider two cases depending upon the values of $b(\Pi)$.

Case 1: $b(\Pi) \leq \frac{m}{2} + 1$. Then we have $2b(\Pi) - 1 \leq m + 1$, which implies

$$\frac{1}{N^2} \sum_{i, j \in H_1(\Pi)} \mathbb{E}(\tilde{P}_i \tilde{P}_j) \leq O(N^{-m-2} N^{m+1}) = O(N^{-1}) \xrightarrow{N \rightarrow \infty} 0$$

Case 2: $b(\Pi) > \frac{m}{2} + 1$. Observe

$$\mathbb{E}(\tilde{P}_i \tilde{P}_j) = \frac{1}{N^m} \mathbb{E} \left[\prod_{l=1}^m G_{i_l \wedge i_{l+1}, i_l \vee i_{l+1}} \prod_{l=1}^m G_{j_l \wedge j_{l+1}, j_l \vee j_{l+1}} \right]$$

This case is exactly similar to Case 2 as before. Using the same estimates as in (5.2.36) we can show

$$\frac{1}{N^2} \sum_{i, j \in H_1(\Pi)} \mathbb{E}(\tilde{P}_i \tilde{P}_j) \xrightarrow{N \rightarrow \infty} 0.$$

Similarly can conclude the same for any $k \in \{1, 2, \dots, b(\Pi)\}$. Hence we have

$$\frac{1}{N^2} \sum_{\substack{i, j \in \Psi(\Pi, N) \\ \text{are not distinct}}} \mathbb{E}(\tilde{P}_i \tilde{P}_j) \xrightarrow{N \rightarrow \infty} 0$$

Then we have shown that (5.2.40) converges to 0 in L^2 , that is

$$\frac{1}{N} \text{Tr}(\tilde{\Delta}_N^k) \xrightarrow{L_2} 0.$$

5.2.2. *Existence of unique limiting distribution.* Before proceeding let us define

$$\beta_k = \begin{cases} \sum_{\mathcal{P}_k} \sum_{\sigma \in NC_2(\Sigma m_p)} \beta(\sigma) \mathcal{E}(\sigma) & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{cases}$$

then by Lemma B.1 from Bai and Silverstein [5] we can easily see that there exists a probability measure ν identified by the moments β_k . To show uniqueness of ν we invoke again Theorem 1 of Lin [25], that is show that

$$\limsup_{k \rightarrow \infty} \frac{1}{2k} \beta_{2k}^{1/2k} < \infty \quad (5.2.42)$$

For all $l \in \{1, 2, \dots, \sum_{p=1}^k m_p\}$, define

$$\alpha_l = \begin{cases} n_j & \text{if } l = 1 + \sum_{p=1}^j m_p, \forall j = 1, \dots, k-1, \\ n_k & \text{if } l = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Recall that we use $1 + \sum_{p=1}^k m_p$ as an identifier for 1. Then it is easy to observe that

$$\prod_{j=1}^k Z_i^{n_j} = \prod_{u \in K(\sigma)} Z_u^{\sum_{l \in u} \alpha_l}, \quad \forall i \in \Psi(K(\sigma), N)$$

where the product is taken over the blocks in $K(\sigma)$. Hence by definition of $\mathcal{E}(\sigma)$ we have

$$\mathcal{E}(\sigma) = \mathbb{E} \left[\prod_{u \in K(\sigma)} Z_u^{\sum_{l \in u} \alpha_l} \right] = \prod_{u \in K(\sigma)} \mathbb{E} \left[Z_1^{\sum_{l \in u} \alpha_l} \right].$$

The above equality follows since Z_u 's are all i.i.d. $N(0, 1)$ for all $u \in K(\sigma)$. Using the expression for moments of the Gaussian distribution,

$$\begin{aligned} \prod_{u \in K(\sigma)} \mathbb{E} \left[Z_1^{\sum_{l \in u} \alpha_l} \right] &\leq \prod_{u \in K(\sigma)} \frac{(2 \sum_{l \in u} \alpha_l)!}{2^{\sum_{l \in u} \alpha_l} (\sum_{l \in u} \alpha_l)!} \\ &\leq 2^{\sum_{u \in K(\sigma)} \sum_{l \in u} \alpha_l} \left(\sum_{u \in K(\sigma)} \sum_{l \in u} \alpha_l \right)! \\ &\leq 2^k k! \end{aligned}$$

The last inequality follows from $\sum_{j=1}^k n_j \leq k$. Observe that by definition of $\beta(\sigma)$ it can be easily seen that there exists a constant $C > 0$ such that $\beta(\sigma) \leq C^k$. Recall C_l is the l^{th} catalan number and $C_a \leq C_b$ whenever $a \leq b$, then taking $k \in 2\mathbb{N}$

$$\begin{aligned} \beta_k &\leq \sum_{\mathcal{P}_k} \sum_{\sigma \in NC_2(\Sigma m_p)} C^k 2^k k! \leq \sum_{\mathcal{P}_k} C_{\Sigma m_p} (2C)^k k! \\ &\leq \sum_{\mathcal{P}_k} C_k (2C)^k k! \leq (4C)^k C_k k! \end{aligned}$$

The above inequalities follows since $\sum_{p=1}^k m_p \leq k$ and the sum over \mathcal{P}_k can have at most 2^k many terms. Using Sterling's approximation we have

$$\frac{1}{k} \beta_k^{\frac{1}{k}} \leq 4C \frac{1}{(k+1)^{\frac{1}{k}}} \frac{4e^{-(1+\frac{1}{k})}}{\pi^{\frac{1}{k}}}$$

and subsequently we conclude that

$$\limsup_{k \rightarrow \infty} \frac{1}{2k} \beta_{2k}^{\frac{1}{2k}} < \infty. \quad (5.2.43)$$

Hence the measure ν is uniquely identified by the moments β_k . Finally note that the moment generating function of ν is finite around origin and the odd moments vanish it easily follows that ν is symmetric around origin.

5.2.3. ν *has unbounded support*. To prove the unbounded support we shall use the following lemma from Chakrabarty et al. [11].

Lemma 5.5. *Chakrabarty et al. [11, Fact A.5] Suppose that for all $n \geq 1$, $Z_{n1} \geq \dots \geq Z_{nn}$ are random variables such that*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \delta_{Z_{nj}} = \mu \text{ weakly in probability,} \quad (5.2.44)$$

for some probability measure μ on \mathbb{R} , where δ_x is the probability measure that puts mass 1 at x . Then,

$$\lim_{p \rightarrow 0} \limsup_{n \rightarrow \infty} Z_{n[np]} = \sup(\text{supp}(\mu)) \text{ almost surely} \quad (5.2.45)$$

where $[x]$ denotes the smallest integer larger than or equal to x .

Since there exists an open set $U \subseteq [0, 1]^2$ such that $W > 0$ on U , then using Lemma 5.5 and Lemma 5.4 we have

$$\lim_{p \rightarrow 0} \limsup_{N \rightarrow \infty} \lambda_{[Np]}(Y_N) = \sup(\text{supp}(\zeta)) = \infty \text{ almost surely} \quad (5.2.46)$$

Since in the following we would be dealing with almost sure statements, then we refrain from writing almost surely every time. Here $[x]$ denotes the smallest integer larger than or equal to x and $\lambda_k(\Sigma)$ denotes the k^{th} largest eigenvalue of the matrix Σ . By Weyl's inequality we have

$$\lambda_{2[Np]-1}(Y_N) \leq \lambda_{[Np]}(\bar{A}_N + Y_N) + \lambda_{[Np]}(-\bar{A}_N)$$

Now by definition of \bar{A}_N , observe that $-\bar{A}_N$ satisfies the assumptions of Theorem 1 of Zhu [41], and let

$$\text{ESD}(-\bar{A}_N) \rightarrow \kappa_A$$

for some unique measure κ_A on \mathbb{R} . Suppose $\tilde{X} \sim \kappa_A$. Then we have from (2.3.2)

$$\mathbb{E}[\tilde{X}^k] = \begin{cases} 0 & k \text{ is odd} \\ \sum_{j=1}^{C_k/2} t(T_j^{k/2+1}, W) & k \text{ is even} \end{cases}$$

where T_j^{k+1} denotes the j^{th} rooted planar tree of $k+1$ vertices and C_k is the k^{th} Catalan number. Then observe that

$$\|\tilde{X}\|_{2k} = \left(\sum_{j=1}^{C_k} t(T_j^{k+1}, W) \right)^{\frac{1}{2k}} \leq (C_k C^k)^{\frac{1}{2k}}$$

where $\|\cdot\|_{2k}$ denotes the L^{2k} norm and C is from L.1. It follows that there exists a constant $C_1^\circ \in (0, \infty)$ such that

$$\lim_{k \rightarrow \infty} (C_k C^k)^{\frac{1}{2k}} = C^{\frac{1}{2}} C_1^\circ.$$

Hence we can conclude that there exists a constant $C^\circ > 0$ such that

$$\|\tilde{X}\|_{2k} \leq C^\circ, \quad \forall k \geq 1$$

Now since $\|\tilde{X}\|_k \leq \|\tilde{X}\|_{2k}$ we have

$$\|\tilde{X}\|_k \leq C^\circ, \quad \forall k \geq 1$$

Now define $f = \min(|\tilde{X}|, C^\circ + 1) \in L^\infty$. Then $\|f\|_p \leq \|\tilde{X}\|_p \leq C^\circ$. Now since $f \in L^\infty \cap L^p$ for some $p \geq 1$, then $\|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p \leq C^\circ$. Then we have $\min(|\tilde{X}|, C^\circ + 1) \leq C^\circ$ a.e. which implies

$$|\tilde{X}| \leq C^\circ \text{ a.e.}$$

Thus

$$\sup(\text{supp}(\kappa_A)) \leq C^\circ. \quad (5.2.47)$$

Hence using Lemma 5.5 we have

$$\lim_{p \rightarrow 0} \limsup_{N \rightarrow \infty} \lambda_{[Np]}(-\bar{A}_N) = \sup(\text{supp}(\kappa_A)) \leq C^\circ$$

Fix $\varepsilon > 0$, then there exists $\eta > 0$ such that for all $p \in (0, \eta \wedge \frac{1}{2})$, $\limsup_{N \rightarrow \infty} \lambda_{[Np]}(-\bar{A}_N) \leq C^\circ + \varepsilon$. So for all $p \in (0, \eta \wedge \frac{1}{2})$ we have

$$\begin{aligned} \limsup_{N \rightarrow \infty} \lambda_{[Np]}(\bar{A}_N + Y_N) &\geq \limsup_{N \rightarrow \infty} \lambda_{2[Np]-1}(Y_N) - \liminf_{N \rightarrow \infty} \lambda_{[Np]}(-\bar{A}_N) \\ &\geq \limsup_{N \rightarrow \infty} \lambda_{2[Np]-1}(Y_N) - (C^\circ + \varepsilon) \end{aligned}$$

Now from (5.2.46) it can be easily shown that $\lim_{p \rightarrow 0} \limsup_{N \rightarrow \infty} \lambda_{2[Np]-1}(Y_N) = \infty$ and hence we have

$$\sup(\text{supp}(v)) = \lim_{p \rightarrow 0} \limsup_{N \rightarrow \infty} \lambda_{[Np]}(\bar{A}_N + Y_N) = \infty$$

which implies that v has unbounded support.

5.2.4. Identification of moments of v . The proof of the previous result already gave the convergence of the moments. We briefly browse through the expressions to write them in terms of the description used in Section 2.5. First, remember that the odd moments are 0. The case for $k = 0$ is trivial, so fix k in $2\mathbb{N}$. Then it is enough to look at the convergence of the terms (recall (5.2.4))

$$\begin{aligned} \mathbb{E} \left[\frac{1}{N} \text{Tr}(\bar{A}_N^{m_1} Y_N^{n_1} \dots \bar{A}_N^{m_k} Y_N^{n_k}) \right] &= \frac{1}{N} \sum_{i_1 \dots i_{\sum m_p}} \mathbb{E} \left[\prod_{j=1}^{\sum m_p} G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}} \right] \prod_{j=1}^{\sum m_p} \frac{\sigma_{i_j, i_{j+1}}}{\sqrt{N}} \\ &= \prod_{j=1}^k \left(\frac{1}{N} \sum_{\substack{t \neq i \\ 1 + \sum_{p=1}^j m_p, t}} \sigma_{i, 1 + \sum_{p=1}^j m_p, t}^2 \right)^{\frac{n_j}{2}} \mathbb{E} \left[\prod_{j=1}^k Z_{i, 1 + \sum_{p=1}^j m_p}^{n_j} \right]. \end{aligned} \quad (5.2.48)$$

where $\sum_{p=1}^k m_p + n_p = k$. Now if $\sum_{p=1}^k m_p$ is odd then

$$\mathbb{E} \left[\prod_{j=1}^{\sum m_p} G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}} \right] = 0$$

and the corresponding moment becomes 0. So we need only consider $\sum_{p=1}^k m_p$ is even. Once again we look at the closed walk $i_1 \rightarrow \dots \rightarrow i_m \rightarrow i_1$. (Remember that $i_{m+1} = i_1$ and hence we identify 1 by $m+1$ for notational convenience.) Arguing as in the convergence of moments part we can show that it is enough to look at a closed walk on a tree of $m/2 + 1$ vertices where each edge is visited twice. We know that there exists one correspondence between such a walk and a depth first search over a rooted planar tree having vertices chosen from $[N]$. Hence consider the r^{th} labelled rooted planar tree $T_{r,l}^{m/2+1} = (V, E)$ with the labelling $l = (l_1, \dots, l_{m/2+1})$ which corresponds to this walk. Then it is easy to see that

$$\mathbb{E} \left[\prod_{j=1}^{\sum m_p} G_{i_j \wedge i_{j+1}, i_j \vee i_{j+1}} \right] = \mathbb{E} \left[\prod_{e \in E} G_e^2 \right] = 1$$

and using the notations defined in the setup of Section 2.5 we have

$$\prod_{j=1}^k \left(\frac{1}{N} \sum_{t \neq i} \sum_{1+\sum_{p=1}^j m_p} \sigma_{i+\sum_{p=1}^j m_p, t}^2 \right)^{\frac{n_j}{2}} = \prod_{s=1}^{\frac{m}{2}+1} \left(\frac{1}{N} \sum_{t \neq l_s} \sigma_{l_s, t}^2 \right)^{\sum_{j=1}^{n_s} n_{s_j}/2}$$

and

$$\mathbb{E} \left[\prod_{j=1}^k Z_i^{n_j} \right] = \mathbb{E} \left[\prod_{s=1}^{\frac{m}{2}+1} Z_{l_s}^{\sum_{j=1}^{n_s} n_{s_j}} \right] = \prod_{s=1}^{\frac{m}{2}+1} \mathbb{E} \left[Z_{l_s}^{\sum_{j=1}^{n_s} n_{s_j}} \right] = f \left(\tilde{T}_{r,l}^{m/2+1} \right)$$

where f is defined in (2.5.1). Hence the contribution of this term is 0 if $\sum_{j=1}^{n_s} n_{s_j}$ is odd for some $s \in \{1, 2, \dots, m/2 + 1\}$. So we consider the situation where $\sum_{j=1}^{n_s} n_{s_j}$ is even for all $s \in \{1, 2, \dots, m/2 + 1\}$. Then (5.2.48) becomes

$$\frac{1}{N^{m/2+1}} \sum_{r=1}^{C_{m/2}} \sum_{l_1 \neq l_2 \neq \dots \neq l_{m/2+1}} \prod_{e \in E(\tilde{T}_{r,l}^{m/2+1})} \sigma_e^2 \prod_{s=1}^{\frac{m}{2}+1} \left(\frac{1}{N} \sum_{t \neq l_s} \sigma_{l_s, t}^2 \right)^{\sum_{j=1}^{n_s} n_{s_j}/2} f \left(\tilde{T}_{r,l}^{m/2+1} \right) \quad (5.2.49)$$

Observing that f does not depend on the labelling of the tree, and going similarly as in (5.2.18) we can show that

$$\begin{aligned} \frac{1}{N^{m/2+1}} \sum_{l_1 \neq l_2 \neq \dots \neq l_{m/2+1}} \prod_{e \in E(\tilde{T}_{r,l}^{m/2+1})} \sigma_e^2 \prod_{s=1}^{\frac{m}{2}+1} \left(\frac{1}{N} \sum_{t \neq l_s} \sigma_{l_s, t}^2 \right)^{\sum_{j=1}^{n_s} n_{s_j}/2} f \left(\tilde{T}_{r,l}^{m/2+1} \right) \\ \rightarrow t(\tilde{T}_r^{m/2+1}, W) f(\tilde{T}_r^{m/2+1}) \end{aligned}$$

Then combining with (5.2.49) and (5.2.1) we are done. \square

Proof of Corollary 2.1. It is easy to observe that $\lambda_i(A + \alpha I_N) = \lambda_i(A) + \alpha$ and $\lambda_i \left(\frac{A}{\alpha} \right) = \frac{1}{\alpha} \lambda_i(A)$ for any $N \times N$ matrix A and for any $\alpha \neq 0$. Then observe that

$$\frac{\lambda_i(\Delta_N) - N\mu_N}{\sqrt{N}\sigma_N} = \lambda_i \left(\frac{\Delta_N - N\mu_N I_N}{\sqrt{N}\sigma_N} \right)$$

Now observe that the centered laplacian is

$$\Delta_N^0 = \frac{\Delta_N - N\mu_N I_N + \mu_N J_N}{\sqrt{N}\sigma_N}$$

where J_N is the $N \times N$ matrix having all entries equal to 1. Define $\Delta_N^1 = \frac{\Delta_N - N\mu_N I_N}{\sqrt{N}\sigma_N}$. Now by rank inequality (Bai and Silverstein [5, Theorem A.43]) we have

$$\left\| F^{\Delta_N^0} - F^{\Delta_N^1} \right\| \leq \frac{1}{N} \text{rank} \left(\Delta_N^0 - \Delta_N^1 \right) = \frac{1}{N} \text{rank} \left(\frac{\mu_N}{\sqrt{N}\sigma_N} J_N \right) = O \left(\frac{1}{N} \right) \rightarrow 0 \quad (5.2.50)$$

where $F^{A_N}(x) = \frac{1}{N} \sum_{i=1}^N I \{ \lambda_i(A_N) \leq x \}$ for any $N \times N$ symmetric matrix A_N and its eigenvalues $\{ \lambda_i(A_N) : 1 \leq i \leq N \}$. Hence it is enough to look at convergence of ESD of Δ_N^0 .

Now define $B_N = \frac{1}{\sigma_N} (A_N - \mathbb{E}A_N) = \frac{1}{\sigma_N} (A_N - \mu_N J_N)$. Then it is easy to see that $\sqrt{N}\Delta_N^0$ is the laplacian corresponding to B_N . Now observe that for any $\eta > 0$

$$\frac{1}{N^2} \sum_{1 \leq i, j \leq N} \mathbb{E} \left[|B_N(i, j)|^2 \mathbf{1} \left\{ |B_N(i, j)| \geq \eta \sqrt{N} \right\} \right] \leq \frac{1}{N^2} \sum_{1 \leq i, j \leq N} \frac{1}{\eta^\delta N^{\frac{\delta}{2}}} \mathbb{E} \left[\left| \frac{A_N(i, j) - \mu_n}{\sigma_N} \right|^{2+\delta} \right] \quad (5.2.51)$$

$$= O \left(N^{-\frac{\delta}{2}} \right) \rightarrow 0 \quad (5.2.52)$$

Thus B_N satisfies assumption **L.2**. Assumption **L.1** is immediate from the given assumptions in the theorem. Observe that

$$\mathbb{E} \left[B_N(i, j)^2 \right] = 1, \quad \forall 1 \leq i, j \leq N \quad (5.2.53)$$

Taking $W \equiv 1$, a graphon, we can easily see that assumption **L.3** is satisfied. Hence by Theorem 2.1 there exists a symmetric probability measure ν of unbounded support where ESD of Δ_N^0 converges to weakly in probability.

Recall from the proof of Theorem 2.1, the moments which identify ν are given by the limits of $\frac{1}{N} \text{Tr}(\bar{A}_N + Y_N)^k$ for all $k \geq 1$ where \bar{A}_N is as defined in the Lemma 5.2 with $\sigma_{i,j} = 1 \forall i, j$ and Y_N is the diagonal $N \times N$ matrix as defined in Lemma 5.3 with $\sigma_{i,j} = 1$. It is easy to observe that $\frac{1}{N} \text{Tr}(\bar{A}_N + Y_N)^k$ is the k^{th} moment of $\text{ESD}(\bar{A}_N + Y_N)$. Since ν is uniquely identified by it's moments, then using method of moments, we can say that $\text{ESD}(\bar{A}_N + Y_N) \implies \nu$ in probability. By the strong law of large numbers, with probability 1, $\text{ESD}(Y_N) \implies \gamma_1$, where γ_1 denotes the standard normal distribution. Also we know that, with probability 1, $\text{ESD}(\bar{A}_N) \implies \gamma_0$, where γ_0 denotes the semicircle law. Further

$$\sup_N \mathbb{E} \int |x| d\text{ESD}(Y_N) = \sup_N \mathbb{E} \frac{1}{N} \sum_{i=1}^N |Z_i| < \infty$$

where Z_i are i.i.d. standard normal. Also $\mathbb{E} \int |x| d\text{ESD}(\bar{A}_N) \leq \frac{1}{N} \sqrt{\mathbb{E} \text{Tr}(\bar{A}_N^2)} = 1$. Then by Theorem 2.1 of Pastur and Vasilchuk [31] we have $\text{ESD}(\bar{A}_N + Y_N)$ converges weakly in probability to $\gamma_0 \boxplus \gamma_1 = \gamma_M$. Hence by uniqueness of ν we must have $\nu = \gamma_M$. \square

Proof of Theorem 2.2. The case when $r \equiv 0$ is trivially true. Hence we assume that $r > 0$ at some point in $[0,1]$. By the proof of Theorem 2.1

$$\text{ESD}(\bar{A}_N + Y_N) \rightarrow \nu, \text{ weakly in probability}$$

where \bar{A}_N and Y_N are as defined in Lemma 5.3. Define

$$\eta^W \left(\frac{i}{N}, \frac{j}{N} \right) = N^2 \int_{I_i \times I_j} W(x, y) dx dy, \quad \forall 1 \leq i, j \leq N, \quad (5.2.54)$$

where $\{I_i\}$ is the partition of $[0,1]$ as defined in Definition 2.1. Then define the $N \times N$ matrix \bar{Z}_N as

$$\bar{Z}_N(i, j) = \sqrt{\frac{\eta^W \left(\frac{i}{N}, \frac{j}{N} \right)}{N}} G_{i \wedge j, i \vee j}, \quad 1 \leq i, j \leq N$$

where the collection $\{G_{i,j} : 1 \leq i, j \leq N\}$ is defined in (5.1.1). Further define the $N \times N$ diagonal matrix Y_N^Z by

$$Y_N^Z(i, i) = \sqrt{\frac{1}{N} \sum_{j \neq i} \eta^W \left(\frac{i}{N}, \frac{j}{N} \right)} Z_i, \quad 1 \leq i \leq N$$

where Z_i for all $1 \leq i \leq N$ are as defined in Lemma 5.3. Observe that

$$\begin{aligned} \frac{1}{N} \mathbb{E} \text{Tr}(\bar{A}_N - \bar{Z}_N)^2 &= \frac{1}{N^2} \sum_{i,j=1}^N \left(\sigma_{i,j} - \sqrt{\eta^W \left(\frac{i}{N}, \frac{j}{N} \right)} \right)^2 \\ &= \frac{1}{N^2} \sum_{i,j=1}^N \eta^W \left(\frac{i}{N}, \frac{j}{N} \right) + \frac{1}{N^2} \sum_{i,j=1}^N \sigma_{i,j}^2 - \frac{2}{N^2} \sum_{i,j=1}^N \sigma_{i,j} \sqrt{\eta^W \left(\frac{i}{N}, \frac{j}{N} \right)} \end{aligned} \quad (5.2.55)$$

Using (5.2.54) and the fact that $|\sqrt{x} - \sqrt{y}| \leq |x - y|/\sqrt{y}$ for $x, y \in (0, \infty)$, we have

$$\begin{aligned} & \left| \frac{1}{N^2} \sum_{i,j=1}^N \sigma_{i,j} \sqrt{\eta^W \left(\frac{i}{N}, \frac{j}{N} \right)} - \int_{[0,1]^2} W(x,y) dx dy \right| \\ &= \left| \sum_{i,j=1}^N \left(\int_{I_i \times I_j} W(x,y) dx dy \right)^{1/2} \left[\left(\int_{I_i \times I_j} W_N(x,y) dx dy \right)^{1/2} - \left(\int_{I_i \times I_j} W(x,y) dx dy \right)^{1/2} \right] \right| \\ &\leq \sum_{i,j=1}^N \int_{I_i \times I_j} |W_N(x,y) - W(x,y)| dx dy \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

From (5.2.55) we conclude that

$$\frac{1}{N} \mathbb{E} \operatorname{Tr} (\bar{A}_N - \bar{Z}_N)^2 \rightarrow 0. \quad (5.2.56)$$

A similar computation as above shows that

$$\frac{1}{N} \mathbb{E} \operatorname{Tr} (Y_N - Y_N^Z)^2 \rightarrow 0. \quad (5.2.57)$$

Combining (5.2.56) and (5.2.57), along with Bai and Silverstein [5, Corollary A.41] we have

$$L \left(\operatorname{ESD} (\bar{A}_N + Y_N), \operatorname{ESD} (\bar{Z}_N + Y_N^Z) \right) \xrightarrow{P} 0.$$

Define

$$g_i = \int_{I_i} r(x) dx, \quad \forall 1 \leq i \leq N \quad (5.2.58)$$

and define the $N \times N$ diagonal matrix \tilde{Y}_N^Z as

$$\tilde{Y}_N^Z(i,i) = \alpha \sqrt{N g_i} Z_i, \quad \forall 1 \leq i \leq N.$$

By definition in (5.2.58)

$$\frac{1}{N} \sum_{j \neq i} \eta^W \left(\frac{i}{N}, \frac{j}{N} \right) = N g_i (\alpha^2 - g_i), \quad \forall 1 \leq i \leq N.$$

Thus

$$\begin{aligned} \frac{1}{N} \mathbb{E} \operatorname{Tr} \left((\bar{Z}_N + Y_N^Z) - (\bar{Z}_N + \tilde{Y}_N^Z) \right)^2 &= \frac{1}{N} \mathbb{E} \operatorname{Tr} (Y_N^Z - \tilde{Y}_N^Z)^2 \\ &= \frac{1}{N} \sum_{i=1}^N \left(\sqrt{N g_i (\alpha^2 - g_i)} - \alpha \sqrt{N g_i} \right)^2 \mathbb{E} Z_i^2 \\ &= \sum_{i=1}^N g_i \left(\alpha - \sqrt{\alpha^2 - g_i} \right)^2 \\ &\leq \sum_{i=1}^n g_i \frac{1}{\alpha} \left(\alpha^2 - (\alpha^2 - g_i)^2 \right) \\ &\leq \sum_{i=1}^N \frac{1}{\alpha} g_i^3 \rightarrow 0. \end{aligned} \quad (5.2.59)$$

where (5.2.59) follows from the inequality $|\sqrt{x} - \sqrt{y}| \leq \frac{|x-y|}{\sqrt{y}}$ and the final limit follows since $g_i \leq 1/N$ for all $1 \leq i \leq N$. Now for $N \geq 1$, define the $N \times N$ matrices

$$G_N(i,j) = N^{-1/2} G_{i \wedge j, i \vee j}, \quad 1 \leq i, j \leq N,$$

$$R_N = \operatorname{Diag} \left(\sqrt{N g_1}, \sqrt{N g_2}, \dots, \sqrt{N g_N} \right)$$

and

$$U_N = \text{Diag}(Z_1, \dots, Z_N)$$

Then observe that

$$\tilde{Z}_N = R_N G_N R_N,$$

and

$$\tilde{Y}_N^Z = \alpha R_N^{1/2} U_N R_N^{1/2}$$

The proof would follow similarly as the proof of Theorem 1.3 from Chakrabarty et al. [11] if we can show that for any $K \geq 0$ and all $k \geq 1, m_1, \dots, m_k$ and $n_1, \dots, n_k \geq 0$,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr}(R_N^{m_1} U_{NK}^{n_1} \dots R_N^{m_k} U_{NK}^{n_k}) \quad (5.2.60)$$

$$= \int_0^1 r^{\sum_{i=1}^k m_i/4}(u) du \int_{-K}^K \frac{1}{\sqrt{2\pi}} x^{\sum_{i=1}^k n_i} e^{-x^2/2} dx \text{ a.s.} \quad (5.2.61)$$

where for all $K \geq 0$, the $N \times N$ diagonal matrix U_{NK} is given by

$$U_{NK} = \text{Diag}\left(Z_1 \mathbf{1}_{\{|Z_1| \leq K\}}, \dots, Z_N \mathbf{1}_{\{|Z_N| \leq K\}}\right)$$

An application of SLLN shows that it is enough to look at the limit of

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N (Ng_i)^{\sum_{j=1}^k m_j/4} \mathbb{E}\left(Z_i^{\sum_{j=1}^k n_j} \mathbf{1}_{\{|Z_i| \leq K\}}\right) \\ = \frac{1}{N} \sum_{i=1}^N (Ng_i)^{\sum_{j=1}^k m_j/4} \int_{-K}^K \frac{1}{\sqrt{2\pi}} x^{\sum_{i=1}^k n_i} e^{-x^2/2} dx \end{aligned}$$

Define

$$m(x) = \int_0^x r(t) dt, \quad \forall x \in [0, 1]$$

Then observe that

$$\frac{1}{N} \sum_{i=1}^N \left(N \int_{I_i} r(x) dx\right)^{\sum_{j=1}^k m_j/4} = \frac{1}{N} \sum_{i=1}^N \left(\frac{m(i/N) - m(i-1/N)}{1/N}\right)^{\sum_{j=1}^k m_j/4}$$

It is easy to see that m is uniformly differentiable on $[0, 1]$.³ Hence given $\varepsilon > 0$, for large enough N we have

$$\left(\frac{m(i/N) - m(i-1/N)}{1/N}\right)^{\sum_{j=1}^k m_j/4} = r\left(\frac{i}{N}\right)^{\sum_{j=1}^k m_j/4} + O(\varepsilon)$$

Hence taking $N \rightarrow \infty$ we have

$$\frac{1}{N} \sum_{i=1}^N \left(\frac{m(i/N) - m(i-1/N)}{1/N}\right)^{\sum_{j=1}^k m_j/4} \rightarrow \int_0^1 r^{\sum_{i=1}^k m_i/4}(u) du + O(\varepsilon)$$

Since the above is true for all $\varepsilon > 0$, then we have shown (5.2.60). \square

³Let f be defined (and real valued) on $[a, b]$ and the derivative f' exists on $[a, b]$ (considering the left and right derivatives at the boundary points). Then the function f is said to be uniformly differentiable if for all $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $0 \leq |t - x| \leq \delta, a \leq x, t \leq b$

$$\left|\frac{f(x) - f(t)}{x - t} - f'(x)\right| < \varepsilon$$

It follows from Rudin et al. [32, Exercise 5.8] that a function f differentiable on $[a, b]$ is uniformly differentiable if f' is continuous on $[a, b]$.

6. PROOF OF THEOREM 2.3 AND COROLLARY 2.2

Before proving that spectral norm of Δ_N scales as $\sqrt{N \log N}$, we show that the spectral norm of the adjacency matrix would scale slower than $\sqrt{N \log N}$, which is crucially needed for proving the theorem.

Lemma 6.1. *For a $N \times N$ generalised Wigner matrix A_N satisfying the above stated assumptions S.1-S.3,*

$$\frac{\|A_N\|}{\sqrt{N \log N}} \xrightarrow{a.s.} 0$$

where $\|\cdot\|$ is the spectral norm.

Proof. Following Theorem 3.2 from Zhu [41] we know that

$$\lim_{N \rightarrow \infty} \text{ESD}(A_N) = \mu \text{ weakly almost surely} \quad (6.1)$$

where μ is identified by the moments given in (2.3.2) Then going similar as in proof of (5.2.47) we can conclude that there exists $C_0 > 0$ such that

$$\sup(\text{supp}(\mu)) \leq C^0$$

Then by Ding and Jiang [18, Lemma 2.8] there exists $\alpha \in \mathbb{R}$ such that $|\alpha| < \infty$ and

$$\liminf_{N \rightarrow \infty} \frac{\lambda_{\max}(A_N)}{\sqrt{N}} \geq \alpha \text{ almost surely.} \quad (6.2)$$

(Observe that due to above bound on $\sup(\text{supp}(\mu))$ we can conclude that $|\alpha| < \infty$). Thus it is enough to prove the upper bound

$$\limsup_{N \rightarrow \infty} \frac{\lambda_{\max}(A_N)}{\sqrt{N}} \leq 2C_2^{1/2} \text{ almost surely}$$

We omit the proof of the upper bound since it follows similar to the proof of Lemma 2.1 from Ding and Jiang [18]. \square

Proof of Theorem 2.3. The proof follows the line of argument in Bryc et al. [9, Theorem 1.5]. Before proceeding with the proof in order to make notations clearer we will use $A_N = (X_{ij}^{(N)})_{1 \leq i, j \leq N}$ to indicate the dependence on N and similarly $\sigma_N(i, j)$ instead of $\sigma_{i, j}$ in corresponding places. Define

$$D_N = \text{diag} \left(\sum_{j=1}^N X_{ij}^{(N)} \right)_{i=1}^N$$

Then using triangle inequality we have

$$\| \|\Delta_N\| - \|D_N\| \| \leq \|A_N\| \quad (6.3)$$

Hence by Lemma 6.1 it is enough to look at $\frac{\|D_N\|}{\sqrt{N \log N}}$. Define

$$T_N = \max_{1 \leq i \leq N} \left| \sum_{j=1}^N X_{ij}^{(N)} \right| \text{ which implies } \frac{T_N}{\sqrt{N \log N}} = \frac{\|D_N\|}{\sqrt{N \log N}}$$

Fix $1 \leq i \leq N$, then using Lemma 2.1 of Bryc et al. [9] (a result on strong Gaussian approximation) there exists $\{Y_{ij}^{(N)} : 1 \leq j \leq N\}$ where $Y_{ij}^{(N)} \sim N(0, \sigma_N(i, j)^2)$ and are independent such that $\forall \alpha > 0$,

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{j=1}^N X_{ij}^{(N)} - \sum_{j=1}^N Y_{ij}^{(N)} \right| \geq \alpha \sqrt{N \log N} \right) &\leq \frac{C}{1 + (\alpha \sqrt{N \log N})^6} \sum_{j=1}^N \mathbb{E} \left| X_{ij}^{(N)} \right|^6 \\ &\leq \frac{C_0}{N^2 (\log N)^3} \end{aligned} \quad (6.4)$$

The last inequality follows from assumption [S.2](#). Then it can be concluded that $\forall \alpha > 0$,

$$\max_{i=1}^N \mathbb{P} \left(\left| \sum_{j=1}^N X_{ij}^{(N)} - \sum_{j=1}^N Y_{ij}^{(N)} \right| \geq \alpha \sqrt{N \log N} \right) \leq \frac{C_0}{N^2 (\log N)^3} \quad (6.5)$$

Now observe that

$$\left| \sum_{j=1}^N X_{ij}^{(N)} \right| \leq \left| \sum_{j=1}^N Y_{ij}^{(N)} \right| + \left| \sum_{j=1}^N X_{ij}^{(N)} - \sum_{j=1}^N Y_{ij}^{(N)} \right| \quad (6.6)$$

Hence we have

$$\begin{aligned} \mathbb{P} \left(T_N \geq (\alpha + 2\epsilon) \sqrt{A_2 N \log N} \right) &\leq N \max_{i=1}^N \mathbb{P} \left(\left| \sum_{j=1}^N X_{ij}^{(N)} \right| \geq (\alpha + 2\epsilon) \sqrt{A_2 N \log N} \right) \\ &\leq N \max_{i=1}^N \mathbb{P} \left(\left| \sum_{j=1}^N Y_{ij}^{(N)} \right| \geq (\alpha + \epsilon) \sqrt{A_2 N \log N} \right) \\ &\quad + N \max_{i=1}^N \mathbb{P} \left(\left| \sum_{j=1}^N X_{ij}^{(N)} - \sum_{j=1}^N Y_{ij}^{(N)} \right| \geq \epsilon \sqrt{A_2 N \log N} \right) \end{aligned} \quad (6.7)$$

We know that $\sum_{j=1}^N Y_{ij}^{(N)} \sim \mathcal{N} \left(0, \sum_{j=1}^N \sigma_N(i, j)^2 \right)$. Hence $\sum_{j=1}^N Y_{ij}^{(N)} \stackrel{d}{=} \left(\sum_{j=1}^N \sigma_N(i, j)^2 \right)^{1/2} Z$, where $Z \sim \mathcal{N}(0, 1)$. Using Gaussian tail inequality (Mill's ratio) we have

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{j=1}^N Y_{ij}^{(N)} \right| \geq (\alpha + \epsilon) \sqrt{A_2 N \log N} \right) &= \mathbb{P} \left(|Z| \geq \frac{(\alpha + \epsilon) \sqrt{A_2 N \log N}}{\left(\sum_{j=1}^N \sigma_N(i, j)^2 \right)^{1/2}} \right) \\ &\leq \mathbb{P} \left(|Z| \geq \frac{(\alpha + \epsilon) \sqrt{A_2 N \log N}}{(A_2 N)^{1/2}} \right) \\ &= \mathbb{P} \left(|Z| \geq (\alpha + \epsilon) \sqrt{\log N} \right) \\ &\leq \frac{2}{\sqrt{2\pi} (\alpha + \epsilon) \sqrt{\log N}} \exp \left(-\frac{(\alpha + \epsilon)^2 \log N}{2} \right) \\ &\leq C_1 N^{-\frac{(\alpha + \epsilon)^2}{2}} \end{aligned} \quad (6.8)$$

for some constant $C_1 > 0$ and all N sufficiently large. Then by [\(6.7\)](#) we have

$$\mathbb{P} \left(T_N \geq (\alpha + 2\epsilon) \sqrt{A_2 N \log N} \right) \leq C_1 N^{1 - \frac{(\alpha + \epsilon)^2}{2}} + \frac{C_0}{N (\log N)^3} \quad (6.10)$$

Now taking $\alpha = 2$ we find that R.H.S. of above equation is $O(N^{-1} (\log N)^{-3})$. Then

$$\sum_{N \geq 1} \mathbb{P} \left(T_N \geq (\alpha + 2\epsilon) \sqrt{A_2 N \log N} \right) < \infty \quad (6.11)$$

Hence by Borel Cantelli Lemma we have for all $\epsilon > 0$

$$\limsup_{N \rightarrow \infty} \frac{T_N}{\sqrt{A_2 N \log N}} \leq 2 + 2\epsilon \quad a.s.$$

Hence

$$\limsup_{N \rightarrow \infty} \frac{T_N}{\sqrt{2N \log N}} \leq (2A_2)^{1/2} \quad a.s. \quad (6.12)$$

Now define $k_N = \lceil N / \log N \rceil$. Further define

$$V_N = \max_{i=1}^{k_N} \left| \sum_{j=1}^{k_N} X_{ij}^{(N)} \right| \quad (6.13)$$

Then observe that

$$T_N \geq \max_{i=1}^{k_N} \left| \sum_{j=k_N+1}^N X_{ij}^{(N)} \right| - V_N \quad (6.14)$$

Now observe that by similar computations as (6.12) we have $\limsup_{N \rightarrow \infty} V_N / \sqrt{2k_N \log k_N} \leq C$ almost surely for some $C > 0$. Since $\sqrt{N \log N} / \sqrt{k_N \log k_N} \rightarrow \infty$. Hence $\lim_{N \rightarrow \infty} V_N / \sqrt{N \log N} = 0$ almost surely. Then enough to look at

$$\liminf_{N \rightarrow \infty} \frac{1}{\sqrt{2N \log N}} \max_{i=1}^{k_N} \left| \sum_{j=k_N+1}^N X_{ij}^{(N)} \right| \quad (6.15)$$

Now then once again for using a Gaussian approximation lemma (Lemma 2.1 of Bryc et al. [9]), for large enough N we have

$$\max_{i=1}^{k_N} \mathbb{P} \left(\left| \sum_{j=1}^N X_{ij}^{(N)} - \sum_{j=1}^N Y_{ij}^{(N)} \right| \geq \alpha \sqrt{N \log N} \right) \leq \frac{C_0}{N^2 (\log N)^3} \quad (6.16)$$

for $\{Y_{ij}^{(N)} : 1 \leq j \leq k_N\}$ such that the random variables are independent for fixed $1 \leq i \leq k_N$ and $Y_{ij}^{(N)} \sim N(0, \sigma_N(i, j)^2)$. Now

$$\mathbb{P} \left(\max_{i=1}^{k_N} \left| \sum_{j=k_N+1}^N X_{ij}^{(N)} \right| \leq (\beta - \epsilon) \sqrt{A_1 N \log N} \right) \leq \prod_{i=1}^{k_N} \mathbb{P} \left(\left| \sum_{j=k_N+1}^N X_{ij}^{(N)} \right| \leq (\beta - \epsilon) \sqrt{A_1 N \log N} \right) \quad (6.17)$$

Observe the above uses the independence because of $i < j$. Then for fixed $1 \leq i \leq k_N$ we have

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{j=k_N+1}^N X_{ij}^{(N)} \right| \leq (\beta - 2\epsilon) \sqrt{A_1 N \log N} \right) &\leq \mathbb{P} \left(\left| \sum_{j=k_N+1}^N Y_{ij}^{(N)} \right| \leq (\beta - \epsilon) \sqrt{A_1 N \log N} \right) \\ &\quad + \mathbb{P} \left(\left| \sum_{j=k_N+1}^N X_{ij}^{(N)} - \sum_{j=k_N+1}^N Y_{ij}^{(N)} \right| \geq \epsilon \sqrt{A_1 N \log N} \right) \end{aligned} \quad (6.18)$$

Now once again observe that $\sum_{j=k_N+1}^N Y_{ij}^{(N)} \sim N(0, \sum_{j=k_N+1}^N \sigma_N(i, j)^2)$. Then

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{j=k_N+1}^N Y_{ij}^{(N)} \right| \leq (\beta - \epsilon) \sqrt{A_1 N \log N} \right) &= \mathbb{P} \left(|Z| \leq \frac{(\beta - \epsilon) \sqrt{A_1 N \log N}}{\sqrt{\sum_{j=k_N+1}^N \sigma_N(i, j)^2}} \right) \\ &\geq 1 - \mathbb{P} \left(|Z| > \frac{(\beta - \epsilon) \sqrt{A_1 N \log N}}{\sqrt{(N - k_N) A_1}} \right) \end{aligned} \quad (6.19)$$

Now observe that taking $0 < \epsilon < \beta$, we must have for N sufficiently large $(\beta - \epsilon) \sqrt{\frac{N}{N - k_N}} \leq (\beta - (\epsilon/2))$. Thus by using Mill's ratio for sufficiently large N we have

$$\begin{aligned} \mathbb{P} \left(|Z| > \frac{(\beta - \epsilon) \sqrt{A_1 N \log N}}{\sqrt{(N - k_N) A_1}} \right) &\geq \mathbb{P} \left(|Z| > (\beta - \epsilon) \sqrt{\frac{N}{N - k_N}} \sqrt{\log N} \right) \\ &\geq \frac{C}{N^{(\beta - \epsilon/2)^2/2} \log N} \end{aligned} \quad (6.20)$$

for some constant $C > 0$. Then combining the above equation with (6.16) and (6.18) we have

$$\begin{aligned} \mathbb{P} \left(\left| \sum_{j=k_N+1}^N X_{ij}^{(N)} \right| \leq (\beta - 2\epsilon) \sqrt{A_1 N \log N} \right) &\leq 1 - \frac{C}{N^{(\beta-\epsilon/2)^2/2} \log N} + \frac{C_0}{N^2 (\log N)^3} \\ &\leq 1 - \frac{C_2}{N^{(\beta-\epsilon/2)^2/2} \log N} \end{aligned} \quad (6.21)$$

for all sufficiently large enough N and some constant $C_2 > 0$. Then using the inequality $1 - x \leq e^{-x}$ for any $x > 0$, we have

$$\begin{aligned} \mathbb{P} \left(\max_{i=1}^{k_N} \left| \sum_{j=k_N+1}^N X_{ij}^{(N)} \right| \leq (\beta - \epsilon) \sqrt{A_1 N \log N} \right) &\leq \exp \left(-\frac{C_2}{N^{(\beta-\epsilon/2)^2/2} \log N} k_N \right) \\ &= O \left(\exp \left(-C_2 N^{1-(\beta-\epsilon/2)^2/2} \right) \right) \end{aligned} \quad (6.22)$$

Taking $\beta = \sqrt{2}$, it can be seen that

$$\sum_{N \geq 1} \mathbb{P} \left(\max_{i=1}^{k_N} \left| \sum_{j=k_N+1}^N X_{ij}^{(N)} \right| \leq (\sqrt{2} - \epsilon) \sqrt{A_1 N \log N} \right) < \infty \quad (6.23)$$

Then by Borel-Cantelli Lemma we can conclude that

$$\liminf_{N \rightarrow \infty} \frac{1}{\sqrt{N \log N}} \max_{i=1}^{k_N} \left| \sum_{j=k_N+1}^N X_{ij}^{(N)} \right| \geq (\sqrt{2} - \epsilon) \sqrt{A_1}, \text{ a.s. } 0 < \epsilon < \sqrt{2}. \quad (6.24)$$

Hence combining with (6.14) we must have

$$\liminf_{N \rightarrow \infty} \frac{T_N}{\sqrt{2N \log N}} \geq \sqrt{A_1}. \quad (6.25)$$

Then using (6.3) we have (2.6.3). \square

Proof of Corollary 2.2. We use the notation of Theorem 2.3. Define a $N \times N$ diagonal matrix E_N as

$$E_N = \mathbb{E} X_{ij}^{(N)} - \sum_{j=1}^N \mathbb{E} X_{ij}^{(N)} \mathbf{1}_{i=j}$$

Using Theorem 2.3 it is immediate that

$$\lim_{N \rightarrow \infty} \|\tilde{\Delta}_N\|/N \rightarrow 0 \text{ almost surely,}$$

where $\tilde{\Delta}_N = \Delta_N - E_N$. Observe that by assumption 5.5 we have $\frac{\|E_N\|}{N} \rightarrow m$. Hence using triangle inequality we can conclude that

$$\frac{\|\Delta_N\|}{N} \rightarrow m \text{ almost surely.} \quad \square$$

APPENDIX

In the appendix, we provide the proofs of some of the lemma used before. The methods are straightforward and hence they are recalled in the appendix. First, we provide proof of Lemma 5.1. We shall use some notations and results from Chatterjee [12].

Definition 6.1. For any open interval I containing 0, any positive integer n , any function $f : I^n \rightarrow \mathbb{C}$ which is thrice differentiable in each coordinate, and $1 \leq r \leq 3$, let

$$\lambda_r(f) = \sup \left\{ \left| \partial_i^p f(x) \right|^{\frac{r}{p}} : 1 \leq i \leq n, 1 \leq p \leq r, x \in I^n \right\} \quad (6.26)$$

where ∂_i^p denotes p -fold differentiation with respect to the i^{th} co-ordinate.

Lemma 6.2. Chatterjee [12, Theorem 1.1] Let $f : I^n \rightarrow \mathbb{R}$ be thrice differentiable in each argument. If we set $U = f(\mathbf{X})$ and $V = f(\mathbf{Y})$, then for any thrice differentiable $g : \mathbb{R} \rightarrow \mathbb{R}$ and any $K > 0$,

$$\begin{aligned} |\mathbb{E}g(U) - \mathbb{E}g(V)| &\leq C_1(g)\lambda_2(f) \sum_{i=1}^n \left[\mathbb{E} \left(X_i^2 \mathbf{1}_{\{|X_i| > K\}} \right) + \mathbb{E} \left(Y_i^2 \mathbf{1}_{\{|Y_i| > K\}} \right) \right] \\ &\quad + C_2(g)\lambda_3(f) \sum_{i=1}^n \left[\mathbb{E} \left(X_i^3 \mathbf{1}_{\{|X_i| \leq K\}} \right) + \mathbb{E} \left(Y_i^3 \mathbf{1}_{\{|Y_i| \leq K\}} \right) \right] \end{aligned}$$

where $C_1(g) = \|g'\|_\infty + \|g''\|_\infty$ and $C_2(g) = \frac{1}{6}\|g'\|_\infty + \frac{1}{2}\|g''\|_\infty + \frac{1}{6}\|g'''\|_\infty$

Proof of Lemma 5.1. For the proof we shall use a result from Chatterjee [12] which is recalled later as Lemma 6.2. Consider f, r, n as defined in Definition 6.1. Then observe that $|\partial_i^p \mathcal{R}f| = |\mathcal{R}\partial_i^p f| \leq |\partial_i^p f|$.

Now using the fact that $\frac{r}{p} > 0$ we have $|\partial_i^p \mathcal{R}f|^{\frac{r}{p}} \leq |\partial_i^p f|^{\frac{r}{p}}$, hence we have

$$\lambda_r(\mathcal{R}f) \leq \lambda_r(f). \quad (6.27)$$

Let define $\tilde{X} = (A_N^0(i, j))_{1 \leq i < j \leq N}$ and $\tilde{Y} = (A_N^g(i, j))_{1 \leq i < j \leq N}$. Take $n = \frac{N(N-1)}{2}$. Then for all $x = (x_{i,j})_{1 \leq i < j \leq N} \in \mathbb{R}^n$ define a real symmetric $N \times N$ matrix $\Delta(x)$ as

$$\Delta(x)(i, j) = \begin{cases} x_{i \wedge j, i \vee j} & \text{if } i \neq j \\ -\sum_{k \neq i, k=1}^N x_{i \wedge k, i \vee k} & \text{if } i = j \end{cases} \quad (6.28)$$

Define $\Phi(x) = H_N(\Delta(x))$. Observe that $\Delta(\tilde{X}) = \Delta_N^0$ and $\Delta(\tilde{Y}) = \Delta_N^g$. Since $\mathbb{E}(A_N^0(i, j)) = \mathbb{E}(A_N^g(i, j)) = 0$ and $\mathbb{E}((A_N^0(i, j))^2) = \mathbb{E}((A_N^g(i, j))^2) = \sigma_{i,j}^2/N$. Hence the assumptions on \tilde{X} and \tilde{Y} for Lemma 6.2 are satisfied. Note that real part of Φ is thrice differentiable as Φ is thrice differentiable. Observe that $\frac{\partial \Delta(x)}{\partial x_{i,j}}$ is a $N \times N$ matrix having -1 at i^{th} and j^{th} diagonals and 1 at (i, j) and (j, i) positions. Using matrix identities derived in Chatterjee [12] we have

$$\begin{aligned} \frac{\partial \Phi}{\partial x_{i,j}} &= -\frac{1}{N} \text{Tr} \left(\frac{\partial \Delta}{\partial x_{i,j}} K^2 \right) \\ \frac{\partial^2 \Phi}{\partial x_{i,j}^2} &= 2\frac{1}{N} \text{Tr} \left(\frac{\partial \Delta}{\partial x_{i,j}} K \frac{\partial \Delta}{\partial x_{i,j}} K^2 \right) \\ \frac{\partial^3 \Phi}{\partial x_{i,j}^3} &= -6\frac{1}{N} \text{Tr} \left(\frac{\partial \Delta}{\partial x_{i,j}} K \frac{\partial \Delta}{\partial x_{i,j}} K \frac{\partial \Delta}{\partial x_{i,j}} K^2 \right) \end{aligned} \quad (6.29)$$

where $K(x) = (\Delta(x) - zI_N)^{-1}$. These identities along with some standard norm inequalities give us

$$\left\| \frac{\partial \Phi}{\partial x_{i,j}} \right\|_\infty \leq \frac{C_1}{N}, \quad \left\| \frac{\partial^2 \Phi}{\partial x_{i,j}^2} \right\|_\infty \leq \frac{C_2}{N}, \quad \left\| \frac{\partial^3 \Phi}{\partial x_{i,j}^3} \right\|_\infty \leq \frac{C_3}{N}$$

Then by definition we have

$$\begin{aligned} \lambda_2(\Phi) &\leq \sup \left\{ \left\| \frac{\partial \Phi}{\partial x_{i,j}} \right\|_\infty^2, \left\| \frac{\partial^2 \Phi}{\partial x_{i,j}^2} \right\|_\infty \right\} \leq \frac{K_1}{N} \\ \lambda_3(\Phi) &\leq \sup \left\{ \left\| \frac{\partial \Phi}{\partial x_{i,j}} \right\|_\infty^3, \left\| \frac{\partial^2 \Phi}{\partial x_{i,j}^2} \right\|_\infty^{\frac{3}{2}}, \left\| \frac{\partial^3 \Phi}{\partial x_{i,j}^3} \right\|_\infty \right\} \leq \frac{K_2}{N} \end{aligned}$$

for some $K_1, K_2 > 0$. Take $U = \mathcal{R}\Phi(\tilde{X})$ and $V = \mathcal{R}\Phi(\tilde{Y})$. Now using Lemma 6.2, we have, $\forall \epsilon > 0$

$$\begin{aligned} & |\mathbb{E}[h(U)] - \mathbb{E}[h(V)]| \leq \\ & C_1(h)\lambda_2(\Phi) \sum_{1 \leq i < j \leq N} \left[\mathbb{E} \left(|A_N^0(i, j)|^2 \mathbf{1}(|A_N^0(i, j)| > \epsilon) \right) + \mathbb{E} \left(|A_N^g(i, j)|^2 \mathbf{1}(|A_N^g(i, j)| > \epsilon) \right) \right] \\ & + C_2(h)\lambda_3(\Phi) \sum_{1 \leq i < j \leq N} \left[\mathbb{E} \left(|A_N^0(i, j)|^3 \mathbf{1}(|A_N^0(i, j)| \leq \epsilon) \right) + \mathbb{E} \left(|A_N^g(i, j)|^3 \mathbf{1}(|A_N^g(i, j)| \leq \epsilon) \right) \right] \end{aligned}$$

Denote $\mathbb{E}[X_{i,j}] = \mu_{i,j}$ and using $\lambda_2(\Phi) = O(N^{-1})$ we have⁴

$$\begin{aligned} & \lambda_2(\Phi) \sum_{1 \leq i < j \leq N} \mathbb{E} \left[|A_N^0(i, j)|^2 \mathbf{1}(|A_N^0(i, j)| > \epsilon) \right] \\ & \preceq \frac{1}{N^2} \sum_{1 \leq i < j \leq N} \mathbb{E} \left[|X_{i,j} - \mu_{i,j}|^2 \mathbf{1}(|X_{i,j} - \mu_{i,j}| > \epsilon\sqrt{N}) \right] \rightarrow 0, \text{ as } N \rightarrow \infty. \end{aligned} \quad (6.30)$$

The last limit follows from L.2. Using the definition of A_N^g from (5.1.1) and $\sup_{i,j} \sigma_{i,j} < C_0$ for some $C_0 > 0$ (L.1) we have

$$\begin{aligned} & \lambda_2(\Phi) \sum_{1 \leq i < j \leq N} \mathbb{E} \left[|A_N^g(i, j)|^2 \mathbf{1}(|A_N^g(i, j)| > \epsilon) \right] \\ & \preceq \frac{1}{N^2} \sum_{1 \leq i < j \leq N} \mathbb{E} \left[|G_{i,j}|^2 \mathbf{1}(|G_{i,j}| > \frac{\epsilon\sqrt{N}}{C_0}) \right] \preceq \mathbb{E} \left[|G_{i,j}|^2 \mathbf{1}(|G_{i,j}| > \frac{\epsilon\sqrt{N}}{C_0}) \right] \rightarrow 0 \text{ as } N \rightarrow \infty \end{aligned} \quad (6.31)$$

Now we deal with the factor involving the third derivative. Again using bounds from L.1 we have

$$\begin{aligned} & \lambda_3(\Phi) \sum_{1 \leq i < j \leq N} \mathbb{E} \left[|A_N^0(i, j)|^3 \mathbf{1}(|A_N^0(i, j)| \leq \epsilon) \right] \preceq \epsilon \frac{1}{N^2} \sum_{1 \leq i < j \leq N} \mathbb{E} \left[|X_{i,j} - \mu_{i,j}|^2 \right] \\ & \preceq \epsilon \text{ as } N \rightarrow \infty. \end{aligned} \quad (6.32)$$

Similarly for the Gaussian case we have,

$$\begin{aligned} & \lambda_3(\Phi) \sum_{1 \leq i < j \leq N} \mathbb{E} \left[|A_N^g(i, j)|^3 \mathbf{1}(|A_N^g(i, j)| \leq \epsilon) \right] \preceq \epsilon \frac{1}{N^2} \sum_{1 \leq i < j \leq N} \mathbb{E} |G_{i,j}|^2 \\ & \preceq \epsilon \text{ as } N \rightarrow \infty \end{aligned}$$

Hence for any $\epsilon > 0$ we have

$$|\mathbb{E}(h(U)) - \mathbb{E}(h(V))| \preceq \epsilon \text{ as } N \rightarrow \infty. \quad (6.33)$$

We have thus proved 5.1.3. Similarly, one can prove 5.1.4. \square

REFERENCES

- [1] E. Abbe. Community detection and stochastic block models: recent developments. *The Journal of Machine Learning Research*, 18(1):6446–6531, 2017.
- [2] O. H. Ajanki, L. Erdős, and T. Krüger. Universality for general wigner-type matrices. *Probability Theory and Related Fields*, 169(3-4):667–727, 2017.
- [3] G. W. Anderson and O. Zeitouni. A CLT for a band matrix model. *Probab. Theory Related Fields*, 134(2):283–338, 2006. ISSN 0178-8051. doi: 10.1007/s00440-004-0422-3. URL <https://doi.org/10.1007/s00440-004-0422-3>.
- [4] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118. Cambridge university press, 2010.

⁴we use the notation $f_N \preceq g_N$ if $f_N \leq Cg_N$ for some C for all N large enough.

- [5] Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010.
- [6] C. Bordenave, P. Caputo, and D. Chafaï. Spectrum of Markov generators on sparse random graphs. *Comm. Pure Appl. Math.*, 67(4):621–669, 2014. ISSN 0010-3640. doi: 10.1002/cpa.21496. URL <https://doi.org/10.1002/cpa.21496>.
- [7] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs i: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6):1801–1851, 2008.
- [8] C. Borgs, J. Chayes, H. Cohn, and Y. Zhao. An l^p theory of sparse graph convergence i: Limits, sparse random graph models, and power law distributions. *Transactions of the American Mathematical Society*, 372(5):3019–3062, 2019.
- [9] W. Bryc, A. Dembo, and T. Jiang. Spectral measure of large random hankel, markov and toeplitz matrices. *The Annals of Probability*, 34(1):1–38, 2006.
- [10] A. Chakrabarty. The Hadamard product and the free convolutions. *Statistics and Probability Letters*, 127:150–157, 2017.
- [11] A. Chakrabarty, R. S. Hazra, F. den Hollander, and M. Sfragara. Spectra of adjacency and Laplacian matrices of inhomogeneous Erdős–Rényi random graphs. To appear in *Random Matrices: Theory and Applications*, DOI: 10.1142/S201032632150009X, 2019.
- [12] S. Chatterjee. A simple invariance theorem. *arXiv preprint math/0508213*, 2005.
- [13] C. Chen, R. Xi, and N. Lin. Community detection by L_0 -penalized graph Laplacian. *Electron. J. Stat.*, 12(1):1842–1866, 2018. doi: 10.1214/18-EJS1445. URL <https://doi.org/10.1214/18-EJS1445>.
- [14] Z. Chi. Random reversible Markov matrices with tunable extremal eigenvalues. *Ann. Appl. Probab.*, 26(4):2257–2272, 2016. ISSN 1050-5164. doi: 10.1214/15-AAP1146. URL <https://doi.org/10.1214/15-AAP1146>.
- [15] F. R. K. Chung. *Spectral graph theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1997. ISBN 0-8218-0315-8.
- [16] N. Cook, W. Hachem, J. Najim, and D. Renfrew. Non-Hermitian random matrices with a variance profile (I): deterministic equivalents and limiting ESDs. *Electron. J. Probab.*, 23:Paper No. 110, 61, 2018. doi: 10.1214/18-ejp230. URL <https://doi.org/10.1214/18-ejp230>.
- [17] R. Couillet and F. Benaych-Georges. Kernel spectral clustering of large dimensional data. *Electron. J. Stat.*, 10(1):1393–1454, 2016. doi: 10.1214/16-EJS1144. URL <https://doi.org/10.1214/16-EJS1144>.
- [18] X. Ding and T. Jiang. Spectral distributions of adjacency and Laplacian matrices of random graphs. *Ann. Appl. Probab.*, 20(6):2086–2117, 2010. ISSN 1050-5164. doi: 10.1214/10-AAP677. URL <https://doi.org/10.1214/10-AAP677>.
- [19] L. Erdős, H.-T. Yau, and J. Yin. Universality for generalized Wigner matrices with Bernoulli distribution. *J. Comb.*, 2(1):15–81, 2011. ISSN 2156-3527. doi: 10.4310/JOC.2011.v2.n1.a2. URL <https://doi.org/10.4310/JOC.2011.v2.n1.a2>.
- [20] L. Erdős, H.-T. Yau, and J. Yin. Bulk universality for generalized wigner matrices. *Probability Theory and Related Fields*, 154(1-2):341–407, 2012.
- [21] W. Hachem, P. Loubaton, and J. Najim. Deterministic equivalents for certain functionals of large random matrices. *The Annals of Applied Probability*, 17(3):875–930, 2007.
- [22] J. Huang and B. Landon. Spectral statistics of sparse erdős–rényi graph laplacians. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(1):120–154, 2020. ISSN 0246-0203. doi: 10.1214/19-AIHP957. URL <https://doi.org/10.1214/19-AIHP957>.
- [23] T. Jiang. Empirical distributions of laplacian matrices of large dilute random graphs. *Random Matrices: Theory and Applications*, 1(03):1250004, 2012.
- [24] T. Jiang. Low eigenvalues of laplacian matrices of large random graphs. *Probability Theory and Related Fields*, 153(3-4):671–690, 2012.
- [25] G. D. Lin. Recent developments on the moment problem. *Journal of Statistical Distributions and Applications*, 4(1):1–17, 2017.
- [26] L. Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

- [27] L. Lovász and B. Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006.
- [28] J. A. Mingo and R. Speicher. *Free probability and random matrices*, volume 35 of *Fields Institute Monographs*. Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2017. ISBN 978-1-4939-6941-8; 978-1-4939-6942-5. doi: 10.1007/978-1-4939-6942-5. URL <https://doi.org/10.1007/978-1-4939-6942-5>.
- [29] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 13. Cambridge University Press, 2006.
- [30] A. Nica, D. Shlyakhtenko, and R. Speicher. Operator-valued distributions. I. Characterizations of freeness. *Int. Math. Res. Not.*, (29):1509–1538, 2002. ISSN 1073-7928. doi: 10.1155/S1073792802201038. URL <https://doi.org/10.1155/S1073792802201038>.
- [31] L. Pastur and V. Vasilchuk. On the law of addition of random matrices. *Communications in Mathematical Physics*, 214(2):249–286, 2000.
- [32] W. Rudin et al. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1964.
- [33] D. Shlyakhtenko. Random gaussian band matrices and freeness with amalgamation. *International Mathematics Research Notices*, 1996(20):1013–1025, 1996.
- [34] T. Squartini, J. de Mol, F. den Hollander, and D. Garlaschelli. Breaking of ensemble equivalence in networks. *Physical review letters*, 115(26):268701, 2015.
- [35] H. Touchette. Equivalence and nonequivalence of ensembles: Thermodynamic, macrostate, and measure levels. *Journal of Statistical Physics*, 159(5):987–1016, 2015.
- [36] L. V. Tran, V. H. Vu, and K. Wang. Sparse random graphs: Eigenvalues and eigenvectors. *Random Structures & Algorithms*, 42(1):110–134, 2013.
- [37] R. van der Hofstad. *Random graphs and complex networks. Vol. 1*. Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge, 2017. ISBN 978-1-107-17287-6. doi: 10.1017/9781316779422. URL <https://doi.org/10.1017/9781316779422>.
- [38] U. von Luxburg, M. Belkin, and O. Bousquet. Consistency of spectral clustering. *Ann. Statist.*, 36(2):555–586, 2008. ISSN 0090-5364. doi: 10.1214/009053607000000640. URL <https://doi.org/10.1214/009053607000000640>.
- [39] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, 62(3):548–564, 1955. ISSN 0003486X. URL <http://www.jstor.org/stable/1970079>.
- [40] Z. Zhou and A. A. Amini. Analysis of spectral clustering algorithms for community detection: the general bipartite setting. *J. Mach. Learn. Res.*, 20:Paper No. 47, 47, 2019. ISSN 1532-4435.
- [41] Y. Zhu. A graphon approach to limiting spectral distributions of Wigner-type matrices. *Random Structures & Algorithms*, 56(1):251–279, 2020.

STATISTICS DEPARTMENT, THE UNIVERSITY OF PENNSYLVANIA
 Email address: anirbanchatterjee052@gmail.com

MATHEMATICAL INSTITUTE, UNIVERSITEIT LEIDEN, NETHERLANDS
 Email address: rajatmaths@gmail.com