

BULK AND EXTREME OF SPARSE PATTERNED RANDOM MATRICES

DEBAPRATIM BANERJEE

M. Stat. student

Indian Statistical Institute

203 B.T. Road, Kolkata 700108

INDIA

debopratinbanerjee@gmail.com

ARUP BOSE *

Statistics and Mathematics Unit

Indian Statistical Institute

203 B.T. Road, Kolkata 700108

INDIA

bosearu@gmail.com

July 25, 2015
Technical Report No. R4/2015

Abstract

We consider the four $n \times n$ sparse patterned random matrices, namely the Symmetric Circulant, the Reverse Circulant, the Toeplitz and the Hankel matrices. The entries are assumed to be Bernoulli with success probability p_n such that $np_n \rightarrow \xi$ with $0 < \xi < \infty$. We use the word-moment approach to show that the expected empirical spectral distribution (EESD) converges weakly in all cases.

Unlike the Wigner matrices, here the ESD itself converges weakly to a random distribution only in the distribution sense. We give explicit descriptions of the random limits of the ESD for Reverse Circulant and Circulant matrices. Though explicit description of the limits is difficult to obtain in Toeplitz and Hankel cases, we provide some properties of these limits.

We then study the behaviour of the largest eigenvalue of these matrices. We prove that for Reverse Circulant and Symmetric Circulant matrices the limit distribution of the largest eigenvalue is a multiple of Poisson while for Toeplitz and Hankel matrices the limit has a non-degenerate implicit distribution.

Keywords: Eigenvalues, large dimensional random matrix, limiting spectral distribution, semi-circular law, weak convergence, method of moments, Wigner matrix, Symmetric Circulant matrix, Hankel matrix, Reverse Circulant matrix, Toeplitz matrix, largest eigenvalue, spectral norm, word-moment method.

AMS 2000 Subject Classification: 60B20.

*Research supported by J.C. Bose National Fellowship, Dept. of Science and Technology, Govt. of India.

1 Introduction

Suppose A_n is an $n \times n$ real symmetric matrix with eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. The empirical spectral measure μ_n of A_n is the measure on \mathbb{R} given by $\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}$, where δ_x is the Dirac delta measure at x . The cumulative probability distribution function on \mathbb{R} corresponding to μ_n , is the *Empirical Spectral Distribution* (ESD) of A_n . We will denote it by

$$F^{A_n}(x) = n^{-1} \sum_{i=1}^n \mathbb{I}\{\lambda_i \leq x\}, \quad x \in \mathbb{R}.$$

This is also known as the *bulk* of the spectrum. If the entries are random, F^{A_n} is a *random distribution function*. $\mathbb{E}[F^{A_n}(\cdot)] := F_0^{A_n}(\cdot)$ is also a distribution function, called the *expected ESD* (EESD). We shall be concerned with three types of weak convergences of the bulk distribution.

(i) F^{A_n} converges weakly almost surely to a distribution function F —outside a null set of ω , $F^{A_n}(\cdot, \omega) \xrightarrow{\mathcal{D}} F(\cdot, \omega)$. F is called the (almost sure) *limiting spectral distribution* (LSD).

(ii) $F_0^{A_n}(\cdot)$ converges weakly to (say F_0), which is then called the *expected LSD* (ELSD).

(iii) For some (possibly random) $F(\cdot, \omega)$, defined perhaps on another probability space (Ω, \mathcal{F}, P) , and for every bounded continuous function $g : \mathbb{R} \rightarrow \mathbb{R}$,

$$Y_g^n = \int g(x) dF^{A_n}(x, \omega') \xrightarrow{\mathcal{D}} Y_g = \int g(x) dF(x, \omega).$$

We then say that F^{A_n} converges weakly in distribution to F . It is easy to see that weak convergence almost surely of F^{A_n} to F implies weak convergence in distribution to F .

We deal with the symmetric patterned matrices, Toeplitz, Hankel, Symmetric Circulant and Reverse Circulant given below. We also need the symmetric Wigner matrix as a reference point.

$$T_n = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_{n-2} & x_{n-1} \\ x_1 & x_0 & x_1 & \dots & x_{n-3} & x_{n-2} \\ x_2 & x_1 & x_0 & \dots & x_{n-4} & x_{n-3} \\ & & & \ddots & & \\ x_{n-1} & x_{n-2} & x_{n-3} & \dots & x_1 & x_0 \end{bmatrix}, \quad SC_n = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_2 & x_1 \\ x_1 & x_0 & x_1 & \dots & x_3 & x_2 \\ x_2 & x_1 & x_0 & \dots & x_2 & x_3 \\ & & & \ddots & & \\ x_1 & x_2 & x_3 & \dots & x_1 & x_0 \end{bmatrix},$$

$$H_n = \begin{bmatrix} x_2 & x_3 & x_4 & \dots & x_n & x_{n+1} \\ x_3 & x_4 & x_5 & \dots & x_{n+1} & x_{n+2} \\ x_4 & x_5 & x_6 & \dots & x_{n+2} & x_{n+3} \\ & & & \ddots & & \\ x_{n+1} & x_{n+2} & x_{n+3} & \dots & x_{2n-1} & x_{2n} \end{bmatrix}, \quad R_n = \begin{bmatrix} x_2 & x_3 & x_4 & \dots & x_0 & x_1 \\ x_3 & x_4 & x_5 & \dots & x_1 & x_2 \\ x_4 & x_5 & x_6 & \dots & x_2 & x_3 \\ & & & \ddots & & \\ x_1 & x_2 & x_3 & \dots & x_{n-1} & x_0 \end{bmatrix}.$$

$$W_n = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1(n-1)} & x_{1n} \\ x_{12} & x_{22} & x_{23} & \dots & x_{2(n-1)} & x_{2n} \\ & & & \ddots & & \\ x_{1n} & x_{2n} & x_{3n} & \dots & x_{(n-1)n} & x_{nn} \end{bmatrix}.$$

Let A_n denote any of these $n \times n$ five patterned matrices. If the input random variables are i.i.d. with mean zero and variance one, then the LSD and ELSD of $n^{-1/2}A_n$ and $n^{-1/2}A_n$ exist and are equal. For a unified proof that uses the *word-moment method*, see Bose and Sen (2008) [8]. For example the LSD of $n^{-1/2}W_n$ is semicircle while the LSD of the Symmetric Circulant and Reverse Circulant are standard normal and the symmetrized square root of $2\chi_2^2$ respectively. However, no nice descriptions are known for the Toeplitz and Hankel cases. If the input random variables are i.i.d. Gaussian with mean 0, then the limiting distribution of the (appropriately normalised) maximum eigenvalue of the Wigner matrix is the Tracy-Widom distribution (see Tracy and Widom (2000) [15]). Bose, Hazra and Saha (2011) [6] proved that under certain moment assumptions, the properly normalized spectral norms of Reverse Circulant and Symmetric Circulant matrices are asymptotically Gumbel distributed. No non-degenerate limit distribution results are known for the maximum eigenvalue in the Hankel and Toeplitz cases.

We wish to study the sparse case of the first four matrices. So assume that the input random variables are a triangular sequence of i.i.d. Bernoulli random variables $\{x_i^{(n)}\}_{i \in I}$ with success probability p_n , $np_n \rightarrow \xi \in (0, \infty)$. Here I is \mathbb{N} or \mathbb{N}^2 as required. The sparse Wigner is equivalent to the incidence matrix of the sparse Erdos-Renyi graph. In this case the LSD exists in the sense (i) and has unbounded support, see Khorunzhy et al. (2004) [12] and Bauer et al. (2001) [3]. The LSD has a dense set of atoms, see Chayes et al. (1986) [10]. Bordenave et al. (2013) [5] proved that the LSD has a continuous part if and only if $\xi > 1$. Among other authors, the largest eigenvalue of the sparse Wigner was studied by Krivelevich and Sudakov (2003) [14]. They proved that for any $\xi > 0$, $\lambda_{max} = (1 + o(1))\sqrt{\frac{\log n}{\log \log n}}$.

We first prove the convergence of the EESD by the word-moment method and describe the contribution of each word for the Reverse Circulant and the Symmetric Circulant cases. In these cases, the ELSD are Poisson mixture of bounded random variables and have infinite support. Any reasonable description eludes us in the Toeplitz and the Hankel cases.

We then show that the LSD exists for all the matrices in the distribution sense (iii). For the Reverse Circulant and Symmetric Circulant matrices, conditioned on the (limiting) number of successes in the Bernoulli sequence, these LSD have bounded support. The Toeplitz and Hankel matrices are much harder to deal with. We relate the moments of the ESD to path counts of certain restricted symmetric random walks with finite jump lengths. This is used to establish the existence of the LSD. As expected, the ELSD are the expectation of the LSD.

It is known that the almost sure LSD of the sparse Wigner has a dense set of atoms. The sparse Symmetric Circulant ELSD has an atom only at 0 and the sparse Reverse Circulant ELSD has atoms at 0 and ± 1 . The proof of Theorem 5.2 (c) implies that the Toeplitz ELSD has atoms at ± 1 and 0, even though it is structurally similar to the Symmetric Circulant. It would be interesting to find the location and mass of all the atoms for the Hankel and Toeplitz ELSD.

We also establish the convergence of the expected trace of any polynomial in independent copies of sparse matrices of a fixed pattern. This implies that the ELSD exists for any symmetric matrix polynomial in independent copies of a fixed sparse patterned matrix. As in the usual case, independent sparse Symmetric Circulant matrices are also independent in the limiting sense. However, unlike the usual case, independent sparse Wigners are not asymptotically free.

Then we study the largest eigenvalue of these sparse matrices. We prove that the limit of the spectral norm is a multiple of the Poisson random variable in case of Reverse and Symmetric Circular matrices and is a non-degenerate implicit distribution in case of Toeplitz and Hankel matrices. Apparently, from the available evidence in the literature, the non-degenerate limiting

distribution of the (normalized) maximum eigenvalue of sparse Wigner does not exist.

Different aspects of some other sparse random matrix models have been studied by others. For example Bai and Silverstein (2006) [2] essentially considered a sequence random matrices A_n and a sequence of sparse random matrices D_n and studied their Hadamard product. They considered D_n to be symmetric having Bernoulli $(0, 1)$ entries with probability of the (i, j) th entry being 1 is $p_{n,i,j}$ where $\sum_i p_{n,i,j} = p + o(p)$ and $p \rightarrow \infty$ as $n \rightarrow \infty$.

Section 2 gives the preliminaries of the word-moment approach. Section 3 contains the results on the ELSD. Section 4 gives the results on the LSD in the sense (iii). Section 5 contains the result on the maximum eigenvalue. All proofs are given in Section 6. A few simulation results complete the paper.

2 Preliminaries of the word-moment method

For any distribution F or any random variable X , $\beta_h(F)$ and $\beta_h(X)$ respectively will denote their h -th moment. The word-moment method to establish the LSD is well-known (see for example Bose and Sen (2008) [8]). We shall use this method in this article. The h -th moments of the ESD and the expected ESD of A_n have the following nice forms:

$$\begin{aligned} h\text{-th moment of } F^{A_n} &= \frac{1}{n} \sum_{i=1}^n \lambda_i^h = \frac{1}{n} \text{Tr}(A_n^h) = \beta_h(A_n) \text{ (say),} & (2.1) \\ h\text{-th moment of } F_0^{A_n} &= \frac{1}{n} E[\text{Tr}(A_n^h)] = E(\beta_h(A_n)) \end{aligned}$$

where Tr is the trace. Lemma 2.1 links convergence of moments and LSD. We omit its proof.

Lemma 2.1. *The expected ESD $\{F_0^{A_n}\}$ converges weakly to F determined by $\{\beta_h\}$ provided the following two conditions hold.*

(M1) For every $h \geq 1$, $E[\beta_h(A_n)] \rightarrow \beta_h$

(C) The sequence $\{\beta_h\}$ satisfies Carleman's condition $\sum_{h=1}^{\infty} \beta_{2h}^{-1/2h} = \infty$.

A sequence of variables $\{x_i^{(n)}\}$ or $\{x_{(i,j)}^{(n)}\}$ used to construct A_n will be called an *input sequence*. Let \mathbb{Z} be the set of all integers and let \mathbb{Z}_+ denote the set of all non-negative integers. Let

$$L_n : \{1, 2, \dots, n\}^2 \rightarrow \mathbb{Z}^2 \text{ (or } \mathbb{Z})$$

be a sequence of functions. We shall write $L_n = L$ and call it the *link function* and by abuse of notation we write \mathbb{Z}_+^2 as the common domain of $\{L_n\}$. A_n is a *patterned matrix* with the link function L_n if $A_n(i, j) = x_{L_n(i,j)}^{(n)}$ is its (i, j) th element. The link function for the four matrices are

$$\begin{aligned} L_T(i, j) &= |i - j|, \\ L_H(i, j) &= i + j, \\ L_{RC}(i, j) &= (i + j) \pmod{n}, \text{ and} \\ L_{SC}(i, j) &= n/2 - |n/2 - |i - j||. \end{aligned}$$

It is easy to see that $E(\text{Tr}(A_n^2))$ converges. Thus, to investigate the LSD of these matrices, unlike the standard case, no scaling is required. The moment trace formula gives us

$$\frac{1}{n} \text{Tr}(A_n^h) = \frac{1}{n} \sum_{1 \leq i_1, i_2, \dots, i_h \leq n} x_{L(i_1, i_2)}^{(n)} x_{L(i_2, i_3)}^{(n)} \cdots x_{L(i_h, i_1)}^{(n)}. \quad (2.2)$$

To write the above in a more convenient form, define a *circuit* to be any function $\pi : \{0, 1, 2, \dots, h\} \rightarrow \{1, 2, \dots, n\}$ with $\pi(0) = \pi(h)$. The dependence of a circuit on h and n will be suppressed. Clearly, the convergence in (M1) may be written in terms of circuits since

$$E[\beta_h(A_n)] = E\left[\frac{1}{n} \text{Tr}(A_n^h)\right] = \frac{1}{n} \sum_{\pi} E[x_{L(\pi(0), \pi(1))}^{(n)} x_{L(\pi(1), \pi(2))}^{(n)} \cdots x_{L(\pi(h-1), \pi(h))}^{(n)}] \quad (2.3)$$

$$= \frac{1}{n} \sum_{\pi} E[X_{\pi}] \quad (\text{say}). \quad (2.4)$$

Circuits π_1 and π_2 are said to be *equivalent* if and only if their L values respectively match at the same locations. That is, for all i, j ,

$$L(\pi_1(i-1), \pi_1(i)) = L(\pi_1(j-1), \pi_1(j)) \Leftrightarrow L(\pi_2(i-1), \pi_2(i)) = L(\pi_2(j-1), \pi_2(j)).$$

Any equivalence class is indexed by a partition of $\{1, 2, \dots, h\}$. Each block of a given partition identifies the positions of the L -matches. We label these partitions by *words* w of length h of letters where the first occurrence of each letter is in alphabetic order. For example if $h = 5$ then the partition $\{\{1, 3, 5\}, \{2, 4\}\}$ is represented by the word *ababa*. This identifies the circuits π such that $L(\pi(0), \pi(1)) = L(\pi(2), \pi(3)) = L(\pi(4), \pi(5))$ and $L(\pi(1), \pi(2)) = L(\pi(3), \pi(4))$. The set of all words of length h will be denoted by \mathcal{W}_h . Define the following classes of circuits:

$$\begin{aligned} \Pi(w) &= \{\pi : w(i) = w(j) \Leftrightarrow L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j))\}, \\ \Pi^*(w) &= \{\pi : w(i) = w(j) \Rightarrow L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j))\}. \end{aligned}$$

Any i (or $\pi(i)$ by abuse of notation) will be called a *vertex*. It is *generating* if either $i = 0$ or $w[i]$ is the first occurrence of a letter. For example, if $w = abbcab$ then $\pi(0), \pi(1), \pi(2), \pi(4)$ are generating. In our case, a circuit will be completely determined, up to a finitely many choices, by its generating vertices. For any word w , let

$$l(w) = \# \text{ of letters in } w \quad \text{and} \quad d(w) = \# \text{ of distinct letters in } w.$$

Then the number of generating vertices is $d(w) + 1$ and it follows easily that

$$\#\Pi^*(w) = O(n^{d(w)+1}) \quad \text{and} \quad \#\Pi^*(w) \setminus \Pi(w) = O(n^{d(w)}).$$

Note that for a word w , $\forall \pi \in \Pi(w)$, $E[X_{\pi}]$ is same. We denote this common expectation by $E[X_w]$. From the definitions of word and circuit we get

$$\begin{aligned} E(\beta_h(A_n)) &= E\left\{\frac{1}{n} \sum_{i=1}^n \lambda_i^h\right\} = \frac{1}{n} \sum_w [\#\Pi(w)] E X_w \\ &= \sum_w \frac{1}{n^{d(w)+1}} [\#\Pi(w)] (np_n)^{d(w)}. \end{aligned}$$

For any w , let (whenever the limit below exists)

$$p(w) = \lim_{n \rightarrow \infty} \frac{1}{n^{d(w)+1}} \#\Pi(w). \quad (2.5)$$

The strategy to establish the existence of the ELSD for the four matrices, is to first prove that $p(w)$ exists in each case. Then

$$E(\beta_h(A_n)) \rightarrow \beta_h = \sum_{w \in \mathcal{W}_h} p(w) \xi^{d(w)}. \quad (2.6)$$

That would verify the (M1) condition. The verification of Condition (C) is common to all the four matrices and is done as follows.

Lemma 2.2. *If $p(w)$ exists for each w , then $\{\beta_h\}$ given in (2.6) satisfies Condition (C).*

Proof. There are only finitely many choices of the non-generating vertices once the generating vertices are fixed. So, $\#\Pi(w) \leq O(n^{d(w)+1})$. As a result, for some integer Δ independent of $d(w)$, $p(w) \leq \Delta^{d(w)}$. Hence

$$\beta_h \leq \Delta^h \sum_{i=1}^h S(h, i) \xi^i \quad (h \text{ even}).$$

where $S(h, i)$ = number of all possible partitions of h distinct objects into i groups. But the right side is the h th moment of ΔP where P is a Poisson random variable with mean ξ . It is easy to check that the the moments of ΔP satisfies condition (C) and hence so does $\{\beta_h\}$. \square

3 Convergence of EESD

We are now in a position to state our results on the convergence of the EESD.

Theorem 3.1. (a) $F_0^{SC_n}$ converges to a distribution with (2k)th moment $\sum_{w \in \mathcal{W}_{2k}} p(w) \xi^{d(w)}$ where

$$p(w) = \begin{cases} \prod_{i=1}^d \binom{2k_i-1}{k_i}, & \text{if } w \text{ has } d \text{ distinct letters, appearing } 2k_i \text{ times, } 1 \leq i \leq d \\ 0, & \text{otherwise.} \end{cases}$$

(b) $F_0^{RC_n}$ converges to a distribution with (2k)th moment $\sum_{w \in \mathcal{W}_{2k}} p(w) \xi^{d(w)}$ where

$$p(w) = \begin{cases} 1, & \text{if each letter appears equally often at odd and even positions} \\ 0, & \text{otherwise.} \end{cases}$$

In both cases, the odd moments are zero.

For Toeplitz and Hankel Matrices, we are only able to prove the existence of $p(w)$ for each word w . The explicit calculation of $p(w)$ for a given w is difficult. It may be noted that for the mean 0 variance 1 case too no explicit formula for the moments are known in the literature.

Theorem 3.2. (a) $F_0^{T_n}$ converges to a distribution which has all odd moments zero and (2k)th moment $\beta_{2k} = \sum_{w \in \mathcal{W}_{2k}} p(w) \xi^{d(w)}$. Here $p(w)$ exists for each word $w \in \mathcal{W}_{2k}$.

(b) $F_0^{H_n}$ converges to a distribution which has all odd moments zero and (2k)th moment $\beta_{2k} = \sum_{w \in \mathcal{W}_{2k}} p(w) \xi^{d(w)}$ where $p(w)$ exists for each w and $p(w) = 0$ if each letter does not appear equally often in even and odd positions of w .

Remark 3.1. Let A_n be any of the four sparse patterned matrices and $np_n \rightarrow \xi$. Then the above four results say that $F_0^{A_n}$ converges to F_ξ^A . It is natural to ask what happens if we use Binomial inputs instead of Bernoulli inputs or if we let $\xi \rightarrow \infty$. We can prove the following:

(a) Let $\{B_n\}_{n=1}^\infty$ be any of the four patterned matrices with independent $\text{Bin}(k, p_n)$ entries, $np_n \rightarrow \xi$. Then $F_0^{B_n} \xrightarrow{\mathcal{D}} F_{k\xi}^A$.

(b) If the entries are i.i.d. with mean zero variance one, then as discussed in the introduction, $F_0^{n^{-1/2}A_n}$ converges to some F^A . Let X_ξ^A and X^A be random variables with distribution F_ξ^A and F^A respectively. Then

$$X_\xi^A / \sqrt{\xi} \xrightarrow{\mathcal{D}} X^A \quad \text{as } \xi \rightarrow \infty. \quad (3.1)$$

We now consider the joint convergence of i.i.d. copies of the same patterned matrix. We need a few definitions. A non-commutative probability space (NCP) for our purposes is a pair (\mathcal{A}, ϕ) , where \mathcal{A} is a unital algebra (with unity 1) and $\phi : \mathcal{A} \rightarrow \mathbb{C}$ is a linear functional satisfying $\phi(1) = 1$. The typical example for us is

$$\mathcal{A}_n := \{\text{All } n \times n \text{ random matrices with entries from } (\Omega, \mathcal{F}, \mathbb{P}) \text{ with all moments finite}\}$$

and $\phi_n(A) = \frac{1}{n} \mathbb{E}[\text{Tr}[A]] \quad \forall A \in \mathcal{A}_n$. Let (\mathcal{A}_n, ϕ_n) be a sequence of non-commutative probability spaces and (\mathcal{A}, ϕ) be another non-commutative probability space. Let J be a finite subset of \mathbb{N} . We say $\{A_{n,i} \in \mathcal{A}_n \mid i \in J\}$ converges jointly in law to $\{A_i \in \mathcal{A} \mid i \in J\}$, if for every monomial $(A_{n,i_1} A_{n,i_2} \dots A_{n,i_l})$ with $i_m \in J \quad \forall m \in \mathbb{N}$,

$$\lim_{n \rightarrow \infty} \phi_n(A_{n,i_1} A_{n,i_2} \dots A_{n,i_l}) = \phi(A_{i_1} A_{i_2} \dots A_{i_l}).$$

Theorem 3.3. Let $\{A_{i,n}\}_{1 \leq i \leq l}$ be l i.i.d. sparse random matrices of any of the four kinds where the entries are i.i.d Bernoulli(p_n) and $np_n \rightarrow \xi$. Then the NCP $(\text{Span}\{A_{i,n}\}, \frac{1}{n} \mathbb{E} \text{Tr})$ converges.

The theorem implies that the average trace of any polynomial of independent sparse identically patterned matrices, converges. Fix a polynomial which is also a symmetric matrix. Then the theorem implies that this matrix satisfies Condition (M1). By extending the argument given for a single matrix, it is not hard to show that Condition (C) for the limiting moments is also satisfied. We omit the details. Thus the ESD of any symmetric polynomial exists.

Note that Symmetric Circulant matrices commute. It can be shown that the corresponding limit variables in this case have a valid joint probability distribution where the components are i.i.d. with the marginal distribution being as given uniquely by the moments in Theorem 3.1.

When the input sequence is i.i.d. with all moments finite, independent Wigner matrices are *asymptotically free*, see page 380 of Anderson, Guionnet and Zeitouni (2010) [1]. This does not remain true in the sparse case. This can be shown by considering specific lower order joint moments and showing that the limiting freeness condition is violated. We omit the details.

4 Convergence of the ESD

As already discussed, the sparse Wigner ESD converges weakly almost surely to a non-random distribution with infinite support. In the other four cases, we show that the ESD converges

weakly in distribution to a random distribution. For the Reverse Circulant and the Symmetric Circulant matrices we use explicit expressions for the eigenvalues to identify the limits. For the Toeplitz and Hankel matrices we only prove the existence of the limit.

Reverse Circulant matrix. The eigenvalues are given by

$$\begin{cases} \lambda_0 &= \sum_{i=0}^{n-1} x_i \\ \lambda_{n/2} &= \sum_{i=0}^{n-1} (-1)^i x_i, \text{ if } n \text{ is even} \\ \lambda_k = -\lambda_{n-k} &= \sqrt{\left(\sum_{i=0}^{n-1} x_i \cos\left(\frac{2\pi ik}{n}\right)\right)^2 + \left(\sum_{i=0}^{n-1} x_i \sin\left(\frac{2\pi ik}{n}\right)\right)^2}, \quad 1 \leq k \leq \lfloor \frac{n-1}{2} \rfloor. \end{cases} \quad (4.1)$$

Symmetric Circulant matrix. The eigenvalues are given by

(i) If n is odd

$$\begin{cases} \lambda_0 &= x_0 + 2 \left(\sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} x_j \right) \\ \lambda_k &= x_0 + 2 \left(\sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} x_j \cos\left(\frac{2\pi jk}{n}\right) \right), \quad \forall 1 \leq k \leq \lfloor \frac{n}{2} \rfloor. \end{cases} \quad (4.2)$$

(ii) If n is even

$$\begin{cases} \lambda_0 &= x_0 + 2 \left(\sum_{j=1}^{\frac{n}{2}-1} x_j \right) + x_{\frac{n}{2}} \\ \lambda_k &= x_0 + 2 \left(\sum_{j=1}^{\frac{n}{2}-1} x_j \cos\left(\frac{2\pi jk}{n}\right) \right) + (-1)^k x_{\frac{n}{2}} \quad \forall 1 \leq k \leq \frac{n}{2}. \end{cases}$$

For $k \geq \frac{n}{2}$, $\lambda_k = \lambda_{n-k}$ in both cases (i) and (ii).

Let $\{U_i\}_{i=1}^{\infty}$ be i.i.d. $U(0, 1)$ and independent of all other variables.

Theorem 4.1. (a) Let $Q \sim \text{Poisson}(\xi)$. Let $F(\cdot, \omega)$ be the distribution of the symmetric square root of $\left(\sum_{i=1}^{Q(\omega)} \cos(2\pi U_i)\right)^2 + \left(\sum_{i=1}^{Q(\omega)} \sin(2\pi U_i)\right)^2$. Then $F^{RC_n}(\cdot, \omega)$ converges weakly to $F(\cdot, \omega)$ in distribution.

(b) Let $Q \sim \text{Poisson}\left(\frac{\xi}{2}\right)$. Then $F^{SC_n}(\cdot, \omega)$ converges weakly to $F(\cdot, \omega)$ in distribution where $F(x, \omega) := \mathbb{P}\left[2 \sum_{i=1}^{Q(\omega)} \cos(2\pi U_i) \leq x\right]$.

(c) Both $F^{T_n}(\cdot, \omega)$ and $F^{H_n}(\cdot, \omega)$ converge weakly in distribution.

Let A_n be any of the four sparse patterned matrices with LSD and ELSD as $F(x, \omega)$ and $F_0(x)$ respectively. Then it can be shown that $\mathbb{E}[F(x, \omega)] = F_0(x) \forall x \in \mathbb{R}$.

5 Extreme eigenvalues

For any matrix A , let $\lambda_{\max, A}$ and $\lambda_{\min, A}$ denote respectively the largest and the smallest eigenvalues of A . The largest eigenvalue of sparse Wigner matrices were studied by Sudakov and Krivelevich (2003) [14]. In particular they proved the following.

Theorem 5.1. (Sudakov and Krivelevich) Let $G = G(n, p)$ be a random graph and let Δ and A be the maximum degree of G and the adjacency matrix of G respectively. Then almost surely $\lambda_{\max, A} = (1 + o(1)) \max\{\sqrt{\Delta}, np\}$.

However, when $np \rightarrow \xi \in (0, \infty)$, the asymptotic value of Δ is known (see Lemma 2.2 of Sudakov and Krivelevich (2003) [14]). As a consequence one obtains the following corollary.

Corollary. (Sudakov and Krivelevich) For any $\xi > 0$, almost surely $\lambda_{\max, A} = (1 + o(1)) \sqrt{\frac{\log n}{\log \log n}}$.

Interestingly the largest eigenvalue asymptotically does not depend on the value of ξ . We believe that normalized $\lambda_{\max, A}$ does not converge to any non-degenerate proper distribution.

The situation is completely different in our case. We prove that the largest eigenvalue of the other four matrices have non-degenerate limiting distribution without any further requirement of centering and scaling. The spectral norm of A is defined as

$$\text{sp}_A = \max\{|\lambda_{\max, A}|, |\lambda_{\min, A}|\}.$$

For a random probability distribution $F(\cdot, \omega)$ on \mathbb{R} , which is a measurable function of ω , let

$$\text{supp}(F(\cdot, \omega)) = |\inf\{x : F(x, \omega) = 1\}| \vee |\sup\{x : F(x, \omega) = 0\}|$$

be the largest absolute value (possibly infinite) of the support of the random probability measure.

Theorem 5.2. (a) $\text{sp}_{RC_n} \xrightarrow{\mathcal{D}} Q \sim \text{Poisson}(\xi)$.

(b) $\text{sp}_{SC_n} \xrightarrow{\mathcal{D}} 2Q$ where $Q \sim \text{Poisson}(\frac{\xi}{2})$.

(c) $\text{sp}_{T_n} \xrightarrow{\mathcal{D}} \lambda_T \stackrel{\mathcal{D}}{=} \text{supp}(F^T(\cdot, \omega))$ where $F^T(\cdot, \omega)$ is the LSD of T_n .

(d) $\text{sp}_{H_n} \xrightarrow{\mathcal{D}} \lambda_H \stackrel{\mathcal{D}}{=} \text{supp}(F^H(\cdot, \omega))$ where $F^H(\cdot, \omega)$ is the LSD of H_n .

(e) In each of the above cases, λ_{\max, A_n} and $-\lambda_{\min, A_n}$ have the same limit distributions.

6 Proofs

Proof of Theorem 3.1. (a) We need to establish the formula for $p(w)$. Define the slopes

$$s(l) = \pi(l) - \pi(l-1), \quad 1 \leq l \leq k. \quad (6.1)$$

Note that the vertices (i, j) match if and only if

$$s(i) - s(j) = 0, \pm n \text{ (} q \text{ match, say)} \text{ or} \quad (6.2)$$

$$s(i) + s(j) = 0, \pm n \text{ (} m \text{ match, say)}. \quad (6.3)$$

Now consider an arbitrary word w and an arbitrary letter x in it. Suppose x appears at positions $p_1 < \dots < p_l$. We label x in these positions as x_1, \dots, x_l . We say that p_1 is in q position

(convention). Any $x_j, 1 < j \leq l$ is in q or m position according as (x_1, x_j) is a q or an m match. In that case we also say x_j (or j) is a q or an m vertex. Let

$$q_x = \text{number of } q \text{ positions of } x, \text{ and } m_x = \text{number of } m \text{ positions of } x.$$

We shall show that

$$p(w) \neq 0 \text{ if and only if } q_x = m_x \text{ for all } x. \quad (6.4)$$

Hence, for a word to make a non-trivial contribution, every letter must appear an even number of times. Further, we shall show that all possible rearrangements of the q and m positions (except the first q position) will give exactly one valid word constraint provided the π values at the generating vertices are fixed. This will then establish Theorem 3.1 (a).

Before we prove (6.4), let us consider two examples. First consider the word $aabb$. Then it can be checked that the only combinations possible are

$$s(2) + s(1) = 0, \pm n \text{ and } s(4) + s(3) = 0, \pm n,$$

yielding $p(aabb) = 1$. Now consider the word $w = aaaa$. It can be checked that then three sets of slope constraints are possible:

(i) $s(2) + s(1) = 0, \pm n, s(3) + s(2) = 0, \pm n$ and $s(4) + s(3) = 0, \pm n$.

(In this case x_1, x_3 are in q positions and x_2, x_4 are in m positions).

(ii) $s(2) + s(1) = 0, \pm n, s(3) - s(2) = 0, \pm n$ and $s(4) + s(3) = 0, \pm n$.

(In this case x_1, x_4 are in q positions and x_2, x_3 are in m positions).

(iii) $s(2) - s(1) = 0, \pm n, s(3) + s(2) = 0, \pm n$ and $s(4) - s(3) = 0, \pm n$.

(In this case x_1, x_2 are in q positions and x_3, x_4 are in m positions). This yields $p(aaaa) = 3$.

Proof of only if part. First suppose x is the *last* distinct letter of w . For $p(w)$ to be non-zero, keeping in mind equation (2.5), the generating vertex $\pi(p_1)$ should not depend on the other generating vertices. Let us see where that leads us to.

Fix the value of all generating vertices except $\pi(p_1)$. Let us start from the extreme right vertex and move left. By circuit condition, $\pi(k) = \pi(0)$. Now suppose $k > p_l$. Note that $\pi(k)$ matches with some vertex to the left of p_1 . Hence $\pi(k-1)$ is determined in terms of generating vertices *before* p_1 . Moving step at a time, $\pi(v)$ for all $v \geq p_l$ do not depend on $\pi(p_1)$.

Now consider $\pi(p_1 + 1)$. Assume $p_1 + 1 < p_2$. Since this vertex matches with a vertex to the left of p_1 , it follows that,

$$\pi(p_1 + 1) = \pi(p_1) + c \pmod{n},$$

where c is a constant depending on the fixed π values. Inductively this holds for all vertices *between* p_1 and p_2 . So $\pi(p) = \pi(p_1) + c \pmod{n}$ for all $p_1 < p < p_2$. If p_2 is an m position then

$$\pi(p_2) = c \pmod{n}. \quad (6.5)$$

If p_2 is a q position, then

$$\begin{aligned} s(p_2) - s(p_1) &= 0 \pmod{n} \\ \Rightarrow \pi(p_2) - \pi(p_2 - 1) - [\pi(p_1) - \pi(p_1 - 1)] &= 0 \pmod{n} \\ \Rightarrow \pi(p_2) &= 2\pi(p_1) + c \pmod{n}. \end{aligned} \quad (6.6)$$

Proceeding inductively, for any $j < l$,

$$\pi(p_j) = c_j \pi(p_1) + c \pmod{n}, \quad (6.7)$$

$$c_{j+1} = \begin{cases} m_j + 1 & \text{if } p_{j+1} \text{ is a } q \text{ vertex} \\ c_j - 1 & \text{if } p_{j+1} \text{ is an } m \text{ vertex.} \end{cases} \quad (6.8)$$

Hence eventually

$$\pi(p_l) = (q_x - m_x) \pi(p_1) + c \pmod{n}.$$

On the other hand we have already argued that $\pi(p_l)$ must be independent of $\pi(p_1)$. Hence we must have $q_x = m_x$ when x is the last distinct letter.

Conversely, it is also clear from the above argument that if $q_x = m_x$ then for any possible combination of q and m vertices, $\pi(p_i)$ is independent of $\pi(p_1)$. Thus (6.4) is proved when x is the last distinct letter.

Now we proceed to establish the only if part of (6.4) for all letters by induction. Suppose w has d distinct letters and (6.4) has been proved for the d th, \dots , $d - (m - 1)$ th distinct letters. Assume now that x is the $(d - m)$ th distinct letter.

Unlike before, the L values appearing after position p_1 , may or may not depend on $\pi(p_1)$. For instance, consider the word $w = abcdacbb$. Take the $(d - m)$ th distinct letter to be b . Here $d = 4, m = 2$. $\pi(0), \pi(1), \pi(2), \pi(3), \pi(4)$ are the generating vertices. Fix all these vertices except $\pi(2)$ (first appearance of b). Then $L(\pi(3), \pi(4)) = |\frac{q}{2} - |\frac{q}{2} - \pi(3) - \pi(4)||$ does not depend on $\pi(2)$.

Letters at vertices *after* $\pi(p_1)$ whose L value depend on $\pi(p_1)$ will be called *dependent* letters. Once the values of other generating vertices are fixed, only the dependent letters have an effect on the value of $\pi(p_1)$. In particular if a letter x does not have any dependent letter except itself, the above argument implies $q_x = m_x$. So instead of considering all the letters we may perform the induction on the number of dependent letters. Note that the last letter always has only itself as its dependent letter for which we have already proved the result. Now assume that we $\gamma > 1$ distinct dependent letters. Let l be the total number of occurrence of these dependent letters and $p_1 < \dots < p_l$ be their positions respectively. Among $p_1 < \dots < p_l$, let f_1, \dots, f_γ be the positions of the first occurrence of these letters. Observe that $p_1 = f_1$ and the letter at this position is x but the same cannot be said about the position p_l as it is the position of the last occurrence of dependent letters, not necessarily the letter x . By successive applications of (6.2) and (6.3), it is easy to observe that for any $\tau > p_1$,

$$\pi(\tau) = c_\tau \pi(p_1) + c \pmod{n}.$$

It is enough to prove that $c_{p_l} = q_x - m_x$. We start from the position f_γ and move in the forward direction to track the values of c_τ for $\tau \geq f_\gamma$. Let for $i \leq \gamma$,

$$q_{i,j} = \#q \text{ positions of the letter } i \text{ in between } f_j \text{ and } f_{j+1},$$

$$m_{i,j} = \#m \text{ positions of the letter } i \text{ in between } f_j \text{ and } f_{j+1}.$$

Note that unlike $i = 1$, for any $i > 1$, $\pi(f_i - 1)$ depends on $\pi(p_1)$. So in contrast to the first letter, an m match of the $i (> 1)$ th letter at position τ will now increase c_τ by c_{f_i-1} and a q match will decrease c_τ by c_{f_i-1} . As a consequence,

$$c_{p_l} = - \sum_{i=2}^{\gamma} c_{f_i-1} (q_{i,\gamma} - m_{i,\gamma}) + q_{1\gamma} - m_{1\gamma}. \quad (6.9)$$

Now by the same argument as above,

$$c_{f_{\gamma-1}} = q_{1,\gamma-1} - m_{1,\gamma-1} - \sum_{i=2}^{\gamma-1} c_{f_{i-1}}(q_{i,\gamma-1} - m_{i,\gamma-1}). \quad (6.10)$$

Observe that $q_{\gamma,\gamma} - m_{\gamma,\gamma} = -1$ since we have excluded f_{γ} in our count. Putting the value $c_{f_{\gamma-1}}$ and using $q_{\gamma,\gamma} - m_{\gamma,\gamma} = -1$, R.S. of (6.9) reduces to

$$- \sum_{i=2}^{\gamma-1} c_{f_{i-1}}(q_{i\gamma} + q_{i\gamma-1} - m_{i\gamma} - m_{i\gamma-1}) + q_{1,\gamma} - m_{1,\gamma} + q_{1,\gamma-1} - m_{1,\gamma-1}. \quad (6.11)$$

Applying this argument successively for each $i > 2$, we get the following form of R.S. of (6.9).

$$- \sum_{j=2}^{\gamma} c_{f_{j-1}}(q_{2,j} - m_{2,j}) + \sum_{j=2}^{\gamma} (q_{1,j} - m_{1,j}). \quad (6.12)$$

Now $\sum_{j=2}^{\gamma} (q_{2,j} - m_{2,j}) = -1$ by induction. However in this case

$$c_{f_{2-1}} = q_{1,1} + 1 - m_{1,1}.$$

So we get the following final reduction of (6.9).

$$c_{p_l} = \sum_{j=1}^{\gamma} q_{1,j} - \sum_{j=1}^{\gamma} m_{1,j} + 1 = q_x - m_x.$$

This completes the proof of the *only if part*.

Proof of if part. We essentially retrace the proof of the only if part. Consider any word in which d distinct letters appear $2k_i$ times, $1 \leq i \leq d$, each appearing an equal number of times at q and m positions. The L values are determined from the slope constraints once we fix the q and the m positions. It then remains to show that the circuit condition $\pi(k) = \pi(0)$ is satisfied. Note that using all the slope constraints etc.,

$$\pi(k) = \sum_{i=0}^d f_i \pi(\gamma_i) \pmod{n} \quad (6.13)$$

for some $\{f_i\}$ and $\{0 < \gamma_0 < \dots < \gamma_d\}$ are all the $(d+1)$ generating vertices. In the only if part, we have proved that except $\pi(0)$, all other generating vertices cancel out at the position of the last dependent letter corresponding to that generating vertex and the coefficients thereafter do not depend on the value of the particular generating vertex. Now the last occurrence of the dependent letter of a particular generating vertex has to be at or before the last position of the word. Hence $\pi(k) = f_0 \pi(0) \pmod{n}$. By arguments similar to the only if part, one can prove that $f_0 = 1$. Hence the circuit condition is satisfied. This completes the proof of Part (a).

The proof of (b) proceeds along similar lines. Here also we at first prove the result for the last distinct letter x and then generalize to previously occurring letters by induction. Let $p_1 < p_2 < \dots < p_l$ be the positions of x in the word w . Without loss of generality we assume that p_1 is

odd. The argument for the even case goes exactly along the same line as that of the odd case and hence shall be omitted. Note that $\forall p_1 < p < p_2$,

$$\begin{aligned}\pi(p) &= (-1)^{p-p_1} \pi(p_1) + c \pmod{n} \text{ and} \\ \pi(p_1 - 1) + \pi(p_1) &= \pi(p_2 - 1) + \pi(p_2) \pmod{n}.\end{aligned}$$

This implies

$$\pi(p_2) = (1 - (-1)^{p_2-p_1-1}) \pi(p_1) + c \pmod{n}$$

and more generally, if $\pi(p_m) = q_m \pi(p_1) + c \pmod{n}$, then

$$\pi(p_{m+1}) = (1 - (-1)^{p_{m+1}-p_m-1}) q_m \pi(p_1) + c \pmod{n}. \quad (6.14)$$

Now fix a position p_m . Define,

$$\begin{aligned}o_m &:= \text{Number of odd positions coming before the position } p_m \\ e_m &:= \text{Number of even positions coming before the position } p_m.\end{aligned}$$

Then application of (6.14) gives

$$q_m = \begin{cases} o_m + 1 - e_m & \text{if } p_m \text{ is odd} \\ e_m + 1 - o_m & \text{if } p_m \text{ is even.} \end{cases} \quad (6.15)$$

In particular $q_l = 0$ iff x occurs at even and odd positions equally often. The next steps are essentially similar to those in the proof of Theorem 3.1. The only difference is, here we have to consider even and odd positions instead of considering q and m positions. We omit the details. \square

Proof of Theorem 3.2. This proof is essentially based on the volume method introduced by Bryc, Dembo and Jiang (2006).

(a) First observe that a Toeplitz match implies a Symmetric Circulant match. Hence if a word w doesn't contribute in the Symmetric Circulant case, it will not contribute in the Toeplitz case. So from the proof of Theorem 3.1, we can restrict our attention to all words in \mathcal{W}_{2k} for which $q_x = m_x$ for every letter x . Note that vertices (i, j) match if and only if

$$s(i) - s(j) = 0, \text{ (} q \text{ match, say) or} \quad (6.16)$$

$$s(i) + s(j) = 0, \text{ (} m \text{ match, say)} \quad (6.17)$$

where $s(l)$ is as in (6.1). Let

$$v_i = \frac{\pi(i)}{n} \quad \text{and} \quad U_n := \left\{0, \frac{1}{n}, \dots, \frac{n-1}{n}\right\}.$$

Fix a word $w \in \mathcal{W}_{2k}$ and fix a sequence for q and m matches for each letter x such that $q_x = m_x$ for each x . We denote this sequence by \mathcal{I} . Denote the set of positions of the generating vertices as $S := \{p_0, p_1, \dots, p_{d(w)}\}$ and $v_S := \{v_i : i \in S\}$. Let

$$\Pi_{\mathcal{I}}^{**}(w) := \text{Set of circuits corresponding to a word } w \text{ and the match sequence } \mathcal{I}.$$

Now let us fix the values of $v_{p_0}, \dots, v_{p_{d(w)}}$. As we have fixed the sequences of q and m matches, we can find the values of each non-generating v_i by solving the set of linear equations:

$$\begin{aligned} v_{f(i)} - v_{f(i)-1} &= v_i - v_{i-1} \quad \text{for a } q \text{ match} \\ &\text{or} \\ v_{f(i)} - v_{f(i)-1} &= -(v_i - v_{i-1}) \quad \text{for an } m \text{ match.} \end{aligned} \tag{6.18}$$

Here $f(i) \in S$ is the position of the first occurrence of the letter corresponding to position i . So, $v_i = L_i^T(v_S)$ where L_i^T is a linear function corresponding to the i^{th} letter and the word w . As we have fixed $q_x = m_x$ for each letter x , from the proof of Theorem 3.1 (a) the circuit condition $v_{2k} = v_0$ automatically holds. Hence total number of circuits of this type equals $\#\Pi_{\mathcal{I}}^{**} = \#\{v_S : L_i^T(v_S) \in U_n \forall i\}$. So,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \#\Pi_{\mathcal{I}}^{**} \mathbb{E}[X_w] = \lim_{n \rightarrow \infty} \frac{1}{n^{d(w)} + 1} \#\Pi_{\mathcal{I}}^{**} \xi^{d(w)}.$$

This limit equals

$$\underbrace{\int_0^1 \cdots \int_0^1}_{d(w)+1 \text{ fold integral}} \mathbb{I}(0 \leq L_i^T(v_S) \leq 1, \forall i \notin S \cup \{2k\}) dv_S \xi^{d(w)} := p_{\mathcal{I}}(w)$$

and hence $p(w) = \sum_{\mathcal{I}} p_{\mathcal{I}}(w)$. This proves part (a) of the Theorem.

For part (b) one may note that any Hankel match implies a Reverse Circulant match. Hence we can restrict to words that contribute only in the the Reverse Circulant case. Rest of the proof is exactly similar to the Toeplitz part. We omit the details. \square

Proof of Remark 3.1. (a) First observe that if $x_n \sim \text{Bin}(k, p_n)$ such that $np_n \rightarrow \xi$, then

$$P[x_n = 1] = \binom{k}{1} p_n (1 - p_n)^{k-1} = O(n^{-1}) \text{ and } P[x_n > 1] = o(n^{-1}).$$

Let

$$p(w) = \lim_{n \rightarrow \infty} \frac{1}{n^{d(w)+1}} \#\Pi(w).$$

Then

$$\begin{aligned} \lim_{n \rightarrow \infty} \int x^h dF_0^{B_n}(x) &= \lim_{n \rightarrow \infty} \sum_{w:l(w)=h} p(w) n^{d(w)} \mathbb{E}[X_w] \\ &= \lim_{n \rightarrow \infty} \sum_{w:l(w)=h} p(w) \frac{\xi^{d(w)}}{P_n} \mathbb{E}[X_w]. \end{aligned}$$

Now write

$$\mathbb{E}[X_w] = \mathbb{E}[x_1^{l_1}] \cdots \mathbb{E}[x_{d(w)}^{l_{d(w)}}]$$

where l_i denotes the number of occurrences of the i^{th} letter in the word. Now

$$\mathbb{E}[x_1^j] = kp_n(1 - p_n)^{k-1} + \sum_{m=2}^k m^j \binom{k}{m} p_n^m (1 - p_n)^{(k-m)} = kp_n + o(p_n).$$

Hence the result follows by observing that

$$\mathbb{E}[X_w] = (kp_n)^{d(w)} + o(p_n^{d(w)}).$$

(b) Let A_n be a sequence of sparse patterned matrices. Then

$$\beta_{2k} = \lim_{n \rightarrow \infty} \mathbb{E} \left[\frac{1}{n} \text{Tr}[A_n]^{2k} \right] = \sum_{w \in \mathcal{W}_{2k}} p(w) \xi^{d(w)}.$$

As a result,

$$\frac{\beta_{2k}}{\xi^k} = \sum_{w \in \mathcal{W}_{2k}} p(w) \xi^{d(w)-k}.$$

Observe $d(w) \leq k$ as each letter in a contributing word comes even number of times. Hence

$$\lim_{\xi \rightarrow \infty} \frac{\beta_{2k}}{\xi^k} = \sum_{w \in \mathcal{W}_{2k} \text{ and } d(w)=k} p(w). \quad (6.19)$$

So only pair-matched words survive. It is a well known fact that the $(2k)$ th moment of the LSD in mean 0 variance 1 regime is exactly same as the R.S. of equation (6.19) (see Bose and Sen (2008) [8]). As a consequence, the result follows. \square

Proof of Theorem 3.3. Fix $M_n = A_{i_1,n} \dots A_{i_k,n}$. The moment trace formula (2.2) gives,

$$\phi_n(M_n) = \frac{1}{n} \mathbb{E} \left[\sum_{1 \leq i_1, i_2, \dots, i_h \leq n} A_{i_1,n}(L(j_1, j_2)) A_{i_2,n}(L(j_2, j_3)) \dots A_{i_h,n}(L(j_h, j_1)) \right].$$

For notational convenience we drop the index n . We define indexed matched circuits and indexed matched words subordinate to this monomial M in the following manner. A circuit π will be called indexed-matched if each L value is repeated only within the same index. Indexed-matched circuits π_1 and π_2 are said to be *equivalent* if and only if, $i_m = i_d$ and

$$\begin{aligned} A_{i_m,n}(L(\pi_1(m-1), \pi_1(m))) &= A_{i_d,n}(L(\pi_1(d-1), \pi_1(d))) \\ \Leftrightarrow A_{i_m,n}(L(\pi_2(m-1), \pi_2(m))) &= A_{i_d,n}(L(\pi_2(d-1), \pi_2(d))). \end{aligned} \quad (6.20)$$

Any equivalence class, subordinate to an indexed set induces a partition of $\{1, 2, \dots, k\}$ and each block of the partition is associated with an index. Any such class can be expressed as an indexed-word w where letters appear in alphabetic order of their first occurrence and with a subscript to distinguish the index. For example the partition $(\{\{1, 3, \}; 1\}, \{\{2, 4\}; 2\}, \{\{5, 7\}; 1\}, \{\{6, 8\}; 3\})$ is identified with the word $a_1 b_2 a_1 b_2 c_1 d_3 c_1 d_3$. A typical position in an indexed word would be referred to as $w_{i_j}[j]$. Let

$$H_M = \{w : w \text{ is an indexed word subordinate to the monomial } M\}.$$

Let the class of all indexed circuits corresponding to an index-matched word w be denoted by

$$\Pi_{I_M}(w) = \{\pi : w_{i_m}[m] = w_{i_d}[d] \Leftrightarrow A_{i_m,n}(L(\pi(m-1), \pi(m))) = A_{i_d,n}(L(\pi(d-1), \pi(d)))\}.$$

Now recall that

$$\begin{aligned}
\phi_n(M_n) &= \frac{1}{n} \sum_{w \in H_M} [\#\Pi_{I_M}(w)] EX_w \\
&= \sum_{w \in H_M} \frac{1}{n^{d(w)+1}} [\#\Pi_{I_M}(w)] (np_n)^{d(w)},
\end{aligned}$$

where EX_w denotes the common expectation of

$$E[A_{i_1,n}(L(\pi(0), \pi(1))) \dots A_{i_k,n}(L(\pi(k-1), \pi(k)))] \quad \forall \pi \in \Pi_{I_M}(w).$$

As $A_{i_1,n}, \dots, A_{i_k,n}$ have the same pattern, we can drop the indices from the words and look at the corresponding non-indexed words. The contribution of an indexed-word $P_M(w)$ will be exactly as its non-indexed counterpart, $p(w)$ which we have proved to exist for all the four matrices. Now, the words subordinate to a particular index set M is always finite being a collection of partitions of $\{1, 2, \dots, k\}$. Hence, $\lim_{n \rightarrow \infty} \phi_n(M_n)$ exists. This proves the theorem. \square

We now prove the claim about independent Symmetric Circulants. Let us consider a monomial $M_n = SC_{i_1,n} \dots SC_{i_k,n}$. Also, let l denote the total number of distinct indices in M_n and m_j be the number of times the j th index is repeated. It is enough to prove that

$$\lim_{n \rightarrow \infty} \phi_n(M_n) = \phi(SC_{i_1} \dots SC_{i_k}) = \prod_{j=1}^l \phi(SC_j^{m_j}).$$

We have proved earlier that $\lim_{n \rightarrow \infty} \phi_n(M_n) = \sum_{w \in H_M} p(w) \xi^{d(w)}$. Fix $w \in H_M$. From Theorem 3.1,

$$p(w) = \begin{cases} \prod_{i=1}^d \binom{2k_i-1}{k_i}, & \text{if } w \text{ has } d \text{ distinct letters, appearing } 2k_i \text{ times, } 1 \leq i \leq d \\ 0, & \text{otherwise.} \end{cases}$$

From the above expression, if a word contributes, all the letters have to come even number of times. Hence m_j 's are even for each j . Let d_j denotes the number of distinct letters subordinate to index j and $2k_{i,j}$ be the number of times the i th letter occurs in the j th index. Then

$$m_j = \sum_{i=1}^{d_j} 2k_{i,j}, \quad d = \sum_{j=1}^l d_j \quad \text{and} \quad p(w) = \prod_{j=1}^l \prod_{i=1}^{d_j} \binom{2k_{i,j}-1}{k_{i,j}}.$$

Hence

$$\lim_{n \rightarrow \infty} \phi_n(M_n) = \sum_{w \in H_M} \prod_{j=1}^l \left(\prod_{i=1}^{d_j} \binom{2k_{i,j}-1}{k_{i,j}} \right) \xi^{d_j}.$$

Now fixing j , the choices for d_j 's are independent of all the other indices. Hence,

$$\lim_{n \rightarrow \infty} \phi_n(M_n) = \prod_{j=1}^l \left(\sum_{w \in \mathcal{W}_{\frac{m_j}{2}}} p(w) \xi^{d(w)} \right) = \prod_{j=1}^l \phi(SC_j^{m_j}). \quad (6.21)$$

\square

Proof of Theorem 4.1. We need the following Lemma. We omit its proof which is based on calculating the finite dimensional joint distribution.

Lemma 6.1. Let $\{x_i^{(n)}, 1 \leq i \leq n-1\}$ be i.i.d. Bernoulli (p_n) such that $np_n \rightarrow \xi$. Let

$$\begin{aligned} Q_n &= \sum_{i=1}^{n-1} x_i^{(n)} \\ U_i^{(n)} &= \begin{cases} \frac{i^{\text{th}} \text{ largest index such that } x_i^{(n)}=1}{n}, & \text{if } Q_n \geq i \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (6.22)$$

Then $U^{(n)} = (Q_n, U_1^{(n)}, \dots, U_n^{(n)}, 0, \dots, 0, \dots) \xrightarrow{\mathcal{D}} (Q, U_1, U_2, \dots) = V$ say where

$$Q \sim \text{Poisson}(\xi), \text{ and } V|Q = k \sim (k, V_{(1)}, \dots, V_{(k)}, \dots),$$

$(V_{(1)}, \dots, V_{(k)})$ are distributed as order statistics of an i.i.d. sample of size k from the uniform $(0, 1)$ distribution.

The proof of the next Lemma follows from Theorem 11.7.1. of Dudley (2002) [11].

Lemma 6.2. There exists a suitable probability space $(\Omega, \mathcal{F}, \mathbb{P})$ in which all the relevant random variables are defined and

$$U^{(n)} = (Q_n, U_1^{(n)}, \dots, U_n^{(n)}, 0, \dots, 0, \dots) \xrightarrow{a.s.} (Q, U_1, U_2, \dots). \quad (6.23)$$

The proof of the next Lemma is available in Knight (1991) [13].

Lemma 6.3. Let $(a_{n1}, \dots, a_{nk}), n \geq 1$ be a vector sequence of real numbers where $(\frac{a_{n1}}{n}, \dots, \frac{a_{nk}}{n}) \rightarrow (u_1, u_2, \dots, u_k) \in (0, 1)^k$ such that for integers $v_1, v_2, \dots, v_k, \sum_{i=1}^k v_i u_i = 0 \pmod{1}$ iff $v_1 = v_2 = \dots = v_k = 0$. Then the measure μ_n which puts mass $1/n$ at each $(\{\frac{j a_{n1}}{n}\}, \dots, \{\frac{j a_{nk}}{n}\}), 1 \leq j \leq n$, converges weakly to the continuous uniform measure on $(0, 1)^k$.

Assume that all random variables, including $\{x_i^n\}$, are defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which Lemma 6.2 holds. We first prove (b). Since each eigenvalue has mass $1/n$, we can ignore the eigenvalues λ_0 and $\lambda_{n/2}$ (where n is even). Moreover, since $\lambda_k(\omega) = -\lambda_{n-k}(\omega), 1 \leq k \leq \lfloor \frac{n-1}{2} \rfloor$, it is enough to consider only the empirical measure on $\{\lambda_1(\omega), \dots, \lambda_f(\omega)\}$ where $f = \lfloor \frac{n-1}{2} \rfloor$. Note that

$$\sum_{i=0}^{n-1} x_i^{(n)} \cos\left(\frac{2\pi i j}{n}\right) = \sum I\{U_i^{(n)} \neq 0\} \cos(2\pi\{U_i^{(n)} j\})$$

where $\{U_i^{(n)}\}$'s are as in (6.22). Since Q_n, Q are discrete, for any given ω ,

$$Q(\omega) = k \Rightarrow Q_n(\omega) = k \quad \forall n \geq N(\omega), \text{ say.} \quad (6.24)$$

Hence for all $n \geq N(\omega)$,

$$\left(\{jU_1^{(n)}(\omega)\}, \dots, \{jU_n^{(n)}(\omega)\}\right) = \left(\{jU_1^{(n)}(\omega)\}, \dots, \{jU_k^{(n)}(\omega)\}, 0, \dots, 0\right).$$

By Lemma 6.3, the empirical measure of

$$\left(\{jU_1^{(n)}(\omega)\}, \dots, \{jU_k^{(n)}(\omega)\}\right), 1 \leq j \leq \lfloor \frac{n-1}{2} \rfloor \xrightarrow{\mathcal{D}} (U_1, \dots, U_k)$$

since for integers v_1, \dots, v_k , for almost all ω such that $Q(\omega) = k$, $\sum_{i=1}^k v_i \cdot U_i(\omega) = 0 \pmod{1} \Rightarrow v_1 = \dots = v_k = 0$. Hence conditional on $Q(\omega) = k$, this empirical measure converges to the measure induced by $\left(\sum_{i=1}^k \cos(2\pi U_i)\right)^2 + \left(\sum_{i=1}^k \sin(2\pi U_i)\right)^2$ and the LSD is the symmetrized version of this. This completes the proof of (b). The proof of (a) is essentially the same after using the formula for eigenvalues.

(c) Suppose we can establish the (M1) condition. Recall Equation (6.24). Also, conditioning on $Q(\omega) = k$, both the Reverse Circulant and the Symmetric Circulant have bounded LSD. Now, it is easy to observe that if the matrices are constructed with same input random variables, the moments of the ESD of Toeplitz and Hankel are bounded by the moments of the ESD of Reverse Circulant and Symmetric Circulant respectively. As a consequence, Carleman's Condition (C) will hold trivially. Thus it is enough to establish the (M1) condition.

First consider the Toeplitz matrix T_n . It has n upper diagonals labelled $\{0, 1, \dots, (n-1)\}$ starting from the main diagonal. Let $T_n(\omega)$ be the realization of the random Toeplitz matrix for a particular ω . Then $T_n(\omega)$ has $Q_n(\omega)$ diagonals each only with 1's and all the other diagonals are 0. Let $U_1^{(n)}, \dots, U_{Q_n}^{(n)}$ be as defined in (6.22).

For example, suppose $n = 4$, $Q_n(\omega) = 2$, $U_1^{(n)}(\omega) = \frac{1}{4}$ and $U_2^{(n)}(\omega) = \frac{2}{4}$. Then $T_n(\omega)$ equals

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

We prove that for all ω outside a \mathbb{P} -null set, $\frac{1}{n} \text{Tr}(T_n^k(\omega))$ converges for all $k \in \mathbb{N}$. Observe that

$$\frac{1}{n} \text{Tr}(T_n^k(\omega)) = \frac{1}{n} \sum_{i_0, i_1, \dots, i_{k-1}} T_n(\omega)(i_0, i_1) T_n(\omega)(i_1, i_2) \dots T_n(\omega)(i_{k-1}, i_0) \text{ where}$$

$$T_n(\omega)(i_0, i_1) \dots T_n(\omega)(i_{k-1}, i_0) = \begin{cases} 1 & \text{if } |i_{j+1} - i_j| = nU_m^{(n)}(\omega) \quad \forall j \text{ for some } m \\ 0 & \text{otherwise.} \end{cases}$$

So this is equivalent to studying a walk starting from i_0 and ending at i_0 in $2k$ steps, having possible jumps of length $\pm nU_1^{(n)}, \pm nU_2^{(n)}, \dots, \pm nU_{Q_n}^{(n)}$ and staying within the range $(0, n-1)$. We scale down the whole process by n so that the jump values become $\pm U_1^{(n)}, \pm U_2^{(n)}, \dots, \pm U_{Q_n}^{(n)}$ and the range is $(0, 1 - \frac{1}{n})$. Then

$$\frac{1}{n} \text{Tr}(T_n^k(\omega)) = \frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{P}_{\frac{i_0}{n}}^n(\omega) \text{ where}$$

$$\mathcal{P}_{\frac{i_0}{n}}^n(\omega) := \text{Collection paths starting from } \frac{i_0}{n} \text{ satisfying the above criteria.}$$

We know that $Q_n(\omega) \xrightarrow{a.s.} Q(\omega) = S$ (say). Hence for almost all ω , $\exists N(\omega)$ such that $\forall n \geq N(\omega)$, $Q_n(\omega) = S$. Now for almost all ω $(U_1^{(n)}(\omega), \dots, U_S^{(n)}(\omega)) \rightarrow (U_1(\omega), \dots, U_S(\omega))$. Let us consider a walk on the $X - Y$ plane with possible jump lengths $(\pm U_1(\omega), \dots, \pm U_S(\omega))$. As we have to deal with several classes of paths of this walk, we introduce few more notations. For any $s \in (0, 1)$

$\mathcal{P}_s(\omega)$:= collection of all paths of length k returning to the starting point s
and which do not cross either of the boundaries $Y = 0$ and $Y = 1$.

$C_s(\omega)$:= collection of all paths of length k returning to the starting point s .

$\mathcal{D}_s(\omega)$:= collection of all paths of length k returning to the starting point s
which touch or cross at least one of the boundaries $Y = 0$ and $Y = 1$ at least once.

Similarly $C_s^n(\omega)$ and $\mathcal{D}_s^n(\omega)$ denote the same collection of paths when the jump lengths are taken to be $(\pm U_1^{(n)}(\omega), \dots, \pm U_S^{(n)}(\omega))$ instead. Note that

$$\#\mathcal{P}_s(\omega) = \#C_s(\omega) - \#\mathcal{D}_s(\omega) \quad \text{and} \quad \#\mathcal{P}_s^n(\omega) = \#C_s^n(\omega) - \#\mathcal{D}_s^n(\omega).$$

It is now enough to prove the convergence of $\frac{1}{n} \sum_{i_0=0}^{n-1} \#C_{\frac{i_0}{n}}^n(\omega)$ and $\frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{D}_{\frac{i_0}{n}}^n(\omega)$.

Step 1. *Convergence of $\frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{D}_{\frac{i_0}{n}}^n(\omega)$.* Note that except for finitely many points,

$$0 \neq s + H_1 U_1(\omega) + \dots + H_S U_S(\omega) \neq 1 \quad \forall H_1, \dots, H_S \in \{-k, -(k-1), \dots, k\}. \quad (6.25)$$

We denote the set of points satisfying (6.25) by G . Consider $\cup_{x \in G} \mathcal{D}_x(\omega)$. Two paths $q, q' \in \cup_{x \in G} \mathcal{D}_x$ are said to be \sim' equivalent if

1. The jump values at each time instant are exactly same for q and q' and
2. q and q' cross any of the boundaries $Y = 0$ or $Y = 1$ at the same instants.

It is easy to check that \sim' is an equivalence relation. Next, for $x, y \in G$, say that $x \sim y$ if and only if for every path $q \in \mathcal{D}_x(\omega)$, the path q' starting from y with same sequence of jumps is \sim' equivalent to q and vice versa. It is easy to observe that this is an equivalence relation and divides G into finitely many disjoint intervals, say I_1, \dots, I_I and points. Hence

$$[0, 1] = G \cup \{\text{finitely many points}\} \quad \text{where} \quad G = I_1 \cup \dots \cup I_I.$$

Observe that $\#\mathcal{D}_x(\omega)$ and $\#\mathcal{P}_x(\omega)$ are constants over any I_i . Let

$$\#\mathcal{D}_{I_i}(\omega) := \#\mathcal{D}_x(\omega) \quad \text{and} \quad \#\mathcal{P}_{I_i}(\omega) := \#\mathcal{P}_x(\omega) \quad \text{such that} \quad x \in I_i. \quad (6.26)$$

Let \mathcal{L} denote the Lebesgue measure. We claim that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{D}_{\frac{i_0}{n}}^n(\omega) = \sum_{i=1}^I \#\mathcal{D}_{I_i} \mathcal{L}(I_i),$$

Now,

$$\frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{D}_{\frac{i_0}{n}}^n(\omega) = \int f^{(n)}(x, \omega) dx \quad \text{where}$$

$$f^{(n)}(x, \omega) = \#\mathcal{D}_{\frac{i}{n}}^n(\omega) \quad \text{if} \quad x \in \left[\frac{i}{n}, \frac{i+1}{n}\right).$$

Fix $x \in \bigcup_{i=1}^I I_i$. We prove that for any such x , $f^{(n)}(x, \omega) \rightarrow \#\mathcal{D}_x(\omega)$. For $x \in I_i$, let

$$\epsilon_x(\omega) = \min_{q \in \mathcal{D}_x(\omega)} \{\text{minimum distance of the path } q \text{ from the boundaries } Y = 0 \text{ and } Y = 1\} > 0.$$

As $(U_1^{(n)}(\omega), \dots, U_S^{(n)}(\omega)) \rightarrow (U_1(\omega), \dots, U_S(\omega))$, choose large enough N_0 such that $\forall n > N_0$,

$$\max_{i \in \{1, 2, \dots, S\}} |U_i(\omega) - U_i^{(n)}(\omega)| < \frac{\epsilon_x}{4kS}.$$

For these n 's, take any path $q \in \mathcal{D}_x(\omega)$ and construct another path q' by replacing jumps of length $U_j(\omega)$ by $U_j^{(n)}(\omega)$. Then whenever q crosses one of the boundaries, q' also crosses the same boundary at the same instant. Moreover, the minimum distance of the path q' from the boundaries is greater than $\frac{3\epsilon_x}{4}$. Hence, the map $q \mapsto q'$ forms a bijection of paths. Thus,

$$\#\mathcal{D}_x^{(n)}(\omega) = \#\mathcal{D}_x(\omega).$$

On the other hand taking $n \geq \frac{8}{3\epsilon_x}$ and repeating the previous argument, we get

$$\#\mathcal{D}_x^{(n)}(\omega) = \#\mathcal{D}_{\frac{i}{n}}^{(n)}(\omega) \text{ when } \frac{i}{n} \leq x < \frac{i+1}{n}.$$

As a result, we get $f^{(n)}(x, \omega) = \#\mathcal{D}_x(\omega)$ eventually. As there are at most $(2S)$ possible jumps in any step, $|f^{(n)}(x, \omega)|$ is uniformly bounded by $(2S)^k$. So we can apply dominated convergence theorem to conclude that

$$\int f^{(n)}(x, \omega) dx = \frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{D}_{\frac{i_0}{n}}^{(n)}(\omega) \rightarrow \sum_{i=1}^I \#\mathcal{D}_{I_i} \mathcal{L}(I_i) = \int \#\mathcal{D}_x(\omega) dx. \quad (6.27)$$

Step 2. *Convergence of $\frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{C}_{\frac{i_0}{n}}^{(n)}(\omega)$.* Note that $\#\mathcal{C}_x^{(n)}(\omega)$ is invariant with respect to the starting point and the jump lengths if the jump lengths are linearly independent in the sense that

$$v_1 U_1^{(n)}(\omega) + \dots + v_S U_S^{(n)}(\omega) = 0 \Leftrightarrow v_1 = \dots = v_S = 0 \quad \forall (v_1, \dots, v_S) \in \{-k, -k+1, \dots, k\}^S. \quad (6.28)$$

We denote this value by $C(\omega)$. Since $(U_1^{(n)}(\omega), \dots, U_S^{(n)}(\omega))$ converge to $(U_1(\omega), \dots, U_S(\omega))$, which are linearly independent almost surely we can assume (6.28) almost surely. As a consequence, $\#\mathcal{C}_x^{(n)}(\omega) = \#\mathcal{C}_x(\omega)$ eventually. Using dominated convergence again we obtain

$$\frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{C}_{\frac{i_0}{n}}^{(n)}(\omega) \rightarrow C(\omega). \quad (6.29)$$

Combining (6.27) and (6.29) completes the proof for T_n . The above argument also proves that $\frac{1}{n} \text{Tr}(T^{2k+1}(\omega)) \rightarrow 0$ almost surely as $\#\mathcal{C}_x(\omega) = 0$ for all x in this case.

We now consider the Hankel case. In this case, there are exactly $(2n - 1)$ cross diagonals. By arguing as in the proof of (b) we get that $U^{(n)} \xrightarrow{\mathcal{D}} (Q, U_1, U_2 \dots) = V$ say where $Q \sim \text{Poisson}(2\xi)$, and $V|Q = S \sim (S, V_{(1)}, \dots, V_{(S)}, \dots)$, and $(V_{(1)}, \dots, V_{(S)})$ are distributed as the S order statistics of an i.i.d. sample of size S from the continuous uniform $(0, 1)$ distribution. Again we assume this convergence to happen almost surely by Lemma 6.2.

Now fix $\omega \in \Omega$, and suppose $H_n(\cdot, \omega)$ has $Q_n(\omega)$ non-zero cross diagonal at positions, $nU_1^{(n)}(\omega), nU_2^{(n)}(\omega), \dots, nU_{Q_n(\omega)}^{(n)}(\omega)$. Then arguing as in the Toeplitz case gives us:

$$\text{Tr}(H_n^k(\omega)) = \frac{1}{n} \sum_{i_0, i_1, \dots, i_{k-1}} H_n(\omega)(i_0, i_1) H_n(\omega)(i_1, i_2) \dots H_n(\omega)(i_{k-1}, i_0) \text{ where}$$

$$H_n(\omega)(i_0, i_1) \dots H_n(\omega)(i_{k-1}, i_0) = \begin{cases} 1 & \text{if } i_{j+1} + i_j = nU_m^{(n)} \quad \forall j \text{ for some } m \\ 0 & \text{otherwise.} \end{cases}$$

Unlike the Toeplitz case, here $i_{j+1} + i_j = nU_m^{(n)}$. Hence we cannot directly apply the earlier random walk argument. We modify that argument slightly. Let i'_0, \dots, i'_{k-1} be defined as:

$$i'_j = \begin{cases} i_j & \text{if } j \text{ is even} \\ -i_j & \text{if } j \text{ is odd.} \end{cases}$$

Now we can see that i'_0, \dots, i'_{k-1} are the points on a path of length k of the random walk with possible jumps size $\pm nU_1^{(n)}, \pm nU_2^{(n)}, \dots, \pm nU_{Q_n}^{(n)}$ with the following restrictions.

1. The path returns to the starting point after k steps.
2. The jump at every odd (even) step is < 0 (> 0).
3. The position of the walk at every odd (even) step is below (above) the X axis.
4. The walk is bounded between -1 and 1 .

Now choose $s \in (0, 1)$ such that

$$s + H_1 U_1(\omega) + \dots + H_S U_S(\omega) \notin \{-1, 0, 1\}, \quad \forall H_1, \dots, H_S \in \{-k, -(k-1), \dots, k\}.$$

$\mathcal{P}_s(\omega)$:= collection of all paths of length k that return to the starting point s , do not cross $Y = -1$, $Y = 1$ and stays below (above) the X axis at all odd (even) time points.

$\mathcal{C}_s(\omega)$:= collection of all paths of length k that return to the starting point s and are below (above) the X axis at all odd (even) time points,

$\mathcal{D}_s(\omega)$:= collection of all paths length k that return to the starting point s , cross one of the boundaries at least once and are below (above) X axis at all odd (even) time points.

So, $\#\mathcal{P}_s(\omega) = \#\mathcal{C}_s(\omega) - \#\mathcal{D}_s(\omega)$. Unlike the Toeplitz case, now $\#\mathcal{C}_s(\omega)$ varies with s . Using the earlier arguments, it follows that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{C}_{\frac{i_0}{n}}^n(\omega) \text{ and } \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{D}_{\frac{i_0}{n}}^n(\omega)$$

exist. This essentially proves that $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i_0=0}^{n-1} \#\mathcal{P}_{\frac{i_0}{n}}^n(\omega)$ exists and the proof for the Hankel case is complete and hence that of Theorem 4.1 too. \square

Proof of Theorem 5.2. We need the following two results. We omit the proof of the second one.

Lemma 6.4. *Suppose (6.23) holds. Let A_n be any one of the four random sparse patterned matrices. Then $\liminf_{n \rightarrow \infty} |\text{sp}_{A_n}(\omega)| \geq \text{supp}(F^A(\cdot, \omega))$ a. s. where $F^A(\cdot, \omega)$ is the weak in distribution limit of $F^{A_n}(\cdot, \omega)$.*

Proof. We know that $\beta_{2k}(A_n)(\omega) = \frac{1}{n} \sum_{i=1}^n (\lambda_i(\omega))^{2k} \leq (\text{sp}_{A_n}(\omega))^{2k}$. Hence,

$$\lim_{n \rightarrow \infty} \beta_{2k}(A_n)(\omega) \stackrel{\text{a.s.}}{=} \beta_{2k}(\omega) \leq \liminf_{n \rightarrow \infty} (\text{sp}_{A_n}(\omega))^{2k}.$$

As a result,

$$[\beta_{2k}(\omega)]^{\frac{1}{2k}} \leq \liminf_{n \rightarrow \infty} \text{sp}_{A_n}(\omega).$$

Hence

$$\text{supp} F^A(\cdot, \omega) = \lim_{k \rightarrow \infty} [\beta_{2k}(\omega)]^{\frac{1}{2k}} \leq \liminf_{n \rightarrow \infty} \text{sp}_{A_n}(\omega).$$

□

Claim 6.1. *Let $f_n(x_1, \dots, x_n) = \sqrt{\left(\sum_{i=1}^n \cos(2\pi x_i)\right)^2 + \left(\sum_{i=1}^n \sin(2\pi x_i)\right)^2}$. Then $\sup_{(x_1, x_2, \dots, x_n) \in [0, 1]^n} (f_n(x_1, \dots, x_n)) = n$.*

As before assume that all variables including $\{x_i^n\}$ are defined on a probability space where (6.23) holds. Theorem 4.1 (a), implies that $F^A(\cdot, \omega)$ depends only on $Q(\omega)$. Let $Q_n = \sum_{i=0}^{n-1} x_i^{(n)}$. Then $Q_n \xrightarrow{\text{a.s.}} Q \sim \text{Poisson}(\xi)$. Recall the eigenvalue formula (4.1). Applying Claim 6.1, we get

$$\mathbb{P}[\text{sp}_{RC_n} \leq S \mid Q_n = S] = 1 \text{ and hence } \mathbb{P}[\text{sp}_{RC_n} \leq Q_n] = 1.$$

On the other hand, since, $Q_n \xrightarrow{\text{a.s.}} Q$, $\limsup_{n \rightarrow \infty} \text{sp}_{RC_n}(\omega) \leq Q(\omega)$ a.s. From Lemma 6.4,

$$\liminf_{n \rightarrow \infty} |\text{sp}_{RC_n}(\omega)| \geq \text{supp}(F^{RC}(\cdot, \omega)) = Q(\omega) \text{ a.s.}$$

Thus, $\lim_{n \rightarrow \infty} \text{sp}_{RC_n}(\omega) = Q(\omega)$ a.s. So, $\text{sp}_{RC_n} \xrightarrow{\mathcal{D}} Q$, proving (a). Proof of (b) is similar.

Now we come to the more difficult proofs of (c) and (d).

Lemma 6.5. *Let $U_1, \dots, U_S \sim U(0, 1)$ be i.i.d. uniform random variables and $a \in [0, 1]$ be any number. Let $B_n := \{-[n^{\frac{1}{4S}}], \dots, [n^{\frac{1}{4S}}]\} \setminus \{0\}$. Then the following set is \mathbb{P} null.*

$$A := \left\{ \omega \mid \sum_{i=1}^S v_i U_i(\omega) \in \left[a - \frac{1}{n}, a + \frac{1}{n} \right] \text{ for some } v_1, \dots, v_S \in B_n, \text{ for infinitely many } n \right\}.$$

Proof. Let us define,

$$A_n := \bigcup_{v_1, \dots, v_S} B_{\{v_1, \dots, v_S\}, n} \text{ where}$$

$$B_{\{v_1, \dots, v_S\}, n} := \begin{cases} \left\{ \omega \mid \sum_{i=1}^S v_i U_i(\omega) \in \left[a - \frac{1}{n}, a + \frac{1}{n} \right] \right\} & \text{for } \max_S \{v_i^{4S}\} \leq n, \\ \emptyset & \text{otherwise.} \end{cases}$$

We have to prove, $\mathbb{P} \left[\bigcap_{k=1}^{\infty} \bigcup_{n \geq k} A_n \right] = 0$. We prove this in the following two steps.

Step 1. At first, we prove

$$\mathbb{P}[B_{\{v_1, \dots, v_S\}, n}] \leq \frac{6}{n} \quad \forall v_1, \dots, v_S. \quad (6.30)$$

To prove this, we fix the values of $U_1(\omega), \dots, U_{S-1}(\omega)$, and prove

$$\mathbb{P}[B_{\{v_1, \dots, v_S\}, n} \mid U_1(\omega), \dots, U_{S-1}(\omega)] \leq \frac{6}{n} \quad \mathbb{P} \text{ almost surely.}$$

Then, taking the expectation of the conditional probability gives us (6.30). Now, fix the values of v_1, v_2, \dots, v_S and $U_1(\omega), \dots, U_{S-1}(\omega)$ and calculate

$$\mathbb{P}\left[v_S U_S \in \left[L + a - \sum_{i=1}^{S-1} v_i U_i(\omega) - \frac{1}{n}, L + a - \sum_{i=1}^{S-1} v_i U_i(\omega) + \frac{1}{n}\right]\right] \text{ where } L \text{ is an integer.}$$

To get a feasible value of U_S , we need

$$L \in \left[-a + \sum_{i=1}^{S-1} v_i U_i(\omega) - \frac{1}{n}, -a + \sum_{i=1}^{S-1} v_i U_i(\omega) + v_S + \frac{1}{n}\right].$$

It is easy to note that there are at most $v_S + 2$ choices for L . Fixing L , we get:

$$U_S \in \frac{1}{v_S} \left[L + a - \sum_{i=1}^{S-1} v_i U_i(\omega) - \frac{1}{n}, L + a - \sum_{i=1}^{S-1} v_i U_i(\omega) + \frac{1}{n}\right].$$

The length of this interval is $\frac{2}{nv_S}$. Hence $\mathbb{P}[B_{\{v_1, \dots, v_S\}, n}] \leq \frac{2(v_S+2)}{nv_S} \leq \frac{6}{n}$.

Step 2. For fixed v_1, v_2, \dots, v_S , $B_{\{v_1, \dots, v_S\}, n}$ are decreasing for $n \geq \max_S \{v_S^{4S}\}$ and $B_{\{v_1, \dots, v_S\}, n} = \emptyset$ for $n \leq \max_S \{v_S^{4S}\}$. In particular, we can write, $\bigcup_{n=1}^{\infty} B_{\{v_1, \dots, v_S\}, n} = B_{\{v_1, \dots, v_S\}, \max_S \{v_S^{4S}\}}$. Now,

$$\begin{aligned} \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} A_n &= \bigcap_{k=1}^{\infty} \bigcup_{n \geq k} \bigcup_{v_1, \dots, v_S} B_{\{v_1, \dots, v_S\}, n} \\ &= \bigcap_{k=1}^{\infty} \bigcup_{v_1, \dots, v_S} \bigcup_{n \geq \max\{k, \max_i \{v_i^{4S}\}\}} B_{\{v_1, \dots, v_S\}, n} \\ &= \bigcap_{k=1}^{\infty} \bigcup_{v_1, \dots, v_S} B_{\{v_1, \dots, v_S\}, \max\{k, \max_i \{v_i^{4S}\}\}}. \end{aligned}$$

Enough to prove that $\mathbb{P}\left[\bigcup_{v_1, \dots, v_S} B_{\{v_1, \dots, v_S\}, \max\{k, \max_i \{v_i^{4S}\}\}}\right] \rightarrow 0$ as $k \rightarrow \infty$. Now,

$$\bigcup_{v_1, \dots, v_S} B_{\{v_1, \dots, v_S\}, \max\{k, \max_i \{v_i^{4S}\}\}} = C_k \bigcup D_k,$$

where

$$C_k = \bigcup_{\max_i \{v_i^{4S}\} \leq k} B_{\{v_1, \dots, v_S\}, k} \quad \text{and} \quad D_k = \bigcup_{\max_i \{v_i^{4S}\} > k} B_{\{v_1, \dots, v_S\}, \max_i \{v_i^{4S}\}}.$$

Now, $\mathbb{P}[C_k] \leq 6 \frac{(k^{\frac{1}{4S}})^S}{k} = 6k^{-\frac{3}{4}} \rightarrow 0$. Also,

$$\mathbb{P}[D_k] \leq \sum_{v_1, v_2, \dots, v_S \text{ and } \max_i \{v_i^{4S}\} \geq k} \frac{6}{\max_i \{v_i^{4S}\}} \leq 6 \left[\sum_{v_1, v_2, \dots, v_S \text{ and } \max_i \{v_i^{4S}\} \geq k} \prod_{i=1}^S \frac{1}{v_i^4} \right].$$

We know, $\sum_{(v_1, \dots, v_S) \in \mathbb{N}^S} \prod_{i=1}^S \frac{1}{v_i^4} < \infty$. As a result,

$$\lim_{k \rightarrow \infty} \mathbb{P}(D_k) = 6 \lim_{k \rightarrow \infty} \left| \sum_{(v_1, v_2, \dots, v_S) \in \mathbb{N}^S} \prod_{i=1}^S \frac{1}{v_i^4} - \sum_{(v_1, v_2, \dots, v_S) \in \{1, \dots, [k^{\frac{1}{4S}}]\}^S} \prod_{i=1}^S \frac{1}{v_i^4} \right| = 0.$$

This completes the proof of the Lemma.

We need the following generalization of Lemma 6.2.

Lemma 6.6. *Suppose $(Q_n, U^{(n)}) = (Q_n, U_1^{(n)}, \dots, U_n^{(n)}, 0, \dots)$, where Q_n is $\text{Bin}(n, p_n)$ and $(U_1^{(n)}, \dots, U_S^{(n)} | Q_n = S)$ is distributed as SRSWOR of size S from $\{0, \frac{1}{n}, \dots, \frac{n-1}{n}\}$. Then, there exists $(\hat{Q}_n, \hat{U}^{(n)}) = (\hat{Q}_n, \hat{U}_1^{(n)}, \dots)$ and $(Q, U) = (Q, U_1, \dots)$ on $([0, 1]^\infty, \mathcal{B}([0, 1]^\infty))$ such that, $Q \sim \text{Poisson}(\xi)$, $(Q, U | Q = S) \sim (S, W_1, \dots, W_S, 0, 0 \dots)$, where $\{W_i\}$ are i.i.d. $U(0, 1)$ and*

(a) $(\hat{Q}_n, \hat{U}^{(n)}) \xrightarrow{a.s.} (Q, U)$,

(b) $(\hat{Q}_n, \hat{U}^{(n)}) \stackrel{\mathcal{D}}{=} (Q_n, U^{(n)})$,

(c) For some c , $|\hat{U}_k^{(n)}(\omega) - U_k(\omega)| \leq \frac{cQ(\omega)}{n}$ for all n and k .

Proof. Consider the probability space $([0, 1]^\infty, \mathcal{B}([0, 1]^\infty), \mathcal{L}^\infty)$, where \mathcal{L}^∞ denotes the Lebesgue measure. At first, we define $\hat{Q}_n(\omega)$. Then given $\hat{Q}_n(\omega)$ we define $\hat{U}_1^{(n)}(\omega)$ and so on. Fix $\omega = (\omega_0, \omega_1, \dots) \in [0, 1]^\infty$. Let

$$\hat{Q}_n(\omega) = g \text{ for } \omega_0 \in [\mathbb{P}[Q_n \leq g], \mathbb{P}[Q_n \leq g + 1]).$$

If $\hat{Q}_n(\omega) = 0$, define $\hat{U}_k^{(n)}(\omega) = 0$ for all k . If $\hat{Q}_n(\omega) > 0$, let,

$$H_n := \{0, \frac{1}{n}, \dots, \frac{n-1}{n}\} \text{ and,}$$

$$H_n^{-B_n}(i) := \{i \text{ th minimum value of the set } H_n \setminus B_n\}, \quad B_n \subset H_n.$$

Define

$$\hat{U}_1^{(n)}(\omega) := \frac{j}{n} \text{ for } \omega_1 \in \left[\frac{j}{n}, \frac{j+1}{n} \right) \quad \forall \frac{j}{n} \in H_n$$

and for $1 < i \leq \hat{Q}_n(\omega)$, we define

$$\hat{U}_i^{(n)}(\omega) := \frac{H_n^{-B_{n,i}}(j)}{n} \text{ for } \omega_i \in \left[\frac{j-1}{n - \#B_{n,i}}, \frac{j}{n - \#B_{n,i}} \right) \quad 1 \leq j \leq n - \#B_{n,i}.$$

where $B_{n,i} := \{U_1^{(n)}(\omega), \dots, U_{i-1}^{(n)}(\omega)\}$. We know for SRSWOR, if W_i denotes the outcome of the i th draw then

$$\mathbb{P}[W_i = H_n^{-B_{n,i}}(j) \mid W_1, \dots, W_{i-1}] = \frac{1}{n - \#B_{n,i}} \quad \forall j \in H_n \setminus B_{n,i},$$

where $B_{n,i} := \{W_1, \dots, W_{i-1}\}$. Hence,

$$\hat{Q}_n \stackrel{\mathcal{D}}{=} Q_n, \quad \hat{U}_1^{(n)} | \hat{Q}_n \stackrel{\mathcal{D}}{=} U_1^{(n)} | Q_n \quad \text{and} \quad \hat{U}_i^{(n)} | (\hat{Q}_n, \hat{U}_1^{(n)}, \dots, \hat{U}_{i-1}^{(n)}) \stackrel{\mathcal{D}}{=} U_i | (Q_n, U_1^{(n)}, \dots, U_{i-1}^{(n)}), \quad Q_n \geq i > 1.$$

Hence

$$(\hat{Q}_n, \hat{U}^{(n)}) \stackrel{\mathcal{D}}{=} (Q_n, U^{(n)}).$$

Similarly, let

$$Q(\omega) = g \quad \text{for } \omega_0 \in [\mathbb{P}[Q \leq g], \mathbb{P}[Q \leq g + 1]) \quad \text{and} \quad U_i(\omega) = \omega_i I(Q(\omega) \geq i).$$

Clearly (Q, U) has the required distribution and (a) and (b) are proved. To prove (c) it is enough to observe that

$$|\hat{U}_i^{(n)}(\omega) - U_i(\omega)| \leq \frac{1}{n - \hat{Q}_n(\omega)},$$

$\hat{Q}_n(\omega) \xrightarrow{a.s.} Q(\omega)$ and since \hat{Q}_n and Q are integers, $\hat{Q}_n = Q$ eventually. \square

We now prove Theorem 5.2 (c). The proof of (d) is similar. Assume that all variables, including $\{x_i^n\}$ are defined on a probability space such that Lemma 6.6 holds where $Q_n(\omega)$ is the number of non-zero (upper) diagonals present in $T_n(\cdot, \omega)$ and $nU_1^{(n)}(\omega), \dots, nU_{Q_n(\omega)}^{(n)}(\omega)$ are their positions. In particular, $Q_n(\omega) \rightarrow Q(\omega)$ a.s.. By Lemma 6.4, it is enough to prove

$$\limsup_{n \rightarrow \infty} \text{sp}_{T_n}(\omega) \leq \text{supp}(F^T(\cdot, \omega)) \quad \mathcal{L}^\infty \text{ a.s..}$$

If $Q(\omega) = 0$, then for large n , $Q_n(\omega) = 0$ and T_n is null. Again, $F^T(\cdot, \omega)$ is degenerate at 0. So, the result follows immediately. If $Q(\omega) > 0$, we carry out the proof in the following five steps.

Step 1. Now fix a value of $Q = S > 0$ (say). From Lemma 6.6,

$$(U_1, \dots, U_S | Q = S) \stackrel{i.i.d.}{\sim} U[0, 1].$$

Now, given any $\epsilon > 0$, we find a $\delta > 0$ such that

$$(1 - 2\delta)^S > (1 - \epsilon). \tag{6.31}$$

Let $N_{S,\epsilon} := \{\omega \mid Q(\omega) = S \text{ and } (U_1(\omega), \dots, U_S(\omega)) \in ([0, \delta] \cup [\delta, 1])^S\}$. Then

$$\mathcal{L}^\infty [N_{S,\epsilon} | Q = S] < \epsilon.$$

Take any $\omega \notin N_{S,\epsilon}$. By Lemma 6.6, we know that,

$$|U_i^{(n)}(\omega) - U_i(\omega)| \leq \frac{c(Q(\omega))}{n} \quad \forall i \leq S. \tag{6.32}$$

Hence,

$$1 - \frac{c(Q(\omega))}{n\delta} \leq \frac{U_i^{(n)}(\omega)}{U_i(\omega)} \leq 1 + \frac{c(Q(\omega))}{n\delta}. \quad (6.33)$$

Consider the sequence $k_n = \lfloor n^{\frac{1}{8S}} \rfloor$. Let

$$\begin{aligned} \mathcal{P}_{\eta,a,b}(\omega) := & \text{set of all paths of length } 2k_n \text{ returning to the starting point } \eta \text{ with possible jumps} \\ & \pm U_1(\omega), \dots, \pm U_S(\omega) \text{ and which do not cross the boundaries } Y = a \text{ and } Y = b. \end{aligned} \quad (6.34)$$

Similarly, define $\mathcal{P}_{\eta,a,b}^n(\omega)$ for paths with jump sizes $\pm U_1^{(n)}(\omega), \dots, \pm U_S^{(n)}(\omega)$. As $\mathcal{P}_{\eta,1,0}(\omega)$ and $\mathcal{P}_{\eta,1-\frac{1}{n},0}^n(\omega)$ will appear frequently, we use the notations $\mathcal{P}_\eta(\omega)$ and $\mathcal{P}_\eta^n(\omega)$ for them respectively.

Let $F^T(\cdot, \omega)$ be the random LSD and $\beta_{k_n}(\omega)$ be its k_n th moment. Theorem 4.1 (c) implies,

$$\beta_{2k_n}(\omega) = \int \#\mathcal{P}_x(\omega) d\mathcal{L}(x)$$

where $\mathcal{L}(x)$ is the Lebesgue measure on the interval $[0, 1]$. From the proof of Theorem 4.1 (c),

$$\frac{1}{n} \text{Tr}(T_n^{2k_n}(\omega)) = \frac{1}{n} \sum_{i=0}^{n-1} \#\mathcal{P}_{\frac{i}{n}}^n(\omega). \quad (6.35)$$

We shall prove the following relations for sufficiently large n . Let $a' = 1 + 2\frac{c(Q(\omega))}{n\delta}$ and $b' = -2\frac{c(Q(\omega))}{n\delta}$. Then

$$\#\mathcal{P}_{\frac{i}{n}}^n(\omega) \leq \#\mathcal{P}_{\frac{i}{n},a',b'}(\omega). \quad (6.36)$$

$$\sum_{i=0}^{n-1} \#\mathcal{P}_{\frac{i}{n},a',b'}(\omega) \leq 2 \sum_{i=0}^{n-1} \#\mathcal{P}_{\frac{i}{n}}(\omega). \quad (6.37)$$

$$\frac{1}{n} \sum_{i=0}^{n-1} \#\mathcal{P}_{\frac{i}{n}}(\omega) \leq 2 \int \#\mathcal{P}_x(\omega) d\mathcal{L}(x). \quad (6.38)$$

Combining the above relations, $n^{-1} \text{Tr}(T_n^{2k_n}(\cdot, \omega)) \leq 4 \int \#\mathcal{P}_x(\omega) d\mathcal{L}(x) = 4\beta_{2k_n}$. Now,

$$\left[\text{sp}_{T_n}(\omega) \right]^{2k_n} \leq n \text{Tr}(T_n^{2k_n}(\omega)) \leq 4n\beta_{2k_n} \quad \forall k_n \in \mathbb{N}.$$

As $k_n = \lfloor n^{\frac{1}{8S}} \rfloor$, we get

$$\limsup_{n \rightarrow \infty} \text{sp}_{T_n}(\omega) \leq \lim_{n \rightarrow \infty} [4n]^{\frac{1}{2k_n}} \beta_{2k_n}^{\frac{1}{2k_n}} = \text{supp}(F^T(\cdot, \omega)).$$

Also, this happens for any arbitrary $\epsilon > 0$ in (6.31) and for any $S \in \mathbb{N}$. Hence,

$$\limsup_{n \rightarrow \infty} \text{sp}_{T_n}(\omega) \leq \lim_{n \rightarrow \infty} [4n]^{\frac{1}{2k_n}} \beta_{2k_n}^{\frac{1}{2k_n}} = \text{supp}(F^T(\cdot, \omega)) \quad \mathbb{P} \text{ a.s.}$$

We now establish the above relations. Then we shall prove the non-degeneracy of the limit.

Step 2. *Proof of (6.36).* Using any path in $\mathcal{P}_{\frac{i}{n}}^n(\omega)$, we construct a new path where the jumps $\pm U_i^{(n)}(\omega)$ are replaced by $\pm U_i(\omega)$. As $U_1(\omega), \dots, U_S(\omega)$'s are all distinct, this map is an injection. From (6.33), this new path always belong to $\mathcal{P}_{\frac{i}{n}, a, b}^i$, where

$$a = \frac{i}{n} + \frac{(1 - \frac{i}{n})}{(1 - \frac{c(Q(\omega))}{n\delta})} \quad \text{and} \quad b = \frac{i}{n} - \frac{\frac{i}{n}}{1 - \frac{c(Q(\omega))}{n\delta}}.$$

Hence,

$$\#\mathcal{P}_{\frac{i}{n}}^n(\omega) \leq \#\mathcal{P}_{\frac{i}{n}, a, b}^i.$$

Also observe that,

$$\begin{aligned} \frac{i}{n} + \frac{(1 - \frac{i}{n})}{(1 - \frac{c(Q(\omega))}{n\delta})} &\leq \frac{1}{(1 - \frac{c(Q(\omega))}{n\delta})} \quad \forall 0 \leq i \leq n-1 \quad \text{and} \\ \frac{i}{n} - \frac{\frac{i}{n}}{(1 - \frac{c(Q(\omega))}{n\delta})} &\geq 1 - \frac{1}{(1 - \frac{c(Q(\omega))}{n\delta})} \quad \forall 0 \leq i \leq n-1. \end{aligned}$$

Now, we know

$$1 + x \leq \frac{1}{1 - x} \leq 1 + 2x \quad \forall 0 \leq x \leq \frac{1}{2}.$$

Choose, n so large that $\frac{c(Q(\omega))}{n\delta} < \frac{1}{2}$ and in this case,

$$1 + \frac{c(Q(\omega))}{n\delta} \leq \frac{1}{1 - \frac{c(Q(\omega))}{n\delta}} \leq 1 + 2\frac{c(Q(\omega))}{n\delta}.$$

We relax the boundaries further to $Y = 1 + 2\frac{c(Q(\omega))}{n\delta}$ and $Y = -2\frac{c(Q(\omega))}{n\delta}$ respectively with $a' = 1 + 2\frac{c(Q(\omega))}{n\delta}$ and $b' = -2\frac{c(Q(\omega))}{n\delta}$. Now (6.36) follows trivially.

Step 3. *Proof of (6.37).* Let $\zeta_n = \lfloor \sqrt{n} \rfloor$. Now $k_n = \lfloor n^{\frac{1}{8S}} \rfloor$. Hence, $k_n \leq \zeta_n^{\frac{1}{4S}}$. By Lemma 6.5,

$$\sum_{i=1}^S v_i U_i(\omega) \notin [a - \frac{1}{\zeta_n}, a + \frac{1}{\zeta_n}] \quad \text{for almost all } \omega, \quad \text{eventually.}$$

Here, $v_i \in \{-2k_n, \dots, -1, 1, \dots, 2k_n\}$ and $a \in [0, 1]$. As $\frac{\zeta_n}{n} \rightarrow 0$, the boundaries $Y = 1 + 2\frac{c(Q(\omega))}{n\delta}$ and $Y = -2\frac{c(Q(\omega))}{n\delta}$ are stricter than the boundaries $Y = 1 + \frac{1}{\zeta_n}$ and $Y = -\frac{1}{\zeta_n}$. By taking $a = 1$ and letting the initial point of a path be $Y = 0$, we can conclude that any path of the random walk with step lengths lesser than $2k_n$ will not be in the interval $[1 - 2\frac{c(Q(\omega))}{n\delta}, 1 + 2\frac{c(Q(\omega))}{n\delta}]$. The same argument holds true for the starting point $Y = 1$.

Clearly for any i , $\mathcal{P}_{\frac{i}{n}, a', b'}^i(\omega) \supseteq \mathcal{P}_{\frac{i}{n}}^i(\omega)$. Suppose at the n th step, for some integer $\tau = \tau(n)$, there exists a path in $\mathcal{P}_{\frac{\tau}{n}, a', b'}^{\tau}(\omega)$ but not in $\mathcal{P}_{\frac{\tau}{n}}^{\tau}(\omega)$. Without loss of generality, suppose that this path crosses $Y = 1$ but lies within the relaxed boundary $Y = 1 + 2\frac{c(Q(\omega))}{n\delta}$. Any such τ cannot be bounded. Otherwise, $\exists m \in \mathbb{N}$ such that $\tau < m$ for infinitely many n . In this case, any path of length $2k_n$ starting from $\frac{\tau}{n}$ will not belong in the region $[1 - \frac{1}{\zeta_n} + \frac{m}{n}, 1 + \frac{1}{\zeta_n}]$ which is more relaxed than $[1 - 2\frac{c(Q(\omega))}{n\delta}, 1 + 2\frac{c(Q(\omega))}{n\delta}]$. This contradicts the fact that this path is in $\mathcal{P}_{\frac{\tau}{n}, a', b'}^{\tau}(\omega)$. As a consequence, we start from the point $\frac{\tau}{n} - \frac{\tau_0}{n}$ and exactly retrace the same jumps, where

$\tau_0 = \lceil \frac{c(Q(\omega))}{\delta} \rceil + 1$ is a fixed number. We claim that this path is a member of $\mathcal{P}_{\frac{\tau-\tau_0}{n}}(\omega)$. According to our assumption, this path always lies above the line

$$Y = -\left(\frac{\tau_0}{n} + 2\frac{c(P(\omega))}{n\delta}\right) \geq -\frac{1}{\zeta_n} \text{ eventually}$$

and at some point of time the path is within the values $1 - \frac{\tau_0}{n}$ and 1 . On the other hand, from Lemma 6.5, any path starting within the interval $[1, 1 - \frac{\tau_0}{n}]$ with lesser than $2k_n$ many steps can't be within $[-\frac{1}{\zeta_n}, \frac{1}{\zeta_n} - \frac{\tau_0}{n}]$. As a result, it lies above the line $Y = \frac{1}{\zeta_n} - \frac{\tau_0}{n} > 0$ for large enough n . This proves the claim. A similar argument is true for the boundary $Y = 0$. As a consequence,

$$\sum_{i=0}^{n-1} \# \left[\mathcal{P}_{\frac{i}{n}, a', b'}(\omega) \setminus \mathcal{P}_{\frac{i}{n}}(\omega) \right] \leq \sum_{i=0}^{n-1} \# \mathcal{P}_{\frac{i}{n}}(\omega).$$

Hence, (6.37) holds.

Step 4. *Proof of (6.38).* Recall from the proof of Theorem 4.1 (c) that,

$$\int \# \mathcal{P}_x(\omega) d\mathcal{L}(x) = \sum_{i=1}^I \# \mathcal{P}_{I_i}(\omega) \mathcal{L}(I_i)$$

where $\# \mathcal{P}_{I_i}(\omega)$ denote the common value for $\# \mathcal{P}_x(\omega)$ for any $x \in I_i$ as in (6.26). Fix any interval I_i . Let $r_{1, I_i} < r_{2, I_i}$ be the boundaries of this interval. Further, let $l_i := \min\{\frac{j}{n} \mid \frac{j}{n} \in I_i\}$ and $u_i := \max\{\frac{j}{n} \mid \frac{j}{n} \in I_i\}$ respectively. It is easy to see that

$$\begin{aligned} \# \mathcal{P}_{I_i}(\omega) \mathcal{L}(I_i) &= \# \mathcal{P}_{I_i}(\omega) ((r_{2, I_i} - u_i) + (l_i - r_{1, I_i}) + (u_i - l_i)) \\ &= \# \mathcal{P}_{I_i}(\omega) \left(\frac{1}{n} \sum_{\frac{j}{n} \in I_i \setminus \{u_i\}} 1 \right) + (r_{2, I_i} - u_i) \# \mathcal{P}_{I_i}(\omega) + (l_i - r_{1, I_i}) \# \mathcal{P}_{I_i}(\omega) \\ &= \frac{1}{n} \sum_{\frac{j}{n} \in I_i \setminus \{u_i\}} \# \mathcal{P}_{\frac{j}{n}}(\omega) + (r_{2, I_i} - u_i) \# \mathcal{P}_{I_i}(\omega) + (l_i - r_{1, I_i}) \# \mathcal{P}_{I_i}(\omega). \end{aligned}$$

As a result,

$$\begin{aligned} \sum_{i=1}^I \# \mathcal{P}_{I_i}(\omega) \mathcal{L}(I_i) &= \frac{1}{n} \sum_{i=1}^I \sum_{\frac{j}{n} \in I_i \setminus \{u_i\}} \# \mathcal{P}_{\frac{j}{n}}(\omega) + \sum_{i=1}^I (r_{2, I_i} - u_i) \# \mathcal{P}_{I_i}(\omega) + \sum_{i=1}^I (l_i - r_{1, I_i}) \# \mathcal{P}_{I_i}(\omega) \\ &= \frac{1}{n} \sum_{i=1}^I \sum_{\frac{j}{n} \in I_i \setminus \{u_i\}} \# \mathcal{P}_{\frac{j}{n}}(\omega) + \sum_{i=1}^I (r_{2, I_i} - u_i) \# \mathcal{P}_{I_i}(\omega) + \sum_{i=2}^I (u_{i-1} + \frac{1}{n} - r_{2, I_{i-1}}) \# \mathcal{P}_{I_i}(\omega) \\ &= \sum_{\frac{j}{n} \in I_i \setminus \{u_i\}} \# \mathcal{P}_{\frac{j}{n}}(\omega) + \sum_{i=1}^I (r_{2, I_i} - u_i) \# \mathcal{P}_{I_i}(\omega) + \sum_{i=1}^{I-1} (u_i + \frac{1}{n} - r_{2, I_i}) \# \mathcal{P}_{I_{i+1}}(\omega). \end{aligned}$$

Since for $i = 1$, $r_{1, 1} = 0$, in this case $l_1 = r_{1, 1}$ and for any $i > 1$, $l_i = u_{i-1} + \frac{1}{n}$ and $r_{1, i} = r_{2, i-1}$. Hence, the only differences between $\int \# \mathcal{P}_x(\omega) d\mathcal{L}(x)$ and $\frac{1}{n} \sum_{j=0}^{n-1} \# \mathcal{P}_{\frac{j}{n}}(\omega)$ are at

the points u_i . In this case, the integral, $\int \#P_x(\omega) d\mathcal{L}(x)$ considers $\#P_{u_i}(\omega) (= \#P_{I_i}(\omega))$ on the interval $[u_i, r_{2,I_i}]$ and $\#P_{u_i + \frac{1}{n}}(\omega) (= \#P_{I_{i+1}}(\omega))$ on the interval $[r_{2,I_i}, u_i + \frac{1}{n}]$. On the other hand, the sum, $\frac{1}{n} \sum_{j=0}^{n-1} \#P_{\frac{j}{n}}(\omega)$ considers $\#P_{\frac{j}{n}}(\omega)$ on the entire interval $[u_i, u_i + \frac{1}{n}]$. Now,

$$\frac{1}{n} \#P_{u_i}(\omega) \geq (r_{2,I_i} - u_i) \#P_{u_i}(\omega) + (u_i + \frac{1}{n} - r_{2,I_i}) \#P_{u_i + \frac{1}{n}}(\omega) \text{ if and only if } \#P_{u_i}(\omega) \geq \#P_{u_i + \frac{1}{n}}(\omega).$$

Now, let $A := \{j \mid \#P_{\frac{j}{n}}(\omega) \geq \#P_{\frac{j+1}{n}}(\omega)\}$. Obviously

$$\frac{1}{n} \sum_{j \in \{0, \dots, n-1\} \setminus A} \#P_{\frac{j}{n}} \leq \int \#P_x(\omega) d\mathcal{L}(x).$$

Observe, the random walk is symmetric around $Y = \frac{1}{2}$. Hence, for any $j \in \{0, \dots, n-1\}$, $\#P_{\frac{j}{n}}(\omega) = \#P_{1 - \frac{j}{n}}(\omega)$. So,

$$\#P_{\frac{j}{n}}(\omega) \geq \#P_{\frac{j+1}{n}}(\omega) \Leftrightarrow \#P_{1 - \frac{j}{n}}(\omega) \geq \#P_{1 - \frac{j+1}{n}}(\omega).$$

Hence, for each deleted point we have another point with the same number of paths which is not deleted. So,

$$\frac{1}{n} \sum_{i \in \{0, \dots, n-1\} \setminus A} \#P_{\frac{i}{n}}(\omega) \geq \frac{1}{2n} \sum_{i \in 0}^{n-1} \#P_{\frac{i}{n}}(\omega).$$

As a result, (6.38) holds.

Step 5. Proof of non-degeneracy. Recall that $Q \sim \text{Poisson}(\xi)$. If $Q(\omega) = 0$, then $\text{supp}(F^T(\cdot, \omega)) = 0$. From the random walk argument it trivially follows that if $Q(\omega) = 1$ and $U_1(\omega) > \frac{1}{2}$, then $\text{supp}(F^T(\cdot, \omega)) = 1$ as there is exactly one feasible path from any point in the set $[0, 1 - U_1(\omega)) \cup (U_1(\omega), 1]$ and no feasible path otherwise. Hence, $\mathbb{P}[\text{supp}(F^T(\cdot, \omega)) = 1] \geq \frac{1}{2} \xi e^{-\xi}$ and the limiting distribution is non-degenerate.

Proof of (e). The result is trivially true if $\text{supp}(F^A(\cdot, \omega)) = 0$. We now consider the case when $\text{supp}(F^A(\cdot, \omega)) > 0$. For any distribution function $F(\cdot)$, let

$$\text{supp}_1(F) := \inf\{x \mid F(x) = 1\}.$$

Without loss of generality, we assume Lemma 6.6 holds. Hence,

$$\limsup_{n \rightarrow \infty} \lambda_{\max, A_n}^{(n)}(\omega), \limsup_{n \rightarrow \infty} -\lambda_{\min, A_n}^{(n)}(\omega) \leq \lim_{n \rightarrow \infty} \text{sp}_{A_n}(\omega) = \text{supp}(F^A(\cdot, \omega))$$

Now we define a truncated distribution function $\tilde{F}^{A_n}(\cdot, \omega)$ as follows.

$$\tilde{F}^{A_n}(x, \omega) = \begin{cases} 0 & \text{for } x < 0 \\ F^{A_n}(x, \omega) & \text{otherwise.} \end{cases}$$

Similarly, define $\tilde{F}^A(\cdot, \omega)$. It is easy to prove that, $\tilde{F}^{A_n}(\cdot, \omega) \xrightarrow{\mathcal{D}} \tilde{F}^A(\cdot, \omega)$ for a fixed ω . Since $F^A(\cdot, \omega)$ is symmetric almost surely,

$$\text{supp}_1(F^A(\cdot, \omega)) = \text{supp}(F^A(\cdot, \omega)).$$

Again

$$\text{supp}_1(\tilde{F}^A(\cdot, \omega)) = \text{supp}_1(F^A(\cdot, \omega))$$

as $\text{supp}_1(F^A(\cdot, \omega)) > 0$. As a result, $\text{supp}_1(\tilde{F}^A(\cdot, \omega)) = \text{supp}(\tilde{F}^A(\cdot, \omega))$. Hence, $\lambda_{\max, A_n}^{(n)}(\omega) \geq 0$ eventually for almost all ω . Arguments similar to Lemma 6.4 shows that

$$\liminf_{n \rightarrow \infty} \lambda_{\max, A_n}^{(n)}(\omega) \geq \text{supp}(\tilde{F}^A(\cdot, \omega)) = \text{supp}(F^A(\cdot, \omega)).$$

Hence, $\lim_{n \rightarrow \infty} \lambda_{\max, A_n}^{(n)}(\omega) = \text{supp}(F^A(\cdot, \omega))$. Similarly we can prove the result for $\lambda_{\min, A_n}^{(n)}$ also. We omit the details. \square

Acknowledgement. We are grateful to Arusharka Sen for interesting discussions on the sparse Wigner matrix.

References

- [1] G. Anderson, A. Guionnet and O. Zeitouni. *An introduction to random matrices*. Cambridge University Press, 2010.
- [2] Zhidong Bai and Jack W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices, Second Edition* Springer 2006.
- [3] M. Bauer and O. Golinelli. Random incidence matrices: moments of the spectral density. *Journal of Statistical Physics.*, 103, 1/2, 301–337, 2001.
- [4] C. Bordenave and M. Lelarge. Resolvent of large random graphs. *Random Structures Algorithms*, 37, 3, 332-352, 2010.
- [5] C. Bordenave, A. Sen and B. Virag. Mean quantum percolation. <http://arxiv.org/pdf/1308.3755v1.pdf>
- [6] A. Bose, R. S. Hazra and K. Saha. Spectral norm of Circulant-type matrices. *Journal of Theoretical Probability*, 24, 2, 479–516.
- [7] A. Bose, R. S. Hazra and K. Saha. Patterned random matrices and method of moments. In *Proceedings of the International Congress of Mathematicians, Hyderabad*, pages 2203–2230. World Scientific, Singapore and Imperial College Press, UK., 2010.
- [8] A. Bose and A. Sen. Another look at the moment method for large dimensional random matrices. *Electron. J. Probab.*, 13, 21, 588–628, 2008.
- [9] T. Carleman. *Les fonctions quasi analytiques* (in French). Paris: Gauthier-Villars, 1926.
- [10] J. T. Chayes, L. Chayes, J. R. Franz, J. P. Sethna, and S. A. Trugman. On the density of states for the quantum percolation problem. *J. Phys. A*, 19, 18, L1173-L1177, 1986.
- [11] R. M. Dudley. *Real Analysis and Probability*. Cambridge University Press, 2002.
- [12] O. Khorunzhy, M. Shcherbina, V. Vengerovsky. Eigenvalue distribution of large weighted random graphs. *Journal of Mathematical Physics* 45, 4, 1648–1672, 2004.

- [13] K. Knight. On the empirical measure of the Fourier coefficients with infinite variance data. *Statistics and Probability Letters* 12, 2, 109-117, 1991.
- [14] M. Krivelevich and B. Sudakov. The largest eigenvalue of sparse random graphs. *Combinatorics, Probability and Computing*, 12, 61-72, 2003.
- [15] C. A. Tracy and H. Widom. The distribution of the largest eigenvalue in the Gaussian ensembles: $\beta = 1, 2, 4$. *Calogero-Moser-Sutherland models*, Springer, New York, 461-472, 2000.
- [16] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.*, 62, 548-564, 1955.

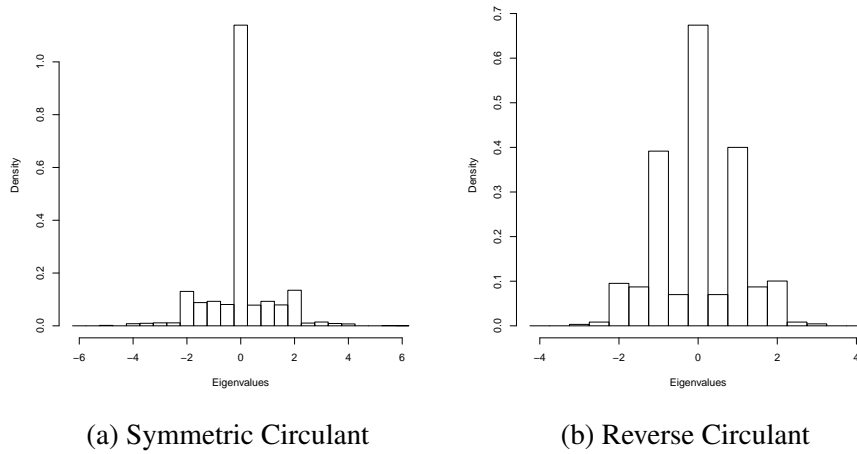


Figure 1: Histogram of eigenvalues with $\xi = 1$, $n = 200$, 50 replications.

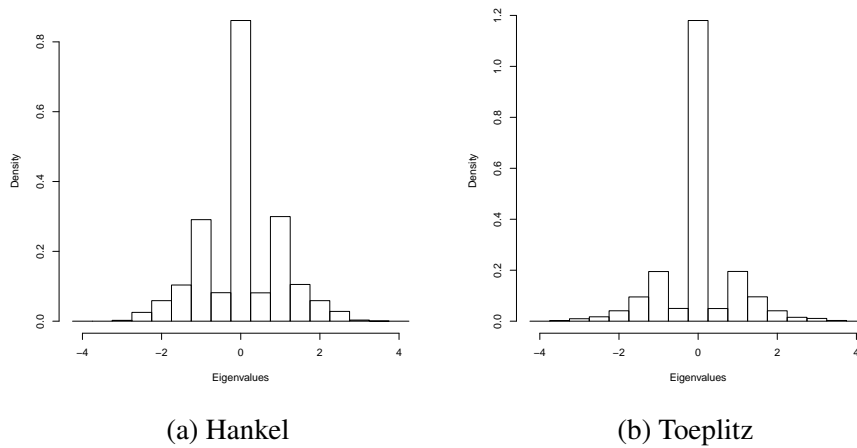


Figure 2: Histogram of eigenvalues with $\xi = 1$, $n = 200$, 50 replications.