

CONVERGENCE OF A CLASS OF HANKEL TYPE MATRICES

ARUP BOSE*

Theoretical Statistics and Mathematics Unit

Indian Statistical Institute

Kolkata, India

bosearu@gmail.com

SREELA GANGOPADHYAY

Theoretical Statistics and Mathematics Unit

Indian Statistical Institute

Kolkata, India

gangopadhyay7@gmail.com

This version December 24, 2014

Abstract

Let H_n be the $n \times n$ symmetric Hankel-type matrix whose (i, j) th element on the k th anti-diagonal (where $k = 0$ denotes the main anti-diagonal) is defined as : $H_{k,n}(i, j) = g_k \left(\frac{i - \lfloor \frac{n+k+1}{2} \rfloor}{n} \right)$ if $i + j = n + 1 + k$ and 0 otherwise. Under suitable symmetry and summability conditions on $\{g_k\}$, we show that as $n \rightarrow \infty$, the limiting spectral distribution of $\{H_n\}$ exists and is given by $\sum_{k=-\infty}^{\infty} g_k(U) a_k$ where U is uniformly distributed on $[-1/2, 1/2]$ and is tensor independent of the non-commuting variables $\{a_k\}$ which are certain symmetric pair-wise free but not completely free Bernoulli variables.

Key words and phrases. Hankel type matrices, non-commutative probability spaces, joint convergence, free independence, pair-wise free, random matrix, limiting spectral distribution.

AMS 2010 Subject Classification Primary 15B05, Secondary 60B10, 60B20.

1 Introduction

Let $\{A_{k,n}\}$, $k = 1, 2, \dots, K$ be K sequences of $n \times n$ matrices. Then as elements of the non-commutative probability space of $n \times n$ complex matrices with the state as average trace, they are said to converge jointly (as $n \rightarrow \infty$), if for every polynomial $P(A_{k,n}, A_{k,n}^*, k \leq K)$, the average trace converges. Here A^* denotes the complex conjugate of A . The limit non-commutative (polynomial) $*$ -algebra is defined by

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

the non-commutative indeterminates (limit variables) $\{a_k\}$ where the state ϕ satisfies $\phi(P(a_k, a_k^*, k \leq K)) = \lim \frac{1}{n} \text{Tr}(P(A_{k,n}, A_{k,n}^*, k \leq K))$ for all polynomials P . The limit non-commutative joint distribution of $\{a_k\}$ is defined as the collection of all the joint moments $\phi(a_{i_1}^{\epsilon_1} a_{i_2}^{\epsilon_2} \cdots a_{i_n}^{\epsilon_n})$ for all $1 \leq i_1, i_2, \dots, i_n \leq K, n \geq 1$ and $\epsilon_i \in \{1, *\}$.

When we have only one sequence of matrices, say $\{A_n\}$ (which are, for simplicity, real symmetric), then there is a related notion of convergence. Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A_n . Then the *Empirical Spectral Distribution Function* (ESD) of A_n equals

$$F_{A_n}(x) = n^{-1} \sum_{i=1}^n \mathbb{I}\{\lambda_i \leq x\}.$$

As $n \rightarrow \infty$, the *Limiting Spectral Distribution* (LSD) of $\{A_n\}$ is defined as the weak limit F of $\{F_{A_n}\}$, if it exists. We identify F with any random variable X whose distribution is F . This definition extends to non-symmetric matrices with complex entries in the obvious way.

It is easy to construct examples of real symmetric matrices $\{A_n\}$ where the LSD exists but there is no convergence in the non-commutative sense (that is, $\lim \frac{1}{n} \text{Tr}(A_n^k)$ does not exist for some k). On the other hand, by using the moment-trace formula, it is also easy to see that if the real symmetric $\{A_n\}$ converges in the non-commutative sense (that is, $\lim \frac{1}{n} \text{Tr}(A_n^k) = \mu_k$ exists for all positive integers k), and if $\{\mu_k\}$ is the moment sequence of a unique probability distribution F , then the LSD of A_n equals F .

Let U_1 and U_2 be i.i.d. random variables, uniformly distributed on the interval $(0, 1)$. The famous Szegő's theorem implies that if $T_n := ((t_{|i-j|}))_{1 \leq i, j \leq n}$ is the Toeplitz matrix and $\{t_k\}$ is square summable, then the LSD of T_n equals $t_0 + 2 \sum_{k=1}^{\infty} t_k \cos(2\pi k U_2)$. This result was extended to the Toeplitz-type matrix $T_{n,g}$ say, where the elements of the k th upper and lower diagonals, instead of being the constant t_k , are of the form $g_k(i/n)$ in the i th row for some suitable functions g_k (see [4], [2]). The limit in this case equals $g_0(2\pi U_1) + 2 \sum_{k=1}^{\infty} g_k(2\pi U_1) \cos(2\pi k U_2)$.

The related Hankel matrix $H_n = ((h_{i+j}))$ and the corresponding Hankel operator has been extensively treated in the literature. See [3], [7], [8] and [6] for detailed information. Note that the elements on each *anti-diagonal* of H_n are identical. However, while in T_n the constant on the main diagonal does not change with n , the main anti-diagonal in H_n is h_{n+1} . We take a cue from this observation and the matrix $T_{n,g}$, to consider the following class of Hankel-type matrices.

In our convention of labelling the anti-diagonals, $k = 0$ refers to the main anti-diagonal and $k = 1, 2, \dots$ denote the successive anti-diagonals below the main anti-diagonal and similarly the negative integers label the upper anti-diagonals. For each k , first consider the Hankel matrix $D_{k,n}$ whose k th anti-diagonal elements equal one and the rest of the elements are zero. These matrices converge jointly. The non-commutative joint distribution of the limit variables $\{a_k\}$ can be described in terms of the non-commutative moments as

$$\phi(a_{i_1} \dots a_{i_k}) = \begin{cases} \mathbb{I}_{\{i_1+i_3+\dots+i_{2m-1}=i_2+i_4+\dots+i_{2m}\}} & \text{if } k = 2m \text{ for some } m \geq 1 \\ 0 & \text{otherwise.} \end{cases} \quad (1.1)$$

Interestingly, the above $\{a_k\}$ are symmetric Bernoulli and are pair-wise free but not completely free. This is easy to check by using the above description.

Now generalise $D_{k,n}$ as follows. Let $g_k : [-\frac{1}{2}, \frac{1}{2}] \rightarrow \mathbb{R}$, be continuous and symmetric about 0; let $H_{k,n}$ be the $n \times n$ Hankel-type matrix whose (i, j) th element is defined as :

$$H_{k,n}(i, j) = \begin{cases} g_k \left(\frac{i - \lfloor \frac{n+k+1}{2} \rfloor}{n} \right) & \text{if } i + j = n + 1 + k \\ 0 & \text{otherwise} \end{cases} \quad (1.2)$$

Note that unlike the Hankel matrices considered usually in the literature, for example [8], [6], where the main anti-diagonal has the variable h_{n+1} which changes as n changes, in our case the main anti-diagonal has elements of $g_0(\cdot)$ which is a fixed function. So the labelling is different.

We show that $\{H_{k,n}\}$ converge jointly and the limit variables are $\{g_k(U)a_k\}$ where U is uniformly distributed on $[-1/2, 1/2]$ and is tensor independent of $\{a_k\}$. As a consequence, for any $K \geq 1$, $\sum_{|k| \leq K} H_{k,n}$ converges in the (algebraic sense) and the LSD of this real symmetric matrix exists and equals $\sum_{|k| \leq K} g_k(U)a_k$ with distribution \tilde{F}_K say.

Finally, consider the full Hankel-type matrix $H_n = \sum_{|k| \leq n} H_{k,n}$. By imposing suitable restrictions on the functions $\{g_k\}$, H_n is approximated by the finite-diagonal matrix $\sum_{|k| \leq K} H_{k,n}$ in an appropriate metric and this helps us to conclude that the LSD of H_n exists under these conditions on $\{g_k\}$. The limit distribution function is the weak limit of \tilde{F}_K as $K \rightarrow \infty$ and may be formally expressed as $\sum_{k=-\infty}^{\infty} g_k(U)a_k$. There does not seem to be any analytic description of the limit distribution function.

The case when the $\{g_k\}$ are not symmetric, leads to a non-symmetric H_n . Studying the LSD of this matrix is an extremely difficult problem. We have made some elementary remarks on some special cases at the end of the article.

2 Preliminaries

A *non-commutative probability space* is a pair (\mathcal{A}, ϕ) where \mathcal{A} is a unital algebra (with unity $\mathbf{1}$) and $\phi : \mathcal{A} \rightarrow \mathbb{C}$ is a linear functional satisfying $\phi(\mathbf{1}) = 1$. Elements of a non-commutative probability space will also be called (*non-commutative*) *random variables*. If an appropriate $*$ operation is defined on \mathcal{A} , then (\mathcal{A}, ϕ) is called a $*$ -probability space¹. A random variable $a \in \mathcal{A}$ is said to be self-adjoint if $a = a^*$ and unitary if $aa^* = a^*a = \mathbf{1}$. It is called Haar unitary if $\phi(a^k) = \mathbb{I}_{\{k=0\}}$.

For our purposes we need the following $*$ -probability space. Let \mathcal{A}_n be the space of $n \times n$ symmetric random matrices with elements which are real numbers or are random variables with all moments finite. Then ϕ_n equal to $\frac{1}{n} E_\mu[\text{Tr}(\cdot)]$ or $\frac{1}{n} [\text{Tr}(\cdot)]$ both yield a $*$ -probability space.

For $\{b_i\}_{i \in J} \subset \mathcal{A}$, their *joint moments* is the collection $\{\phi(b_{i_1} b_{i_2} \dots b_{i_k}), k \geq 1\}$, where each $b_{i_j} \in \{b_i\}_{i \in J}$.

¹Since we shall, mostly, be dealing with only real symmetric matrices, all our algebras, unless otherwise stated, are real.

Random variables $\{b_{i,n}\}_{i \in J} \in (\mathcal{A}_n, \phi_n)$ are said to *converge in law* to $\{b_i\}_{i \in J} \in (\mathcal{A}, \phi)$ (as $n \rightarrow \infty$) if each joint moment of $\{b_{i,n}\}_{i \in J}$ converges to the corresponding joint moment of $\{b_i\}_{i \in J}$. That is, if for $k \geq 1$,

$$\phi_n[P(b_{i_1,n}, b_{i_2,n}, \dots, b_{i_k,n})] \rightarrow \phi[P(b_{i_1}, b_{i_2}, \dots, b_{i_k})]$$

for all polynomials P . If this happens, we write

$$\{b_{i,n}\}_{i \in J} \xrightarrow{\phi_n} \{b_i\}_{i \in J}.$$

If the random variables $\{b_{i,n}\}_{i \in J}$ are $n \times n$ (non-random) matrices, then the above convergence is assumed to be with respect to $\phi_n = \frac{1}{n} \text{Tr}$. If instead they are random matrices, then the above convergence is in one of the following two senses:

(i) We say that $\{b_{i,n}\}_{i \in J}$ converges to $\{b_i\}_{i \in J}$ if the convergence holds with respect to $\phi_n = \frac{1}{n} \text{E Tr}$.

(ii) We say $\{b_{i,n}\}_{i \in J}$ converges almost surely to $\{b_i\}_{i \in J}$ if the convergence holds with respect to $\phi_n = \frac{1}{n} \text{Tr}$, almost surely.

3 Hankel-type finite-diagonal matrices

Let $\{g_k\}_{-\infty < k < \infty}$ be a two-sided sequence of functions, such that for each k , $g_k : [-\frac{1}{2}, \frac{1}{2}] \rightarrow \mathbb{R}$, g_k is continuous and symmetric about 0. Let $H_{k,n}$ be the $n \times n$ Hankel-type matrix defined in (1.2). When $g_k \equiv 1$, $H_{k,n}$ is the Hankel matrix with all entries 0, except the entries on the k th anti-diagonal which are all assumed to be 1. Note that counted from the main anti-diagonal, k positive (negative) refers to the lower (respectively upper) anti-diagonal. We call this matrix $D_{k,n}$.

To describe the joint limit of $H_{k,n}$, let (\mathcal{A}, ϕ) be a $*$ -probability space, and let $\{a_i\}_{i \in \mathbb{Z}} \subset \mathcal{A}$ be a sequence of self-adjoint and unitary elements such that $\phi(a_{i_1} \dots a_{i_k})$ is as defined in (1.1).

It is then not hard to see that, a_i 's are distributed as symmetric Bernoulli and are pair-wise freely independent but not totally free.

Theorem 3.1. *For any K , $(H_{k,n}, |k| \leq K)$ jointly converge to $(g_k(U)a_k, |k| \leq K)$ where $\{a_k\}$ are elements of a $*$ -probability space (\mathcal{A}, ϕ) , ϕ as defined in (1.1) and U is a random variable, uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$ and independent (in the classical sense) of \mathcal{A} . In particular the LSD of $\sum_{|j| \leq K} H_{j,n}$ equals $\sum_{|j| \leq K} g_j(U)a_j$.*

Before we prove the above theorem, we state and prove a corollary.

Corollary 1. *$(D_{k,n}, |k| \leq K)$ converge jointly to $(a_k, |k| \leq K)$ where a_k are as in (1.1). In particular, for real numbers $\{h_k, |k| \leq K\}$, the LSD of $\sum_{|k| \leq K} h_k D_{k,n}$ equals $\sum_{|k| \leq K} h_k a_k$. For any $s \neq t$, the LSD of $D_{s,n} + D_{t,n}$ is the arc-sine law and $D_{s,n} D_{t,n}$ is asymptotically Haar unitary.*

Proof. The joint convergence follows from Theorem 3.1. By that Theorem, all moments of the ESD converge. Note that these moments determine a distribution uniquely which is as given in the statement of the corollary. Finally it just suffices to observe that for any $s \neq t$, the $*$ -distribution of $a_s + a_t$ is the arc-sine law as it is a free convolution of symmetric Bernoulli (see, [5], pp.200-202) and that $a_s a_t$ is Haar unitary.

Proof of Theorem 3.1. First note that if $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ are the eigenvalues of $D_{k,n}$, then $\forall i = 1, \dots, n$, $\lambda_i \in \{-1, 0, 1\}$ and 0 has algebraic multiplicity $|k|$ and multiplicity of 1 and -1 are equal as $n \rightarrow \infty$. So ESD of $D_{k,n}$ converges to the random variable $a_k = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, i.e., a_k is symmetric Bernoulli.

Let, for any fixed s , $T_{s,n}$ denote the $n \times n$ Toeplitz matrix whose entries are all zero except those on the s th diagonal which equal 1, index s being counted from the main diagonal ($s = 0$) and $s = \pm 1, \dots$ above and below the main diagonal respectively.

If r and s are any two integers, then (for large enough n), the product $D_{r,n} D_{s,n}$ equals $T_{s-r,n}$ except for s many rows and r many columns which are zero. Consequently $D_{k,n}^2$ is an identity matrix except whose k rows and k columns are zero. Thus for asymptotic purposes, we may treat $D_{k,n}^2$ as an identity matrix.

Now consider $T_{s,n}$ and $D_{r,n}$. Then the (i, j) th entry of the product $T_{s,n} D_{r,n}$ equals $= \sum_{j_1} t_{i,j_1} h_{j_1,j}$ which is $\neq 0$ iff $j = n + r + 1 - i - s$. Since there are only finitely many

such possibilities, $\lim_{n \rightarrow \infty} \frac{1}{n} \text{Tr}(T_s D_r) = 0$

Finally note that $T_{r,n} T_{s,n} = T_{r+s,n}$ except for a finitely many entries.

Using the above facts repeatedly, it is easy to see that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{Tr}(H_{k_1,n} H_{k_2,n} \dots H_{k_s,n}) = 0, \text{ if } s = 2m - 1 \text{ for some } m \geq 1.$$

So assume, $s = 2m$. For convenience, we will write n_k for $n + k + 1$ for any integer k .

$$\begin{aligned} & \frac{1}{n} \text{Tr}(H_{k_1,n} H_{k_2,n} \dots H_{k_{2m},n}) \\ &= \frac{1}{n} \sum_{i, j_1, \dots, j_{2m-1}} (H_{k_1,n}(i, j_1) H_{k_2,n}(j_1, j_2) \dots H_{k_{2m},n}(j_{2m-1}, i)) \\ &= \frac{1}{n} \sum_i H_{k_1,n}(i, n_{k_1} - i) \dots H_{k_{2m},n} \left(\sum_{j=1}^{2m-1} (-1)^{j+1} n_{k_j} - i, \sum_{j=1}^{2m} (-1)^j n_{k_j} + i \right) \\ & \quad (\text{to satisfy trace condition the last index must be } i, \text{ i.e., } k_1 + k_3 + \dots = k_2 + k_4 + \dots) \\ &= \frac{1}{n} \sum_i g_{k_1} \left(\frac{i - \lfloor \frac{n_{k_1}}{2} \rfloor}{n} \right) \dots g_{k_{2m}} \left(\frac{i - \lfloor \frac{n_{k_{2m}}}{2} \rfloor}{n} \right) \mathbb{I}_{k_1+k_3+\dots+k_{2m-1}=k_2+k_4+\dots+k_{2m}} \\ & \quad (\text{this is a Riemann sum and using uniform continuity of } g_k \text{'s}) \\ & \rightarrow \left(\int_{-\frac{1}{2}}^{\frac{1}{2}} (g_{k_1} \dots g_{k_{2m}})(x) dx \right) \mathbb{I}_{k_1+k_3+\dots+k_{2m-1}=k_2+k_4+\dots+k_{2m}}. \end{aligned}$$

Thus $\lim_{n \rightarrow \infty} \frac{1}{n} \text{Tr}(H_{k_1,n} H_{k_2,n} \dots H_{k_s,n}) = \mathbb{E}_U \otimes \phi(g_{k_1}(U) a_1, g_{k_2}(U) a_2, \dots, g_{k_s}(U) a_s)$, where \mathbb{E}_U is the usual expectation with respect to Lebesgue measure on $[-\frac{1}{2}, \frac{1}{2}]$ and ϕ is a linear

functional on \mathcal{A} as defined in (1.1) and they act independently (classical sense) on

$$C\left[-\frac{1}{2}, \frac{1}{2}\right] \otimes \mathcal{A} := \{f(U)a : f \text{ continuous real-valued function on } [-\frac{1}{2}, \frac{1}{2}], a \in \mathcal{A}\}.$$

This completes the proof of the theorem.

Remark 3.1. Let us define an $n \times n$ k -diagonal random Hankel matrix $\tilde{H}_{k,n}$ whose (i, j) th entry is $g_k(U)\mathbb{I}_{i+j=n+k+1}$ where U is a random variable uniformly distributed on $I := [-\frac{1}{2}, \frac{1}{2}]$. Suppose $g_k(\cdot)$ are continuous even functions on I . Following arguments similar to that given in the proof of Theorem 3.1, one can show that

(i) For fixed $K > 0$,

$$\left(\tilde{H}_{-K,n}, \tilde{H}_{-K+1,n}, \dots, \tilde{H}_{K-1,n}, \tilde{H}_{K,n}\right) \xrightarrow{\frac{1}{n} \text{ETr}} (g_{-K}(U)a_{-K}, \dots, g_K(U)a_K)$$

where $\{a_i\}_{|i| \leq K} \in (\mathcal{A}, \phi)$ are as defined in (1.1) and U is independent of (\mathcal{A}, ϕ) . As a consequence the expected ESD of $\sum_{i=-K}^K \tilde{H}_{i,n}$ converges weakly to $\sum_{j=-K}^K g_j(U)a_j$.

(ii) For almost every value of U ,

$$\left(\tilde{H}_{-K,n}, \tilde{H}_{-K+1,n}, \dots, \tilde{H}_{K,n}\right) \xrightarrow{\frac{1}{n} \text{Tr}} (g_{-K}(U)a_{-K}, \dots, g_K(U)a_K)$$

and hence for fixed $K > 0$, for almost every given ω , the ESD of $\sum_{i=-K}^K \tilde{H}_{i,n}$ converges weakly to $\sum_{j=-K}^K g_j(U(\omega))a_j$. Note that this is a random limit depending on ω (a typical point in the probability space where U is defined).

4 When all diagonals are present

Now, for U as previously defined, let $(C(U), \mathbb{E}_U)$ be a classical probability space where $C(U) := \{f(U) : f \text{ continuous on } I\}$ and \mathbb{E}_U is the usual expectation on I with respect to Lebesgue measure. Then consider the non-commutative probability space $(\tilde{\mathcal{A}}, \tilde{\phi})$ where $\tilde{\mathcal{A}}$ is the algebra generated by $\{f(U)a : f(U) \in C(U), a \in \mathcal{A}\}$ and $\tilde{\phi}$ acts on $f(U)a \in \tilde{\mathcal{A}}$ as $\tilde{\phi}(f(U)a) = \left(\int_{-\frac{1}{2}}^{\frac{1}{2}} f(x)dx\right)\phi(a)$ which is extended linearly on $\tilde{\mathcal{A}}$. $(\tilde{\mathcal{A}}, \tilde{\phi})$ is a $*$ -probability space where $(f^{\hat{}}(U)a)^* = f(U)a^*$.

Let

$$\sum_{j=-k}^k g_j(U)a_j =: b_k \in \tilde{\mathcal{A}}.$$

We have seen that $\sum_{j=-k}^k H_{j,n}$ converges to b_k which is self-adjoint. It is also easy to see that $\{\tilde{\phi}(b_k^m)\}_{m \geq 1}$ defines a unique distribution function \tilde{F}_k (say) which is the LSD of $\sum_{j=-k}^k H_{j,n}$.

To deal with matrices which may have all diagonals non-zero, we need some additional conditions on $\{g_j\}$ and an appropriate metric which will allow such matrices to be approximated by Hankel-type matrices with finitely many non-zero anti-diagonals.

The Mallow's metric is defined on the space of all probability distributions with finite second moment. Let F and G be two distribution functions with finite second moment. Then the Mallow's distance between F and G is defined as

$$d_M^2(F, G) := \inf_{X \sim F, Y \sim G} \mathbb{E} |X - Y|^2. \quad (4.1)$$

It is known that $d_M(F_n, F) \rightarrow 0$ if and only if $\int x^2 dF_n(x) \rightarrow \int x^2 dF(x)$ and F_n converges to F weakly.

We need the following upper bound of this metric between the ESD of two matrices: let A, B be two $n \times n$ real symmetric matrices with eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ and $\beta_1 \leq \beta_2 \leq \dots \leq \beta_n$, respectively. Then

$$d_M^2(F_A, F_B) \leq \frac{1}{n} \sum_{j=1}^n (\lambda_j - \beta_j)^2 \leq \frac{1}{n} \text{Tr}(A - B)^2. \quad (4.2)$$

The first inequality is obvious and the last inequality above is a standard result in matrix algebra; one can see a proof of this in Lemma 2.3 of [1].

Theorem 4.1. *Suppose $\{g_j\}$ are continuous even functions on $I := [-\frac{1}{2}, \frac{1}{2}]$. Suppose*

- (i) $\sum_{k=-\infty}^{\infty} \int_{-\frac{1}{2}}^{\frac{1}{2}} g_j^2(x) dx < \infty$ and
- (ii) $\sum_{j=-n+1}^{n-1} \sum_{i=1}^n \int_{\frac{i-1-\lfloor \frac{n_j}{2} \rfloor}{n}}^{\frac{i-\lfloor \frac{n_j}{2} \rfloor}{n}} \left(g_j^2 \left(\frac{i - \lfloor \frac{n_j}{2} \rfloor}{n} \right) - g_j^2(x) \right) dx \rightarrow 0$ as $n \rightarrow \infty$, where $n_j := n + j + 1$.

Then LSD of $H_n := \sum_{j:|j| \leq n} H_{j,n}$ exists and equals $\lim_{k \rightarrow \infty} \tilde{F}_k =: \tilde{F}_\infty$. The limiting random variable may be written as $\sum_{-\infty}^{\infty} g_j(U) a_j$ where U and a_j are as defined in Theorem 3.1 and has distribution function \tilde{F}_∞ .

Proof. Let F_n and $F_{k,n}$ denote respectively the ESD of H_n and $\sum_{j=-k}^k H_{j,n}$. First we will show that $\{\tilde{F}_k\}$ is weakly convergent. For that, it is enough to show that $\{\tilde{F}_k\}$ is Cauchy in d_M . Let $n > k_2 > k_1$. Then

$$d_M(\tilde{F}_{k_1}, \tilde{F}_{k_2}) \leq d_M(\tilde{F}_{k_1}, F_{k_1,n}) + d_M(\tilde{F}_{k_2}, F_{k_2,n}) + d_M(F_{k_1,n}, F_{k_2,n}) = d_1 + d_2 + d_3 \text{ (say)}.$$

But by Theorem 3.1, $d_1 + d_2 \rightarrow 0$ as $n \rightarrow \infty$. For d_3 , observe that

$$d_3^2 := d_M^2(F_{k_1,n}, F_{k_2,n}) \leq \frac{1}{n} \text{Tr} \left(\sum_{k_1 < |j| \leq k_2} H_{j,n} \right)^2 \text{ (by (4.2))}$$

$$\begin{aligned}
&\leq \frac{1}{n} \sum_{j:k_1 < |j| \leq k_2} \sum_i g_j^2 \left(\frac{i - \lfloor \frac{n_j}{2} \rfloor}{n} \right) \\
&\leq \sum_{j:k_1 < |j| \leq k_2} \left[\frac{1}{n} \sum_{i=1}^n g_j^2 \left(\frac{i - \lfloor \frac{n_j}{2} \rfloor}{n} \right) - \int_{-\frac{1}{2}}^{\frac{1}{2}} g_j^2(x) dx \right] + \sum_{j:k_1 < |j| \leq k_2} \int_{-\frac{1}{2}}^{\frac{1}{2}} g_j^2(x) dx, \\
&\rightarrow 0 \text{ as } k_1, k_2 \rightarrow \infty \text{ (by Conditions (i) and (ii)).}
\end{aligned}$$

This implies that \tilde{F}_k converges weakly to a distribution function \tilde{F}_∞ (say). Now, to prove the Theorem, consider

$$d_M(F_n, \tilde{F}_\infty) \leq d_M(F_n, F_{k,n}) + d_M(F_{k,n}, \tilde{F}_k) + d_M(\tilde{F}_k, \tilde{F}_\infty).$$

Since \tilde{F}_k converges weakly to \tilde{F}_∞ , for a fixed $\varepsilon > 0$, there exists a $K \in \mathbb{N}$ such that

$$d_M(\tilde{F}_k, \tilde{F}_\infty) \leq \varepsilon \text{ for all } k \geq K.$$

Now for any fixed $k \geq K$, by Theorem 3.1 $d_M(F_{k,n}, \tilde{F}_k) \rightarrow 0$ as $n \rightarrow \infty$. Finally, again using (4.2) we have

$$\begin{aligned}
d_M^2(F_n, F_{k,n}) &\leq \frac{1}{n} \text{Tr} \left(H_n - \sum_{j=-k}^k H_{j,n} \right)^2 \\
&\leq \frac{1}{n} \sum_{j:|j|>k} \sum_i g_j^2 \left(\frac{i - \lfloor \frac{n_j}{2} \rfloor}{n} \right) \\
&\leq \sum_{j:|j|>k} \left[\frac{1}{n} \sum_{i=1}^n g_j^2 \left(\frac{i - \lfloor \frac{n_j}{2} \rfloor}{n} \right) - \int_{-\frac{1}{2}}^{\frac{1}{2}} g_j^2(x) dx \right] + \sum_{j:|j|>k} \int_{-\frac{1}{2}}^{\frac{1}{2}} g_j^2(x) dx,
\end{aligned}$$

and due to Conditions (i), (ii), right side goes to zero as $n \rightarrow \infty$. Hence $d_M(F_n, \tilde{F}_\infty) \rightarrow 0$ as $n \rightarrow \infty$. This completes the proof of the Theorem.

Remark 4.1. *The study of the LSD of $H_{k,n}$ and H_n , when the symmetry assumption is removed, does not seem to be easy. This requires further investigation. The following simple observations though, may be made. Let $N_{k,n} = (a_{ij})_{1 \leq i, j \leq n}$, be the $n \times n$ non-symmetric Hankel type matrix,*

$$a_{i,j} = \begin{cases} 1 & \text{if } i + j = n + k + 1 \text{ and } i \leq \lfloor (n + k)/2 \rfloor; \\ 0 & \text{otherwise} \end{cases}$$

Then clearly the LSD of both $N_{k,n}$ and $N_{k,n}^$ are the point mass at zero, δ_0 . The LSD of the symmetric matrix $N_{k,n} N_{k,n}^*$ converges in distribution to the Bernoulli random variable $(1/2)\delta_0 + (1/2)\delta_1$. Since $N_{k,n} + N_{k,n} r^* = D_{k,n}$, its LSD is the symmetric Bernoulli $(1/2)\delta_{-1} + (1/2)\delta_1$. The limiting moment of any monomial in $(N_{k,n}, N_{k,n}^*)$ is zero unless $N_{k,n}$ and $N_{k,n}^*$ appear alternately in the monomial. However, the limiting joint free cumulants of $(N_{k,n}, N_{k,n}^*)$ may be non-zero even if they do not appear alternately. For example, it can be checked that*

$$\lim \kappa_4(N_{k,n}, N_{k,n}, N_{k,n}^*, N_{k,n}^*) = 1/4.$$

References

- [1] Z. D. Bai (1999). Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica*, 9(3):611–677, 1999. ISSN 1017-0405. With comments by G. J. Rodgers and Jack W. Silverstein; and a rejoinder by the author.
- [2] A. Bose, S. Gangopadhyay, and K. Saha (2013). Convergence of a class of Toeplitz type matrices. *Random Matrices Theory Appl.*, Vol. 2 (3):21 pp.
- [3] A. Böttcher, and B. Silbermann (1997). *Introduction to large truncated Toeplitz matrices*. Springer-Verlag, New York.
- [4] M. Kac, W. L. Murdock, and G. Szegö (1953). On the eigenvalues of certain Hermitian forms. *J. Rational Mech. Anal.*, 2:767–800.
- [5] A. Nica and R. Speicher (2006). *Lectures on the combinatorics of free probability*. London Mathematical Society Lecture Notes Series: 353 Cambridge University Press.
- [6] Vladimir V. Peller (1998). An excursion into the theory of Hankel operators. *Holomorphic Spaces, MSRI Publications*, 33:66-120.
- [7] U. Grenander and G. Szegö (1958). *Toeplitz forms and their applications* University of California Press, Berkeley and Los Angeles.
- [8] H. Widom (1966). Hankel matrices. *Trans. American Math. Soc.*, 121:1-35.