



Limiting spectral distribution of Gram matrices associated with functionals of β -mixing processes



Marwa Banna

Université Paris Est, LAMA (UMR 8050), UPEMLV, CNRS, UPEC, 5 Boulevard Descartes,
77454 Marne La Vallée, France

ARTICLE INFO

Article history:

Received 23 June 2014

Available online 31 July 2015

Submitted by M. Peligrad

Keywords:

Random matrices

Sample covariance matrices

Stieltjes transform

Absolutely regular sequences

Limiting spectral distribution

Spectral density

ABSTRACT

We give asymptotic spectral results for Gram matrices of the form $n^{-1}\mathcal{X}_n\mathcal{X}_n^T$ where the entries of \mathcal{X}_n are dependent across both rows and columns. More precisely, they consist of short or long range dependent random variables having moments of second order and that are functionals of an absolutely regular sequence. We also give a concentration inequality of the Stieltjes transform and we prove that, under an arithmetical decay condition on the β -mixing coefficients, it is almost surely concentrated around its expectation. Applications to examples of positive recurrent Markov chains and dynamical systems are also given.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

For a random matrix $\mathcal{X}_n \in \mathbb{R}^{N \times n}$, the study of the asymptotic behavior of the eigenvalues of the $N \times N$ Gram matrix $n^{-1}\mathcal{X}_n\mathcal{X}_n^T$ gained interest as it is employed in many applications in statistics, signal processing, quantum physics, finance, etc. In order to describe the distribution of the eigenvalues, it is convenient to introduce the empirical spectral measure defined by $\mu_{n^{-1}\mathcal{X}_n\mathcal{X}_n^T} = N^{-1} \sum_{k=1}^N \delta_{\lambda_k}$, where $\lambda_1, \dots, \lambda_N$ are the eigenvalues of $n^{-1}\mathcal{X}_n\mathcal{X}_n^T$. This type of study was actively developed after the pioneering work of Marčenko and Pastur [11], who proved that under the assumption $\lim_{n \rightarrow +\infty} N/n = c \in (0, +\infty)$, the empirical spectral distribution of large dimensional Gram matrices with i.i.d. centered entries having finite variance converges almost surely to a non-random distribution. The limiting spectral distribution (LSD) obtained, i.e. the Marčenko–Pastur distribution, is given explicitly in terms of c and depends on the distribution of the entries of \mathcal{X}_n only through their common variance. The original Marčenko–Pastur theorem is stated for random variables having moments of fourth order; for the proof with second moments only, we refer to Yin [18].

Since then, a large amount of study has been done aiming to relax the independence structure between the entries of \mathcal{X}_n . For example, Bai and Zhou [2] treated the case where the columns of \mathcal{X}_n are independent

E-mail address: marwa.banna@u-pem.fr.

with their coordinates having a very general dependence structure and moments of fourth order. Recently, Banna and Merlevède [3] extended along another direction the Marčenko–Pastur theorem to a large class of weakly dependent sequences of real random variables having moments of second order. Letting $(X_k)_{k \in \mathbb{Z}}$ be a stationary process of the form $X_k = g(\cdots, \varepsilon_{k-1}, \varepsilon_k, \varepsilon_{k+1}, \cdots)$, where the ε_k 's are i.i.d. real-valued random variables and $g : \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}$ is a measurable function, they consider the $N \times N$ sample covariance matrix $\mathbf{A}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{X}_k \mathbf{X}_k^T$ with the \mathbf{X}_k 's being independent copies of the vector $\mathbf{X} = (X_1, \dots, X_N)^T$. Assuming that X_0 has just a moment of second order, then provided that $\lim_{n \rightarrow \infty} N/n = c \in (0, \infty)$, they prove, under a mild dependence condition, that almost surely, $\mu_{\mathbf{A}_n}$ converges weakly to a non-random probability measure μ whose Stieltjes transform satisfies an integral equation depending on c and on the spectral density of the underlying stationary process $(X_k)_{k \in \mathbb{Z}}$. In a recent paper, Merlevède and Peligrad [12] extend this result to stationary sequences satisfying mild regularity conditions and prove that the empirical spectral measure of a sample covariance matrix generated by independent copies of a stationary regular sequence has a limiting distribution depending only on the spectral density of the sequence.

In the above mentioned model, the random vector $\mathbf{X} = (X_1, \dots, X_N)^T$ can be viewed as an N -dimensional process repeated independently n times to obtain the \mathbf{X}_k 's. However, in practice it is not always possible to observe a high dimensional process several times. In the case where only one observation of length Nn can be recorded, it seems reasonable to partition it into n dependent observations of length N , and to treat them as n dependent observations. Up to our knowledge this was first done by Pfaffel and Schlemm [13] who showed that this approach is valid and leads to the correct asymptotic eigenvalue distribution of the sample covariance matrix if the components of the underlying process are modeled as short memory linear filters of independent random variables. They consider Gram matrices having the same form as in (2.3) and associated with a stationary linear process $(X_k)_{k \in \mathbb{Z}}$ with independent innovations having finite fourth moments and such that the coefficients decay with an arithmetical rate, and they derive its LSD.

In this work, we study the same model of random matrices as in [13] but considering the case where the entries come from a non-causal stationary process $(X_k)_{k \in \mathbb{Z}}$ of the form $X_k = g(\cdots, \varepsilon_{k-1}, \varepsilon_k, \varepsilon_{k+1}, \cdots)$ where $(\varepsilon_k)_{k \in \mathbb{Z}}$ is an absolutely regular sequence and $g : \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}$ is a measurable function such that X_k is a proper centered random variable having finite moments of second order. We prove in Theorem 2.1 a concentration inequality for the Stieltjes transform which allows us to prove that, under an arithmetical decay condition on the β -mixing coefficients, the Stieltjes transform is concentrated almost surely around its expectation as n tends to infinity. Having reduced the study to the expectation of the Stieltjes transform, it is enough to show that the latter converges to the Stieltjes transform of a non-random probability measure. This can be achieved by approximating it by the expectation of the Stieltjes transform of a Gaussian matrix having a close covariance structure as shown in Theorem 2.2. Finally, provided that the spectral density of $(X_k)_k$ exists, we prove in Theorem 2.3 that almost surely, $\mu_{\mathbf{B}_n}$ converges weakly to the same non-random limiting probability measure μ obtained in the cases mentioned before.

We recall now that the absolutely regular (β -mixing) coefficient between two σ -algebras \mathcal{A} and \mathcal{B} is defined by

$$\beta(\mathcal{A}, \mathcal{B}) = \frac{1}{2} \sup \left\{ \sum_{i \in I} \sum_{j \in J} |\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)| \right\},$$

where the supremum is taken over all finite partitions $(A_i)_{i \in I}$ and $(B_j)_{j \in J}$ that are respectively \mathcal{A} and \mathcal{B} measurable (see Rozanov and Volkonskii [16]). The coefficients $(\beta_n)_{n \geq 0}$ of a sequence $(\varepsilon_i)_{i \in \mathbb{Z}}$ are defined by

$$\beta_0 = 1 \quad \text{and} \quad \beta_n = \sup_{k \in \mathbb{Z}} \beta(\sigma(\varepsilon_\ell, \ell \leq k), (\varepsilon_{\ell+n}, \ell \geq k)) \quad \text{for } n \geq 1. \quad (1.1)$$

Moreover, $(\varepsilon_i)_{i \in \mathbb{Z}}$ is said to be absolutely regular or β -mixing if $\beta_n \rightarrow 0$ as $n \rightarrow \infty$.

Outline. In Section 2, we specify the model studied and state the limiting results for the Gram matrix associated with the process defined in (2.1). The proofs shall be deferred to Section 4, whereas applications to examples of Markov chains and dynamical systems shall be introduced in Section 3.

Notation. For any real numbers x and y , $x \wedge y := \min(x, y)$ whereas $x \vee y := \max(x, y)$. Moreover, the notation $[x]$ denotes the integer part of x . For any non-negative integer q , a null row vector of dimension q will be denoted by $\mathbf{0}_q$. For a matrix A , we denote by A^T its transpose matrix and by $\text{Tr}(A)$ its trace. Finally, we shall use the notation $\|X\|_r$ for the \mathbb{L}^r -norm ($r \geq 1$) of a real-valued random variable X .

For any square matrix A of order N having only real eigenvalues, its empirical spectral measure and distribution are respectively defined by

$$\mu_A = \frac{1}{N} \sum_{k=1}^N \delta_{\lambda_k} \text{ and } F^A(x) = \frac{1}{N} \sum_{k=1}^N \mathbf{1}_{\{\lambda_k \leq x\}},$$

where $\lambda_1, \dots, \lambda_N$ are the eigenvalues of A . The Stieltjes transform of μ_A is given by

$$S_A(z) := S_{\mu_A}(z) = \int \frac{1}{x - z} d\mu_A(x) = \frac{1}{N} \text{Tr}(A - z\mathbf{I})^{-1},$$

where \mathbf{I} is the identity matrix of order N and $z = u + iv \in \mathbb{C}_+$ with \mathbb{C}_+ being the set of complex numbers with positive imaginary part.

2. Results

We consider a non-causal stationary process $(X_k)_{k \in \mathbb{Z}}$ defined as follows: let $(\varepsilon_i)_{i \in \mathbb{Z}}$ be an absolutely regular process with a β -mixing sequence $(\beta_k)_{k \geq 0}$ and let for any $k \in \mathbb{Z}$,

$$X_k = g(\xi_k) \quad \text{with} \quad \xi_k = (\dots, \varepsilon_{k-1}, \varepsilon_k, \varepsilon_{k+1}, \dots) \quad (2.1)$$

where $g: \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}$ is a measurable function such that X_k is a proper centered random variable having finite moment of second order; that is, $\mathbb{E}(X_k) = 0$ and $\|X_k\|_2 < \infty$.

Now, let $N := N(n)$ be a sequence of positive integers and consider the $N \times n$ random matrix \mathcal{X}_n defined by

$$\mathcal{X}_n = ((\mathcal{X}_n)_{i,j}) = (X_{(j-1)N+i}) = \begin{pmatrix} X_1 & X_{N+1} & \cdots & X_{(n-1)N+1} \\ X_2 & X_{N+2} & \cdots & X_{(n-1)N+2} \\ \vdots & \vdots & & \vdots \\ X_N & X_{2N} & \cdots & X_{nN} \end{pmatrix} \in \mathcal{M}_{N \times n}(\mathbb{R}) \quad (2.2)$$

and note that its entries are dependent across both rows and columns. Let \mathbf{B}_n be its corresponding Gram matrix given by

$$\mathbf{B}_n = \frac{1}{n} \mathcal{X}_n \mathcal{X}_n^T. \quad (2.3)$$

In what follows, \mathbf{B}_n will be referred to as the Gram matrix associated with $(X_k)_{k \in \mathbb{Z}}$. We note that it can be written as the sum of dependent rank one matrices. Namely, $\mathbf{B}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^T$, where for any $i = 1, \dots, n$, $\mathbf{X}_i = (X_{(i-1)N+1}, \dots, X_{iN})^T$. Our purpose is to study the limiting distribution of the empirical spectral measure $\mu_{\mathbf{B}_n}$ defined by

$$\mu_{\mathbf{B}_n}(x) = \frac{1}{N} \sum_{k=1}^N \delta_{\lambda_k},$$

where $\lambda_1, \dots, \lambda_N$ are the eigenvalues of \mathbf{B}_n . We start by showing that if the β -mixing coefficients decay arithmetically then the Stieltjes transform of \mathbf{B}_n concentrates almost surely around its expectation as n tends to infinity.

Theorem 2.1. *Let \mathbf{B}_n be the matrix defined in (2.3) and associated with $(X_k)_{k \in \mathbb{Z}}$ defined in (2.1). If $\lim_{n \rightarrow \infty} N/n = c \in (0, \infty)$ and*

$$\sum_{n \geq 1} \log(n)^{\frac{3\alpha}{2}} n^{-\frac{1}{2}} \beta_n < \infty \quad \text{for some } \alpha > 1, \quad (2.4)$$

the following convergence holds: for any $z \in \mathbb{C}_+$,

$$S_{\mathbf{B}_n}(z) - \mathbb{E}(S_{\mathbf{B}_n}(z)) \rightarrow 0 \quad \text{almost surely, as } n \rightarrow +\infty.$$

The above convergence shall be proved via a concentration inequality of the Stieltjes transform. However, as the \mathbf{X}_i 's are dependent, classical arguments as those in Theorem 1(ii) of [8] are no longer sufficient. In fact, we shall use Berbee's maximal coupling for absolutely regular sequences which will allow us to break the dependence structure between the matrix columns and eventually get a concentration inequality in terms of the β -mixing coefficients. The proof of Theorem 2.1 shall be postponed to Section 4.

In the following theorem, we shall approximate the Stieltjes transform of the LSD of \mathbf{B}_n with that of a sample covariance matrix \mathbf{G}_n which is the sum of i.i.d. rank one matrices associated with a Gaussian process $(Z_k)_{k \in \mathbb{Z}}$ having the same covariance structure as $(X_k)_{k \in \mathbb{Z}}$. Namely, for any $k, \ell \in \mathbb{Z}$,

$$\text{Cov}(Z_k, Z_\ell) = \text{Cov}(X_k, X_\ell). \quad (2.5)$$

For $i = 1, \dots, n$, we denote by $(Z_k^{(i)})_{k \in \mathbb{Z}}$ an independent copy of $(Z_k)_k$ that is also independent of $(X_k)_k$ and we define the $N \times N$ sample covariance matrix \mathbf{G}_n by

$$\mathbf{G}_n = \frac{1}{n} \mathbf{Z}_n \mathbf{Z}_n^T = \frac{1}{n} \sum_{i=1}^n \mathbf{Z}_i \mathbf{Z}_i^T, \quad (2.6)$$

where for any $i = 1, \dots, n$, $\mathbf{Z}_i = (Z_1^{(i)}, \dots, Z_N^{(i)})^T$ and \mathbf{Z}_n is the matrix whose columns are the \mathbf{Z}_i 's. Namely, $\mathbf{Z}_n := ((\mathbf{Z}_n)_{u,v}) = (Z_u^{(v)})$.

Theorem 2.2. *Let \mathbf{B}_n and \mathbf{G}_n be the matrices defined in (2.3) and (2.6) respectively. Provided that $\lim_{n \rightarrow \infty} N/n = c \in (0, \infty)$ and $\lim_{n \rightarrow \infty} \beta_n = 0$, then for any $z \in \mathbb{C}_+$,*

$$\lim_{n \rightarrow \infty} |\mathbb{E}(S_{\mathbf{B}_n}(z)) - \mathbb{E}(S_{\mathbf{G}_n}(z))| = 0.$$

The above theorem allows us to reduce the study of the expectation of the Stieltjes transform of \mathbf{B}_n to that of a Gram matrix, being the sum of independent rank-one matrices associated with a Gaussian process, without requiring any rate of convergence to zero of the correlation between the entries nor of the β -mixing coefficients.

Theorem 2.3. *Let \mathbf{B}_n and \mathbf{G}_n be the matrices defined in (2.3) and (2.6) respectively. Provided that $\lim_{n \rightarrow \infty} N/n = c \in (0, \infty)$ and that (2.4) is satisfied, then for any $z \in \mathbb{C}_+$,*

$$\lim_{n \rightarrow \infty} |S_{\mathbf{B}_n}(z) - S_{\mathbf{G}_n}(z)| = 0 \quad \text{a.s.} \quad (2.7)$$

Moreover, if $(X_k)_{k \in \mathbb{Z}}$ admits a spectral density f , then with probability one, $\mu_{\mathbf{B}_n}$ converges weakly to a probability measure μ whose Stieltjes transform $S = S(z)$ ($z \in \mathbb{C}_+$) satisfies the equation

$$z = -\frac{1}{\underline{S}} + \frac{c}{2\pi} \int_0^{2\pi} \frac{1}{\underline{S} + (2\pi f(\lambda))^{-1}} d\lambda, \quad (2.8)$$

where $\underline{S}(z) := -(1-c)/z + cS(z)$.

The convergence (2.7) in Theorem 2.3 is a combination of Theorems 2.1 and 2.2 of this paper and Theorem 1(ii) of [8]. To prove the second part of the theorem, it suffices to notice that, provided that the spectral density f of $(X_k)_{k \in \mathbb{Z}}$ exists, then according to the proof of Theorem 1 by Merlevède and Peligrad [12], we have that, almost surely, the empirical spectral measure of \mathbf{G}_n converges weakly to a non-random probability measure whose Stieltjes transform $S = S(z)$ ($z \in \mathbb{C}_+$) is uniquely determined by (2.8).

Remark 2.4. The spectral density function f of $(X_k)_{k \in \mathbb{Z}}$ is the discrete Fourier transform of the autocovariance function. If $\sum_{k \in \mathbb{Z}} |\text{Cov}(X_0, X_k)| < \infty$ then f exists, is continuous and bounded on $[0, 2\pi)$. It also follows from Proposition 1 by Yao [17] that the limiting distribution is compactly supported.

3. Applications

In this section we shall apply the results of Section 2 to a Harris recurrent Markov chain and some uniformly expanding maps in dynamical systems.

3.1. Harris recurrent Markov chain

The following example is a symmetrized version of the Harris recurrent Markov chain defined by Doukhan et al. [7]. Let $(\varepsilon_n)_{n \in \mathbb{Z}}$ be a stationary Markov chain taking values in $E = [-1, 1]$ and let K be its Markov kernel defined by

$$K(x, \cdot) = (1 - |x|)\delta_x + |x|\nu,$$

where ν is a symmetric atomless law on E and δ_x is the Dirac measure at point x . Assuming that $\theta = \int_E |x|^{-1} \nu(dx) < \infty$ then $(\varepsilon_n)_{n \in \mathbb{Z}}$ is positively recurrent and the unique invariant measure π is given by

$$\pi(dx) = \theta^{-1} |x|^{-1} \nu(dx).$$

We shall assume in what follows that ν satisfies for any $x \in [0, 1]$,

$$\frac{d\nu}{dx}(x) \leq c x^a \quad \text{for some } a, c > 0. \quad (3.1)$$

Finally, let g be a measurable function defined on E such that

$$X_k = g(\varepsilon_k) \quad (3.2)$$

is a centered random variable having a finite second moment.

Corollary 3.1. *Let $(X_k)_{k \in \mathbb{Z}}$ be defined as in (3.2) and assume that ν satisfies (3.1) and that for any $x \in E$, $g(-x) = -g(x)$ and $|g(x)| \leq C|x|^{1/2}$ with C being a positive constant. Then, provided that $N/n \rightarrow c \in (0, \infty)$, the conclusion of Theorem 2.2 holds and the limiting measure μ has a compact support. In addition, if (3.1) holds with $a > 1/2$ then Theorems 2.1 and 2.3 follow as well.*

Proof. Doukhan et al. prove in Section 4 of [7] that if (3.1) is satisfied then $(\varepsilon_k)_{k \in \mathbb{Z}}$ is an absolutely regular sequence with $\beta_n = O(n^{-a})$ as $n \rightarrow \infty$ and thereby, Theorem 2.2 follows. Now, noting that g is an odd function we have

$$\mathbb{E}(g(\varepsilon_k)|\varepsilon_0) = (1 - |\varepsilon_0|)^k g(\varepsilon_0) \quad \text{a.s.}$$

Therefore, by the assumption on g and (3.1), we get for any $k \geq 0$,

$$\begin{aligned} \gamma_k := \mathbb{E}(X_0 X_k) &= \mathbb{E}(g(\varepsilon_0) \mathbb{E}(g(\varepsilon_k)|\varepsilon_0)) = \theta^{-1} \int_E g^2(x) (1 - |x|)^k |x|^{-1} \nu(dx) \\ &\leq c C^2 \theta^{-1} \int_E \frac{x^{a+1}}{|x|} (1 - |x|)^k dx. \end{aligned} \quad (3.3)$$

By the properties of the Beta and Gamma functions, $|\gamma_k| = O(\frac{1}{k^{a+1}})$ which implies $\sum_k |\gamma_k| < \infty$ and thus the spectral density f is continuous and bounded over $[0, 2\pi)$ and the limiting distribution has a compact support as mentioned in Remark 2.4. However, if in addition $a > 1/2$ then (2.4) is also satisfied and Theorems 2.1 and 2.3 follow as well. \square

3.2. Uniformly expanding maps

Functionals of absolutely regular sequences occur naturally as orbits of chaotic dynamical systems. For instance, for a uniformly expanding map $T : [0, 1] \rightarrow [0, 1]$ with an absolutely continuous invariant measure ν , one can write T^k as a measurable function $g : \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}$ of an absolutely regular sequence $(\varepsilon_k)_{k \geq 0}$. Namely,

$$T^k = g(\varepsilon_k, \varepsilon_{k+1}, \dots).$$

We refer to Section 2 of [10] and Example 1.4 of [6] for more details and for a precise definition of such maps. Hofbauer and Keller prove in Theorem 4 of [10] that the mixing rate of $(\varepsilon_k)_{k \geq 0}$ decreases exponentially, i.e.

$$\beta_k \leq C e^{-\lambda k}, \quad \text{for some } C, \lambda > 0, \quad (3.4)$$

and thus (2.4) holds. Setting for any $k \geq 0$,

$$X_k = h \circ T^k - \nu(f), \quad (3.5)$$

where $h : [0, 1] \rightarrow \mathbb{R}$ is a continuous Hölder function, the theorems in Section 2 hold for the associated matrix \mathbf{B}_n . Moreover, Hofbauer and Keller prove in Theorem 5 of [10] that $\sum_k |\text{Cov}(X_0, X_k)| < \infty$ which implies that the spectral density f exists, is continuous and bounded on $[0, 2\pi)$ and that the limiting measure μ is compactly supported.

4. Proof of Theorem 2.1

Let m be a positive integer, fixed for the moment, such that $m \leq \sqrt{N}/2$ and let $(X_{k,m})_{k \in \mathbb{Z}}$ be the sequence defined for any $k \in \mathbb{Z}$ by

$$X_{k,m} = \mathbb{E}(X_k | \varepsilon_{k-m}, \dots, \varepsilon_{k+m}). \quad (4.1)$$

Consider the $N \times n$ matrix $\mathcal{X}_{n,m} = ((\mathcal{X}_{n,m})_{i,j}) = (X_{(j-1)N+i,m})$ and finally set

$$\mathbf{B}_{n,m} = \frac{1}{n} \mathcal{X}_{n,m} \mathcal{X}_{n,m}^T. \quad (4.2)$$

The proof will be done in two principal steps. First, we shall prove that

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} |S_{\mathbf{B}_n}(z) - S_{\mathbf{B}_{n,m}}(z)| = 0 \quad \text{a.s.} \quad (4.3)$$

and then

$$\lim_{n \rightarrow \infty} |S_{\mathbf{B}_{n,m}}(z) - \mathbb{E}(S_{\mathbf{B}_{n,m}}(z))| = 0 \quad \text{a.s.} \quad (4.4)$$

We note that for any two $N \times n$ random matrices \mathbf{A} and \mathbf{B} , we have

$$|S_{\mathbf{A}\mathbf{A}^T}(z) - S_{\mathbf{B}\mathbf{B}^T}(z)| \leq \frac{\sqrt{2}}{Nv^2} |\text{Tr}(\mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T)|^{1/2} |\text{Tr}(\mathbf{A} - \mathbf{B})(\mathbf{A} - \mathbf{B})^T|^{1/2}. \quad (4.5)$$

For a proof, the reader can check inequalities (4.18) and (4.19) of [3]. Thus, we get, for any $z = u + iv \in \mathbb{C}_+$,

$$|S_{\mathbf{B}_n}(z) - S_{\mathbf{B}_{n,m}}(z)|^2 \leq \frac{2}{v^4} \left(\frac{1}{N} \text{Tr}(\mathbf{B}_n + \mathbf{B}_{n,m}) \right) \left(\frac{1}{Nn} \text{Tr}(\mathcal{X}_n - \mathcal{X}_{n,m})(\mathcal{X}_n - \mathcal{X}_{n,m})^T \right). \quad (4.6)$$

Recall that mixing implies ergodicity and note that as $(\varepsilon_k)_{k \in \mathbb{Z}}$ is an ergodic sequence of real-valued random variables then $(X_k)_{k \in \mathbb{Z}}$ is also so. Therefore, by the ergodic theorem,

$$\lim_{n \rightarrow +\infty} \frac{1}{N} \text{Tr}(\mathbf{B}_n) = \lim_{n \rightarrow +\infty} \frac{1}{Nn} \sum_{k=1}^{Nn} X_k^2 = \mathbb{E}(X_0^2) \quad \text{a.s.} \quad (4.7)$$

Similarly,

$$\lim_{n \rightarrow \infty} \frac{1}{N} \text{Tr}(\mathbf{B}_{n,m}) = \mathbb{E}(X_{0,m}^2) \quad \text{a.s.} \quad (4.8)$$

Starting from (4.6) and noticing that $\mathbb{E}(X_{0,m}^2) \leq \mathbb{E}(X_0^2)$, it follows that (4.3) holds if we prove

$$\lim_{m \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \left| \frac{1}{Nn} \text{Tr}(\mathcal{X}_n - \mathcal{X}_{n,m})(\mathcal{X}_n - \mathcal{X}_{n,m})^T \right| = 0 \quad \text{a.s.} \quad (4.9)$$

By the construction of \mathcal{X}_n and $\mathcal{X}_{n,m}$ and again the ergodic theorem, we get

$$\lim_{n \rightarrow \infty} \frac{1}{Nn} \text{Tr}(\mathcal{X}_n - \mathcal{X}_{n,m})(\mathcal{X}_n - \mathcal{X}_{n,m})^T = \lim_{n \rightarrow \infty} \frac{1}{Nn} \sum_{k=1}^{Nn} (X_k - X_{k,m})^2 = \mathbb{E}(X_0 - X_{0,m})^2 \quad \text{a.s.}$$

(4.9) follows by applying the usual martingale convergence theorem in \mathbb{L}^2 (Corollary 2.2 in [9]), from which we infer that $\lim_{m \rightarrow +\infty} \|X_0 - \mathbb{E}(X_0 | \varepsilon_{-m}, \dots, \varepsilon_m)\|_2 = 0$.

We turn now to the proof of (4.4). With this aim, we shall prove that for any $z = u + iv$ and $x > 0$,

$$\mathbb{P}(|S_{\mathbf{B}_{n,m}}(z) - \mathbb{E}S_{\mathbf{B}_{n,m}}(z)| > 4x) \leq 4 \exp \left\{ -\frac{x^2 v^2 N^2 (\log n)^\alpha}{256 n^2} \right\} + \frac{32 n^2 (\log n)^\alpha}{x^2 v^2 N^2} \beta_{\lfloor \frac{n}{(\log n)^\alpha} \rfloor N}, \quad (4.10)$$

for some $\alpha > 1$. Noting that

$$\sum_{n \geq 2} (\log n)^\alpha \beta_{\lfloor \frac{n^2}{(\log n)^\alpha} \rfloor} < +\infty \text{ is equivalent to (2.4)}$$

and applying the Borel–Cantelli Lemma, (4.4) follows from (2.4) and the fact that $\lim_{n \rightarrow \infty} N/n = c$. Now, to prove (4.10), we start by noting that

$$\begin{aligned} & \mathbb{P}\left(|S_{\mathbf{B}_{n,m}}(z) - \mathbb{E}(S_{\mathbf{B}_{n,m}}(z))| > 4x\right) \\ & \leq \mathbb{P}\left(|\Re(S_{\mathbf{B}_{n,m}}(z)) - \mathbb{E}\Re(S_{\mathbf{B}_{n,m}}(z))| > 2x\right) + \mathbb{P}\left(|\Im(S_{\mathbf{B}_{n,m}}(z)) - \mathbb{E}\Im(S_{\mathbf{B}_{n,m}}(z))| > 2x\right). \end{aligned}$$

For a row vector $\mathbf{x} \in \mathbb{R}^{Nn}$, we partition it into n elements of dimension N and write $\mathbf{x} = (x_1, \dots, x_n)$ where x_1, \dots, x_n are row vectors of \mathbb{R}^N . Now, let $A(\mathbf{x})$ and $B(\mathbf{x})$ be respectively the $N \times n$ and $N \times N$ matrices defined by

$$A(\mathbf{x}) = (x_1^T | \dots | x_n^T) \quad \text{and} \quad B(\mathbf{x}) = \frac{1}{n} A(\mathbf{x}) A(\mathbf{x})^T. \quad (4.11)$$

Also, let $h_1 := h_{1,z}$ and $h_2 := h_{2,z}$ be the functions defined from \mathbb{R}^{Nn} into \mathbb{R} by

$$h_1(\mathbf{x}) = \int f_{1,z} d\mu_{B(\mathbf{x})} \quad \text{and} \quad h_2(\mathbf{x}) = \int f_{2,z} d\mu_{B(\mathbf{x})},$$

where $f_{1,z}(\lambda) = \frac{\lambda - u}{(\lambda - u)^2 + v^2}$ and $f_{2,z}(\lambda) = \frac{v}{(\lambda - u)^2 + v^2}$ and note that $S_{B(\mathbf{x})}(z) = h_1(\mathbf{x}) + ih_2(\mathbf{x})$. Now, denoting by $\mathbf{X}_{1,m}^T, \dots, \mathbf{X}_{n,m}^T$ the columns of $\mathcal{X}_{n,m}$ and setting \mathbf{A} to be the row random vector of \mathbb{R}^{Nn} given by

$$\mathbf{A} = (\mathbf{X}_{1,m}, \dots, \mathbf{X}_{n,m}),$$

we note that $B(\mathbf{A}) = \mathbf{B}_{n,m}$ and $h_1(\mathbf{A}) = \Re(S_{\mathbf{B}_{n,m}}(z))$. Moreover, letting q be a positive integer less than n , we set $\mathcal{F}_i = \sigma(\varepsilon_k, k \leq iN + m)$ for $1 \leq i \leq [n/q]q$ with the convention that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and that $\mathcal{F}_s = \mathcal{F}_n$ for any $s \in \{[n/q]q, \dots, n\}$. Noting that $\mathbf{X}_{1,m}, \dots, \mathbf{X}_{i,m}$ are \mathcal{F}_i -measurable, we write the following decomposition:

$$\begin{aligned} \Re(S_{\mathbf{B}_{n,m}}(z)) - \mathbb{E}\Re(S_{\mathbf{B}_{n,m}}(z)) &= h_1(\mathbf{X}_{1,m}, \dots, \mathbf{X}_{n,m}) - \mathbb{E}h_1(\mathbf{X}_{1,m}, \dots, \mathbf{X}_{n,m}) \\ &= \sum_{i=1}^{[n/q]} (\mathbb{E}(h_1(\mathbf{A})|\mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A})|\mathcal{F}_{(i-1)q})). \end{aligned}$$

Now, let $(\mathbf{A}_i)_i$ be a family of row random vectors of \mathbb{R}^{Nn} defined for any $i \in \{1, \dots, [n/q] - 1\}$ by

$$\mathbf{A}_i = (\mathbf{X}_{1,m}, \dots, \mathbf{X}_{(i-1)q,m}, \underbrace{\mathbf{0}_N, \dots, \mathbf{0}_N}_{2q \text{ times}}, \mathbf{X}_{(i+1)q+1,m}, \dots, \mathbf{X}_{n,m}),$$

and for $i = [n/q]$ by

$$\mathbf{A}_{[n/q]} = (\mathbf{X}_{1,m}, \dots, \mathbf{X}_{([n/q]-1)q,m}, \mathbf{0}_N, \dots, \mathbf{0}_N).$$

Noting that $\mathbb{E}(h_1(\mathbf{A}_{[n/q]})|\mathcal{F}_n) = \mathbb{E}(h_1(\mathbf{A}_{[n/q]})|\mathcal{F}_{([n/q]-1)q})$, we write

$$\begin{aligned} & \Re(S_{\mathbf{B}_{n,m}}(z)) - \mathbb{E} \Re(S_{\mathbf{B}_{n,m}}(z)) \\ &= \sum_{i=1}^{[n/q]} \left(\mathbb{E}(h_1(\mathbf{A}) - h_1(\mathbf{A}_i)|\mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A}) - h_1(\mathbf{A}_i)|\mathcal{F}_{(i-1)q}) \right) \\ & \quad + \sum_{i=1}^{[n/q]-1} \left(\mathbb{E}(h_1(\mathbf{A}_i)|\mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A}_i)|\mathcal{F}_{(i-1)q}) \right) \\ &:= M_{[n/q],q} + \sum_{i=1}^{[n/q]-1} \left(\mathbb{E}(h_1(\mathbf{A}_i)|\mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A}_i)|\mathcal{F}_{(i-1)q}) \right). \end{aligned} \quad (4.12)$$

Thus, we get

$$\begin{aligned} & \mathbb{P} \left(\left| \Re(S_{\mathbf{B}_{n,m}}(z)) - \mathbb{E} \Re(S_{\mathbf{B}_{n,m}}(z)) \right| > 2x \right) \\ & \leq \mathbb{P}(|M_{[n/q],q}| > x) + \mathbb{P} \left(\left| \sum_{i=1}^{[n/q]-1} (\mathbb{E}(h_1(\mathbf{A}_i)|\mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A}_i)|\mathcal{F}_{(i-1)q})) \right| > x \right). \end{aligned} \quad (4.13)$$

Note that $(M_{k,q})_k$ is a centered martingale with respect to the filtration $(\mathcal{G}_{k,q})_k$ defined by $\mathcal{G}_{k,q} = \mathcal{F}_{kq}$. Moreover, for any $k \in \{1, \dots, [n/q]\}$,

$$\begin{aligned} \|M_{k,q} - M_{k-1,q}\|_\infty &= \|\mathbb{E}(h_1(\mathbf{A}) - h_1(\mathbf{A}_k)|\mathcal{F}_{kq}) - \mathbb{E}(h_1(\mathbf{A}) - h_1(\mathbf{A}_k)|\mathcal{F}_{(k-1)q})\|_\infty \\ &\leq 2\|h_1(\mathbf{A}) - h_1(\mathbf{A}_k)\|_\infty. \end{aligned}$$

Noting that $\|f'_{1,z}\|_1 = 2/v$ then by integrating by parts, we get

$$\begin{aligned} |h_1(\mathbf{A}) - h_1(\mathbf{A}_k)| &= \left| \int f_{1,z} d\mu_{B(\mathbf{A})} - \int f_{1,z} d\mu_{B(\mathbf{A}_k)} \right| \leq \|f'_{1,z}\|_1 \|F^{B(\mathbf{A})} - F^{B(\mathbf{A}_k)}\|_\infty \\ &\leq \frac{2}{vN} \text{Rank}(A(\mathbf{A}) - A(\mathbf{A}_k)), \end{aligned} \quad (4.14)$$

where the second inequality follows from Theorem A.44 in [1]. As for any $k \in \{1, \dots, [n/q]-1\}$, $\text{Rank}(A(\mathbf{A}) - A(\mathbf{A}_k)) \leq 2q$ and $\text{Rank}(A(\mathbf{A}) - A(\mathbf{A}_{[n/q]})) \leq q$, then overall we derive that almost surely

$$\|M_{k,q} - M_{k-1,q}\|_\infty \leq \frac{8q}{vN} \quad \text{and} \quad \|M_{[n/q],q} - M_{[n/q]-1,q}\|_\infty \leq \frac{4q}{vN}$$

and hence applying the Azuma–Hoeffding inequality for martingales we get for any $x > 0$,

$$\mathbb{P}(|M_{[n/q],q}| > x) \leq 2 \exp \left\{ -\frac{x^2 v^2 N^2}{128 q n} \right\}. \quad (4.15)$$

Now to control the second term of (4.12), we have, by Markov's inequality and orthogonality, for any $x > 0$,

$$\begin{aligned} & \mathbb{P} \left(\left| \sum_{i=1}^{[n/q]-1} (\mathbb{E}(h_1(\mathbf{A}_i)|\mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A}_i)|\mathcal{F}_{(i-1)q})) \right| > x \right) \\ & \leq \frac{1}{x^2} \sum_{i=1}^{[n/q]-1} \|\mathbb{E}(h_1(\mathbf{A}_i)|\mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A}_i)|\mathcal{F}_{(i-1)q})\|_2^2. \end{aligned} \quad (4.16)$$

Fixing $i \in \{1, \dots, [n/q] - 1\}$, one can construct by Berbee's maximal coupling lemma [5], a sequence $(\varepsilon'_k)_{k \in \mathbb{Z}}$ distributed as $(\varepsilon_k)_{k \in \mathbb{Z}}$ and independent of \mathcal{F}_{iq} such that for any $j > iqN + m$,

$$\mathbb{P}(\varepsilon'_k \neq \varepsilon_k, \text{ for some } k \geq j) = \beta_{j-iqN-m}. \quad (4.17)$$

Let $(X'_{k,m})_{k \geq 1}$ be the sequence defined for any $k \geq 1$ by $X'_{k,m} = \mathbb{E}(X_k | \varepsilon'_{k-m}, \dots, \varepsilon'_{k+m})$ and let $\mathbf{X}'_{i,m}$ be the row vector of \mathbb{R}^N defined by $\mathbf{X}'_{i,m} = (X'_{(i-1)N+1,m}, \dots, X'_{iN,m})$. Finally, we define for any $i \in \{1, \dots, [n/q] - 1\}$ the row random vector \mathbf{A}'_i of \mathbb{R}^{Nn} by

$$\mathbf{A}'_i = (\mathbf{X}_{1,m}, \dots, \mathbf{X}_{(i-1)q,m}, \underbrace{\mathbf{0}_N, \dots, \mathbf{0}_N}_{2q \text{ times}}, \mathbf{X}'_{(i+1)q+1,m}, \dots, \mathbf{X}'_{n,m}).$$

As $\mathbf{X}'_{(i+1)q+1,m}, \dots, \mathbf{X}'_{n,m}$ are independent of \mathcal{F}_{iq} then $\mathbb{E}(h_1(\mathbf{A}'_i) | \mathcal{F}_{iq}) = \mathbb{E}(h_1(\mathbf{A}'_i) | \mathcal{F}_{(i-1)q})$. Thus we write

$$\mathbb{E}(h_1(\mathbf{A}_i) | \mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A}_i) | \mathcal{F}_{(i-1)q}) = \mathbb{E}(h_1(\mathbf{A}_i) - h_1(\mathbf{A}'_i) | \mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A}_i) - h_1(\mathbf{A}'_i) | \mathcal{F}_{(i-1)q}),$$

and infer that

$$\begin{aligned} & \|\mathbb{E}(h_1(\mathbf{A}_i) | \mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A}_i) | \mathcal{F}_{(i-1)q})\|_2 \\ & \leq \|\mathbb{E}(h_1(\mathbf{A}_i) - h_1(\mathbf{A}'_i) | \mathcal{F}_{iq})\|_2 + \|\mathbb{E}(h_1(\mathbf{A}_i) - h_1(\mathbf{A}'_i) | \mathcal{F}_{(i-1)q})\|_2 \\ & \leq 2\|h_1(\mathbf{A}_i) - h_1(\mathbf{A}'_i)\|_2. \end{aligned} \quad (4.18)$$

Similarly as in (4.14), we have

$$\begin{aligned} |h_1(\mathbf{A}_i) - h_1(\mathbf{A}'_i)| & \leq \frac{2}{vN} \text{Rank}(A(\mathbf{A}_i) - A(\mathbf{A}'_i)) \leq \frac{2}{vN} \sum_{\ell=(i+1)q+1}^n \mathbf{1}_{\{\mathbf{X}'_{\ell,m} \neq \mathbf{X}_{\ell,m}\}} \\ & \leq \frac{2n}{vN} \mathbf{1}_{\{\varepsilon'_k \neq \varepsilon_k, \text{ for some } k \geq (i+1)qN+1-m\}}. \end{aligned}$$

Hence by (4.17), we infer that

$$\|h_1(\mathbf{A}_i) - h_1(\mathbf{A}'_i)\|_2^2 \leq \frac{4n^2}{v^2 N^2} \beta_{qN+1-2m} \leq \frac{4n^2}{v^2 N^2} \beta_{(q-1)N}. \quad (4.19)$$

Starting from (4.16) together with (4.18) and (4.19), it follows that

$$\mathbb{P}\left(\left|\sum_{i=1}^{[n/q]-1} (\mathbb{E}(h_1(\mathbf{A}_i) | \mathcal{F}_{iq}) - \mathbb{E}(h_1(\mathbf{A}_i) | \mathcal{F}_{(i-1)q}))\right| > x\right) \leq \frac{16n^3}{x^2 v^2 q N^2} \beta_{(q-1)N}. \quad (4.20)$$

Therefore, considering (4.13) and gathering the upper bounds in (4.15) and (4.20), we get

$$\mathbb{P}(|\Re(S_{\mathbf{B}_{n,m}}(z)) - \mathbb{E} \Re(S_{\mathbf{B}_{n,m}}(z))| > 2x) \leq 2 \exp \left\{ -\frac{x^2 v^2 N^2}{128 q n} \right\} + \frac{16n^3}{x^2 v^2 q N^2} \beta_{(q-1)N}.$$

Finally, noting that $\mathbb{P}(|\Im(S_{\mathbf{B}_{n,m}}(z)) - \mathbb{E} \Im(S_{\mathbf{B}_{n,m}}(z))| > 2x)$ also admits the same upper bound and choosing $q = [n/(\log n)^\alpha] + 1$, (4.10) follows. This ends the proof of Theorem 2.1.

5. Proof of Theorem 2.2

The proof, being technical, will be divided into three major steps (Sections 5.1 to 5.3).

5.1. A first approximation

Let m be a fixed positive integer and set $p := p(m) = a_m m$ with $(a_n)_{n \geq 1}$ being a sequence of positive integers such that $\lim_{n \rightarrow \infty} a_n = \infty$. Setting $k_N = \left\lfloor \frac{N}{p+3m} \right\rfloor$, we write the subset $\{1, \dots, Nn\}$ as a union of disjoint subsets of \mathbb{N} as follows:

$$[1, Nn] \cap \mathbb{N} = \bigcup_{i=1}^n [(i-1)N + 1, iN] \cap \mathbb{N} = \bigcup_{i=1}^n \bigcup_{\ell=1}^{k_N+1} I_\ell^i \cup J_\ell^i,$$

where, for $i \in \{1, \dots, n\}$ and $\ell \in \{1, \dots, k_N\}$,

$$\begin{aligned} I_\ell^i &:= [(i-1)N + (\ell-1)(p+3m) + 1, (i-1)N + (\ell-1)(p+3m) + p] \cap \mathbb{N}, \\ J_\ell^i &:= [(i-1)N + (\ell-1)(p+3m) + p + 1, (i-1)N + \ell(p+3m)] \cap \mathbb{N}, \end{aligned}$$

and, for $\ell = k_N + 1$, $I_{k_N+1}^i = \emptyset$ and

$$J_{k_N+1}^i = [(i-1)N + k_N(p+3m) + 1, iN] \cap \mathbb{N}.$$

Note that for all $i \in \{1, \dots, n\}$, $J_{k_N+1}^i = \emptyset$ if $k_N(p+3m) = N$. Now, let M be a fixed positive number not depending on (n, m) and let φ_M be the function defined by $\varphi_M(x) = (x \wedge M) \vee (-M)$. Setting

$$B_{i,\ell} = (\varepsilon_{(i-1)N + (\ell-1)(p+3m) + 1 - m}, \dots, \varepsilon_{(i-1)N + (\ell-1)(p+3m) + p + m}), \quad (5.1)$$

we define the sequences $(\tilde{X}_{k,m,M})_{k \geq 1}$ and $(\bar{X}_{k,m,M})_{k \geq 1}$ as follows:

$$\tilde{X}_{k,m,M} = \begin{cases} \mathbb{E}(\varphi_M(X_k) | B_{i,\ell}) & \text{if } k \in I_\ell^i \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\bar{X}_{k,m,M} = \tilde{X}_{k,m,M} - \mathbb{E}(\tilde{X}_{k,m,M}). \quad (5.2)$$

To soothe the notations, we shall write $\tilde{X}_{k,m}$ and $\bar{X}_{k,m}$ instead of $\tilde{X}_{k,m,M}$ and $\bar{X}_{k,m,M}$ respectively. Note that for any $k \geq 1$,

$$\|\bar{X}_{k,m}\|_2 \leq 2\|\tilde{X}_{k,m}\|_2 = 2\|\mathbb{E}(\varphi_M(X_k) | B_{i,\ell})\|_2 \leq 2\|\varphi_M(X_k)\|_2 \leq 2\|X_k\|_2 = 2\|X_0\|_2, \quad (5.3)$$

and

$$\|\bar{X}_{k,m}\|_\infty \leq 2\|\tilde{X}_{k,m}\|_\infty \leq 2M, \quad (5.4)$$

where the last equality in (5.3) follows from the stationarity of $(X_k)_k$. As $\bar{X}_{k,m}$ is $\sigma(B_{i,\ell})$ -measurable then it can be written as a measurable function h_k of $B_{i,\ell}$, i.e.

$$\bar{X}_{k,m} = h_k(B_{i,\ell}). \quad (5.5)$$

Finally, let $\bar{\mathcal{X}}_{n,m} = ((\bar{X}_{n,m})_{i,j}) = (\bar{X}_{(j-1)N+i,m})$ and set

$$\bar{\mathbf{B}}_{n,m} = \frac{1}{n} \bar{\mathcal{X}}_{n,m} \bar{\mathcal{X}}_{n,m}^T. \quad (5.6)$$

We shall approximate \mathbf{B}_n by $\bar{\mathbf{B}}_{n,m}$ by applying the following proposition:

Proposition 5.1. *Let \mathbf{B}_n and $\bar{\mathbf{B}}_{n,m}$ be the matrices defined in (2.3) and (5.6) respectively then if $\lim_{n \rightarrow \infty} N/n = c \in (0, \infty)$, we have for any $z \in \mathbb{C}_+$,*

$$\lim_{m \rightarrow +\infty} \limsup_{M \rightarrow +\infty} \limsup_{n \rightarrow +\infty} |\mathbb{E}(S_{\mathbf{B}_n}(z)) - \mathbb{E}(S_{\bar{\mathbf{B}}_{n,m}}(z))| = 0. \quad (5.7)$$

Proof. By (4.5) and Cauchy–Schwarz’s inequality, it follows that

$$\begin{aligned} & |\mathbb{E}(S_{\mathbf{B}_n}(z)) - \mathbb{E}(S_{\bar{\mathbf{B}}_{n,m}}(z))| \\ & \leq \frac{\sqrt{2}}{v^2} \left\| \frac{1}{N} \text{Tr}(\mathbf{B}_n + \bar{\mathbf{B}}_{n,m}) \right\|_1^{1/2} \left\| \frac{1}{Nn} \text{Tr}(\mathcal{X}_n - \bar{\mathcal{X}}_{n,m})(\mathcal{X}_n - \bar{\mathcal{X}}_{n,m})^T \right\|_1^{1/2}. \end{aligned} \quad (5.8)$$

By the definition of \mathbf{B}_n , $N^{-1}\mathbb{E}|\text{Tr}(\mathbf{B}_n)| = \|X_0\|_2^2$. Similarly and due to the fact that $pk_N \leq N$ and (5.3),

$$\frac{1}{N} \mathbb{E}|\text{Tr}(\bar{\mathbf{B}}_{n,m})| = \frac{1}{Nn} \sum_{i=1}^n \sum_{\ell=1}^{k_N} \sum_{k \in I_\ell^i} \|\bar{X}_{k,m}\|_2^2 \leq 4\|X_0\|_2^2. \quad (5.9)$$

Moreover, by the construction of \mathcal{X}_n and $\bar{\mathcal{X}}_{n,m}$, we have

$$\frac{1}{Nn} \mathbb{E}|\text{Tr}(\mathcal{X}_n - \bar{\mathcal{X}}_{n,m})(\mathcal{X}_n - \bar{\mathcal{X}}_{n,m})^T| = \frac{1}{Nn} \sum_{i=1}^n \sum_{\ell=1}^{k_N} \sum_{k \in I_\ell^i} \|X_k - \bar{X}_{k,m}\|_2^2 + \frac{1}{Nn} \sum_{i=1}^n \sum_{\ell=1}^{k_N+1} \sum_{k \in J_\ell^i} \|X_k\|_2^2.$$

Now, since X_k is centered, we write for $k \in I_\ell^i$,

$$\begin{aligned} \|X_k - \bar{X}_{k,m}\|_2 &= \|X_k - \tilde{X}_{k,m} - \mathbb{E}(X_k - \tilde{X}_{k,m})\|_2 \leq 2\|X_k - \tilde{X}_{k,m}\|_2 \\ &\leq 2\|X_k - \mathbb{E}(X_k|B_{i,\ell})\|_2 + 2\|\tilde{X}_{k,m} - \mathbb{E}(X_k|B_{i,\ell})\|_2. \end{aligned} \quad (5.10)$$

Analyzing the second term of the last inequality, we get

$$\|\tilde{X}_{k,m} - \mathbb{E}(X_k|B_{i,\ell})\|_2 = \|\mathbb{E}(X_k - \varphi_M(X_k)|B_{i,\ell})\|_2 \leq \|X_k - \varphi_M(X_k)\|_2 = \|(|X_0| - M)_+\|_2. \quad (5.11)$$

As X_0 belongs to \mathbb{L}^2 , then $\lim_{M \rightarrow +\infty} \|(|X_0| - M)_+\|_2 = 0$. Now, we note that for $k \in I_\ell^i$, $\sigma(\varepsilon_{k-m}, \dots, \varepsilon_{k+m}) \subset \sigma(B_{i,\ell})$ which implies that

$$\begin{aligned} \|X_k - \mathbb{E}(X_k|B_{i,\ell})\|_2 &\leq \|X_k - \mathbb{E}(X_k|\varepsilon_{k-m}, \dots, \varepsilon_{k+m})\|_2 \\ &= \|X_0 - \mathbb{E}(X_0|\varepsilon_{-m}, \dots, \varepsilon_m)\|_2 = \|X_0 - X_{0,m}\|_2, \end{aligned} \quad (5.12)$$

where the first equality is due to the stationarity. Therefore, by (5.11), (5.12), the fact that $pk_N \leq N$ and

$$\text{Card}\left(\bigcup_{i=1}^n \bigcup_{\ell=1}^{k_N+1} J_\ell^i\right) \leq Nn - npk_N,$$

we infer that

$$\begin{aligned} \frac{1}{Nn} \mathbb{E}|\text{Tr}(\mathcal{X}_n - \bar{\mathcal{X}}_{n,m})(\mathcal{X}_n - \bar{\mathcal{X}}_{n,m})^T| &\leq 8\|X_0 - X_{0,m}\|_2^2 + 8\|(|X_0| - M)_+\|_2^2 \\ &\quad + (3(a_m + 3)^{-1} + a_m m N^{-1})\|X_0\|_2^2. \end{aligned} \quad (5.13)$$

Thus starting from (5.8), considering the upper bounds (5.9) and (5.13), we derive that there exists a positive constant C not depending on (n, m, M) such that

$$\limsup_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} |\mathbb{E}(S_{\mathbf{B}_n}(z)) - \mathbb{E}(S_{\mathbf{B}_{n,m}}(z))| \leq \frac{C}{v^2} \left(\|X_0 - X_{0,m}\|_2^2 + \frac{3}{a_m} \right)^{1/2}.$$

Taking the limit on m , Proposition 5.1 follows by applying the martingale convergence theorem in \mathbb{L}^2 and that fact that a_m converges to infinity. \square

5.2. Approximation by a Gram matrix with independent blocks

By Berbee's classical coupling lemma [5], one can construct by induction a sequence of random variables $(\varepsilon_k^*)_{k \geq 1}$ such that:

- for any $1 \leq i \leq n$ and $1 \leq \ell \leq k_N$,

$$B_{i,\ell}^* = (\varepsilon_{(i-1)N+(\ell-1)(p+3m)+1-m}^*, \dots, \varepsilon_{(i-1)N+(\ell-1)(p+3m)+p+m}^*)$$

has the same distribution as $B_{i,\ell}$ defined in (5.1),

- the array $(B_{i,\ell}^*)_{1 \leq i \leq n, 1 \leq \ell \leq k_N}$ is i.i.d.,
- for any $1 \leq i \leq n$ and $1 \leq \ell \leq k_N$, $\mathbb{P}(B_{i,\ell} \neq B_{i,\ell}^*) \leq \beta_m$.

We refer to page 484 of [15] for more details concerning the construction of the array $(B_{i,\ell}^*)_{i,\ell \geq 1}$. We define now the sequence $(\bar{X}_{k,m}^*)_{k \geq 1}$ as follows:

$$\bar{X}_{k,m}^* = h_k(B_{i,\ell}^*) \quad \text{if } k \in I_\ell^i, \quad (5.14)$$

where the functions h_k are defined in (5.5).

We construct the $N \times n$ random matrix $\bar{\mathcal{X}}_{n,m}^* = ((\bar{\mathcal{X}}_{n,m}^*)_{i,j}) = (\bar{X}_{(j-1)N+i,m}^*)$. Note that the block of entries $(\bar{X}_{k,m}^*, k \in I_\ell^i)$ is independent of $(\bar{X}_{k,m}^*, k \in I_{\ell'}^{i'})$ if $(i, \ell) \neq (i', \ell')$. Thus, $\bar{\mathcal{X}}_{n,m}^*$ has independent blocks of dimension p separated by null blocks whose dimension is at least $3m$. Setting

$$\bar{\mathbf{B}}_{n,m}^* := \frac{1}{n} \bar{\mathcal{X}}_{n,m}^* \bar{\mathcal{X}}_{n,m}^{*T}, \quad (5.15)$$

we approximate $\bar{\mathbf{B}}_{n,m}$ by the Gram matrix $\bar{\mathbf{B}}_{n,m}^*$ as shown in the following proposition.

Proposition 5.2. *Let $\bar{\mathbf{B}}_{n,m}$ and $\bar{\mathbf{B}}_{n,m}^*$ be defined in (5.6) and (5.15) respectively. Assuming that $\lim_{n \rightarrow \infty} N/n = c \in (0, \infty)$ and $\lim_{n \rightarrow \infty} \beta_n = 0$ then for any $z \in \mathbb{C}_+$,*

$$\lim_{m \rightarrow +\infty} \limsup_{M \rightarrow +\infty} \limsup_{n \rightarrow +\infty} |\mathbb{E}(S_{\bar{\mathbf{B}}_{n,m}}(z)) - \mathbb{E}(S_{\bar{\mathbf{B}}_{n,m}^*}(z))| = 0. \quad (5.16)$$

Proof. By (4.5) and Cauchy–Schwarz's inequality, it follows that

$$\begin{aligned} & |\mathbb{E}(S_{\bar{\mathbf{B}}_{n,m}}(z)) - \mathbb{E}(S_{\bar{\mathbf{B}}_{n,m}^*}(z))| \\ & \leq \frac{\sqrt{2}}{v^2} \left\| \frac{1}{N} \text{Tr}(\bar{\mathbf{B}}_{n,m} + \bar{\mathbf{B}}_{n,m}^*) \right\|_1^{1/2} \left\| \frac{1}{Nn} \text{Tr}(\bar{\mathcal{X}}_{n,m} - \bar{\mathcal{X}}_{n,m}^*)(\bar{\mathcal{X}}_{n,m} - \bar{\mathcal{X}}_{n,m}^*)^T \right\|_1^{1/2}. \end{aligned} \quad (5.17)$$

Notice that $\|\bar{X}_{k,m}^*\|_2 = \|h_k(B_{i,\ell}^*)\|_2 = \|h_k(B_{i,\ell})\|_2 = \|\bar{X}_{k,m}\|_2 \leq 2\|X_0\|_2$, where the second equality follows from the fact that $B_{i,\ell}^*$ is distributed as $B_{i,\ell}$ whereas the last inequality follows from (5.3). Thus, we get from the definition of $\mathbf{B}_{n,m}^*$ and the fact that $pk_N \leq N$,

$$\frac{1}{N} \mathbb{E} |\text{Tr}(\bar{\mathbf{B}}_{n,m}^*)| = \frac{1}{Nn} \sum_{i=1}^n \sum_{\ell=1}^{k_N} \sum_{k \in I_\ell^i} \|\bar{X}_{k,m}^*\|_2^2 \leq 4\|X_0\|_2^2. \quad (5.18)$$

Considering (5.17), (5.9) and (5.18), we infer that Proposition 5.2 follows once we prove that

$$\lim_{m \rightarrow +\infty} \limsup_{M \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \frac{1}{Nn} \mathbb{E} |\text{Tr}(\bar{\mathcal{X}}_{n,m} - \bar{\mathcal{X}}_{n,m}^*)(\bar{\mathcal{X}}_{n,m} - \bar{\mathcal{X}}_{n,m}^*)^T| = 0. \quad (5.19)$$

By the construction of $\bar{\mathcal{X}}_{n,m}$ and $\bar{\mathcal{X}}_{n,m}^*$, we write

$$\frac{1}{Nn} \mathbb{E} |\text{Tr}(\bar{\mathcal{X}}_{n,m} - \bar{\mathcal{X}}_{n,m}^*)(\bar{\mathcal{X}}_{n,m} - \bar{\mathcal{X}}_{n,m}^*)^T| = \frac{1}{Nn} \sum_{i=1}^n \sum_{\ell=1}^{k_N} \sum_{k \in I_\ell^i} \|\bar{X}_{k,m} - \bar{X}_{k,m}^*\|_2^2. \quad (5.20)$$

Now, let L be a fixed positive real number strictly less than M and not depending on (n, m, M) . To control the term $\|\bar{X}_{k,m} - \bar{X}_{k,m}^*\|_2^2$, we write for $k \in I_\ell^i$,

$$\begin{aligned} \|\bar{X}_{k,m} - \bar{X}_{k,m}^*\|_2^2 &= \|(h_k(B_{i,\ell}) - h_k(B_{i,\ell}^*))\mathbf{1}_{B_{i,\ell} \neq B_{i,\ell}^*}\|_2^2 \\ &\leq 4\|h_k(B_{i,\ell})\mathbf{1}_{B_{i,\ell} \neq B_{i,\ell}^*}\|_2^2 = 4\|\bar{X}_{k,m}\mathbf{1}_{B_{i,\ell} \neq B_{i,\ell}^*}\|_2^2 \\ &\leq 12\|\bar{X}_{k,m} - \mathbb{E}(X_k|B_{i,\ell})\|_2^2 + 12\|\mathbb{E}(X_k|B_{i,\ell}) - \mathbb{E}(\varphi_L(X_k)|B_{i,\ell})\|_2^2 \\ &\quad + 12\|\mathbb{E}(\varphi_L(X_k)|B_{i,\ell})\mathbf{1}_{B_{i,\ell} \neq B_{i,\ell}^*}\|_2^2. \end{aligned}$$

Since $\mathbb{P}(B_{i,\ell} \neq B_{i,\ell}^*) \leq \beta_m$ and $\varphi_L(X_k)$ is bounded by L , we get

$$\|\mathbb{E}(\varphi_L(X_k)|B_{i,\ell})\mathbf{1}_{B_{i,\ell} \neq B_{i,\ell}^*}\|_2^2 \leq L^2\beta_m.$$

Moreover, it follows from the fact that X_k is centered and (5.11) that

$$\|\bar{X}_{k,m} - \mathbb{E}(X_k|B_{i,\ell})\|_2^2 \leq 4\|\tilde{X}_{k,m} - \mathbb{E}(X_k|B_{i,\ell})\|_2^2 \leq 4\|(|X_0| - M)_+\|_2^2$$

and

$$\|\mathbb{E}(X_k|B_{i,\ell}) - \mathbb{E}(\varphi_L(X_k)|B_{i,\ell})\|_2^2 \leq \|X_k - \varphi_L(X_k)\|_2^2 = \|(|X_0| - L)_+\|_2^2.$$

Hence gathering the above upper bounds we get

$$\|\bar{X}_{k,m} - \bar{X}_{k,m}^*\|_2^2 \leq 48\|(|X_0| - M)_+\|_2^2 + 12\|(|X_0| - L)_+\|_2^2 + 12L^2\beta_m. \quad (5.21)$$

As $pk_N \leq N$, the right-hand side of (5.20) converges to zero by letting first M , then m and finally L tend to infinity. Therefore, (5.19) and thus the proposition follow. \square

5.3. Approximation with a Gaussian matrix

In order to complete the proof of the theorem, it suffices, in view of (5.7) and (5.16), to prove the following convergence: for any $z \in \mathbb{C}_+$,

$$\lim_{m \rightarrow +\infty} \limsup_{n \rightarrow +\infty} |\mathbb{E}(S_{\bar{\mathbf{B}}_{n,m}^*}(z)) - \mathbb{E}(S_{\mathbf{G}_n}(z))| = 0. \quad (5.22)$$

With this aim, we shall first consider a sequence $(Z_{k,m})_{k \in \mathbb{Z}}$ of centered Gaussian random variables such that for any $k, \ell \in \{1, \dots, N\}$,

$$\text{Cov}(Z_{k,m}, Z_{\ell,m}) = \text{Cov}(\bar{X}_{k,m}^*, \bar{X}_{\ell,m}^*) \quad (5.23)$$

and let $(Z_{k,m}^{(i)})_k$, $i = 1, \dots, n$, be n independent copies of $(Z_{k,m})_k$. We then define the $N \times n$ matrix $\mathcal{Z}_{n,m} = ((\mathcal{Z}_{n,m})_{u,v}) = (Z_{u,m}^{(v)})$ and finally set

$$\mathbf{G}_{n,m} = \frac{1}{n} \mathcal{Z}_{n,m} \mathcal{Z}_{n,m}^T. \quad (5.24)$$

Now, we shall construct a matrix $\tilde{\mathcal{Z}}_{n,m}$ having the same block structure as the matrix $\bar{\mathcal{X}}_{n,m}^*$. With this aim, we let for $\ell = 1, \dots, k_N$,

$$I_\ell = \{(\ell - 1)(p + 3m) + 1, \dots, (\ell - 1)(p + 3m) + p\}$$

and let $\tilde{Z}_{k,m}^{(i)}$ be defined for any $1 \leq i \leq n$ and $1 \leq k \leq N$ by

$$\tilde{Z}_{k,m}^{(i)} = \begin{cases} Z_{k,m}^{(i)} & \text{if } k \in I_\ell \text{ for some } \ell \in \{1, \dots, k_N\} \\ 0 & \text{otherwise.} \end{cases}$$

We define now the $N \times n$ matrix $\tilde{\mathcal{Z}}_{n,m} = ((\tilde{\mathcal{Z}}_{n,m})_{u,v}) = (\tilde{Z}_{u,m}^{(v)})$ and we note that $\tilde{\mathcal{Z}}_{n,m}$, as $\bar{\mathcal{X}}_{n,m}^*$, consists of independent blocks of dimension p separated by null blocks whose dimension is at least $3m$. We finally set

$$\tilde{\mathbf{G}}_{n,m} = \frac{1}{n} \tilde{\mathcal{Z}}_{n,m} \tilde{\mathcal{Z}}_{n,m}^T. \quad (5.25)$$

Provided that $\lim_{n \rightarrow \infty} n/N = c \in (0, \infty)$, we have by Proposition 4.2 in [3] that for any $z \in \mathbb{C}_+$,

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} |\mathbb{E}(S_{\mathbf{G}_{n,m}}(z)) - \mathbb{E}(S_{\tilde{\mathbf{G}}_{n,m}}(z))| = 0.$$

In order to prove (5.22), we shall prove for any $z \in \mathbb{C}_+$,

$$\lim_{n \rightarrow +\infty} |\mathbb{E}(S_{\bar{\mathbf{B}}_{n,m}^*}(z)) - \mathbb{E}(S_{\tilde{\mathbf{G}}_{n,m}}(z))| = 0$$

and then

$$\lim_{m \rightarrow \infty} \limsup_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} |\mathbb{E}(S_{\mathbf{G}_n}(z)) - \mathbb{E}(S_{\mathbf{G}_{n,m}}(z))| = 0.$$

The technique followed to prove the first convergence is based on Lindeberg's method by blocks, whereas, for the second, it is based on the Gaussian interpolation technique.

Proposition 5.3. *Provided that $N/n \rightarrow c \in (0, \infty)$, then for any $z = x + iy \in \mathbb{C}_+$,*

$$\lim_{n \rightarrow +\infty} |\mathbb{E}(S_{\tilde{\mathbf{B}}_{n,m}^*}(z)) - \mathbb{E}(S_{\tilde{\mathbf{G}}_{n,m}}(z))| = 0. \quad (5.26)$$

Sketch of the proof. We don't give a full proof of this proposition because the computation involved has been almost done in the proof of Proposition 4.3 of [3]. Indeed as $\tilde{\mathcal{X}}_{n,m}^*$, the matrix $\tilde{\mathcal{X}}_n$ considered in [3] has independent blocks separated by blocks of zero entries. The main difference is that in our case even the Gaussian blocks are mutually independent and thus the terms of first and second order in the Taylor expansion vanish simplifying the proof even more. Therefore, a careful analysis of the proof of Proposition 4.3 in [3] gives for any $z = x + iy \in \mathbb{C}_+$,

$$|\mathbb{E}(S_{\tilde{\mathbf{B}}_{n,m}^*}(z)) - \mathbb{E}(S_{\tilde{\mathbf{G}}_{n,m}}(z))| \leq \frac{Cp^2(1+M^3)N^{1/2}}{y^3(1 \wedge y)n},$$

which converges to 0 as n tends to infinity since $\lim_{n \rightarrow \infty} N/n = c \in (0, \infty)$. \square

To end the proof of Theorem 2.2, it remains to prove the following convergence.

Proposition 5.4. *Provided that $\lim_{n \rightarrow \infty} N/n = c \in (0, \infty)$ and $\lim_{n \rightarrow \infty} \beta_n = 0$ then for any $z \in \mathbb{C}_+$,*

$$\lim_{m \rightarrow \infty} \limsup_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} |\mathbb{E}(S_{\mathbf{G}_n}(z)) - \mathbb{E}(S_{\mathbf{G}_{n,m}}(z))| = 0.$$

Proof. Let $n' = N + n$ and $\mathbb{G}_{n'}$ be the symmetric matrix of order n defined by

$$\mathbb{G}_{n'} = \frac{1}{\sqrt{n}} \begin{pmatrix} \mathbf{0} & \mathbf{Z}_n^T \\ \mathbf{Z}_n & \mathbf{0} \end{pmatrix}.$$

It is well known that for any $z \in \mathbb{C}^+$,

$$S_{\mathbf{G}_n}(z) = z^{-1/2} \frac{n}{2N} S_{\mathbb{G}_{n'}}(z^{1/2}) + \frac{n-N}{2Nz}.$$

We refer, for instance, to page 549 in Rashidi Far et al. [14] for arguments leading to the relation above. Since the same relation also holds for the symmetric matrix $\mathbb{G}_{n',m}$ associated with $\mathbf{G}_{n,m}$ and since $n'/N \rightarrow 1 + c^{-1}$, it is equivalent to prove for any $z \in \mathbb{C}^+$,

$$\lim_{m \rightarrow \infty} \limsup_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} |\mathbb{E}(S_{\mathbb{G}_{n'}}(z)) - \mathbb{E}(S_{\mathbb{G}_{n',m}}(z))| = 0.$$

This can be proved in the same way as the convergence (32) in [4]. Indeed, noting that $(\mathbb{G}_{n'})_{k,\ell} = n^{-1/2} Z_{k-n}^{(\ell)} \mathbf{1}_{k>n} \mathbf{1}_{\ell \leq n}$ and $(\mathbb{G}_{n',m})_{k,\ell} = n^{-1/2} Z_{k-n,m}^{(\ell)} \mathbf{1}_{k>n} \mathbf{1}_{\ell \leq n}$ if $1 \leq \ell \leq k \leq n'$ and keeping in mind the independence structure between the columns, we apply Lemma 16 in [4] and we get for any $z \in \mathbb{C}_+$,

$$\begin{aligned} & \mathbb{E}(S_{\mathbb{G}_{n'}}(z)) - \mathbb{E}(S_{\mathbb{G}_{n',m}}(z)) \\ &= \frac{n'}{2n} \sum_{j=1}^n \sum_{k,\ell=n+1}^{n+N} \int_0^1 \left(\mathbb{E}(Z_{k-n} Z_{\ell-n}) - \mathbb{E}(Z_{k-n,m} Z_{\ell-n,m}) \right) \mathbb{E} \left(\frac{\partial^2}{\partial x_{k,j} \partial x_{\ell,j}} f(\mathbf{g}(t)) \right) dt \\ &= \frac{N+n}{2n} \sum_{j=1}^n \sum_{k,\ell=1}^N \int_0^1 \left(\mathbb{E}(Z_k Z_\ell) - \mathbb{E}(Z_{k,m} Z_{\ell,m}) \right) \mathbb{E} \left(\frac{\partial^2}{\partial x_{k+n,j} \partial x_{\ell+n,j}} f(\mathbf{g}(t)) \right) dt \end{aligned} \quad (5.27)$$

where, for $t \in [0, 1]$, $\mathbf{g}(t) = \sqrt{N+n}(\sqrt{t}(\mathbb{G}_{n'})_{k,\ell} + \sqrt{1-t}(\mathbb{G}_{n',m})_{k,\ell})_{1 \leq \ell \leq k \leq n'}$ and f is the function that allows us to write the Stieltjes transform of a symmetric matrix in terms of its entries (for a precise definition, see (49) in [4]). Then, by using (2.5) and (5.23), we write the following decomposition

$$\begin{aligned}\mathbb{E}(Z_k Z_\ell) - \mathbb{E}(Z_{k,m} Z_{\ell,m}) &= \mathbb{E}(X_k X_\ell) - \mathbb{E}(\bar{X}_{k,m}^* \bar{X}_{\ell,m}^*) \\ &= \mathbb{E}(X_k (X_\ell - \bar{X}_{\ell,m}^*)) + \mathbb{E}(\bar{X}_{\ell,m}^* (X_k - \bar{X}_{k,m}^*)). \end{aligned} \quad (5.28)$$

We shall decompose the left-hand side of (5.27) into two sums according to the decomposition (5.28) and treat them separately. With this aim, we shall use Lemma 13 by Merlevède and Peligrad [12] which is very useful for controlling the second order partial derivatives of the Stieltjes transform of symmetric matrices. Applying this lemma, we get for any two sequences $(a_k)_k$ and $(b_k)_k$ of real numbers

$$\sum_{j=1}^n \left| \sum_{k,\ell=1}^N a_k b_\ell \frac{\partial^2}{\partial x_{k+n,j} \partial x_{\ell+n,j}} f(\mathbf{g}(t)) \right| \leq \frac{C}{N+n} \left(\sum_{k=1}^N a_k^2 \sum_{\ell=1}^N b_\ell^2 \right)^{1/2}, \quad (5.29)$$

where C is a universal constant depending only on the imaginary part of z and might change from a line to another. Then, applying (5.29) with $a_k = X_k$ and $b_\ell = X_\ell - \bar{X}_{\ell,m}^*$ and another time with $a_k = X_k - \bar{X}_{k,m}^*$ and $b_\ell = \bar{X}_{\ell,m}^*$, we get for any $z \in \mathbb{C}_+$,

$$\left| \mathbb{E}(S_{\mathbb{G}_{n'}}(z)) - \mathbb{E}(S_{\mathbb{G}_{n',m}}(z)) \right|^2 \leq \frac{C}{n^2} \sum_{k=1}^N (\|X_k\|_2^2 + \|\bar{X}_{k,m}^*\|_2^2) \sum_{\ell=1}^N \|X_\ell - \bar{X}_{\ell,m}^*\|_2^2. \quad (5.30)$$

In view of (5.10), (5.11), (5.12) and (5.21), we get

$$\begin{aligned}\|X_\ell - \bar{X}_{\ell,m}^*\|^2 &\leq 2\|X_\ell - \bar{X}_{\ell,m}\|^2 + 2\|\bar{X}_{\ell,m} - \bar{X}_{\ell,m}^*\|^2 \\ &\leq 16\|X_0 - X_{0,m}\|_2^2 + 112\|(|X_0| - M)_+\|_2^2 + 24\|(|X_0| - L)_+\|_2^2 + 24L^2\beta_m, \end{aligned}$$

where L is a fixed positive real number strictly less than M and not depending on (n, m, M) . Taking into account the stationarity of $(X_k)_k$ and (5.3), we then infer that

$$\left| \mathbb{E}(S_{\mathbb{G}_{n'}}(z)) - \mathbb{E}(S_{\mathbb{G}_{n',m}}(z)) \right|^2 \leq \frac{CN^2}{n^2} \left(\|X_0 - X_{0,m}\|_2^2 + \|(|X_0| - M)_+\|_2^2 + \|(|X_0| - L)_+\|_2^2 + L^2\beta_m \right),$$

which converges to zero by letting first n , then M followed by m and finally L tend to infinity. This ends the proof of the proposition. \square

Acknowledgments

I would like to thank the referee for carefully reading the manuscript and my thesis advisor Florence Merlevède who was generous with her time, knowledge and assistance. I would also like to thank my co-advisor Emmanuel Rio for valuable discussions.

References

- [1] Z. Bai, J.W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, second ed., Springer Ser. Statist., Springer, New York, 2010.
- [2] Z. Bai, W. Zhou, Large sample covariance matrices without independence structures in columns, *Statist. Sinica* 18 (2008) 425–442.
- [3] M. Banna, F. Merlevède, Limiting spectral distribution of large sample covariance matrices associated with a class of stationary processes, *J. Theoret. Probab.* (2013) 1–39.

- [4] M. Banna, F. Merlevède, M. Peligrad, On the limiting spectral distribution for a large class of symmetric random matrices with correlated entries, *Stochastic Process. Appl.* 125 (7) (2015) 2700–2726.
- [5] H.C.P. Berbee, *Random Walks with Stationary Increments and Renewal Theory*, Math. Centre Tracts, vol. 112, Mathematisch Centrum, Amsterdam, 1979.
- [6] S. Borovkova, R. Burton, H. Dehling, Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation, *Trans. Amer. Math. Soc.* 353 (11) (2001) 4261–4318.
- [7] P. Doukhan, P. Massart, E. Rio, The functional central limit theorem for strongly mixing processes, *Ann. Inst. Henri Poincaré Probab. Stat.* 30 (1994) 63–82.
- [8] A. Guntuboyina, H. Leeb, Concentration of the spectral measure of large Wishart matrices with dependent entries, *Electron. Commun. Probab.* 14 (2009) 334–342.
- [9] P. Hall, C.C. Heyde, *Martingale Limit Theory and Its Application*, Probab. Math. Statist., Academic Press, New York, London, 1980.
- [10] F. Hofbauer, G. Keller, Ergodic properties of invariant measures for piecewise monotonic transformations, in: *The Theory of Chaotic Attractors*, Springer, New York, 1982, pp. 120–141.
- [11] V. Marčenko, L. Pastur, Distribution of eigenvalues for some sets of random matrices, *Mat. Sb.* 72 (114) (1967) 507–536.
- [12] F. Merlevède, M. Peligrad, On the empirical spectral distribution for matrices with long memory and independent rows, [arXiv:1406.1216](https://arxiv.org/abs/1406.1216), 2014.
- [13] O. Pfaffel, E. Schlemm, Limiting spectral distribution of a new random matrix model with dependence across rows and columns, *Linear Algebra Appl.* 436 (9) (2012) 2966–2979.
- [14] R. Rashidi Far, T. Oraby, W. Bryc, R. Speicher, On slow-fading MIMO systems with nonseparable correlation, *IEEE Trans. Inform. Theory* 54 (2008) 544–553.
- [15] G. Viennet, Inequalities for absolutely regular sequences: application to density estimation, *Probab. Theory Related Fields* 107 (1997) 467–492.
- [16] V.A. Volkonskii, Y.A. Rozanov, Some limit theorems for random functions. I, *Theory Probab. Appl.* 4 (1959) 178–197.
- [17] J. Yao, A note on a Marčenko–Pastur type theorem for time series, *Statist. Probab. Lett.* 82 (2012) 22–28.
- [18] Y.Q. Yin, Limiting spectral distribution for a class of random matrices, *J. Multivariate Anal.* 20 (1986) 50–68.