DOI: 10.1214/ECP.v15-1537

Elect. Comm. in Probab. **15** (2010), 134–148

ELECTRONIC COMMUNICATIONS in PROBABILITY

BALANCED RANDOM TOEPLITZ AND HANKEL MATRICES

ANIRBAN BASAK[1]
Department of Statistics, Stanford University, Stanford, CA 94305, USA
email: anirbanb@stanford.edu

ARUP BOSE[2]
Statistics and Mathematics Unit, Indian Statistical Institute, 203 B. T. Road Kolkata 700108, INDIA
email: bosearu@gmail.com

Submitted January 10, 2010, *accepted in final form* April 08, 2010

AMS 2000 Subject classification: Primary 60B20, Secondary 60F05, 62E20, 60G57, 60B10.
Keywords: Random matrix, eigenvalues, balanced matrix, moment method, bounded Lipschitz metric, Carleman condition, almost sure convergence, convergence in distribution, uniform integrability.

Abstract

Except for the Toeplitz and Hankel matrices, the common patterned matrices for which the limiting spectral distribution (LSD) are known to exist share a common property–the number of times each random variable appears in the matrix is (more or less) the same across the variables. Thus it seems natural to ask what happens to the spectrum of the Toeplitz and Hankel matrices when each entry is scaled by the square root of the number of times that entry appears in the matrix instead of the uniform scaling by $n^{-1/2}$. We show that the LSD of these balanced matrices exist and derive integral formulae for the moments of the limit distribution. Curiously, it is not clear if these moments define a unique distribution.

1 Introduction and main results

For any (random and symmetric) $n \times n$ matrix B, let $\mu_1(B), \ldots, \mu_n(B) \in \mathbb{R}$ denote its eigenvalues including multiplicities. Then the *empirical spectral distribution* (ESD) of B is the (random) distribution function on $\mathbb{R}$ given by

$$F^B(x) = n^{-1}\#\Big\{j : \mu_j(B) \in (-\infty, x],\ 1 \le j \le n\Big\}.$$

[1]SUPPORTED BY MELVIN AND JOAN LANE ENDOWED STANFORD GRADUATE FELLOWSHIP FUND
[2]RESEARCH SUPPORTED BY J.C.BOSE NATIONAL FELLOWSHIP, DEPARTMENT OF SCIENCE AND TECHNOLOGY, GOVERNMENT OF INDIA

For a sequence of random $n \times n$ matrices $\{B_n\}_{n\geq 1}$ if as $n \to \infty$, the corresponding ESDs F^{B_n} converge weakly (either almost surely or in probability) to a (nonrandom) distribution F in the space of probability measures on $\mathbb{R}$, then F is called the *limiting spectral distribution* (LSD) of $\{B_n\}_{n\geq 1}$. See Bai (1999)[1], Bose and Sen (2007)[6] and Bose, Sen and Gangopadhyay (2009)[5] for description of several interesting patterned matrices whose LSD exist. Examples include the Wigner, the circulants, the Hankel and the Toeplitz matrices. For the Wigner and circulant matrices, the number of times each random variable appears in the matrix is same across most variables. We may call them *balanced* matrices. For them, the LSDs exist after the eigenvalues are scaled by $n^{-1/2}$.

Consider the $n \times n$ symmetric Toeplitz and Hankel matrices with an i.i.d. *input* sequence $\{x_i\}$. For these matrices when scaled by $n^{-1/2}$, the LSDs exist. The limits are symmetric about 0, are non-Gaussian and have unbounded support with LSD for the Hankel matrices being not unimodal. Further the ratio of the (even) moments of the LSD to the standard Gaussian moments tend to 0 as the order increases. See Bryc, Dembo and Jiang (2006)[7] and Hammond and Miller (2005)[8]. Not much more is known about these LSDs. However, these matrices are unbalanced. It seems natural to consider the balanced versions of the Toeplitz and Hankel matrices where each entry is scaled by the square root of the number of times that entry appears in the matrix instead of the uniform scaling by $n^{-1/2}$. Define the (symmetric) balanced Hankel and Toeplitz matrices BH_n and BT_n with input $\{x_i\}$ as follows:

$$BH_n = \begin{bmatrix} \frac{x_1}{\sqrt{1}} & \frac{x_2}{\sqrt{2}} & \frac{x_3}{\sqrt{3}} & \cdots & \frac{x_{n-1}}{\sqrt{n-1}} & \frac{x_n}{\sqrt{n}} \\ \frac{x_2}{\sqrt{2}} & \frac{x_3}{\sqrt{3}} & \frac{x_4}{\sqrt{4}} & \cdots & \frac{x_n}{\sqrt{n}} & \frac{x_{n+1}}{\sqrt{n-1}} \\ \frac{x_3}{\sqrt{3}} & \frac{x_4}{\sqrt{4}} & \frac{x_5}{\sqrt{5}} & \cdots & \frac{x_{n+1}}{\sqrt{n-1}} & \frac{x_{n+2}}{\sqrt{n-2}} \\ & & & \vdots & & \\ \frac{x_n}{\sqrt{n}} & \frac{x_{n+1}}{\sqrt{n-1}} & \frac{x_{n+2}}{\sqrt{n-2}} & \cdots & \frac{x_{2n-2}}{\sqrt{2}} & \frac{x_{2n-1}}{\sqrt{1}} \end{bmatrix}. \tag{1.1}$$

$$BT_n = \begin{bmatrix} \frac{x_0}{\sqrt{n}} & \frac{x_1}{\sqrt{n-1}} & \frac{x_2}{\sqrt{n-2}} & \cdots & \frac{x_{n-2}}{\sqrt{2}} & \frac{x_{n-1}}{\sqrt{1}} \\ \frac{x_1}{\sqrt{n-1}} & \frac{x_0}{\sqrt{n}} & \frac{x_1}{\sqrt{n-1}} & \cdots & \frac{x_{n-3}}{\sqrt{3}} & \frac{x_{n-2}}{\sqrt{2}} \\ \frac{x_2}{\sqrt{n-2}} & \frac{x_1}{\sqrt{n-1}} & \frac{x_0}{\sqrt{n}} & \cdots & \frac{x_{n-4}}{\sqrt{4}} & \frac{x_{n-3}}{\sqrt{3}} \\ & & & \vdots & & \\ \frac{x_{n-1}}{\sqrt{1}} & \frac{x_{n-2}}{\sqrt{2}} & \frac{x_{n-3}}{\sqrt{3}} & \cdots & \frac{x_1}{\sqrt{n-1}} & \frac{x_0}{\sqrt{n}} \end{bmatrix}. \tag{1.2}$$

Strictly speaking BT_n is not completely balanced, with the main diagonal being unbalanced compared to the rest of the matrix. The main diagonal has all identical elements and making BT_n balanced will shift its eigenvalues by $x_0/\sqrt{2n}$ which in the limit will go to zero and hence this does not affect the asymptotic behavior of the eigenvalues. We use the above version because of the convenience in writing out the calculations later. Figures 1 and 2 exhibit the simulation results for the ESDs of the above matrices. We prove the following theorem. A primitive version of this result appears in Basak [3].

Theorem 1. *Suppose $\{x_i\}$ are i.i.d. with mean 0 and variance 1. Then almost surely the LSDs, say BT and BH of the matrices BT_n and BH_n respectively, exist and are free of the underlying distribution of the $\{x_i\}$.*

Remark Our proofs will imply that the same limit continues to hold if $\{x_i\}$ are independent, uniformly bounded with mean 0 and variance 1 and are not necessarily identically distributed.

BT and BH have unbounded support, are symmetric about zero and have all moments finite. Both LSDs are non-Gaussian. The integral formulae for the moments are given in (2.10) and (2.12) in Section 2.5 after we develop the requisite notation to write them out. It does not seem to be apparent if these moments define a distribution uniquely. Establishing further properties of the limits is a difficult problem.

The main steps in the proof may be described as follows:

(1) In Section 2.1 we first show that we may restrict attention to bounded $\{x_i\}$ and hence in the later sections, we assume this to be the case.

(2) In Sections 2.2–2.4 we develop the trace formula for moments, some related notions and results to reduce the number of terms in the trace formula. In Section 2.5 we show that the *expected* moments of the ESD of $\{BT_n\}$ converge. However, it does not seem to be straightforward to show that this limiting sequence uniquely determines a distribution. Even if it did, it is not clear how the convergence of the expected moments can be sharpened to convergence of the ESD itself. If we pull out the usual scaling $n^{-1/2}$, the scaling for the (i,j)th entry is $[1-\frac{|i-j|}{n}]^{-1/2}$ whose maximum is $n^{1/2}$. This unboundedness creates problems in the usual argument.

(3) In Section 2.6 we discuss a known approximation result.

(4) Fix any $\varepsilon > 0$. Let BT_n^{ε} denote the top-left $\lfloor n(1-\varepsilon)\rfloor \times \lfloor n(1-\varepsilon)\rfloor$ principal sub-matrix of BT_n. The Lévy distance between F^{BT_n} and $F^{BT_n^{\varepsilon}}$ is less than ε. Since these truncated matrices are well behaved, we have convergence of $F^{BT_n^{\varepsilon}}$ to a non-random distribution BT^{ε} almost surely, and also the corresponding expected moments follow by the same arguments as for the usual Toeplitz matrices. This limit is uniquely determined by its moments. This is done in Section 2.7.

(5) In Section 2.8 we show by using results derived in (3) that as $\varepsilon \to 0$, the spectral measures of BT^{ε} converge to some F^{BT} and this is the LSD of $\{BT_n\}$. We finally use a uniform integrability argument to conclude that the moments of F^{BT} are same as those obtained in (2) above.

A similar proof works for the balanced Hankel matrix by defining the truncated Hankel matrix obtained by deleting the first $\lfloor n\varepsilon/2\rfloor$ and last $\lceil n\varepsilon/2\rceil$ rows and columns from BH_n.

2 Proof of Theorem 1

2.1 Reduction to the uniform bounded case

Lemma 1. *Suppose for every bounded, mean zero and variance one i.i.d. input sequence* $\{x_0, x_1, x_2, \ldots\}$, $\{F^{BT_n}\}$ *converges to some non-random distribution* F *a.s. Then the same limit continues to hold even if the* $\{x_i\}$ *are not bounded. All the above hold for* $\{F^{BH_n}\}$ *as well.*

We make use of the *bounded Lipschitz metric*. It is defined on the space of probability measures as:

$$d_{BL}(\mu,\ \nu) = \sup\left\{\int f\,d\mu - \int f\,d\nu : ||f||_\infty + ||f||_L \le 1\right\}$$

where $||f||_\infty = \sup_x |f(x)|,\ ||f||_L = \sup_{x\neq y} |f(x)-f(y)|/|x-y|$. Recall that convergence in d_{BL} implies the weak convergence of measures and vice versa.

We also need the following fact. This fact is an estimate of the metric distance d_{BL} in terms of trace. A proof may be found in Bai and Silverstein (2006)[2] or Bai (1999)[1].

Fact 1. Suppose A, B are $n \times n$ symmetric real matrices. Then

$$d_{BL}^2(F^A, F^B) \leq \left(\frac{1}{n}\sum_{i=1}^{n}|\lambda_i(A) - \lambda_i(B)|\right)^2 \leq \frac{1}{n}\sum_{i=1}^{n}(\lambda_i(A) - \lambda_i(B))^2 \leq \frac{1}{n}\operatorname{Tr}(A-B)^2. \tag{2.1}$$

Proof of Lemma 1. For brevity, we deal with only the balanced Toeplitz case. The same arguments work for the balanced Hankel matrices. For any event A, define $I(A)$ as the indicator of the set A so that it is zero or one according as the event A does not or does happen. For $t > 0$ define

$$\mu(t) \stackrel{def}{=} E[x_0(I|x_0| \leq t)], \quad \sigma^2(t) \stackrel{def}{=} Var(x_0 I(|x_0| \leq t)) = E[x_0^2 I(|x_0| \leq t)] - \mu(t)^2,$$

$$x_i^* = \frac{x_i I(|x_i| \leq t) - \mu(t)}{\sigma(t)} = \frac{x_i - \bar{x}_i}{\sigma(t)}, \text{ where } \bar{x}_i = x_i I(|x_i| > t) + \mu(t) = x_i - \sigma(t)x_i^*.$$

Let $\{BT_n^*\}$ be the balanced Toeplitz matrix for the input sequence $\{x_i^*\}$ and $\{\widetilde{\overline{BT}}_n\}$ be the same for the input sequence $\{\bar{x}_i\}$. It is clear that $\{x_i^*\}$ is a bounded, mean zero, variance one i.i.d. sequence. Hence by our assumption, $F^{BT_n^*}$ converges to a non-random distribution function F a.s. Using Fact 1,

$$\begin{aligned} d_{BL}^2(F^{BT_n}, F^{BT_n^*}) &\leq 2d_{BL}^2(F^{BT_n}, F^{\sigma(t)BT_n^*}) + 2d_{BL}^2(F^{BT_n^*}, F^{\sigma(t)BT_n^*}) \\ &\leq \frac{2}{n}Tr[(BT_n - \sigma(t)BT_n^*)^2] + \frac{2}{n}(1-\sigma(t))^2 Tr[(BT_n^*)^2]. \end{aligned}$$

Now using the strong law of large numbers, we get

$$\begin{aligned} \frac{1}{n}Tr[(BT_n^*)^2] &= \frac{1}{n}\sum_{i,j}\left(\frac{x^*_{|i-j|}}{\sqrt{n-|i-j|}}\right)^2 \\ &= \frac{1}{n}\left(n \times \frac{x_0^{*2}}{n} + 2(n-1) \times \frac{x_1^{*2}}{(n-1)} + \cdots + 2 \times \frac{{x^*_{n-1}}^2}{1}\right) \\ &\leq \frac{2}{n}(x_0^{*2} + x_1^{*2} + \cdots + {x^*_{n-1}}^2) \stackrel{a.s.}{\to} 2E(x_0^{*2}) = 2. \end{aligned}$$

Note that $1 - \sigma(t) \to 0$ as $t \to \infty$. Similarly,

$$\begin{aligned} \frac{1}{n}Tr[(BT_n - \sigma(t)BT_n^*)^2] &= \frac{1}{n}Tr[\widetilde{\overline{BT}}_n^2] \\ &= \frac{1}{n}\sum_{i,j}\left(\frac{\bar{x}_{|i-j|}}{\sqrt{n-|i-j|}}\right)^2 \\ &= \frac{1}{n}\left(n \times \frac{\bar{x}_0^2}{n} + 2(n-1) \times \frac{\bar{x}_1^2}{(n-1)} + \cdots + 2 \times \frac{\bar{x}_{n-1}^2}{1}\right) \\ &\leq \frac{2}{n}(\bar{x}_0^2 + \cdots + \bar{x}_{n-1}^2) \stackrel{a.s.}{\to} 2E[\bar{x}_0^2] = 1 - 2\mu(t)^2 - \sigma^2(t) \to 0 \quad \text{as} \quad t \to \infty. \end{aligned}$$

Hence combining the above arguments, we get $\limsup_n d_{BL}(F^{BT_n}, F^{BT_n^*}) \to 0$ a.s. as $t \to \infty$. This completes the proof of this lemma. $\square$

2.2 Moment and trace formula

We need some notation to express the moments of the ESD in a way which is convenient for further analysis.

Circuit and vertex: A *circuit* is any function $\pi : \{0, 1, 2, \ldots, h\} \to \{1, 2, \ldots, n\}$ such that $\pi(0) = \pi(h)$. Any $\pi(i)$ is a *vertex*. A circuit depends on h and n but we will suppress this dependence.

Define two functions L^T and L^H, which we call *link functions*, by

$$L^T(i, j) = |i - j| \text{ and } L^H(i, j) = i + j - 1. \tag{2.2}$$

Also for $L = L^H$ or L^T, as the case may be, define

$$\mathbb{X}_\pi = x_{L(\pi(0),\pi(1))} x_{L(\pi(1),\pi(2))} \cdots x_{L(\pi(h-2),\pi(h-1))} x_{L(\pi(h-1),\pi(h))}.$$

Also define

$$\phi_T(i, j) = n - |i - j| \text{ and } \phi_H(i, j) = \min(i + j - 1, 2n - i - j + 1), \tag{2.3}$$

$$\phi_T^n(x, y) = \phi_T^\infty(x, y) = 1 - |x - y|, \tag{2.4}$$

$$\phi_H^n(x, y) = \min(x + y - \frac{1}{n}, 2 - x - y + \frac{1}{n}), \phi_H^\infty(x, y) = \lim_{n\to\infty} \phi_H^n(x, y). \tag{2.5}$$

Finally, for any matrix B, let $\beta_h(B)$ denote the hth moment of its ESD. Then the *trace formula* implies

$$\frac{1}{n}\operatorname{Tr}[BT_n]^h = \frac{1}{n}\sum_{1\le i_1,i_2,\ldots,i_h\le n}\left(\prod_{1\le j\le h-1}\frac{x_{L^T(i_j,i_{j+1})}}{\sqrt{\phi_T(i_j,i_{j+1})}}\right)\times\frac{x_{L^T(i_h,i_1)}}{\sqrt{\phi_T(i_h,i_1)}} \tag{2.6}$$

$$\mathrm{E}[\beta_h(BT_n)] = \mathrm{E}[\frac{1}{n}\operatorname{Tr}(BT_n)^h] = \frac{1}{n}\sum_{\pi:\ \pi \text{ circuit}}\frac{\mathrm{E}\,\mathbb{X}_\pi}{\prod_{1\le i\le h}\sqrt{\phi_T(\pi(i-1),\pi(i))}}. \tag{2.7}$$

$$\frac{1}{n}\operatorname{Tr}[BH_n]^h = \frac{1}{n}\sum_{1\le i_1,i_2,\ldots,i_h\le n}\left(\prod_{1\le j\le h-1}\frac{x_{L^H(i_j,i_{j+1})}}{\sqrt{\phi_H(i_j,i_{j+1})}}\right)\times\frac{x_{L^H(i_h,i_1)}}{\sqrt{\phi_H(i_h,i_1)}} \tag{2.8}$$

$$\mathrm{E}[\beta_h(BH_n)] = \mathrm{E}[\frac{1}{n}\operatorname{Tr}(BH_n)^h] = \frac{1}{n}\sum_{\pi:\ \pi \text{ circuit}}\frac{\mathrm{E}\,\mathbb{X}_\pi}{\prod_{1\le i\le h}\sqrt{\phi_H(\pi(i-1),\pi(i))}}. \tag{2.9}$$

Matched circuits: Any value $L(\pi(i-1), \pi(i))$ is an L value of π and π has an *edge of order* e $(1 \le e \le h)$ if it has an *L-value* repeated exactly e times. If π has at least one edge of order one then $\mathrm{E}(\mathbb{X}_\pi) = 0$. Thus only those π with all $e \ge 2$ are relevant. Such circuits will be said to be *matched*. π is *pair matched* if all its edges are of order two.

Equivalence relation on circuits: Two circuits π_1 and π_2 are equivalent iff their L-values agree at exactly the same pairs (i, j). That is, iff $\{L(\pi_1(i-1), \pi_1(i)) = L(\pi_1(j-1), \pi_1(j)) \Leftrightarrow L(\pi_2(i-1), \pi_2(i)) = L(\pi_2(j-1), \pi(j))\}$. This defines an equivalence relation between the circuits.

Words: Equivalence classes may be identified with partitions of $\{1, 2, \cdots, h\}$: to any partition we associate a *word* w of length $l(w) = h$ of letters where the first occurrence of each letter is in alphabetical order. For example, if $h = 6$, then the partition $\{\{1,3,6\},\{2,5\},\{4\}\}$ is associated with $w = abacba$. For a word w, let $w[i]$ denote the letter in the ith position. The notion of matching and order e edges carries over to words. For instance, $abacabc$ is matched. $abcadbaa$ is non-matched, has edges of order 1, 2 and 4 and the corresponding partition is $\{\{1,4,7,8\},\{2,6\},\{3\},\{5\}\}$.

Independent vertex: If $w[i]$ is the first occurrence of a letter then $\pi(i)$ is called an *independent vertex*. We make the convention that $\pi(0)$ is also an independent vertex. The other vertices will be called *dependent* vertices. If a word has d distinct letters then there are $d+1$ independent vertices.

2.3 Reduction in the number of terms

Fix an integer h. Define

$$\Pi_h^{3+} = \{\pi : \pi \text{ is matched, of length } h \text{ and has an edge of order greater than equal to } 3\}.$$

$$S_h^{A_n} = \frac{1}{n}\sum_{\pi:\pi\in\Pi_h^{3+}} \frac{1}{\prod_{i=1}^{h}\sqrt{\phi_A(\pi(i-1),\pi(i))}}, \quad A = H \text{ or } T.$$

Lemma 2. *$S_h^{A_n} \to 0$ as $n \to \infty$ for $A_n = T_n$ or H_n. Hence, only pair matched circuits are relevant while calculating* $\lim E[\beta_h(A_n)]$.

Proof. We provide the proof only for T_n. The proof for H_n is similar and the details are omitted. Note that

$$S_h^{A_n} = \sum_w \frac{1}{n}\sum_{\pi\in\Pi(w)\bigcap\Pi_h^{3+}} \frac{1}{\prod_{i=1}^{h}\sqrt{n-|\pi(i-1)-\pi(i)|}} = \sum_w S_{h,w} \text{ say.}$$

It is enough to prove that for each w, $S_{h,w} \to 0$. We first restrict attention to w which have only one edge of order 3 and all other edges of order 2. Note that this forces h to be odd. Let $h = 2t+1$ and $|w| = t$. Fix the L-values at say $k_1, k_2, \ldots, k_t$ where k_1 is the L-value corresponding to the order 3 edge and let i_0 be such that $L(\pi(i_0-1),\pi(i_0)) = k_1$. We start counting the number of possible π's from the edge $(\pi(i_0-1),\pi(i_0))$. Clearly the number of possible choices of that edge is at most $2(n-k_1)$. Having chosen the vertex i_0, the number of possible choices of the vertex (i_0+1) is at most 2. Carrying on with this argument, we may conclude that the total number of π's having L values $k_1, k_2, \ldots, k_t$ is at most $C \times (n-k_1)$. Hence for some generic constant C,

$$\begin{aligned} S_{h,w} = \frac{1}{n}\sum_{0\le k_i\le n-1}\ \sum_{\substack{\pi:\pi \text{has } L \text{ values} \\ k_1,k_2,\ldots,k_t}} \frac{1}{(n-k_1)^{\frac{3}{2}}\prod_{i=2}^{t}(n-k_i)} \quad &\le \quad \frac{1}{n}\sum_{0\le k_i\le n-1}\frac{C\times(n-k_1)}{(n-k_1)^{\frac{3}{2}}\prod_{i=2}^{t}(n-k_i)} \\ &= \quad \frac{O(\sqrt{n})O((\log n)^{t-1})}{n} \to 0 \text{ as } n\to\infty. \end{aligned}$$

In the last step, we have used the facts that $\sum_{k=1}^{n}\frac{1}{k} = O(\log n)$ and for $0 < s < 1$, $\sum_{k=1}^{n}\frac{1}{k^s} = O(n^{1-s})$. It is easy to see that when w contains more than one edge of order 3 or more, the order of the sum will be even smaller. This completes the proof of the first part. The second part is immediate since $E(X_\pi) = 1$ for every pair matched circuit and $E(|X_\pi|) < \infty$ uniformly over all π. $\square$

2.4 Slope in balanced Toeplitz matrices

Since the Toeplitz matrices have link function $L(i,j) = |i-j|$, given an L-value and a vertex there are at most two possible choices of the other vertex. Bryc, Dembo and Jiang (2006)[7] and Hammond and Miller (2005)[8] showed that out of these two possible choices of vertices only one choice counts in the limit. We show now that the same is true for the balanced matrices. Let

$$\Pi_{h,+} = \{\pi \text{ pair matched : there exists at least one pair } (i_0, j_0) \text{ with } \pi(i_0-1)-\pi(i_0)+\pi(j_0-1)-\pi(j_0) \neq 0\},$$

Define $\Pi_{h,+}(w) = \Pi_{h,+} \cap \Pi(w)$ and let $\pi(i-1) - \pi(i)$ be the i^{th} slope value.

Lemma 3. *Let $\pi \in \Pi_{h+}$ and $k_1, k_2, \ldots, k_h$ be the L-values of π. Then there exists $j_0 \in \{1, 2, \ldots, h\}$ such that $k_{j_0} = \Lambda(k_1, k_2, \ldots, k_{j_0-1}, k_{j_0+1}, \ldots, k_h)$ for some linear function Λ.*

Proof. Note that the sum of all the slope-values of π is zero. Now the sum of the slope-values from the j^{th} matched pair is 0 if the L values have opposite signs while it is $2k_j$ or $-2k_j$ if the L values have the same sign. Hence we have $f(k_1, k_2, \ldots, k_h) = 0$ for some linear function f where the coefficient of k_j equals 0 if the L values corresponding to k_j have opposite signs while it is ± 2 if the L values corresponding to k_j are of the same sign and the slope-values are positive (negative). Since $\pi \in \Pi_{h,+}$, $\exists\, k_j \neq 0$ such that the L-values corresponding to the j^{th} pair have the same sign. Let

$$\{i_1, i_2, \ldots, i_l\} = \{j : \text{ coefficient of } k_j \neq 0\} \quad \text{and} \quad j_0 = \max\{j : \text{ coefficient of } k_j \neq 0\}.$$

Then k_{j_0} can be expressed as a linear combination $k_{j_0} = \Lambda(k_{i_1}, k_{i_2}, \ldots, k_{i_l})$. □

Lemma 4. $S_{h+} \stackrel{def}{=} \frac{1}{n} \sum_{\pi \in \Pi_{h,+}} \frac{1}{\prod_{i=1}^{h} \sqrt{n - |\pi(i-1) - \pi(i)|}} \to 0$ *as $n \to \infty$. Hence, to calculate $\lim E[\beta_h(BT_n)]$ we may restrict attention to pair matched circuits where each edge has oppositely signed L-value.*

Proof. As in Lemma 2, write $S_{h+} = \sum_w S_{h+,w}$ where $S_{h+,w}$ is the sum restricted to $\pi \in \Pi_{h,+}(w)$. It is enough to show that this tends to zero for each w. Let the corresponding L values to this w be $k_1, k_2, \ldots, k_h$. Hence

$$S_{h+,w} = \frac{1}{n} \sum_{\substack{k_1, k_2, \ldots, k_h \\ \in \{0,1,2,\ldots,n-1\}}} \frac{\#\{\pi \in \Pi_{h+}(w) \text{ such that } L \text{ values of } \pi \text{ are } \{k_1, k_2, \ldots, k_h\}\}}{\prod_{i=1}^{h} (n - k_i)}.$$

For this fixed set of L values, there are at most 2^{2h} sets of slope-values. It is enough to prove the result for any one such set. Now we start counting the number of possible π's having those slope-values.

By the previous lemma there exists j_0 such that $k_{j_0} = \Lambda(k_{i_1}, k_{i_2}, \ldots, k_{i_l})$. We start counting the number of possible π from the edge corresponding to the L value k_{j_0}, say $(\pi(i_*-1), \pi(i_*))$. Clearly the number of ways to choose vertices $\pi(i_*-1)$ and $\pi(i_*)$ is $(n-k_{j_0})$. Having chosen $\pi(i_*)$, there is only one choice of $\pi(i_*+1)$ (since the slope-values have been fixed). We continue this procedure to choose all the vertices of the circuit π and hence the number of π's having the fixed set of slope-values is at most $(n - k_{j_0})$. Note that since w and the slope signs are fixed, the linear function Λ

and the index j_0 are determined as well. Thus for that fixed set we have

$$S_{h+,w}^{set} \leq \frac{1}{n} \sum_{0 \leq k_i \leq n-1} \frac{n-k_{j_0}}{\prod_{i=1}^{h}(n-k_i)} = \frac{1}{n} \sum_{\substack{0 \leq k_i \leq n-1 \\ i \neq j_0}} \frac{1}{\prod_{\substack{i=1 \\ i \neq j_0}}^{h}(n-k_i)}.$$

As $k_{j_0} = \Lambda(k_{i_1}, k_{i_2}, \ldots, k_{i_l})$, in the above sum k_{j_0} should be kept fixed which implies that $S_{h+,w} \leq \frac{O\left((\log n)^{h-1}\right)}{n} \to 0$ as $n \to \infty$, proving the first part. The second part now follows immediately. $\square$

2.5 Convergence of the moments $\mathrm{E}[\beta_h(BT_n)]$ and $\mathrm{E}[\beta_h(BH_n)]$

We need to first establish a few results on moments of a truncated uniform random variable. For a given random variable X (to be chosen), define (whenever it is finite)

$$g_T(x) = \mathrm{E}[\phi_T^n(X,x)^{-(1+\alpha)}] \text{ and } g_H(x) = \mathrm{E}[\phi_H^n(X,x)^{-(1+\alpha)}].$$

Lemma 5. *Let $x \in \mathbb{N}_n = \{1/n, 2/n, \ldots, 1\}$, $\alpha > 0$ and X be discrete uniform on $\mathbb{N}_n$. Then, for some constants C_1, C_2,*

$$\max\{g_T(x), g_H(x)\} \leq C_1 x^{-\alpha} + C_2(1-x+1/n)^{-\alpha} + 1/n.$$

Proof. Note that

$$g(x) = \frac{1}{n}\sum_{y=1}^{n} \frac{1}{[1-|x-\frac{y}{n}|]^{1+\alpha}} = \frac{1}{n}\sum_{y<j} \frac{1}{(1-\frac{j-y}{n})^{1+\alpha}} + \frac{1}{n}\sum_{y>j} \frac{1}{(1-\frac{y-j}{n})^{1+\alpha}} + \frac{1}{n},$$

where $x = j/n$ and $1 < j < n$. For $j = 1$ or n similar arguments will go through. Now,

$$\begin{aligned} \frac{1}{n}\sum_{y<j}\left(1-\frac{j-y}{n}\right)^{-(1+\alpha)} = \frac{1}{n}\sum_{t=1}^{j-1}\left(1-\frac{t}{n}\right)^{-(1+\alpha)} &= n^{\alpha} \sum_{t=n-j+1}^{n-1} t^{-(1+\alpha)} \\ &\leq n^{\alpha} \times \frac{C_1}{(n-j+1)^{\alpha}} = C_1(1-x+1/n)^{-\alpha}. \end{aligned}$$

By similar arguments, $\frac{1}{n}\sum\limits_{y>j} \frac{1}{(1-\frac{y-j}{n})^{1+\alpha}} \leq C_2 x^{-\alpha}$.

By similar calculations $g_H(x) \leq C_1 x^{-\alpha} + C_2(1-x+1/n)^{-\alpha} + 1/n$ and thus the result follows. $\square$

Lemma 6. *Suppose $U_{i,n}$ are i.i.d. discrete uniform on $\mathbb{N}_n$. Let $a_i \in \mathbb{Z}, 1 \leq i \leq m$ be fixed and $0 < \beta < 1$. Let $Y_n = \sum\limits_{i=1}^{m} a_i U_{i,n}$ and $Z_n = 1 - Y_n + 1/n$. Then*

$$\sup_n \mathrm{E}[|Y_n|^{-\beta} I(|Y_n| \geq 1/n)] + \sup_n \mathrm{E}[|Z_n|^{-\beta} I(|Z_n| \geq 1/n)] < \infty.$$

Proof. First note that

$$\begin{aligned}\mathbb{P}(|Y_n| \le M/n) &= \mathrm{E}\Big[\mathbb{P}\Big(-M/n \le \sum_{i=1}^{m} a_i U_{i,n} \le M/n \Big| U_{i,n}, j \ne i_0\Big)\Big] \\ &= \mathrm{E}\Big[\mathbb{P}\Big(-M/n - \sum_{\substack{i=1\\ i\ne i_0}}^{m} a_i U_{i,n} \le a_{i_0} U_{i_0,n} \le M/n - \sum_{\substack{i=1\\ i\ne i_0}}^{m} a_i U_{i,n}\Big)\Big] \le (2M+1)/n.\end{aligned}$$

Let $U_1, U_2, \dots, U_m$ be m i.i.d $U(0,1)$ random variables. We note that

$$\Big(U_{1,n}, U_{2,n}, \dots, U_{m,n}\Big) \stackrel{\mathscr{D}}{=} \Big(\frac{\lceil nU_1 \rceil}{n}, \frac{\lceil nU_2 \rceil}{n}, \dots, \frac{\lceil nU_m \rceil}{n}\Big).$$

Define

$$\hat{Y}_n = \sum_{i=1}^{m} a_i \frac{\lceil nU_i \rceil}{n} \;\;,\;\; Y = \sum_{i=1}^{m} a_i U_i \;\text{ and }\; K = \sum_{i=1}^{m} |a_i|.$$

Then

$$\hat{Y}_n \stackrel{\mathscr{D}}{=} Y_n \;\text{ and }\; |\hat{Y}_n - Y| \le K/n.$$

$$\begin{aligned}\mathrm{E}[|Y_n|^{-\beta} I(|Y_n| \ge 1/n)] &= \mathrm{E}[|Y_n|^{-\beta} I(1/n \le |Y_n| \le 2K/n)] + \mathrm{E}[|\hat{Y}_n|^{-\beta} I(|\hat{Y}_n| > 2K/n)] \\ &\le n^{\beta} \frac{4K+1}{n} + \mathrm{E}[(|Y| - K/n)^{-\beta} I(|\hat{Y}_n| > 2K/n)] \\ &\le o(1) + \mathrm{E}[(|Y| - K/n)^{-\beta} I(|Y| > K/n)] \\ &\le o(1) + \int_{x > K/n} (x - K/n)^{-\beta} f_{|Y|}(x) dx.\end{aligned}$$

Now

$$\int_{x>K/n} (x - K/n)^{-\beta} f_{|Y|}(x) dx = \int_0^{\infty} x^{-\beta} f_{|Y|}(x + K/n) dx.$$

It is easy to see that f_Y vanishes outside $[-K, K]$. Using induction one can also prove that $f_Y(x) \le 1$ for all x. These two facts yield

$$\int_0^{\infty} x^{-\beta} f_{|Y|}(x + K/n) dx \le \int_0^{K+K/n} x^{-\beta} 2 dx = O(1).$$

Hence

$$\sup_n \mathrm{E}[|Y_n|^{-\beta} I(|Y_n| \ge 1/n)] < \infty.$$

The proof of the finiteness of the other supremum is similar and we omit the details. $\square$

Lemma 7. *Suppose $\{x_i\}$ are i.i.d. bounded with mean zero and variance 1. Then* $\lim \mathrm{E}[\beta_h(BT_n)]$ *and* $\mathrm{E}[\beta_h(BH_n)]$ *exist for every h.*

Proof. From Lemma 2 it follows that $\mathrm{E}[\beta_{2k+1}(BA_n)] \to 0$ as $n \to \infty$ where $A_n = T_n$ or H_n. From Lemma 2 and Lemma 4 (if limit exists) we have

$$\begin{aligned}\lim_{n\to\infty} \mathrm{E}[\beta_{2k}(BT_n)] &= \sum_{w \text{ pair matched}} \lim_{n\to\infty} \frac{1}{n} \sum_{\pi \in \Pi^*(w)} \frac{\mathrm{E}\,\mathbb{X}_\pi}{\prod_{i=1}^{h} \sqrt{\phi_T(\pi(i-1), \pi(i))}} \\ &= \sum_{w \text{ pair matched}} \lim_{n\to\infty} \frac{1}{n} \sum_{\pi \in \Pi^{**}(w)} \frac{1}{\prod_{i=1}^{h} \sqrt{\phi_T(\pi(i-1), \pi(i))}}\end{aligned}$$

where $\Pi^{**}(w) = \{\pi : \ w[i] = w[j] \Rightarrow \pi(i-1) - \pi(i) + \pi(j-1) - \pi(j) = 0\}$. Denote $x_i = \pi(i)/n$. Let $S = \{0\} \cup \{\min(i,j) : w[i] = w[j], i \neq j\}$ be the set of all independent vertices of the word w and let $\max S$ be the maximum of the elements present in S. Define $x_S = \{x_i : i \in S\}$. Each x_i can be expressed as a unique linear combination $L_i^T(x_S)$. L_i^T depends on the word w but for notational convenience we suppress its dependence. Note that $L_i^T(x_S) = x_i$ for $i \in S$ and also summing k equations we get $L_{2k}^T(x_S) = x_0$. If $w[i] = w[j]$ then $|L_{i-1}^T(x_S) - L_i^T(x_S)| = |L_{j-1}^T(x_S) - L_j^T(x_S)|$. Thus using this equality and proceeding as in Bose and Sen [6] and Bryc, Dembo and Jiang [7] we have,

$$\lim_{n\to\infty} \mathrm{E}[\beta_{2k}(BT_n)] = \sum_{w \text{ pair matched}} \lim_{n\to\infty} \mathrm{E}\left[\frac{\mathbb{I}(L_i^T(U_{n,S}) \in \mathbb{N}_n, i \notin S \cup \{2k\})}{\prod\limits_{i\in S\setminus\{0\}} \phi_T^n(L_{i-1}(U_{n,S}), U_i))}\right],$$

where for each $i \in S$, $U_{n,i}$ is discrete uniform on $\mathbb{N}_n$ and $U_{n,S}$ is the random vector on $\mathbb{R}^{k+1}$ whose co-ordinates are $U_{n,i}$ and $U_{n,i}$'s are independent of each other. We claim that

$$\lim_{n\to\infty} \mathrm{E}[\beta_{2k}(BT_n)] = m_{2k}^T = \sum_{w \text{ pair matched}} m_{2k,w}^T = \sum_{w \text{ pair matched}} \mathrm{E}\left[\frac{\mathbb{I}(L_i^T(U_S) \in (0,1), i \notin S \cup \{2k\})}{\prod\limits_{i\in S\setminus\{0\}} \phi_T^\infty(L_{i-1}(U_S), U_i))}\right], \tag{2.10}$$

where for each $i \in S$, $U_i \sim U(0,1)$ and U_S is an $\mathbb{R}^{k+1}$ dimensional random vector whose co-ordinates are U_i and they are independent of each other. Note that to prove (2.10) it is enough to show that for each pair matched word w and for each k there exists $\alpha_k > 0$ such that

$$\sup_n \mathrm{E}\left[\left(\frac{\mathbb{I}(L_i^T(U_{n,S}) \in \mathbb{N}_n, i \notin S \cup \{2k\})}{\prod\limits_{i\in S\setminus\{0\}} \phi_T^n(L_{i-1}(U_{n,S}), U_i))}\right)^{1+\alpha_k}\right] < \infty. \tag{2.11}$$

We will prove that for each pair matched word w

$$\sup_n \mathrm{E}\left[\left(\frac{\mathbb{I}(L_i^T(U_{n,S}) \in \mathbb{N}_n, i \notin S \cup \{2k\}, i < \max S)}{\prod\limits_{i\in S\setminus\{0\}} \phi_T^n(L_{i-1}(U_{n,S}), U_{n,i}))}\right)^{1+\alpha_k}\right] < \infty.$$

We prove the above by induction on k. For $k = 1$ the expression reduces to $\mathrm{E}\left[\left(\frac{1}{1-|U_{n,0}-U_{n,1}|}\right)^{1+\alpha}\right]$. Now

$$\begin{aligned}
\mathrm{E}\left[\left(\frac{1}{1-|U_{n,0}-U_{n,1}|}\right)^{1+\alpha}\right] &= \mathrm{E}\left[\mathrm{E}\left\{\left(\frac{1}{1-|U_{n,0}-U_{n,1}|}\right)^{1+\alpha}\Big| U_{n,0}\right\}\right] \\
&= \mathrm{E}[g_T(U_{n,0})] \leq C_1 \mathrm{E}[U_{n,0}^{-\alpha}] + C_2 \mathrm{E}[(1-U_{n,0})^{-\alpha}] + 1/n \ \text{ by Lemma 5.}
\end{aligned}$$

Hence by Lemma 6 we have $\sup_n \mathrm{E}\left[\left(\frac{1}{1-|U_{n,0}-U_{n,1}|}\right)^{1+\alpha}\right] < \infty$ for all $0 < \alpha < 1$.

Suppose that the result is true for $k = 1, 2, \ldots, t$. We then show that it is true for $k = t+1$. Fix any pair matched word w_0. Note that the random variable corresponding to the generating vertex of the last letter appears only once and hence we can do the following calculations. Let

$$B_{t+1} = \left[\frac{I\left(L_i^T(U_{n,S}) \in \mathbb{N}_n, i \notin S \cup \{2(t+1)\}, i < \max S\right)}{\prod\limits_{i\in S\setminus\{0\}} (1 - |L_{i-1}^T(U_{n,S}) - U_{n,i}|)}\right]^{1+\alpha}.$$

Then

$$\begin{aligned}
\mathrm{E}[B_{t+1}] &= \mathrm{E}\Big[\,\mathrm{E}[B_{t+1}\Big|U_{n,i}, i\in S\setminus\{i_{t+1}\}]\Big] \\
&= \mathrm{E}\Bigg[\underbrace{\Bigg(\frac{I\Big(L_i^T(U_{n,S})\in\mathbb{N}_n, i\notin S, i<\max S\setminus\{i_{t+1}\}\Big)}{\prod\limits_{i\in S\setminus\{0,i_{t+1}\}}(1-|L_{i-1}^T(U_{n,S})-U_{n,i}|)}\Bigg)^{1+\alpha}}_{\Phi_n} \\
&\qquad\qquad \times \underbrace{g_T(U_{n,i_{t+1}-1})I[L_{i_{t+1}-1}^T(U_{n,S})\in\mathbb{N}_n]}_{\Psi_n}\Bigg].
\end{aligned}$$

Letting $||\cdot||_q$ denote the L_q norm, by Lemma 5 and Lemma 6, $\sup\limits_n ||\Psi_n||_q < \infty$ whenever $\alpha q < 1$.

Let us now consider the word w_0^* obtained from w_0 by removing both occurrences of the last used letter. We note that the quantity Φ_n is the candidate for the expectation expression corresponding to the word w_0^*. Now by the induction hypothesis, there exists an $\alpha_t > 0$ such that

$$\sup_n \mathrm{E}\Bigg[\Bigg(\frac{I\Big(L_i^T(U_{n,S})\in\mathbb{N}_n, i\notin S, i<\max S\setminus\{i_{t+1}\}\Big)}{\prod\limits_{i\in S\setminus\{0,i_{t+1}\}}(1-|L_{i_1}^T(U_{n,S})-U_{n,i}|)}\Bigg)^{1+\alpha_t}\Bigg] < \infty.$$

Hence $\sup\limits_n ||\Phi_n||_p < \infty$ if $(1+\alpha)p \le (1+\alpha_t)$. Therefore

$$\alpha_{t+1} + \frac{1+\alpha_{t+1}}{1+\alpha_t} < \frac{1}{p} + \frac{1}{q} = 1 \Rightarrow \sup_n \mathrm{E}[\Phi_n\Psi_n] \le \sup_n ||\Phi_n||_p ||\Psi_n||_q < \infty.$$

This proves the claim for balanced Toeplitz matrices. For balanced Hankel matrices we again use Lemma 5 and Lemma 6 and proceed in an exactly similar way to get

$$\lim_{n\to\infty}\mathrm{E}[\beta_{2k}(BH_n)] = m_{2k}^H = \sum_{\substack{w \text{ pair matched}\\ \text{and symmetric}}} m_{2k,w}^H = \sum_{\substack{w \text{ pair matched}\\ \text{and symmetric}}} \mathrm{E}\left[\frac{\mathbb{I}(L_i^H(U_S)\in(0,1), i\notin S\cup\{2k\}}{\prod\limits_{i\in S\setminus\{0\}}\phi_H^\infty(L_{i-1}(U_S),U_i))}\right]. \quad (2.12)$$

It may be noted that the above sum is over *symmetric* pari matched words. These are words in which every letter appears once each in an odd position and an even position. Using ideas of Bose and Sen (2008) [6] and Bryc, Dembo and Jiang [7] it can be shown that for any pair matched non-symmetric word w, $m_{2k,w}^H = 0$. So the above summation is taken over only pair matched symmetric words. □

2.6 An approximation result

Even though the limit of the moments have been established, it does not seem to be easy to show that this moment sequence determines a probability distribution uniquely (which would then be the candidate LSD). We tackle this issue by using approximating matrices whose scalings are not unbounded. We shall use the *Lévy metric* to develop this approximation. Recall that this metric metrizes weak convergence of probability measures on $\mathbb{R}$. Let μ_i, $i = 1, 2$ be two probability measures on $\mathbb{R}$. The Lévy distance between them is given by

$$\rho(\mu_1, \mu_2) = \inf\{\varepsilon > 0 : \ F_1(x - \varepsilon) - \varepsilon < F_2(x) < F_1(x + \varepsilon) + \varepsilon, \ \forall \ x \in \mathbb{R}\},$$

where F_i, $i = 1, 2$, are the distribution functions corresponding to the measures $\mu_i, i = 1, 2$.

Proposition 1. *(Bhamidi, Evans and Sen (2009)) Suppose $A_{n\times n}$ is a real symmetric matrix and $B_{m\times m}$ is a principal sub-matrix of $A_{n\times n}$. Then*

$$\rho(F^A, F^B) \leq \min\left(\frac{n}{m} - 1, 1\right).$$

Let $(A_k)_{k=1}^{\infty}$ be a sequence of $n_k \times n_k$ real symmetric matrices. For each $\varepsilon > 0$ and each k, let $(B_k^{\varepsilon})_{k=1}^{\infty}$ be an $n_k^{\varepsilon} \times n_k^{\varepsilon}$ principal sub-matrix of A_k. Suppose that for each $\varepsilon > 0$, $F_{\infty}^{\varepsilon} = \lim_{k\to\infty} F^{B_k^{\varepsilon}}$ exists and $\limsup_{k\to\infty} n_k/n_k^{\varepsilon} \leq 1 + \varepsilon$. Then $F_{\infty} = \lim_{k\to\infty} F^{A_k}$ exists and is given by $F_{\infty} = \lim_{\varepsilon \downarrow 0} F_{\infty}^{\varepsilon}$.

Consider the principal submatrix BT_n^{ε} of BT_n obtained by retaining the first $\lfloor n(1-\varepsilon)\rfloor$ rows and columns of BT_n. Then for this matrix, since $|i - j| \leq \lfloor n(1-\varepsilon)\rfloor$, the balancing factor becomes bounded. We shall show that LSD of $\{F^{BT_n^{\varepsilon}}\}$ exists for every ε and then invoke the above result to obtain the LSD of $\{F^{BT_n}\}$. A similar argument holds for $\{BH_n\}$, by considering the principal sub-matrix obtained by removing the first $\lfloor n\varepsilon/2\rfloor$ and last $\lceil n\varepsilon/2\rceil$ rows and columns.

2.7 Existence of limit of $\{F^{BT_n^{\varepsilon}}\}$ and $\{F^{BH_n^{\varepsilon}}\}$ almost surely

Clearly, for any fixed $\varepsilon > 0$, we may write

$$\frac{1}{n}\operatorname{Tr}[BT_n^{\varepsilon}]^h = \frac{1}{n}\sum_{\pi:\pi \text{ circuit}} \frac{\mathbb{X}_{\pi}}{\prod\limits_{1\leq i\leq h}\sqrt{n - |\pi(i-1) - \pi(i)|}} \times \prod_{1\leq i\leq h}\mathbb{I}\Big[\pi(i) \leq \lfloor n(1-\varepsilon)\rfloor\Big] \tag{2.13}$$

and similarly

$$\frac{1}{n}\operatorname{Tr}[BH_n^{\varepsilon}]^h = \frac{1}{n}\sum_{\pi:\pi \text{ circuit}} \frac{\mathbb{X}_{\pi}}{\prod\limits_{1\leq i\leq h}\phi_H(\pi(i-1), \pi(i))} \times \prod_{1\leq i\leq h}\mathbb{I}\Big[\lfloor n\varepsilon/2\rfloor \leq \pi(i) \leq \lfloor n(1-\varepsilon/2)\rfloor\Big]. \tag{2.14}$$

Since for every $\varepsilon > 0$ the scaling is bounded, the proof of the following lemma is exactly the same as the proof of Lemma 1 and Theorem 6 of Bose and Sen (2008) [6]. Hence we skip the proof.

Recall that a symmetric word is a pair-matched word where each letter appears exactly once each in an odd and an even position.

Lemma 8. *(i) If h is odd, $\mathrm{E}[\beta_h(BT_n^{\varepsilon})] \to 0$ and $\mathrm{E}[\beta_h(BH_n^{\varepsilon})] \to 0$.*

(ii) If h is even $(= 2k)$, then

$$\lim_{n\to\infty}\mathrm{E}[\beta_h(BT_n^{\varepsilon})] = \sum_w p_{BT^{\varepsilon}}(w) = \sum_w \underbrace{\int_0^{1-\varepsilon}\cdots\int_0^{1-\varepsilon}}_{k+1} \frac{\prod\limits_{i\notin S\cup\{2k\}}\mathbb{I}(0 \leq L_i^T(x_S) \leq 1-\varepsilon)}{\prod\limits_{i\in S\setminus\{0\}}(1 - |L_{i-1}^T(x_S) - x_i|)}\, dx_S = \beta_{2k}^{T^{\varepsilon}} \ \textit{say}$$

where the sum is over pair-matched words w. Similarly,

$$\mathrm{E}[\beta_h(BH_n^{\varepsilon})] \to \sum_w p_{BH^{\varepsilon}}(w) = \sum_w \underbrace{\int_{\frac{\varepsilon}{2}}^{1-\frac{\varepsilon}{2}} \cdots \int_{\frac{\varepsilon}{2}}^{1-\frac{\varepsilon}{2}}}_{k+1} \frac{\prod_{i \notin S \cup \{2k\}} \mathbb{I}(\frac{\varepsilon}{2} \le L_i^H(x_S) \le 1 - \frac{\varepsilon}{2})}{\prod_{i \in S \setminus \{0\}} \phi_H^{\infty}(L_{i-1}^H(x_S), x_i)} dx_S = \beta_{2k}^{H^{\varepsilon}} \text{ say}$$

where the sum is over all symmetric pair-matched words w. Further, $\max\{\beta_{2k}^{T^{\varepsilon}}, \beta_{2k}^{H^{\varepsilon}}\} \le \frac{2k!}{k!2^k} \times \varepsilon^{-k}$. *Hence there exists unique probability distributions $F^{T^{\varepsilon}}$ and $F^{H^{\varepsilon}}$ with $\beta_k^{H^{\varepsilon}}$ and $\beta_k^{H^{\varepsilon}}$ (respectively) as their moments.*

The almost sure convergence of $\{F^{BT_n^{\varepsilon}}\}$ and $\{F^{BH_n^{\varepsilon}}\}$ now follow from the following Lemma. We omit its proof since it is essentially a repetition of arguments given in the proofs of Propositions 4.3 and 4.9 of Bryc, Dembo and Jiang [7] who established it for the usual Toeplitz matrix T_n and the usual Hankel matrix H_n.

Lemma 9. *Fix any $\varepsilon > 0$ and let $A_n = T_n$ or H_n. If the input sequence is uniformly bounded, independent, with mean zero and variance one then*

$$\mathrm{E}\left[\frac{1}{n}\operatorname{Tr}(BA_n^{\varepsilon})^h - \mathrm{E}\,\frac{1}{n}\operatorname{Tr}(BA_n^{\varepsilon})^h\right]^4 = O\Big(\frac{1}{n^2}\Big). \tag{2.15}$$

As a consequence, the ESD of BA_n^{ε} converges to $F^{A^{\varepsilon}}$ almost surely.

2.8 Connecting limits of $\{BT_n^{\varepsilon}\}$ (resp. BH_n^{ε}) and $\{BT_n\}$ (resp. BH_n)

From Lemma 8 and Lemma 9, given any $\varepsilon > 0$, there exists B_{ε} such that $\mathbb{P}(B_{\varepsilon}) = 1$ and on B_{ε}, $F^{BT_n^{\varepsilon}} \Rightarrow F^{T^{\varepsilon}}$.

Fix any sequence $\{\varepsilon_m\}_{m=1}^{\infty}$ decreasing to 0. Define $B = \cap B_{\varepsilon_m}$. Using Proposition 1, on B, $F^{BT_n} \Rightarrow F^T$ for some non-random distribution function F^T where F^T is the weak limit of $\{F^{T^{\varepsilon_m}}\}_{m=1}^{\infty}$.

Let X^{ε_m} (resp. X) be a random variable with distribution $F^{T^{\varepsilon_m}}$ (resp. F^T) with kth moments $\beta_k^{T^{\varepsilon_m}}$ (resp. β_k^T). From Lemma 8, and (2.10) it is clear that for all $k \ge 1$,

$$\lim_{m\to\infty} \beta_{2k+1}^{T^{\varepsilon_m}} = 0 \text{ and } \lim_{m\to\infty} \beta_{2k}^{T^{\varepsilon_m}} = m_{2k}^T = \sum_{w \text{ pair matched}} m_{2k,w}^T.$$

From Lemma 7, m_{2k}^T is finite for every k. Hence $\{(X^{\varepsilon_m})^k\}_{m=1}^{\infty}$ is uniformly integrable for every k and $\lim_{m\to\infty} \beta_k^{T^{\varepsilon_m}} = \beta_k^T$. This proves that $m_k^T = \beta_k^T$ and so $\{m_k^T\}$ are the moments of F^T. The argument for BH_n is exactly same and hence details are omitted. The proof of Theorem 1 is now complete. □

Acknowledgement. We learnt the idea of the truncation that we have used from personal communication with Arnab Sen, who credits them to Steven Evans. We are grateful to Steven Miller for his detailed comments on the earlier version of this article. We also thank the Referee and the Associate Editor for their detailed, thoughtful and constructive comments.

References

[1] Bai, Z. D. (1999). Methodologies in spectral analysis of large dimensional random matrices, a review. *Statistica Sinica* 9, 611–677 (with discussions). MR1711663

[2] Bai, Z. D. and Silverstein, J. (2006). *Spectral Analysis of Large Dimensional Random Matrices.* Science Press, Beijing.

[3] Basak, Anirban (2009). Large dimensional random matrices. *M.Stat. Project Report*, May 2009. Indian Statitstical Institute, Kolkata.

[4] Bhamidi, Shankar; Evans, Steven N. and Sen, Arnab (2009). Spectra of large random Trees. Available at *http://front.math.ucdavis.edu/0903.3589*.

[5] Bose, Arup; Gangopadhyay, Sreela and Sen, Arnab (2009). Limiting spectral distribution of XX' matrices. *Annales de l'Institut Henri Poincaré*. To appear. Currently available at http://imstat.org/aihp/accepted.html.

[6] Bose, Arup and Sen, Arnab (2008). Another look at the moment method for large dimensional random matrices. *Electronic Journal of Probability*, 13, 588–628. MR2399292

[7] Bryc, Włodzimierz; Dembo, Amir and Jiang, Tiefeng (2006). Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.*, 34, no. 1, 1–38. Also available at *http://arxiv.org/abs/math.PR/0307330* MR2206341

[8] Hammond, C. and Miller, S. J. (2005). Distribution of eigenvalues for the ensemble of real symmetric Toeplitz matrices. *J. Theoret. Probab.* 18, no. 3, 537–566. MR2167641

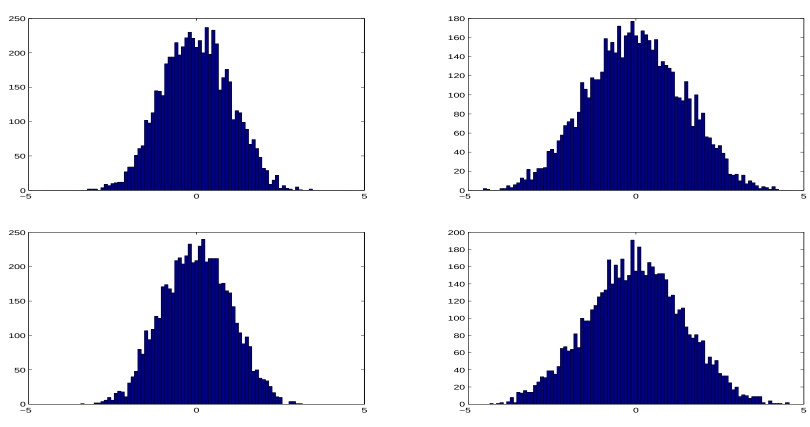

Figure 1: Histograms of the ESD of 15 realizations of the Toeplitz matrix (left) and the balanced Toeplitz matrix (right) of order 400 with standardized Normal(0, 1) (`top row`), and Bernoulli(0.5) (`bottom row`) entries.

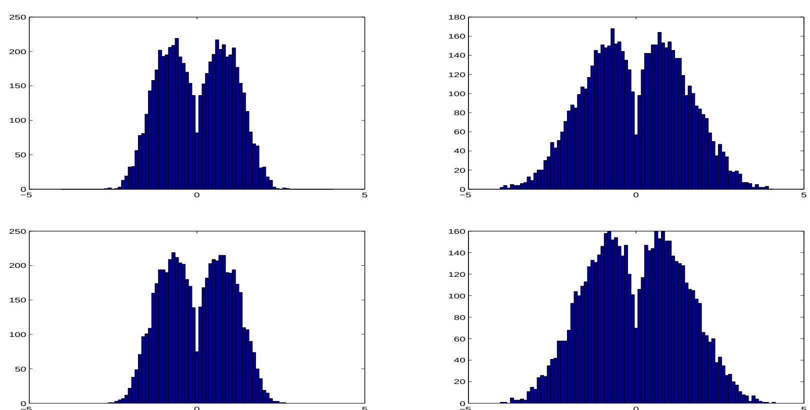

Figure 2: Histograms of the ESD of 15 realizations of the Hankel matrix (left) and the balanced Hankel matrix (right) of order 400 with standardized Normal$(0,1)$ (`top row`), and Bernoulli(0.5) (`bottom row`) entries.