LIMITING SPECTRAL DISTRIBUTION OF SOME BAND MATRICES

ANIRBAN BASAK
Stanford Graduate Fellow
Department of Statistics
Stanford University
USA

ARUP BOSE*
Statistics and Mathematics Unit
Indian Statistical Institute
203 B. T. Road, Kolkata 700108
INDIA

Original technical report, September 11, 2009. This version, November 04, 2009

Abstract

We use the method of moments to establish the limiting spectral distribution (LSD) of appropriately scaled large dimensional random symmetric circulant, reverse circulant, Toeplitz and Hankel matrices which have suitable band structures. The input sequence used to construct these matrices is assumed to be either i.i.d. with mean zero and variance one or independent and appropriate finite fourth moment. The class of LSD includes the normal and the symmetrized square root of chi square with two degrees of freedom. In several other cases, explicit forms of the limit do not seem to be obtainable but the limits can be shown to be symmetric and their second and the fourth moments can be calculated with some effort. Simulations suggest some further properties of the limits.

Keywords: Large dimensional random matrix, eigenvalues, Toeplitz matrix, Hankel matrix, circulant matrix, reverse circulant matrix, band matrix, empirical spectral distribution, bounded Lipschitz metric, limiting spectral distribution, moment method.

AMS Subject Classification: 60F05, 60F15, 62E20, 60G57.

1 INTRODUCTION

Let $\{x_0, x_1, \ldots\}$ be random variables, called the *input sequence*. Let $A_n = ((x_{L(i,j)}))$ be a sequence of $n \times n$ *patterned* random matrices where L is a *link function*. Examples of link function are: the Toeplitz link function $L(i,j) = |i-j|$, the Hankel link function $L(i,j) = i + j$ and the reverse circulant link function $L(i,j) = i + j \mod n$. Note that L may depend on n but we suppress this in our notation.

Let $F^{a_n A_n}$ be the *empirical spectral distribution* (ESD) which puts mass $1/n$ at each of the eigenvalues of the scaled matrix $a_n A_n$. With appropriate a_n, under suitable moment and independence conditions on $\{x_i\}$, the *limiting spectral distribution* (LSD) of $F^{a_n A_n}$ exists for several random matrices. That is, in each of these cases, there exists some nonrandom distribution function F such that $F^{a_n A_n}$ converges to F weakly (almost surely or in probability). See for example Bai (1999)[1], Bose and Sen (2008)[6] and Bose, Gangopadhyay and Sen (2009)[4].

Band matrices are usually defined as matrices where the top right corner and the bottom left corner elements are zeroes. With increasing dimension, the band of non zero elements around the main diagonal

*Research supported by J.C.Bose Fellowship, Department of Science and Technology, Government of India. Part of the work was done while the author was visiting Dept. of Economics, University of Cincinnati.

may also be assumed to be increasing. Banding may change the LSD of the original matrix drastically. See for example Popescu (2009)[13] for some interesting limits for tridiagonal Wigner matrices. We study the LSD of the following matrices with suitable banding.

1. Symmetric circulant matrtix SC_n. Link function $L_{SC}(i,j) = n/2 - |n/2 - |i-j||$.

$$SC_n = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_2 & x_1 \\ x_1 & x_0 & x_1 & \dots & x_3 & x_2 \\ x_2 & x_1 & x_0 & \dots & x_4 & x_3 \\ & & & \vdots & & \\ x_1 & x_2 & x_3 & \dots & x_1 & x_0 \end{bmatrix}. \tag{1.1}$$

2. Reverse circulant matrix, RC_n. Link function $L_{RC}(i,j) = i + j \mod n$.

$$RC_n = \begin{bmatrix} x_2 & x_3 & x_4 & \dots & x_0 & x_1 \\ x_3 & x_4 & x_5 & \dots & x_1 & x_2 \\ x_4 & x_5 & x_6 & \dots & x_2 & x_3 \\ & & & \vdots & & \\ x_1 & x_2 & x_3 & \dots & x_{n-1} & x_0 \end{bmatrix}. \tag{1.2}$$

3. (Symmetric) Toeplitz matrix, T_n. Link function $L_T(i,j) = |i-j|$.

$$T_n = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_{n-2} & x_{n-1} \\ x_1 & x_0 & x_1 & \dots & x_{n-3} & x_{n-2} \\ x_2 & x_1 & x_0 & \dots & x_{n-4} & x_{n-3} \\ & & & \vdots & & \\ x_{n-1} & x_{n-2} & x_{n-3} & \dots & x_1 & x_0 \end{bmatrix}. \tag{1.3}$$

4. (Symmetric) Hankel matrix, H_n. Link function $L_H(i,j) = i + j$.

$$H_n = \begin{bmatrix} x_2 & x_3 & x_4 & \dots & x_n & x_{n+1} \\ x_3 & x_4 & x_5 & \dots & x_{n+1} & x_{n+2} \\ x_4 & x_5 & x_6 & \dots & x_{n+2} & x_{n+3} \\ & & & \vdots & & \\ x_{n+1} & x_{n+2} & x_{n+3} & \dots & x_{2n-1} & x_{2n} \end{bmatrix}. \tag{1.4}$$

5. Palindromic matrices. These may be defined as symmetric matrices where the first row is a palindrome. We will consider Palindromic Toeplitz and Palindromic Hankel matrices.

(a) Palindromic Toeplitz matrix, PT_n. This matrix was introduced by Massey, Miller and Sinsheimer (2006) [MMS] [10]. It may be noted that the $n \times n$ principal minor of PT_{n+1} is SC_n.

$$PT_n = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_2 & x_1 & x_0 \\ x_1 & x_0 & x_1 & \dots & x_3 & x_2 & x_1 \\ x_2 & x_1 & x_0 & \dots & x_4 & x_3 & x_2 \\ & & & \vdots & & & \\ x_1 & x_2 & x_3 & \dots & x_1 & x_0 & x_1 \\ x_0 & x_1 & x_2 & \dots & x_2 & x_1 & x_0 \end{bmatrix}. \tag{1.5}$$

(b) Palindromic Hankel matrix, PH_n. We follow the definition of [MMS] [10], who use a slightly different indexing:

$$PH_n = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_2 & x_1 & x_0 \\ x_1 & x_2 & x_3 & \dots & x_1 & x_0 & x_1 \\ x_2 & x_3 & x_4 & \dots & x_0 & x_1 & x_2 \\ & & & \vdots & & & \\ x_1 & x_0 & x_1 & \dots & x_3 & x_2 & x_1 \\ x_0 & x_1 & x_2 & \dots & x_2 & x_1 & x_0 \end{bmatrix}. \tag{1.6}$$

6. Doubly symmetric Hankel matrix, DH_n. The symmetric circulant with link function $n/2-|n/2-|i-j||$ may also be considered as a "doubly symmetric" Toeplitz matrix. Likewise we may define the doubly symmetric Hankel matrix DH_n with link function $L_{DH}(i,j) = n/2 - |n/2 - (i+j) \bmod n|,\ 0 \le i, j \le n-1$.

$$DH_n = \begin{bmatrix} x_0 & x_1 & x_2 & \dots & x_3 & x_2 & x_1 \\ x_1 & x_2 & x_3 & \dots & x_2 & x_1 & x_0 \\ x_2 & x_3 & x_4 & \dots & x_1 & x_0 & x_1 \\ & & & \vdots & & & \\ x_2 & x_1 & x_0 & \dots & x_5 & x_4 & x_3 \\ x_1 & x_0 & x_1 & \dots & x_4 & x_3 & x_2 \end{bmatrix}. \tag{1.7}$$

We use two types of banding. Let $m_n \to \infty$ be a sequence of integers. For notational simplicity we write m for m_n.

(i) **Type I banding**. For any of the above matrices, say A_n, the Type I band matrix A_n^b is the matrix A_n but with input $\{x_i^*\}$ where,

$$x_i^* = \begin{cases} x_i & \text{if } i \le m, \\ 0 & \text{otherwise}. \end{cases} \tag{1.8}$$

(ii) **Type II banding**. It seems natural also to define the following band versions of H_n, T_n and RC_n. The Type II band version H_n^B of H_n is defined with the input sequence $\{\hat{x}_i\}$ where,

$$\hat{x}_i = \begin{cases} x_i & \text{if } n-m \le i \le n+m, \\ 0 & \text{otherwise}. \end{cases} \tag{1.9}$$

The Type II band versions RC_n^B of RC_n and T_n^B of T_n are defined with the input sequence $\{\bar{x}_i\}$ where

$$\bar{x}_i = \begin{cases} x_i & \text{if } i \le m \text{ or } i \ge n-m, \\ 0 & \text{otherwise}. \end{cases} \tag{1.10}$$

Let,

$$k_n = \#\{L_{A_n}(i,j) : 1 \le i, j \le n\}.$$

To obtain a nontrivial Type I band matrix we must have, $m_n \le k_n$. We will assume that $m_n \to \infty$ and $\frac{m_n}{n} \to \alpha < \infty$. Note that Type II banding does not yield any nontrivial situations for the symmetric circulant, the doubly symmetric and the palindromic matrices. Thus the Type II versions of these matrices are not considered in this article. Further, when $\alpha > 2$, all the above band matrices reduce to the full matrix with no zeroes. Further restrictions are required on α, for the individual matrices to yield nontrivial situations. These are specified in the statement of our main theorem later on.

The Toeplitz matrix has been an object of great interest in mathematics for a long time. See for example Grenander and Szego (1958) [9]. The sample autocovariance matrix in time series analysis also has the Toeplitz structure. If $\{x_i\}$ is square summable, then as the dimension increases, the finite dimensional circulant approximates the corresponding Toeplitz in various senses. Indeed this approximating property is exploited to obtain the limiting spectral distribution of the Toeplitz matrix in the nonrandom case. See Gray (2009) [8] for a recent account of this approximation idea.

Now, when the $\{x_i\}$ are i.i.d. with mean zero and variance one, then clearly $\{x_i\}$ is not square summable. For such an input sequence, the LSD of the symmetric circulant, after scaling, is relatively easily shown to be the standard normal distribution. Bose and Mitra (2002) [5] provide a proof based on normal approximation. A moment method proof is available in Bose and Sen (2008) [6]. The LSD of the Toeplitz matrix is much harder to obtain since approximation by a circulant, even after scaling, fails. A proof of the LSD of the Toeplitz matrix is available in Bryc, Dembo and Jiang (2006) [7]. It turns out that, when $\alpha = 0$, the Type I and Type II band symmetric Toeplitz matrices can still be approximated by the corresponding band symmetric circulant matrices in a suitable metric. We show that the LSD of the latter is normal and hence the LSD of the former two matrices are also normal. The situation changes drastically when $\alpha \neq 0$. The LSD of Type I symmetric circulant continues to be the normal distribution. However, the LSD of Type I and Type II Toeplitz matrices depend on α and no explicit forms seem to be obtainable.

Likewise, after appropriate scaling, the LSD of Type I and Type II reverse circulants is the symmetrized square root of chi square with two degrees of freedom irrespective of the value of α. When $\alpha = 0$, the Type II Hankel can be approximated by the Type II reverse circulant and hence has the same LSD. The Type I Hankel on the other hand has a degenerate distribution when $\alpha = 0$. When $\alpha \neq 0$, both Type I and Type II Hankel matrices have complicated limit distributions which depend on the value of α.

Table 1 summarizes our limit results. The proofs are given in details in Section 3. Let $N(0, 2)$ denote a normal random variable with mean zero and variance 2. Let R denote a random variable with density

$$f_R(x) = |x|\,\mathrm{e}^{-x^2}, \ \ -\infty < x < \infty.$$

Note that for all nonnegative integers k,

$$\mathrm{E}[N(0,2)^{2k}] = 2^k \frac{(2k)!}{2^k k!}, \ \ \mathrm{E}(R^{2k}) = k!, \ \ \mathrm{E}[N(0,2)^{2k+1}] = \mathrm{E}(R^{2k+1}) = 0.$$

In cases where the limit cannot be explicitly described, we later provide the second and fourth moments of the limits.

To prove our results, we mostly use the method of moments except, as mentioned above, for the Type I Toeplitz matrix and Type II Hankel matrix when $\alpha = 0$. The results of a few simulations is given in the figures. Figures 3, 4 and 5 show that the LSD of (scaled) T_n^b and H_n^b depend on α. From the proofs it will follow that the even moments of the LSD of $n^{-1/2}T_n^b$ and $n^{-1/2}H_n^b$ increase with α. Figures 4 and Figure 5 suggest that the LSD of $n^{-1/2}H_n^b$ is bimodal if $\alpha \geq 4/3$ and there is a positive mass at zero if $\alpha \leq 5/4$.

Some of the results reported in this paper first appeared in Basak (2009)[2]. The first version of the present article appeared in Bose and Basak (2009) [3]. We then discovered that results similar to ours have been obtained independently by Kargin (2009) [11] and Liu and Wang (2009) [12].

Kargin (2009) [11] deals with band Toeplitz matrices, which corresponds to Type I banding. When $\alpha = 0$, he proves that the LSD is Gaussian in probability, under the assumptions that the entries are independent mean zero and have uniformly bounded fourth moment. For this he uses the closeness of the circulant and the Toeplitz and the normal approximation as in Bose and Mitra (2002) [5]. When $\alpha \neq 0$, he assumes that, the entries have symmetric distribution and moments of any order are uniformly bounded, and proves that

Table 1: Matrices and the nature of their limits.

α	Matrix	Scaling	Limit
$\alpha = 0$	T_n^b, T_n^B	$m_n^{-1/2}$	$N(0,2)$
$0 \leq \alpha \leq 1/2$	SC_n^b, DH_n^b, PT_n^b, PH_n^b	$m_n^{-1/2}$	$N(0,2)$
$\alpha = 0$	H_n^B	$(2m_n)^{-1/2}$	R
$0 \leq \alpha \leq 1$	RC_n^b	$m_n^{-1/2}$	R
$0 \leq \alpha \leq 1/2$	RC_n^B	$(2m_n)^{-1/2}$	R
$0 \leq \alpha \leq 1$	T_n^b, H_n^B	$m_n^{-1/2}$	symmetric
$0 \leq \alpha \leq 1/2$	T_n^B	$m_n^{-1/2}$	symmetric
$0 < \alpha \leq 2$	H_n^b	$m_n^{-1/2}$	symmetric
$\alpha = 0$	H_n^b	$m_n^{-1/2}$	degenerate at 0

the expected spectral measure converges and that the limit is non Gaussian when $\alpha \neq \frac{4}{7}$. For this he uses the moment method and establishes the non Gaussianity by showing that the kurtosis of the limit is not 3.

Liu and Wang (2009) [12] deal with self adjoint Toeplitz and real symmetric Hankel band matrices and prove some general results on products of such matrices. Specializing to band matrices, they deal with only Type I banding for the Toeplitz matrix and Type II banding for the Hankel matrix. They assume the input sequence to be independent mean zero variance one, with uniformly bounded moment of all orders and prove the existence of the limit of the expected empirical spectral distribution in the above two cases. They also obtain the fourth moment of the limit distribution. Their method of proof is significantly different from ours. Their primary tool being the representation of Toeplitz and Hankel matrices as linear combinations of backward and forward shift matrices.

We have been able to provide a more comprehensive picture of the limits for both types of banding, for a larger class of matrices. All the results of Kargin (2009) [11] and Liu and Wang [12] on the LSD for the band matrices listed earlier follow from our main result.

In Section 2 we state our assumptions and the main theorem. Section 3 contains the details of the proofs of which the major part is verifying the (M1) condition. Section 4 contains the details of the second and fourth moments of the LSD which are not known in explicit forms. This section also discusses some limited simulations, which raise quite a few interesting questions.

2 MAIN RESULT

Assumption I. $\{x_i\}$ are independent with mean zero and variance one which are either (i) uniformly bounded or (ii) identically distributed.

When Assumption I (i) holds, the moments can be calculated without any reservation. We shall show in Lemma 1 that, (almost sure) LSD results which hold under Assumption I (i), continue to hold Assumption I (ii).

Assumption I* $\{x_i\}$ are independent with mean zero and variance one which satisfy

(i) $\sup \mathrm{E}\,|x_i|^{2+\delta} < \infty$ for some $\delta > 0$, and

(ii) For all large t, $\lim n^{-2} \sum_{i=0}^{n} \mathrm{E}\,|x_i|^4 I(|x_i| > t) = 0$.

We shall show in Lemma 1 that the LSD results which hold under Assumption I (i) continue to hold in probability under Assumptions I* (i), (ii).

Assumption II. $\{m_n\} \to \infty$ and $\lim_{n\to\infty} m_n/n = \alpha$ exists.

Assumption III. $\sum_{n=1}^{\infty} m_n^{-2} < \infty$. (Holds trivially when $\alpha \neq 0$).

We now state our main theorem.

Theorem 1. *Suppose one of the following hold: (A) Assumption I (i) and Assumption II, (B) Assumption I (ii) and Assumption II or, (C) Assumption I *(i) (ii) and Assumption II. Then the following hold in probability.*

(i) If $m_n \leq n/2$ *then ESD of* $m_n^{-1/2} SC_n^b, m_n^{-1/2} DH_n^b, m_n^{-1/2} PT_n^b$ *and* $m_n^{-1/2} PH_n^b \Rightarrow N(0,2)$.

(ii) If $m_n \leq n$ *then* $m_n^{-1/2} RC_n^b \Rightarrow R$.

(iii) If $m_n \leq 2n$ *then* $m_n^{-1/2} H_n^b \Rightarrow H_\alpha^b$ *(say), which is symmetric.* H_0^b *is the degenerate distribution at zero.*

(iv) If $m_n \leq n$ *then* $m_n^{-1/2} T_n^b \Rightarrow T_\alpha^b$ *(say), which is symmetric. In particular,* $T_0{}^b = N(0,2)$.

(v) If $m_n \leq n/2$ *then* $(2m_n)^{-1/2} RC_n^B \Rightarrow R$.

(vi) If $m_n \leq n$ *then* $(2m_n)^{-1/2} H_n^B \Rightarrow H_\alpha^B$ *(say), which is symmetric. In particular,* $H_0^B = R$.

(vii) If $m_n \leq n/2$ *then* $m_n^{-1/2} T_n^B \Rightarrow T_\alpha^B$ *(say), which is symmetric. In particular,* $T_0^B = N(0,2)$.

If Assumption III holds, then all the above convergence are almost sure in cases (A) and (B).

3 PROOF OF THEOREM

The method of moments may be described as follows. Suppose $\{Y_n\}$ is a sequence of random variables with distribution functions $\{F_n\}$ such that $\mathrm{E}(Y_n^h) \to \beta_h$ for every positive integer h, and $\{\beta_h\}$ satisfies *Carleman's condition* (see (C) below). Then there exists a distribution function F, such that for all h,

$$\beta_h = \beta_h(F) = \int x^h dF(x) \tag{3.1}$$

and $\{Y_n\}$ (or equivalently $\{F_n\}$) converges to F in distribution. This method is applied to the spectral distribution of random matrices as follows. Suppose $\{A_n\}$ is a sequence of matrices, and let, by a slight abuse of notation, $\beta_h(A_n)$, for $h \geq 1$, denote the h-th moment of the ESD of A_n. Suppose there is a sequence of nonrandom $\{\beta_h\}_{h=1}^{\infty}$ such that

(M1) **First moment condition:** For every $h \geq 1$, $\mathrm{E}[\beta_h(A_n)] \to \beta_h$

(M2) **Second moment condition:** $\text{Var}[\beta_h(A_n)] \to 0$

(C) **Carleman condition:** $\{\beta_h\}$ satisfies $\sum_{h=1}^{\infty} \beta_{2h}^{-1/2h} = \infty$.

Then the LSD, say A, is identified by $\{\beta_h\}_{h=1}^{\infty}$ and the convergence to LSD holds in probability. We denote this weak convergence by $A_n \Rightarrow A$. This convergence may be strengthened to almost sure convergence by strengthening (M2), for example, by showing that

(M4) **Fourth moment condition:** $\displaystyle\sum_{n=1}^{\infty} \text{E}[\beta_h(A_n) - \text{E}(\beta_h(A_n))]^4 < \infty$.

Here is an outline of the main steps in the proof. (i) In Section 3.1 we introduce the bounded Lipschitz metric d_{BL}. (ii) In Section 3.2 we use this first to show that it is enough to prove the theorem under the additional and convenient assumption that the input sequence is uniformly bounded. (iii) Then in Section 3.3 we introduce the trace formula for calculating moments and, the concepts of circuits, words etc which are useful in verifying the (M1) condition. (iv) In Sections 3.4 it is shown that certain terms are asymptotically negligible in the trace formula. (v) In Section 3.5 the (M1) condition is verified for each matrix, except in a few cases for which we use an indirect approach described below. Verification of the (M1) condition essentially yield the corresponding LSD. (vi) The Carleman condition and (M4) condition are then finally verified in a unified way in Section 3.6 and Section 3.7 for all the matrices except for the ones which are discussed below.

We bypass verifying the (M1) condition for the following situations. Suppose $\alpha = 0$. In this case, we do not verify (M1) for T_n^b, T_n^B and H_n^B. Instead, we establish their LSD by showing that, almost surely, the first two matrices are close in d_{BL} metric to SC_n^b, (see Sections 3.5.2 and 3.5.9) and the last one is close to RC_n^B (see Section 3.5.7). Further, the results for DH_n^b, PT_n^b and PH_n^b follow using their interrelations and their relations with SC_n^b and no detailed arguments are necessary (see Section 3.5.10).

3.1 A useful metric: the bounded Lipschitz metric

Weak convergence of measure is equivalent to convergence in the metric d_{BL} which is defined as

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

where $||f||_\infty = \sup_x |f(x)|$, $||f||_L = \sup_{x \neq y} |f(x) - f(y)|/|x - y|$.

The following facts will be used. Fact 1 is an estimate of the metric distance d_{BL} in terms of trace. See Bai (1999) for proof. Fact 2 is the well known Cauchy's interlacing inequality and its consequence.

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

$$d_{BL}^2(F^A, F^B) \leq \left(\frac{1}{n}\sum_{i=1}^{n} |\lambda_i(A) - \lambda_i(B)|\right)^2 \leq \frac{1}{n}\sum_{i=1}^{n}(\lambda_i(A) - \lambda_i(B))^2 \leq \frac{1}{n}\text{Tr}(A - B)^2. \quad (3.2)$$

Fact 2. Suppose C is an $n \times n$ symmetric real matrix with eigenvalues $\lambda_n \geq \lambda_{n-1} \geq \cdots \geq \lambda_1$. Let D be the $(n-1) \times (n-1)$ principal submatrix of C with eigenvalues $\mu_{n-1} \geq \mu_{n-2} \geq \cdots \geq \mu_1$. Then

$$\lambda_n \geq \mu_{n-1} \geq \lambda_{n-1} \geq \mu_{n-2} \geq \cdots \geq \lambda_2 \geq \mu_1 \geq \lambda_1.$$

As a consequence, if A, B are $n \times n$ symmetric real matrices then

$$||F^A - F^B||_\infty \leq \frac{\text{rank}(A - B)}{n}. \quad (3.3)$$

3.2 Reduction to uniformly bounded input

In this section we will show that in general, it is enough to consider only input sequences which are uniformly bounded. Hence the moment method can be applied without any further reservation. Define, α_n and β_n as follows. Note that $\beta_n \leq \alpha_n$.

$$\begin{aligned}
\alpha_n &= \max_k \#\{(i,j) : L_{A_n}(i,j) = k, 1 \leq i,j \leq n\} \\
\beta_n &= \max_{k \leq m_n} \#\{(i,j) : L_{A_n}(i,j) = k, 1 \leq i,j \leq n\} \text{ for matrices } A_n^b, \\
&= \max_{n-m_n \leq k \leq n+m_n} \#\{(i,j) : L_{H_n}(i,j) = k, 1 \leq i,j \leq n\} \text{ for } H_n^B, \\
&= \max_{|k-n/2| \geq n/2-m_n} \#\{(i,j) : L(i,j) = k, 1 \leq i,j \leq n\} \text{ for } A_n^B, \ A_n = RC_n \text{ or } T_n .
\end{aligned}$$

Lemma 1. *Let A_n^b be a sequence of Type I or Type II band matrices defined as above where $m_n \to \infty$ and $\beta_n = O(n)$. If for every bounded, mean zero and variance one input sequence, $F^{m_n^{-1/2}A_n^b}$ converges to some fixed nonrandom distribution F a.s.. Then,*

(i) the same limit continues to hold almost surely if $\{x_i\}$ is i.i.d. with mean zero and variance one.

(ii) the same limit continues to hold in probability if $\{x_i\}$ satisfies Assumptions I (i) and (ii).*

In particular, for all our band matrices, $\alpha_n = O(n)$ and the conclusions are valid.

Proof. We prove the result only for Type I banding. The proof is similar for Type II banding and is omitted.

(i) Let $\{x_0, x_1, \ldots\}$ be an i.i.d. input sequence for the matrices $\{A_n^b\}$. For $t > 0$, denote

$$\mu(t) \stackrel{def}{=} \mathrm{E}[x_0 \mathbb{I}(|x_0|x > t)] = -\mathrm{E}[x_0 \mathbb{I}(|x_0| \leq t)], \qquad \sigma^2(t) \stackrel{def}{=} \mathrm{Var}(x_0 \mathbb{I}(|x_0| \leq t)).$$

Note that $\mu(t) \to 0$ and $\sigma(t) \to 1$ as $t \to \infty$ and $\sigma^2(t) \leq 1$. Define bounded random variables

$$\hat{x}_i^t = \frac{x_i \mathbb{I}(|x_i| \leq t) + \mu(t)}{\sigma(t)} = \frac{x_i - \bar{x}_i^t}{\sigma(t)} \quad \text{where } \bar{x}_i^t = x_i \mathbb{I}(|x_i| > t) - \mu(t) = x_i - \sigma(t)\hat{x}_i^t. \tag{3.4}$$

It is easy to see that $\mathrm{E}((\bar{x}_0^t)^2) = 1 - \sigma^2(t) - \mu(t)^2 \to 0$ as $t \to \infty$. Further, $\{\hat{x}_i^t\}$ are i.i.d. bounded, with mean zero and variance one. Let us denote the matrix A_n^b constructed from this sequence $\{\hat{x}_i^t\}_{i \geq 0}$ by $\hat{A}_n^b$. Then

$$\begin{aligned}
d_{BL}^2\left(F^{m_n^{-1/2}A_n^b}, F^{m_n^{-1/2}\hat{A}_n^b}\right) &\leq 2d_{BL}^2\left(F^{m_n^{-1/2}A_n^b}, F^{m_n^{-1/2}\sigma(t)\hat{A}_n^b}\right) + 2d_{BL}^2\left(F^{m_n^{-1/2}\hat{A}_n^b}, F^{m_n^{-1/2}\sigma(t)\hat{A}_n^b}\right) \\
&\leq \frac{2}{m_n n}\mathrm{Tr}(A_n^b - \sigma(t)\hat{A}_n^b)^2 + \frac{2(1-\sigma(t))^2}{m_n n}\mathrm{Tr}(\hat{A}_n^b)^2.
\end{aligned}$$

Now by the hypotheses on A_n, using strong law of large numbers, we get

$$\frac{1}{m_n n}\mathrm{Tr}(\hat{A}_n^b)^2 = \frac{1}{m_n n}\sum_{1 \leq i,j \leq n}(\hat{a}_{i,j,n}^b)^2 \leq \frac{\beta_n}{m_n n}\sum_{h=0}^{m_n}(\hat{x}_h^t)^2 \leq \frac{C}{m_n}\sum_{h=0}^{m_n}(\hat{x}_h^2)^2 \stackrel{a.s.}{\to} C\,\mathrm{E}((\hat{x}_0^t)^2) = C$$

and we have $1 - \sigma(t) \to 0$ as $t \to \infty$. Similarly,

$$\frac{1}{m_n n}\mathrm{Tr}(A_n^b - \sigma(t)\hat{A}_n^b)^2 \leq \frac{C}{m_n}\sum_{h=0}^{m_n}(\bar{x}_h^t)^2 \stackrel{a.s.}{\to} C\,\mathrm{E}((\bar{x}_0^t)^2) \to 0 \text{ as } t \to \infty.$$

This proves the result for i.i.d. input sequence.

(ii) Since $\{x_i\}$ are not necesarily i.i.d., let $\sigma_i^2(t) = \mathrm{Var}(x_i \mathbb{I}(|x_i| \le t)$ and now define,

$$\hat{x}_i^t = \frac{x_i \mathbb{I}(|x_i| \le t) - \mathrm{E}[x_i \mathbb{I}(|x_i| \le t)}{\sigma_i(t)}, \; \bar{x}_i^t = x_i \mathbb{I}(|x_i| > t) - \mathrm{E}(x_i \mathbb{I}(|x_i| > t) = x_i - \sigma_i(t)\hat{x}_i^t.$$

Observe that

$$[\mathrm{E}[x_i \mathbb{I}(|x_i| \le t)]]^2 = [\mathrm{E}[x_i \mathbb{I}(|x_i| > t)]]^2 \le \mathrm{E}\, x_i^2 \mathbb{I}(|x_i| > t) \le \frac{1}{t^\delta} \mathrm{E}[|x_i|^{2+\delta}].$$

Thus under Assumption I* (i),

$$1 - \sigma_i^2(t) = 1 - \mathrm{E}[x_i^2 \mathbb{I}(|x_i| \le t)] + [\mathrm{E}[x_i \mathbb{I}(|x_i| \le t)]]^2 = \mathrm{E}[x_i^2 \mathbb{I}(|x_i| > t)] + [\mathrm{E}[x_i \mathbb{I}(|x_i| \le t)]]^2 = O(t^{-\delta}),$$

uniformly over i and $\forall\, t > 0$. Thus $\sup_i(1 - \sigma_i^2(t)) = O(t^{-\delta})$. It also easily follows that $\sup_i(1 - \sigma_i(t)) = O(t^{-\delta})$. Now let $\hat{A}_n^b$ the matrix A_n^b constructed from the input sequence $\{\hat{x}_i^t\}_{i \ge 0}$. Then

$$\begin{aligned}
d_{BL}^2\Big(F^{m_n^{-1/2} A_n^b}, F^{m_n^{-1/2} A_n}\Big) &\le \frac{1}{m_n n} \mathrm{Tr}(A_n^b - \hat{A}_n^b)^2 \\
&\le \frac{\beta_n}{m_n n} \sum_{h=0}^{m_n} (x_h - \hat{x}_h^t)^2 \\
&\le \frac{C}{m_n} \sum_{h=0}^{m_n} (x_h - \hat{x}_h^t)^2 \\
&= \frac{C}{m_n} \sum_{h=0}^{m_n} (x_h - \sigma_h(t)\hat{x}_h^t + \sigma_h(t)\hat{x}_h^t - \hat{x}_h^t)^2 \\
&= \frac{2C}{m_n} \sum_{h=0}^{m_n} (x_h - \sigma_h(t)\hat{x}_h^t)^2 + \frac{2C}{m_n} \sum_{h=0}^{m_n} (\sigma_h(t)\hat{x}_h^t - \hat{x}_h^t)^2 = T_1 + T_2.
\end{aligned}$$

Now $T_2 \le \frac{2C}{m_n} \sup_h (1 - \sigma_h(t))^2 \sum_{h=0}^{m_n} (\hat{x}_h^t)^2$. Since $\{\hat{x}_h^t\}$ are uniformly bounded, a simple application of Borel Cantelli Lemma implies that $\frac{1}{m_n} \sum_{h=0}^{m_n} (\hat{x}_i^t)^2 \stackrel{a.s.}{\to} 1$. So

$$T_2 \stackrel{a.s.}{\le} \sup_h (1 - \sigma_h(t))^2 = O(t^{-2\delta}) \to 0 \;\text{ as }\; t \to \infty.$$

Now we deal with T_1. First observe $T_1 = \frac{2C}{m_n} \sum_{h=0}^{m_n} (x_h - \sigma_h(t)\hat{x}_h^t)^2 = \frac{2C}{m_n} \sum_{h=0}^{m_n} (\bar{x}_h^t)^2$.

Now,

$$\mathrm{E}[\frac{1}{m_n} \sum_{h=0}^{m_n} (\bar{x}_h^t)^2] \le \frac{2}{m_n} \sum_{h=0}^{m_n} \Big[\mathrm{E}[x_h^2 \mathbb{I}(|x_h| > t)] + [\mathrm{E}[x_h \mathbb{I}(|x_h| > t)]^2 \Big] \le \frac{4}{t^\delta} \sup_h \mathrm{E}[|x_h^{2+\delta}|],$$

and

$$\mathrm{Var}[\frac{1}{m_n} \sum_{h=0}^{m_n} (\bar{x}_h^t)^2] \;\; \le \;\; \frac{1}{m_n^2} \sum_{h=0}^{m_n} \mathrm{E}[(\bar{x}_h^t)^4]$$

$$= \frac{1}{m_n^2}\sum_{h=0}^{m_n} \mathrm{E}[(x_h \mathbb{I}(|x_h|>t) - \mathrm{E}[x_h \mathbb{I}(|x_h|>t])^4]$$
$$\leq Const. \times \frac{1}{m_n^2}\sum_{h=0}^{m_n} \Big[\mathrm{E}[(x_h^4 \mathbb{I}(|x_h|>t)] \to 0,$$

using Assumption I* (ii). Thus $\limsup_n T_1 \leq \frac{4}{t^\delta}\mathrm{E}[|x_h^{2+\delta}|] \to 0$ as $t \to \infty$. As a consequence, $\limsup_n d_{BL}^2\Big(F^{m_n^{-1/2}A_n^b}, F^{m_n^{-1/2}A_n}\Big) \xrightarrow{P} 0$ and the proof is complete. □

3.3 Trace formula, circuits, words and matches

As preparatory to using the moment method, we introduce the trace formula and a few other concepts.

$$\frac{1}{n}\mathrm{Tr}[A_n^b]^h = \frac{1}{n}\sum_{1\leq i_1,i_2,\ldots,i_h\leq n}\prod_{j=1}^{h-1} x_{L_A(i_j,i_{j+1})}\times x_{L_A(i_h,i_1)}\times\prod_{j=1}^{h-1}\mathbb{I}(L_A(i_j,i_{j+1})\leq m)\times\mathbb{I}(L_A(i_h,i_1)\leq m). \tag{3.5}$$

The following concepts from Bose and Sen (2008) [6] will be needed to take advantage of the trace formula.

Circuit and vertices: Any function $\pi : \{0,1,2,\ldots,h\} \to \{1,2,\ldots,n\}$ is said to be a *circuit* if $\pi(0) = \pi(h)$. Any $\pi(i)$ will be called a *vertex*. The **length** $l(\pi)$ of π is taken to be h. A circuit depends on h and n but we will suppress this dependence.

Let,

$$\mathbb{X}_\pi = \prod_{i=1}^{h} x_{L(\pi(i-1),\pi(i))}, \quad \mathbb{I}_{\pi,L}^h = \mathbb{I}_\pi^h = \prod_{i=1}^{h}\mathbb{I}(L(\pi(i-1),\pi(i))\leq m).$$

Hence

$$\mathrm{E}[\beta_h(n^{-1/2}A_n)] = \mathrm{E}\left[\frac{1}{n}\mathrm{Tr}\left(\frac{A_n}{\sqrt{n}}\right)^h\right] = \frac{1}{n^{1+h/2}}\sum_{\pi:\pi\ circuit}\mathrm{E}\,\mathbb{X}_\pi,$$

and

$$\mathrm{E}[\beta_h(m^{-1/2}A_n^b)] = \mathrm{E}\left[\frac{1}{n}\mathrm{Tr}\left(\frac{A_n^b}{\sqrt{m}}\right)^h\right] = \frac{1}{nm^{h/2}}\sum_{\pi:\pi\ \text{circuit}}\mathrm{E}\,\mathbb{X}_\pi\times\mathbb{I}_\pi^h.$$

Matched circuits: Any value $L(\pi(i-1),\pi(i))$ is an L value of π and π has an **edge of order** e $(1 \leq e \leq h)$ 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 $e \geq 2$ are relevant. Such circuits will be said to be *matched*. For any such π, given any i, there is at least one $j \neq i$ such that $L(\pi(i-1),\pi(i)) = L(\pi(j-1),\pi(j))$. π is *pair matched* if all its edges are of order two. To deal with (M2) and (M4), we need multiple circuits and the following related notions.

Jointly Matched: k circuits $\pi_1,\pi_2,\ldots,\pi_k$ are said to be *jointly matched* if given $1 \leq l \leq k$ and $1 \leq i \leq h$ there exists $1 \leq t \leq k$ and $1 \leq j \leq h$ such that $(l,i) \neq (t,j)$ and $L(\pi_l(i-1),\pi_l(i)) = L(\pi_t(j-1),\pi_t(j))$.

Cross Matched: k circuits $\pi_1,\pi_2,\ldots,\pi_k$ are said to be *cross matched* if each circuit has one L-value which occurs in at least one of the remaining circuits.

Equivalence relation on circuits: Two circuits π_1 and π_2 (of same length) are said to be equivalent if their L-values agree at exactly the same pairs (i,j). That is, iff $\big\{L(\pi_1(i-1),\pi_1(i)) = L(\pi_1(j-1),\pi_1(j)) \Leftrightarrow L(\pi_2(i-1),\pi_2(i)) = L(\pi_2(j-1),\pi(j))\big\}$. 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 the word $abacba$. 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 nonmatched, 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 a *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.

Define,

$$\begin{aligned}
\Pi(w) &= \{\pi : w[i] = w[j] \Leftrightarrow L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j))\},\\
\Pi^b(w) &= \{\pi : w[i] = w[j] \Leftrightarrow L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j)) \text{ and } \mathbb{I}^h_\pi = 1\}.
\end{aligned}$$

Note these depend on the underlying L funciton but we suppress this dependence. We can rewrite $\mathrm{E}[\beta_h(n^{-1/2}A_n)]$ and $\mathrm{E}[\beta_h(m^{-1/2}A^b_n)]$ as

$$\mathrm{E}[\beta_h(n^{-1/2}A_n)] = \sum_{w \text{ matched}} \frac{1}{n^{1+h/2}} \sum_{\Pi(w)} \mathrm{E}\,\mathbb{X}_\pi \text{ and } \mathrm{E}[\beta_h(m^{-1/2}A^b_n)] = \sum_{w \text{ matched}} \frac{1}{nm^{h/2}} \sum_{\Pi^b(w)} \mathrm{E}\,\mathbb{X}_\pi.$$

Note that for some of the matrices m will be replaced by $(2m)$ above. Since the outer sum is a finite sum, to verify condition (M1), it is enough to show the existence of the inner limit for every w. For any w, let (whenever the limit exists)

$$p(w) = \sum_{\Pi^b(w)} \mathrm{E}\,\mathbb{X}_\pi.$$

Note that this depends on the link function but we do show this dependence in our notation. While deriving limits, later, we show that it is enough to consider only pair matched words. The total number of pair matched words with $|w| = k$ equals $\frac{(2k)!}{2^k k!}$. We also show later that $\Pi^{b*}(w) \supseteqq \Pi^b(w)$ defined below is equivalent to $\Pi^b(w)$ for asymptotic considerations, but is easier to work with.

$$\Pi^{b*}(w) = \{\pi : w[i] = w[j] \Rightarrow L(\pi(i-1), \pi(i)) = L(\pi(j-1), \pi(j)) \text{ and } \mathbb{I}^h_\pi = 1\}.$$

Similarly, for RC^B_n, T^B_n and H^B_n, define,

$$\begin{aligned}
\mathbb{I}^h_{\pi,RC} &= \prod_{i=1}^{h} \mathbb{I}\Big(L_{RC}(\pi(i-1), \pi(i)) \le m \text{ or } L_{RC}(\pi(i-1), \pi(i)) \ge n-m\Big),\\
\mathbb{I}^h_{\pi,H} &= \prod_{i=1}^{h} \mathbb{I}(n-m \le L_H(\pi(i-1), \pi(i)) \le n+m),\\
\mathbb{I}^h_{\pi,T} &= \mathbb{I}(|\pi(0)-\pi(1)| \le m \text{ or } |\pi(0)-\pi(1)| \ge n-m) \times\\
&\quad \mathbb{I}(|\pi(1)-\pi(2)| \le m \text{ or } |\pi(1)-\pi(2)| \ge n-m) \times \cdots \times\\
&\quad \mathbb{I}(|\pi(h-1)-\pi(h)| \le m \text{ or } |\pi(h-1)-\pi(h)| \ge n-m),
\end{aligned}$$

and $\Pi^B(w)$ and $\Pi^{B*}(w)$ are same as $\Pi^b(w)$ and $\Pi^{b*}(w)$ above with the above indicators fot Type II banded matrices. We shall continue to use $p(w)$ to denote the corresponding limits as above.

3.4 Negligibility of edges of order ≥ 3

The following Lemma helps to show that contribution from $\Pi^b(w)$ and $\Pi^B(w)$ are zero if w has at least one $e \geq 3$ and, also helps to verify the (M1) condition later.

Lemma 2. *(a) For SC^b, given $\pi(0), \pi(1), \ldots, \pi(i-1)$, for any dependent vertex $\pi(i)$, $\#\pi(i) \leq 6$.*

(b) For any independent vertex $\pi(i),\ i \geq 1$, of $SC^b,\ 2m \leq \#\pi(i) \leq 2m+1$.

(c) For any independent vertex $\pi(i),\ i \geq 1$, of $RC_n^b,\ m \leq \#\pi(i) \leq m+1$.

(d) For any independent vertex $\pi(i),\ i \geq 1$, of $RC_n^B,\ 2m \leq \#\pi(i) \leq 2m+1$.

Proof. Part (a) follows easily from the relation $L_{SC}(\pi(j-1), \pi(j) = L_{SC}(\pi(i-1), \pi(i))$ where $j < i$, which in turn implies $|n/2 - |\pi(j-1) - \pi(j)|| = |n/2 - |\pi(i-1) - \pi(i)||$.

We now prove part (b). The proofs of (c) and (d) are similar and shall be omitted.

First assume $i = 1$. Since $0 \leq L_{SC}(\pi(0), \pi(1)) \leq m$ and $1 \leq \pi(1) \leq n$, we obtain,

$$0 \leq n/2 - |n/2 - |\pi(0) - \pi(1)|| \leq m \quad \Rightarrow \quad n/2 - m \leq |n/2 - |\pi(0) - \pi(1)|| \leq n/2.$$

It then easily follows that one of the following three holds.
(i) $\pi(0) - m \leq \pi(1) \leq \pi(0) + m$,
(ii) $\pi(1) \leq \pi(0) + m - n$ or,
(iii) $\pi(1) \geq \pi(0) + n - m$.

If $m = n/2$ then irrespective of the value of $\pi(1)$, we have $0 \leq n/2 - |n/2 - |\pi(0) - \pi(1)|| \leq m$. Hence number of choices of $\pi(1)$ is $n = 2m$.

Now assume $m < n/2$. Note that the three regions in (i), (ii) and (iii) are disjoint in this case.

We consider three cases and establish (b) in each case separately.

Case A. $\pi(0) \in \{1, 2, \ldots, m\}$. Note that $\pi(0)+m-n \leq 2m-n < 0$ implies $\pi(0)+m \leq 2m < n$ and the range of choices in (i) is $1 \leq \pi(1) \leq \pi(0)+m$. Hence number of choices from (i) is $\pi(0)+m$. The number of choices from (ii) is 0. As $\pi(0)+n-m \leq n$, range of choices from (iii) is $\pi(0)+n-m \leq \pi(1) \leq n$. Hence number of choices from (iii) is $m - \pi(0) + 1$. In all, that is total of $(2m+1)$ choices for $\pi(1)$.

Case B. $\pi(0) \in \{m+1, m+2, \ldots, n-m-1\}$. As $\pi(0) - m > 0$ and $\pi(0) + m < n - m + m = n$, now the range of choices from (i) is $\pi(0) - m \leq \pi(1) \leq \pi(0) + m$. whcih is $2m$ many choices. As $\pi(0) + m - n < 0$ and $\pi(0) + n - m > m + n - m = n$, there are no choices from (ii) and (iii).

Case C. $\pi(0) \in \{n-m, n-m+1, \ldots, n\}$. Note that $\pi(0)+m \geq n-m+m = n$ and $\pi(0) \geq n-m > m$. Now the range of choices for $\pi(1)$ from (i) is $\pi(0)-m \leq \pi(1) \leq n$ which gives $n+m-\pi(0)+1$ choices. Since $\pi(0)+n-m > n$, the number of choices from (iii) is 0. As $\pi(0)+m-n \geq 0$, number of choices from (ii) is $\pi(0)+m-n$. In all, that is a total of choices $(2m+1)$ choices.

Now move to $\pi(2)$. If it is a dependent vertex then there are at most six ways to choose it and if it is an independent vertex then the above argument may be repeated. Continuing the process proves (b) completely. $\square$

The next Lemma shows that the contribution from any word having edge(s) of order ≥ 3 is zero in the limit. Let $N_{h,3+}^L$ be the number of L matched circuits on $\{1, 2, \ldots, n\}$ of length h with at least one $e \geq 3$ and $N_{h,3+}^{bL}$ be the same with added restriction $\mathbb{I}_\pi^h = 1$. Define $N_{h,3+}^{BL}$ similarly for Type II banding.

Lemma 3. *(a) Let $L = L_T$ or L_H. There exists a constant C_L such that, $N_{h,3+}^{bL} \leq C_L n^{\lfloor (h+1)/2 \rfloor}$ and hence $n^{-(1+h/2)} N_{h,3+}^{bL} \to 0$.*

(b) Let $L = L_{SC}$ or L_{RC}. Then there exists a constant $C_L$such that, $N_{h,3+}^{bL} \leq C_L n m^{\lfloor (h-1)/2 \rfloor}$ and hence $n^{-1} m^{-h/2} N_{h,3+}^{bL} \to 0$.

(c) (a) holds for $N_{h,3+}^{BL_H}$ and $N_{h,3+}^{BL_T}$ and, (b) holds for $N_{h,3+}^{BL_{RC}}$.

(d) For $\alpha = 0$, $n^{-1} m^{-h/2} N_{h,3+}^{bL_H} \to$ as $n \to \infty$.

Proof. (a) First observe that, $N_{h,3+}^{bL} \leq N_{h,3+}^{L}$. Now note that, if $h = 2k$ or $2k-1$ then there are at most $k-1$ distinct L-values, and hence, at most k independent vertices. Each independent vertex can be chosen in at most n ways and the dependent vertices can be chosen in $O(1)$ ways. This completes the proof of (a).

(b) Let $h = 2k$ or $2k-1$. Since there is at least one $e \geq 3$, there must be at most $(k-1)$ distinct L values and hence at most k independent vertices including $\pi(0)$. Note that $\#\pi(0) = n$, and by Lemma 2 $\#\pi(i) = O(1)$ or $\#\pi(i) = O(m)$ depending on whether or not $\pi(i)$ is dependent or independent. Hence, $N_{h,3+}^{bL} \leq Cnm^{k-1} \leq Cnm^{\lfloor (h-1)/2 \rfloor}$ and (b) follows.

The proof of (c) is similar and we omit the details.

(d) Since $\alpha = 0$, for sufficiently large n, $m < n$. To achieve $\mathbb{I}_\pi^h = 1$, every independent vertex $\pi(i)$, $i \geq 1$, can be chosen in at most m ways (as the link function is $L(i,j) = i + j$). Let $h = 2k$ or $h = 2k-1$, since there is at least one edge of order ≥ 3, we must have at most $k-1$ independent vertices excluding $\pi(0)$. So total number of vertex choices is $O(nm^{k-1})$. Hence we obtain the result. $\square$

3.5 (M1) condition

Before we verify the (M1) condition for all the cases, we prove a Lemma which reduces the number of terms in the trace formula asymptotically. It also implies that the odd moments of the LSD equal zero.

Lemma 4. *(a) Suppose $\{A_n^b = RC_n^b,\ SC_n^b\}$ (for any α) or $\{A_n^b = T_n^b$ or $H_n^b\}$ (for $\alpha \neq 0$) with a uniformly bounded, independent mean zero and variance one input sequence $\{x_i\}$.*

(i) If h is odd, $\lim_{n\to\infty} \mathrm{E}[\beta_h(m^{-1/2}A_n^b)] = \lim_{n\to\infty} \mathrm{E}\left[\frac{1}{n}\mathrm{Tr}\left(\frac{A_n^b}{\sqrt{m}}\right)^h\right] = 0$.

(ii) If $h = 2k$ then, $\sum_{w \text{ pair matched}} \lim_{n\to\infty} \frac{1}{nm^k}|\Pi^{b*}(w) - \Pi^b(w)| = 0$ and if any of the last two limits below exists,

$$\lim_{n\to\infty} \mathrm{E}[\beta_{2k}(m^{-1/2}A_n^b)] = \sum_{w \text{ pair matched}} \lim_{n\to\infty} \frac{1}{nm^k}|\Pi^b(w)| = \sum_{w \text{ pair matched}} \lim_{n\to\infty} \frac{1}{nm^k}|\Pi^{b*}(w)|. \quad (3.6)$$

(b) All conclusions of (a) above holds for $\{m^{-1/2}RC_n^B\}$ (for any α) and for $\{n^{-1/2}H_n^B\}$ and $\{n^{-1/2}T_n^B\}$ (when in either case $\alpha \neq 0$) with Π^b and Π^{b*} replaced by Π^B and Π^{B*} respectively.

(c) Similar conclusions hold for $\{m^{-1/2}H_n^b\}$ when $\alpha = 0$.

Proof. Note that from mean zero and independence assumption (provided the limit below exists),

$$\lim_{n\to\infty} \mathrm{E}[\beta_h(m^{-1/2}A_n^b)] = \lim_{n\to\infty} \frac{1}{nm^{h/2}} \sum_{\pi \text{ circuit}} \mathrm{E}\,\mathbb{X}_\pi \times \mathbb{I}_\pi^h = \sum_{w \text{ matched}} \lim_{n\to\infty} \frac{1}{nm^{h/2}} \sum_{\pi\in\Pi^b(w)} \mathrm{E}\,\mathbb{X}_\pi. \quad (3.7)$$

By Holder's inequality, $|\sum_{\pi:\pi\in\Pi^b(w)} \mathrm{E}\,\mathbb{X}_\pi| \le |\Pi^b(w)|\,\mathrm{E}(|x_{L(1,1)}|^h)$. Therefore, from Lemma 3, matched circuits which have edges of order ≥ 3 do not contribute to the limit in (3.7). Further, if w is pair matched and $\pi \in \Pi^{b*}(w) \setminus \Pi^b(w)$ then π must have an edge of order ≥ 4 and hence $\sum_{w \text{ pair matched}} \lim_{n\to\infty} \frac{1}{nm^k}|\Pi^{b*}(w) - \Pi^b(w)| = 0$. So provided the limits exist, (3.6) holds for $h = 2k$, proving (a) (ii).

To prove (a) (i), note that at least one edge is of order ≥ 3. Hence by Lemma 3 $\lim_{n\to\infty} \mathrm{E}[\beta_h(m^{-1/2}A_n^b)] = \lim_{n\to\infty} E\left[\frac{1}{n}\operatorname{Tr}\left(\frac{A_n^b}{\sqrt{m}}\right)^h\right] = 0$, proving (a) (i).

To prove (b) and (c) we may proceed similarly. Details are omitted. □

Thus from the above result, to establish (M1), we may restrict consideration to only pair matched words. In the next few subsections we show that the existence of the limit in (3.6) (and hence of $p(w)$) for all relevant choices of the band matrices, by identifying the inner limit for each fixed word and then summing over all pair matched words.

For the cases when $\alpha \neq 0$, it will be convenient to obtain expressions for $\lim_{n\to\infty} \mathrm{E}[\beta_h(n^{-1/2}A_n^b)]$ where $A_n^b = T_n^b, H_n^b, T_n^B$ or H_n^B. Observe that then the expression for $\lim_{n\to\infty} \mathrm{E}[\beta_h(m^{-1/2}A_n^b)]$ is given by the relation

$$\lim_{n\to\infty} \mathrm{E}[\beta_h(m^{-1/2}A_n^b)] = \alpha^{-h/2} \times \lim_{n\to\infty} \mathrm{E}[\beta_h(n^{-1/2}A_n^b)].$$

Similar comment holds for Type II banding.

3.5.1 (M1) condition for SC_n^b

In this case, for any pair matched word w, (3.6) equals 2^k (which eventually leads to the $N(0,\ 2)$ LSD). Define $u_i = \pi(i-1) - \pi(i)$. We call them "slopes". Note that $w[i] = w[j]$ implies that $u_i \pm u_j = 0, \pm n$. We first show that only those matchings with $u_i + u_j = 0, \pm n$ contribute in the limit.

Lemma 5. *Fix an L_{SC} matched word w of length $2k$, with $|w| = k$. Let $\mathcal{N}^b$ be the number of matched circuits π on $\{1, 2, \ldots, n\}$ of length $2k$ such that $\mathbb{I}_\pi^h = 1$ and it has at least one pair $i < j$ with $w[i] = w[j]$ such that $u_i - u_j = 0, \pm n$. Then, $\mathcal{N}^b = O(nm^{k-1})$ and hence $n^{-1}m^{-k}\mathcal{N}^b \to 0$.*

Proof. Let $(i_1, j_1), (i_2, j_2), \ldots (i_k, j_k)$ denote the pair partition corresponding to the word w, i.e. $w[i_l] = w[j_l], 1 \leq l \leq k$. Suppose, w.l.o.g. $u_{i_k} - u_{j_k} = 0, \pm n$. Clearly a circuit π becomes completely specified if we know $\pi(0)$ and all the 'slopes' u_i's.

By Lemma 2(a), if we fix some value for u_{i_l}, each u_{j_l} has six possible values. Now, $L_{SC}(\pi(i-1), \pi(i)) \leq m$ implies either $|u_i| \leq m$ or $|u_i| \geq n - m$. Hence for every i, each u_i has $O(m)$ possible values. Further, $\pi(0)$ has at most n possible values. For any such valid choice, from the sum restriction $\sum_{i=1}^{2k} u_i = \pi(2k) - \pi(0) = 0$ we know $u_{i_k} + u_{j_k}$ and on the other hand by hypothesis, $u_{i_k} - u_{j_k} = 0, +n, -n$. Thus the pair (u_{i_k}, u_{j_k}) has 3 possibilities. Thus there are at most $O(nm^{k-1})$ circuits with the given restrictions and the proof of the Lemma is complete. □

As a consequence of the above lemma we have,

$$\sum_{w \text{ pair matched}} \lim_{n\to\infty} \frac{1}{nm^k}|\Pi^{b*}(w)| = \sum_{w \text{ pair matched}} \lim_{n\to\infty} \frac{1}{nm^k}|\Pi^{b'}(w)|, \tag{3.8}$$

where

$$\Pi^{b'}(w) = \{\pi : \pi \text{ is a circuit}, w[i] = w[j] \Rightarrow u_i + u_j = 0, \pm n \text{ and } I_\pi^h = 1\}.$$

Lemma 6. *If w is pair matched with $|w| = k$, $\frac{1}{nm^k}|\Pi^{b'}(w)| \to 2^k$. Hence*

$$\lim \mathrm{E}[\beta_{2k}(m^{-1/2}SC_n^b)] = 2^k \frac{(2k)!}{2^k k!} = \mathrm{E}[N(0,2)^{2k}].$$

Proof. Suppose for some $i < j, u_i + u_j = 0, \pm n$. If we know the circuit up to position $(j-1)$ then $\pi(j)$ has to take one of the values $A - n, A, A + n$ where $A = \pi(j-1) - \pi(i) + \pi(i-1)$. Noting that $-(n-2) \leq A \leq (2n-1)$, exactly one of the three values will fall within 1 and n and be a valid choice for $\pi(j)$. From Lemma 2 (b) if i is an independent vertex and the circuit is known up to position $i-1$ then there are $2m + c_i$ many choices for $\pi(i)$ lying between 1 and n, where $c_i = 0$ or 1, depending on the value of the previous vertices. Assume for the moment that such a choice of all $\pi(i)$ satisfies the circuit condition. Then $n \times (2m)^k \leq |\Pi^{b'}(w)| \leq n \times (2m+1)^k$. So we have, $\lim_{n\to\infty} \frac{1}{nm^k}|\Pi^{b'}(w)| = 2^k$.

Now we show why the circuit condition $\pi(0) = \pi(2k)$ is satisfied. Observe that

$$\pi(2k) - \pi(0) = \sum_{i=1}^{2k} u_i = dn$$

for some $d \in \mathbb{Z}$. But since $|\pi(2k) - \pi(0)| \leq n - 1$, we must have $d = 0$. Thus the circuit condition is satisfied. Since the total number of pair matched words is $\frac{(2k)!}{2^k k!}$, the proof is complete. □

3.5.2 Closeness of T_n^b and SC_n^b when $\alpha = 0$

We now show that for $\alpha = 0$, T_n^b and SC_n^b have the same limit behaviour.

Lemma 7. *(i) Suppose Assumption I (i) or (ii) hold and $\alpha = 0$. Then $d_{BL}^2(F^{m^{-1/2}T_n^b}, F^{m^{-1/2}SC_n^b}) \xrightarrow{a.s.} 0$.*

(ii) Suppose Assumption I (i) and (ii) hold and $\alpha = 0$. Then $d_{BL}^2(F^{m^{-1/2}T_n^b}, F^{m^{-1/2}SC_n^b}) \xrightarrow{P} 0$.*

Proof. Define the upper (lower) k^{th} diagonal as those indices (i,j) such that $i - j = k$ (respectively, $i - j = -k$). Note that both the matrices SC_n^b and T_n^b have all upper and lower k^{th} diagonal entries equal to x_k, for $k = 1, 2, \ldots, m$ and moreover SC_n^b has all the upper and lower k^{th} diagonal entries equal to x_{n-k}, for $k = n-1, n-2, \ldots, n-m$. Hence

$$\begin{aligned}
d_{BL}^2(F^{m^{-1/2}T_n^b}, F^{m^{-1/2}SC_n^b}) &\leq \frac{1}{n}\,\mathrm{Tr}\left[\frac{(T_n^b - SC_n^b)^2}{m}\right] \\
&= \frac{2}{mn}(mx_m^2 + (m-1)x_{m-1}^2 + \cdots + x_1^2) \\
&\leq \frac{2}{n}(x_m^2 + x_{m-1}^2 + \cdots + x_1^2) \\
&= 2 \times \left(\frac{m}{n}\right) \times \frac{x_1^2 + x_2^2 + \cdots + x_m^2}{m}.
\end{aligned}$$

(i) First suppose $\{x_i\}$ are i.i.d. Then $\dfrac{x_1^2 + x_2^2 + \cdots + x_m^2}{m} \xrightarrow{a.s.} \mathrm{E}[x_0^2] = 1$ and $\dfrac{m}{n} \to 0$ and the result follows. If, instead, the $\{x_i\}$ are uniformly bounded, a simple modification in the above proof does the job.

(ii) Suppose Assumption I* (i) and (ii) hold. Then $\mathrm{E}[\frac{x_1^2+x_2^2+\cdots+x_m^2}{n}] = \frac{m}{n} \to 0$ as $n \to \infty$ and

$$\mathrm{Var}\left[\frac{x_1^2 + x_2^2 + \cdots + x_m^2}{n}\right] \leq \frac{\sum_{i=1}^{m} \mathrm{E}[x_i]^4}{n^2} \leq \frac{1}{n^2} \times [mt^4 + \sum_{i=1}^{m} \mathrm{E}[x_i^4 \mathbb{I}(|x_i| > t)]]$$

$$= \quad \frac{mt^4}{n^2} + \left(\frac{m^2}{n^2}\right) \times \frac{\sum_{i=1}^{m} \mathrm{E}[x_i^4 \mathbb{I}(|x_i| > t)]}{m^2} \to 0.$$

Thus $\frac{x_1^2+x_2^2+\cdots+x_m^2}{n} \xrightarrow{P} 0$. This completes the proof. □

3.5.3 (M1) condition for RC_n^b

A pair matched word, of length $2k$, will be said to be *symmetric* if each letter appears once in an even position and once in an odd position. There are $k!$ symmetric words of length $2k$. We shall show that for any nonsymmetric word (3.6) equals zero and for any symmetric word, (3.6) equals 1.

Lemma 8. *(a) For L_H, if w is nonsymmetric, pair matched, then there exists a linear function $\Lambda(\pi(i),\ i \in S)$ such that all its coefficients are not zero and $\Lambda(\pi(i),\ i \in S) = 0$.*

(b) For L_{RC}, (a) holds with the modification $\Lambda(\pi(i),\ i \in S) = dn$ for some integer d.

Proof. Let $t_i = \pi(i-1) + \pi(i)$ and for any word, let S be the set of independent vertices. Following Bose and Sen (2008) [6] or BDJ (2006) [7], write $\pi(i)$ as a linear combination $L_i^H(\pi_S)$ of the independent vertices. Observe that, if $i \in S$ then $L_i^H(\pi_S) = \pi(i)$. Since the RC_n link function is $L(i,j) = i + j \mod n$, one can check that for every dependent vertex i, one must have $\pi(i) = L_i^H(\pi_S) + a_i n$ for some integer a_i.

(a) First note that, $\pi(0) - \pi(2k) = (t_1 + t_3 + t_5 + \cdots + t_{2k-1}) - (t_2 + t_4 + t_6 + \cdots + t_{2k}) = 0$. Since w is pair matched let $\{(i_s, j_s),\ s \geq 1\}$ be such that $w[i_s] = w[j_s],\ \ i_s < j_s$ and $i_s,\ s = 1, 2, \ldots, k$ are arranged in ascending order of i_s. Since w is not a symmetric word we have at least one pair (i_*, j_*) such that both of them are either at odd places or at even places. Now

$$(t_1 + t_3 + \cdots + t_{2k-1}) - (t_2 + t_4 + \cdots + t_{2k}) = 2 \times \Bigg(\sum_{\substack{i_s : (i_s, j_s) \\ \text{are both odd}}} t_{i_s} - \sum_{\substack{i_s : (i_s, j_s) \\ \text{are both even}}} t_{i_s} \Bigg). \tag{3.9}$$

If there is any dependent vertex $\pi(i)$ in the above expression we replace it by the linear combination $L_i^H(\pi_S)$. Let i_* be the largest index that appears in the above equation. Note that i^* is an independent vertex. As any dependent vertex can be expressed as a linear combination of independent vertices to its left, the coefficient of $\pi(i_*)$ is non-zero. Hence we obtain a linear function $\Lambda(\pi(i),\ i \in S)$ such that $\Lambda(\pi(i),\ i \in S) = 0$.

(b) For RC_n matrix,

$$(t_1 + t_3 + \cdots + t_{2k-1}) - (t_2 + t_4 + \cdots + t_{2k}) = 2 \times \Bigg(\sum_{\substack{i_s : (i_s, j_s) \\ \text{are both odd}}} t_{i_s} - \sum_{\substack{i_s : (i_s, j_s) \\ \text{are both even}}} t_{i_s} \Bigg) + cn \tag{3.10}$$

for some integer c. The proof can now be completed by repeating the arguments of part (a). □

Lemma 9. *For any nonsymmetric pair matched word w, $\lim_{n\to\infty} \frac{1}{nm^k} |\Pi^{b*}(w)| = 0$.*

Proof. By Lemma 8 there exists a non zero linear function $\Lambda = \Lambda(\pi(i),\ i \in S)$ such that $\Lambda = dn$ for some integer d. Let i_* be the largest index having non-zero coefficient in Λ. Now we choose vertices of π from left. First $\pi(0)$ has n choices. Next by Lemma 2(c), for any $1 \leq i \leq i^* - 1$, $\pi(i)$ has at most $m+1$ choices or at most one choice, depending on whether or not the vertex is independent. Having chosen all these vertices, the number of choices of $\pi(i_*)$ is $O(1)$. Next we move on to $\pi(i_* + 1)$ and proceed as before. Clearly, total number of possible choices is $O(nm^{k-1})$. This completes the proof. □

Lemma 10. *For any symmetric pair matched word, given the independent vertices, each dependent vertex has exactly one value.*

Proof. We determine the dependent vertices from left to right. Suppose a typical restriction due to (matched) word looks like $L_{RC}(\pi(i-1), \pi(i)) = L_{RC}(\pi(j-1), \pi(j))$ where $i < j$ and $\pi(i-1), \pi(i)$ and $\pi(j-1)$ have already been determined. We show that $\#\pi(j) = 1$. The above restriction can be rewritten as

$$\pi(j) = Z + dn \text{ for some integer } d \text{ and } Z = \pi(i-1) + \pi(i) - \pi(j-1).$$

Clearly $\pi(j)$ can be determined uniquely from the above equation since $1 \le \pi(j) \le n$. We proceed inductively from left to right to complete the proof. □

Lemma 11. *If w is pair matched symmetric, then $\lim_{n\to\infty} \frac{1}{nm^k}|\Pi^{b*}(w)| = 1$. Hence*

$$\lim \mathrm{E}[\beta_{2k}(m^{-1/2}RC_n^b)] = k!.$$

Proof. Obviously there are n valid choices for $\pi(0)$. From Lemma 2(c), for any other independent vertex $\pi(i)$, $m \le \#\pi(i) \le m+1$. From Lemma 10, given the independent vertices, every dependent vertex has only one value. We show that such a choice automatically satisfies the circuit condition. Let the partition generated by the word w be $\{(i_s, j_s),\ s \ge 1\}$, $i_s < j_s$. Let $t_i = \pi(i-1)+\pi(i)$. We will have $t_{i_s} = t_{j_s} + d_s n$, for some integer d_s. Since the word is symmetric, one of (i_s, j_s) will occur in an odd place and another will occur in an even place. Now for any choice of the independent vertices we will have,

$$\pi(0) - \pi(2k) = (t_1 + t_3 + t_5 + \cdots + t_{2k-1}) - (t_2 + t_4 + t_6 + \cdots + t_{2k}) = dn \text{ for some integer } d.$$

As $|\pi(0) - \pi(2k)| \le n-1$, d must be 0, proving that the circuit condition is satisfied. Hence $n \times m^k \le |\Pi^{b*}(w)| \le n \times (m+1)^k$. So we have, $\lim_{n\to\infty} \frac{1}{nm^k}|\Pi^{b*}(w)| = 1$. Since there are exactly $k!$ symmetric words, the Lemma is proved completely. □

3.5.4 (M1) condition for T_n^b, $\alpha \neq 0$

The following result is similar to result Lemma 11 and is essentially proved as Proposition 4.4 in [BDJ] [7]

Lemma 12. *Fix $k \in \mathbb{N}$. Let $\mathcal{M}$ be the number of L_T matched circuits π on $\{1, 2, \ldots, n\}$ of length $2k$ with at least one pair of L_T matched edges $(\pi(i-1), \pi(i))$ and $(\pi(j-1), \pi(j))$ such that $\pi(i) - \pi(i-1) + \pi(j) - \pi(j-1) \neq 0$. Let $\mathcal{M}^b$ be the same count with the extra condition $\mathbb{I}_\pi^{2k} = 1$. Then $n^{-(k+1)}\mathcal{M}^b \le n^{-(k+1)}\mathcal{M} \to 0$.*

Earlier we have shown that, $E[\beta_{2k+1}(n^{-1/2}T_n^b)] \to 0$. By Lemma 4 and 12, *if* the second limit below exists,

$$\lim E[\beta_{2k}(n^{-1/2}T_n^b)] = \lim \frac{1}{n^{1+k}} \sum_{\substack{w \text{ matched},\\ |w|=k}} |\Pi^{b*}(w)| \tag{3.11}$$

$$= \lim \frac{1}{n^{1+k}} \sum_{\substack{w \text{ matched},\\ |w|=k}} |\Pi^{b**}(w)| \tag{3.12}$$

$$= \lim \left(\frac{m}{n}\right)^k \times \frac{1}{nm^k} \sum_{\substack{w \text{ matched},\\ |w|=k}} |\Pi^{b**}(w)| = \alpha^k \times \beta_{2k}(T_\alpha^b) \text{ say,} \tag{3.13}$$

where

$$\Pi^{b**}(w) = \left\{\pi : w[i] = w[j] \Rightarrow \pi(i-1) - \pi(i) + \pi(j-1) - \pi(j) = 0 \text{ and } \mathbb{I}_\pi^{2k} = 1\right\}.$$

Taking $x_i = \pi(i)/n$, $|\Pi^{b**}(w)|$ can be expressed as $\#\Big\{(x_0, x_1, \cdots, x_{2k}) : x_0 = x_{2k}, x_i \in \{j/n,\ 1 \le j \le n,$
$|x_{i-1} - x_i| \le m$ and $x_{i-1} - x_i + x_{j-1} - x_j = 0$ if $w[i] = w[j]\Big\}$.

Let $S = \{0\} \cup \{\min(i,j) : w[i] = w[j], i \neq j\}$ be the set of all indices of the independent vertices of word w and let $x_S = \{x_i : i \in S\}$. Each x_i is a unique linear combination $L_i^T(x_S)$. L_i^T depend on the word w but for notational ease we suppress this dependence. Clearly $L_i^T(x_S) = x_i$ if $i \in S$ and also summing the k equations would imply $L_{2k}^T(x_S) = x_0$. So,

$$|\Pi^{b**}(w)| = \#\{x_S : L_i^T(x_S) \in \mathbb{N}_n \text{ and } |L_{i-1}^T(x_S) - L_i^T(x_S)| \le m/n \text{ for all } i = 0, 1, \cdots, 2k\}$$

where $\mathbb{N}_n = \{1/n, 2/n, \ldots, 1\}$. If $w[i] = w[j]$ then, $|L_{i-1}^T(x_S) - L_i^T(x_S)| = |L_{j-1}^T(x_S) - L_j^T(x_S)|$. So we can replace the restriction $|L_{i-1}^T(x_S) - L_i^T(x_S)| \le m/n,\ 1 \le i \le 2k$ by $|L_{i-1}^T(x_S) - L_i^T(x_S)| \le m/n$ for $i \in S$. Hence, as in Bose and Sen [6] and [BDJ] [7],

$$\frac{1}{n^{1+k}}|\Pi^{b**}(w)| = \mathrm{E}\left[\mathbb{I}\Big(0 \le L_i^T(U_{n,S}) \le 1,\ \forall\, i \notin S \cup \{2k\}\Big) \times \mathbb{I}\Big(|L_{i-1}^T(U_{n,S}) - U_{n,i}| \le \alpha_n\ \forall\, i \in S\Big)\right], \tag{3.14}$$

where $\alpha_n = m/n$ and for each $i \in S$, $U_{n,i}$ is a 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. Taking limits,

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n^{1+k}}|\Pi^{b**}(w)| &= \underbrace{\int_0^1\int_0^1\cdots\int_0^1}_{k+1} \prod_{i\notin S\cup\{2k\}} \mathbb{I}(0 \le L_i^T(x_S) \le 1) \times \prod_{i\in S} \mathbb{I}(|L_{i-1}^T(x_S) - x_i| \le \alpha) dx_S \\ &= p_{T_\alpha^b}(w) \ \text{(say)},\end{aligned}$$

and

$$\lim E[\beta_{2k}(m^{-1/2}T_n^b)] = \beta_{2k}(T_\alpha^b) = \alpha^{-k} \sum_{w:w \text{ pair matched}} p_{T_\alpha^b}(w)$$

where $p_{T_\alpha^b}$ is as above.

3.5.5 (M1) condition for H_n^b, $\alpha \neq 0$

Similar arguments as in the previous section lead to $\beta_{2k}(H_\alpha^b) = \alpha^{-k} \lim_{n\to\infty} \frac{1}{n^{1+k}} \sum_{w:w \text{ pair matched}, |w|=k} |\Pi^{b*}(w)|$

where

$$\Pi^{b*}(w) = \{\pi : w[i] = w[j] \Rightarrow \pi(i-1) + \pi(i) = \pi(j-1) + \pi(j) \text{ and } \mathbb{I}_\pi^{2k} = 1\}.$$

As before, each x_i is a unique linear combination $L_i^H(x_S)$ of the independent vertices and $L_i^H(x_S) = x_i$ for all $i \in S$ and $n^{-(k+1)}|\Pi^{b*}(w)|$ may be written as an expectation with independent discrete uniform on $\mathbb{R}^{k+1}$ and then we take limit. Unlike the previous case, we do not automatically have $L_{2k}^H(x_S) = x_{2k}$ for every word w and that explains the extra indicator in the integrand below.

$$p_{H_\alpha^b}(w) = \underbrace{\int_0^1\int_0^1\cdots\int_0^1}_{k+1} \prod_{i\notin S\cup\{2k\}} \mathbb{I}(0 \le L_i^H(x_S) \le 1) \times \prod_{i\in S} \mathbb{I}(L_{i-1}^H(x_S) + x_i \le \alpha) \times \mathbb{I}(x_0 = L_{2k}^H(x_S)) dx_S, \tag{3.15}$$

and $\beta_{2k}(H^b_\alpha) = \alpha^{-k} \sum_{w:w \text{ pair matched}} p_{H^b_\alpha}(w)$.

From the work of Bose and Sen [6] and [BDJ] [7], $\mathbb{I}(x_0 = x_{2k}) = 1$ iff w is a symmetric word. Since for all other words, the restriction $\mathbb{I}(x_0 = x_{2k}) = 1$ is one extra linear restriction, the integral above is zero and hence $p_{H^b_\alpha}(w) = 0$ for all non symmetric words. Thus we can write

$$\beta_{2k}(H^b_\alpha) = \alpha^{-k} \sum_{\substack{w:w \text{ symmetric} \\ \text{pair matched}}} p_{H^b_\alpha}(w).$$

3.5.6 (M1) condition for H^b_n, $\alpha = 0$

Lemma 13. *Fix any word w such that $l(w) = 2k$ and $|w| = k$. Then, $\lim_{n\to\infty} \frac{1}{nm^k}|\Pi^b(w)| = 0$.*

Proof. Since the link function is $L(i, j) = i + j$, to have $\mathbb{I}^h_\pi = 1$, every independent vertex, including $\pi(0)$, has at most m choices and every dependent vertices has at most one choice. This proves that $|\Pi^b(w)| = O(m^{k+1})$ and hence we get the result. □

Hence we get $\lim E[\beta_h(m^{-1/2}H^b_n)] = 0$ for all h and the limit is degenerate.

3.5.7 (M1) condition for RC^B_n and H^B_n

The first lemma shows that nonsymmetric words do not contribute to the limit of RC^B_n.

Lemma 14. *For RC^B_n if w is non-symmetric pair matched of length $2k$ then, $\lim_{n\to\infty} \frac{1}{nm^k}|\Pi^{B*}(w)| = 0$.*

The next Lemma yields the LSD for RC^B_n.

Lemma 15. *For RC^B_n, if w is symmetric pair matched with $|w| = k$, then, $\lim_{n\to\infty} \frac{1}{nm^k}|\Pi^{B*}(w)| = 2^k$. Hence*

$$\lim \mathrm{E}[\beta_{2k}((2m_n)^{-1/2}RC^B_n)] = k! = \mathrm{E}[R^{2k}].$$

The proofs of the above two Lemma are omitted.

For H^B_n, when $\alpha \neq 0$, proceeding as in Section 3.5.5,

$$\beta_{2k}(H^B_\alpha) = \lim_{n\to\infty} \frac{1}{n(2m_n)^k} \sum_{w:\ w \text{ matched, } |w|=k} |\Pi^{B**}(w)|, \quad \text{where,}$$

$|\Pi^{B**}(w)| = \#\{(x_0, x_1, \ldots, x_{2k}) : x_0 = x_{2k}, x_i \in \{1/n, 2/n, \ldots, (n-1)/n, 1\} n - m \leq x_{i-1} + x_i \leq n + m$ and $x_{i-1} + x_i = x_{j-1} + x_j$ if $w[i] = w[j]\}$.

Hence

$$\frac{1}{n^{1+k}}|\Pi^{B*}(w)| = \mathrm{E}\Big[\mathbb{I}\big(0 \leq L^H_i(U_{n,S}) \leq 1,\ \forall\, i \notin S \cup \{2k\}\big) \times \mathbb{I}\big(|L^H_{i-1}(U_{n,S}) + U_{n,i} - 1| \leq \frac{m}{n}\ \forall\, i \in S\big) \times \mathbb{I}(x_0 = x_{2k})\Big], \tag{3.16}$$

$$\frac{1}{n^{1+k}}|\Pi^{B*}(w)| \quad \to \quad \underbrace{\int_0^1 \int_0^1 \cdots \int_0^1}_{k+1} \prod_{i \notin S \cup \{2k\}} \mathbb{I}(0 \leq L^H_i(x_S) \leq 1) \times \prod_{i \in S} \mathbb{I}(|L^H_{i-1}(x_S) + x_i - 1| \leq \alpha)$$

$$\times \mathbb{I}(x_0 = x_{2k}) dx_S = p_{H^B_\alpha}(w) \text{ say.}$$

Hence $\beta_{2k}(H^B_\alpha) = (2\alpha)^{-k} \sum_{\substack{w:w \text{ matched}, \\ |w|=k}} p_{H^B_\alpha}(w)$ for $\alpha \neq 0$.

Now we will deal with the case $\alpha = 0$.

Lemma 16. *(i) Suppose Assumption I (i) or (ii) holds and $\alpha = 0$. Then $d^2_{BL}(F^{m^{-1/2}H^B_n}, F^{m^{-1/2}RC^B_n}) \xrightarrow{a.s.} 0$.*
(i) Suppose Assumption I (i) and (ii) holds and $\alpha = 0$. Then $d^2_{BL}(F^{m^{-1/2}H^B_n}, F^{m^{-1/2}RC^B_n}) \xrightarrow{P} 0$.*

Proof. We construct another pair of RC^B_n and H^B_n matrices as follows: Let

$$y_{i+j,n} = \begin{cases} x_{i+j} & \text{if } i+j \leq n-1 \\ x_{i+j-n} & \text{if } i+j \geq n. \end{cases} \tag{3.17}$$

$$\widetilde{H}^B_n(i,j) = \begin{cases} y_{i+j,n} & \text{if } |i+j-n| \leq m \\ 0 & \text{otherwise .} \end{cases} \tag{3.18}$$

$$RC^B_n(i,j) = \begin{cases} y_{i+j,n} & \text{if } i+j \mod n \leq m \text{ or } i+j \mod n \geq n-m \\ 0 & \text{otherwise.} \end{cases} \tag{3.19}$$

As $m/n \to 0$ for sufficiently large n, $m < n/2$. Hence $y_{i_1+j_1,n} \neq y_{i_2+j_2,n}$ for any (i_1, j_1) and (i_2, j_2) such that $i_1 + j_1 \leq n-1 < i_2 + j_2$.

It is not hard to see that for sufficiently large n,

$$\mathrm{E}[\beta_h(m^{-1/2}H^B_n)] = \mathrm{E}[\beta_h(m^{-1/2}\widetilde{H}^B_n)]$$

and

$$\mathrm{E}[\beta_h(m^{-1/2}H^B_n) - \mathrm{E}(\beta_h(m^{-1/2}H^B_n))]^4 = \mathrm{E}[\beta_h(m^{-1/2}\widetilde{H}^B_n) - \mathrm{E}(\beta_h(m^{-1/2}\widetilde{H}^B_n))]^4.$$

for each h. This proves that if LSD of one of these sequences exist, then so does the other and moreover if, almost sure convergence (via (M4)) holds or, in probability convergence (via (M2)) holds, for one sequence, then the same holds for the other. Hence it is enough to show that

$$d^2_{BL}(F^{m^{-1/2}\widetilde{H}^B_n}, F^{m^{-1/2}RC^B_n}) \xrightarrow{a.s.} 0.$$

On the other hand the LSD of the above RC matrix is same as obtained before.

Now

$$\begin{aligned} d^2_{BL}(F^{m^{-1/2}\widetilde{H}^B_n}, F^{m^{-1/2}RC^B_n}) &\leq \frac{1}{n} \mathrm{Tr}\left[\frac{(\widetilde{H}^B_n - RC^B_n)^2}{m}\right] \\ &= \frac{1}{mn}[(m-1)x^2_m + \cdots + x^2_2] + \frac{1}{mn}[x^2_0 + x^2_{n-1} + \cdots + (m+1)x^2_{n-m}] \\ &\leq \frac{1}{n}(x^2_0 + x^2_2 + \cdots + x^2_m) + \frac{2}{n}(x^2_{n-m} + \cdots + x^2_{n-1}) \\ &= \left(\frac{m}{n}\right) \times \frac{x^2_0 + x^2_2 + \cdots + x^2_m}{m} + \frac{2}{n}(x^2_{n-m} + \cdots + x^2_{n-1}). \end{aligned}$$

(i) First assume $\{x_i\}$ are i.i.d. As $m/n \to 0$, by SLLN, the first term in the above expression $\to 0$ and, the second term is,

$$\begin{aligned}\frac{x_{n-m}^2 + \cdots + x_{n-1}^2}{n} &= \frac{x_0^2 + x_1^2 + \cdots + x_{n-1}^2}{n} - \frac{x_0^2 + x_1^2 + \cdots + x_{n-m-1}^2}{n-m}\frac{n-m}{n} \\ &\stackrel{a.s.}{\to} \mathrm{E}[x_0^2] - \mathrm{E}[x_0^2](1) = 0.\end{aligned}$$

If the input sequence is uniformly bounded, a slight modification of the above argument yields the result.

(ii) Now assume that the input sequence $\{x_i\}$ satisfy Assumption I* (i) and (ii). We then have,

$$d_{BL}^2\Big(F^{m^{-1/2}\tilde{H}_n^B}, F^{m^{-1/2}RC_n^B}\Big) \le \frac{x_0^2 + x_1^2 + \cdots + x_m^2}{n} + \frac{2}{n}(x_{n-m}^2 + \cdots + x_{n-1}^2).$$

Using ideas in Lemma 7 we infer easily that, $\frac{x_0^2+x_1^2+\cdots+x_m^2}{n} \to 0$ in probability.

We will show that $\frac{1}{n}(x_{n-m}^2 + \cdots + x_{n-1}^2) \stackrel{P}{\to} 0$. First note that $\frac{1}{n}\,\mathrm{E}(x_{n-m}^2 + \cdots + x_{n-1}^2) \to 0$. On the other hand,

$$\frac{1}{n^2}\,\mathrm{Var}(x_{n-m}^2 + \cdots + x_{n-1}^2) \le \frac{\sum_{i=n-m}^{n-1} \mathrm{E}[x_i^4]}{n^2} \le \frac{\sum_{i=1}^{n-1} \mathrm{E}[x_i^4 \mathbb{I}(|x_i| > t)]}{n^2} + \frac{1}{n}t^4 \to 0,$$

using Assumption I* (ii). So the proof is complete. □

3.5.8 (M1) Condition for T_n^B ($\alpha \neq 0$)

The following Lemma follows immediately from the work of Bryc, Dembo and Jiang (2006)[7]. This lemma on slopes is similar to Lemma 12 proved earlier for T_n^b.

Lemma 17. *Let $\mathcal{M}$ be the number of L_T matched circuits π on $\{1, 2, \ldots, n\}$ of length $2k$ with at least one pair of L^T matched edges $(\pi(i-1), \pi(i))$ and $(\pi(j-1), \pi(j))$ such that $\pi(i) - \pi(i-1) + \pi(j) - \pi(j-1) \neq 0$. Let $\mathcal{M}^B$ be the same count with the extra condition $\mathbb{I}_{\pi,T}^{2k} = 1$. Then $n^{-(k+1)}\mathcal{M}^B \le n^{-(k+1)}\mathcal{M} \to 0$.*

Now we proceed to verify the (M1) condition and drive the limit. Earlier we have shown that, $\mathrm{E}[\beta_{2k+1}(n^{-1/2}T_n^B)] \to 0$. Further, *if* the limit exists,

$$\beta_{2k}(T_\alpha^B) = \lim_{n\to\infty} \mathrm{E}[\beta_{2k}(m^{-1/2}T_n^B)] = \alpha^{-k} \lim_{n\to\infty} \frac{1}{n^{1+k}} \sum_{\substack{w:w \text{ matched,} \\ |w|=k}} |\Pi^B(w)|.$$

Using Lemma 4 and Lemma 17,

$$\lim_{n\to\infty} \frac{1}{n^{1+k}} |\Pi^B(w)| = \lim_{n\to\infty} \frac{1}{n^{1+k}} |\Pi^{B*}(w)| = \lim_{n\to\infty} \frac{1}{n^{1+k}} |\Pi^{B**}(w)|.$$

Taking $x_i = \pi(i)/n$, $|\Pi^{b**}(w)| = \#\Big\{(x_0, x_1, \cdots, x_{2k}) : x_0 = x_{2k}, x_i \in \{1/n, 2/n, \ldots, (n-1)/n, 1\}, |x_{i-1} - x_i| \le m$ or $|x_{i-1} - x_i| \ge n - m$ and $x_{i-1} - x_i + x_{j-1} - x_j = 0$ if $w[i] = w[j]\Big\}$.

Let $S = \{0\} \cup \{\min(i, j) : w[i] = w[j], i \neq j\}$ be the set of all indices corresponding to the independent vertices. Let $x_S = \{x_i : i \in S\}$. Since x_i's satisfy k equations, each is a unique linear combination $L_i^T(x_S)$

of x_j's, $j \in S, \quad j \leq i$. These L_i^T's depend on the word w but we suppress this dependence. Clearly $L_i^T(x_S) = x_i$ if $i \in S$ and also summing the k equations would imply $L_{2k}^T(x_S) = x_0$. So,

$$|\Pi^{B**}(w)| = \# \quad \big\{x_S : L_i^T(x_S) \in \mathbb{N}_n \text{ and } |L_{i-1}^T(x_S) - L_i^T(x_S)| \leq m/n \\ \text{or } |L_{i-1}^T(x_S) - L_i^T(x_S)| \geq 1 - m/n \text{ for all } i = 0, 1, \cdots, 2k\big\}$$

where $\mathbb{N}_n = \{1/n, 2/n, \ldots, 1\}$. Since,

$$m \leq n/2 \times |L_{i-1}^T(x_S) - L_i^T(x_S)| \leq m/n \tag{3.20}$$

$$\text{or} \qquad |L_{i-1}^T(x_S) - L_i^T(x_S)| \geq 1 - m/n \tag{3.21}$$

can be written as $||L_{i-1}^T(x_S) - L_i^T(x_S)| - 1/2| \geq 1/2 - m/n$, we can write $|\Pi^{B**}(w)|$ as

$$\#\big\{x_S : L_i^T(x_S) \in \mathbb{N}_n \text{ and } ||L_{i-1}^T(x_S) - L_i^T(x_S)| - 1/2| \geq 1/2 - m/n \text{ for all } 0 \leq i \leq 2k\big\}.$$

Now we note, if $w[i] = w[j]$ then we have, $|L_{i-1}^T(x_S) - L_i^T(x_S)| = |L_{j-1}^T(x_S) - L_j^T(x_S)|$. So we can replace the restriction $||L_{i-1}^T(x_S) - L_i^T(x_S)| - 1/2| \geq 1/2 - m/n$, for all $i = 1, 2, \ldots, 2k$ by $||L_{i-1}^T(x_S) - L_i^T(x_S)| - 1/2| \geq 1/2 - m/n$ for $i \in S$. Hence

$$\frac{1}{n^{1+k}}|\Pi^{B**}(w)| = \mathrm{E}\Big[\mathbb{I}\Big(0 \leq L_i^T(U_{n,S}) \leq 1, \ \forall\, i \notin S \cup \{2k\}\Big) \tag{3.22}$$

$$\times \mathbb{I}\Big(||L_{i-1}^T(U_{n,S}) - U_{n,i}| - 1/2| \geq 1/2 - \alpha_n \ \forall\, i \in S\Big)\Big], \tag{3.23}$$

where $\alpha_n = m/n$ and taking limits,

$$\frac{1}{n^{1+k}}|\Pi^{B**}(w)| \rightarrow \underbrace{\int_0^1 \int_0^1 \cdots \int_0^1}_{k+1} \prod_{i \notin S \cup \{2k\}} \mathbb{I}(0 \leq L_i^T(x_S) \leq 1) \times \prod_{i \in S} \mathbb{I}(||L_{i-1}^T(x_S) - x_i| - 1/2| \geq 1/2 - \alpha) dx_S$$

$$= p_{T_\alpha^B}(w) \text{ say.}$$

Hence $\beta_{2k}(T_\alpha^B) = \alpha^{-k} \sum_{w:w \text{ pair matched}} p_{T_\alpha^B}(w)$.

3.5.9 Closeness of T_n^B and SC_n^b when $\alpha = 0$

The next Lemma and its proof are similar to Lemma 7 and its proof.

Lemma 18. *(i) Suppose Assumption I (i) or (ii) holds and* $\alpha = 0$*. Then,* $\Delta_n \equiv d_{BL}^2(F^{m^{-1/2}T_n^B}, F^{m^{-1/2}SC_n^b}) \xrightarrow{a.s.} 0$.

(ii) Suppose Assumption I (i) and (ii) holds and* $\alpha = 0$*. Then,* $\Delta_n \equiv d_{BL}^2(F^{m^{-1/2}T_n^B}, F^{m^{-1/2}SC_n^b}) \xrightarrow{P} 0$.

Proof. Both SC_n^b and T_n^B have all upper and lower k^{th} diagonal entries equal to x_k, for $0 \leq k \leq m$. Also note that SC_n^b have all the upper and lower k^{th} diagonal entries equal to x_{n-k} whereas T_n^B have all the upper and lower k^{th} diagonal entries equal to x_k for $n - m \leq k \leq n - 1$. Hence,

$$\Delta_n \leq \frac{1}{n}\operatorname{Tr}\left[\frac{(T_n^B - SC_n^b)^2}{m}\right]$$

$$= \frac{2}{mn}(m(x_{n-m} - x_m)^2 + (m-1)(x_{n-m+1} - x_{m-1})^2 + \cdots + (x_{n-1} - x_1)^2)$$

$$
\begin{aligned}
&\leq \frac{2}{n}((x_{n-m}-x_m)^2+(x_{n-m+1}-x_{m-1})^2+\cdots+(x_{n-1}-x_1)^2) \\
&\leq \frac{4}{n}(x_m^2+x_{m-1}^2+\cdots+x_1^2)+\frac{4}{n}(x_{n-m}^2+x_{n-m+1}^2+\cdots+x_{n-1}^2) \\
&= 4\times\left(\frac{m}{n}\right)\times\frac{x_1^2+x_2^2+\cdots+x_m^2}{m}+4\times\frac{x_{n-m}^2+x_{n-m+1}^2+\cdots+x_{n-1}^2}{n}.
\end{aligned}
$$

If $\{x_i\}$ are i.i.d., using SLLN, both the terms in the above expression go to zero a.s.. If $\{x_i\}$ is uniformly bounded, a slight modification of this argument yields the required result. If $\{x_i\}$ follow Assumption I* (i) and (ii) then using exactly same ideas as in Lemma 16 a proof can be obtained. Detailes are omitted. □

3.5.10 DH_n^b, PH_n^b and PT_n^b matrices

Note that $PH_n^b J_n = J_n PH_n^b = PT_n^b$ where J_n is the matrix with entries 1 in the main anti-diagonal and zero elsewhere. Since $J_n^2 = I_n$ $(PT_n^b)^{2k} = (PH_n^b)^{2k}$. Hence LSD of PT_n^b and PH_n^b, if they exist, are same. Moreover one can easily check that (i) the $n \times n$ principal minor of PH_{n+1}^b is DH_n^b and (ii) the $n \times n$ principal minor of PT_{n+1}^b is SC_n^b. Hence by Fact 2 all these four Type I band matrices have the same LSD. Since ESD of $m_n^{-1/2} SC_n^b \Rightarrow N(0,2)$ (provided (M2) or (M4) condition is satisfied), the same is true for the other three matrices.

3.6 Carleman condition

Lemma 19. *Suppose L_{2k} stands for any of the limits obtained in the previous section. Then there exists constants C, such that $L_{2k} \leq C^k \frac{(2k)!}{k!}$ and hence Carleman condition is satisfied.*

Proof. The proofs for all the matrices are similar. Here is an outline. Note that $\pi(0)$ has at most n possible values and each remaining independent vertex has at most $C_1 m$ (even when $\alpha \neq 0$, m and n are of the same order) possible values where C_1 is a constant. Each dependent vertex has at most C_2 possible values for some constant C_2. Hence we get,

$$
\sum_{w \text{ pair matched}} \lim_{n\to\infty} \frac{1}{nm^k}|\Pi^b(w)| \leq \sum_{w \text{ pair matched}} \lim_{n\to\infty} \frac{(C_1C_2)^k n(m)^k}{nm^k} = \frac{2k!}{k!2^k} \times (C_1C_2)^k. \tag{3.24}
$$

The right side are the moments of $N(0, C_1C_2)$ variable. This completes the proof. □

3.7 (M2) and (M4) conditions

Let $Q_{h,4}$ be the number of circuits $(\pi_1,\pi_2,\pi_3,\pi_4)$ of length h which are jointly and cross matched with respect to link function L and let $Q_{h,4}^b$ be the same count with the added restriction $\mathbb{I}_{\pi_i}^h = 1,\ i = 1,2,3,4$.

Lemma 20. *(a) For L_{SC} or L_{RC}, there exists a constant K depending on L such that, $Q_{h,4}^b \leq Kn^4m^{2h-2}$.*

(b) If the input sequence is uniformly bounded, independent, with mean zero and variance one, and $\{A_n^b = SC_n^b$ or $RC_n^b\}$ then,

$$
\mathrm{E}\left[\frac{1}{n}\mathrm{Tr}\left(\frac{A_n^b}{\sqrt{m}}\right)^h - \mathrm{E}\frac{1}{n}\mathrm{Tr}\left(\frac{A_n^b}{\sqrt{m}}\right)^h\right]^4 = O\Big(\frac{1}{m^2}\Big). \tag{3.25}
$$

Hence if $m_n \to \infty$, (M2) holds. If $\sum\limits_{n=1}^{\infty} \frac{1}{m_n^2} < \infty$, then (M4) holds. Same conclusions hold for $\{RC_n^B\}$.

(c) If the input sequence is uniformly bounded, independent, with mean zero and variance one, and $\{A_n^b = SC_n^b \text{ or } RC_n^b\}$ then,

$$\mathrm{Var}[\beta_h(A_n^b)] = \mathrm{E}\Big[\frac{1}{n}\,\mathrm{Tr}\Big(\frac{A_n^b}{\sqrt{m}}\Big)^h - \mathrm{E}\,\frac{1}{n}\,\mathrm{Tr}\Big(\frac{A_n^b}{\sqrt{m}}\Big)^h\Big]^2 = O\Big(\frac{1}{m}\Big). \tag{3.26}$$

Hence (M2) holds. Same conclusions hold for $\{RC_n^B\}$.

(d) For L_T or L_H, there exists a constant K depending on L such that, $Q_{h,4}^b \le Kn^{2h+2}$.

(e) If the input sequence is uniformly bounded, independent, with mean zero and variance one, and $\{A_n^b = T_n^b \text{ or } H_n^b\}$ then,

$$\mathrm{E}\Big[\frac{1}{n}\,\mathrm{Tr}\Big(\frac{A_n^b}{\sqrt{n}}\Big)^h - \mathrm{E}\,\frac{1}{n}\,\mathrm{Tr}\Big(\frac{A_n^b}{\sqrt{n}}\Big)^h\Big]^4 = O\Big(\frac{1}{n^2}\Big). \tag{3.27}$$

Hence (M4) holds. Same conclusion holds for $\{T_n^B\}$ and $\{H_n^B\}$.

(f) For $\alpha = 0$ and Hankel link function, there exists a constant K depending on the link function such that for sufficiently large n, $Q_{h,4}^b \le K_m^{2h+2}$.

Hence for sufficiently large n,

$$\mathrm{E}\Big[\frac{1}{n}\,\mathrm{Tr}\Big(\frac{A_n^b}{\sqrt{m}}\Big)^h - \mathrm{E}\,\frac{1}{n}\,\mathrm{Tr}\Big(\frac{A_n^b}{\sqrt{m}}\Big)^h\Big]^4 = O\Big(\frac{1}{n^2}\Big). \tag{3.28}$$

Proof. (a) This proof is exactly same as the proof of Proposition 4.3 in [BDJ] [7]), the only difference being that, excluding $\pi(0)$, all remaining independent vertices can be chosen in at most $m+1$ ($2m+1$ ways) for SC_n^b (RC_n^b matrices). We omit the details.

(b) We write the fourth moment as

$$\frac{1}{n^4 m^{2h}}\,\mathrm{E}[\mathrm{Tr}(A_n^b)^h - \mathrm{E}(\mathrm{Tr}(A_n^b)^h)]^4 = \frac{1}{n^4 m^{2h}} \sum_{\pi_1,\pi_2,\pi_3,\pi_4} \mathrm{E}[\prod_{i=1}^{4}(\mathbb{X}_{\pi_i} - \mathrm{E}\,\mathbb{X}_{\pi_i})] \times \prod_{i=1}^{4} \mathbb{I}_{\pi_i}^h.$$

Using independence, it follows easily that whenever $(\pi_1, \pi_2, \pi_3, \pi_4)$ (i) are not jointly matched, or (ii) are jointly matched but is not cross cross-matched, the corresponding expectation in the above expression is zero. Since the entries are bounded, so is $\mathrm{E}[\prod_{i=1}^4(\mathbb{X}_{\pi_i} - \mathrm{E}\,\mathbb{X}_{\pi_i})]$. Therefore by part (a),

$$\mathrm{E}\left[\frac{1}{n}\,\mathrm{Tr}\left(\frac{A_n^b}{\sqrt{m}}\right)^h - \mathrm{E}\,\frac{1}{n}\,\mathrm{Tr}\left(\frac{A_n^b}{\sqrt{m}}\right)^h\right]^4 \le K\frac{n^4 m^{2h-2}}{n^4 (m^{h/2})^4} = O\Big(\frac{1}{m^2}\Big). \tag{3.29}$$

(c) This is an easy consequence of Cauchy Schwartz inequality and part (b). Observe,

$$\mathrm{E}\Big[\frac{1}{n}\,\mathrm{Tr}\Big(\frac{A_n^b}{\sqrt{m}}\Big)^h - \mathrm{E}\,\frac{1}{n}\,\mathrm{Tr}\Big(\frac{A_n^b}{\sqrt{m}}\Big)^h\Big]^2 \le \left(\mathrm{E}\Big[\frac{1}{n}\,\mathrm{Tr}\Big(\frac{A_n^b}{\sqrt{m}}\Big)^h - \mathrm{E}\,\frac{1}{n}\,\mathrm{Tr}\Big(\frac{A_n^b}{\sqrt{m}}\Big)^h\Big]^4\right)^{\frac{1}{2}}. \tag{3.30}$$

Now using part (b) completes the proof of (c).

(d) First note that $Q_{h,4}^b \le Q_{h,4}$. Following the proof of Propostion 4.3 in [BDJ] [7], we obntain $Q_{h,4} \le Kn^{2h+2}$. This proves (d).

To prove (e) we need to repeat the arguments of (c). Details are omitted.

(f) Proof is exactly same as the proof of Proposition 4.3 in [BDJ] [7], the difference being that, as $\alpha = 0$, for sufficiently large n we will have $m < n$ and hence each independent vertex can be chosen in at most m ways. The proof of (3.28) follows by repeating the arguments in (c). □

This completes all the steps in the proof of Theorem 1. □

4 Some properties of the LSD

Tables 2–5 consider the situations where the LSD are not known explicitly. Using the results obtain the results in Section 3.5, and performing the necessary integrations (details of which are available on request), we provide the contribution from the different words for moments of order two and four and adding these contributions provide the moments of the LSD. It may also be noted that if we follow the threads of proofs in Section 3.5, then it is clear that the even moments of the LSD of $n^{-1/2}T_n^b$ and $n^{-1/2}H_n^b$ increase with α.

We also did some modest amount of simulations. The results of these simulations are exhibited in the figures. Figures 3, 4 and 5 show that the LSD of (scaled) T_n^b and H_n^b clearly depend on α. More interestingly, Figures 4 and Figure 5 suggest that the LSD of $n^{-1/2}H_n^b$ is bimodal if $\alpha \geq 4/3$ and there is a positive mass at zero if $\alpha \leq 5/4$. It will be interesting to investigate theoretically if this is indeed true and if any further properties of the LSD may be derived in the cases where the LSD are not known explicitly.

Acknowledgement. We thank Rajat Subhra Hazra and Koushik Saha for helpful discussions.

References

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

[2] Basak, Anirban (2009). Large dimensional random matrices. *M.Stat. Project Report*, May 2009. Indian Statitstical Institute, Kolkata. *Math. Notes* 50 (1991), 12321242.

[3] Bose, Arup and Basak, Anirban (2009). Limiting spectral distribution of some band matrices. Technical report No R16/2009, September 11, 2009, Stat-Math Unit, Indian Statistical Institute, Kolkata. Available at www.isical.ac.in/∼statmath.

[4] Bose, Arup; Gangopadhyay, Sreela and Sen, Arnab (2009). Limiting spectral distribution of XX' matrices. To appear in *Annales de l'Institut Henri Poincaré*.

[5] Bose, Arup and Mitra, J. (2002). Limiting spectral distribution of a special circulant. *Stat. Probab. Letters*, 60, 1, 111-120.

[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.

[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*

[8] Gray, Robert M. (2009). *Toeplitz and Circulant Matrices: A review*. Now Publishers, Norwell, Massachusetts.

[9] Grenander, Ulf and Szego, Gabor (1958). *Toeplitz forms and their applications*. California Monographs in Mathematical Sciences University of California Press, Berkeley-Los Angeles.

[10] 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.

[11] Kargin, Vladislav (2009). Spectrum of random Toeplitz matrices with band structures. *Elect. Comm. in Probab*. 14 (2009), 412421.

[12] Liu, Dang-Zheng and Wang, Zheng-Dong (2009). Limit Distributions for Random Hankel, Toeplitz Matrices and Independent Products. Available at *http://arxiv.org/abs/0904.2958*

[13] Popescu, Ionel (2009). General tridiagonal random matrix models, limiting distributions and fluctuations. *Probab. Theory Relat. Fields*, 144, 179220.

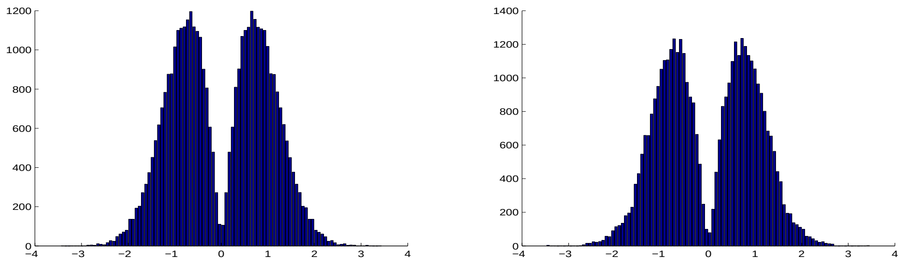

Figure 1: Histograms of the ESD of 100 realizations of RC_n^b, $n = 400$, with Normal$(0,1)$ entries for $\alpha = 0.5$ (left) and $\alpha = 0$ (right).

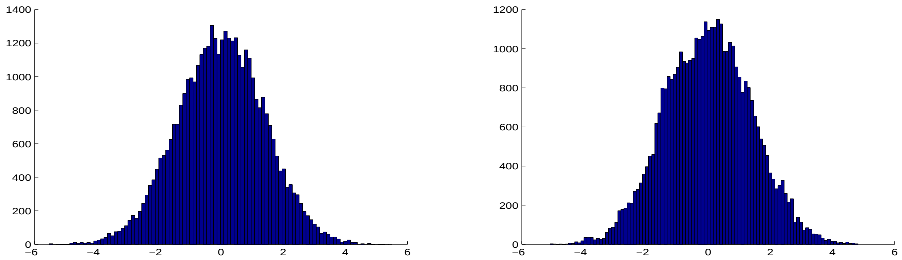

Figure 2: Histograms of the ESD of 100 realizations of SC_n^b, $n = 400$ with Normal$(0,1)$ entries for $\alpha = 1/3$ (left) and $\alpha = 0$ (right).

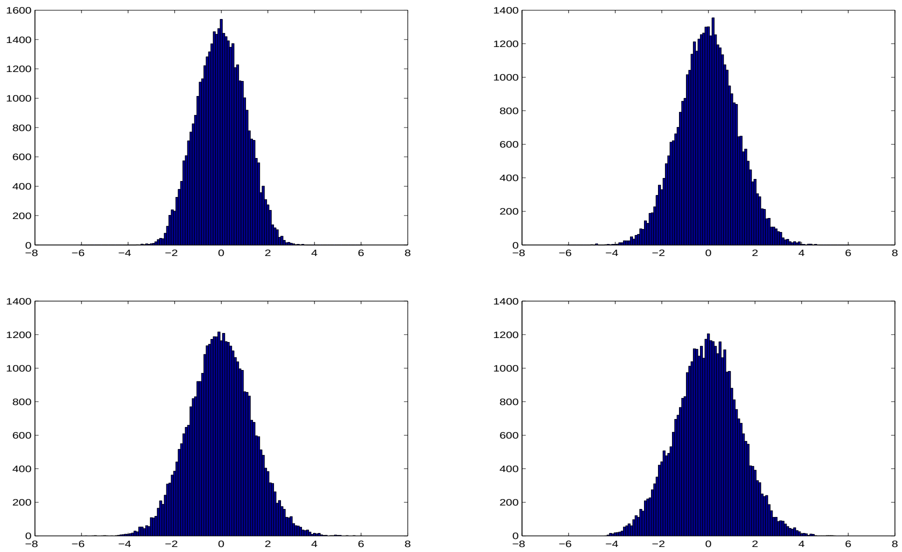

Figure 3: Histograms of the ESD of 100 realizations of T_n^b, $n = 400$, with Normal$(0, 1)$ entries for $\alpha = 1$ (`top left`), $\alpha = 0.5$ (`top right`), $\alpha = 0.25$ (`bottom left`) and $\alpha = 0$ (`bottom right`).

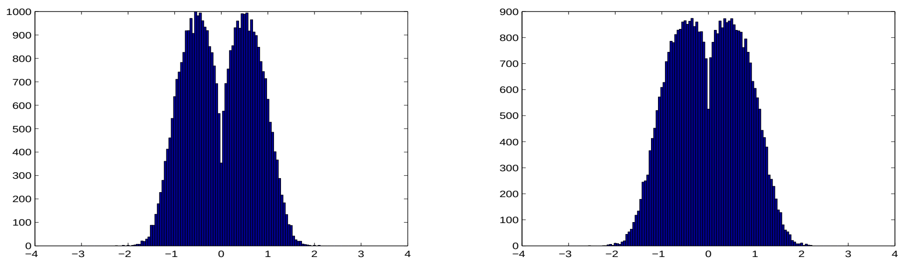

Figure 4: Histograms of the ESD of 100 realizations of H_n^b, $n = 400$, with Normal$(0, 1)$ entries for $\alpha = 2$ (`left`) and $\alpha = 4/3$ (`right`).

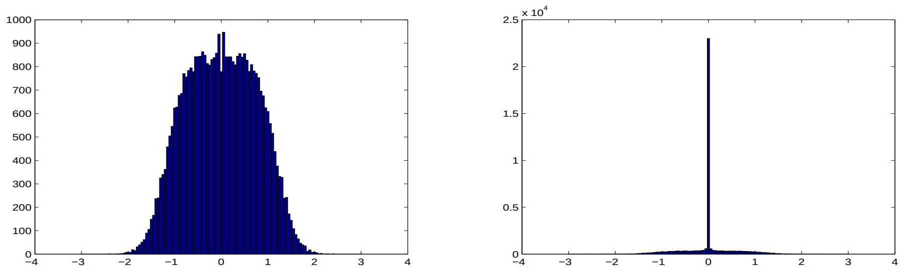

Figure 5: Histograms of the ESD of 100 realizations of H_n^b, $n = 400$, with Normal$(0, 1)$ entries for $\alpha = 5/4$ (`left`) and $\alpha = 1/2$ (`right`).

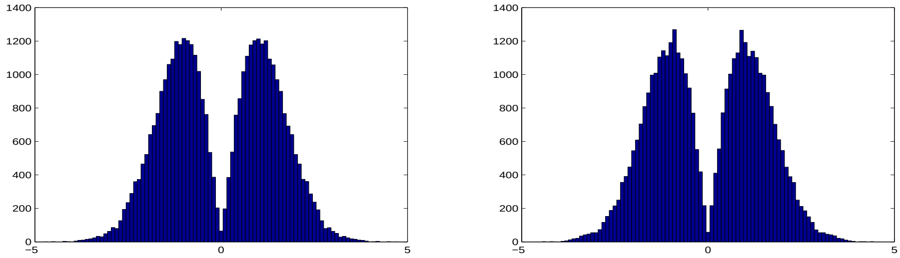

Figure 6: Histograms of the ESD of 100 realizations of RC_n^B, $n = 400$ with Normal$(0,1)$ entries for $\alpha = 1/2$ (`left`) and $\alpha = 0$ (`right`).

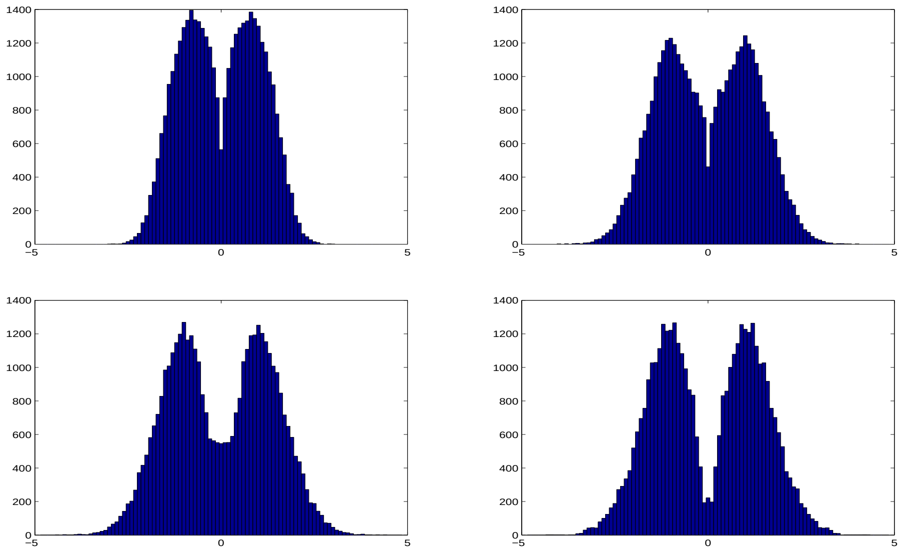

Figure 7: Histograms of the ESD of 100 realizations of H_n^B, $n = 400$, with Normal$(0,1)$ entries for $\alpha = 1$ (`top left`), $\alpha = 0.5$ (`top right`), $\alpha = 0.25$ (`bottom left`) and $\alpha = 0$ (`bottom right`).

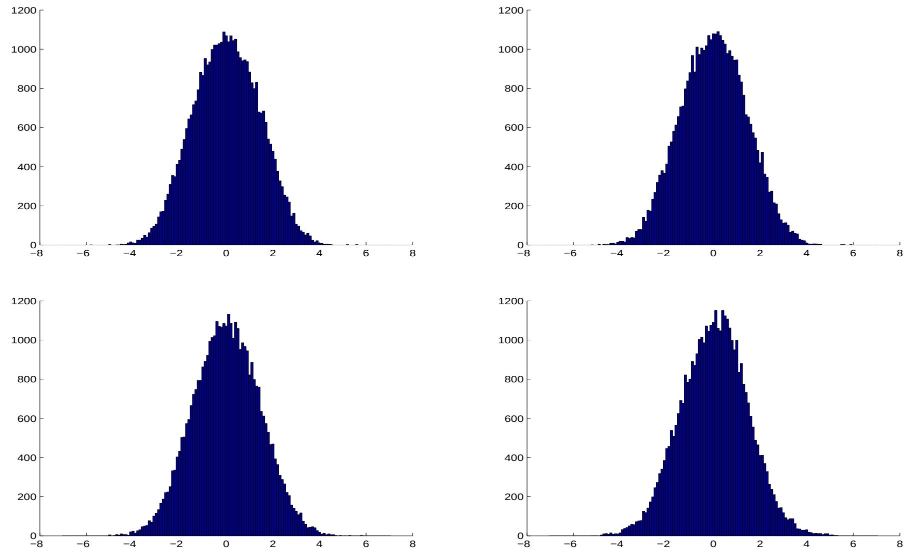

Figure 8: Histograms of the ESD of 100 realizations of T_n^B, $n = 400$, with Normal$(0,1)$ entries for $\alpha = 1/2$ (`top left`), $\alpha = 1/3$ (`top right`), $\alpha = 1/4$ (`bottom left`) and $\alpha = 0$ (`bottom right`).

Table 2: Type I Toeplitz limit ($\alpha \neq 0$). Word contributions and moments of order 2 and 4.

α	w	Indicator set	$p_{T^b_\alpha}(w)$	Moments
$\alpha \in (0,1)$	aa	$\lvert X_0 - X_1 \rvert \leq \alpha$	$\alpha(2-\alpha)$	$\beta_2(T^b_\alpha) = (2-\alpha)$
$\alpha \in (0,1/2]$	$abba$ $aabb$ $abab$	$\lvert X_0 - X_1 \rvert \leq \alpha, \lvert X_1 - X_2 \rvert \leq \alpha$ $\lvert X_0 - X_1 \rvert \leq \alpha, \lvert X_1 - X_2 \rvert \leq \alpha,$ $0 \leq X_0 - X_1 + X_2 \leq 1$	$\frac{2}{3}\alpha^2(6-5\alpha)$ same as above $4\alpha^2(1-\alpha)$	$\beta_4(T^b_\alpha) = \frac{4}{3}(9-8\alpha)$
$\alpha \in [1/2,1]$	$abba$ $aabb$ $abab$		$\frac{-1+6\alpha-2\alpha^3}{3}$ same as above $\frac{2}{3}(2\alpha^3-6\alpha^2+6\alpha-1)$	$\beta_4(T^b_\alpha) = \frac{4(-3\alpha^2+6\alpha-1)}{3\alpha^2}$

Table 3: Type I Hankel limit ($\alpha \neq 0$). Word contributions and moments of order 2 and 4.

α	w	Indicator set	$p_{H^b_\alpha}(w)$	Moments
$\alpha \in (0,1]$	aa	$X_0 + X_1 \leq \alpha$	$\frac{1}{2}\alpha^2$	$\beta_2(H^b_\alpha) = \frac{1}{2}\alpha$
$\alpha \in [1,2]$	aa		$\alpha - 1 + \frac{\alpha(2-\alpha)}{2}$	$\beta_2(H^b_\alpha) = 1 - \frac{1}{\alpha} + \frac{(2-\alpha)}{2}$
$\alpha \in (0,1]$	$abab$ $abba$ $aabb$	$X_0 + X_1 \leq \alpha, X_1 + X_2 \leq \alpha$	0 $\frac{\alpha^3}{3}$ same as above	$\beta_4(H^b_\alpha) = \frac{2\alpha}{3}$
$\alpha \in [1,2]$	$abab$ $abba$ $aabb$		0 $\alpha - 1 + \frac{1-(\alpha-1)^3}{3}$ same as above	$\beta_4(H^b_\alpha) = 2(\frac{\alpha-1}{\alpha^2} + \frac{1-(\alpha-1)^3}{3\alpha^2})$

Table 4: Type II Hankel limit ($\alpha \neq 0$). Word contributions and moments of order 2 and 4.

α	w	Indicator set	$p_{H^B_\alpha}(w)$	Moments
$\alpha \in (0,1]$	aa	$\lvert X_0 + X_1 - 1\rvert \leq \alpha$	$\alpha(2-\alpha)$	$\beta_2(H^B_\alpha) = 1 - \frac{\alpha}{2}$
$\alpha \in (0,1/2]$	$abab$ $abba$ $aabb$	$\lvert X_0 + X_1 - 1\rvert \leq \alpha,$ $\lvert X_1 + X_2 - 1\rvert \leq \alpha$	0 $\frac{2}{3}\alpha^2(6-5\alpha)$ same as above	$\beta_4(H^B_\alpha) = \frac{1}{3}(6-5\alpha)$
$\alpha \in [1/2,1]$	$abab$ $abba$ $aabb$		0 $\frac{-1+6\alpha-2\alpha^3}{3}$ same as above	$\beta_4(H^B_\alpha) = \frac{(-1+6\alpha-2\alpha^3)}{6\alpha^2}$

Table 5: Type II Toeplitz limit ($\alpha \neq 0$). Word contributions and moments of order 2 and 4.

α	w	Indicator set	$p_{T_\alpha^B}(w)$	Moments
$\alpha \in (0, 1/2]$	aa	$\lvert X_0 - X_1\rvert \leq \alpha$ or $\lvert X_0 - X_1\rvert \geq 1 - \alpha$	2α	$\beta_2(T_\alpha^B) = 2$
$\alpha \in (0, 1/2]$	$abba$	$\lvert X_0 - X_1\rvert \leq \alpha$ or $\lvert X_0 - X_1\rvert \geq 1 - \alpha$, $\lvert X_1 - X_2\rvert \leq \alpha$ or $\lvert X_1 - X_2\rvert \geq 1 - \alpha$	$4\alpha^2$	
	$aabb$		same as above	
	$abab$	$\lvert X_0 - X_1\rvert \leq \alpha$ or $\lvert X_0 - X_1\rvert \geq 1 - \alpha$, $\lvert X_1 - X_2\rvert \leq \alpha$ or $\lvert X_1 - X_2\rvert \geq 1 - \alpha$, $0 \leq X_0 - X_1 + X_2 \leq 1$	$4\alpha^2(1 - \frac{2}{3}\alpha)$	$\beta_4(T_\alpha^B) = 4(3 - \frac{2}{3}\alpha)$