LIMITING SPECTRAL DISTRIBUTION FOR WIGNER MATRICES WITH DEPENDENT ENTRIES*

Arijit Chakrabarty[a], Rajat Subhra Hazra[b], Deepayan Sarkar[a]

[a]Theoretical Statistics and Mathematics Division, Indian Statistical Institute
7, S.J.S. Sansanwal Marg, New Delhi 110 016, India
[b]Theoretical Statistics and Mathematics Division, Indian Statistical Institute
203 B.sT. Road, Kolkata 700108, India

(Received December 19, 2014)

In this article, we show the existence of limiting spectral distribution of a symmetric random matrix whose entries come from a stationary Gaussian process with covariances satisfying a summability condition. We provide an explicit description of the moments of the limiting measure. We also show that in some special cases the Gaussian assumption can be relaxed. The description of the limiting measure can also be made via its Stieltjes transform which is characterized as the solution of a functional equation. In two special cases, we get a description of the limiting measure — one as a free product convolution of two distributions, and the other one as a dilation of the Wigner semicircular law.

DOI:10.5506/APhysPolB.46.1637
PACS numbers: 02.10.Yn, 05.40.–a

1. Introduction

In his seminal paper, Wigner [1] showed that for a symmetric random matrix with independent on and off diagonal entries satisfying some moment conditions, the empirical spectral distribution (henceforth ESD) converges to the Wigner semicircle law (defined in (5.2), henceforth WSL). Subsequent work has tried to obtain a better understanding of the spectrum of such matrices, which plays an important role in physics as well as other branches of mathematics such as operator algebras. Recently, there has been interest in how far the independence assumption and the moment conditions can be relaxed. The reader may refer to the recent review article by Ben Arous and Guionnet [2], and the references therein, for an overview of currently available results.

* Presented at the Conference "Random Matrix Theory: Foundations and Applications", Kraków, Poland, July 1–6, 2014.

Relaxation of the independence assumption has been investigated by Götze and Tikhomirov [3], Chatterjee [4], Rashidi Far *et al.* [5], and Hofmann-Credner and Stolz [6]. The sample covariance matrix has been studied, imposing some dependence on the rows and columns, by Hachem *et al.* [7], Adamczak [8], and Pfaffel and Schlemm [9]. However, the limiting spectral distributions (henceforth LSD) obtained by considering symmetric matrices with the independence assumption weakened have stayed within the WSL regime for the most part. One exception is Anderson and Zeitouni [10], who considered the LSD of Wigner matrices where on and off diagonal elements form a finite-range dependent random field; in particular, the entries are assumed to be independent beyond a finite range, and within the finite range the correlation structure is given by a kernel function.

1.1. Motivation

We begin with a few examples to motivate the problem studied in this article. In each of the following examples, a random field $\{Z_{i,j} : i, j \geq 1\}$ is developed. For $n \geq 1$, let A_n be the $n \times n$ matrix whose $(i,j)^{\text{th}}$ entry is $Z_{i\wedge j, i\vee j}$. The question is whether the ESD of $A_n/\sqrt{n}$ converges as $n \to \infty$, and if so, can one identify the limit.

Example 1. Let $\{Z_{i,j} : i, j \geq 1\}$ be a mean zero Gaussian process such that $E\left[Z_{i,j} Z_{i+k,j+l}\right] = \rho^{|k|+|l|}$ for integers i, j, k, l such that $i, j, i+k, j+l \geq 1$, where $|\rho| < 1$ is fixed. This process can be thought of as a "two-dimensional AR(1) process", because $\{Z_{i,j} : j \geq 1\}$ is an AR(1) process for fixed i, as is $\{Z_{i,j} : i \geq 1\}$ for fixed j.

Example 2. Assume that $\{G_{i,j} : i, j \geq 1\}$ are i.i.d. standard Gaussian random variables, and N is a fixed positive integer. Define

$$Z_{i,j} := \sum_{k=0}^{N} \sum_{l=0}^{N} G_{i+k,j+l},\ i, j \geq 1\,.$$

Example 3. Suppose that $(G_n : n \in \mathbb{Z})$ is a mean zero variance one stationary Gaussian process. Let $(G_n^i : n \in \mathbb{Z})$ be i.i.d. copies of $(G_n : n \in \mathbb{Z})$ for $i = \ldots, -2, -1, 0, 1, 2, \ldots$. Set $Z_{i,j} := G_i^{i-j}$, $i, j \in \mathbb{Z}$.

Example 4. Let $\{c_{k,l}\}$ be real numbers such that

$$\begin{aligned} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} c_{k,l}^2 &< \infty\,, \\ c_{k,l} &= c_{l,k} \qquad \text{for all } k, l \in \mathbb{Z}\,, \\ \sum_{l=-\infty}^{\infty} c_{k,l} c_{k',l} &= 0 \qquad \text{for all } k \neq k'\,. \end{aligned}$$

As in Example 2, let $\{G_{i,j} : i,j \geq 1\}$ be i.i.d. standard Gaussian random variables. Define

$$Z_{i,j} := \sum_{k,l\in\mathbb{Z}} c_{k,l} G_{i-k,j-l}\,, \qquad i,j \in \mathbb{Z}\,.$$

It is shown later in Section 5 that for Examples 1 and 2, the LSD of $A_n/\sqrt{n}$ is the free product convolution of the WSL with a distribution supported on a compact subset of $[0,\infty)$, and for Examples 3 (under the additional assumption that $\sum_{n=1}^{\infty} |E(G_0 G_n)| < \infty$) and 4, the LSD is a dilation of the WSL. To the best of our understanding, Example 2 is the only one of the above examples where the result follows from the work of Anderson and Zeitouni [10], because that is the only example where two entries are independent if their distance is above a threshold.

1.2. Outline of our contribution

Motivated by these examples, this article considers a random matrix model where on and off diagonal entries form a stationary Gaussian field, with the covariance of the entries being summable. Specifically, let $(Z_{i,j} : i,j \in \mathbb{Z})$ be a stationary, mean zero, variance one Gaussian process. Stationarity here means that for $k,l \in \mathbb{Z}$,

$$(Z_{i+k,j+l} : i,j \in \mathbb{Z}) \stackrel{d}{=} (Z_{i,j} : i,j \in \mathbb{Z})\,.$$

For $i,j \geq 1$, set

$$X_{i,j} := Z_{i\wedge j, i\vee j}\,,$$

and let

$$A_n := ((X_{i,j}))_{n\times n}\,, \qquad n \geq 1\,. \tag{1.1}$$

Let $\lambda_1 \leq \cdots \leq \lambda_n$ be the eigenvalues of A_n, which are real because A_n is symmetric, and denote

$$\mu_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{\{\lambda_i/\sqrt{n}\}}\,. \tag{1.2}$$

The main result of this article is Theorem 2.1, stated in Section 2 along with an outline of the proof, which gives a set of conditions on the covariance of $\{X_{i,j}\}$ under which the ESD μ_n converges weakly in probability. The assumption of Gaussianity, although important in the proof, can be relaxed to allow for a fairly general class of input sequences using the Lindeberg type argument developed in Chatterjee [11]. This is done in Section 3, where we show that by specializing on an infinite order moving average process with independent inputs satisfying the Pastur condition (see (3.1)), Theorem 2.1 and an invariance principle can be used to establish the convergence of the ESD. An interpretation of the limiting moments in terms of functions of non-crossing pair partitions is used to derive the Stieltjes transform of the measure in Section 4. There, an explicit description of the Stieltjes transform is provided using the moment formula and some properties of the Kreweras complement. The form of the Stieltjes transform indicates a relationship with operator-valued semicircular variables studied in Speicher [12] (for the application of free probability to random matrices, see the recent review by Speicher [13]). Finally, in Section 5, two explicit examples are described where we get better descriptions of the limit: Theorem 5.1 gives conditions under which the LSD is the free multiplicative convolution of the WSL and another distribution. Theorem 5.2 gives conditions under which the LSD is the WSL. Detailed proofs of these results have been omitted for brevity and can be found in a longer version of this paper available on `arXiv.org` [14].

2. The main result

In this section, we state the main result, and give an outline of the proof. Let the $n \times n$ random symmetric matrix A_n be as in (1.1), and set μ_n to be ESD of $A_n/\sqrt{n}$, as defined in (1.2). Before stating the main result, we need a few more notations and assumptions. Define

$$R(u,v) := E\left[Z_{0,0} Z_{-u,v}\right] , \qquad u,v \in \mathbb{Z} . \tag{2.1}$$

The assumptions are the following.

Assumption 1. $R(\cdot,\cdot)$ is symmetric, that is,

$$R(u,v) = R(v,u) \quad \text{for all} \quad u,v \in \mathbb{Z} . \tag{2.2}$$

Assumption 2. $R(\cdot,\cdot)$ is absolutely summable, that is,

$$\bar{R} := \sum_{u,v \in \mathbb{Z}} |R(u,v)| < \infty . \tag{2.3}$$

An immediate consequence of Assumption 1 and stationarity is that

$$R(u,v) = R(-v,-u) , \qquad u,v \in \mathbb{Z} . \tag{2.4}$$

Fix $\sigma \in NC_2(2m)$, the set of non-crossing pair partitions of $\{1,\ldots,2m\}$. Let $(V_1,\ldots,V_{m+1})$ denote the Kreweras complement of σ, which is the maximal partition $\overline{\sigma}$ of $\{\overline{1},\ldots,\overline{2m}\}$ such that $\sigma \cup \overline{\sigma}$ is a non-crossing partition of $\{1,\overline{1},\ldots,2m,\overline{2m}\}$. For $1 \le i \le m+1$, denote

$$V_i := \left\{v_1^i,\ldots,v_{l_i}^i\right\} . \tag{2.5}$$

Denote

$$S(\sigma) := \left\{ (k_1, \ldots, k_{2m}) \in \mathbb{Z}^{2m} : \sum_{j=1}^{l_s} k_{v_j^s} = 0\,,\ s = 1, \ldots, m+1 \right\}, \tag{2.6}$$

and define

$$\beta_{2m} := \sum_{\sigma \in NC_2(2m)} \sum_{k \in S(\sigma)} \prod_{(u,v) \in \sigma} R(k_u, k_v)\,, \qquad m \geq 1\,. \tag{2.7}$$

Indeed, if $\sigma = \{(u_1, u_{m+1}), \ldots, (u_m, u_{2m})\}$, then

$$\sum_{k \in S(\sigma)} \prod_{(u,v) \in \sigma} |R(k_u, k_v)| < \infty\,.$$

That is, as a consequence of Assumption 2, (2.7) makes sense.

The main result of this article is the following.

Theorem 2.1. *Under Assumptions 1 and 2, μ_n converges weakly in probability to a distribution μ. The distribution μ is even and compactly supported.*

Remark 1. As is common in the literature, the phrase "μ_n converges weakly in probability to a distribution μ" means that

$$L(\mu_n, \mu) \xrightarrow{P} 0\,,$$

as $n \to \infty$, where L, the Lévy distance, is defined by

$$\begin{aligned} L(\nu_1, \nu_2) \ := \ & \inf \{\varepsilon > 0 : \nu_1((-\infty, x - \varepsilon]) - \varepsilon \leq \nu_2((-\infty, x]) \\ & \leq \nu_1((-\infty, x + \varepsilon]) + \varepsilon \quad \text{for all} \quad x \in \mathbb{R}\}\,, \end{aligned} \tag{2.8}$$

for probability measures ν_1, ν_2 on $\mathbb{R}$.

We end this section with a brief outline of the proof of the above result.

Sketch of the proof of Theorem 2.1. As is standard in a proof by the method of moments, what needs to be shown is that for fixed $m \geq 1$,

$$\lim_{n\to\infty} n^{-(m+1)} \sum_{i_1,\ldots,i_{2m}=1}^{n} E\left[X_{i_1,i_2} \ldots X_{i_{2m-1},i_{2m}} X_{i_{2m},i_1}\right] = \beta_{2m}\,. \tag{2.9}$$

Note that the expectation of the odd moments vanish anyway, and hence it suffices to consider only the even ones. As in the proof of the classical Wigner's result, the first step is to get rid of the "non-pair matched" tuples $i = (i_1, \dots i_{2m})$ in the above sum. Fix $N \geq 1$, and say that a tuple $i \in \{1, \dots, n\}^{2m}$ is N-pair matched if there exists a pair partition π of $\{1, \dots, 2m\}$ such that for all $(u, v) \in \pi$,

$$|i_{u-1} \wedge i_u - i_{v-1} \wedge i_v| \vee |i_{u-1} \vee i_u - i_{v-1} \vee i_v| \leq N \,,$$

with the convention $i_0 := i_{2m}$. It needs to be shown that if $C_{N,n}$ denotes the set of tuples in $\{1, \dots, n\}^{2m}$ which are not N-pair matched, then

$$\lim_{N\to\infty} \limsup_{n\to\infty} n^{-(m+1)} \left| \sum_{i \in C_{N,n}}^{n} E\left[X_{i_1,i_2} \dots X_{i_{2m-1},i_{2m}} X_{i_{2m},i_1}\right] \right| = 0 \,. \tag{2.10}$$

Unlike in the classical Wigner's result, this is a non-trivial step in our situation because not only does the above sum not vanish for N large, even showing that the expectation in modulus is less than some ε is not enough because $\# C_{N,n} \sim n^{2m}$ as $n \to \infty$, and the sum is scaled only by $n^{(m+1)}$. This is precisely the step where Assumption 2 plays an important role.

Once (2.10) is established, what remains to be shown for (2.9) is that for fixed N,

$$\begin{aligned} &\lim_{n\to\infty} n^{-(m+1)} \sum_{i \in C_{N,n}^c}^{n} E\left[X_{i_1,i_2} \dots X_{i_{2m-1},i_{2m}} X_{i_{2m},i_1}\right] \\ &= \sum_{\sigma \in NC_2(2m)} \sum_{k \in S(\sigma) : \max_j |k_j| \leq N} \prod_{(u,v) \in \sigma} R(k_u, k_v) \,. \end{aligned} \tag{2.11}$$

By standard combinatorial arguments, the sum over $C_{N,n}^c$ can be shown to be asymptotically equivalent to the sum over all tuples that are *Catalan with respect to some* $\sigma \in NC_2(2m)$, that is, whenever $(j, k) \in \sigma$,

$$|i_{j-1} - i_k| \vee |i_j - i_{k-1}| \leq N \,.$$

The final step is to show that for fixed $\sigma \in NC_2(2m)$, if D_σ denotes the set of tuples in $\{1, \dots, n\}^{2m}$ which are Catalan with respect to σ, then

$$\begin{aligned} &\lim_{m\to\infty} \sum_{i \in D_\sigma} E\left[X_{i_1,i_2} \dots X_{i_{2m-1},i_{2m}} X_{i_{2m},i_1}\right] \\ &= \sum_{k \in S(\sigma) : \max_j |k_j| \leq N} \prod_{(u,v) \in \sigma} R(k_u, k_v) \,. \end{aligned}$$

This follows by computing the expectation via Wick's formula, and observing that in that formula, the contribution of all the pair partitions excluding σ is asymptotically negligible. This final step establishes (2.11). $\square$

3. The linear process

In this section, we study the ESD of a random matrix whose entries are generated from a linear process with independent random variables as the input sequence. In particular, let $\{\epsilon_{i,j} : i, j \in \mathbb{Z}\}$ be independent, mean zero, variance one random variables which satisfy the Pastur condition

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i,j=1}^{n} E\left[\epsilon_{i,j}^2 \mathbf{1}\left(|\epsilon_{i,j}| > \varepsilon\sqrt{n}\right)\right] = 0 \quad \text{for all} \quad \varepsilon > 0\,. \tag{3.1}$$

Let $\{c_{k,l} : k, l \in \mathbb{Z}\}$ be a collection of deterministic real numbers such that

$$0 < \sum_{k,l\in\mathbb{Z}} |c_{k,l}| < \infty \tag{3.2}$$

and

$$c_{k,l} = c_{l,k}\,, \qquad k, l \in \mathbb{Z}\,. \tag{3.3}$$

Define

$$Z_{i,j} := \sum_{k,l\in\mathbb{Z}} c_{k,l}\epsilon_{i-k,j-l}\,, \qquad i, j \in \mathbb{Z}\,, \tag{3.4}$$

where the sum on the right-hand side converges in L^2 because $c_{k,l}$ are square summable, which is a consequence of (3.2). While the family of random variables $\{Z_{i,j} : i, j \in \mathbb{Z}\}$ need not be stationary because the distributions of $\epsilon_{i,j}$ are not necessarily identical, it is easy to see that

$$\begin{aligned} E(Z_{i,j}) &= 0\,, \qquad i, j \in \mathbb{Z}\,, \\ E(Z_{i,j}Z_{i-u,j+v}) &= \sum_{k,l\in\mathbb{Z}} c_{k,l}c_{k-u,l+v} =: R(u, v)\,, \qquad i, j, u, v \in \mathbb{Z}\,. \end{aligned} \tag{3.5}$$

Define the $n \times n$ symmetric random matrix A_n and μ_n, the ESD of $A_n/\sqrt{n}$ by (1.1) and (1.2) respectively. The assumption (3.2) ensures that

$$\sum_{u\in\mathbb{Z}}\sum_{v\in\mathbb{Z}} |R(u, v)| \le \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} \left[|c_{k,l}| \sum_{u\in\mathbb{Z}}\sum_{v\in\mathbb{Z}} |c_{k-u,l+v}|\right] = \left[\sum_{k,l\in\mathbb{Z}} |c_{k,l}|\right]^2 < \infty\,.$$

Therefore, we define β_{2m} by (2.7). Let μ be the unique probability measure whose odd moments are all zero, and for $m \ge 1$, the $2m^{\text{th}}$ moment equals β_{2m}.

The content of this section is the following result.

Theorem 3.1. *Under assumptions (3.1) to (3.3), μ_n converges weakly in probability to μ.*

Sketch of the proof. Fix $m \geq 1$ and let

$$Z_{i,j}^{(m)} = \sum_{k,l=-m}^{m} c_{k,l}\epsilon_{i-k,j-l} \quad \text{for} \quad i,j \geq 1\,.$$

Define

$$A_n^{(m)} := \left(\left(Z_{i,j}^{(m)}\right)\right)_{n\times n}, \qquad n \geq 1\,.$$

We next define a similar random matrix model, but with Gaussian entries. Let $(G_{i,j} : i,j \in \mathbb{Z})$ be i.i.d. standard Gaussian, and set

$$Y_{i,j}^{(m)} = \sum_{k,l=-m}^{m} c_{k,l}G_{i-k,j-l} \quad \text{for} \quad i,j \geq 1\,.$$

Denote

$$B_n^{(m)} := \left(\left(Y_{i,j}^{(m)}\right)\right)_{n\times n}, \qquad n \geq 1\,.$$

For a finite linear process, it can be shown that the Stieltjes transform of the ESD of a matrix made up of Gaussian random variables and another with general entries satisfying (3.1) are close to each other using the Lindeberg type argument, developed in [11], of "replacing $\epsilon_{i,j}$ by $G_{i,j}$ one at a time". That is,

$$\frac{1}{n}\left[\mathrm{Tr}\left(\left(zI_n - \frac{A_n^{(m)}}{\sqrt{n}}\right)^{-1}\right) - \mathrm{Tr}\left(\left(zI_n - \frac{B_n^{(m)}}{\sqrt{n}}\right)^{-1}\right)\right] \xrightarrow{P} 0\,, \tag{3.6}$$

as $n \to \infty$, for all z in the complex plane with non-zero imaginary part. The arguments for the above are very similar to those in Subsections 2.3 and 2.4 of [11] and hence are omitted.

It is easy to see that

$$R^{(m)}(u,v) := E\left(Y_{i,j}^{(m)}Y_{i-u,j+v}^{(m)}\right) = \sum_{k,l=-m}^{m} c_{k,l}c_{k-u,l+v}\,, \qquad u,v \in \mathbb{Z}\,.$$

By Theorem 2.1, it follows that for fixed m, as $n \to \infty$, the ESD of $B_n^{(m)}/\sqrt{n}$ converges weakly in probability to the probability measure $\mu^{(m)}$ whose odd moments are all zero and for $l \geq 1$, the $2l^{\text{th}}$ moment is $\beta_{2l}^{(m)}$ defined by

$$\beta_{2l}^{(m)} := \sum_{\sigma\in NC_2(2l)}\sum_{k\in S(\sigma)}\prod_{(u,v)\in\sigma} R^{(m)}(k_u,k_v)\,, \qquad l \geq 1\,,$$

with $S(\sigma)$ being as in (2.6). This, along with (3.6), implies that as $n \to \infty$,

$$\frac{1}{n} \operatorname{Tr} \left(\left(zI_n - \frac{A_n^{(m)}}{\sqrt{n}} \right)^{-1} \right) \xrightarrow{P} \int \frac{1}{z-x} \mu^{(m)}(dx) \,, \qquad z \in \mathbb{C} \setminus \mathbb{R} \,.$$

Recalling from (2.8) the definition of L, a restatement of the above is that

$$L\left(\mu_n^{(m)}, \mu^{(m)}\right) \xrightarrow{P} 0 \,, \tag{3.7}$$

as $n \to \infty$, where $\mu_n^{(m)}$ denotes the ESD of $A_n^{(m)}/\sqrt{n}$. Notice that

$$\lim_{m\to\infty} R^{(m)}(u,v) = R(u,v) \,, \qquad u, v \in \mathbb{Z} \,.$$

By using (3.2) to interchange limit and sum, it follows that

$$\lim_{m\to\infty} \beta_{2l}^{(m)} = \beta_{2l} \,, \qquad l \geq 1 \,.$$

Therefore,

$$\lim_{m\to\infty} L\left(\mu^{(m)}, \mu\right) = 0 \,. \tag{3.8}$$

In view of (3.7) and (3.8), to complete the proof of the result, it suffices to show that

$$\lim_{m\to\infty} \limsup_{n\to\infty} E\left[L^3\left(\mu_n^{(m)}, \mu_n\right)\right] = 0 \,, \tag{3.9}$$

recalling that μ_n is the ESD of $A_n/\sqrt{n}$.

To that end, we shall use the fact that for $n \times n$ (deterministic) symmetric matrices C and D with ESD ν_C and ν_D respectively,

$$L^3(\nu_C, \nu_D) \leq \frac{1}{n} \operatorname{Tr}\left((C-D)^2\right) \,,$$

which is a consequence of the Hoffman–Wielandt inequality; see Corollary A.41, page 502 in Bai and Silverstein [15]. Using this inequality, it is immediate that

$$\begin{aligned} E\left[L^3\left(\mu_n^{(m)}, \mu_n\right)\right] &\leq \frac{1}{n} E\left[\operatorname{Tr}\left[\left(A_n/\sqrt{n} - A_n^{(m)}/\sqrt{n} \right)^2 \right]\right] \\ &= \sum_{k,l \in \mathbb{Z}: |k| \vee |l| > m} c_{k,l}^2 \,. \end{aligned} \tag{3.10}$$

The assumption (3.2) ensures, of course, that $\{c_{k,l}\}$ is square summable, and thus establishes (3.9). Combining this with (3.7) and (3.8) completes the proof. □

4. Stieltjes transform

In this section, a characterization of the Stieltjes transform of μ, the LSD in Theorems 2.1 and 3.1, is given via a functional equation. As the reader may have already noticed, in both the above results, μ is defined via the correlations $R(u, v)$ which is as in (2.1) or (3.5). For this section, let $R(\cdot, \cdot)$ be the correlations of a weakly stationary mean zero variance one process $(Y_{ij} : i, j \in \mathbb{Z})$, that is,

$$\begin{aligned} E(Y_{i,j}) &= 0\,, & i, j \in \mathbb{Z}\,, \\ E\left(Y_{i,j}^2\right) &= 1\,, & i, j \in \mathbb{Z}\,, \\ E(Y_{i,j}Y_{i-u,j+v}) &=: R(u, v)\,, & i, j, u, v \in \mathbb{Z}\,. \end{aligned}$$

As before, we assume (2.2) and (2.3). As before, let μ be the unique even probability measure whose $2m^{\text{th}}$ moment equals β_{2m} which is as defined in (2.7). Recall that the Stieltjes transform of the probability measure μ on $\mathbb{R}$ is denoted by

$$\mathcal{G}(z) = \int_{\mathbb{R}} \frac{1}{z - x}\mu(dx)\,, \qquad z \in \mathbb{C} \setminus \mathbb{R}\,.$$

The main result of this section is Theorem 4.1 below.

Let the Fourier transform of covariance function $\{R(k, l)\}_{k,l\in\mathbb{Z}}$ be given by

$$f(x, y) = \sum_{k,l\in\mathbb{Z}} R(k, l) \exp(2\pi i(kx + ly)) \quad \text{for} \quad (x, y) \in [0, 1] \times [0, 1]\,.$$

Note that by (2.2), it follows that $f(x, y)$ is a real, symmetric function. For stating the main result, we need the following proposition.

Proposition 4.1. *Suppose that $\mathcal{H}_1$ and $\mathcal{H}_2$ are functions from $\mathbb{C} \times [0, 1]$ to $\mathbb{C}$ satisfying the following for $i = 1, 2$:*

1. *for all $x \in [0, 1]$ and $z \in \mathbb{C}$,*
$$z\mathcal{H}_i(z, x) = 1 + \mathcal{H}_i(z, x) \int_0^1 \mathcal{H}_i(z, y) f(x, y) dy\,, \tag{4.1}$$
2. *there exists a neighbourhood N_i (independent of x) of infinity such that for all $x \in [0, 1]$, $\mathcal{H}_i(\cdot, x)$ is analytic on N_i,*
3. *for all $x \in [0, 1]$,*
$$\lim_{z\to\infty} z\mathcal{H}_i(z, x) = 1\,, \tag{4.2}$$

4. and

$$\mathcal{H}(-z,x) = -\mathcal{H}(z,x)\,, \qquad z \in \mathbb{C}\,, \qquad x \in [0,1]\,. \tag{4.3}$$

Then

$$\mathcal{H}_1 \equiv \mathcal{H}_2 \quad on \quad N_1 \cap N_2\,.$$

The following is the main result.

Theorem 4.1. *There exists a function $\mathcal{H}$ satisfying the assumptions of the Proposition 4.1. The Stieltjes transform $\mathcal{G}$ of the LSD μ is given by*

$$\mathcal{G}(z) := \left[\int_0^1 \mathcal{H}(z,x)dx\right]\,, \qquad z \in \mathbb{C}\,.$$

Sketch of the proof. Fix $\sigma \in NC_2(2m)$, and denote its Kreweras complement by $(V_1,\ldots,V_{m+1})$. Although the Kreweras complement is a partition of $\{\overline{1},\ldots,\overline{2m}\}$, for the ease of notation, $V_1,\ldots,V_{m+1}$ will be thought of as subsets of $\{1,\ldots,2m\}$, that is, the overline will be suppressed. In order to ensure uniqueness in the notation, we impose the requirement that the blocks $V_1,\ldots,V_{m+1}$ are ordered in the following way. If $1 \le i < j \le m+1$, then the **maximal** element of V_i is strictly less than that of V_j. Let $\mathcal{T}_\sigma$ be the unique function from $\{1,\ldots,2m\}$ to $\{1,\ldots,m+1\}$ satisfying

$$i \in V_{\mathcal{T}_\sigma(i)}\,, \qquad 1 \le i \le 2m\,.$$

For example, if $\sigma := \{(1,4),(2,3),(5,6)\}$, then $\mathcal{T}_\sigma(1) = 2, \mathcal{T}_\sigma(2) = 1, \mathcal{T}_\sigma(3) = 2, \mathcal{T}_\sigma(4) = 4, \mathcal{T}_\sigma(5) = 3, \mathcal{T}_\sigma(6) = 4$. Define the function L_σ from $\mathbb{R}^{m+1}$ to $\mathbb{R}$ by

$$L_\sigma(x) := \prod_{(u,v)\in\sigma} f\left(x_{\mathcal{T}_\sigma(u)}, x_{\mathcal{T}_\sigma(v)}\right)\,, \qquad x \in \mathbb{R}^{m+1}\,.$$

Finally, set

$$h_\sigma(y) := \int_0^1 \ldots \int_0^1 L_\sigma(x_1,\ldots,x_m,y)dx_m \ldots dx_1\,, \qquad y \in \mathbb{R}\,.$$

We start with defining the functions

$$H_0(x) = 1\,, \qquad H_{2m}(x) = \sum_{\sigma\in NC_2(2m)} h_\sigma(x)\,.$$

Since $\#NC_2(2m) \le 4^m$, it can be shown that the power series

$$\mathcal{H}(z,x) := \sum_{m=0}^{\infty} \frac{H_{2m}(x)}{z^{2m+1}}$$

converges on $\{z \in \mathbb{C} : |z| > 2\bar{R}^{1/2}\}$ for every fixed $x \in [0,1]$. Note that this neighborhood around infinity is independent of $x \in [0,1]$. It is easy to see that $z\mathcal{H}(z,x)$ has a power series expansion with the leading term as 1 and hence $z\mathcal{H}(z,x) \to 1$ as $|z| \to \infty$. It follows from the definition of $\mathcal{H}(z,x)$ that $\mathcal{H}(-z,x) = -\mathcal{H}(z,x)$. It is easy to check that the Stieltjes transform $\mathcal{G}$ of μ satisfies

$$\mathcal{G}(z) = \int_0^1 \mathcal{H}(z,x)dx\,.$$

Equation (4.1) with $\mathcal{H}_i$ replaced by $\mathcal{H}$ is all that remains to be checked.

To that end, we derive a recursion for $\mathcal{H}(z,x)$ using the properties of $h_\sigma(x)$. Recall that there is a natural one–one correspondence between $NC_2(2m)$ and the set of Catalan words of the length of $2m$ with the understanding that two words will be considered identical if one can be obtained from the other by a relabelling of letters. Keeping this correspondence in mind, by an abuse of notation, we shall now consider $h_w(x)$ for Catalan words w, and denote by $NC_2(2m)$ the set of Catalan words of the length of $2m$. Note that any Catalan word w of the length of $2m$ can be written as $w = aw_1aw_2$, for some $w_1 \in NC_2(2k-2)$ and $w_2 \in NC_2(2m-2k)$. So if

$$H_{2m,k}(x) := \sum_{w_1 \in NC_2(2k-2)} \sum_{w_2 \in NC_2(2m-2k)} h_{aw_1aw_2}(x)\,,$$

then

$$H_{2m}(x) = \sum_{k=1}^{m} H_{2m,k}(x)\,.$$

Notice that

$$\begin{aligned}
H_{2m,k}(x) &= \sum_{w_1 \in NC_2(2k-2)} h_{aw_1a}(x) \sum_{w_2 \in NC_2(2m-2k)} h_{w_2}(x) \\
&= \sum_{w_1 \in NC_2(2k-2)} \int_0^1 \left[f(x,y)h_{w_1}(y)\right] dy \sum_{w_2 \in NC_2(2m-2k)} h_{w_2}(x) \\
&= \int_0^1 \left[f(x,y)H_{2(k-1)}(y)H_{2(m-k)}(x)\right] dy\,,
\end{aligned}$$

the equalities in the first two lines following from standard tricks with Catalan words. From here, an easy computation completes the proof. One can refer to the `arXiv` version for the complete details [14]. □

5. Special cases and examples

In this section, we attempt to give a better description of the probability measure μ, which appears as the LSD in Theorems 2.1 and 3.1, in some special cases. As in Section 4, $R(\cdot,\cdot)$ are the correlations of a weakly stationary mean zero variance one process $(Y_{ij} : i,j \in \mathbb{Z})$. As always, (2.2) and (2.3) are assumed, and μ is the unique even probability measure whose $2m^{\text{th}}$ moment equals β_{2m} which is as defined in (2.7). The first main result of this section is the following.

Theorem 5.1. *Assume that*

$$R(u,v) = R(u,0)R(0,v)\,, \qquad u,v \in \mathbb{Z}\,. \tag{5.1}$$

Then, the function $r(\cdot)$ *defined on* $[-\pi,\pi]$ *by*

$$r(x) := \sum_{k=-\infty}^{\infty} R(k,0)e^{-ikx}\,, \qquad -\pi \le x \le \pi\,,$$

is a well-defined function, that is the sum on the right-hand side converges absolutely, and its range is a compact subset of $[0,\infty)$. *Furthermore,*

$$\mu = \mu_r \boxtimes \mu_s\,,$$

where μ_r *denotes the law of* $r(U)$, U *being a Uniform* $(-\pi,\pi)$ *random variable,* μ_s *denotes the WSL whose density is given by*

$$\mu_s(dx) := \frac{\sqrt{4-x^2}}{2\pi}\mathbf{1}(|x|\le 2)dx\,, \tag{5.2}$$

and " $\boxtimes$*" denotes the free product convolution.*

Remark 2. In order that $\mu_r \boxtimes \mu_s$ be defined, it is required that both μ_r and μ_s are compactly supported, and the support of at least one of them is a subset of the positive half line. Hence, in the above result, the claim that the range of $r(\cdot)$ is a compact subset of $[0,\infty)$ is needed. For this and other results on the free product convolution, the reader is referred to Chapter 14 of [16].

The next result is the other main result of this section.

Theorem 5.2. *If*

$$R(k,0) = 0 \quad \text{for all} \quad k \neq 0\,, \tag{5.3}$$

then $\mu = \mu_s$, *where* μ_s *is the WSL as defined in (5.2).*

Now, we shall see the relevance of the two main results proved above in the light of Theorem 3.1.

5.1. Corollary of Theorem 5.1

Let $\{\epsilon_{i,j} : i, j \in \mathbb{Z}\}$ be as in Section 3; in particular, the Pastur condition (3.1) holds. Let $\{c_k : k \in \mathbb{Z}\}$ be a sequence of real numbers such that

$$\sum_{k=-\infty}^{\infty} |c_k| < \infty \tag{5.4}$$

and

$$\sum_{k=-\infty}^{\infty} c_k^2 = 1\,. \tag{5.5}$$

Set

$$c_{k,l} := c_k c_l\,, \qquad k, l \in \mathbb{Z}\,.$$

Define $Z_{i,j}$ and $R(\cdot, \cdot)$ by (3.4) and (3.5) respectively. Clearly, (3.2) and (3.3) hold, and the process $(Z_{i,j} : i, j \in \mathbb{Z})$ is weakly stationary with mean zero and variance one. Also,

$$\begin{aligned} R(u,v) &= \left(\sum_k c_k c_{k-u}\right)\left(\sum_l c_l c_{l+v}\right) \\ &= \left(\sum_k \sum_{k'} c_k c_{k-u} c_{k'}^2\right)\left(\sum_l \sum_{l'} c_l c_{l+v} c_{l'}^2\right) \\ &= R(u,0)R(0,v)\,, \end{aligned}$$

the second equality following from (5.5). Let A_n and μ_n be as in (1.1) and (1.2) respectively, that is, the former is the $n \times n$ matrix whose $(i,j)^{\text{th}}$ entry is $Z_{i\wedge j, i\vee j}$, and the latter is the ESD of $A_n/\sqrt{n}$. Let μ_r and μ_s be as in the statement of Theorem 5.1. Then, as a corollary of the result mentioned above and Theorem 3.1, it follows that, μ_n converges weakly in probability to $\mu_r \boxtimes \mu_s$.

5.2. Corollary of Theorem 5.2

Once again, let $\{\epsilon_{i,j} : i, j \in \mathbb{Z}\}$ be as in Section 3 satisfying (3.1). Assume that $\{c_{k,l} : k, l \in \mathbb{Z}\} \subset \mathbb{R}$ is such that (3.2) and (3.3) hold, and furthermore

$$\sum_{l=-\infty}^{\infty} c_{k,l} c_{k',l} \;=\; \mathbf{1}\left(k = k'\right) \qquad \text{for all} \qquad k, k' \in \mathbb{Z}\,. \tag{5.6}$$

As before, let $Z_{i,j}$, A_n and μ_n be as in (3.4), (1.1) and (1.2) respectively. It is easy to see that the conditions imposed above ensure that $(Z_{i,j} : i, j \in \mathbb{Z})$

is a mean zero variance one weakly stationary process, and that (5.3) holds. Then by Theorem 3.1 and Theorem 5.2, it follows that μ_n converges weakly in probability to μ_s which is the WSL defined in (5.2).

We end this section by revisiting Examples 1 to 4 mentioned in Section 1.

Example 1. To start with, one needs to argue the existence of a stationary centred Gaussian process $\{Z_{i,j} : i, j \in \mathbb{Z}\}$ satisfying

$$E[Z_{0,0}Z_{u,v}] = \rho^{|u|+|v|}\,, \qquad u, v \in \mathbb{Z}\,.$$

That, however, is obvious from the observation that

$$\rho^{|u|+|v|} = \int\limits_{(-\pi,\pi]^2} e^{i(ux+vy)} F(dx)F(dy)\,, \qquad u, v \in \mathbb{Z}\,,$$

where F is the spectral measure of the autocovariance function $(\rho^{|h|} : h \in \mathbb{Z})$; see Herglotz theorem (Theorem 4.3.1 in Brockwell and Davis [17]). By Theorem 5.1 and results about the AR(1) process, it follows that μ_n converges in probability to $\mu_r \boxtimes \mu_s$, where μ_r is the law of $\frac{1-\rho^2}{1-2\rho\cos U+\rho^2}$, U being an Uniform $(-\pi, \pi)$ random variable.

Example 2. Notice that the $(i, j)^{\text{th}}$ entry of A_n is given by

$$(N + 1) \sum_{k,l\in\mathbb{Z}} c_k c_l G_{i-k,j-l} =: (N + 1)Y_{i,j}\,,$$

where $c_k := (N + 1)^{-1/2}\mathbf{1}(-N \le k \le 0)$. Then (5.4) and (5.5) hold, and therefore, the ESD of $((Y_{i,j}/\sqrt{n}))_{n\times n}$ converges to $\mu_r \boxtimes \mu_s$, where μ_r is the law of

$$1 + 2(N + 1)^{-2} \sum_{k=1}^{N} (N - k + 1)^2 \cos(kU)\,,$$

U being distributed as Uniform $(-\pi, \pi)$. Hence, the LSD of $A_n/\sqrt{n}$ is the free product convolution of μ_s with the law of

$$N + 1 + 2(N + 1)^{-1} \sum_{k=1}^{N} (N - k + 1)^2 \cos(kU)\,.$$

Example 3. By Theorem 5.2, it follows that under the additional assumption that $\sum_{n=1}^{\infty} |E(G_0G_n)| < \infty$, the LSD of $A_n/\sqrt{n}$ is μ_s.

Example 4. Setting

$$\sigma := \left(\sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} c_{k,l}^2 \right)^{1/2}\,,$$

it is easy to see from the Corollary of Theorem 5.2 that the LSD of $\sigma^{-1}A_n/\sqrt{n}$ is μ_s. Therefore, the LSD of $A_n/\sqrt{n}$ is $\tilde{\mu}_s$ given by

$$\tilde{\mu}_s(dx) := \frac{\sqrt{4 - x^2/\sigma^2}}{2\pi\sigma}\mathbf{1}(|x| \leq 2\sigma)dx\,,$$

which is a dilation of the WSL.

The authors are grateful to Parthanil Roy for helpful discussions.

REFERENCES

[1] E.P. Wigner, *Ann. Math.* **67**, 325 (1958).

[2] G. Ben Arous, A. Guionnet, *Wigner Matrices*, in: *The Oxford Handbook of Random Matrix Theory*, pp. 433–451, Oxford Univ. Press, Oxford 2011.

[3] F. Götze, A.N. Tikhomirov, *Limit Theorems for Spectra of Random Matrices with Martingale Structure*, in: *Stein's Method and Applications*, Vol. 5 of Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., pp. 181–193, Singapore Univ. Press, Singapore 2005.

[4] S. Chatterjee, *Ann. Probab.* **34**, 2061 (2006).

[5] R. Rashidi Far, T. Oraby, W. Bryc, R. Speicher, *IEEE Trans. Inform. Theory* **54**, 544 (2008).

[6] K. Hofmann-Credner, M. Stolz, *Electron. Commun. Probab.* **13**, 401 (2008).

[7] W. Hachem, P. Loubaton, J. Najim, *Markov Process. Related Fields* **11**, 629 (2005).

[8] R. Adamczak, *Electron. J. Probab.* **16**, 1065 (2011).

[9] O. Pfaffel, E. Schlemm, *Linear Algebra Appl.* **436**, 2966 (2012).

[10] G.W. Anderson, O. Zeitouni, *Commun. Pure Appl. Math.* **61**, 1118 (2008).

[11] S. Chatterjee, `arXiv:math/0508213 [math.PR]`.

[12] R. Speicher, *Mem. Amer. Math. Soc.* **132**, 1 (1998).

[13] R. Speicher, *Free Probability Theory*, in: *The Oxford Handbook of Random Matrix Theory*, pp. 452–470, Oxford Univ. Press, Oxford 2011.

[14] A. Chakrabarty, R.S. Hazra, D. Sarkar, `arXiv:1304.3394v2 [math.PR]`.

[15] Z. Bai, J.W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Springer Series in Statistics, New York 2010, second edition.

[16] A. Nica, R. Speicher, *Lectures on the Combinatorics of Free Probability*, Cambridge University Press, New York 2006.

[17] P.J. Brockwell, R.A. Davis, *Time Series: Theory and Methods*, Springer-Verlag, New York 1991, second edition.