LARGEST EIGENVALUE OF LARGE RANDOM BLOCK MATRICES: A COMBINATORIAL APPROACH

DEBAPRATIM BANERJEE
Dept. of Statistics
The Wharton School
University of Pennsylvania
3730 Walnut Street, PA 19104
USA
dban@wharton.upenn.edu

ARUP BOSE [*]
Statistics and Mathematics Unit
Indian Statistical Institute
203 B.T. Road, Kolkata 700108
INDIA
bosearu@gmail.com

January 9, 2016

Technical Report R1/2016
Statistics and Mathematics Unit
Indian Statistical Institute

Abstract

We study the largest eigenvalue of certain block matrices where the number of blocks and the block size both increase with suitable conditions on their relative growth. In one of them, we employ a symmetric block structure with large independent Wigner blocks and in the other we have the Wigner block structure with large independent symmetric blocks. The entries are assumed to be i.i.d. mean 0 variance 1 with an appropriate growth condition on the moments. Under our conditions the limit spectral distribution of these matrices is the standard semi-circle law. It is natural to ask if the extreme eigenvalues converge to the extreme points of its support, namely ± 2. We exhibit models where this indeed happens as well as models where the spectral norm converges to $2\sqrt{2}$. Our proofs are based on combinatorial analysis of the behaviour of the trace of large powers of the matrix.

---

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

1 Introduction

Let A_n be an $n \times n$ real symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. The *empirical spectral measure* μ_n of A_n is the measure on $\mathbb{R}$ given by

$$\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i}, \tag{1.1}$$

where δ_x is the Dirac delta measure at x. The corresponding probability distribution function F^{A_n} on $\mathbb{R}$ is known as the *empirical spectral distribution* (ESD) of A_n. For us the entries of A_n shall be random, and hence F^{A_n} is a *random distribution function*. If F^{A_n} converges weakly almost surely to a non-random distribution function F, then F is called the *limiting spectral distribution* (LSD) of A_n almost surely. When a sequence of real symmetric random matrices has an LSD with compact support then it is a very natural question to ask whether the extreme eigenvalues converge to values that lie outside the support of the LSD. This question has been well studied in the literature.

Consider the standard $n \times n$ Wigner matrix $n^{-1/2} W_n$ where W_n is a symmetric matrix with all entries i.i.d. and mean 0 and variance 1. Then as $n \to \infty$, its LSD is the semi-circle law with support $[-2,\ 2]$. Bai and Yin (1988) [5] proved that the largest eigenvalue of $n^{-1/2} W_n$ converges almost surely to 2 if and only if the entries have finite fourth moments. Similarly, Bai and Silverstein (1998) [4] proved that for Wishart matrices almost surely there exists no eigenvalue outside the ϵ neighbourhood of the support of the LSD under the finiteness of the fourth moment. Haagerup, Schultz and Thorbjørnsen (2005) [12] showed the non-existence of any eigenvalue outside the ϵ neighbourhood of the support for self-adjoint polynomials in independent GUE matrices. For similar works in GOE and GSE cases, see Haagerup and Thorbjørnsen (2006) [13] and Schultz (2005) [20].

Anderson (2013) [2] considered a class of independent $N \times N$ Wigner matrices $X^N := \{X_l^N\}_{l=1}^{\infty}$ and a self-adjoint matrix $f \in \mathrm{Mat}_n \mathbb{C}\langle X^N \rangle$ of fixed dimension n. Here $\mathbb{C}\langle X^N \rangle$ is the non-commutative polynomial algebra generated over $\mathbb{C}$ by the sequence of non-commutative random variables $X^N = \{X_l^N\}_{l=1}^{\infty}$. Suppose the fourth moment is finite. He showed that as $N \to \infty$, there does not exist any eigenvalue outside the ϵ neighbourhood of the support of the LSD for the matrices $f(\frac{X^N}{\sqrt{N}})$.

Loubaton (2015) [15] dealt with this question for certain Gaussian block Hankel random matrices. He considered the matrices $W_N = (W_N^{(1)T}, \ldots, W_N^{(M)T})^T$ where $(W_N^{(m)})_{m=1}^{M}$ is a class of independent $L \times N$ block Hankel matrices with properly scaled i.i.d. complex Gaussian entries when $\frac{LM}{N} \to c_* \in (0, \infty)$ and $M \to \infty$. Here for any matrix A, A^T denotes its transpose. He showed that the LSD of $W_N W_N^*$ is almost surely the Marchenko-Pastur law with parameter c_*. Moreover, if $L = O(N^{\alpha})$ with $\alpha < \frac{2}{3}$, almost surely all the eigenvalues of $W_N W_N^*$ are located in the ϵ neighbourhood of the support of this Marchenko-Pastur law.

We consider two classes of block matrices with $m \times m$ blocks where each block is an $n \times n$ matrix. Unlike Anderson (2013) [2], we assume both m and n grow to infinity. The entries are assumed to be i.i.d. mean 0 variance 1. We further assume that there exists $0 < C_B < \infty$ such that $\mathrm{E}[|X|^h] \leq h^{C_B h}\ \forall h \in \mathbb{N}$. This assumption is in the spirit of Anderson, Guionnet and Zeitouni (2010, page 24) [1] and precludes the use of truncation arguments that would be necessary (see for example Bai and Yin (1988) [5]) if lesser moment conditions were to be used. This is done for convenience so that we can avoid the truncation arguments.

First consider the following two models: (i) the blocks are arranged in a Wigner pattern and the blocks are symmetric independent and (ii) the block structure is symmetric while the blocks are independent Wigner. We first observe that for all the above models, under suitable conditions on the symmetry, the LSD is the standard semi-circular law with support $[-2, 2]$. In addition, when a particular random variable appears at most once along any row of the symmetric blocks or the block structure, then in Theorem 2.1 we claim that there is no eigenvalue outside $(-2 - \epsilon, 2 + \epsilon)$.

Observe that with the Symmetric Circulant blocks or block structure some indices appears twice in a row and that is enough to violate the result of Theorem 2.1. In Theorem 2.2 we consider two models: (i) the Wigner block structure with independent Symmetric Circulant blocks and (ii) the Symmetric Circulant structure with independent Wigner blocks. We prove that in either case when the Wigner dimension (structure or blocks) is sufficiently larger than the Symmetric circulant dimension (structure or blocks), the largest eigenvalue converge in probability to $2\sqrt{2}$. In particular, for models (i) and (ii) stated above, we require $\frac{n}{m} \to 0$ and $\frac{m}{n} \to 0$ respectively. Our proofs are combinatorial and Lemma 2.1.23 of Anderson, Guionnet and Zeitouni (2010, page 24-28) [1] lies at the heart of our proofs.

2 Main results

We shall work in the framework of patterned matrices. A *patterned matrix* is defined through a *link function* L_n. For each n, let $L_n : \{0, 1, \ldots, n\}^2 \to \mathbb{Z}^d$ be a function ($d = 1$ or 2). Then $A_n = ((x_{L_n(i,j)}))_{1 \le i,j \le n}$ is said to be a patterned matrix with link function L_n. Here $\{x_{i,j}\}$ or $\{x_i\}$ is the *input sequence* of random variables. For notational convenience we write L for L_n. Some common link functions are given by

$$\begin{aligned} L_W(i,j) &= (\min(i,j), \max(i,j)), \text{ (Wigner)} \qquad (2.1)\\ L_T(i,j) &= |i-j|, \text{ (Toeplitz)}\\ L_H(i,j) &= i+j, \text{ (Hankel)}\\ L_{RC}(i,j) &= (i+j) \mod n, \text{ (Reverse Circulant)}\\ L_{SC}(i,j) &= n/2 - |n/2 - |i-j||, \text{ (Symmetric Circulant).} \end{aligned}$$

We now introduce two important classes of link functions that have been used earlier in the literature by Bose and Sen (2008) [10] and Banerjee and Bose (2013) [7].

Definition 2.1. (Properties B and C)

(a) The link function L is said to satisfy Property B if along any row or column, the maximum number of occurrences of any particular random variable is uniformly bounded. That is,

$$\sup_n \sup_{t \in \mathbb{Z}^d} \sup_{1 \le k \le n} \#\{l : 1 \le l \le n, L(k,l) = t\} \vee \#\{l : 1 \le l \le n, L(l,k) = t\} \le \Delta < \infty. \qquad (2.2)$$

(b) The link function L is said to satisfy the stronger Property C if in (a) above,

$$\Delta = 1. \qquad (2.3)$$

All the five link functions given above satisfy Property B and out of these, only the Wigner, Hankel and Reverse Circulant links satisfy Property C.

Let L_m be a link function and $I = \mathbb{Z}$ or $\mathbb{Z}^2$ be the range of L. Also let $\{A_{n,i}\}_{i \in I}$ be a class of $n \times n$ matrices indexed by I. Let $\mathbb{B}_n(L_m)$ be the block matrix

$$\mathbb{B}_n(L_m) := \begin{pmatrix} A_{n,L_m(1,1)} & A_{n,L_m(1,2)} & \dots & A_{n,L_m(1,m-1)} & A_{n,L_m(1,m)} \\ A_{n,L_m(2,1)} & A_{n,L_m(2,2)} & \dots & A_{n,L_m(2,m-1)} & A_{n,L_m(2,m)} \\ & & \vdots & & \\ A_{n,L_m(m,1)} & A_{n,L_m(m,2)} & \dots & A_{n,L_m(m,m-1)} & A_{n,L_m(m,m)} \end{pmatrix}. \tag{2.4}$$

For convenience, we write $\mathbb{B}_n(L)$ instead of $\mathbb{B}_n(L_m)$.

The concepts of circuits, words and related notions available in Bose and Sen (2008) [10] will be crucial for us. We give a short overview. A *circuit* of length $l(\pi) := h$ corresponding to a generic link function $L_n = L_n(i,j), 1 \le i,j \le n$ is any function $\pi : \{0,1,\dots,h\} \to \{1,2,\dots,n\}$ with $\pi(0) = \pi(h)$. The link function will be clear from the context so often we don't mention the link function and call it just a circuit. Let Tr denote the trace of a matrix. Then,

$$\begin{aligned} \operatorname{Tr}\left[(\mathbb{B}_n(L))^h\right] &= \sum_{\pi:l(\pi)=h} \operatorname{Tr}(A_{n,L(\pi(0),\pi(1))} \dots A_{n,(L(\pi(h-1),\pi(h)))}) \\ &= \sum_{\pi:l(\pi)=h} \sum_{\pi':l(\pi')=h} A_{n,L(\pi(0),\pi(1))}(\pi'(0),\pi'(1)) \dots A_{n,(L(\pi(h-1),\pi(h)))}(\pi'(h-1),\pi'(h)). \end{aligned} \tag{2.5}$$

Here any π is a function from $\{0,1,\dots,h\} \to \{1,2,\dots,m\}$ and any π' is a function from $\{0,1,\dots,h\} \to \{1,2,\dots,n\}$.

k circuits $\pi_1,\dots,\pi_k$ are *jointly-matched* if each L-value occurs at least twice across all circuits. They are *cross-matched* if each circuit has at least one L-value which occurs in at least one of the other circuits.

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

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

This defines an equivalence relation for circuits. Any equivalence class can be indexed by a partition of $\{1,2,\dots,h\}$. Each block of a given partition identifies the positions where the L-matches take place. We can 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 the word *ababa* represents the partition $\{\{1,3,5\}\{2,4\}\}$ and identifies the circuits π for which $L(\pi(0),\pi(1)) = L(\pi(2),\pi(3)) = L(\pi(4),\pi(5)) \ne L(\pi(1),\pi(2)) = L(\pi(3),\pi(4))$. Let $\Pi(w)$ denote the equivalence class induced by w. For ease of calculations, we also need the following class of circuits corresponding to a word w.

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

The set of words of length $2k$ where each letter is repeated at least (respectively exactly) twice is denoted by $\mathcal{W}_{2k}$ (respectively $\mathcal{P}_{2k}$). Each $w \in \mathcal{P}_{2k}$ is said to be *pair-matched.* A word $w \in \mathcal{P}_{2k}$ is called *Catalan* (or *non-crossing*) if there are no four positions $i_1 < i_2 < i_3 < i_4$ such that $w[i_1] = w[i_3]$ and $w[i_2] = w[i_4]$. The set of all Catalan words of length $2k$ is denoted by C_{2k}.

For example, *abba* is Catalan whereas *abab* is not. For any word w, let $d(w)$ denote the total number of distinct letters in w. It is well known (see Bose and Sen (2008) [10]) that if an $n \times n$ patterned matrix satisfies Property B, then

$$\#\Pi^*(w) = O(n^{d(w)+1}), \text{ and } \#(\Pi^*(w)\backslash\Pi(w)) = O(n^{d(w)}) \text{ for any word } w. \tag{2.6}$$

Let

$$p(w) := \lim_{n\to\infty} \frac{\#\Pi^*(w)}{n^{d(w)+1}} = \lim_{n\to\infty} \frac{\#\Pi(w)}{n^{d(w)+1}} \text{ whenever the limit exists (they exist or not together).} \tag{2.7}$$

The following assumptions will remain in force throughout our article. Assumption (a) is borrowed from Anderson, Guionnet and Zeitouni (2010, page 24) [1]. Its main consequence is that the truncation arguments as in Bai and Yin (1988) [5] can be bypassed. Assumption (b) specifies the relation between the block size, n and the dimension of the block structure, m.

Throughout the paper, we shall always assume that the input random variables satisfy Assumption2.1 (a).

Assumption 2.1. (a) The input random variables are real valued i.i.d. with mean 0 and variance 1. Let X be a random variable with this common distribution. Suppose $\exists\ 0 < C_B < \infty$ such that

$$\mathrm{E}[|X|^h] \leq h^{C_B h}, \quad \forall h \in \mathbb{N}.$$

(b) $m, n \to \infty$ and there exist constants $0 < C, D < \infty$ such that $n = O(m^C)$ and $m = O(n^D)$.

For any matrix A, let $\lambda_{\max,A}$ and $\lambda_{\min,A}$ denote respectively the largest and the smallest eigenvalues of A. The spectral norm of A is defined as

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

Theorem 2.1. *Suppose Assumption 2.1 (b) holds. Then the LSD of* $\frac{1}{\sqrt{nm}}\mathbb{B}_n(L)$ *is standard semi-circular. The quantities* $\mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}$, $\lambda_{\max,\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}$ *and* $\lambda_{\min,\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}$ *converge to* 2, 2 *and* -2 *almost surely in each of the following two cases:*

(a) The symmetric blocks $\{A_{n,i}\}_{i\in\mathbb{Z}^2}$ *are i.i.d. for all* $i \neq j$*, their common pattern satisfies Property C and L is the Wigner link.*

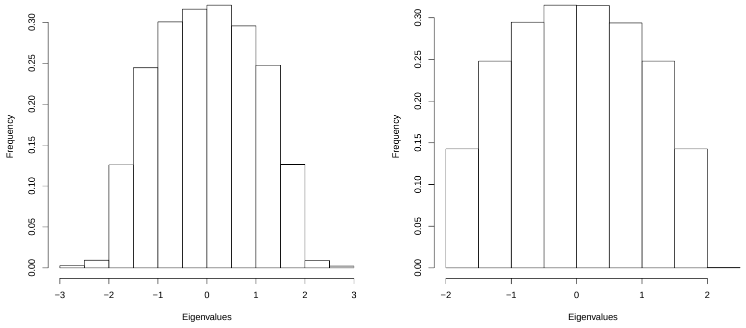
(b) $\{A_{n,i}\}_{i\in\mathbb{Z}^d}$ *are i.i.d. Wigner matrices whose entries satisfy Assumption 2.1 and L satisfies Property C.*

In the above theorem, the semi-circularity of the LSD of $\frac{1}{\sqrt{nm}}\mathbb{B}_n(L)$ is straightforward: suppose $\{A_{n,i}\}_{i\in\mathbb{Z}^2}$, are i.i.d. symmetric matrices for which $p(w) = 1$ for every $w \in C_{2k}$ for $\frac{1}{\sqrt{n}}A_{n,(1,1)}$. Suppose further that L is the Wigner link function. Then from Theorem 3.1 of Banerjee and Bose (2015) [6] it is clear that the LSD of $\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)$ is the semi-circular law. On the other hand, from part (ii) of Theorem 3 of Banerjee and Bose (2013) [7], the semi-circular law also holds when we have certain general link functions L and independent Wigner blocks. These results also indicate that for the semi-circularity of the LSD of block matrices, one only requires $p(w) = 1$ for every $w \in C_{2k}$ for the non-Wigner link functions. A sufficient condition for this is Property C. So, in the above theorem we see that this Property is also sufficient for the extreme eigenvalues to converge to the boundary of the support of the LSD. In the next theorem we give two block matrix models where the LSD is semi-circle but the extreme eigenvalues do not converge to the boundaries of the LSD.

Theorem 2.2. *Suppose Assumption 2.1 (b) holds. Then the LSD of* $\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)$ *is standard semi-circular but* $\mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}$ *converges to* $2\sqrt{2}$ *in probability in each of the following cases:*

(a) $\{A_{n,i}\}_{i\in\mathbb{Z}^2}$ *are i.i.d. Symmetric Circulant matrices and L is the Wigner link and* $\frac{n}{m} \to 0$.

(b) $\{A_{n,i}\}_{i\in\mathbb{Z}^d}$ *are i.i.d. Wigner matrices and L is the Symmetric Circulant link function and* $\frac{m}{n} \to 0$.



(a) Symmetric Circulant blocks (b) Hankel blocks

Figure 1: ESD of $\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)$, L is the Wigner link, entries are i.i.d. $N(0, \mathbf{1})$, $m/5 = n = 30$.

We performed simulations for the cases Theorem 2.1 (a) and Theorem 2.2 (a). Figure 1 provides the histograms of the ESD $\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)$ when L is the Wigner link and the blocks are either Hankel (then Theorem 2.1 (a) applies) or Symmetric Circulant (then Theorem 2.2 (a) applies). Note that the histograms look close to the semi-circle law. However, one may observe that the histogram corresponding to the Symmetric Circulant case has heavier concentration of eigenvalues in the regions $[2,\ 3]$. and $[-3,\ -2]$. Here $m = 150$ and $n = 30$ as we need $\frac{n}{m} \to 0$ for Theorem 2.2 (a) . For the Hankel case, the ESD is way more concentrated in $[-2,\ 2]$. For the same two cases, in Figure 2 we plot the spectral norm against the dimension n (and m is taken to be $n[\sqrt{n}]$ in order to ensure $m >> n$). It is quite clear that for the Hankel case the spectral norm stays more or less around 2. However, the figure clearly suggests that the spectral norm is beyond 2 in the Symmetric Circulant case.

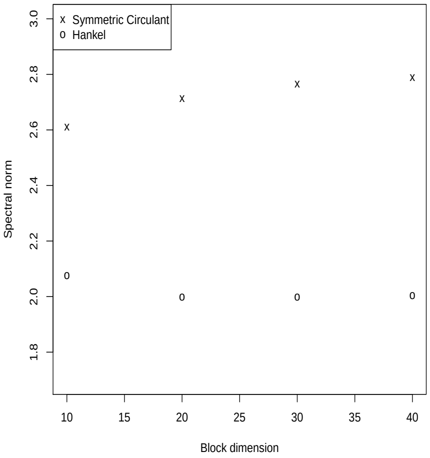


Figure 2: Spectral norm of $\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)$, L is the Wigner link, entries are i.i.d. $N(0, \mathbf{1})$,$m \approx n^{\frac{3}{2}}$.

3 Proofs

We first present a brief introduction to partitions. Let S be a finite set. A typical partition of S is $\sigma = \{V_1, \ldots, V_r\}$ where the blocks $\{V_i\}$ are non-empty disjoint such that $\cup V_i = S$. For $p, q \in S$, we write $p \sim_\sigma q$ if p and q belong to the same block of σ. $\mathcal{P}(S)$ denotes the set of all partitions of S and if $S = \{1, \ldots, n\}$ we simply denote it by $\mathcal{P}(n)$. Note that a word w of length n is nothing but a partition of $S = \{1, \ldots, n\}$ and $d(w)$ is the number of blocks in w.

Let $S(n, r)$ be the set of all possible partitions of $\{1, 2, \ldots, n\}$ into r blocks. We need the following important bounds for $\#S(n, r)$.

Lemma 3.1. *(Rennie and Dobson (1969) [19]) For any $r < n$,*

$$\frac{1}{2}(r^2 + r + 2)r^{n-r-1} - 1 \;\leq\; \#S(n, r) \;\leq\; \frac{1}{2}\binom{n}{r}r^{n-r}, \;\; \forall 1 \leq r \leq n - 1.$$

$\mathcal{P}(n)$ is a partially ordered set in the following way. Let $\sigma_1, \sigma_2 \in \mathcal{P}(n)$. Then $\sigma_1 \leq \sigma_2$ if every block of σ_1 is contained in some block of σ_2. For example, let $n = 5$. Then $\sigma_1 = \{\{1, 2\}; \{3, 4\}; \{5\}\} \leq \sigma_2 = \{\{1, 2, 3, 4\}; \{5\}\}$.

Let P be a finite partially ordered set. Let σ_1 and σ_2 be any two elements in P. Define sets $U := \{\tau \in P : \tau \geq \sigma_1 \text{ and } \tau \geq \sigma_2\}$ and $V := \{\tau \in P : \tau \leq \sigma_1 \text{ and } \tau \leq \sigma_2\}$. P is called a lattice if U and V are non-empty and have a minimum and maximum element in P respectively for any two $\sigma_1, \sigma_2 \in P$.

Let P be a lattice and $\sigma_1, \sigma_2 \in P$. Then the join and meet of σ_1 and σ_2 are defined as the minimum of U and maximum of V respectively. They are denoted by $\sigma_1 \vee \sigma_2$ and $\sigma_1 \wedge \sigma_2$ respectively. For any n, $\mathcal{P}(n)$ is a lattice with respect to the partial ordering $\leq$.

Definition 3.1. (Graph associated with a circuit) Let π be circuit of length k. Let $G_\pi = (V_\pi, E_\pi)$ be the graph corresponding to the circuit π as follows. The vertex set V_π is the collection of all distinct values among $\{\pi(0),\ldots,\pi(k-1)\}$. The edge set $E_\pi := \{\{\pi(i),\pi(i+1)\} : i = 0,1,\ldots,k-1\}$. The set of self-edges and connecting edges are defined respectively as $E^s_\pi := \{\{u,u\} \in E_\pi\}$ and $E^c_\pi := E_\pi \backslash E^s_\pi$. For $e \in E_\pi$ let N^π_e denote the number of times the edge e appears in $\{\{\pi(i),\pi(i+1)\} : 0 \le i \le k-1\}$.

For the Wigner matrix it is easy to note that $\#E_\pi$ is exactly same as $d(w)$ where w is the unique word such that $\pi \in \Pi(w)$.

We now introduce the "class representative of closed word" as in Anderson, Guionnet and Zeitouni (2010, page 13) [1]. Since the terminology "word" has a different meaning in the current paper, we opt to call them "companion words".

Definition 3.2. A companion word $\mathfrak{w}$ of length k is a sequence of Greek letters $\alpha_0\alpha_1\ldots\alpha_k$ with $\alpha_0 = \alpha_k$.

Using these companion words, we introduce another equivalent class of circuits of length k as follows. We say a circuit π belongs to the equivalence class induced by $\mathfrak{w}$ if and only if

$$\pi(i) = \pi(j) \Leftrightarrow \mathfrak{w}[i] = \mathfrak{w}[j].$$

For example, π_1, π_2 defined by $\pi_1(0) = 1, \pi_1(1) = 2, \pi_1(2) = 2, \pi_1(3) = 1$ and $\pi_2(0) = 1, \pi_2(1) = 3, \pi_2(2) = 3, \pi_2(3) = 1$ will belong to the same equivalence class of the companion word $\alpha_0\alpha_1\alpha_1\alpha_0$.

Definition 3.3. For any companion word $\mathfrak{w}$, we denote $\Gamma(\mathfrak{w})$ to be the set of all equivalent circuits corresponding to $\mathfrak{w}$.

Let $\mathcal{W}_{k,t}$ be the class of all companion words of length k with t distinct letters such that the graph of the corresponding circuits has each edge repeated at least twice. That is,

$$\mathcal{W}_{k,t} := \{\mathfrak{w} : \forall\, \pi \in \Gamma(\mathfrak{w}),\ l(\pi) = k,\ \#V_\pi = t \text{ and } N^\pi_e \ge 2\ \forall e \in E_\pi\}. \tag{3.1}$$

For a proof of the following Lemma, see Anderson, Guionnet and Zeitouni (2010, page 23) [1].

Lemma 3.2. *For all integers $k > 2t - 2$,*

$$\#\mathcal{W}_{k,t} \le 2^k k^{3(k-2t+2)}.$$

Let w and w' be any two elements in $\mathcal{P}(k)$ for some $k \in \mathbb{N}$ and consider $w \wedge w'$. It is easy to observe $d(w \wedge w') \ge \max\{d(w), d(w')\}$.

Lemma 3.3. *Given two partitions $w' \ge w''$, any partition w satisfying $w \wedge w' = w''$ is uniquely determined upto a partition of* $\{1,2,\ldots,d(w'')\}$.

Proof. Let $C_1,\ldots,C_{d(w'')}$ be the blocks of w''. Let $\sigma := \{V_1,\ldots,V_r\}$ be any partition of $\mathcal{S} := \{1,\ldots,d(w'')\}$ where $r \le d(w'')$. Now observe that for any given σ, w' and w'', a partition w can be constructed as $w := \{B_1,\ldots,B_r\}$ where

$$B_i = \cup_{\{j : j \in V_i\}} C_j.$$

It is easy to observe that $w \wedge w' = w''$.

On the other hand suppose w is any partition such that $w \wedge w' = w''$. Then consider $\sigma \in S(d(w''), d(w)) := \sigma(w, w')$ in the following manner. $\sigma = \{V_1,\ldots,V_{d(w)}\}$ where

$$V_i = \{j : C_j \subset B_i\}.$$

It is easy to check that this definition of σ is consistent with the earlier one. Hence the proof is complete. □

3.1 Proof of Theorem 2.1.

Proofs of (a) and (b) are similar. We prove only (a). Observe that we can directly apply Theorem 2.2 of Banerjee and Bose (2015) [6] to conclude that the LSD of $\frac{1}{\sqrt{nm}}\mathbb{B}_n(L)$ is the standard semi-circular law.

Let $\lambda_1, \dots, \lambda_{mn}$ be the eigenvalues of $\frac{1}{\sqrt{nm}}\mathbb{B}_n(L)$. From their proof, it also follows that the $(2k)$ moment of the ESD of this matrix equals

$$\beta_{2k}\Big(\frac{\mathbb{B}_n(L)}{\sqrt{mn}}\Big) := \frac{1}{mn}\sum_{i=1}^{mn}\lambda_i^{2k} \to \beta_{2k} \text{ almost surely}$$

where β_{2k} is the $(2k)$th moment of the semi-circular law, and equals

$$\beta_{2k} = \frac{1}{k+1}\binom{2k}{k}. \tag{3.2}$$

We first focus on the spectral norm. Then

$$\beta_{2k}\Big(\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)\Big) \le \Big(\text{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}\Big)^{2k}.$$

Hence,

$$\lim_{n\to\infty}\beta_{2k}\Big(\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)\Big) \overset{a.s.}{=} \beta_{2k} \le \liminf_{n\to\infty}(\text{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)})^{2k}.$$

As a result,

$$[\beta_{2k}]^{\frac{1}{2k}} \le \liminf_{n\to\infty}\text{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}(\omega).$$

Hence

$$2 = \lim_{k\to\infty}[\beta_{2k}]^{\frac{1}{2k}} \le \liminf_{n\to\infty}\text{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}(\omega).$$

So it remains to prove that

$$\limsup \text{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \le 2 \text{ almost surely.}$$

For any $k \in \mathbb{N}$

$$\Big[\text{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}\Big]^{2k} \le \sum_{i=1}^{mn}\lambda_i^{2k} = \text{Tr}\Big[\frac{1}{(mn)^k}\left(\mathbb{B}_n(L)\right)^{2k}\Big]. \tag{3.3}$$

In order to prove the Theorem, we offer an efficient upper bound to the expectation of the R.S. of (3.3). Observe that

$$\begin{aligned}\text{Tr}\Big[(\mathbb{B}_n(L))^{2k}\Big] &= \sum_{\pi:l(\pi)=2k}\text{Tr}(A_{n,L(\pi(0),\pi(1))}\dots A_{n,L(\pi(2k-1),\pi(2k))})\\ &= \sum_{\pi:l(\pi)=2k}\sum_{\pi':l(\pi')=2k}A_{n,L(\pi(0),\pi(1))}(\pi'(0),\pi'(1))\dots A_{n,L(\pi(2k-1),\pi(2k))}(\pi'(2k-1),\pi'(2k)).\end{aligned} \tag{3.4}$$

Let

$$A_\pi := A_{n,L(\pi(0),\pi(1))}\dots A_{n,L(\pi(2k-1),\pi(2k))}.$$

Fix any π. Let $\mathfrak{w}$ be the companion word such that $\pi \in \Gamma(\mathfrak{w})$. As the blocks are independent, it is easy to observe that if $\mathfrak{w} \notin \cup_t \mathcal{W}_{2k,t}$ then $\mathrm{E}\left[\mathrm{Tr}(A_\pi)\right] = 0$. So in order to get a non-trivial contribution of $\mathrm{E}\left[\mathrm{Tr}(A_\pi)\right]$ we need $\mathfrak{w} \in \cup_t \mathcal{W}_{2k,t}$.

We at first prove that $t \leq k+1$. Consider the graph $G_\pi = (V_\pi, E_\pi)$ induced by the circuit π. Consider the set E_π. If for any $j \geq 1$, $\pi(j) \notin \{\pi(0), \ldots, \pi(j-1)\}$, then the unordered tuple $\{\pi(j-1), \pi(j)\}$ is an edge which has never appeared in the graph G_π when the vertices $\{\pi(0), \ldots, \pi(j-1)\}$ were considered. As a consequence, apart from the first vertex $\pi(0)$, the introduction of a new vertex always introduces a new edge in the graph G_π.

Now observe that $\#E_\pi = 2k$. However, our assumption that $\mathfrak{w} \in \cup_t \mathcal{W}_{2k,t}$ will enforce each edge to appear at least twice. Hence we get

$$2(t-1) \leq 2k \ \text{ or } \ t \leq k+1. \tag{3.5}$$

Now let l be number of edges repeated exactly twice in the graph G_π. We prove $2k - 2l < 6(k-t+1)$. By definition G_π is connected. As a result, we have $\#E_\pi^c \geq \#V_\pi - 1 = t-1$. Also the graph has $2k$ many edges in all and so

$$\begin{aligned} &(2k) \geq 3(\#E_\pi^c - l) + 2l + 2\#E_\pi^s \\ \Rightarrow &3k + 3 - 3t \geq k - l \\ \Leftrightarrow &2k - 2l \leq 6(k+1-t). \end{aligned} \tag{3.6}$$

Let $\{i_{p,q}\}_{1 \leq p \leq l, 1 \leq q \leq 2}$ be the positions of the l double edges in the circuit π. So we get

$$\{\pi(i_{p,1}-1), \pi(i_{p,1})\} = \{\pi(i_{p,2}-1), \pi(i_{p,2})\} \ \forall \ 1 \leq p \leq l.$$

If we consider the monomial A_π, then

$$\mathrm{E}\,\mathrm{Tr}(A_\pi) = \sum_{\pi' : l(\pi') = 2k} \mathrm{E}\left[A_{n,L(\pi(0),\pi(1))}(\pi'(0), \pi'(1)) \ldots A_{n,(L(\pi(2k-1),\pi(2k)))}(\pi'(2k-1), \pi'(2k))\right]. \tag{3.7}$$

Now observe that if we need

$$\mathrm{E}\left[A_{n,L(\pi(0),\pi(1))}(\pi'(0), \pi'(1)) \ldots A_{n,(L(\pi(2k-1),\pi(2k)))}(\pi'(2k-1), \pi'(2k))\right] \neq 0 \tag{3.8}$$

for some π', then we need to have

$$\{\pi'(i_{p,1}-1), \pi'(i_{p,1})\} = \{\pi'(i_{p,2}-1), \pi'(i_{p,2})\} \ \forall \ 1 \leq p \leq l.$$

We now apply Holder's inequality on the L.S. of (3.8) to obtain

$$\mathrm{E}\left[A_{n,L(\pi(0),\pi(1))}(\pi'(0), \pi'(1)) \ldots A_{n,(L(\pi(2k-1),\pi(2k)))}(\pi'(2k-1), \pi'(2k))\right] \leq 1^{2l}\,\mathrm{E}[|X|^{2k-2l}]. \tag{3.9}$$

Since $\mathrm{E}[X^2] = 1$, we get that $\mathrm{E}[X^{2k}] \geq \mathrm{E}[X^2]^k = 1$ for $k \in \mathbb{N}$. Hence

$$\begin{aligned} &\left[\mathrm{E}[X^{2k+2}]\right]^{\frac{1}{2k+2}} \geq \left[\mathrm{E}[X^{2k}]\right]^{\frac{1}{2k}} \\ \Rightarrow &\,\mathrm{E}[X^{2k+2}] \geq \left[\mathrm{E}[X^{2k}]\right]^{\frac{2k+2}{2k}} \geq \mathrm{E}[X^{2k}]. \end{aligned} \tag{3.10}$$

So using (3.6) we get that

$$\mathrm{E}[|X|^{2k-2l}] \leq \mathrm{E}[|X|^{6(k+1-t)}] \leq (6(k+1-t))^{6C_B(k+1-t)} \ \ \forall \ k.$$

Hence R.S. of (3.9) is bounded by $(6(k+1-t))^{6C_B(k+1-t)}$.

We now give an upper bound to the number of π' such that (3.8) holds. Let w and w' be the unique words induced by π and π' respectively. As we have fixed the circuit π, the word w is also fixed. In addition to this we also fix another partition $w'' \leq w$ and find out all such w' s such that $w \wedge w' = w''$.

Applying Lemma 3.3 it is easy to observe that given the partition w'', the number of such w's is bounded by

$$\sum_{d(w')=1}^{d(w'')} \#S(d(w''), d(w')).$$

Let $A_1, \ldots, A_{d(w)}$ and $B_1, \ldots, B_{d(w')}$ be the blocks corresponding to the partitions w and w'. It is easy to observe that for (3.8) to hold, we require that either $A_i \cap B_j = \emptyset$ or $\#(A_i \cap B_j) \geq 2$ for all i, j. As a consequence, each block of the partition w'' has cardinality at least 2. Hence $d(w'') \leq k$.

Now given w, we give an upper bound to the number of possible choices of w''.

Note that l is exactly same as the number of double letters in w. As a result, when we chose $w'' \leq w$ such that each block has cardinality at least 2, we are compelled to take the blocks of w with cardinality exactly 2 as they are. Hence we are left with at most $2k-2l$ many choices to complete the partition w''. We use the crude upper bound $(2k-2l)^{2k-2l+1}$ for all possible partitions of $2k-2l$ objects. As a consequence, using (3.6) again we get

$$\#(w \wedge w') \leq (2k-2l)^{2k-2l+1} \leq (6(k+1-t))^{6(k+1-t)+1} \leq (6(k+1-t))^{12(k+1-t)}. \tag{3.11}$$

As $d(w'') \leq k$, $\#S(d(w''), d(w')) \leq \#S(k, d(w'))$. Combining these, given w, the number of partitions w' is bounded by

$$(6(k+1-t))^{12(k+1-t)} \sum_{d(w')=1}^{k} (6(k+1-t))^{12(k+1-t)} \#S(k, d(w')). \tag{3.12}$$

Also, by Property C the number of circuits π' corresponding to w' is bounded by $n^{d(w')+1}$. Hence

$$\begin{aligned}
&\mathrm{E}\left[\mathrm{Tr}(\mathbb{B}_n(L))^{2k}\right] \\
&\leq \sum_{t=1}^{k+1} \sum_{\mathfrak{w} \in \mathcal{W}_{2k,t}} \sum_{\pi \in \Gamma(\mathfrak{w})} (6(k+1-t))^{6C_B(k+1-t)} (6(k+1-t))^{12(k+1-t)} \sum_{d(w')=1}^{k} \#S(k, d(w')) n^{d(w')+1} \\
&= \sum_{t=1}^{k+1} \sum_{\mathfrak{w} \in \mathcal{W}_{2k,t}} \#\Gamma(\mathfrak{w}) (6(k+1-t))^{6C_B(k+1-t)} (6(k+1-t))^{12(k+1-t)} \sum_{d(w')=1}^{k} \#S(k, d(w')) n^{d(w')+1}.
\end{aligned} \tag{3.13}$$

Now we use Lemma 3.1 for a bound of $\#S(k, d(w'))$ and the fact that $\#\Gamma(\mathfrak{w}) \leq m^t$ for any $\mathfrak{w} \in \mathcal{W}_{2k,t}$. As a consequence, we arrive at the following upper bound to (3.13).

$$\begin{aligned}
&\sum_{t=1}^{k+1} \sum_{d(w')=1}^{k} \#\mathcal{W}_{2k,t} m^t (6(k+1-t))^{6C_B(k+1-t)} (6(k+1-t))^{12(k+1-t)} \\
&\qquad \frac{1}{2}\binom{k}{d(w')} (d(w'))^{k-d(w')} n^{d(w')+1}.
\end{aligned} \tag{3.14}$$

Now observe that

$$\binom{k}{d(w)} \le \frac{k!}{(d(w))!} \le k^{k-d(w)} \text{ and } (d(w))^{k-d(w)} \le k^{k-d(w)}.$$

All these bounds and Lemma 3.2 lead us to the inequality

$$\begin{aligned}
&\frac{1}{(mn)^{k+1}} \mathrm{E}\left[\mathrm{Tr}(\mathbb{B}_n(L))^{2k}\right] \\
&\le \sum_{t=1}^{k+1} \sum_{d(w')=1}^{k} \frac{1}{(nm)^{k+1}} 2^{2k-1}(2k)^{6(k-t+1)} m^t \,(6(k+1-t))^{6C_B(k+1-t)}\,(6(k+1-t))^{12(k+1-t)} \left(\frac{k^2}{n}\right)^{k-d(w')}. \\
&= 2^{2k-1} \sum_{t=1}^{k+1} \sum_{d(w')=1}^{k} \left(\frac{C_1 k^{C_2}}{m}\right)^{k+1-t} \left(\frac{k^2}{n}\right)^{k-d(w')}
\end{aligned} \tag{3.15}$$

where C_1 and C_2 are some known constants.

Choose $\delta > 0$ arbitrary. Now

$$\begin{aligned}
P\left[\left(\mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}\right)^{2k} \ge (2+\delta)^{2k}\right] &\le \frac{\mathrm{E}\left[(\mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)})^{2k}\right]}{(2+\delta)^{2k}} \\
&\le \frac{2^{2k}}{(2+\delta)^{2k}} mn \sum_{t=1}^{k+1} \sum_{d(w')=1}^{k} \left(\frac{C_1 k^{C_2}}{m}\right)^{(k+1-t)} \left(\frac{k^2}{n}\right)^{k-d(w')} \qquad (3.16) \\
&\le \frac{2^{2k}}{(2+\delta)^{2k}} \alpha m^{C+1} \sum_{t=1}^{k+1} \sum_{d(w')=1}^{k} \left(\frac{C_1 k^{C_2}}{m}\right)^{(k+1-t)} \left(\frac{k^2}{n}\right)^{k-d(w')}
\end{aligned}$$

for some known α. Now we can choose $k = k(m,n) \to \infty$ such that

$$\frac{C_1 k^{C_2+1}}{m} \to 0, \quad \frac{k^2}{n} \to 0 \text{ and } \frac{m^C}{(1+\frac{\delta}{2})^{2k}} = O(\frac{1}{(nm)^2}).$$

For this choice of $k(m,n)$, the R.S. of (3.16) is $O(\frac{1}{(nm)^2})$. By use of Borel-Cantelli lemma, this implies

$$P[\limsup(\mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}) \le 2+\delta] = 1 \text{ for every } \delta > 0.$$

This in turn clearly implies that

$$\limsup \mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \le 2 \text{ almost surely}$$

and the proof is complete for the spectral norm.

Now we prove the claims for $\lambda_{\max,\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}$ and $\lambda_{\min,\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}$. From the definition of $\mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}$ it is clear that

$$\lambda_{\max,\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \le \mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \text{ and } -\lambda_{\min,\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \le \mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}.$$

As a consequence, we get

$$\limsup \lambda_{\max, \frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \leq 2 \text{ and } \limsup -\lambda_{\min, \frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \leq 2 \text{ a.s.}$$

Let $F^{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}(x,\omega)$ and $F(x)$ be the empirical distribution function of the eigenvalues of $\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)$ and the distribution function of the semi-circular law respectively. We now consider the truncated distribution function

$$\tilde{F}^{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}(x,\omega) = \begin{cases} 0 & \text{for } x < 0 \\ F^{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}(x,\omega) & \text{otherwise.} \end{cases}$$

similarly we consider $\tilde{F}(x)$. It is easy to observe that

$$\int x^{2k} d\tilde{F}^{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}(x,\omega) \overset{a.s.}{\to} \int x^{2k} d\tilde{F}(x) \ \forall k \in \mathbb{N}.$$

Now applying an argument similar to above we get that

$$\liminf \lambda_{\max, \frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \geq 2 \text{ a.s.}$$

which proves that

$$\lim \lambda_{\max, \frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} = 2 \text{ a.s.}$$

We can repeat the same arguments for $-\lambda_{\min, \frac{1}{\sqrt{mn}}\mathbb{B}_n(L)}$ and hence the proof is complete. □

3.2 Proof of Theorem 2.2.

The proofs of (a) and (b) are similar. So we only prove (a). We need the following result. Its proof can be found in Banerjee and Bose (2013) [7].

Lemma 3.4. *Let $L(i,j), 1 \leq j \leq n$ be the Symmetric Circulant link function. Then for every $w \in \mathcal{C}_{2k}$,*

$$\#\Pi^*(w) = \begin{cases} (n-1)^{k+1} + (2n)^k & \textit{if } n \textit{ is odd} \\ (n-2)^{k+1} + 2(2n-2)^k & \textit{otherwise.} \end{cases}$$

We have two link functions namely the Wigner link function $L = (\min(i,j), \max(i,j))$ and the Symmetric Circulant link function $L_{SC}(i,j) = n/2 - |n/2 - |i-j||$. We need to consider the equivalence of circuits with respect to both L and L_{SC}. Let $\Pi_{L_{SC}}(w)$, $p_{L_{SC}}(w)$, $\Pi_L(w)$ and $p_L(w)$ be the equivalence classes induced by w and the contributions of w, for the link functions L_{SC} and L respectively.

One can directly apply Theorem 3.1 of Banerjee and Bose (2015) [6] to conclude that in this case the LSD of $\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)$ is the semi-circular law. Also by appropriate modification of the proof of Theorem 2.1, one gets

$$\begin{aligned} \frac{1}{(mn)^{k+1}} \mathrm{E}\left[\mathrm{Tr}(\mathbb{B}_n(L))^{2k}\right] &\leq 2^{2k-1} \sum_{t=1}^{k+1} \sum_{d(w')=1}^{k} \left(\frac{C_1 k^{C_2}}{m}\right)^{k+1-t} \left(\frac{k^2}{n}\right)^{k-d(w')} 2^{2k-d(w')} \\ &= \frac{(2\sqrt{2})^{2k}}{2} \sum_{t=1}^{k+1} \sum_{d(w')=1}^{k} \left(\frac{C_1 k^{C_2}}{m}\right)^{k+1-t} \left(\frac{2k^2}{n}\right)^{k-d(w')}. \end{aligned} \tag{3.17}$$

This implies that

$$\limsup \mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \leq 2\sqrt{2} \text{ almost surely.}$$

As a consequence, it is enough to prove that for any $\epsilon > 0$,

$$P\left[\mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \leq 2\sqrt{2} - \epsilon\right] \to 0 \text{ as } m, n \to \infty.$$

Let

$$\beta_{2k} := \frac{1}{(mn)^{k+1}} \operatorname{Tr}(\mathbb{B}_n(L))^{2k}.$$

Now

$$\begin{aligned}
\mathrm{P}\left[\mathrm{sp}_{\frac{1}{\sqrt{mn}}\mathbb{B}_n(L)} \leq (2\sqrt{2} - \epsilon)\right] &\leq \mathrm{P}\left[\beta_{2k} \leq (2\sqrt{2} - \epsilon)^{2k}\right] \\
&\leq P\left[(\beta_{2k} - \mathrm{E}[\beta_{2k}])^2 \geq \left(\mathrm{E}[\beta_{2k}] - (2\sqrt{2} - \epsilon)^{2k}\right)^2\right] \\
&\leq \frac{\mathrm{Var}(\beta_{2k})}{(\mathrm{E}[\beta_{2k}])^2} \frac{(\mathrm{E}[\beta_{2k}])^2}{\left(\mathrm{E}[\beta_{2k}] - (2\sqrt{2} - \epsilon)^{2k}\right)^2}.
\end{aligned} \tag{3.18}$$

Our eventual goal is to find a sequence $k = k(m, n)$ such that

$$\liminf\, [\mathrm{E}[\beta_{2k}]]^{\frac{1}{2k}} \geq 2\sqrt{2} \text{ and } \frac{\beta_{2k}}{\mathrm{E}[\beta_{2k}]} \xrightarrow{P} 1. \tag{3.19}$$

In particular, we exhibit a choice $k = k(m, n) \to \infty$ such that

$$\liminf\, [\mathrm{E}[\beta_{2k}]]^{\frac{1}{2k}} \geq 2\sqrt{2} \tag{3.20}$$

and

$$\frac{\mathrm{Var}(\beta_{2k})}{(\mathrm{E}[\beta_{2k}])^2} \to 0. \tag{3.21}$$

The rest of the argument would then follow from appropriate use of Chebyshev inequality to be discussed in the last step of this proof.

We break the proof into five steps.

Claim 1. Suppose $k = k(m, n) \to \infty$ such that $\frac{k}{\log n} \to \infty$, $\frac{k^{p_1}}{m} \to 0$ and $\frac{k^{p_2}}{n} \to 0$ for suitably chosen constants p_1 and p_2 as $m, n \to \infty$. Then (3.20) holds.

Proof: This is achieved by obtaining an appropriate lower bound for $\mathrm{E}[\beta_{2k}]$.

$$\begin{aligned}
\mathrm{E}[\beta_{2k}] &= \frac{1}{(mn)^{k+1}} \sum_{\pi} \mathrm{E}[\operatorname{Tr}[A_{L(\pi)}]] \\
&= \frac{1}{(mn)^{k+1}} \sum_{\pi} \sum_{\pi'} \mathrm{E}[A_{L(\pi)}(\pi')].
\end{aligned} \tag{3.22}$$

Here

$$A_{L(\pi)} := A_{n,L(\pi(0),\pi(1))} \ldots A_{n,L(\pi(2k-1),\pi(2k))}$$

and

$$A_{L(\pi)}(\pi') := A_{n,L(\pi(0),\pi(1))}(\pi'(0),\pi'(1))\ldots A_{n,(L(\pi(2k-1),\pi(2k)))}(\pi'(2k-1),\pi'(2k)).$$

Let

$$\mathrm{Ne} := \{(\pi,\pi') : \mathrm{E}\left[A_{L(\pi)}(\pi')\right] < 0\} \text{ and } T_1 = \frac{1}{(mn)^{k+1}} \sum_{(\pi,\pi')\in \mathrm{Ne}} \mathrm{E}\left[A_{L(\pi)}(\pi')\right].$$

Fix circuits π and π' and let w and w' be the unique words such that $\pi \in \Pi_L(w)$ and $\pi' \in \Pi_{L_{SC}}(w')$. Also let $\mathfrak{w}$ denote the companion word such that $\pi \in \Gamma(\mathfrak{w})$. Now observe that for any $(\pi,\pi') \in \mathrm{Ne}$, one needs both the words w and w' to have *at least one letter appearing more than thrice*. In this case $d(w') < k$ and $t < k+1$ where $\mathfrak{w} \in \mathcal{W}_{2k,t}$. One can repeat the arguments of Theorem 2.1 to conclude that

$$|T_1| \le \sum_{t=1}^{k} \sum_{d(w')=1}^{k-1} \frac{(2\sqrt{2})^{2k}}{2} \left(\frac{C_1 k^{C_2}}{m}\right)^{k+1-t} \left(\frac{2k^2}{n}\right)^{k-d(w')}. \tag{3.23}$$

Using the formula for geometric series, the R.S. of (3.23) becomes

$$\frac{2^{3k}}{2}\left(\frac{C_1 k^{C_2}}{m}\right)\left(\frac{2k^2}{n}\right)\left(1-\left(\frac{C_1 k^{C_2}}{m}\right)^{k}\right)\left(1-\left(\frac{2k^2}{n}\right)^{k-1}\right). \tag{3.24}$$

Now consider,

$$\mathrm{Po} := \{(\pi,\pi') : \mathrm{E}\left[A_{L(\pi)}(\pi')\right] > 0\} \text{ and } T_2 = \frac{1}{(mn)^{k+1}} \sum_{(\pi,\pi')\in \mathrm{Po}} \mathrm{E}\left[A_{L(\pi)}(\pi')\right].$$

Clearly

$$\mathrm{E}[\beta_{2k}] = T_1 + T_2.$$

Observe that

$$T_2 \ge \sum_{w\in C_{2k}} \frac{\#\Pi_L(w)}{m^{k+1}} \frac{\#\Pi^*_{L_{SC}}(w)}{n^{k+1}}.$$

We now prove that $\lim_{m,n\to\infty} \frac{|T_1|}{T_2} \to 0$ for a suitable choice of $k = k(m,n)$.

One can follow carefully the proof of Theorem 4 of Bose and Sen (2008) [10] to see that for any $w \in C_{2k}$,

$$\frac{\#\Pi_L(w)}{m^{k+1}} \ge \prod_{i=1}^{k}(1-\frac{i}{m}) \ge (1-\frac{k}{m})^k. \tag{3.25}$$

As a consequence,

$$1 \ge \left[\prod_{i=1}^{k}(1-\frac{i}{m})\right]^{\frac{1}{2k}} \ge \sqrt{(1-\frac{k}{m})} \to 1$$

whenever, $\frac{k}{m} \to 0$. Stirling's approximation of $k! \sim \alpha k^{k+\frac{1}{2}} e^{-k}$ where α is some known constant, implies

$$\frac{(2k)!}{k!(k+1)!} \sim \frac{(2k)^{2k+\frac{1}{2}}}{k^{2k+1}} = 2^{2k+\frac{1}{2}} \frac{1}{\sqrt{k}}. \tag{3.26}$$

Applying Lemma 3.4 and the above approximation,

$$T_2 \geq \begin{cases} \frac{(n-1)^{k+1}+(2n)^k}{n^{k+1}} \frac{(2k)!}{k!(k+1)!} \prod_{i=1}^{k}(1-\frac{i}{m}) & \geq \frac{1}{\sqrt{k}}\frac{1}{n}2^{3k+\frac{1}{2}} \prod_{i=1}^{k}(1-\frac{i}{m}) \quad \text{if } n \text{ is odd} \\ \frac{(n-2)^{k+1}+2(2n-2)^k}{n^{k+1}} \frac{(2k)!}{k!(k+1)!} \prod_{i=1}^{k}(1-\frac{i}{m}) & \geq \frac{1}{\sqrt{k}}2^{3k+1+\frac{1}{2}} \frac{(n-1)^k}{n^{k+1}} \prod_{i=1}^{k}(1-\frac{i}{m}) \quad \text{if } n \text{ is even.} \end{cases} \tag{3.27}$$

Without loss of generality we consider n to be odd. The case when n is even follows similarly. Choose a sequence $k = k(m,n)$ such that $\frac{C_1 k^{C_2+2}}{m} \to 0$ and $\frac{k^2}{n} \to 0$. Now apply the lower bound of T_2 as given in (3.27) to get the following upper bound on $\frac{|T_1|}{T_2}$.

$$\begin{aligned} \frac{|T_1|}{T_2} &\leq \frac{\frac{2^{3k}}{2}\left(\frac{C_1 k^{C_2}}{m}\right)\left(\frac{2k^2}{n}\right)\left(1-\left(\frac{C_1 k^{C_2}}{m}\right)^k\right)\left(1-\left(\frac{2k^2}{n}\right)^{k-1}\right)}{\frac{1}{\sqrt{k}}\frac{1}{n}2^{3k+\frac{1}{2}} \prod_{i=1}^{k}(1-\frac{i}{m})} \\ &= \frac{\frac{1}{\sqrt{2}}\frac{C_1 k^{C_2+2}}{m}\left(1-\left(\frac{C_1 k^{C_2}}{m}\right)^k\right)\left(1-\left(\frac{2k^2}{n}\right)^{k-1}\right)}{\prod_{i=1}^{k}(1-\frac{i}{m})}. \end{aligned} \tag{3.28}$$

Recall (3.25) to conclude that $\prod_{i=1}^{k}(1-\frac{i}{m}) \to 1$ whenever $\frac{k^2}{m} \to 0$. Now all the terms except $\frac{C_1 k^{C_2+2}}{m}$ in the numerator and denominator of the R.S. of (3.28) are finite in the limit. As a consequence, we conclude that the R.S. of (3.28) goes to 0. Hence we can choose a sufficiently large m and n such that $\mathrm{E}[\beta_{2k}] \geq \frac{1}{2}T_2$. So we choose $p_1 = C_2 + 2$ and $p_2 = 2$ to satisfy all the conditions.

Now it is easy to observe from (3.27), (3.26) and (3.28) that if further $\frac{k}{\log n} \to \infty$ as $m, n \to \infty$, then $\liminf\, [\mathrm{E}[\beta_{2k}]]^{\frac{1}{2k}} \geq 2\sqrt{2}$. This completes the proof of Claim 1.

Now we focus on the variance term. Note that

$$\mathrm{Var}(\beta_{2k}) = \frac{1}{(mn)^{2(k+1)}} \sum_{\pi_1}\sum_{\pi_2} \mathrm{E}\left[\left(\mathrm{Tr}(A_{L(\pi_1)}) - \mathrm{E}\,\mathrm{Tr}(A_{L(\pi_1)})\right)\left(\mathrm{Tr}(A_{L(\pi_2)}) - \mathrm{E}\,\mathrm{Tr}(A_{L(\pi_2)})\right)\right]. \tag{3.29}$$

We show that several terms in the above sum are negligible.

Claim 2. The sum in (3.29) reduces to that over all circuits (π_1, π_2) such that π_1 and π_2 are cross-matched and jointly-matched.

Proof. It is enough to prove

$$\mathrm{E}\left[\left(\mathrm{Tr}(A_{L(\pi_1)}) - \mathrm{E}\,\mathrm{Tr}(A_{L(\pi_1)})\right)\left(\mathrm{Tr}(A_{L(\pi_2)}) - \mathrm{E}\,\mathrm{Tr}(A_{L(\pi_2)})\right)\right] = 0 \tag{3.30}$$

when the circuits π_1 and π_2 are either not cross-matched or not jointly-matched.

Suppose π_1 and π_2 are not cross-matched. Then the collection of matrix entries in $A_{L(\pi_1)}$ and $A_{L(\pi_2)}$ are completely disjoint. Also by our assumption, the entries of different matrices are independent. Hence the random variables $\mathrm{Tr}(A_{L(\pi_1)})$ and $\mathrm{Tr}(A_{L(\pi_2)})$ are independent. Hence (3.30) holds.

Now we consider the case when π_1 and π_2 are not jointly-matched. Without loss of generality let us assume that the matrix $A_{n,L(\pi_1(j-1),\pi_1(j))}$ appears exactly once in the monomials $A_{L(\pi_1)}$ and $A_{L(\pi_2)}$. As a consequence, for the two circuits π_1 and π_2 one can write L.S. of (3.30) as

$$\sum_{\pi_1'}\sum_{\pi_2'} \mathrm{E}\left[\left(A_{L(\pi_1)}(\pi_1') - \mathrm{E}(A_{L(\pi_1)}(\pi_1'))\right)\left(A_{L(\pi_2)}(\pi_2') - \mathrm{E}(A_{L(\pi_2)}(\pi_2'))\right)\right]. \tag{3.31}$$

Now for any π'_1 and π'_2 the random variable $A_{n,L(\pi_1(j-1),\pi_1(j))}(\pi'_1(j-1),\pi'_1(j))$ occurs exactly once. Hence $\mathrm{E}(A_{L(\pi_1)}(\pi'_1)) = 0$ and we have

$$\begin{aligned}\mathrm{E}\left[\left(A_{L(\pi_1)}(\pi'_1) - \mathrm{E}(A_{L(\pi_1)}(\pi'_1))\right)\left(A_{L(\pi_2)}(\pi'_2) - \mathrm{E}(A_{L(\pi_2)}(\pi'_2))\right)\right] \\ = \mathrm{E}\left[A_{n,L(\pi_1(j-1),\pi_1(j))}(\pi'_1(j-1),\pi'_1(j))B\right]\end{aligned} \tag{3.32}$$

where B is a random variable independent of $A_{n,L(\pi_1(j-1),\pi_1(j))}(\pi'_1(j-1),\pi'_1(j))$. Hence (3.32) is 0. As a consequence, (3.30) holds. This completes the proof of Claim 2.

Now to upper bound the variance, we first give an upper bound for the number of terms in the variance formula. Then we provide a uniform upper bound for these terms.

Let

$$I := \{(\pi_1,\pi_2) : \pi_1 \text{ and } \pi_2 \text{ are jointly-matched and crossed-matched}\}$$

and

$$J := \cup_{(\pi_1,\pi_2)\in I}\{(\mathfrak{w}_1,\mathfrak{w}_2) : \pi_1 \in \Gamma(\mathfrak{w}_1) \text{ and } \pi_2 \in \Gamma(\mathfrak{w}_2)\}.$$

Claim 3. There exists a function $f : J \to \cup_{t=1}^{2k+1}\mathcal{W}_{4k,t}$ such that for some known constants C_1 and C_2,

$$\#\{\mathfrak{w}_1,\mathfrak{w}_2 : f(\mathfrak{w}_1,\mathfrak{w}_2) = \mathfrak{w}_3\} \le C_1 k^{C_2}.$$

Proof. At first fix any two jointly-matched and cross-matched circuit π_1,π_2 and let $\mathfrak{w}_1,\mathfrak{w}_2$ be the companion words such that $\pi_1 \in \Gamma(\mathfrak{w}_1)$ and $\pi_2 \in \Gamma(\mathfrak{w}_2)$.

We construct the function f in the following way.

We place $\mathfrak{w}_1$ and $\mathfrak{w}_2$ side by side. The circuits are cross-matched and L is the Wigner link function. As a consequence, there exists at least two positions j and l such that $\mathfrak{w}_1[j] = \mathfrak{w}_2[l]$. Without loss of generality, we assume that j is the minimum value such that the letter $\mathfrak{w}_1[j]$ appears in $\mathfrak{w}_2$ and l is the minimum value for which $\mathfrak{w}_2[l] = \mathfrak{w}_1[j]$. Now we construct $\mathfrak{w}_3 = f(\mathfrak{w}_1,\mathfrak{w}_2) \in \cup_{t=1}^{2k+1}\mathcal{W}_{4k,t}$ as follows.

$$\mathfrak{w}_3[i] = \begin{cases} \mathfrak{w}_1[i] & 0 \le i \le j \\ \mathfrak{w}_2[l+i-j] & j < i \le j+2k-l \\ \mathfrak{w}_2[i-j-2k+l] & j+2k-l < i \le j+2k \\ \mathfrak{w}_1[i-2k] & j+2k < i \le 4k. \end{cases} \tag{3.33}$$

It is easy to check that every edge in $\mathfrak{w}_1$ and $\mathfrak{w}_2$ is present in $\mathfrak{w}_3$. So $\mathfrak{w}_3 \in \cup_{t=1}^{2k+1}\mathcal{W}_{4k,t}$.

Now given any $\mathfrak{w}_3 \in \cup_{t=1}^{2k+1}\mathcal{W}_{4k,t}$, we give an upper bound to the number of possible $(\mathfrak{w}_1,\mathfrak{w}_2)$ such that $f(\mathfrak{w}_1,\mathfrak{w}_2) = \mathfrak{w}_3$.

Observe that $\mathfrak{w}_3[l] = \mathfrak{w}_3[2k+l]$. So if $\mathfrak{w}_3 = f(\mathfrak{w}_1,\mathfrak{w}_2)$, then it has to have two positions j' and l' such that $\mathfrak{w}_3[j'] = \mathfrak{w}_3[l']$. The total possible choices of the unordered pair $\{j',l'\}$ is bounded by $\binom{4k+1}{2}$. If j' and l' are fixed, then we need to find any position $j' \le b \le l'$. This can be done in $l'-j' \le 4k$ ways. Now we can define $\mathfrak{w}_1$ and $\mathfrak{w}_2$ in the following manner.

$$\mathfrak{w}_1[i] = \begin{cases} \mathfrak{w}_3[i] & 0 \le i \le j' \\ \mathfrak{w}_3[i-j'+l'] & j' < i \le 4k-l' \end{cases} \tag{3.34}$$

and

$$\mathfrak{w}_2[i] = \begin{cases} \mathfrak{w}_3[b+i] & 0 \le i \le l'-b \\ \mathfrak{w}_3[i+b-l'+j] & l'-b' < i \le l'-j'. \end{cases} \tag{3.35}$$

Observe that $l(\mathfrak{w}_1) = 4k + j' - l'$ and $l(\mathfrak{w}_2) = l' - j'$. So $l(\mathfrak{w}_1) = 2k$ and $l(\mathfrak{w}_2) = 2k$ if and only if $l' - j' = 2k$. However from the construction of f it is clear that given $\mathfrak{w}_3$, $\mathfrak{w}_1$ and $\mathfrak{w}_2$ must be of the forms in (3.34) and (3.35). As a consequence,

$$\#\{\mathfrak{w}_1, \mathfrak{w}_2 : f(\mathfrak{w}_1, \mathfrak{w}_2) = \mathfrak{w}_3\} \leq 4k\binom{4k+1}{2} \leq C_1 k^{C_2} \tag{3.36}$$

for some known constants C_1, C_2. This completes the proof of Claim 3.

Observe that from Step 3, we can write (3.29) as

$$\frac{1}{(mn)^{2(k+1)}} \sum_{t=1}^{2k+1} \sum_{\mathfrak{w}_3 \in \mathcal{W}_{4k,t}} \sum_{f(\mathfrak{w}_1,\mathfrak{w}_2)=\mathfrak{w}_3} \sum_{(\pi_1,\pi_2)\in(\Gamma(\mathfrak{w}_1),\Gamma(\mathfrak{w}_2))} \mathrm{E}\left[\left(\mathrm{Tr}(A_{L(\pi_1)}) - \mathrm{E}\,\mathrm{Tr}(A_{L(\pi_1)})\right)\left(\mathrm{Tr}(A_{L(\pi_2)}) - \mathrm{E}\,\mathrm{Tr}(A_{L(\pi_2)})\right)\right]. \tag{3.37}$$

Claim 4. For any jointly-matched and cross-matched π_1 and π_2,

$$\mathrm{E}\left[\left(\mathrm{Tr}(A_{L(\pi_1)}) - \mathrm{E}\,\mathrm{Tr}(A_{L(\pi_1)})\right)\left(\mathrm{Tr}(A_{L(\pi_2)}) - \mathrm{E}\,\mathrm{Tr}(A_{L(\pi_2)})\right)\right] \tag{3.38}$$

is bounded above by

$$2(6(2k-t+1))^{(12+C_B)(2k-t+1)}\left[\sum_{d(w')=1}^{2k-1} \#S(2k, d(w'))2^{4k-d(w')}n^{d(w')+2} + 2^{2k}n^{2k+1}\right].$$

Proof. Let w_1 and w_2 be the unique words such that $\pi_1 \in \Pi_L(w_1)$ and $\pi_2 \in \Pi_L(w_2)$. We place the words w_1 and w_2 side by side honouring the mutual matches between π_1 and π_2 to get the new word $w := (w_1 w_2)$. It is easy to observe that w has all letters repeated at least twice. Let l be the number of double letters in w.

Now recall Claim 3 and consider $\mathfrak{w}_3 = f(\mathfrak{w}_1, \mathfrak{w}_2)$. Let π_3 be any circuit in $\Gamma(\mathfrak{w}_3)$ and w_3 be the word such that $\pi_3 \in \Pi_L(w_3)$. From the construction of the function f in Claim 3, it is clear that for any given letter s in the word w, there is a letter g in w_3 which appears exactly same number of times as the letter s. Hence w_3 also has l doubly appearing letters. Again $\mathfrak{w}_3 \in \mathcal{W}_{4k,t}$ for some t. Hence from the proof of Theorem 2.1 we get $4k - 2l \leq 6(2k - t + 1)$.

Recall L.S. of (3.31) to conclude that we need to consider only those π'_1 and π'_2 for which

$$\mathrm{E}\left[\left(A_{L(\pi_1)}(\pi'_1) - \mathrm{E}(A_{L(\pi_1)}(\pi'_1))\right)\left(A_{L(\pi_2)}(\pi'_2) - \mathrm{E}(A_{L(\pi_2)}(\pi'_2))\right)\right] \neq 0.$$

Now define w'_1 and w'_2 to be the unique words such that $\pi'_1 \in \Pi_{L_{SC}}(w'_1)$ and $\pi'_2 \in \Pi_{L_{SC}}(w'_2)$ and $w' := (w'_1 w'_2)$. Here w' is also obtained honouring the mutual matches between π'_1 and π'_2. One can check that even in this case each block of the partition $w \wedge w'$ has cardinality greater or equal to 2. So one can apply Lemma 3.3 and argue along the same lines as in the proof of Theorem 2.1 to conclude that

$$\#w' \leq \sum_{d(w')=1}^{2k} (6(2k+1-t))^{12(2k+1-t)}\#S(2k, d(w')). \tag{3.39}$$

Now arguments similar to those in the proof of Theorems 2.1 will imply that for any π'_1, π'_2 such that

$$\mathrm{E}\left[A_{L(\pi_1)}(\pi'_1) - \mathrm{E}\left(A_{L(\pi_1)}(\pi'_1)\right)\right] \mathrm{E}\left[A_{L(\pi_2)}(\pi'_2) - \mathrm{E}\left(A_{L(\pi_2)}(\pi'_2)\right)\right] \neq 0, \tag{3.40}$$

we have

$$\mathrm{E}\left[\left|A_{L(\pi_1)}(\pi'_1) A_{L(\pi_2)}(\pi'_2)\right|\right] \leq (4k-2l)^{C_B(4k-2l)}.$$

As $A_{L(\pi_1)}(\pi'_1)$ and $A_{L(\pi_2)}(\pi'_2)$ are both product of powers of i.i.d. random variables, it is easy to observe from Holder's inequality

$$\mathrm{E}\left[\left|A_{L(\pi_1)}(\pi'_1) A_{L(\pi_2)}(\pi'_2)\right|\right] \geq \mathrm{E}\left[\left|A_{L(\pi_1)}(\pi'_1)\right|\right] \mathrm{E}\left[\left|A_{L(\pi_2)}(\pi'_2)\right|\right].$$

So L.S. of (3.40) is bounded by $2(4k-2l)^{C_B(4k-2l)}$.

Once w', w'_1 and w'_2 are fixed, let $B(w'_1, w'_2)$ be the class of all circuits (π'_1, π'_2) such that $\pi'_1 \in \Pi_{L_{SC}}(w'_1)$ and $\pi'_2 \in \Pi_{L_{SC}}(w'_2)$. It is easy to observe that $\#B(w'_1, w'_2) \leq 2^{4k-d(w')} n^{d(w')+2}$. Here one extra power of n comes due to the free choice of $\pi'_2(0)$.

However, for the special case $d(w') = 2k$, one can show that $\#B(w'_1, w'_2) \leq 2^{2k} n^{2k+1}$. The argument essentially follows from the Property B of Symmetric Circulant matrices and is similar to the proof of Lemma 2(a) of Bose and Sen (2008) [10].

Combining all these results we get the following upper bound to (3.38).

$$2(6(2k-t+1))^{(12+C_B)(2k-t+1)} \left[\sum_{d(w')=1}^{2k-1} \#S(2k, d(w')) 2^{4k-d(w')} n^{d(w')+2} + 2^{2k} n^{2k+1}\right] \tag{3.41}$$

This completes the proof of Claim 4.

Claim 5. First we derive an upper bound to $\mathrm{Var}(\beta_{2k})$. Recall the expression of $\mathrm{Var}(\beta_{2k})$ in (3.37) and apply (3.41) to conclude, for some given C_3 and C_4,

$$\begin{aligned} \mathrm{Var}(\beta_{2k}) \leq \frac{1}{(mn)^{2(k+1)}} \sum_{t=1}^{2k+1} \sum_{\mathfrak{w}_3 \in \mathcal{W}_{4k,t}} & \#\Gamma(\mathfrak{w}_3) \#\{f^{-1}(\mathfrak{w}_3)\} (C_3 k)^{C_4(2k-t+1)} \times \\ & \left[\sum_{d(w')=1}^{2k-1} 2^{4k-d(w')} \#S(2k, d(w')) n^{d(w')+2} + 2^{2k} n^{2k+1}\right] \end{aligned} \tag{3.42}$$

Now $\#\Gamma(\mathfrak{w}_3)$ is bounded by m^t. Applying Lemma 3.2, Lemma 3.1 and the bound of $\#\{f^{-1}(\mathfrak{w}_3)\}$ in (3.36) simultaneously, we get the following bound to R.S. of (3.42):

$$\begin{aligned} & \frac{1}{(mn)^{2(k+1)}} \sum_{t=1}^{2k+1} 2^{4k} (4k)^{6(2k-t+1)} m^t (C_1 k)^{C_2} 2^{4k-d(w')} (C_3 k)^{C_4(2k-t+1)} \times \\ & \qquad \left[\sum_{d(w')=1}^{2k-1} \frac{1}{2} \binom{2k}{d(w')} (d(w'))^{4k-d(w')} n^{d(w')+2} + 2^{4k} n^{2k+1}\right] \\ & \leq \frac{1}{m} \sum_{t=1}^{2k+1} \sum_{d(w')=1}^{2k-1} 2^{6k} \left(\frac{C_5 k^{C_6}}{m}\right)^{2k-t+1} \left(\frac{(4k^2)}{n}\right)^{2k-d(w')} + \frac{1}{mn} \sum_{t=1}^{2k+1} 2^{6k} \left(\frac{C_5 k^{C_6}}{m}\right)^{2k-t+1} \end{aligned} \tag{3.43}$$

where C_5 and C_6 are some known constants.

Without loss of generality let us assume n to be odd. The even case is similar. Now using (3.27) one gets

$$\begin{aligned}\frac{\mathrm{Var}(\beta_{2k})}{(\mathrm{E}[\beta_{2k}])^2} &\leq n^2 \frac{1}{2^{6k+1}} \left(\frac{1}{\prod_{i=1}^{k}(1-\frac{i}{m})} \right)^2 \frac{1}{m} \sum_{t=1}^{2k+1} 2^{6k} \left(\frac{C_5 k^{C_6}}{m} \right)^{2k-t+1} \left[\sum_{d(w')=1}^{2k-1} \left(\frac{(4k^2)}{n} \right)^{2k-d(w')} + \frac{1}{n} \right] \\ &\leq n^2 \frac{1}{2^{6k+1}} \left(\frac{1}{1-\frac{k}{m}} \right)^{2k} \frac{1}{m} \sum_{t=1}^{2k+1} 2^{6k} \left(\frac{C_5 k^{C_6}}{m} \right)^{2k-t+1} \times \left[\sum_{d(w')=1}^{2k-1} \left(\frac{(4k^2)}{n} \right)^{2k-d(w')} + \frac{1}{n} \right] \\ &= \frac{n}{2m} \left(\frac{1}{1-\frac{k}{m}} \right)^{2k} \sum_{t=1}^{2k+1} \left(\frac{C_5 k^{C_6}}{m} \right)^{2k-t+1} \times \left[\sum_{d(w')=1}^{2k-1} (4k^2) \left(\frac{4k^2}{n} \right)^{2k-d(w')-1} + 1 \right] \\ &= \frac{n}{2m} \left(\frac{1}{1-\frac{k}{m}} \right)^{2k} \sum_{t=1}^{2k+1} \left(\frac{C_7 k^{C_8}}{m} \right)^{2k-t+1} \times \left[\sum_{d(w')=1}^{2k-1} \left(\frac{4k^2}{n} \right)^{2k-d(w')-1} + 1 \right]. \end{aligned} \tag{3.44}$$

Combining all these we get that

$$\frac{\mathrm{Var}(\beta_{2k})}{(\mathrm{E}[\beta_{2k}])^2} \to 0$$

if $\frac{k^2}{m} \to 0$, $\frac{k^2}{n} \to 0$, $C_7 \frac{k^{C_8}}{m} \to 0$ and $\frac{n}{m} \to 0$. Now recall from the proof of Claim 1 that if $\frac{k}{\log n} \to \infty$, we have $\mathrm{E}[\beta_{2k}]^{\frac{1}{2k}} \to 2\sqrt{2}$. We choose a sequence $k = k(m, n)$ satisfying all these properties.

Now

$$\lim_{k\to\infty} \frac{\mathrm{E}[\beta_{2k}]^{\frac{1}{2k}}}{(2\sqrt{2}-\epsilon)} \to \frac{2\sqrt{2}}{2\sqrt{2}-\epsilon} > 1.$$

So $\lim_{k\to\infty} \frac{\mathrm{E}[\beta_{2k}]}{(2\sqrt{2}-\epsilon)^{2k}} \to \infty$. As a consequence,

$$\frac{(\mathrm{E}[\beta_{2k}])^2}{\left(\mathrm{E}[\beta_{2k}] - (2\sqrt{2}-\epsilon)^{2k}\right)^2} \to 1.$$

Hence we conclude that the R.S. of (3.18) goes to 0. This completes the proof. □

References

[1] G. W. Anderson, A. Guionnet and O. Zeitouni. *An introduction to random matrices*. Cambridge University Press, 2010.

[2] G. W. Anderson. Convergence of the largest singular value of a polynomial in independent Wigner matrices. *Ann. Probab.* 41, 3B, 2103-2181, 2013.

[3] Zhidong Bai and Jack W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices, Second Edition, Springer 2006.*

[4] Z. D. Bai and J. W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* 26, 316-345, 1998.

[5] Z. D. Bai and Y. Q. Yin. Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. *Ann. Probab*. 16, 1729-1741, 1988.

[6] D. Banerjee and A. Bose. Bulk behaviour of some patterned block matrices. To appear in *Indian Jour. of Pure and Appl. Math.* Preprint available as *Tech. Report R3/2015 Stat-Math Unit, Indian Statistical Institute, Kolkata http://www.isical.ac.in/~statmath/*

[7] S. Banerjee and A. Bose. Noncrossing partitions, Catalan words and the semicircular law. *J. Theoret. Probab., 26, 386-409, 2013.*

[8] R. Basu, A. Bose, S. Ganguly, R. S. Hazra. Limiting spectral distribution of block matrices with Toeplitz block structure. *Stat. & Probab. Lett. , 82, 7, 1430-1438, 2012.*

[9] R. Bhatia. *Matrix Analysis.* Springer, New York, 1997.

[10] A. Bose and A. Sen. Another look at the moment method for large dimensional random matrices. *Electron. J. Probab.*, 13, 21, 588–628, 2008.

[11] H. Gazzah, P. A. Regalia, and J. P. Delmas. Asymptotic eigenvalue distribution of block Toeplitz matrices and application to blind SIMO channel identification. *IEEE Trans. Inform. Theory, 47, 3, 1243-1251, 2001.*

[12] U. Haagerup , H. Schultz, and S. Thorbjørnsen (2006). A random matrix approach to the lack of projections in $C^*_{\text{red}}(\mathbb{F}_2)$. *Adv. Math.* 204, 1-83, 2006.

[13] U. Haagerup and S. Thorbjørnsen (2005). A new application of random matrices: $\text{Ext}(C^*_{\text{red}}(\mathbb{F}_2))$ is not a group. *Ann. of Math*. 162, 2, 711-775, 2005.

[14] Y. T. Li, D. Z. Liu, and Z. D. Wang. Limit distributions of eigenvalues for random block Toeplitz and Hankel matrices. *J. Theoret. Probab., 24, 4, 1063-1086, 2011.*

[15] P. Loubaton. On the almost sure location of singular values of certain Gaussian block-Hankel large random matrices.. To appear in *Jour. Theoret. Probab.*. Preprint available as *http://arxiv.org/pdf/1405.2006.pdf*

[16] A. Nica and R. Speicher. *Lectures on the combinatorics of Free Probability. London Mathematical Society Lecture Note Series:335,* Cambridge University Press 2006.

[17] T. Oraby. The spectral laws of Hermitian block-matrices with large random blocks. *Electron. Comm. Probab., 12, 465-476, 2007.*

[18] R. Rashidi Far, T. Oraby, W. Bryc, and R. Speicher. On slow-fading MIMO systems with non- separable correlation. *IEEE Trans. Inform. Theory, 54, 2, 544-553, 2008.*

[19] B. C. Rennie and A. J. Dobson. On the Stirling number of second kind. *Journal of Combinatorial Theory 7,* 116–121, 1969.

[20] H. Schultz. Non–commutative polynomials of independent Gaussian random matrices. The real and symplectic cases. *Probab. Theory Related Fields* 131, 261- 309, 2005.