Markov Processes Relat. Fields **11**, 629–648 (2005)

Markov Processes and Related Fields

The Empirical Eigenvalue Distribution of a Gram Matrix: from Independence to Stationarity*

W. Hachem[1] **, P. Loubaton**[2] **and J. Najim**[3]

[1] Supélec (Ecole Supérieure d'Electricité), Plateau de Moulon, 3, rue Joliot-Curie, 91192 Gif Sur Yvette Cedex, France. E-mail: walid.hachem@supelec.fr
[2] IGM LabInfo, UMR 8049, Institut Gaspard Monge, Université de Marne La Vallée, 5, Bd Descartes, Champs sur Marne, 77454 Marne La Vallée Cedex 2, France. E-mail: loubaton@univ-mlv.fr
[3] CNRS, Télécom Paris 46, rue Barrault, 75013 Paris, France. E-mail: najim@tsi.enst.fr

Received May 4, 2005, revised November 15, 2005

Abstract. Consider a $N \times n$ matrix $Z_n = (Z^n_{j_1 j_2})$ where the individual entries are a realization of a properly rescaled stationary Gaussian random field:

$$Z^n_{j_1 j_2} = \frac{1}{\sqrt{n}} \sum_{(k_1,k_2)\in \mathbb{Z}^2} h(k_1,k_2)\, U(j_1-k_1, j_2-k_2),$$

where $h \in \ell^1(\mathbf{Z}^2)$ is a deterministic complex summable sequence and $(U(j_1,j_2); (j_1,j_2)\in \mathbf{Z}^2)$ is a sequence of independent complex Gaussian random variables with mean zero and unit variance.

The purpose of this article is to study the limiting empirical distribution of the eigenvalues of Gram random matrices $Z_n Z_n^*$ and $(Z_n+A_n)(Z_n+A_n)^*$ where A_n is a deterministic matrix with appropriate assumptions in the case where $n \to \infty$ and $N/n \to c \in (0,\infty)$.

The proof relies on related results for matrices with independent but not identically distributed entries and substantially differs from related works in the literature (Boutet de Monvel et al. [4], Girko [9], etc.).

Keywords: random matrix, empirical eigenvalue distribution, Stieltjes transform

AMS Subject Classification: Primary 15A52, Secondary 15A18, 60F15

*This work was partially supported by the Fonds National de la Science (France) via the ACI program "Nouvelles Interfaces des Mathématiques", project MALCOM n° 205

1. Introduction

The model

Let $Z_n = (Z^n_{j_1 j_2}, 0 \le j_1 < N,\ 0 \le j_2 < n)$ be a $N \times n$ random matrix with entries

$$Z^n_{j_1 j_2} = \frac{1}{\sqrt{n}} \sum_{(k_1,k_2)\in \mathbb{Z}^2} h(k_1,k_2)\, U(j_1 - k_1, j_2 - k_2),$$

where $(U(j_1,j_2),\ (j_1,j_2) \in \mathbf{Z}^2)$ is a sequence of independent complex Gaussian random variables (r.v.) such that $\mathsf{E}\, U(j_1,j_2) = 0$, $\mathsf{E}\, U(j_1,j_2)^2 = 0$ and $\mathsf{E}\, |U(j_1,j_2)|^2 = 1$, and $(h(k_1,k_2), (k_1,k_2) \in \mathbf{Z}^2)$ is a deterministic complex sequence satisfying

$$\sum_{(k_1,k_2)\in \mathbb{Z}^2} |h(k_1,k_2)| < \infty.$$

The bidimensional process $Z^n_{j_1 j_2}$ is a stationary Gaussian field. Indeed,

$$\mathrm{cov}(Z^n_{j_1 j_2}, Z^n_{j_1' j_2'}) = n^{-1} C(j_1 - j_1', j_2 - j_2')$$

where

$$C(j_1,j_2) = \sum_{(k_1,k_2)\in \mathbb{Z}^2} h(k_1,k_2) h^*(k_1 - j_1, k_2 - j_2) \tag{1.1}$$

(we denote by a^* the complex conjugate of $a \in \mathbf{C}$ — we also denote by A^* the hermitian adjoint of matrix A).

The main results

The purpose of this article is to establish the convergence of the empirical distribution of the eigenvalues of various Gram matrices based on Z_n in the large limit $n \to \infty$, $N/n \to c \in (0,\infty)$. More precisely, we shall study the convergence of the spectral distribution of $Z_n Z_n^*$ and $(Z_n + A_n)(Z_n + A_n)^*$ where A_n is a deterministic matrix with a given structure. In particular, if Z_n is square, we take A_n to be Toeplitz. The contribution of this article is to provide a new method to study Gram matrices based on Gaussian fields. The main idea is to approximate the matrix Z_n by a matrix $\tilde{Z}_n$ unitarily congruent to a matrix with independent but not identically distributed entries. This method will allow us to revisit the centered case $Z_n Z_n^*$, already studied by Boutet de Monvel et al. in [4] and to establish the limiting spectral distribution of the non-centered case $(Z_n + A_n)(Z_n + A_n)^*$ for some deterministic matrix A_n.

Motivations

The motivations for such a work are twofold. First of all, we believe that this line of proof is new. Let us briefly describe the three main elements of it.

The first one is a periodization scheme popular in signal processing and described as follows:

$$\tilde{Z}_n = (\tilde{Z}^n_{j_1 j_2})$$

where

$$\tilde{Z}^n_{j_1 j_2} = \frac{1}{\sqrt{n}} \sum_{(k_1,k_2)\in \mathbb{Z}^2} h(k_1,k_2)\, U\big((j_1-k_1) \bmod N, (j_2-k_2) \bmod n\big),$$

where mod denotes modulo.

The second element is an inequality due to Bai [2] involving the Lévy distance $\mathcal{L}$ between distribution functions:

$$\mathcal{L}^4(F^{AA^*}, F^{BB^*}) \le \frac{2}{N^2} \operatorname{Tr}(A-B)(A-B)^* \operatorname{Tr}(AA^* + BB^*), \tag{1.2}$$

where F^{AA^*} denotes the empirical distribution function of the eigenvalues of the matrix AA^* and $\operatorname{Tr}(X)$ denotes the trace of matrix X. With the help of this inequality, we shall prove that $Z_n Z_n^*$ and $\tilde{Z}_n \tilde{Z}_n^*$ have the same limiting spectral distribution.

The third element comes from the advantage of considering $\tilde{Z}_n$. In fact, $\tilde{Z}_n$ is congruent (via Fourier unitary transforms) to a random matrix with independent but not identically distributed entries. Therefore, we can (and will) rely on results established in [11] for Gram matrices with independent but not identically distributed entries.

The second motivation comes from the field of wireless communications. In a communication system employing antenna arrays at the transmitter and at the receiver sides, random matrices extracted from Gaussian fields are often good models for representing the radio communication channel. In this course, the stationary model as considered above is often a realistic channel model. The computations of popular receiver performance indexes such as Signal to Interference plus Noise Ratio or Shannon channel capacity heavily rely on the knowledge of the limiting spectral distribution of matrices of the type $Z_n Z_n^*$ (see [6, 13] and also the tutorial [16] for further references).

About the literature

Various Gram matrices based on Gaussian fields have already been studied in the literature. The study of the general case $(Z_n + A_n)(Z_n + A_n)^*$ has been undertaken by Girko in [9, 10]. His approach is based on more general results valid in the case of a Gram matrix with asymptotically independent entries. In this context, Girko shows that the normalized trace of its resolvent has the same asymptotic behavior as the normalized trace of a deterministic matrix verifying a certain non-linear "canonical equation". Since no assumptions are done on the structure of A_n, there might not be any limiting spectral distribution. In the case where Z_n is a stationary field and A_n is Toeplitz, the equations have a simpler form, and depend on the spectral measure of Z_n and on the Fourier transform of the entries of A_n. Note that the Gaussianity is not necessary in this approach.

Boutet de Monvel et al. [4] have also studied Gram matrices based on stationary Gaussian fields in the case where the matrix has the form $V_n + Z_n Z_n^*$, V_n being a deterministic Toeplitz matrix. Their line of proof is based on a direct study of the resolvent, taking advantage of the Gaussianity of the entries.

Disclaimer

In this paper, we study in detail the case where the entries of matrix Z_n are complex. In the real case, the general framework of the proof works as well if one considers the real counterpart of the Fourier unitary transforms, however the computations are more involved. We provide some details in Section 5.

2. Assumptions and useful results

2.1. Notation, assumptions, Stieltjes transforms and Stieltjes kernels

Let $N = N(n)$ be a sequence of integers such that

$$\lim_{n\to\infty} \frac{N(n)}{n} = c \in (0, \infty).$$

We denote by $\mathbf{i}$ the complex number $\sqrt{-1}$, by $\mathbf{1}_A(x)$ the indicator function over set A and by $\delta_{x_0}(x)$ the Dirac measure at point x_0. A sum will be equivalently written as $\sum_{k=1}^{n}$ or $\sum_{k=1:n}$. We denote by $\mathcal{CN}(0,1)$ the distribution of the Gaussian complex random variable U satisfying $\mathsf{E}\,U = 0$, $\mathsf{E}\,U^2 = 0$, and $\mathsf{E}\,|U|^2 = 1$ (equivalently, $U = A + \mathbf{i}B$ where A and B are real independent Gaussian r.v.'s with mean 0 and standard deviation $1/\sqrt{2}$ each).

Assumption A-1. The entries $(Z^n_{j_1 j_2},\ 0 \le j_1 < N,\ 0 \le j_2 < n,\ n \ge 1)$ of the $N \times n$ matrix Z_n are random variables defined as:

$$Z^n_{j_1 j_2} = \frac{1}{\sqrt{n}} \sum_{(k_1,k_2)\in \mathbf{Z}^2} h(k_1,k_2)\, U(j_1 - k_1, j_2 - k_2),$$

where $(h(k_1,k_2),\ (k_1,k_2) \in \mathbf{Z}^2)$ is a deterministic complex sequence satisfying

$$h_{\max} \stackrel{\triangle}{=} \sum_{(k_1,k_2)\in \mathbf{Z}^2} |h(k_1,k_2)| < \infty$$

and $(U(j_1,j_2), (j_1,j_2) \in \mathbf{Z}^2)$ is a sequence of independent random variables with distribution $\mathcal{CN}(0,1)$.

Remark 2.1. Assumption (A-1) is a bit more restrictive than the related assumption [4], which only relies on the summability of the covariance function of the stationary process.

For every matrix A, we denote by $F^{A\,A^*}$ the empirical distribution function of the eigenvalues of $A\,A^*$. Since we will study at the same time the limiting spectrum of the matrices $Z_n Z_n^*$ (resp. $(Z_n + A_n)(Z_n + A_n)^*$) and $Z_n^* Z_n$ (resp. $(Z_n + A_n)^*(Z_n + A_n)$), we can assume without loss of generality that $c \leq 1$. We also assume for simplicity that $N \leq n$.

When dealing with vectors, the norm $\|\cdot\|$ will denote the Euclidean norm. In the case of matrices, the norm $\|\cdot\|$ will refer to the spectral norm. Denote by $\mathbf{C}^+$ the set $\mathbf{C}^+ = \{z \in \mathbf{C},\ \mathrm{Im}(z) > 0\}$ and by $C(\mathcal{X})$ the set of bounded continuous functions over a given topological space $\mathcal{X}$ endowed with the supremum norm $\|\cdot\|_\infty$.

Let μ be a probability measure over $\mathbf{R}$. Its Stieltjes transform f is defined by

$$f(z) = \int_{\mathrm{R}} \frac{\mu(d\lambda)}{\lambda - z}, \quad z \in \mathbf{C}^+.$$

We list below the main properties of the Stieltjes transforms that will be needed in the sequel.

Proposition 2.1. *The following properties hold true:*

(1) *Let f be the Stieltjes transform of μ, then*

- *the function f is analytic over $\mathbf{C}^+$,*
- *the function f satisfies: $|f(z)| \leq 1/\mathrm{Im}(z)$,*
- *if $z \in \mathbf{C}^+$, then $f(z) \in \mathbf{C}^+$,*
- *if $\mu(-\infty, 0) = 0$, then $z \in \mathbf{C}^+$ implies $z\,f(z) \in \mathbf{C}^+$.*

(2) *Conversely, let f be a function analytic over $\mathbf{C}^+$ such that $f(z) \in \mathbf{C}^+$ if $z \in \mathbf{C}^+$ and $|f(z)|\,|\mathrm{Im}(z)|$ bounded on $\mathbf{C}^+$. If $\lim_{y \to +\infty} -\mathbf{i}y\,f(\mathbf{i}y) = 1$, then f is the Stieltjes transform of a probability measure μ and the following inversion formula holds:*

$$\mu([a,b]) = \lim_{\eta \to 0^+} \frac{1}{\pi} \int_a^b \mathrm{Im}\, f(\xi + \mathbf{i}\eta)\, d\xi,$$

where a and b are continuity points of μ. If moreover $zf(z) \in \mathbf{C}^+$ if $z \in \mathbf{C}^+$ then, $\mu(\mathbf{R}^-) = 0$.

(3) *Let P_n and P be probability measures over $\mathbf{R}$ and denote by f_n and f their Stieltjes transforms. Then*

$$\Big(\forall z \in \mathbf{C}^+,\ f_n(z) \xrightarrow[n\to\infty]{} f(z)\Big) \Longrightarrow \mathsf{P}_n \xrightarrow[n\to\infty]{\mathcal{D}} \mathsf{P}\,.$$

Denote by $\mathcal{M}_{\mathbf{C}}(\mathcal{X})$ the set of complex measures over the topological set $\mathcal{X}$. In the sequel, we will call Stieltjes kernel every application

$$\pi : \mathbf{C}^+ \to \mathcal{M}_{\mathbf{C}}(\mathcal{X})$$

either denoted $\pi(z, dx)$ or $\pi_z(dx)$ and satisfying:

(1) $\forall z \in \mathbf{C}^+,\ \forall g \in C(\mathcal{X}),\ \left| \int g\, d\pi_z \right| \le \|g\|_\infty / \mathrm{Im}(z)$.

(2) $\forall g \in C(\mathcal{X})$, $\int g\, d\pi_z$ is analytic over $\mathbf{C}^+$,

(3) $\forall z \in \mathbf{C}^+$, $\forall g \in C(\mathcal{X})$ and $g \ge 0$, $\mathrm{Im}\big(\int g\, d\pi_z \big) \ge 0$,

(4) $\forall z \in \mathbf{C}^+$, $\forall g \in C(\mathcal{X})$ and $g \ge 0$, $\mathrm{Im}\big(z \int g\, d\pi_z \big) \ge 0$.

2.2. A quick review of the results for matrices with independent entries

In order to establish the convergence of the empirical distribution of the eigenvalues, we will rely on the results based on matrices with independent but not identically distributed entries. Let us recall here those of interest (the assumptions and the statements are based on [11]).

Remark 2.2. The Wigner case (Hermitian matrix with independent but not identically distributed entries) is also of interest since one can relate the eigenvalues of ZZ^* to the eigenvalues of the Wigner matrix $\left(\begin{smallmatrix} 0 & Z^* \\ Z & 0 \end{smallmatrix} \right)$. This case has been studied by Casati and Girko [7], Shlyakhtenko [14,15], Anderson and Zeitouni [1] among others.

Consider a $N \times n$ random matrix Y_n where the entries are given by

$$Y^n_{j_1 j_2} = \frac{\Phi(j_1/N, j_2/n)}{\sqrt{n}} X^n_{j_1 j_2}$$

where $X^n_{j_1 j_2}$ and Φ are defined below.

Assumption A-2. The complex random variables

$$(X^n_{j_1 j_2};\, 0 \le j_1 < N,\, 0 \le j_2 < n,\, n \ge 1)$$

are independent and identically distributed (i.i.d.). They are centered with $\mathsf{E}\, |X^n_{j_1 j_2}|^2 = 1$ and there exists $\varepsilon > 0$ such that $\mathsf{E}\, |X^n_{j_1 j_2}|^{4+\varepsilon} < \infty$.

Assumption A-3. The function $\Phi : [0,1] \times [0,1] \to \mathbf{C}$ is such that $|\Phi|^2$ is continuous and therefore there exists a non-negative constant $\Phi_{\max}$ such that

$$\forall (t_1, t_2) \in [0,1]^2, \quad 0 \le |\Phi(t_1, t_2)|^2 \le \Phi^2_{\max} < \infty. \tag{2.1}$$

Theorem 2.1 (Independent entries, the centered case [8]). *If* (A-2) *and* (A-3) *hold and* $n \to \infty$, *then the empirical distribution of the eigenvalues of the matrix* $Y_n Y_n^*$ *converges a.s. to a non-random probability measure* μ *whose Stieltjes transform* f *is given by* $f(z) = \int_{[0,1]} \pi_z(dx)$, *where* π_z *is the unique Stieltjes kernel with support included in* $[0,1]$ *and satisfying for all* $g \in C([0,1])$,

$$\int g\, d\pi_z = \int_0^1 \frac{g(u)}{-z + \int_0^1 \left(|\Phi|^2(u,t) \big/ \left[1 + c \int_0^1 |\Phi|^2(x,t)\, \pi_z(dx)\right]\right) dt}\, du. \tag{2.2}$$

If one adds a deterministic pseudo-diagonal matrix Λ_n to the matrix Y_n, the limiting equation is modified and in fact becomes a system of equations.

Assumption A-4. Let $\Lambda_n = (\Lambda_{ij}^n)$ be a complex deterministic $N \times n$ matrix whose non-diagonal entries are zero. We assume moreover that there exists a probability measure $H(\,du, d\lambda)$ over the set $[0,1] \times \mathbf{R}$ with compact support $\mathcal{H}$ such that

$$\frac{1}{N} \sum_{i=1}^N \delta_{(i/N,\, |\Lambda_{ii}^n|^2)}(du, d\lambda) \xrightarrow[n \to \infty]{\mathcal{D}} H(\,du, d\lambda). \tag{2.3}$$

Denote by $\mathcal{H}_c$ the support of the image of probability measure H under the application $(u,\lambda) \to (cu, \lambda)$ and by $\mathcal{R}$ the support of the measure $\mathbf{1}_{[c,1]}(du) \otimes \delta_0(d\lambda)$ where $\otimes$ denotes the product of measures. The set $\tilde{\mathcal{H}} = \mathcal{H}_c \cup \mathcal{R}$ will be of importance in the sequel (see also Remarks 2.4 and 2.5 in [11] for more information).

Theorem 2.2 (Independent entries, the non-centered case [11]). *Assume that* (A-2), (A-3) *and* (A-4) *hold and let* $n \to \infty$. *Then the empirical distributions of the eigenvalues of matrices* $(Y_n + \Lambda_n)(Y_n + \Lambda_n)^*$ *and* $(Y_n + \Lambda_n)^*(Y_n + \Lambda_n)$ *converge a.s. to non-random probability measures* μ *and* $\tilde{\mu}$ *whose Stieltjes transforms* f *and* $\tilde{f}$ *are given by*

$$f(z) = \int_{\mathcal{H}} \pi_z(dx) \quad \text{and} \quad \tilde{f}(z) = \int_{\tilde{\mathcal{H}}} \tilde{\pi}_z(dx)$$

where π_z *and* $\tilde{\pi}_z$ *are the unique Stieltjes kernels with supports included in* $\mathcal{H}$ *and* $\tilde{\mathcal{H}}$ *and satisfying*

$$\int g\, d\pi_z = \int \frac{g(u,\lambda)}{-z(1 + \int |\Phi|^2(u,t) \tilde{\pi}(z, dt, d\zeta)) + \lambda / [1 + c \int |\Phi|^2(t, cu) \pi(z, dt, d\zeta)]} \times H(du, d\lambda), \tag{2.4}$$

$$\int g\, d\tilde{\pi}_z = c\int \frac{g(cu,\lambda)}{-z(1+c\int |\Phi|^2(t,cu)\pi(z,dt,d\zeta))+\lambda/[1+\int |\Phi|^2(u,t)\tilde{\pi}(z,dt,d\zeta)]}$$
$$\times H(du,d\lambda) + (1-c)\int_c^1 \frac{g(u,0)}{-z(1+c\int |\Phi|^2(t,u)\pi(z,dt,d\zeta))}\, du \quad (2.5)$$

where (2.4) *and* (2.5) *hold for every* $g \in C(\mathcal{H})$.

3. The limiting distribution in the centered stationary case

We first introduce the following complex-valued function $\Phi : [0,1]\times[0,1] \to \mathbf{C}$ defined by

$$\Phi(t_1,t_2) = \sum_{(\ell_1,\ell_2)\in \mathrm{Z}^2} h(\ell_1,\ell_2) e^{2\pi \mathbf{i}(\ell_1 t_1 - \ell_2 t_2)}. \quad (3.1)$$

We also introduce the $p\times p$ Fourier matrix $F_p = (F^p_{j_1,j_2})_{0\le j_1,j_2<p}$ defined by

$$F^p_{j_1,j_2} = \frac{1}{\sqrt{p}} \exp\Big(2\mathbf{i}\pi\Big(\frac{j_1 j_2}{p}\Big)\Big). \quad (3.2)$$

Note that matrix F_p is a unitary matrix.

Theorem 3.1 (Stationary entries, the centered case [4, 9]). *Let Z_n be a $N\times n$ matrix satisfying* (A-1) *and let $n\to\infty$. Then the empirical distribution of the eigenvalues of the matrix $Z_n Z_n^*$ converges in probability to the non-random probability measure μ defined in Theorem* 2.1.

3.1. Proof of Theorem 3.1

Recall that

$$Z^n_{j_1 j_2} = \frac{1}{\sqrt{n}} \sum_{(k_1,k_2)\in \mathrm{Z}^2} h(k_1,k_2)\, U(j_1-k_1, j_2-k_2).$$

We introduce the $N\times n$ matrix $\tilde{Z}_n$ whose entries are defined by

$$\tilde{Z}^n_{j_1 j_2} = \frac{1}{\sqrt{n}} \sum_{(k_1,k_2)\in \mathrm{Z}^2} h(k_1,k_2)\, U(j_1-k_1 \bmod N, j_2-k_2 \bmod n).$$

For simplicity, we shall write $\tilde{U}^n(j_1,j_2)$ instead of $U(j_1 \bmod N,\, j_2 \bmod n)$. Recall that $\mathcal{L}$ stands for the Lévy distance between distribution functions. The main interest in dealing with matrix $\tilde{Z}_n$ lies in the following two lemmas.

Lemma 3.1. *Consider the $N \times n$ matrix $Y_n = F_N \tilde{Z}_n F_n^*$. Then the entries $Y_{\ell_1\ell_2}^n$ of Y_n can be written as*

$$Y_{\ell_1\ell_2}^n = \frac{1}{\sqrt{n}} \Phi\Big(\frac{\ell_1}{N}, \frac{\ell_2}{n}\Big) X_{\ell_1\ell_2}^n$$

where Φ is defined in (3.1) *and the complex random variables $\{X_{\ell_1\ell_2}^n, 0 \le \ell_1 < N, 0 \le \ell_2 < n\}$ are independent with distribution $\mathcal{CN}(0,1)$.*

Proof of Lemma 3.1. We first compute the individual entries of matrix $Y_n = F_N \tilde{Z}_n F_n^*$:

$$\begin{aligned}
Y_{\ell_1\ell_2}^n &= \sum_{\substack{j_1=0:N-1 \\ j_2=0:n-1}} \frac{\exp\{2\mathbf{i}\pi(j_1\ell_1/N - j_2\ell_2/n)\}}{\sqrt{Nn}} \tilde{Z}_{j_1j_2}^n \\
&= \frac{1}{\sqrt{n}} \sum_{\substack{j_1=0:N-1 \\ j_2=0:n-1}} \frac{\exp\{2\mathbf{i}\pi(j_1\ell_1/N - j_2\ell_2/n)\}}{\sqrt{Nn}} \\
&\quad \times \sum_{(k_1,k_2)\in\mathbb{Z}^2} h(k_1,k_2)\, \tilde{U}^n(j_1-k_1, j_2-k_2) \\
&= \frac{1}{\sqrt{n}} \sum_{\substack{j_1=0:N-1 \\ j_2=0:n-1}} \frac{\exp\{2\mathbf{i}\pi(j_1\ell_1/N - j_2\ell_2/n)\}}{\sqrt{Nn}} \sum_{\substack{m_1=0:N-1 \\ m_2=0:n-1}} U(m_1,m_2) \\
&\quad \times \sum_{(k_1,k_2)\in\mathbb{Z}^2} h(j_1-m_1+k_1N, j_2-m_2+k_2n) \\
&= \frac{1}{\sqrt{n}} \Phi\Big(\frac{\ell_1}{N}, \frac{\ell_2}{n}\Big) \sum_{\substack{m_1=0:N-1 \\ m_2=0:n-1}} U(m_1,m_2) \frac{\exp\{2\mathbf{i}\pi(m_1\ell_1/N - m_2\ell_2/n)\}}{\sqrt{Nn}}.
\end{aligned}$$

Let $X_{\ell_1\ell_2}^n$ be the random variable defined as

$$X_{\ell_1\ell_2}^n = \sum_{\substack{m_1=0:N-1 \\ m_2=0:n-1}} U(m_1,m_2) \frac{\exp\{2\mathbf{i}\pi(m_1\ell_1/N - m_2\ell_2/n)\}}{\sqrt{Nn}}$$

for $0 \le \ell_1 \le N-1$ and $0 \le \ell_2 \le n-1$. Denoting by X_n and U_n the $N \times n$ matrices with entries $X_{\ell_1\ell_2}^n$ and $U(\ell_1,\ell_2)$ respectively, we then have $X_n = F_N U_n F_n^*$. Define $\mathrm{vec}(A)$ to be the vector obtained by stacking the columns of matrix A. Then the $Nn \times 1$ vectors $\mathbf{X} = \mathrm{vec}(X_n)$ and $\mathbf{U} = \mathrm{vec}(U_n)$ are related by the equation $\mathbf{X} = (F_n^* \otimes F_N)\mathbf{U}$ (Lemma 4.3.1 in [12]), where $\otimes$ denotes the Kronecker product of matrices. The vector $\mathbf{X}$ is a complex Gaussian random vector that satisfies

$$\mathsf{E}\,\mathbf{X} = (F_n^* \otimes F_N)\, \mathsf{E}\,\mathbf{U} = 0$$

and
$$\mathsf{E}\,\mathbf{X}\mathbf{X}^T = (F_n^* \otimes F_N)\ \mathsf{E}\,\mathbf{U}\mathbf{U}^T (F_n^* \otimes F_N) = 0.$$
After noticing that the matrix $(F_n^* \otimes F_N)$ is unitary, we furthermore have
$$\mathsf{E}\,\mathbf{X}\mathbf{X}^* = (F_n^* \otimes F_N)\ \mathsf{E}\,\mathbf{U}\mathbf{U}^* (F_n^* \otimes F_N)^* = I_{nN}$$
where I_p is the $p\times p$ identity matrix. In short, the entries of X_n are independent and have the distribution $\mathcal{CN}(0,1)$. Lemma 3.1 is proved. □

Lemma 3.2. *Let B_n be a $N \times n$ deterministic matrix such that the sequence $(1/n)\mathrm{Tr} B_n B_n^*$ is bounded over n. Then*
$$\mathcal{L}\big(F^{(Z_n+B_n)(Z_n+B_n)^*}, F^{(\tilde{Z}_n+B_n)(\tilde{Z}_n+B_n)^*}\big) \xrightarrow[n\to\infty]{\mathrm{P}} 0,$$
where $\xrightarrow{\mathrm{P}}$ denotes convergence in probability.

Proof of Lemma 3.2. Bai's inequality (1.2) yields:
$$\begin{aligned}
&\mathcal{L}^4\big(F^{(Z_n+B_n)(Z_n+B_n)^*},\, F^{(\tilde{Z}_n+B_n)(\tilde{Z}_n+B_n)^*}\big) \\
&\quad \le \frac{2}{n^2}\,\mathrm{Tr}(Z_n-\tilde{Z}_n)(Z_n-\tilde{Z}_n)^* \\
&\qquad \times \mathrm{Tr}\big((Z_n+B_n)(Z_n+B_n)^* + (\tilde{Z}_n+B_n)(\tilde{Z}_n+B_n)^*\big).
\end{aligned} \tag{3.3}$$
We introduce the following notation:
$$\begin{aligned}
\alpha_n &= \frac{1}{n}\,\mathrm{Tr}(Z_n-\tilde{Z}_n)(Z_n-\tilde{Z}_n)^*, \\
\beta_n &= \frac{1}{n}\,\mathrm{Tr}(Z_n+B_n)(Z_n+B_n)^*, \quad \tilde{\beta}_n = \frac{1}{n}\,\mathrm{Tr}(\tilde{Z}_n+B_n)(\tilde{Z}_n+B_n)^*.
\end{aligned}$$
With this notation, inequality (3.3) becomes:
$$\mathcal{L}^4\big(F^{(Z_n+B_n)(Z_n+B_n)^*}, F^{(\tilde{Z}_n+B_n)(\tilde{Z}_n+B_n)^*}\big) \le 2\alpha_n(\beta_n+\tilde{\beta}_n).$$
In order to prove that
$$\mathcal{L}(F^{(Z_n+B_n)(Z_n+B_n)^*}, F^{(\tilde{Z}_n+B_n)(\tilde{Z}_n+B_n)^*}) \xrightarrow{\mathrm{P}} 0,$$
it is sufficient to prove that $\alpha_n(\beta_n+\tilde{\beta}_n) \xrightarrow{\mathrm{P}} 0$, which follows from $\alpha_n \xrightarrow{\mathrm{P}} 0$ and β_n and $\tilde{\beta}_n$ being tight. Indeed,
$$\begin{aligned}
\mathsf{P}\{\alpha_n(\beta_n+\tilde{\beta}_n)\ge\varepsilon\} &\le \mathsf{P}\{\alpha_n\beta_n \ge \varepsilon/2\} + \mathsf{P}\{\alpha_n\tilde{\beta}_n \ge \varepsilon/2\} \\
&\le \mathsf{P}\left\{\alpha_n \ge \frac{\varepsilon}{2K}\right\} + \mathsf{P}\{\beta_n \ge 2K\} \\
&\quad + \mathsf{P}\left\{\alpha_n \ge \frac{\varepsilon}{2\tilde{K}}\right\} + \mathsf{P}\{\tilde{\beta}_n \ge 2\tilde{K}\}.
\end{aligned}$$

Let us first prove that

$$\alpha_n \xrightarrow{\mathrm{P}} 0. \tag{3.4}$$

Since α_n is non-negative, it is sufficient by Markov's inequality to prove that $\mathsf{E}\,\alpha_n \to 0$.

$$\begin{aligned}
\alpha_n &= \frac{1}{n}\operatorname{Tr}(Z_n - \tilde{Z}_n)(Z_n - \tilde{Z}_n)^* \\
&= \frac{1}{n^2}\sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} \left| Z^n_{j_1,j_2} - \tilde{Z}^n_{j_1,j_2}\right|^2 \\
&= \frac{1}{n^2}\sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} \Big| \sum_{(k_1,k_2)\in \mathrm{Z}^2} h(k_1,k_2)V(j_1-k_1,j_2-k_2)\Big|^2,
\end{aligned}$$

where $V(j_1,j_2)$ stands for $U(j_1,j_2) - \tilde{U}^n(j_1,j_2)$. Thus

$$\begin{aligned}
\mathsf{E}\,\alpha_n = \frac{1}{n^2}\sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} \sum_{\substack{(k_1,k_2)\in \mathrm{Z}^2\\ (k_1',k_2')\in \mathrm{Z}^2}} & h(k_1,k_2)h^*(k_1',k_2') \\
&\times \mathsf{E}\,V(j_1-k_1,j_2-k_2)V^*(j_1-k_1',j_2-k_2').
\end{aligned}$$

Introduce the set $\mathcal{J} = \{0,\cdots,N-1\}\times\{0,\cdots,n-1\}$. Then

$$\begin{aligned}
\mathsf{E}\,V(\ell_1,\ell_2)V^*(\ell_1',\ell_2') = \mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(\ell_1,\ell_2)\,\mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(\ell_1',\ell_2') & \Big(\mathbf{1}_{(\ell_1,\ell_2)}(\ell_1',\ell_2') \\
&+ \sum_{(m_1,m_2)\in \mathrm{Z}^2} \mathbf{1}_{(\ell_1,\ell_2)}(\ell_1'+m_1N,\ell_2'+m_2n)\Big)
\end{aligned}$$

and $\mathsf{E}\,\alpha_n$ becomes $\mathsf{E}\,\alpha_n = \mathsf{E}\,\alpha_{n,1} + \mathsf{E}\,\alpha_{n,2}$ where

$$\begin{aligned}
\mathsf{E}\,\alpha_{n,1} &= \frac{1}{n^2}\sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} \sum_{(k_1,k_2)\in \mathrm{Z}^2} |h(k_1,k_2)|^2\,\mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(j_1-k_1,j_2-k_2), \\
\mathsf{E}\,\alpha_{n,2} &= \frac{1}{n^2}\sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} \sum_{\substack{(k_1,k_2)\in \mathrm{Z}^2\\ (k_1',k_2')\in \mathrm{Z}^2}} h(k_1,k_2)h^*(k_1',k_2')\,\mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(j_1-k_1,j_2-k_2) \\
&\quad \times \mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(j_1-k_1',j_2-k_2') \sum_{(m_1,m_2)\in \mathrm{Z}^2} \mathbf{1}_{(k_1,k_2)}(k_1'+m_1N,k_2'+m_2n).
\end{aligned}$$

Let us first deal with $\mathsf{E}\,\alpha_{n,2}$.

$$\mathsf{E}\,\alpha_{n,2} \le \frac{1}{n^2} \sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} \sum_{(k_1,k_2)\in \mathrm{Z}^2} |h(k_1,k_2)|\, \mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(j_1-k_1, j_2-k_2)$$
$$\times \sum_{(k_1',k_2')\in\mathrm{Z}^2} |h(k_1',k_2')|\, \mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(j_1-k_1', j_2-k_2')$$
$$\times \sum_{(m_1,m_2)\in\mathrm{Z}^2} \mathbf{1}_{(k_1,k_2)}(k_1'+m_1N, k_2'+m_2 n).$$

Since h is summable over $\mathbf{Z}^2$ by (A-1),

$$\sum_{(k_1',k_2')\in\mathrm{Z}^2} |h(k_1',k_2')|\, \mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(j_1-k_1', j_2-k_2') \sum_{(m_1,m_2)\in\mathrm{Z}^2} \mathbf{1}_{(k_1,k_2)}(k_1'+m_1N, k_2'+m_2n)$$

is bounded by $h_{\max}$ and

$$\mathsf{E}\,\alpha_{n,2} \le \frac{h_{\max}}{n^2} \sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} \sum_{(k_1,k_2)\in \mathrm{Z}^2} |h(k_1,k_2)|\, \mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(j_1-k_1, j_2-k_2). \tag{3.5}$$

Since

$$\mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(j_1-k_1, j_2-k_2) = 1 \iff \left\{ \begin{array}{lll} j_1-k_1<0 & \text{or} & j_1-k_1 \ge N, \\ j_2-k_2<0 & \text{or} & j_2-k_2\ge n, \end{array}\right.$$

we get:

$$\sum_{(k_1,k_2)\in\mathrm{Z}^2} |h(k_1,k_2)|\, \mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(j_1-k_1, j_2-k_2)$$
$$= \sum_{\substack{k_1=-\infty:j_1-N;\\ k_2=-\infty:j_2-n}} |h(k_1,k_2)| + \sum_{\substack{k_1=-\infty:j_1-N;\\ k_2=j_2+1:\infty}} |h(k_1,k_2)|$$
$$+ \sum_{\substack{k_1=j_1+1:\infty;\\ k_2=-\infty:j_2-n}} |h(k_1,k_2)| + \sum_{\substack{k_1=j_1+1:\infty;\\ k_2=j_2+1:\infty}} |h(k_1,k_2)|.$$

The change of variables $\left\{ \begin{array}{l} j_1' = N-1-j_1 \\ k_1' = -k_1 \end{array}\right.$ and $\left\{ \begin{array}{l} j_2' = n-1-j_2 \\ k_2' = -k_2 \end{array}\right.$ yields

$$\sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} \sum_{\substack{k_1=-\infty:j_1-N;\\ k_2=-\infty:j_2-n}} |h(k_1,k_2)| = \sum_{\substack{j_1'=0:N-1\\ j_2'=0:n-1}} \sum_{\substack{k_1'=j_1'+1:\infty;\\ k_2'=j_2'+1:\infty}} |h(-k_1',-k_2')|.$$

By performing similar change of variables, one gets:

$$\begin{aligned}
&\sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} \sum_{(k_1,k_2)\in \mathrm{Z}^2} |h(k_1,k_2)|\, \mathbf{1}_{\mathrm{Z}^2-\mathcal{J}}(j_1-k_1, j_2-k_2)\\
&\quad = \sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} \sum_{\substack{k_1=j_1+1:\infty;\\ k_2=j_2+1:\infty}} |h(-k_1,-k_2)| + |h(-k_1,k_2)|\\
&\qquad + |h(k_1,-k_2)| + |h(k_1,k_2)|.
\end{aligned}$$

Let us denote the inner sum in the right-hand side by $S(j_1,j_2)$. In order to check that

$$\frac{1}{n^2} \sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} S(j_1,j_2) \xrightarrow[n\to\infty\,;\, N/n\to c]{} 0, \tag{3.6}$$

we introduce

$$T(j) = \sum_{k_1+k_2\geq j+2} |h(-k_1,-k_2)| + |h(-k_1,k_2)| + |h(k_1,-k_2)| + |h(k_1,k_2)|.$$

Is is straightforward to check that $T(j) \xrightarrow[j\to\infty]{} 0$ and that $S(j_1,j_2) \leq T(j_1+j_2)$. We prove (3.6) by a Césaro-like argument: Let n_0 be such that $T(n_0+1) \leq \varepsilon$ and take $N \geq n_0$. We have

$$\frac{1}{n^2} \sum_{\substack{j_1=0:N-1\\ j_2=0:n-1}} S(j_1,j_2) = \frac{1}{n^2} \sum_{0\leq j_1+j_2\leq n_0} S(j_1,j_2) + \frac{1}{n^2} \sum_{\substack{n_0+1\leq j_1+j_2;\\ j_1\leq N-1,\, j_2\leq n-1}} S(j_1,j_2). \tag{3.7}$$

If n is large enough, then the first part of the right-hand side of (3.7) is smaller than ε. Moreover,

$$\frac{1}{n^2} \sum_{\substack{n_0+1\leq j_1+j_2;\\ j_1\leq N-1,\, j_2\leq n-1}} S(j_1,j_2) \leq \frac{1}{n^2} \sum_{\substack{n_0+1\leq j_1+j_2;\\ j_1\leq N-1,\, j_2\leq n-1}} T(n_0+1) \leq \varepsilon$$

and (3.6) is proved. By plugging (3.6) into (3.5), we prove that $\mathsf{E}\,\alpha_{n,2} \to 0$. Using the same kind of arguments, one proves that $\mathsf{E}\,\alpha_{n,1} \to 0$. Finally, (3.4) is proved: $\alpha_n \xrightarrow{\mathrm{P}} 0$.

Let us now check that

$$\exists\, K>0, \quad \mathsf{E}\,\beta_n \leq K \quad \text{and} \quad \exists\, \tilde{K}>0, \quad \mathsf{E}\,\tilde{\beta}_n \leq \tilde{K} \tag{3.8}$$

for n large enough. This will imply the tightness of β_n and $\tilde{\beta}_n$.

Recall that by assumption there exists $B_{\max}$ such that $\sup_n (1/n)\mathrm{Tr} B_n B_n^* \le B_{\max}$. Consider now

$$\begin{aligned}\frac{1}{n}\,\mathrm{Tr}(Z_n+B_n)(Z_n+B_n)^* &\le \Big(\Big(\frac{1}{n}\,\mathrm{Tr} Z_n Z_n^*\Big)^{1/2} + \Big(\frac{1}{n}\,\mathrm{Tr} B_n B_n^*\Big)^{1/2}\Big)^2 \\ &\le \Big(\Big(\frac{1}{n}\,\mathrm{Tr} Z_n Z_n^*\Big)^{1/2} + B_{\max}^{1/2}\Big)^2 .\end{aligned}$$

In particular,

$$\begin{aligned}\mathsf{E}\,\frac{\mathrm{Tr}(Z_n+B_n)(Z_n+B_n)^*}{n} &\le \mathsf{E}\,\frac{\mathrm{Tr} Z_n Z_n^*}{n} + 2B_{\max}^{1/2}\,\mathsf{E}\Big(\frac{\mathrm{Tr} Z_n Z_n^*}{n}\Big)^{1/2} + B_{\max} \\ &\overset{(a)}{\le} \mathsf{E}\,\frac{\mathrm{Tr} Z_n Z_n^*}{n} + 2B_{\max}^{1/2}\Big(\mathsf{E}\Big(\frac{\mathrm{Tr} Z_n Z_n^*}{n}\Big)\Big)^{1/2} + B_{\max}\end{aligned} \tag{3.9}$$

where (a) follows from Jensen's inequality. Notice that (3.9) still holds if one replaces Z_n by $\tilde{Z}_n$. Therefore in order to prove (3.8), it is sufficient to prove that

$$\exists\, K' > 0, \quad \mathsf{E}\Big(\frac{\mathrm{Tr} Z_n Z_n^*}{n}\Big) \le K' \quad \text{and} \quad \exists\, \tilde{K}' > 0, \quad \mathsf{E}\Big(\frac{\mathrm{Tr} \tilde{Z}_n \tilde{Z}_n^*}{n}\Big) \le \tilde{K}'.$$

Consider

$$\mathsf{E}\Big(\frac{\mathrm{Tr} Z_n Z_n^*}{n}\Big) = \frac{1}{n} \sum_{\substack{j_1=1:N\\ j_2=1:n}} \mathsf{E}\,|Z^n_{j_1 j_2}|^2 = N\,\mathsf{E}\,|Z^n_{11}|^2 = \frac{N}{n} C(0,0),$$

where C is defined by (1.1). This quantity is asymptotically bounded. From Lemma 3.1, we have

$$\mathsf{E}\Big(\frac{\mathrm{Tr} \tilde{Z}_n \tilde{Z}_n^*}{n}\Big) = \mathsf{E}\Big(\frac{\mathrm{Tr} Y_n Y_n^*}{n}\Big) = \frac{1}{n^2} \sum_{\substack{j_1=1:N\\ j_2=1:n}} \Big|\Phi\Big(\frac{j_1}{N}, \frac{j_2}{n}\Big)\Big|^2\,\mathsf{E}\,|X^n_{j_1 j_2}|^2 \le \frac{N}{n}\Phi^2_{\max},$$

which is also asymptotically bounded. Thus (3.8) is proved and so is Lemma 3.2. □

Proof of Theorem 3.1. Lemma 3.2 implies that

$$\mathsf{P}\left\{\mathcal{L}\big(F^{Z_n Z_n^*}, F^{\tilde{Z}_n \tilde{Z}_n^*}\big) \ge \varepsilon\right\} \xrightarrow[n\to\infty]{} 0 \quad \text{for every} \ \ \varepsilon > 0. \tag{3.10}$$

By Lemma 3.1, $F_N \tilde{Z}_n \tilde{Z}_n^* F_N^* = Y_n Y_n^*$. Since F_N is unitary, $\tilde{Z}_n \tilde{Z}_n^*$ and $Y_n Y_n^*$ have the same eigenvalues. Moreover, matrix Y_n fulfills (A-2) and the variance profile

Φ defined in (3.1) satisfies (A-3) since $(h(k_1,k_2),\ (k_1,k_2)\in \mathbf{Z}^2)$ is summable; therefore one can apply Theorem 2.1. In particular,

$$F^{\tilde{Z}_n\tilde{Z}_n^*} \xrightarrow[n\to\infty]{} \mu \ a.s. \quad \Longrightarrow \quad \forall\varepsilon>0,\quad \mathsf{P}\left\{\mathcal{L}\big(F^{\tilde{Z}_n\tilde{Z}_n^*},\mu\big)\geq\varepsilon\right\} \xrightarrow[n\to\infty]{} 0 \quad (3.11)$$

where μ is the probability distribution defined in Theorem 2.1. Now (3.10) together with (3.11) imply that $F^{Z_nZ_n^*} \xrightarrow{\mathrm{P}} \mu$ and Theorem 3.1 is proved. □

4. The limiting distribution in the non-centered stationary case

Recall the definitions of function Φ and matrix F_p (respectively defined in (3.1) and (3.2)).

Theorem 4.1 (Stationary entries, the non-centered case). *Let Z_n be a $N\times n$ matrix satisfying* (A-1)*; let A_n be a $N\times n$ matrix such that $\Lambda_n=F_NA_nF_n^*$ is $N\times n$ pseudo-diagonal and satisfies* (A-4)*. Then the empirical distributions of the eigenvalues of matrices $(Z_n+A_n)(Z_n+A_n)^*$ and $(Z_n+A_n)^*(Z_n+A_n)$ converge in probability to the non-random probability measures μ and $\tilde{\mu}$ defined in Theorem* 2.2 *as $n\to\infty$.*

Proof of Theorem 4.1. Denote $F^{(Z_n+A_n)(Z_n+A_n)^*}$ by F^n and $F^{(\tilde{Z}_n+A_n)(\tilde{Z}_n+A_n)^*}$ by $\tilde{F}^n$. Since Λ_n satisfies (A-4), $(1/n)\mathrm{Tr}A_nA_n^*=(1/n)\mathrm{Tr}\Lambda_n\Lambda_n^*$ is bounded and Lemma 3.2 implies that

$$\mathsf{P}\left\{\left|\mathcal{L}(F^n,\tilde{F}^n)\right|\geq\varepsilon\right\} \xrightarrow[n\to\infty]{} 0 \quad \text{for every} \quad \varepsilon>0. \tag{4.1}$$

By Lemma 3.1 and the assumption over A_n,

$$(\tilde{Z}_n+A_n)(\tilde{Z}_n+A_n)^* = F_N(Y_n+\Lambda_n)(Y_n+\Lambda_n)^*F_N^*.$$

Since the Fourier matrix F_N is unitary, $(\tilde{Z}_n+A_n)(\tilde{Z}_n+A_n)^*$ and $(Y_n+\Lambda_n)(Y_n+\Lambda_n)^*$ have the same eigenvalues. Since Φ defined in (3.1) satisfies (A-3), the matrices Y_n and Λ_n fulfill assumptions (A-2), (A-3) and (A-4), therefore one can apply Theorem 2.2. In particular,

$$\tilde{F}^n \xrightarrow[n\to\infty]{} \mu \ a.s. \quad \Longrightarrow \quad \forall\varepsilon>0,\quad \mathsf{P}\left\{\left|\mathcal{L}(\tilde{F}^n,\mu)\right|\geq\varepsilon\right\} \xrightarrow[n\to\infty]{} 0 \tag{4.2}$$

where μ is the probability distribution defined in Theorem 2.2. Relation (4.1) together with (4.2) imply that $F^n \xrightarrow{\mathrm{P}} \mu$ and Theorem 4.1 is proved. □

In the square case $n\times n$, we can deal with slightly more general matrices A_n.

Assumption A-5. The $n\times n$ matrix A_n is a Toeplitz matrix defined as $A_n=(a(j_1-j_2))_{0\leq j_1,j_2<n}$ where $(a(j))_{j\in\mathrm{Z}}$ is a deterministic sequence of complex numbers satisfying:

$$\sum_{j\in\mathrm{Z}}|a(j)|<\infty.$$

Let $\psi : [0,1] \mapsto \mathbf{C}$ be the so called symbol of A_n defined as

$$\psi(t) = \sum_{j \in \mathbb{Z}} a(j) e^{2\mathbf{i}\pi j t}. \tag{4.3}$$

Due to (A-5), ψ is bounded and continuous.

Theorem 4.2 (Stationary entries, the non-centered square case).
Let Z_n be a $n \times n$ matrix satisfying (A-1)*; let A_n be a $n \times n$ matrix satisfying* (A-5) *and let $n \to \infty$. Then the empirical distributions of the eigenvalues of matrices $(Z_n + A_n)(Z_n + A_n)^*$ and $(Z_n + A_n)^*(Z_n + A_n)$ converge in probability to non-random probability measures μ and $\tilde{\mu}$ whose Stieltjes transforms f and $\tilde{f}$ are given by*

$$f(z) = \int_{[0,1]} \pi_z(dx) \quad \text{and} \quad \tilde{f}(z) = \int_{[0,1]} \tilde{\pi}_z(dx)$$

where π_z and $\tilde{\pi}_z$ are the unique Stieltjes kernels with supports included in $[0,1]$ and satisfying the system of equations:

$$\int g \, d\pi_z = \int_0^1 \frac{g(u)}{-z(1 + \int |\Phi(u,\cdot)|^2 d\tilde{\pi}_z) + |\psi(u)|^2/[1 + \int |\Phi(\cdot,u)|^2 \, d\pi_z]} \, du \tag{4.4}$$

$$\int g \, d\tilde{\pi}_z = \int_0^1 \frac{g(u)}{-z(1 + \int |\Phi(\cdot,u)|^2 d\pi_z) + |\psi(u)|^2/[1 + \int |\Phi(u,\cdot)|^2 \, d\tilde{\pi}_z]} \, du \tag{4.5}$$

for every function $g \in C([0,1])$.

Proof. The proof is based on the fact that a Toeplitz matrix A_n is very close to a Toeplitz circulant matrix $\tilde{A}_n$ defined in such a way that the diagonal matrix $\Lambda_n = F_n \tilde{A}_n F_n^*$ satisfies assumption (A-4). Denoting by ψ_n the truncated function $\psi_n(t) = \sum_{j=-n}^{n} a(j) \exp\{2\mathbf{i}\pi j t\}$, we choose $\tilde{A}_n$ to be the matrix whose entries are defined by

$$\tilde{a}^n_{j_1 j_2} = \frac{1}{n} \sum_{k=0}^{n-1} \psi_n\Big(\frac{k}{n}\Big) \exp\Big(\frac{-2\pi \mathbf{i} k (j_1 - j_2)}{n}\Big).$$

Notice that in this case, $\Lambda_n = F_n \tilde{A}_n F_n^*$ is given by $\Lambda_n = \mathrm{diag}([\psi_n(0), \psi_n(1/n), \ldots, \psi_n((n-1)/n)])$ where $\mathrm{diag}(v)$ is the diagonal matrix bearing the entries of the vector v on its diagonal.

One can also prove that the complex number $\tilde{a}^n(j_1 - j_2) = \tilde{a}^n_{j_1 j_2}$ satisfies $\tilde{a}^n(0) = a(0) + a(n) + a(-n)$ and

$$\tilde{a}^n(j) = \begin{cases} a(j) + a(j-n) & \text{if} \quad n-1 \geq j > 0, \\ a(j) + a(j+n) & \text{if} \quad -n+1 \leq j < 0. \end{cases}$$

We denote by F^n and $\breve{F}^n$ the distribution functions $F^n = F^{(Z_n+A_n)(Z_n+A_n)^*}$ and $\breve{F}^n = F^{(Z_n+\tilde{A}_n)(Z_n+\tilde{A}_n)^*}$. We shall prove that $\mathcal{L}(F^n, \breve{F}^n) \to 0$ as $n \to \infty$.

Bai's inequality yields:

$$\mathcal{L}^4(F^n, \breve{F}^n) \le \frac{2}{n^2} \operatorname{Tr}(A_n - \tilde{A}_n)(A_n - \tilde{A}_n)^* \operatorname{Tr}(A_n A_n^* + \tilde{A}_n \tilde{A}_n^*). \tag{4.6}$$

We first prove that $n^{-1}\operatorname{Tr}(A_n A_n^*)$ and $n^{-1}\operatorname{Tr}(\tilde{A}_n \tilde{A}_n^*)$ are bounded:

$$\frac{1}{n} \operatorname{Tr} A_n A_n^* = \frac{1}{n} \sum_{j_1, j_2 = 0}^{n-1} |a(j_1 - j_2)|^2 = \sum_{j=-n+1}^{n-1} |a(j)|^2 \Big(1 - \frac{|j|}{n}\Big) \le \Big(\sum_{j \in \mathbb{Z}} |a(j)|\Big)^2. \tag{4.7}$$

Moreover,

$$\frac{1}{n} \operatorname{Tr} \tilde{A}_n \tilde{A}_n^* = \frac{1}{n} \operatorname{Tr} \Lambda_n \Lambda_n^* = \frac{1}{n} \sum_{j=0}^{n-1} \Big| \psi_n\Big(\frac{j}{n}\Big)\Big|^2 \le \Big(\sum_{j \in \mathbb{Z}} |a(j)|\Big)^2. \tag{4.8}$$

We now prove that

$$\frac{1}{n} \operatorname{Tr}(A_n - \tilde{A}_n)(A_n - \tilde{A}_n)^* \xrightarrow[n\to\infty]{} 0. \tag{4.9}$$

Indeed,

$$\begin{aligned}
&\frac{1}{n} \operatorname{Tr}(A_n - \tilde{A}_n)(A_n - \tilde{A}_n)^* \\
&= \frac{1}{n} \sum_{j_1, j_2 = 0}^{n-1} |a(j_1 - j_2) - \tilde{a}^n(j_1 - j_2)|^2 = \sum_{j=-(n-1)}^{n-1} |a(j) - \tilde{a}^n(j)|^2 \Big(1 - \frac{|j|}{n}\Big) \\
&= |a(-n) + a(n)|^2 + \sum_{j=1}^{n-1} \Big(|a(j-n)|^2 + |a(n-j)|^2\Big)\Big(1 - \frac{j}{n}\Big) \\
&= |a(-n) + a(n)|^2 + \sum_{j=1}^{n-1} \frac{j}{n} \Big(|a(j)|^2 + |a(-j)|^2\Big) \\
&\le |a(-n) + a(n)|^2 + \frac{1}{n} \sum_{j=1}^{J} j \Big(|a(j)|^2 + |a(-j)|^2\Big) + \sum_{j=J+1}^{\infty} \Big(|a(j)|^2 + |a(-j)|^2\Big).
\end{aligned}$$

By first taking J large enough then n large enough, the claim is proved by a 2ε-argument.

Inequality (4.6) together with the arguments provided by (4.7), (4.8) and (4.9) imply that

$$\mathcal{L}(F^n, \breve{F}^n) \xrightarrow[n\to\infty]{} 0.$$

It remains to prove that $\breve{F}^n$ converges towards the non-random probability distribution characterized by equations (4.4) and (4.5). As previously, the variance profile Φ defined in (3.1) satisfies (A-3). Moreover, we have

$$\frac{1}{n}\sum_{i=1}^{n}\delta_{(i/n,\,|\psi_n((i-1)/n)|^2)} \xrightarrow[n\to\infty]{} H(du,d\lambda)$$

where $H(du,d\lambda)$ is the image of the Lebesgue measure over $[0,1]$ under $u\mapsto(u,|\psi(u)|^2)$. Therefore Λ_n satisfies (A-4) and Theorem 4.1 can be applied. This completes the proof of Theorem 4.2. □

5. Remarks on the real case

In the case where the entries of matrix Z_n are given by

$$Z^n_{j_1j_2}=\frac{1}{\sqrt{n}}\sum_{(k_1,k_2)\in\mathrm{Z}^2}h(k_1,k_2)\,U(j_1-k_1,j_2-k_2),$$

where $(h(k_1,k_2),\ (k_1,k_2)\in\mathbf{Z}^2)$ is a deterministic real and summable sequence and where $U(j_1,j_2)$ are real standard independent Gaussian r.v.'s, the conclusion of Lemma 3.1 is no longer valid. In fact the entries of $Y_n=F_N\tilde{Z}_nF_n^*$ are far from being independent since straightforward computation yields:

$$Y^n_{\ell_1,\ell_2}=Y^{n^*}_{N-\ell_1,n-\ell_2}\quad\text{for}\quad 0<\ell_1<N\ \text{ and }\ 0<\ell_2<n.$$

We introduce the $p\times p$ orthogonal matrix $Q_p=(Q^p_{j_1j_2})_{0\le j_1,j_2<p}$ defined as follows.

$$Q^p_{0,j_2}=\frac{1}{\sqrt{p}},\quad 0\le j_2<p.$$

In the case where p is even, the entries $Q^p(j_1,j_2)$ $(j_1\ge 1)$ are defined by

$$\begin{cases} Q^p_{2j_1-1,j_2}=\sqrt{\dfrac{2}{p}}\cos\left(\dfrac{2\pi j_1j_2}{p}\right) & \text{if } 1\le j_1\le \dfrac{p}{2}-1,\ 0\le j_2<p;\\ Q^p_{2j_1,j_2}=\sqrt{\dfrac{2}{p}}\sin\left(\dfrac{2\pi j_1j_2}{p}\right) & \text{if } 1\le j_1\le \dfrac{p}{2}-1,\ 0\le j_2<p;\\ Q^p_{p-1,j_2}=\dfrac{(-1)^{j_2}}{\sqrt{p}} & \text{if } 0\le j_2<p.\end{cases}$$

In the case where p is odd, they are defined by

$$\begin{cases} Q^p_{2j_1-1,j_2}=\sqrt{\dfrac{2}{p}}\cos\left(\dfrac{2\pi j_1j_2}{p}\right) & \text{if } 1\le j_1\le \dfrac{p-1}{2},\ 0\le j_2<p;\\ Q^p_{2j_1,j_2}=\sqrt{\dfrac{2}{p}}\sin\left(\dfrac{2\pi j_1j_2}{p}\right) & \text{if } 1\le j_1\le \dfrac{p-1}{2},\ 0\le j_2<p.\end{cases}$$

In the sequel, $\lfloor x \rfloor$ stands for the integer part of x. The following result is the counterpart of Lemma 3.1 in the real case.

Lemma 5.1. *Consider the $N \times n$ matrix $W_n = Q_N \tilde{Z}_n Q_n^{\mathrm{T}}$ where A^{T} is the transpose of matrix A. Then the entries $W_{\ell_1 \ell_2}^n$ of W_n can be written as*

$$W_{\ell_1 \ell_2}^n = \frac{1}{\sqrt{n}} \left| \Phi\left(\frac{1}{N} \left\lfloor \frac{\ell_1 + 1}{2} \right\rfloor, \frac{1}{n} \left\lfloor \frac{\ell_2 + 1}{2} \right\rfloor \right) \right| X_{\ell_1 \ell_2}^n$$

where Φ is defined in (3.1) *and the real random variables $\{X_{\ell_1 \ell_2}^n, 0 \le \ell_1 < N, 0 \le \ell_2 < n\}$ are independent standard Gaussian r.v.'s.*

The proof is computationally more involved but similar in spirit to that of Lemma 3.1. It is thus omitted.

As a consequence of this lemma, Theorems 3.1 and 4.1 remain true with the following minor modification: In (2.2), (2.4) and (2.5), the quantity $|\Phi|^2$ must be replaced by Φ_{R}^2 where

$$\Phi_{\mathrm{R}}(u, v) = |\Phi(u/2, v/2)|.$$

Similarly, in the case where the Toeplitz matrix A_n introduced in (A-5) is real, Theorem 4.2 remains true if one replaces in (4.4) and (4.5) the quantities $|\Phi|^2$ and $|\psi|^2$ by Φ_{R}^2 and ψ_{R}^2 where

$$\psi_{\mathrm{R}}(u) = |\psi(u/2)|.$$

The proof of Theorem 4.2 can be modified by replacing the Fourier matrices F_p by Q_p (see also [5, chap. 4], for elements about the pseudo-diagonalization of a real Toeplitz matrix via real orthogonal matrices Q_p).

References

[1] G. Anderson and O. Zeitouni (2005) A CLT for a band matrix model. To appear in *PTRF*.

[2] Z.D. Bai (1993) Convergence rate of expected spectral distributions of large random matrices. II. Sample covariance matrices. *Ann. Prob.* **21** (2), 649–672.

[3] Z.D. Bai (1999) Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica* **9** (3), 611–677.

[4] A. Boutet de Monvel, A. Khorunzhy and V. Vasilchuk (1996) Limiting eigenvalue distribution of random matrices with correlated entries. *Markov Processes Relat. Fields* **2** (4), 607–636.

[5] P.J. Brockwell and R.A. Davis (1991) *Time Series: Theory and Methods.* Springer Series in Statistics, Springer-Verlag, New York.

[6] C.-N. Chuah, D.N.C. Tse, J.M. Kahn and R.A. Valenzuela (2002) Capacity scaling in MIMO wireless systems under correlated fading. *IEEE Trans. Inform. Theory* **48** (3), 637–650.

[7] G. Casati and V. Girko (1993) Generalized Wigner law for band random matrices. *Random Oper. and Stochas. Equations* **1** (3), 279–286.

[8] V.L. Girko (1990) *Theory of Random Determinants*. Volume 45 of Mathematics and its Applications (Soviet Series). Kluwer Academic Publishers Group, Dordrecht.

[9] V.L. Girko (2001) *Theory of Stochastic Canonical Equations*. Vol. I. Volume 535 of Mathematics and its Applications. Kluwer Academic Publishers, Dordrecht.

[10] V.L. Girko (2001) *Theory of Stochastic Canonical Equations*. Vol. II. Volume 535 of Mathematics and its Applications. Kluwer Academic Publishers, Dordrecht.

[11] W. Hachem, P. Loubaton and J. Najim (2004) The empirical distribution of the eigenvalues of a Gram matrix with a given variance profile. Available at `http://arxiv.org/abs/math.PR/0411333`.

[12] R.A. Horn and C.R. Johnson (1994) *Topics in Matrix Analysis*. Cambridge University Press, Cambridge.

[13] K. Liu, V. Raghavan and A.M. Sayeed (2003) Capacity scaling and spectral efficiency in wide-band correlated MIMO channels. *IEEE Trans. Inform. Theory* **49** (10), 2504–2526.

[14] D. Shlyakhtenko (1996) Random Gaussian band matrices and freeness with amalgamation. *Internat. Math. Res. Notices* **20**, 1013–1025.

[15] D. Shlyakhtenko (1998) Gaussian random band matrices and operator-valued free probability theory. In: *Quantum Probability* (Gdańsk, 1997), volume 43 of Banach Center Publ. Polish Acad. Sci., Warsaw., 359–368.

[16] A. Tulino and S. Verdú (2004) Random matrix theory and wireless communications. In: *Foundations and Trends in Communications and Information Theory*, **1**, 1–182. Now Publishers, June 2004.