EINDHOVEN UNIVERSITY OF TECHNOLOGY

Finite diagonal random matrices

Citation for published version (APA):
Bose, A., & Sen, S. (2013). Finite diagonal random matrices. *Journal of Theoretical Probability*, *26*(3), 819-835. https://doi.org/10.1007/s10959-011-0378-z

DOI:
10.1007/s10959-011-0378-z

Document status and date:
Published: 01/01/2013

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 18. Mar. 2025

J Theor Probab (2013) 26:819–835
DOI 10.1007/s10959-011-0378-z

Finite Diagonal Random Matrices

Arup Bose · Sanchayan Sen

Received: 30 April 2011 / Published online: 10 September 2011
© Springer Science+Business Media, LLC 2011

Abstract The goal of this article is to extend some results of Popescu (Probab. Theory Relat. Fields 144:179, 2009) in several directions. We establish the limiting spectral distribution (LSD) for r-diagonal matrices under reduced moment conditions compared to those required by Popescu. We also deal with the joint convergence of several sequences of such matrices. In particular, we show that there is a large class of such matrices where the joint limit is not free while the marginals are semicircular. We also consider matrices of the form $X_n X_n^T$ where X_n is a sequence of nonsymmetric r-diagonal random matrices and establish their limiting spectral distribution.

Keywords Tridiagonal and finite diagonal matrices · Sample covariance type matrices · Limiting spectral distribution · Semicircle law · Free independence

Mathematics Subject Classification (2010) Primary 60B20 · Secondary 60B10 · 46L53 · 46L54

1 Introduction

Suppose $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of a real symmetric matrix $M_{n\times n}$. Then its *Empirical Spectral Distribution (ESD)* is defined as

$$F^{M_n}(x, y) = n^{-1} \sum_{i=1}^{n} \mathbb{I}(\lambda_i \leq x). \tag{1.1}$$

A. Bose (✉)
Statistics and Mathematics Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India
e-mail: bosearu@gmail.com

S. Sen
Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, NY 10012, USA
e-mail: sen@cims.nyu.edu

The *Limiting Spectral Distribution (or measure) (LSD)* of the sequence is defined as the weak limit of the sequence $\{F^{M_n}\}$, if it exists, either almost surely or in probability.

LSD of various random matrices have been studied in the literature. See for example Bose, Hazra and Saha [5] for a list of common symmetric patterned matrices and their LSD. More examples of LSD may be found in Bai [2].

Tridiagonal matrix models are heavily used in numerical recipes. Popescu [7] discussed an interesting class of symmetric tridiagonal matrix models under some nonstandard conditions on the entries. He proved the convergence of the ESD and computed the LSD in some particular cases. He also obtained the limit of the trace of monomials formed by several such independent matrices. Extensions were also provided for the r-diagonal model where the kth superdiagonals and subdiagonals are zero for $k > r$.

The goal of this article is to extend some of Popescu's results in several directions. First, we relax some of the moment conditions that are required by Popescu. We also deal with the joint convergence of sequences of such matrices. In particular, we show that the joint limit need not be free while the marginals may still be semicircular. Matrices of the form $X_n X_n^T$ are generalizations of the sample variance covariance matrix. We also consider such matrices where X_n is a sequence of nonsymmetric r-diagonal random matrices and establish their LSD.

2 Main Results

2.1 Popescu's Result and an Extension

We first restate Popescu's main result in a convenient manner and then state an extension. Let $\{b_{ij}\}$ be a sequence of random variables. Further conditions on this sequence shall be imposed as needed. Consider for any fixed r, the sequence of r-diagonal symmetric random matrices $\{A_n\}$ given by

$$(A_n)_{ij} = \begin{cases} b_{j-i,i}, & \text{if } i \le j \le r+i \\ 0, & \text{if } r+i < j \le n \\ (A_n)_{ji}, & \text{if } j < i \end{cases} \tag{2.1}$$

and let, for some $\alpha > 0$,

$$X_n = \frac{A_n}{n^\alpha}. \tag{2.2}$$

Let

$$\Gamma_k = \bigl\{\gamma = (j_0, \dots, j_k) \mid j_0 = j_k,\ \max j_i = 0,\ |j_{i+1} - j_i| \le r\bigr\}$$

$$l_{i,s}(\gamma) = \#\{u : 0 \le u \le k-1 \text{ such that either } j_u = i,\ j_{u+1} = s \text{ or } j_u = s,\ j_{u+1} = i\}.$$

We often write $\gamma = (\gamma_0, \dots \gamma_k)$. Each element $\gamma \in \Gamma_k$ is a cyclical path and the definition is motivated by the restriction that the indices must satisfy while writing the trace of a matrix.

The definition of $l_{i,s}(\gamma)$ is motivated by the matching that must take place to obtain possible nonzero contribution from a given summand in the (expected) trace formula.

Theorem 1 ([7]) *Suppose* $\{b_{i,j}\}$, $j \geq 0$, $r \geq i \geq 0$ *are independent and for some* $\alpha > 0$,

$$\lim_{n\to\infty} \mathbf{E}\left(\frac{b_{j,n}}{n^\alpha}\right)^k = m_{jk}, \quad \forall k = 1, 2, \ldots.$$

(a) *Then as* $n \to \infty$, *for all* $k = 1, 2, \ldots,$

$$\mathbf{E}\frac{\mathrm{Tr}(X_n^k)}{n} \to L_k \quad and \quad \frac{\mathrm{Tr}(X_n^k)}{n} \to L_k, \ almost\ surely,$$

where

$$L_k = \frac{1}{\alpha k + 1} \sum_{\gamma \in \Gamma_k} \prod_{0 \leq i \leq r,\ j \leq 0} m_{i, l_{j,j+i}(\gamma)}. \tag{2.3}$$

(b) *Further if for every* $0 \leq j \leq r$, $\{m_{jk}\}_{k\geq 0}$ *are the moments of a compactly supported measure* μ_j, *then there exists a unique measure* μ, *which is compactly supported and has moments* $\{L_k\}_{k\geq 0}$. *Hence* $F^{X_n} \longrightarrow \mu$ *almost surely.*

(c) *If* $m_{jk} = a_j^k$, $\forall k \geq 0$, $0 \leq j \leq r$, *then* μ *is the distribution of* $U_1^\alpha[a_0 + 2\sum_{k=1}^r a_k \cos(2\pi k U_2)]$ *where* U_1 *and* U_2 *are i.i.d. uniformly distributed on* $[0, 1]$.

Note that in the above result, convergence of the ESD in (b) is guaranteed if $\{L_k\}$ determines a distribution uniquely. The compact support assumption on $\{\mu_j\}$ achieves this. To state our extension of this result, let $\{a_{in}\}$, $0 \leq i \leq r$, $n \geq 1$ be a sequence of real numbers and let $\{x_{in}\}$, $0 \leq i \leq r$, $n \geq 1$ be a sequence of independent random variables.

Theorem 2 *Let* A_n *and* X_n *be as in* (2.1) *and* (2.2) *with* $b_{i,n} = a_{in} x_{in}$. *Suppose for some* $\alpha > 0$, *and for* $0 \leq i \leq r$, *as* $n \to \infty$,

$$\frac{a_{in}}{n^\alpha} \longrightarrow a_i. \tag{2.4}$$

Further suppose that either of the following holds: (a1) *for each* $0 \leq i \leq r$, $\{x_{in}\}_{n\geq 1}$ *is i.i.d. with finite second moment or* (a2) *there exists a sequence of positive numbers* $\{c_n\}$ *increasing to infinity such that*

(i)

$$\sum_{j=1}^{\infty} \frac{\mathbb{E}(x_{ij}^4)}{j^2} < \infty \quad for\ 0 \leq i \leq r,$$

(ii)

$$\lim_n \mathbb{E}\big(x_{in}\mathbf{1}\big(|x_{in}| < c_l\big)\big)^k$$

exists for each $k, l \geq 1$ *and* $0 \leq i \leq r$,

(iii)

$$\lim_{l\to\infty}\limsup_{n\to\infty}\frac{\sum_{j=1}^{n}\mathbb{E}(x_{ij}^2\mathbf{1}(|x_{ij}|\geq c_l))}{n}=0\quad for\ 0\leq i\leq r.$$

(a) *Then almost surely,* X_n *has a nonrandom limiting spectral distribution, say* μ.
(b) *Suppose* $\mathbb{E}(|x_{in}^k|)<\infty$, $\forall i,n,k$ *and*

$$\lim_n \mathbb{E}\big(n^{-\alpha}b_{i,n}\big)^k=a_i^k\lim_n\mathbb{E}(x_{in})^k=m_{ik}.$$

Then L_k *in* (2.3) *are the moments of* μ. *Moreover, if for some* $0\leq i_0\leq r$, $\{m_{i_0k}\}_{k\geq1}$ *are moments of some distribution having noncompact support and either* m_{ik} *are nonnegative* $\forall i,k$ *or* $m_{ik}=0$ *whenever* $i\neq i_0$ *(i.e. all the diagonals except the* i_0*th diagonal are "switched off") and* $k\geq 1$, *then* μ *also has noncompact support.*

Remark 1

(a) Condition (ii) in part (b) holds whenever there exists random variables x_i for $0\leq i\leq r$ such that $x_{in}\xrightarrow{w}x_i$ as $n\to\infty$.
(b) In Theorem 2 we have managed to remove the compactness assumption of Theorem 1 in exchange for some other conditions. As a consequence, the LSD can now have noncompact support. For instance, if the limiting moments $\{m_{jk}\}$ are nonnegative and for at least one j_0, the moment sequence $\{m_{j_0k}\}$ corresponds to a measure μ_{j_0} with noncompact support then the LSD also has noncompact support.

2.2 r-Diagonal Covariance Type Matrix

Suppose that $\{A_n\}$ is a sequence of nonsymmetric square matrices and consider the symmetric matrices $\{A_nA_n^T\}$. If A_n is the i.i.d. matrix (all entries are i.i.d.), then $n^{-1}A_nA_n^T$ is the well studied sample variance covariance matrix; it is known that under appropriate conditions $F^{n^{-1}A_nA_n^T}$ converges to the Marčenko–Pastur law. Moreover the LSD of the symmetric version of $n^{-1/2}A_n$ is the semicircular law and distribution of the square of a semicircle random variable is the Marčenko–Pastur law. As observed in Bose, Gangopadhyay and Sen [4], this phenomenon of squaring may or may not hold for general patterned matrices.

Now let A_n be the nonsymmetric square r-diagonal matrix defined as

$$(A_n)_{ij}=\begin{cases}b_{j-i,i}, & \text{if } |j-i|\leq r\\ 0, & \text{otherwise.}\end{cases}\tag{2.5}$$

Let

$$S_n=\frac{A_nA_n^T}{n^{2\alpha}}.\tag{2.6}$$

Then we show that under suitable conditions, F^{S_n} converges almost surely (see Theorem 3). The moments of the LSD are equal to the moments of a trigonometric polynomial of a uniform random variable. Further, under suitable conditions, the squaring relation mentioned above also holds. See Corollary 3.1.

We need the following notation to state our next theorem. Define for $\gamma = (j_0, \dots, j_k)$,

$$l_{i\to s,R}(\gamma) = \#\{u : 1 \le u \le k,\ u \text{ is odd } j_{u-1} = i,\ j_u = s\}, \quad \text{and similarly}$$
$$l_{i\to s,B}(\gamma) = \#\{u : 1 \le u \le k,\ u \text{ is even } j_{u-1} = i,\ j_u = s\}.$$

So we visualize this as follows: every path γ is being colored using two colors R and B. Any segment (j_{u-1}, j_u) is colored R if u is odd and is colored B if u is even.

Theorem 3

(a) *Let r be a positive integer, $\alpha > 0$ and let $\{b_{i,j}\}$ be independent such that for some $\alpha > 0$,*

$$\lim_{n\to\infty} \mathbf{E}\left(\frac{b_{j,n}}{n^\alpha}\right)^k = m'_{jk}, \quad \forall k \ge 0,\ |j| \le r.$$

Then as $n \to \infty$,

$$\mathbf{E}\frac{\operatorname{Tr}(S_n^k)}{n} \to L'_k \quad \textit{and} \quad \frac{\operatorname{Tr}(S_n^k)}{n} \to L'_k, \textit{ almost surely},$$

where

$$L'_k = \frac{1}{2\alpha k + 1} \sum_{\gamma\in\Gamma_{2k}} \prod_{|h|\le r,\, i\le 0} m'_{h, l_{i\to i+h,R}(\gamma)+l_{i+h\to i,B}(\gamma)}. \tag{2.7}$$

(b) *If $\{m'_{jk}\}_{k\ge0}$ are the moments of compactly supported measures μ'_j, then there exists a compactly supported measure μ' whose moments are $\{L'_k\}_{k\ge0}$. Hence $F^{S_n} \to \mu'$ almost surely.*

(c) *If $m'_{jk} = a_j^k$, $\forall k \ge 0$, $|j| \le r$, then μ' is the distribution of*

$$U_1^{2\alpha}\left[c_0 + 2\sum_{k=1}^r c_k \cos(2\pi U_2)\right]$$

where U_1 and U_2 are i.i.d. uniformly distributed on $[0, 1]$ and

$$c_h = 2\sum_{i=-r}^{r-h} a_i a_{i+h} \quad \textit{if } 1 \le h \le 2r \textit{ and } c_0 = \sum_{|i|\le r} a_i^2.$$

From this the following corollary follows.

Corollary 3.1 *If $m'_{jk} = a_{|j|}^k$, $\forall k \ge 0$, $|j| \le r$, then $L'_k = L_{2k}$, $\forall k \ge 0$, where $\{L_k\}$ are as in Theorem 1 obtained from $m_{jk} = a_j^k$, $0 \le j \le r$, $k \le 0$ and hence if the random* *variable Z is distributed as μ (given in Theorem 1) then Z^2 is distributed as μ' (given in Theorem 3).*

Remark 2

(a) It is clear from the expressions for the moments that in general we cannot say that if Z is distributed as μ then Z^2 is distributed as μ'. Corollary 3.1 provides a sufficient condition for this relation to hold.
(b) In Theorem 1 (resp. Theorem 3) we can set $m_{jk} = 0$, $\forall k \geq 0$ (resp. $m'_{jk} = 0$, $\forall k \geq 0$) to "switch off" the jth diagonal and we can get different patterns in the matrix models. The corresponding LSDs will be given by the moments $\{L_k\}$ (resp. $\{L'_k\}$) by plugging in zeroes in (2.3) (resp. (2.7)) in place of m_{jk} (resp. m'_{jk}) in the paths $\gamma \in \Gamma_k$ (resp. Γ_{2k}).

We can relax the assumption of existence of all moments as we did in Theorem 2. This is done in the following result.

Corollary 3.2 *Let A_n and S_n be as in (2.5) and (2.6) with $b_{i,n} = a_{in}x_{in}$ (x_{in}'s are independent). Suppose for some $\alpha > 0$, (2.4) holds for $|i| \leq r$, and either (a1) of Theorem 2 holds for $|i| \leq r$ or conditions* (i), (ii) *and* (iii) *of* (a2) *of Theorem 2 hold for $|i| \leq r$.*

(a) *Almost surely, X_n has a nonrandom limiting spectral distribution, say μ'.*
(b) *Suppose $\mathbb{E}(|x_{in}^k|) < \infty$, $\forall i, n, k$ and $\lim_n \mathbb{E}(n^{-\alpha} b_{i,n})^k = a_i^k \lim_n \mathbb{E}(x_{in})^k = m'_{ik}$. Then L'_k in (2.7) are the moments of μ'. Moreover, if for some $0 \leq i_0 \leq r$, $\{m'_{i_0 k}\}_{k\geq 1}$ are moments of some distribution having noncompact support and either m'_{ik} are nonnegative $\forall i, k$ or $m'_{ik} = 0$ whenever $i \neq i_0$ and $k \geq 1$, then μ' also has noncompact support.*

2.3 Joint Convergence

Suppose now that we have J independent sequences of r-diagonal random matrices. The well known appropriate notion of joint convergence now is the convergence of the trace of all monomials formed from this sequence. We now describe this notion precisely.

A *noncommutative probability space* is a pair $(\mathcal{A}, \phi)$, where $\mathcal{A}$ is an algebra over $\mathbb{C}$ having a unit and $\phi : \mathcal{A} \to \mathbb{C}$ is a linear functional such that $\phi(1) = 1$. Elements of the algebra $\mathcal{A}$ are called noncommutative random variables.

Let $\mathcal{A}_n$ be the space of $n \times n$ complex random matrices with entries being random variables defined on a fixed probability space and having all moments finite. We can define two functionals on $\mathcal{A}_n$ as

$$\phi_{1n}(a) = \frac{1}{n}\operatorname{Tr}(a) \quad \text{and} \quad \phi_{2n}(a) = \frac{1}{n}\mathbb{E}\big(\operatorname{Tr}(a)\big).$$

Then $(\mathcal{A}_n, \phi_{in})$, $i = 1, 2$ are noncommutative probability spaces.

Let $\{a_i\}_{i\in I}$ be a set of elements in $(\mathcal{A}, \phi)$. Denote by $\mathbb{C}[\{X_i\}_{i\in I}]$ the polynomial algebra in the noncommutative indeterminates $\{X_i\}_{i\in I}$. The *joint distribution* of $\{a_i\}_{i\in I}$

is given by the linear functional

$$\mu\colon\ \mathbb{C}\big[\{X_i\}_{i\in I}\big]\to\mathbb{C}$$
$$p\mapsto\phi\big(p\{a_i\}\big).$$

Let $(\mathcal{A}_n,\phi_n)_{n\geq 1}$ and $(\mathcal{A},\phi)$ be noncommutative probability spaces and let $\{a_{i,n}\}_{i\in I}\in\mathcal{A}_n$ for each n, $\{a_i\}_{i\in I}\in\mathcal{A}$. Then $\{a_{i,n}\}_{i\in I}$ *converges in law (or in distribution)* to $\{a_i\}_{i\in I}$ if $\forall p\in\mathbb{C}[\{X_i\}_{i\in I}]$,

$$\lim_n \mu_{\{a_{i,n}\}_{i\in I}}(p)=\mu_{\{a_i\}_{i\in I}}(p).$$

For our purposes the algebra $\mathcal{A}_n$ will be as described above with the two functionals ϕ_{in}, $i=1,2$. If there is convergence we say that the matrices converge almost surely and in expectation, respectively.

To state our result on joint convergence, we first generalize the notation used in the proof of Theorem 3. As before, our generic notation for a path $(j_0,\dots,j_l)$ is γ. For a sequence of "colors" $c=\{c_1,\dots,c_l\}$ (some of the c_i's maybe repeated), let

$$\Gamma_l^c=\big\{\gamma\in\Gamma_l\,|\ \forall u \text{ the segment } (j_{u-1},j_u) \text{ is colored by } c_u\big\}.$$

We assign a unique integer t to each of the colors in c, $1\leq t\leq\#\{c-1,\dots,c_l\}$. For $\gamma\in\Gamma_l^c$, let

$$\begin{aligned}
l_{j\to j+s;R,t}(\gamma)=\#\big\{u|(j_{u-1},j_u)=(j,j+s),&\\
u \text{ is odd and } (j_{u-1},j_u) \text{ is a } t \text{ colored segment}\big\},&\\
l_{j\to j+s;B,t}(\gamma)=\#\big\{u|(j_{u-1},j_u)=(j,j+s),&\\
u \text{ is even and } (j_{u-1},j_u) \text{ is a } t \text{ colored segment}\big\},&\\
l_{j,j+s;t}(\gamma)=\#\big\{u|\text{ either } (j_u,j_{u+1})=(j,j+s) \text{ or}&\\
(j_u,j_{u+1})=(j+s,j) \text{ and } (j_u,j_{u+1}) \text{ is a } t \text{ colored segment}\big\}.&
\end{aligned}$$

Let $\{b_{t,s}^{(i)}\}$ be random variables and define the following matrix sequences:

$$\begin{aligned}
\big(A^{(n,i)}\big)_{st}&=\begin{cases} b_{t-s,s}^{(i)}, & \text{if } s\leq t\leq r+s\\ 0, & \text{if } r+s<t\leq n\\ (A^{(n,i)})_{ts}, & \text{if } t<s,\end{cases}\\
X^{(n,i)}&=\frac{A^{(n,i)}}{n^{\alpha_i}},\\
\big(\overline{A}^{(n,i)}\big)_{st}&=\begin{cases} b_{t-s,s}^{(i)}, & \text{if } |s-t|\leq r\\ 0, & \text{otherwise,}\end{cases}\\
S^{(n,i)}&=\frac{\overline{A}^{(n,i)}\overline{A}^{(n,i)T}}{n^{2\alpha_i}}.
\end{aligned}$$

Note that $A^{(n,i)}$ are symmetric matrices while $\overline{A}^{(n,i)}$ are not symmetric.

Theorem 4 *Let r, m be positive integers, $\alpha_i > 0$ for $1 \le i \le m$ and let $\{b_{i,j}^{(s)}\}$, be independent such that, for some $\{\alpha_i\}$,*

$$\lim_{n\to\infty} \mathbf{E}\left(\frac{b_{j,n}^{(i)}}{n^{\alpha_i}}\right)^k = m_{jk}^{(i)}, \quad \forall k \ge 0,\ -r \le j \le r,\ 1 \le i \le m.$$

(a) *The matrices $\{X^{(n,i)}\}_{1\le i\le m}$ converge almost surely and in expectation to $\{a_i\}_{1\le i\le m}$ whose joint distribution is given by*

$$\phi(a_{i_1}\dots a_{i_l}) = \frac{1}{\alpha_{i_1}+\cdots+\alpha_{i_l}+1}\sum_{\gamma\in\Gamma_l^c}\prod_{\substack{1\le t\le m\\ 0\le s\le r,\ j\le 0}} m^{(t)}_{s,l_{j,j+s;t}(\gamma)}, \tag{2.8}$$

where c is the sequence of colors $(i_1,\dots,i_l)$.

(b) *The matrices $\{S^{(n,i)}\}_{1\le i\le m}$ converge almost surely and in expectation to the joint law of $\{a_i'\}_{1\le i\le m}$ whose joint distribution is given by*

$$\phi(a'_{i_1}\dots a'_{i_l}) = \frac{1}{2(\alpha_{i_1}+\cdots+\alpha_{i_l})+1}\sum_{\gamma\in\Gamma_{2l}^c}\prod_{\substack{1\le t\le m\\ |s|\le r,\ j\le 0}} m^{(t)}_{s,l_{j\to j+s;R,t}(\gamma)+l_{j+s\to j;B,t}(\gamma)}, \tag{2.9}$$

where c is the sequence of colors $(i_1,i_1,i_2,i_2,\dots,i_l,i_l)$.

Remark 3 Suppose $(\mathcal{A},\phi)$ is a noncommutative probability space. A family $\{\mathcal{A}_i\}_{i\in I}$ of unital subalgebras of $\mathcal{A}$ are called freely independent (or simple free) if $\forall n$ positive, indices $k_1,\dots,k_n \in I$ with $k_j \ne k_{j+1}$ and $a_j \in \mathcal{A}_{k_j}$, $1 \le j \le n$ with $\phi(a_j)=0$, we have $\phi(a_1\dots a_n)=0$. For positive integers m, $(m_k)_{1\le k\le m}$, the sets $\{a_{1,p},\dots,a_{m_p,p}\}_{1\le p\le m}$ are freely independent if the algebras they generate are freely independent. In the setup of Theorem 1, when $r=1$, $\alpha=\frac{1}{2}$ and $m_{jk}=\delta_{j1}$ for $j=0,1$; we get $L_{2k}=\frac{1}{k+1}\binom{2k}{k}$, i.e. the LSD μ is the semicircular distribution. Under the same parameter values in part (a) of Theorem 4 we get the joint distribution to be

$$\phi(a_{i_1}\dots a_{i_l}) = \begin{cases} 0, & \text{if } l \text{ is odd}\\ \frac{1}{\frac{l}{2}+1}\binom{l}{\frac{l}{2}}, & \text{if } l \text{ is even}\end{cases}$$

hence $\phi(a_{i_1}\dots a_{i_l}) = \phi(a_1^l)$. This is an example where the marginals converge to the semicircular distribution but the joint distribution is not a free product of semicircular distributions.

3 Proof of Theorems

3.1 Auxiliary Results

In the course of our proofs, we will need to estimate the distance between different spectral measures. This shall be done via the bounded Lipschitz metric d_{BL}, which is a complete metric defined on the space of probability measures on any Polish space (X,d), topologising the weak convergence of probability measures (see [6]):

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

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

We shall also need the following lemma. Its proof may be found in [2] or [1] and uses Lidskii's theorem (see [3, p. 69]).

Lemma 1

(a) *Suppose A, B are $n\times n$ real symmetric matrices. Then*

$$d_{BL}^2\big(F^A,F^B\big)\le\frac{1}{n}\operatorname{Tr}(A-B)^2. \tag{3.1}$$

(b) *Suppose A and B are $p\times n$ real matrices. Let $X=AA^T$ and $Y=BB^T$. Then*

$$d_{BL}^2\big(F^X,F^Y\big)\le\frac{2}{p^2}\operatorname{Tr}(X+Y)\operatorname{Tr}\big[(A-B)(A-B)^T\big]. \tag{3.2}$$

3.2 Proof of Theorem 2

Proof of (a): Define $x_{in}^{(l)}=x_{in}\mathbf{1}(|x_{in}|<c_l)$ and let

$$\big(A_n^{(l)}\big)_{ij}=\begin{cases}a_{j-i,i}x_{j-i,i}^{(l)}, & \text{if } i\le j\le r+i\\ 0, & \text{if } r+i<j\le n\\ (A_n^{(l)})_{ji}, & \text{if } j<i.\end{cases}$$

Using (2.4) (and condition (ii) if we are in case (a2)), we see that in both the cases (a1) and (a2),

$$\lim_n\mathbb{E}\left(\frac{a_{in}x_{in}^{(l)}}{n^\alpha}\right)^k=m_{ik}^{(l)}\quad\text{exists and } m_{ik}^{(l)}\le\big(\max_i\{a_i\}c_l\big)^k.$$

Hence for each l and $0\le i\le r$, $\{m_{ik}^{(l)}\}_{k\ge0}$ is the moment sequence of a compactly supported measure. From Theorem 1, $F^{\frac{A_n^{(l)}}{n^\alpha}}$ converges almost surely to a nonrandom compactly supported probability measure $\mu^{(l)}$.

From (2.4), for some $M>0$, $\frac{a_{ij}^2}{j^{2\alpha}} \le M, \forall i, j$. Hence an application of (3.1) gives

$$\begin{aligned} d_{Bl}^2\big(F^{\frac{A_n}{n^\alpha}}, F^{\frac{A_n^{(l)}}{n^\alpha}}\big) &\le \frac{1}{n^{2\alpha+1}} \operatorname{Tr}\big(A_n - A_n^{(l)}\big)^2 \\ &\le 2\sum_{i=1}^r \sum_{j=1}^n \frac{(a_{ij}x_{ij}\mathbf{1}(x_{ij}\ge c_l))^2}{n^{2\alpha+1}} \\ &\le 2M\sum_{i=1}^r \sum_{j=1}^n \frac{(x_{ij}\mathbf{1}(x_{ij}\ge c_l))^2}{n}. \end{aligned}$$

Under condition (a1)

$$\sum_{i=1}^r \sum_{j=1}^n \frac{(x_{ij}\mathbf{1}(x_{ij}\ge c_l))^2}{n} \xrightarrow{a.s.} \sum_{i=1}^r \mathbb{E}\big(\big(x_{i1}\mathbf{1}(x_{i1}\ge c_l)\big)^2\big)$$

and hence $\lim_l \lim_n d_{Bl}(F^{\frac{A_n}{n^\alpha}}, F^{\frac{A_n^{(l)}}{n^\alpha}}) = 0,\ a.s.$

Under condition (a2),

$$\sum_{j=1}^\infty \frac{\mathbf{Var}(x_{ij}^2\mathbf{1}(x_{ij}\ge c_l))}{j^2} \le \sum_{j=1}^\infty \frac{\mathbb{E}(x_{ij}^4)}{j^2} < \infty.$$

By Kolmogorov SLLN,

$$\sum_{j=1}^n \frac{1}{n}\big(x_{ij}^2\mathbf{1}(x_{ij}\ge c_l) - \mathbb{E}\big(x_{ij}^2\mathbf{1}(x_{ij}\ge c_l)\big)\big) \xrightarrow{a.s.} 0.$$

From (iii) it follows that $\lim_l \limsup_n \sum_{j=1}^n \frac{x_{ij}^2\mathbf{1}(x_{ij}\ge c_l)}{n} \xrightarrow{a.s.} 0$ and hence

$$\lim_l \limsup_n d_{Bl}\big(F^{\frac{A_n}{n^\alpha}}, F^{\frac{A_n^{(l)}}{n^\alpha}}\big) = 0, \quad a.s.$$

Since d_{Bl} metrizes weak convergence of probability measures on $\mathbb{R}$,

$$\lim_n d_{BL}\big(F^{\frac{A_n^{(l)}}{n^\alpha}}, \mu^{(l)}\big) = 0, \quad a.s.\ \forall l,$$

i.e. $\{F^{\frac{A_n^{(l)}}{n^\alpha}}\}_n$ is almost surely Cauchy. Hence

$$\begin{aligned} &d_{BL}\big(F^{\frac{A_n}{n^\alpha}}, F^{\frac{A_m}{m^\alpha}}\big) \\ &\quad\le d_{BL}\big(F^{\frac{A_n}{n^\alpha}}, F^{\frac{A_n^{(l)}}{n^\alpha}}\big) + d_{BL}\big(F^{\frac{A_n^{(l)}}{n^\alpha}}, F^{\frac{A_m^{(l)}}{m^\alpha}}\big) + d_{BL}\big(F^{\frac{A_m^{(l)}}{m^\alpha}}, F^{\frac{A_m}{m^\alpha}}\big) \end{aligned}$$

$$\Rightarrow \limsup_{m,n} d_{BL}\big(F^{\frac{A_n}{n^\alpha}}, F^{\frac{A_m}{m^\alpha}}\big) \le 2 \limsup_{n} d_{BL}\big(F^{\frac{A_n}{n^\alpha}}, F^{\frac{A_n^{(l)}}{n^\alpha}}\big), \quad \text{for each } l$$

$$\Rightarrow \limsup_{m,n} d_{BL}\big(F^{\frac{A_n}{n^\alpha}}, F^{\frac{A_m}{m^\alpha}}\big) \le 2 \lim_{l} \limsup_{n} d_{BL}\big(F^{\frac{A_n}{n^\alpha}}, F^{\frac{A_n^{(l)}}{n^\alpha}}\big) = 0,$$

i.e. $\{F^{\frac{A_n}{n^\alpha}}\}$ is almost surely Cauchy. Since d_{BL} is complete on the space of probability measures on $\mathbb{R}$, there exists a probability measure μ such that $F^{\frac{A_n}{n^\alpha}} \xrightarrow{w} \mu$.

Also

$$\lim_{l} d_{BL}\big(\mu, \mu^{(l)}\big) = \lim_{l} \lim_{n} d_{BL}\big(F^{\frac{A_n}{n^\alpha}}, F^{\frac{A_n^{(l)}}{n^\alpha}}\big) = 0,$$

being the weak limit of the nonrandom sequence $\{\mu^{(l)}\}$, μ itself is nonrandom. This completes the proof of (a). □

Proof of (b): We have established that

$$F^{X_n} \xrightarrow{w} \mu \quad w.p.\ 1.$$

Also from Theorem 1, as $n \longrightarrow \infty$, $\int x^k \, d\mu_n \longrightarrow L_k w.p.\ 1$, μ_n being the ESD of X_n. In particular

$$\sup_n \int x^{2k} \, d\mu_n < \infty \quad w.p.\ 1,\ \forall k.$$

Hence if Y_n follows μ_n, then almost surely $\{Y_n\}$ are uniformly integrable. Hence we have

$$\int x^k \, d\mu = \lim_n \int x^k \, d\mu_n = L_k.$$

For the other part, consider the path $\gamma_0 \in \Gamma_{2k}$ given by $\gamma_0 = (-i_0, 0, -i_0, 0, \ldots, -i_0, 0)$. Then

$$\begin{aligned} L_{2k} &= \frac{1}{2\alpha k + 1} \sum_{\gamma \in \Gamma_{2k}} \prod_{0 \le i \le r,\, j \le 0} m_{i, l_{j,j+i}(\gamma)} \\ &\ge \frac{1}{2\alpha k + 1} \prod_{0 \le i \le r,\, j \le 0} m_{i, l_{j,j+i}(\gamma_0)} = \frac{1}{2\alpha k + 1} m_{i_0, l_{-i_0,0}(\gamma_0)} \\ &= \frac{1}{2\alpha k + 1} m_{i_0, 2k}. \end{aligned}$$

This implies

$$\limsup_k (L_{2k})^{\frac{1}{2k}} \ge \limsup_k \left(\frac{1}{2\alpha k + 1} m_{i_0, 2k}\right)^{\frac{1}{2k}} = \infty,$$

and hence μ has unbounded support. This completes the proof. □

3.3 Proof of Theorem 3

(a) First we prove convergence in expectation. Let

$$S=\{\gamma:\gamma=(j_0,\dots,j_{2k}),\ j_0=j_{2k},\ 1\le j_u\le n,\ |j_{u+1}-j_u|\le r\ \forall u\}.$$

Then

$$\begin{aligned}\frac{\operatorname{Tr}(S_n^k)}{n}&=\frac{1}{n^{2\alpha k+1}}\operatorname{Tr}\big(A_nA_n^T\big)^k\\&=\frac{1}{n^{2\alpha k+1}}\sum_{\gamma\in S}a_{j_0j_1}a_{j_2j_1}\dots a_{j_{2k-2}j_{2k-1}}a_{j_{2k}j_{2k-1}}\\&=\frac{1}{n^{2\alpha k+1}}\sum_{p=1}^{n}\sum_{\gamma\in S,\,\max\gamma_j=p}a_{j_0j_1}a_{j_2j_1}\dots a_{j_{2k-2}j_{2k-1}}a_{j_{2k}j_{2k-1}}.\end{aligned}$$

Denoting the inner sum by S_p, we write

$$\frac{\operatorname{Tr}(S_n^k)}{n}=\frac{1}{n^{2\alpha k+1}}\sum_{p=1}^{n}S_p.\tag{3.3}$$

Now $\max\gamma_j=p$, $j_0=j_{2k}$ and $|j_{u+1}-j_u|\le r$ together imply $\min\gamma_j\ge p-rk$. Hence for $p\ge rk+1$, we have $\{\gamma: j_0=j_{2k},\ 1\le j_u\le n,\ |j_{u+1}-j_u|\le r,\ \max\gamma_j=p\}=\{\gamma+p:\gamma\in\Gamma_{2k}\}$.

For $p\ge rk+1$,

$$\begin{aligned}\mathbf{E}\bigg(\frac{S_p}{p^{2\alpha k}}\bigg)&=\frac{1}{p^{2\alpha k}}\sum_{\gamma\in\Gamma_{2k}}\mathbf{E}(a_{j_0+p,j_1+p}a_{j_2+p,j_1+p}\dots a_{j_{2k-2}+p,j_{2k-1}+p}a_{j_{2k}+p,j_{2k-1}+p})\\&=\frac{1}{p^{2\alpha k}}\sum_{\gamma\in\Gamma_{2k}}\mathbf{E}\prod_{|h|\le 2r,\,i\le 0}b_{h,i+p}^{l_{i\to i+h,R}(\gamma)+l_{i+h\to i,B}(\gamma)}.\end{aligned}$$

Now, we have $\mathbb{E}(\frac{b_{h,i+p}}{p})^{l_{i\to i+h,R}(\gamma)+l_{i+h\to i,B}(\gamma)}\to m'_{h,l_{i\to i+h,R}(\gamma)+l_{i+h\to i,B}(\gamma)}$.

Since $\sum_{|h|\le 2r,\,i\le 0}(l_{i\to i+h,R}(\gamma)+l_{i+h\to i,B}(\gamma))=2k$, we have as $p\longrightarrow\infty$,

$$\mathbf{E}\bigg(\frac{S_p}{p^{2\alpha k}}\bigg)\longrightarrow\sum_{\gamma\in\Gamma_{2k}}\prod_{|h|\le r,\,i\le 0}m'_{h,l_{i\to i+h,R}(\gamma)+l_{i+h\to i,B}(\gamma)}.\tag{3.4}$$

Hence using (3.4) and (3.3), it follows that $\mathbf{E}(\frac{\operatorname{Tr}(S_n^k)}{n})\to L'_k$.

To establish almost sure convergence, we shall show that $\frac{1}{n^{2\alpha k+1}}\sum_{p=1}^{n}(S_p-\mathbf{E}(S_p))\stackrel{a.s.}{\to}0$.

Consider $p \geq rk+1$, then we have

$$\begin{aligned}
&\frac{1}{p^{4\alpha k}}\mathbf{Var}(\mathbf{S_p})\\
&\quad\leq \frac{1}{p^{4\alpha k}}\mathbf{E}\big(S_p^2\big)\\
&\quad= \frac{1}{p^{4\alpha k}}\mathbf{E}\bigg(\sum_{\gamma\in\Gamma_{2k}} a_{j_0+p,j_1+p}a_{j_2+p,j_1+p}\dots a_{j_{2k-2}+p,j_{2k-1}+p}a_{j_{2k}+p,j_{2k-1}+p}\bigg)^2\\
&\quad\leq C\frac{1}{p^{4\alpha k}}\sum_{\gamma\in\Gamma_{2k}}\mathbf{E}\big(a^2_{j_0+p,j_1+p}a^2_{j_2+p,j_1+p}\dots a^2_{j_{2k-2}+p,j_{2k-1}+p}a^2_{j_{2k}+p,j_{2k-1}+p}\big)
\end{aligned}$$

where $C=|\Gamma_{2k}|$ is free of p. Hence

$$\begin{aligned}
\frac{1}{p^{4\alpha k}}\mathbf{Var}(S_p) &\leq C\frac{1}{p^{4\alpha k}}\sum_{\gamma\in\Gamma_{2k}}\mathbf{E}\prod_{|h|\leq 2r,\, i\leq 0} b_{h,i+p}^{2l_{i\to i+h,R}(\gamma)+2l_{i+h\to i,B}(\gamma)}\\
&\to \sum_{\gamma\in\Gamma_{2k}}\prod_{|h|\leq r,\, i\leq 0} m'_{h,2l_{i\to i+h,R}(\gamma)+2l_{i+h\to i,B}(\gamma)}.
\end{aligned}$$

Then we have $\frac{1}{p^{4\alpha k+2}}\mathbf{Var}(S_p)=O(\frac{1}{p^2})$ and hence

$$\sum_1^\infty \frac{1}{p^{4\alpha k+2}}\mathbf{Var}(S_p)<\infty. \tag{3.5}$$

Now the paths $\gamma=(j_0,\dots,j_{2k})$ contributing to S_p satisfy $\min\gamma_j\geq p-rk$. Hence $\forall t\geq 0$, S_t and S_{t+rk+1} are independent. Hence by (3.5) we have, for each fixed $t\leq rk$, the sequence $\{Y_n\}$ given by

$$Y_n=\sum_{p=0}^{n}\frac{S_{t+p(rk+1)}-\mathbf{E}(S_{t+p(rk+1)})}{(t+p(rk+1))^{2\alpha k+1}}$$

is an L^2 bounded martingale. Hence for every $t\leq rk$, $\sum_{p=0}^\infty \frac{S_{t+p(rk+1)}-\mathbf{E}(S_{t+p(rk+1)})}{(t+p(rk+1))^{2\alpha k+1}}$ converges almost surely and hence $\sum_{p=1}^\infty \frac{S_p-\mathbf{E}(S_p)}{p^{2\alpha k+1}}$ converges almost surely. Finally we use Kronecker's lemma to conclude that $\frac{1}{n^{2\alpha k+1}}\sum_{p=1}^n(S_p-\mathbf{E}(S_p))\stackrel{a.s.}{\to}0$, which completes the proof of part (a).

(b) Following Popescu's arguments for the tridiagonal model we can easily construct a bounded operator on a Hilbert space such that L'_k's will be the moments of a probability measure concentrated on the spectrum of the operator.

Assume μ'_j is supported on I_j, $|j|\leq r$ and let $I=\prod_{|j|\leq r}I_j$ and that $\mu'=\bigotimes_{|j|\leq r}\mu'_j$.

Let

$$\Omega=\prod_{i\in\mathbb{Z}}I_i,\qquad P=\bigotimes_{\mathbb{Z}}\mu\quad\text{and}\quad H=\bigoplus_{\mathbb{Z}}L^2(\Omega,P).$$

Let us denote a typical $\omega \in \Omega$ as $\omega = (\omega_i)_{i\in\mathbb{Z}}$ where $\omega_i = (\omega_{i,-r}, \dots, \omega_{i,r}) \in I$. We define the operator

$$T\colon\quad H \to H$$
$$x \mapsto y$$

where $y_j(\omega) = \sum_{|h|\le r} \omega_{j,h} x_{j+h}(\omega)$ if $x = (x_j)_{j\in\mathbb{Z}}$. Since each I_j is bounded, T is a bounded operator. It can be represented by the following matrix:

$$A = \begin{bmatrix} & & & \cdots & & & \cdots & & & \\ \cdots & X_{-r,-1} & X_{-r+1,-1} & \cdots & X_{0,-1} & X_{1,-1} & X_{2,-1} & \cdots & 0 & 0 & \cdots \\ \cdots & 0 & X_{-r,0} & \cdots & X_{-1,0} & X_{0,0} & X_{1,0} & \cdots & X_{r,0} & 0 & \cdots \\ \cdots & 0 & 0 & \cdots & X_{-2,1} & X_{-1,1} & X_{0,1} & \cdots & X_{r-1,1} & X_{r,1} & \cdots \\ & & & \cdots & & & \cdots & & & \end{bmatrix} \tag{3.6}$$

where $X_{h,n}(\omega) = \omega_{n,h}$ for $|h| \le r$, $n \in \mathbb{Z}$. We note that $\int X^k_{h,n}\,dP = m'_{hk}$. Define U to be the self-adjoint, bounded linear operator $U = TT^*$. Let $\mathbf{e} \in H$ be the element given by $\mathbf{e_j} \equiv 0$, $\forall j \ne 0$ and $\mathbf{e_0} \equiv 1$. Then

$$\begin{aligned} \langle U^k\mathbf{e}, \mathbf{e}\rangle &= \int \big(AA^T\big)^k_{0,0}\,dP \\ &= \int \sum_{\substack{j_i\in\mathbb{Z}\\ |j_{i+1}-j_i|\le r}} a_{0,j_1}a_{j_2,j_1}\dots a_{j_{2k-2},j_{2k-1}}a_{0,j_{2k-1}}\,dP \\ &= \int \sum_{\gamma\in\Gamma_{2k}} \prod_{\substack{|h|\le r\\ i\le 0}} X_{h,i}^{l_{i\to i+h,R}(\gamma)+l_{i+h\to i,B}(\gamma)}\,dP \\ &= \sum_{\gamma\in\Gamma_{2k}} \prod_{\substack{|h|\le r\\ i\le 0}} m'_{h,l_{i\to i+h,R}(\gamma)+l_{i+h\to i,B}(\gamma)}. \end{aligned}$$

The spectral theorem furnishes a probability measure Λ supported on the spectrum of U such that $\langle U^k\mathbf{e}, \mathbf{e}\rangle = \int x^k\,d\Lambda$ (Λ is a probability measure as $\|\mathbf{e}\| = 1$). Hence if Z is distributed according to Λ and Y is a uniform random variable independent of Z, then the kth moment of $Y^{2\alpha}Z$ is L'_k, $\forall k \ge 0$.

Since $Y^{2\alpha}Z$ is compactly supported, we conclude that the sequence L'_k is the moment sequence of a unique compactly supported probability measure and hence the ESD's of the sequence S_n converges almost surely to μ'.

(c) If $m'_{jk} = a^k_j$ then $I_j = \{a_j\}$, Ω consists of a singleton ω with $\omega_{n,h} = a_h$, $\forall n$, $H = l^2$, U is given by the matrix AA^T whose i,jth entry is $(AA^T)_{ij} = \sum_{-r}^{r-h} a_i a_{i+h} = c_h$ (say) where $h = |j - i|$.

Identifying $L^2(S^1)$ with l^2 via the isometry which sends $f \in L^2$ to the sequence of its Fourier coefficients $(\widehat{f}(n))$, we see that U on l^2 corresponds to the operator M on $L^2(S^1)$ given by $Mf = p'f$ where $p'(x) = c_0 + \sum_{k=1}^r c_k(e^{ikx} + e^{-ikx})$. Since $\mathbf{e}$

corresponds to the function which is identically 1, we get

$$\langle U^k \mathbf{e}, \mathbf{e}\rangle = \langle M^k \mathbf{1}, \mathbf{1}\rangle = \int_{S^1} \left(p'(x)\right)^k d\lambda$$

and the rest is immediate from part (b).

Proof of Corollary 3.1 For a path $\gamma = (j_0, \dots, j_k)$, let $t_h(\gamma) = \#\{u : |j_{u+1} - j_u| = h\}$, for $h \geq 0$, i.e. the number of transitions with jump h.

Then under the assumptions of Corollary 3.1,

$$\begin{aligned}
L_k &= \frac{1}{\alpha k + 1} \sum_{\gamma \in \Gamma_k} \prod_{0 \leq h \leq r,\, j \leq 0} a_h^{l_{j,j+h}(\gamma)} \\
&= \frac{1}{\alpha k + 1} \sum_{\gamma \in \Gamma_k} \prod_{0 \leq h \leq r} a_h^{t_h(\gamma)} \\
L'_k &= \frac{1}{2\alpha k + 1} \sum_{\gamma \in \Gamma_{2k}} \prod_{|h| \leq r,\, i \leq 0} a_{|h|}^{l_{i \to i+h, R}(\gamma) + l_{i+h \to i, B}(\gamma)} \\
&= \frac{1}{2\alpha k + 1} \sum_{\gamma \in \Gamma_{2k}} \prod_{0 \leq h \leq r} a_h^{t_h(\gamma)}.
\end{aligned}$$

Hence $L'_k = L_{2k}$. Since all the measures are compactly supported, the result follows. □

Proof of Corollary 3.2 This can be done in the exact same way as in the proof of Theorem 2. We define $S_n^{(l)} = n^{-2\alpha} A_n^{(l)} A_n^{(l)T}$ ($A_n^{(l)}$ being the matrices with truncated entries). Using (3.2) we get

$$\begin{aligned}
d_{BL}^2\left(F^{S_n}, F^{S_n^{(l)}}\right) &\leq \frac{2}{n^2} \operatorname{Tr}\left(S_n + S_n^{(l)}\right) \frac{\operatorname{Tr}((A_n - A_n^{(l)})(A_n - A_n^{(l)})^T)}{n^{2\alpha}} \\
&\leq 4 \frac{\operatorname{Tr}(S_n)}{n} \frac{\operatorname{Tr}((A_n - A_n^{(l)})(A_n - A_n^{(l)})^T)}{n^{2\alpha+1}}.
\end{aligned}$$

Under the assumptions made, $n^{-1} \operatorname{Tr}(S_n)$ remains bounded almost surely and we manipulate the other factor as in proof of Theorem 2 to get the result. □

3.4 Proof of Theorem 4

Let $c_0 = \{1, \dots, m\}$ be the set of colors corresponding to $X^{(n,1)}, \dots, X^{(n,m)}$. Let $c = (i_1, \dots, i_l)$ be the color sequence corresponding to $X^{(n,i_1)}, \dots, X^{(n,i_l)}$.

$$\begin{aligned}
&\mathbf{E} \frac{1}{n} \operatorname{Tr}\left(X^{(n,i_1)} \dots X^{(n,i_l)}\right) \\
&\quad = \frac{1}{n^{\alpha_{i_1} + \dots + \alpha_{i_l} + 1}} \sum_{\substack{\gamma : j_0 = j_l \\ 1 \leq j_u \leq n \\ |j_{u+1} - j_u| \leq r}} \mathbf{E}\left(a_{j_0 j_1}^{(n,i_1)} \dots a_{j_{l-1} j_l}^{(n,i_l)}\right)
\end{aligned}$$

$$= \frac{1}{n^{\alpha_{i_1}+\cdots+\alpha_{i_l}+1}} \sum_{p=1}^{n} \sum_{\substack{\gamma: j_0=j_l \\ 1\le j_u,\, \max\gamma=p \\ |j_{u+1}-j_u|\le r}} \mathbf{E}\big(a_{j_0 j_1}^{(n,i_1)} \ldots a_{j_{l-1} j_l}^{(n,i_l)}\big)$$

$$= \frac{1}{n^{\alpha_{i_1}+\cdots+\alpha_{i_l}+1}} \sum_{p=1}^{n} S_p \text{ (say)}.$$

As in the proof of Theorem 3, for $p \geq \frac{rl}{2} + 1$,

$$\{\gamma : j_0 = j_l,\ 1 \le j_u,\ \max\gamma = p,\ |j_{u+1} - j_u| \le r\} = \{\gamma + p : \gamma \in \Gamma_l^c\}.$$

So we consider $p \geq \frac{rl}{2} + 1$. Then

$$\frac{S_p}{p^{\alpha_{i_1}+\cdots+\alpha_{i_l}}} = \frac{1}{p^{\alpha_{i_1}+\cdots+\alpha_{i_l}}} \sum_{\gamma\in\Gamma_l^c} \mathbf{E}\big(a_{j_0+p, j_1+p}^{(n,i_1)} \ldots a_{j_{l-1}+p, j_l+p}^{(n,i_l)}\big)$$

$$= \frac{1}{p^{\alpha_{i_1}+\cdots+\alpha_{i_l}}} \sum_{\gamma\in\Gamma_l^c} \prod_{\substack{t\in\{i_1,\ldots,i_l\} \\ 0\le s\le r,\, j\le 0}} {b_{s,j+p}^{(t)}}^{l_{j,j+s;t}(\gamma)}.$$

Since

$$\sum_{0\le s\le r,\, j\le 0} l_{j,j+s;t}(\gamma) = \#\{k | i_k = t\},$$

we get

$$\frac{S_p}{p^{\alpha_{i_1}+\cdots+\alpha_{i_l}}} \to \sum_{\gamma\in\Gamma_l^c} \prod_{\substack{1\le t\le m \\ 0\le s\le r, j\le 0}} m_{s,l_{j,j+s;t}(\gamma)}^{(t)}.$$

Hence by an application of Cesaro sums one establishes convergence in expectation.

The almost sure convergence can be proved using arguments similar to those used in the proof of (a) of Theorem 3. We omit the details.

The proof of part (b) follows along the same line of argument as in part (a) and is omitted. □

Acknowledgements A. Bose's research was supported by J.C. Bose National Fellowship, Dept. of Science and Technology, Govt. of India. Part of the work was done while visiting Dept. of Economics, University of Cincinnati. S. Sen was supported by NYU graduate fellowship under Henry M. MacCracken Program.

References

1. Bai, Z.D.: Methodologies in spectral analysis of large dimensional random matrices: a review. Stat. Sin. **9**(3), 611–677 (1999).
2. Bai, Z.D., Silverstein, J.: Spectral Analysis of Large Dimensional Random Matrices. Science Press, Beijing (2006)
3. Bhatia, R.: Matrix Analysis. Springer, New York (1997)

4. Bose, A., Gangopadhyay, S., Sen, A.: Limiting spectral distribution of XX' matrices. Ann. Inst. Henri Poincaré, B Calc. Probab. Stat. **46**(3), 677–707 (2010). doi:10.1214/09-AIHP329
5. Bose, A., Hazra, R.S., Saha, K.: Patterned random matrices and method of moments. In: Proceedings of the International Congress of Mathematicians, Hyderabad, India, 2010. pp. 2203–2230. World Scientific, Singapore (2010)
6. Dudley, R.M.: Real Analysis and Probability. Cambridge University Press, Cambridge (2002)
7. Popescu, Ionel: General tridiagonal random matrix models, limiting distributions and fluctuations. Probab. Theory Relat. Fields **144**, 179–220 (2009)