On the limiting spectral distribution for a large class of symmetric random matrices with correlated entries.

Marwa Banna[a], Florence Merlevède[b], Magda Peligrad[b1]

[a b] Université Paris Est, LAMA (UMR 8050), UPEMLV, CNRS, UPEC, 5 Boulevard Descartes, 77 454 Marne La Vallée, France.

E-mail: marwa.banna@u-pem.fr; florence.merlevede@u-pem.fr

[c] Department of Mathematical Sciences, University of Cincinnati, PO Box 210025, Cincinnati, Oh 45221-0025, USA.

Email: peligrm@ucmail.uc.edu

Key words: random matrices, correlated entries, sample covariance matrices, weak dependence, limiting spectral distribution.

Mathematics Subject Classification (2010): 60F15, 60G60, 60G10, 62E20.

Abstract

For symmetric random matrices with correlated entries, which are functions of independent random variables, we show that the asymptotic behavior of the empirical eigenvalue distribution can be obtained by analyzing a Gaussian matrix with the same covariance structure. This class contains both cases of short and long range dependent random fields. The technique is based on a blend of blocking procedure and Lindeberg's method. This method leads to a variety of interesting asymptotic results for matrices with dependent entries, including applications to linear processes as well as nonlinear Volterra-type processes entries.

1 Introduction

The limiting spectral distribution for symmetric matrices with correlated entries received a lot of attention in the last two decades. The starting point is deep results for symmetric matrices with correlated Gaussian entries by Khorunzhy and Pastur [13], Boutet de Monvel *et al* [6], Boutet de Monvel and Khorunzhy [5], Chakrabarty *et al* [7] among others. On the other hand there is a sustained effort for studying linear filters of independent random variables as entries of a matrix. For instance, Anderson and Zeitouni [1] considered symmetric matrices with entries that are linear processes of finite range of independent random variables. Hachem *et al* [12] considered large sample covariance matrices whose entries are modeled by a short memory linear process of infinite range with independent Gaussian innovations. Bai and Zhou [3], Yao [18], Banna and Merlevède [4] and Merlevède and Peligrad [14], among others, treated large covariance matrices based on an independent structure between columns and correlated random variables in rows.

In this paper we consider symmetric random matrices whose entries are functions of independent and identically distributed (i.i.d.) real-valued random variables. Such kind of processes provide a very general framework for stationary ergodic random fields. Our main goal is to reduce the study of the limiting spectral distribution to the same problem for a Gaussian matrix having the same covariance structure as the underlying process. In this way we prove the universality and we are able to formulate various limiting results for large classes of matrices. We also treat large sample covariance matrices with correlated entries, known under the name of Gram matrices. Our proofs are based on the large-small block arguments, a method which, in one dimensional setting, is going back to Bernstein. Then, we apply a variant of the so-called

[1]Supported in part by a Charles Phelps Taft Memorial Fund grant, and the NSF grant DMS-1208237.

Lindeberg method, namely we develop a block Lindeberg method, where we replace at one time a big block of random variables with a Gaussian one with the same covariance structure. Lindeberg method is popular with these type of problems. Replacing only one variable at one time with a Gaussian one, Chatterjee [8] treated random matrices with exchangeable entries.

Our paper is organized in the following way. Section 2 contains the main results for symmetric random matrices and sample covariance matrices. As an intermediate step we also treat matrices based on K-dependent random fields, results that have interest in themselves (see Theorem 11 in Section 4.1). In Section 3, we give applications to matrices with entries which are either linear random fields or nonlinear random fields as Volterra-type processes. The main proofs are included in Section 4. In Section 5 we prove a concentration of spectral measure inequality for a row-wise K-dependent random matrix and we also mention some of the technical results used in the paper.

Here are some notations used all along the paper. The notation $[x]$ is used to denote the integer part of any real x. For any positive integers a, b, the notation $\mathbf{0}_{a,b}$ means a matrix with 0 entries of size $a \times b$, whereas the notation $\mathbf{0}_a$ means a row vector of size a. For a matrix A, we denote by A^T its transpose matrix and by $\mathrm{Tr}(A)$ its trace. 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 with only real eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$, its spectral empirical measure and its spectral distribution function are respectively defined by

$$\nu_A = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i} \quad \text{and} \quad F_n^A(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{\lambda_k \leq x\}} .$$

The Stieltjes transform of F^A is given by

$$S_A(z) = \int \frac{1}{x-z} dF^A(x) = \frac{1}{n}\mathrm{Tr}(A - z\mathbf{I}_n)^{-1} ,$$

where $z = u + \mathrm{i}v \in \mathbb{C}^+$ (the set of complex numbers with positive imaginary part), and $\mathbf{I}_n$ is the identity matrix of order n.

The Lévy distance between two distribution functions F and G is defined by

$$L(F, G) = \inf\{\varepsilon > 0 \ : \ F(x-\varepsilon) - \varepsilon \leq G(x) \leq F(x+\varepsilon) + \varepsilon\} .$$

It is well-known that a sequence of distribution functions $F_n(x)$ converges to a distribution function $F(x)$ at all continuity points x of F if and only if $L(F_n, F) \to 0$.

2 Main results

2.1 On the limiting distribution for a large class of symmetric matrices with correlated entries

Let $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ be an array of real-valued random variables, and consider its associated symmetric random matrix $\mathbf{X}_n$ of order n defined by

$$\big(\mathbf{X}_n\big)_{i,j} = X_{i,j} \text{ if } 1 \leq j \leq i \leq n \text{ and } \big(\mathbf{X}_n\big)_{i,j} = X_{j,i} \text{ if } 1 \leq i < j \leq n . \tag{1}$$

Define then

$$\mathbb{X}_n := n^{-1/2}\mathbf{X}_n . \tag{2}$$

The aim of this section is to study the limiting spectral empirical distribution function of the symmetric matrix $\mathbb{X}_n$ defined by (2) when the process $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ has the following dependence structure: for any $(k,\ell) \in \mathbb{Z}^2$,

$$X_{k,\ell} = g(\xi_{k-i,\ell-j} \ ; \ (i,j) \in \mathbb{Z}^2) , \tag{3}$$

where $(\xi_{i,j})_{(i,j)\in\mathbb{Z}^2}$ is an array of i.i.d. real-valued random variables given on a common probability space $(\Omega, \mathcal{K}, \mathbb{P})$, and g is a measurable function from $\mathbb{R}^{\mathbb{Z}^2}$ to $\mathbb{R}$ such that $\mathbb{E}(X_{0,0}) = 0$ and $\|X_{0,0}\|_2 < \infty$. A representation as (3) includes linear as well as many widely used nonlinear random fields models as special cases.

Our Theorem 1 below shows a universality scheme for the random matrix $\mathbb{X}_n$ as soon as the entries of the symmetric matrix $\sqrt{n}\mathbb{X}_n$ have the dependence structure (3). It is noteworthy to indicate that this result does not require rate of convergence to zero of the correlation between the entries.

Theorem 1 *Let $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ be a real-valued stationary random field given by (3). Define the symmetric matrix $\mathbb{X}_n$ by (2). Let $(G_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ be a real-valued centered Gaussian random field, with covariance function given by*

$$\mathbb{E}(G_{k,\ell}G_{i,j}) = \mathbb{E}(X_{k,\ell}X_{i,j}) \text{ for any } (k,\ell) \text{ and } (i,j) \text{ in } \mathbb{Z}^2 . \tag{4}$$

Let $\mathbf{G}_n$ be the symmetric random matrix defined by $\big(\mathbf{G}_n\big)_{i,j} = G_{i,j}$ if $1 \le j \le i \le n$ and $\big(\mathbf{G}_n\big)_{i,j} = G_{j,i}$ if $1 \le i < j \le n$. Denote $\mathbb{G}_n = \frac{1}{\sqrt{n}}\mathbf{G}_n$. Then, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty} \big|S_{\mathbb{X}_n}(z) - \mathbb{E}\big(S_{\mathbb{G}_n}(z)\big)\big| = 0 \quad \text{almost surely.}$$

Theorem 1 is important since it shows that the study of the limiting spectral distribution function of a symmetric matrix whose entries are functions of i.i.d. random variables can be reduced to studying the same problem as for a Gaussian matrix with the same covariance structure. The following corollary is a direct consequence of our Theorem 1 together with Theorem B.9 in Bai-Silverstein [2] (see also the arguments on page 38 in [2], based on Vitali's convergence theorem).

Corollary 2 *Assume that $\mathbb{X}_n$ and $\mathbb{G}_n$ are as in Theorem 1. Furthermore, assume there exists a distribution function F such that*

$$\mathbb{E}\big(F^{\mathbb{G}_n}(t)\big) \to F(t) \text{ for all continuity points } t \in \mathbb{R} \text{ of } F .$$

Then

$$\mathbb{P}(L(F^{\mathbb{X}_n(\omega)}, F) \to 0) = 1 . \tag{5}$$

For instance, Corollary 2 above combined with the proof of Theorem 2 in Khorunzhy and Pastur [13] concerning the asymptotic spectral behavior of certain ensembles with correlated Gaussian entries (see also Theorem 17.2.1 in [15]), gives the following:

Theorem 3 *Let $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ be a real-valued stationary random field given by (3). Define the symmetric matrix $\mathbb{X}_n$ by (2). For any $(k,\ell) \in \mathbb{Z}^2$, let $\gamma_{k,\ell} = \mathbb{E}(X_{0,0}X_{k,\ell})$. Assume that*

$$\sum_{k,\ell\in\mathbb{Z}} |\gamma_{k,\ell}| < \infty , \tag{6}$$

and that the following holds: for any $(k,\ell) \in \mathbb{Z}^2$,

$$\gamma_{k,\ell} = \gamma_{\ell,k} , \tag{7}$$

Then (5) holds, where F is a nonrandom distribution function whose Stieltjes transform $S(z)$ is uniquely defined by the relations:

$$S(z) = \int_0^1 h(x,z)dx , \tag{8}$$

where $h(x,z)$ *is a solution to the equation*

$$h(x,z)=\Big(-z+\int_0^1 f(x,y)h(y,z)dy\Big)^{-1} \quad \textit{with} \;\; f(x,y)=\sum_{k,j\in\mathbb{Z}}\gamma_{k,j}\mathrm{e}^{-2\pi\mathrm{i}(kx+jy)}\,. \tag{9}$$

Equation (9) is uniquely solvable in the class $\mathcal{F}$ of functions $h(x,z)$ with domain $(x,z)\in[0,1]\otimes\mathbb{C}\backslash\mathbb{R}$, which are analytic with respect to z for each fixed x, continuous with respect to x for each fixed z and satisfying the conditions: $\lim_{v\to\infty} v\,\mathrm{Im}\,h(x,\mathrm{i}v)\le 1$ and $\mathrm{Im}(z)\,\mathrm{Im}\,h(x,z)>0$.

Remark 4 *If condition (7) of Theorem 3 is replaced by:* $\gamma_{\ell,k}=V(\ell)V(k)$ *where* V *is an even function, then its conclusion can be given in the following alternative way: the convergence (5) holds where* F *is a nonrandom distribution function whose Stieltjes transform* $S(z)$ *is given by the relation*

$$S(z)=\int_0^\infty \frac{dv(\lambda)}{-z-\lambda h(z)}$$

where $v(t)=\lambda\{x\in[0,1]; f(x)<t\}$, λ *is the Lebesgue measure, for* $x\in[0,1]$, $f(x)=\sum_{k\in\mathbb{Z}}V(k)\mathrm{e}^{2\pi\mathrm{i}kx}$ *and* $h(z)$ *is solution to the equation*

$$h(z)=\int_0^\infty \frac{\lambda dv(\lambda)}{-z-\lambda h(z)},\quad z\in\mathbb{C}\backslash\mathbb{R}\,.$$

This equation is uniquely solvable in the class of analytic functions in $\mathbb{C}\backslash\mathbb{R}$ *satisfying the conditions:* $\lim_{x\to\infty} xh(\mathrm{i}x)<\infty$ *and* $\mathrm{Im}\big(h(z)\big)\,\mathrm{Im}(z)>0$ *for* $z\in\mathbb{C}\backslash\mathbb{R}$. *(See Boutet de Monvel and Khorunzhy [5]).*

2.2 On the limiting distribution for large Gram (sample covariance) matrices with correlated entries

Adapting the proof of Theorem 1, we can also obtain a universality scheme for large sample covariance matrices associated with a process $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ having the representation (3). So, all along this section $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ is assumed to be a random field having the representation (3). To define the Gram matrices associated with this random field, we consider two positive integers N and p, and the $N\times p$ matrix

$$\mathcal{X}_{N,p}=\big(X_{i,j}\big)_{1\le i\le N,1\le j\le p}\,. \tag{10}$$

Define now the symmetric matrix $\mathbb{B}_N$ of order N by

$$\mathbb{B}_N=\frac{1}{p}\mathcal{X}_{N,p}\mathcal{X}_{N,p}^T:=\frac{1}{p}\sum_{k=1}^p \mathbf{r}_k\mathbf{r}_k^T\,, \tag{11}$$

where $\mathbf{r}_k=(X_{1,k},\cdots,X_{N,k})^T$ is the k-th column of $\mathcal{X}_{N,p}$.

The matrix $\mathbb{B}_N$ is usually referred to as the sample covariance matrix associated with the process $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$. It is also known under the name of Gram random matrix.

Theorem 5 *Let* $\mathbb{B}_N$ *be defined by (11) and let* $\mathbb{H}_N=\frac{1}{p}\mathcal{G}_{N,p}\mathcal{G}_{N,p}^T$ *be the Gram matrix associated with a real-valued centered Gaussian random field* $(G_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$, *with covariance function given by (4). Then, provided that* $N,p\to\infty$ *such that* $N/p\to c\in(0,\infty)$, *for any* $z\in\mathbb{C}^+$,

$$\lim_{n\to\infty}\big|S_{\mathbb{B}_N}(z)-\mathbb{E}(S_{\mathbb{H}_N}(z))\big|=0 \quad \textit{almost surely.} \tag{12}$$

Therefore, if $N,p\to\infty$ *such that* $N/p\to c\in(0,\infty)$ *and if there exists a distribution function* F *such that*

$$\mathbb{E}\big(F^{\mathbb{H}_N}(t)\big)\to F(t) \textit{ for all continuity points } t\in\mathbb{R} \textit{ of } F$$

then

$$\mathbb{P}(L(F^{\mathbb{B}_N(\omega)},F)\to 0)=1\,. \tag{13}$$

Theorem 5 together with Theorem 2.1 in Boutet de Monvel *et al* [6] allow then to derive the limiting spectral distribution of large sample covariance matrices associated with a process $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ having the representation (3) and satisfying a short range dependence condition.

Theorem 6 *Let $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ be a real-valued stationary random field given by* (3)*. Assume that* (6) *holds. Then, provided that $N, p \to \infty$ such that $N/p \to c \in (0,\infty)$, $\mathbb{P}(L(F^{\mathbb{B}_N(\omega)}, F) \to 0) = 1$ where F is a nonrandom distribution function whose Stieltjes transform $S(z)$, $z \in \mathbb{C}^+$ is uniquely defined by the relations:*

$$S(z) = \int_0^1 h(x,z)dx\,,$$

where $h(x,z)$ is a solution to the equation

$$h(x,z) = \Big(-z + \int_0^1 \frac{f(x,s)}{1 + c\int_0^1 f(u,s)h(u,z)du} ds\Big)^{-1}\,, \tag{14}$$

with $f(x,y)$ given in (9).

Equation (14) is uniquely solvable in the class $\mathcal{F}$ of functions $h(x,z)$ as described after the statement of Theorem (3).

We refer to the paper by Boutet de Monvel *et al* [6] regarding discussions on the smoothness and boundedness of the limiting density of states. Note that condition (6) is required in the statement of Theorem 6 only because all the estimates in the proof of Theorem 2.1 in [6] require this condition. However using arguments as developed in the paper by Chakrabarty *et al* [7], it can be proved that if the process $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ admits a spectral density then there exists a nonrandom distribution function F such that $\mathbb{P}(L(F^{\mathbb{B}_N(\omega)}, F) \to 0) = 1$ (if $N/p \to c \in (0,\infty)$). Unfortunately the arguments developed in [7] do not allow, in general, to exhibit the limiting equation (14) which gives a lot of information on the limiting spectral distribution. Notice however that if we add the assumption that the lines (resp. the columns) of $\mathcal{X}_{N,p}$ are non correlated (corresponding to the semantically (resp. spatially) "patterns" studied in Section 3 of [6]), condition (6) is not needed to exhibit the limiting equation of the Stieltjes transform. Indeed, in this situation, the lines (resp. the columns) of $\mathcal{G}_{N,p}$ become then independent and the result of Merlevède and Peligrad [14] about the limiting spectral distribution of Gram random matrices associated to independent copies of a stationary process applies. Proving, however, Theorem 6 in its full generality and without requiring condition (6) to hold, remains an open question.

3 Examples

All along this section, $(\xi_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ will designate a double indexed sequence of i.i.d. real-valued random variables defined on a common probability space, centered and in $\mathbb{L}^2$.

3.1 Linear processes

Let $(a_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ be a double indexed sequence of numbers such that

$$\sum_{k,\ell\in\mathbb{Z}} |a_{k,\ell}| < \infty\,. \tag{15}$$

Let then $(X_{i,j})_{(i,j)\in\mathbb{Z}^2}$ be the linear random field in $\mathbb{L}^2$ defined by: for any $(i,j)\in\mathbb{Z}^2$,

$$X_{i,j} = \sum_{k,\ell\in\mathbb{Z}} a_{k,\ell}\xi_{k+i,\ell+j}\,. \tag{16}$$

Corollary 2 (resp. Theorem 5) then applies to the matrix $\mathbb{X}_n$ (resp. $\mathbb{B}_N$) associated with the linear random field $(X_{i,j})_{(i,j)\in\mathbb{Z}^2}$ given in (2).

In case of short dependence, based on our Theorem 6, we can describe the limit of the empirical spectral distribution of the Gram matrix associated with a linear random field.

Corollary 7 *Assume that $X_{i,j}$ is defined by (16) and condition (15) is satisfied. Let N and p be positive integers, such that $N,p \to \infty$, $N/p \to c \in (0,\infty)$. Let $\mathcal{X}_{N,p} = \big(X_{i,j}\big)_{1\le i\le N, 1\le j\le p}$ and $\mathbb{B}_N = N^{-1}\mathcal{X}_{N,p}\mathcal{X}_{N,p}^T$. Then the convergence (13) holds for $F^{\mathbb{B}_N}$, where F is a nonrandom distribution function whose Stieltjes transform satisfies the relations given in Theorem 6 with $\gamma_{k,j} = \|\xi_{0,0}\|_2^2 \sum_{u,v\in\mathbb{Z}} a_{u,v}a_{u+k,v+j}$.*

Concerning now the Wigner-type matrix $\mathbb{X}_n$, by using Remark 4, we obtain the following corollary, describing the limit in a particular case.

Corollary 8 *Let $(a_n)_{n\in\mathbb{Z}}$ be a sequence of numbers such that*

$$\sum_{k\in\mathbb{Z}} |a_k| < \infty.$$

Define $X_{i,j} = \sum_{k,\ell\in\mathbb{Z}} a_k a_\ell \xi_{k+i,\ell+j}$ for any $(i,j)\in\mathbb{Z}^2$. Consider the symmetric matrix $\mathbb{X}_n$ associated with $(X_{i,j})_{(i,j)\in\mathbb{Z}^2}$ and defined by (2). Then (5) holds, where F is a nonrandom distribution function whose Stieltjes transform satisfies the relation given in Remark 4 with $f(x) = \|\xi_{0,0}\|_2 \sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} a_j a_{j+k} \mathrm{e}^{2\pi \mathrm{i} kx}$.

3.2 Volterra-type processes

Other classes of stationary random fields having the representation (3) are Volterra-type processes which play an important role in the nonlinear system theory. For any $\mathbf{k} = (k_1,k_2)\in\mathbb{Z}^2$, define a second-order Volterra expansion as follows:

$$X_{\mathbf{k}} = \sum_{\mathbf{u}\in\mathbb{Z}^2} a_{\mathbf{u}}\xi_{\mathbf{k}-\mathbf{u}} + \sum_{\mathbf{u},\mathbf{v}\in\mathbb{Z}^2} b_{\mathbf{u},\mathbf{v}}\xi_{\mathbf{k}-\mathbf{u}}\xi_{\mathbf{k}-\mathbf{v}}\,, \tag{17}$$

where $a_{\mathbf{u}}$ and $b_{\mathbf{u},\mathbf{v}}$ are real numbers satisfying

$$b_{\mathbf{u},\mathbf{v}} = 0 \text{ if } \mathbf{u}=\mathbf{v}\,,\ \sum_{\mathbf{u}\in\mathbb{Z}^2} a_{\mathbf{u}}^2 < \infty \ \text{ and } \ \sum_{\mathbf{u},\mathbf{v}\in\mathbb{Z}^2} b_{\mathbf{u},\mathbf{v}}^2 < \infty\,. \tag{18}$$

Under the above conditions, the random field $X_{\mathbf{k}}$ exists, is centered and in $\mathbb{L}^2$, and Corollary 2 (resp. Theorem 5) applies to the matrix $\mathbb{X}_n$ (resp. $\mathbb{B}_N$) associated with the Volterra-type random field. Further generalization to arbitrary finite order Volterra expansion is straightforward.

If we reinforced condition (18), we derive the following result concerning the limit of the empirical spectral distribution of the Gram matrix associated with the Volterra-type process:

Corollary 9 *Assume that $(X_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}^2}$ is defined by (17) and that the following additional condition is assumed:*

$$\sum_{\mathbf{u}\in\mathbb{Z}^2} |a_{\mathbf{u}}| < \infty\ ,\ \sum_{\mathbf{v}\in\mathbb{Z}^2}\Big(\sum_{\mathbf{u}\in\mathbb{Z}^2} b_{\mathbf{u},\mathbf{v}}^2\Big)^{1/2} < \infty \ \text{ and } \ \sum_{\mathbf{v}\in\mathbb{Z}^2}\Big(\sum_{\mathbf{u}\in\mathbb{Z}^2} b_{\mathbf{v},\mathbf{u}}^2\Big)^{1/2} < \infty\,. \tag{19}$$

Let N and p be positive integers, such that $N,p\to\infty$, $N/p\to c\in(0,\infty)$. Let $\mathcal{X}_{N,p} = \big(X_{i,j}\big)_{1\le i\le N,1\le j\le p}$ and $\mathbb{B}_N = N^{-1}\mathcal{X}_{N,p}\mathcal{X}_{N,p}^T$. Then (13) holds for $F^{\mathbb{B}_N}$, where F is a nonrandom distribution function whose Stieltjes transform satisfies the relations given in Theorem 6 with

$$\gamma_{\mathbf{k}} = \|\xi_{0,0}\|_2^2 \sum_{\mathbf{u}\in\mathbb{Z}^2} a_{\mathbf{u}}a_{\mathbf{u}+\mathbf{k}} + \|\xi_{0,0}\|_2^4 \sum_{\mathbf{u},\mathbf{v}\in\mathbb{Z}^2} b_{\mathbf{u},\mathbf{v}}(b_{\mathbf{u}+\mathbf{k},\mathbf{v}+\mathbf{k}} + b_{\mathbf{v}+\mathbf{k},\mathbf{u}+\mathbf{k}}) \ \textit{for any } \mathbf{k}\in\mathbb{Z}^2\,. \tag{20}$$

If we impose additional symmetric conditions to the coefficients $a_{\mathbf{u}}$ and $b_{\mathbf{u},\mathbf{v}}$ defining the Volterra random field (17), we can derive the limiting spectral distribution of its associated symmetric matrix $\mathbb{X}_n$ defined by (2). Indeed if for any $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ in $\mathbb{Z}^2$,

$$a_{\mathbf{u}} = a_{u_1} a_{u_2} \, , \; b_{\mathbf{u},\mathbf{v}} = b_{u_1,v_1} b_{u_2,v_2} \, , \tag{21}$$

where the a_i and $b_{i,j}$ are real numbers satisfying

$$b_{i,j} = 0 \text{ if } i = j \, , \; \sum_{i \in \mathbb{Z}} |a_i| < \infty \text{ and } \sum_{(i,j) \in \mathbb{Z}^2} |b_{i,j}| < \infty \, , \tag{22}$$

then $(\gamma_{k,\ell})$ satisfies (6) and (7) . Hence, an application of Theorem 3 leads to the following result.

Corollary 10 *Assume that $(X_{\mathbf{k}})_{\mathbf{k} \in \mathbb{Z}^2}$ is defined by (17) and that conditions* (21) *and* (22) *are satisfied. Define the symmetric matrix $\mathbb{X}_n$ by* (2)*. Then (5) holds, where F is a nonrandom distribution function whose Stieltjes transform is uniquely defined by the relations given in Theorem 3 with $\gamma_{s,t} = A(s)A(t) + B_1(s)B_1(t) + B_2(s)B_2(t)$ with $A(t) = \|\xi_{0,0}\|_2 \sum_{i \in \mathbb{Z}} a_i a_{i+t}$, $B_1(t) = \|\xi_{0,0}\|_2^2 \sum_{(i,r) \in \mathbb{Z}^2} b_{i,r} b_{i+t,r+t}$ and $B_2(t) = \|\xi_{0,0}\|_2^2 \sum_{(i,r) \in \mathbb{Z}^2} b_{i,r} b_{r+t,i+t}$.*

4 Proofs of the main results

The proof of Theorem 1 being based on an approximation of the underlying symmetric matrix by a symmetric matrix with entries that are $2m$-dependent (for m a sequence of integers tending to infinity after n), we shall first prove a universality scheme for symmetric matrices with K-dependent entries. This result has an interest in itself.

4.1 A universality result for symmetric matrices with K-dependent entries

In this section, we are interested by a universality scheme for the spectral limiting distribution of symmetric matrices $\mathbf{X}_n = [X_{k,\ell}^{(n)}]_{k,\ell=1}^n$ normalized by $\sqrt{n}$ when the entries are real-valued random variables defined on a common probability space and satisfy a K-dependence condition (see Assumption $\mathbf{A_3}$). As we shall see later, Theorem 11 below will be a key step to prove Theorem 1 .

Let us start by introducing some assumptions concerning the entries $(X_{k,\ell}^{(n)}, 1 \leq \ell \leq k \leq n)$.

$\mathbf{A_1}$ For all positive integers n, $\mathbb{E}(X_{k,\ell}^{(n)}) = 0$ for all $1 \leq \ell \leq k \leq n$, and

$$\frac{1}{n^2} \sum_{k=1}^{n} \sum_{\ell=1}^{k} \mathbb{E}(|X_{k,\ell}^{(n)}|^2) \leq C < \infty \, .$$

$\mathbf{A_2}$ For any $\tau > 0$,

$$L_n(\tau) := \frac{1}{n^2} \sum_{k=1}^{n} \sum_{\ell=1}^{k} \mathbb{E}(|X_{k,\ell}^{(n)}|^2 \mathbf{1}_{|X_{k,\ell}^{(n)}| > \tau \sqrt{n}}) \to_{n \to \infty} 0 \, .$$

$\mathbf{A_3}$ There exists a positive integer K such that for all positive integers n, the following holds: for all nonempty subsets

$$A, B \subset \{(k, \ell) \in \{1, \ldots, n\}^2 \,|\, 1 \leq \ell \leq k \leq n\}$$

such that

$$\min_{(i,j)\in A} \min_{(k,\ell)\in B} \max(|i-k|, |j-\ell|) > K$$

the σ-fields

$$\sigma\big(X_{i,j}^{(n)}\ ,\ (i,j)\in A\big) \quad \text{and} \quad \sigma\big(X_{k,\ell}^{(n)}\ ,\ (k,\ell)\in B\big)$$

are independent.

Condition $\mathbf{A}_3$ states that variables with index sets which are at a distance larger than K are independent.

In Theorem 11 below, we then obtain a universality result for symmetric matrices whose entries are K-dependent and satisfy $\mathbf{A}_1$ and the traditional Lindeberg's condition $\mathbf{A}_2$. Note that $\mathbf{A}_2$ is known to be a necessary and sufficient condition for the empirical spectral distribution of $n^{-1/2}\mathbf{X}_n$ to converge almost surely to the semi-circle law when the entries $X_{i,j}^{(n)}$ are independent, centered and with common variance not depending on n (see Theorem 9.4.1 in Girko [9]).

Theorem 11 *Let $\mathbf{X}_n = [X_{k,\ell}^{(n)}]_{k,\ell=1}^n$ be a symmetric matrix of order n whose entries $(X_{k,\ell}^{(n)}, 1 \leq \ell \leq k \leq n)$ are real-valued random variables satisfying conditions $\mathbf{A_1}$, $\mathbf{A_2}$ and $\mathbf{A_3}$. Let $\mathbf{G}_n = [G_{i,j}^{(n)}]_{i,j=1}^n$ be a symmetric matrix of order n whose entries $(G_{k,\ell})_{1\leq \ell \leq k\leq n}$ are real-valued centered Gaussian random variables with covariance function given by*

$$\mathbb{E}(G_{k,\ell}^{(n)} G_{i,j}^{(n)}) = \mathbb{E}(X_{k,\ell}^{(n)} X_{i,j}^{(n)})\,. \tag{23}$$

Then, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty} \big|S_{\mathbb{X}_n}(z) - \mathbb{E}(S_{\mathbb{G}_n}(z))\big| = 0 \quad \textit{almost surely,} \tag{24}$$

where $\mathbb{X}_n = n^{-1/2}\mathbf{X}_n$ and $\mathbb{G}_n = n^{-1/2}\mathbf{G}_n$.

The proof of this result will be given in Appendix.

As we mentioned at the beginning of the section, this theorem will be a building block to prove that in the stationary and non triangular setting the K-dependence condition can be relaxed and more general models for the entries can be considered. However, the above theorem has also interest in itself. For instance, for the matrices with real entries, it makes possible to weaken the conditions of Theorems 2.5 and 2.6 in Anderson and Zeitouni [1]. More precisely, due to our Theorem 11, their assumption 2.2.1 (Ib) can be weaken from the boundness of all moments to the boundness of moments of order 2 only plus $\mathbf{A}_2$. Furthermore their result can be strengthened by replacing the convergence in probability to almost sure convergence. Indeed, our Theorem 11 shows that if their assumption 2.2.1 (Ib) is replaced by $\mathbf{A}_1$ plus $\mathbf{A}_2$, then to study the limiting spectral distribution we can actually assume without loss of generality that the entries come from a Gaussian random field with the same covariance structure as the initial entries. If the $X_{k,\ell}^{(n)}$ are Gaussian random variables then boundness of all moments means boundness of moments of order 2.

4.2 Proof of Theorem 1

For m a positive integer (fixed for the moment) and for any (u,v) in $\mathbb{Z}^2$ define

$$X_{u,v}^{(m)} = \mathbb{E}\big(X_{u,v}|\mathcal{F}_{u,v}^{(m)}\big)\,, \tag{25}$$

where $\mathcal{F}_{u,v}^{(m)} := \sigma(\xi_{i,j}\,;\, u-m \leq i \leq u+m,\ v-m\leq j\leq v+m)$.

Let $\mathbf{X}_n^{(m)}$ be the symmetric random matrix of order n associated with $(X_{u,v}^{(m)})_{(u,v)\in\mathbb{Z}^2}$ and defined by $\big(\mathbf{X}_n^{(m)}\big)_{i,j} = X_{i,j}^{(m)}$ if $1 \le j \le i \le n$ and $\big(\mathbf{X}_n^{(m)}\big)_{i,j} = X_{j,i}^{(m)}$ if $1 \le i < j \le n$. Let

$$\mathbb{X}_n^{(m)} = n^{-1/2}\mathbf{X}_n^{(m)} . \tag{26}$$

We first show that, for any $z \in \mathbb{C}^+$,

$$\lim_{m\to\infty}\limsup_{n\to\infty}\big|S_{\mathbb{X}_n}(z) - S_{\mathbb{X}_n^{(m)}}(z)\big| = 0 \ \text{ a.s.} \tag{27}$$

According to Lemma 2.1 in Götze *et al.* [10] (given for convenience in Section 5, Lemma 14),

$$\big|S_{\mathbb{X}_n}(z) - S_{\mathbb{X}_n^{(m)}}(z)\big|^2 \le \frac{1}{n^2 v^4}\mathrm{Tr}\big((\mathbf{X}_n - \mathbf{X}_n^{(m)})^2\big) ,$$

where $v = \mathrm{Im}(z)$. Hence

$$\big|S_{\mathbb{X}_n}(z) - S_{\mathbb{X}_n^{(m)}}(z)\big|^2 \le \frac{2}{n^2 v^4}\sum_{1\le \ell\le k\le n}\big(X_{k,\ell} - X_{k,\ell}^{(m)}\big)^2 .$$

Since the shift is ergodic with respect to the measure generated by a sequence of i.i.d. random variables and the sets of summations are on regular sets, the ergodic theorem entails that

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{1\le k,\ell\le n}\big(X_{k,\ell} - X_{k,\ell}^{(m)}\big)^2 = \mathbb{E}\big((X_{0,0} - X_{0,0}^{(m)})^2\big) \ \text{ a.s. and in } \mathbb{L}^1 .$$

Therefore

$$\limsup_{n\to\infty}\big|S_{\mathbb{X}_n}(z) - S_{\mathbb{X}_n^{(m)}}(z)\big|^2 \le 2v^{-4}\|X_{0,0} - X_{0,0}^{(m)}\|_2^2 \ \text{ a.s.} \tag{28}$$

Now, by the martingale convergence theorem

$$\|X_{0,0} - X_{0,0}^{(m)}\|_2 \to 0 \text{ as } m \to \infty , \tag{29}$$

which combined with (28) proves (27).

Let now $(G_{k,\ell}^{(m)})_{(k,\ell)\in\mathbb{Z}^2}$ be a real-valued centered Gaussian random field, with covariance function given by

$$\mathbb{E}(G_{k,\ell}^{(m)}G_{i,j}^{(m)}) = \mathbb{E}(X_{k,\ell}^{(m)}X_{i,j}^{(m)}) \text{ for any } (k,\ell) \text{ and } (i,j) \text{ in } \mathbb{Z}^2 . \tag{30}$$

Note that the process $(G_{k,\ell}^{(m)})_{(k,\ell)\in\mathbb{Z}^2}$ is then in particular $2m$-dependent. Let now $\mathbf{G}_n^{(m)}$ be the symmetric random matrix of order n defined by $\big(\mathbf{G}_n^{(m)}\big)_{i,j} = G_{i,j}^{(m)}$ if $1 \le j \le i \le n$ and $\big(\mathbf{G}_n^{(m)}\big)_{i,j} = G_{j,i}^{(m)}$ if $1 \le i < j \le n$. Denote $\mathbb{G}_n^{(m)} = \mathbf{G}_n^{(m)}/\sqrt{n}$.

We shall prove that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty}\big|S_{\mathbb{X}_n^{(m)}}(z) - \mathbb{E}\big(S_{\mathbb{G}_n^{(m)}}(z)\big)\big| = 0, \text{ almost surely.} \tag{31}$$

With this aim, we shall apply Theorem 11 and then show in what follows that $(X_{k,\ell}^{(m)}, 1 \le \ell \le k \le n)$ satisfies its assumptions.

Note that the sigma-algebras $\mathcal{F}_{u,v}^{(m)} := \sigma(\xi_{i,j} ; u-m \le i \le u+m,\ v-m \le j \le v+m)$ and $\mathcal{F}_{k,\ell}^{(m)}$ are independent as soon as $|u-k| > 2m$ or $|v-\ell| > 2m$. From this consideration, we then infer that $(X_{k,\ell}^{(m)}, 1 \le \ell \le k \le n)$ satisfies the assumption $\mathbf{A}_3$ of Section 4.1 with $K = 2m$.

On another hand, since $X_{k,\ell}$ is a centered random variable, so is $X^{(m)}_{k,\ell}$. Moreover, $\|X^{(m)}_{k,\ell}\|_2 \le \|X_{k,\ell}\|_2 = \|X_{1,1}\|_2$. Hence $(X^{(m)}_{k,\ell}\,,\, 1 \le \ell \le k \le n)$ satisfies the assumption $\mathbf{A}_1$ of Section 4.1.

We prove now that the assumption $\mathbf{A}_2$ of Section 4.1 holds. With this aim, we first notice that, by Jensen's inequality and stationarity, for any $\tau > 0$,

$$\mathbb{E}((X^{(m)}_{k,\ell})^2 \mathbf{1}_{|X^{(m)}_{k,\ell}|>\tau\sqrt{n}}) \le \mathbb{E}(X^2_{1,1}\mathbf{1}_{|X^{(m)}_{1,1}|>\tau\sqrt{n}})\,.$$

Notice now that if X is a real-valued random variable and $\mathcal{F}$ a sigma-algebra, then for any $\varepsilon > 0$,

$$\mathbb{E}\big(X^2\mathbf{1}_{|\mathbb{E}(X|\mathcal{F})|>2\varepsilon}\big) \le 2\,\mathbb{E}\big(X^2\mathbf{1}_{|X|>\varepsilon}\big)\,.$$

Therefore,

$$\mathbb{E}((X^{(m)}_{k,\ell})^2 \mathbf{1}_{|X^{(m)}_{k,\ell}|>\tau\sqrt{n}}) \le 2\mathbb{E}(X^2_{1,1}\mathbf{1}_{|X_{1,1}|>\tau\sqrt{n}/2})$$

which proves that $(X^{(m)}_{k,\ell}\,,\, 1 \le \ell \le k \le n)$ satisfies $\mathbf{A}_2$ because $\mathbb{E}(X^2_{1,1}) < \infty$.

Since $(X^{(m)}_{k,\ell}\,,\, 1 \le \ell \le k \le n)$ satisfies the assumptions $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{A}_3$ of Section 4.1, applying Theorem 11, (31) follows.

According to (27) and (31), the theorem will follow if we prove that, for any $z \in \mathbb{C}^+$,

$$\lim_{m\to\infty}\limsup_{n\to\infty}\big|\mathbb{E}\big(S_{\mathbb{G}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{G}^{(m)}_n}(z)\big)\big| = 0\,. \tag{32}$$

With this aim, we apply Lemma 16 from Section 5.2 which gives

$$\begin{aligned}&\mathbb{E}\big(S_{\mathbb{G}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{G}^{(m)}_n}(z)\big)\\ &= \frac{1}{2}\sum_{1\le\ell\le k\le n}\sum_{1\le j\le i\le n}\int_0^1 \big(\mathbb{E}(G_{k,\ell}G_{i,j}) - \mathbb{E}(G^{(m)}_{k,\ell}G^{(m)}_{i,j})\big)\mathbb{E}\big(\partial_{k\ell}\partial_{ij}f(\mathbf{g}(t))\big)\,,\end{aligned}$$

where f is defined in (49) and , for $t \in [0,1]$,

$$\mathbf{g}(t) = (\sqrt{t}G_{k,\ell} + \sqrt{1-t}G^{(m)}_{k,\ell})_{1\le\ell\le k\le n}\,.$$

We shall prove that, for any t in $[0,1]$,

$$\begin{aligned}\Big|\sum_{1\le\ell\le k\le n}\sum_{1\le j\le i\le n}\big(\mathbb{E}(G_{k,\ell}G_{i,j}) - \mathbb{E}(G^{(m)}_{k,\ell}G^{(m)}_{i,j})\big)\mathbb{E}\big(\partial_{k\ell}\partial_{ij}f(\mathbf{g}(t))\big)\Big|&\\ \le C\|X^{(m)}_{0,0} - X_{0,0}\|_2\|X_{0,0}\|_2\,.&\end{aligned} \tag{33}$$

where C does not depend on n and t. After integrating on $[0,1]$ and then by taking into account that $\|X_{0,0} - X^{(m)}_{0,0}\|_2^2 \to 0$ as $m \to \infty$, (32) follows by letting n tend to infinity and then m.

To prove (33), using (30) and (4), we write now the following decomposition:

$$\begin{aligned}\mathbb{E}(G_{k,\ell}G_{i,j}) - \mathbb{E}(G^{(m)}_{k,\ell}G^{(m)}_{i,j}) &= \mathbb{E}(X_{k,\ell}X_{i,j}) - \mathbb{E}(X^{(m)}_{k,\ell}X^{(m)}_{i,j})\\ &= \mathbb{E}(X_{k,\ell}(X_{i,j} - X^{(m)}_{i,j})) - \mathbb{E}((X^{(m)}_{k,\ell} - X_{k,\ell})X^{(m)}_{i,j})\,.\end{aligned} \tag{34}$$

We shall decompose the sum on the left-hand side of (33) in two sums according to the decomposition (34) and analyze them separately. Let us prove that there exists a constant C not depending on n and t such that

$$\Big|\sum_{1\le\ell\le k\le n}\sum_{1\le j\le i\le n}\mathbb{E}((X^{(m)}_{k,\ell} - X_{k,\ell})X^{(m)}_{i,j})\mathbb{E}\left(\partial_{k\ell}\partial_{ij}f(\mathbf{g}(t))\right)\Big| \le C\|X^{(m)}_{0,0} - X_{0,0}\|_2\|X_{0,0}\|_2\,. \tag{35}$$

To prove (35), we first notice that without loss of generality $\mathbf{g}(t)$ can be taken independent of $(X_{k,\ell})$ and then

$$\mathbb{E}((X_{k,\ell}^{(m)} - X_{k,\ell})X_{i,j}^{(m)})\mathbb{E}\left(\partial_{k\ell}\partial_{ij} f(\mathbf{g}(t))\right) = \mathbb{E}\left((X_{k,\ell}^{(m)} - X_{k,\ell})X_{i,j}^{(m)}\partial_{k\ell}\partial_{ij} f(\mathbf{g}(t))\right).$$

Next Lemma 15 from from Section 5.2 applied with $a_{k,\ell} = (X_{k,\ell}^{(m)} - X_{k,\ell})$ and $b_{k,\ell} = X_{k,\ell}^{(m)}$ gives: for any $z = u + iv \in \mathbb{C}^+$,

$$\Big| \sum_{1\le \ell\le k\le n} \sum_{1\le j\le i\le n} ((X_{k,\ell}^{(m)} - X_{k,\ell})X_{i,j}^{(m)})\big(\partial_{k\ell}\partial_{ij} f(\mathbf{g}(t))\big)\Big|$$
$$\le \frac{2}{v^3 n^2}\Big(\sum_{1\le \ell\le k\le n} (X_{k,\ell}^{(m)} - X_{k,\ell})^2\Big)^{1/2}\Big(\sum_{1\le j\le i\le n} (X_{i,j}^{(m)})^2\Big)^{1/2}.$$

Therefore, by using Cauchy-Schwarz's inequality, we derive

$$\Big| \sum_{1\le \ell\le k\le n} \sum_{1\le j\le i\le n} \mathbb{E}((X_{k,\ell}^{(m)} - X_{k,\ell})X_{i,j}^{(m)})\mathbb{E}\left(\partial_{k\ell}\partial_{ij} f(\mathbf{g}(t))\right)\Big|$$
$$\le \frac{2}{v^3 n^2}\Big(\sum_{1\le \ell\le k\le n} \mathbb{E}(X_{k,\ell}^{(m)} - X_{k,\ell})^2\Big)^{1/2}\Big(\sum_{1\le j\le i\le n} \mathbb{E}(X_{i,j}^{(m)})^2\Big)^{1/2}.$$

Using stationarity it follows that, for any $z = u + iv \in \mathbb{C}^+$ and any t in $[0, 1]$,

$$\Big| \sum_{1\le \ell\le k\le n} \sum_{1\le j\le i\le n} \mathbb{E}((X_{k,\ell}^{(m)} - X_{k,\ell})X_{i,j}^{(m)})\mathbb{E}\big(\partial_{k\ell}\partial_{ij} f(\mathbf{g}(t))\big)\Big| \le 2v^{-3}\|X_{0,0}^{(m)} - X_{0,0}\|_2\|X_{0,0}\|_2 .$$

Similarly, we can prove that for any $z = u + iv \in \mathbb{C}^+$ and any t in $[0, 1]$,

$$\Big| \sum_{1\le \ell\le k\le n} \sum_{1\le j\le i\le n} \mathbb{E}(X_{k,\ell}(X_{i,j} - X_{i,j}^{(m)}))\mathbb{E}\big(\partial_{k\ell}\partial_{ij} f(\mathbf{g}(t))\big)\Big| \le 2v^{-3}\|X_{0,0}^{(m)} - X_{0,0}\|_2\|X_{0,0}\|_2 .$$

This leads to (33) and then ends the proof of the theorem. $\square$

4.3 Proof of Theorem 3

In order to establish Theorem 3, it suffices to apply Theorem 1 and to derive the limit of $\mathbb{E}(S_{\mathbb{G}_n}(z))$ for any $z \in \mathbb{C}^+$, where $\mathbb{G}_n$ is the symmetric matrix defined in Theorem 1. With this aim, we apply Proposition 20 given in Section 5.2. Proposition 20 is a modification of Theorem 2 in Khorunzhy and Pastur [13] (see also Theorem 17.2.1 in [15]) since in our case, we cannot use directly the conclusion of their theorem: we are not exactly in the situation described there. Their symmetric matrix is defined via a *symmetric* real-valued centered Gaussian random field $(W_{k,\ell})_{k,\ell}$ satisfying the following property: $W_{k,\ell} = W_{\ell,k}$ for any $(k, \ell) \in \mathbb{Z}^2$ and also (2.8) in [13]. In our situation, and if (7) is assumed, the entries $(g_{k,\ell})_{1\le k,\ell\le n}$ of $n^{1/2}\mathbb{G}_n$ have the following covariances

$$\mathbb{E}(g_{i,j}g_{k,\ell}) = \gamma_{i-k,j-\ell}(\mathbf{1}_{i\ge j,k\ge \ell} + \mathbf{1}_{j>i,\ell>k}) + \gamma_{i-\ell,j-k}(\mathbf{1}_{i\ge j,\ell>k} + \mathbf{1}_{j>i,k\ge \ell}), \qquad (36)$$

since by (4) and stationarity

$$g_{k,\ell} = G_{\max(k,\ell),\min(k,\ell)} \quad \text{and} \quad \mathbb{E}(G_{i,j}, G_{k,\ell}) = \gamma_{k-i,\ell-j}.$$

Hence, because of the indicator functions appearing in (36), our covariances do not satisfy the condition (2.8) in [13]. However, the conclusion of Theorem 2 in [13] also holds for $S_{\mathbb{G}_n}(z)$ provided that (6) and (7) are satisfied. We did not find any reference where the assertion above is mentioned so Proposition 20 is proved with this aim. $\square$

4.4 Proof of Theorem 5

Let $n = N + p$ and $\mathbb{X}_n$ the symmetric matrix of order n defined by

$$\mathbb{X}_n = \frac{1}{\sqrt{p}} \begin{pmatrix} \mathbf{0}_{p,p} & \mathcal{X}_{N,p}^T \\ \mathcal{X}_{N,p} & \mathbf{0}_{N,N} \end{pmatrix} .$$

Notice that the eigenvalues of $\mathbb{X}_n^2$ are the eigenvalues of $p^{-1}\mathcal{X}_{N,p}^T\mathcal{X}_{N,p}$ together with the eigenvalues of $p^{-1}\mathcal{X}_{N,p}\mathcal{X}_{N,p}^T$. Since these two latter matrices have the same nonzero eigenvalues, the following relation holds: for any $z \in \mathbb{C}^+$

$$S_{\mathbb{B}_N}(z) = z^{-1/2}\frac{n}{2N}S_{\mathbb{X}_n}(z^{1/2}) + \frac{p-N}{2Nz} . \tag{37}$$

(See, for instance, page 549 in Rashidi Far *et al* [16] for additional arguments leading to the relation above). A similar relation holds for the Gram matrix $\mathbb{H}_N$ associated with the centered Gaussian random field $(G_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ having the same covariance structure as $(X_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$, namely: for any $z \in \mathbb{C}^+$

$$S_{\mathbb{H}_N}(z) = z^{-1/2}\frac{n}{2N}S_{\mathbb{G}_n}(z^{1/2}) + \frac{p-N}{2Nz} , \tag{38}$$

where $\mathbb{G}_n$ is defined as $\mathbb{X}_n$ but with $\mathcal{G}_{N,p}$ replacing $\mathcal{X}_{N,p}$.

In view of the relations (37) and (38), and since $n/N \to 1 + c^{-1}$, to prove (12), it suffices to show that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty} \big|S_{\mathbb{X}_n}(z) - \mathbb{E}\big(S_{\mathbb{G}_n}(z)\big)\big| = 0 \ \text{ almost surely} \tag{39}$$

Note now that $\mathbb{X}_n := n^{-1/2}[x_{ij}^{(n)}]_{i,j=1}^n$ where $x_{ij}^{(n)} = \sqrt{\frac{n}{p}}X_{i-p,j}\mathbf{1}_{i\geq p+1}\mathbf{1}_{1\leq j\leq p}$ if $1 \leq j \leq i \leq n$, and $x_{ij}^{(n)} = x_{ji}^{(n)}$ if $1 \leq i < j \leq n$. Similarly we can write $\mathbb{G}_n := n^{-1/2}[g_{ij}^{(n)}]_{i,j=1}^n$ where the $g_{ij}^{(n)}$'s are defined as the $x_{ij}^{(n)}$'s but with $X_{i-p,j}$ replaced by $G_{i-p,j}$. Following the proof of Theorem 1 , we infer that its conclusion holds (and therefore (39) does) even when the stationarity of entries of $\mathbb{X}_n$ and $\mathbb{G}_n$ is slightly relaxed as above. □

4.5 Proof of Theorem 6

In view of the convergence (12), it suffices to show that when $N, p \to \infty$ such that $N/p \to c \in (0,\infty)$, then for any $z \in \mathbb{C}^+$, $\mathbb{E}\big(S_{\mathbb{H}_N}(z)\big)$ converges to $S(z) = \int_0^1 h(x,z)dx$ where $h(x,z)$ is a solution to the equation (14). This follows by applying Theorem 2.1 in Boutet de Monvel *et al* [6] . Indeed setting $\tilde{\mathbb{H}}_N = \frac{p}{N}\mathbb{H}_N$, this theorem asserts that if (6) holds then, when $N, p \to \infty$ such that $N/p \to c \in (0,\infty)$, $\mathbb{E}\big(S_{\tilde{\mathbb{H}}_N}(z)\big)$ converges to $m(z) = \int_0^1 v(x,z)dx$, for any $z \in \mathbb{C}^+$, where $v(x,z)$ is a solution to the equation

$$v(x,z) = \Big(-z + c^{-1}\int_0^1 \frac{f(x,s)}{1 + \int_0^1 f(u,s)v(u,z)du}ds\Big)^{-1} .$$

This implies that $\mathbb{E}\big(S_{\mathbb{H}_N}(z)\big)$ converges to $S(z)$ as defined in the theorem since the following relation holds: $S(z) = c^{-1}m(z/c)$.

□

5 Technical results

5.1 Concentration of the spectral measure

Next proposition is a generalization to row-wise K-dependent random matrices of Theorem 1 (ii) of Guntuboyina and Leeb [11].

Proposition 12 *Let $(X_{k,\ell}^{(n)})_{1\le\ell\le k\le n}$ be an array of complex-valued random variables defined on a common probability space. Assume that there exists a positive integer K such that for any integer $u \in [1, n-K]$, the σ-fields*

$$\sigma\big(X_{i,j}^{(n)}\,,\,1\le j\le i\le u\big) \quad \text{and} \quad \sigma\big(X_{k,\ell}^{(n)}\,,\;1\le \ell\le k\,,\,u+K+1\le k\le n\big)$$

are independent. Define the symmetric matrix $\mathbb{X}_n$ by (2). Then for every measurable function $f : \mathbb{R}\to\mathbb{R}$ of bounded variation, any $n \ge K$ and any $r \ge 0$,

$$\mathbb{P}\Big(\Big|\int f d\nu_{\mathbb{X}_n} - \mathbb{E}\int f d\nu_{\mathbb{X}_n}\Big| \ge r\Big) \le 2\exp\Big(-\frac{nr^2}{160KV_f^2}\Big)\,, \tag{40}$$

where V_f is the variation of the function f.

Application to the Stieltjes transform. Assume that the assumptions of Proposition 12 hold. Let $z = u + \mathrm{i}v \in \mathbb{C}^+$ and note that

$$S_{\mathbb{X}_n}(z) = \int \frac{1}{x-z} d\nu_{\mathbb{X}_n}(x) = \int f_1(x)d\nu_{\mathbb{X}_n}(x) + \mathrm{i}\int f_2(x)d\nu_{\mathbb{X}_n}(x)\,,$$

where $f_1(x) = \frac{x-u}{(x-u)^2+v^2}$ and $f_2(x) = \frac{v}{(x-u)^2+v^2}$. Now

$$V_{f_1} = \|f_1'\|_1 = \frac{2}{v} \text{ and } V_{f_2} = \|f_2'\|_1 = \frac{2}{v}\,.$$

Therefore, by applying Proposition 12 to f_1 and f_2, we get that for any $n \ge K$ and any $r \ge 0$,

$$\mathbb{P}\big(\big|S_{\mathbb{X}_n}(z) - \mathbb{E}S_{\mathbb{X}_n}(z)\big| \ge r\big) \le 4\exp\Big(-\frac{nr^2v^2}{2560K}\Big)\,. \tag{41}$$

Proof of Proposition 12. It is convenient to start by considering the map A which "constructs" symmetric matrices of order n as in (2). To define it, let $N = n(n+1)/2$ and write elements of $\mathbb{R}^N$ as $\mathbf{x} = (r_1,\dots,r_n)$ where $r_i = (x_{i,j})_{1\le j\le i}$. For any $\mathbf{x}\in\mathbb{R}^N$, let $A(\mathbf{x}) = A(r_1,\dots,r_n)$ be the matrix defined by

$$\big(A(\mathbf{x})\big)_{ij} = \begin{cases} \frac{1}{\sqrt{n}}x_{i,j} = \frac{1}{\sqrt{n}}(r_i)_j & \text{if } i\ge j \\ \frac{1}{\sqrt{n}}x_{j,i} = \frac{1}{\sqrt{n}}(r_j)_i & \text{if } i<j \end{cases} \tag{42}$$

For $1 \le i \le n$, let $R_i = (X_{i,j}^{(n)})_{1\le j\le i}$. By definition, we have that $\mathbb{X}_n = A(R_1,\dots,R_n)$. Let then h be the function from $\mathbb{C}^N$ to $\mathbb{R}$ defined by

$$h(R_1,\dots,R_n) = \int f d\nu_{A(R_1,\dots,R_n)}\,.$$

Let $n \geq K$. Denoting by $\mathcal{F}_k = \sigma(R_1, \dots, R_k)$ for $k \geq 1$, and by $\mathcal{F}_0 = \{\emptyset, \Omega\}$, we then write the following martingale decomposition:

$$\begin{aligned}
\int f d\nu_{\mathbb{X}_n} - \mathbb{E}\int f d\nu_{\mathbb{X}_n} &= h(R_1, \dots, R_n) - \mathbb{E}h(R_1, \dots, R_n) \\
&= \sum_{i=1}^{[n/K]} \Big(\mathbb{E}\big(h(R_1, \dots, R_n)|\mathcal{F}_{iK}\big) - \mathbb{E}\big(h(R_1, \dots, R_n)|\mathcal{F}_{(i-1)K}\big)\Big) \\
&\qquad + \mathbb{E}\big(h(R_1, \dots, R_n)|\mathcal{F}_n\big) - \mathbb{E}\big(h(R_1, \dots, R_n)|\mathcal{F}_{K[n/K]}\big) \\
&:= \sum_{i=1}^{[n/K]+1} d_{i,n} \,.
\end{aligned}$$

Let

$$\mathbf{R}_n = (R_1, \dots, R_n) \quad \text{and} \quad \mathbf{R}_n^{k,\ell} = (R_1, \dots, R_k, 0, \dots, 0, R_{\ell+1}, \dots, R_n) \,.$$

Note now that, for any $i \in \{1, \dots, [n/K]\}$,

$$\mathbb{E}\Big(h\big(\mathbf{R}_n^{(i-1)K,(i+1)K}\big)|\mathcal{F}_{iK}\Big) = \mathbb{E}\Big(h\big(\mathbf{R}_n^{(i-1)K,(i+1)K}\big)|\mathcal{F}_{(i-1)K}\Big) \,. \tag{43}$$

To see this it suffices to apply Lemma 13 with $X = (R_{(i+1)K+1}, \dots, R_n)$, $Y = (R_1, \dots, R_{(i-1)K})$ and $Z = (R_1, \dots, R_{iK})$. Therefore, by taking into account (43), we get that, for any $i \in \{1, \dots, [n/K]\}$,

$$\begin{aligned}
&\mathbb{E}\big(h(\mathbf{R}_n)|\mathcal{F}_{iK}\big) - \mathbb{E}\big(h(\mathbf{R}_n)|\mathcal{F}_{(i-1)K}\big) \\
&\quad = \mathbb{E}\big(h(\mathbf{R}_n) - h\big(\mathbf{R}_n^{(i-1)K,(i+1)K}\big)|\mathcal{F}_{iK}\big) - \mathbb{E}\big(h(\mathbf{R}_n) - h\big(\mathbf{R}_n^{(i-1)K,(i+1)K}\big)|\mathcal{F}_{(i-1)K}\big) \,.
\end{aligned} \tag{44}$$

We write now that

$$\begin{aligned}
&h(\mathbf{R}_n) - h\big(\mathbf{R}_n^{(i-1)K,(i+1)K}\big) \\
&\quad = \sum_{j=iK+1}^{(i+1)K} \Big(h\big(\mathbf{R}_n^{iK,j-1}\big) - h\big(\mathbf{R}_n^{iK,j}\big)\Big) + \sum_{j=(i-1)K+1}^{iK} \Big(h\big(\mathbf{R}_n^{j,(i+1)K}\big) - h\big(\mathbf{R}_n^{j-1,(i+1)K}\big)\Big) \,,
\end{aligned} \tag{45}$$

since $\mathbf{R}_n = \mathbf{R}_n^{iK,iK}$. But if $\mathbb{Y}_n$ and $\mathbb{Z}_n$ are two symmetric matrices of size n, then

$$\Big|\int f d\nu_{\mathbb{Y}_n} - \int f d\nu_{\mathbb{Z}_n}\Big| \leq V_f \|F^{\mathbb{Y}_n} - F^{\mathbb{Z}_n}\|_\infty$$

(see for instance the proof of Theorem 6 in [11]). Hence, from Theorem A.43 in Bai and Silverstein [2],

$$\Big|\int f d\nu_{\mathbb{Y}_n} - \int f d\nu_{\mathbb{Z}_n}\Big| \leq \frac{V_f}{n} \operatorname{rank}\big(\mathbb{Y}_n - \mathbb{Z}_n\big) \,.$$

With our notations, this last inequality implies that for any $0 \leq k \leq \ell \leq n$ and $0 \leq i \leq j \leq n$

$$\big|h\big(\mathbf{R}_n^{k,\ell}\big) - h\big(\mathbf{R}_n^{i,j}\big)\big| \leq \frac{V_f}{n} \operatorname{rank}\big(A(\mathbf{R}_n^{k,\ell}) - A(\mathbf{R}_n^{i,j})\big) \,. \tag{46}$$

Starting from (45) and using (46) together with

$$\operatorname{rank}\big(A(\mathbf{R}_n^{iK,j-1}) - A(\mathbf{R}_n^{iK,j})\big) \leq 2$$

and

$$\operatorname{rank}\big(A(\mathbf{R}_n^{j,(i+1)K}) - A(\mathbf{R}_n^{j-1,(i+1)K})\big) \leq 2 \,,$$

we get that

$$\big|h(\mathbf{R}_n) - h\big(\mathbf{R}_n^{(i-1)K,(i+1)K}\big)\big| \le \frac{4K}{n} V_f\,. \tag{47}$$

Starting from (44) and using (47), it follows that, for any $i \in \{1,\dots,[n/K]\}$,

$$\big|\mathbb{E}\big(h(\mathbf{R}_n)|\mathcal{F}_{iK}\big) - \mathbb{E}\big(h(\mathbf{R}_n)|\mathcal{F}_{(i-1)K}\big)\big| \le \frac{8K}{n} V_f$$

On another hand, since $\mathbf{R}_n^{K[n/K],n}$ is $\mathcal{F}_{K[n/K]}$-measurable,

$$\begin{aligned}\mathbb{E}\big(h(\mathbf{R}_n)|\mathcal{F}_n\big) - \mathbb{E}\big(h(\mathbf{R}_n)|\mathcal{F}_{K[n/K]}\big) &= \mathbb{E}\big(h(\mathbf{R}_n) - h\big(\mathbf{R}_n^{K[n/K],n}\big)|\mathcal{F}_n\big) \\ &\quad - \mathbb{E}\big(h(\mathbf{R}_n) - h\big(\mathbf{R}_n^{K[n/K],n}\big)|\mathcal{F}_{K[n/K]}\big)\,.\end{aligned}$$

Now

$$h(\mathbf{R}_n) - h\big(\mathbf{R}_n^{K[n/K],n}\big) = \sum_{j=K[n/K]+1}^{n} \Big(h\big(\mathbf{R}_n^{j,n}\big) - h\big(\mathbf{R}_n^{j-1,n}\big)\Big)\,.$$

So, proceeding as before, we infer that

$$\big|\mathbb{E}\big(h(\mathbf{R}_n)|\mathcal{F}_n\big) - \mathbb{E}\big(h(\mathbf{R}_n)|\mathcal{F}_{K[n/K]}\big)\big| \le \frac{4K}{n} V_f\,.$$

So, overall we derive that $\|d_{i,n}\|_\infty \le \frac{8K}{n} V_f$ for any $i \in \{1,\dots,[n/K]\}$ and $\|d_{[n/K]+1,n}\|_\infty \le \frac{4K}{n} V_f$. Therefore, the proposition follows by applying the Azuma-Hoeffding inequality for martingales. $\square$

5.2 Other useful technical results

Lemma 13 *If X, Y, Z are three random vectors defined on a probability space $(\Omega, \mathcal{K}, \mathbb{P})$, such that X is independent of $\sigma(Z)$ and $\sigma(Y) \subset \sigma(Z)$. Then, for any measurable function g such that $\|g(X,Y)\|_1 < \infty$,*

$$\mathbb{E}(g(X,Y)|Z) = \mathbb{E}(g(X,Y)|Y) \text{ a.s.} \tag{48}$$

The following lemma is Lemma 2.1 in Götze *et al.* [10] and allows to compare two Stieltjes transforms.

Lemma 14 *Let $\mathbf{A}_n$ and $\mathbf{B}_n$ be two symmetric $n \times n$ matrices. Then, for any $z \in \mathbb{C}\backslash\mathbb{R}$,*

$$\big|S_{\mathbb{A}_n}(z) - S_{\mathbb{B}_n}(z)\big|^2 \le \frac{1}{n^2 |\operatorname{Im}(z)|^4} \operatorname{Tr}\big((\mathbf{A}_n - \mathbf{B}_n)^2\big)\,,$$

where $\mathbb{A}_n = n^{-1/2}\mathbf{A}_n$ and $\mathbb{B}_n = n^{-1/2}\mathbf{B}_n$.

All along the proofs, we shall use the fact that the Stieltjes transform of the spectral measure is a smooth function of the matrix entries. Let $N = n(n+1)/2$ and write elements of $\mathbb{R}^N$ as $\mathbf{x} = (x_{ij})_{1\le j\le i\le n}$. For any $z \in \mathbb{C}^+$, let $f(\cdot) := f_{n,z}(\cdot)$ be the function defined from $\mathbb{R}^N$ to $\mathbb{C}$ by

$$f(\mathbf{x}) = \frac{1}{n}\operatorname{Tr}\big(A(\mathbf{x}) - z\mathbf{I}_n\big)^{-1} \text{ for any } \mathbf{x} \in \mathbb{R}^N\,, \tag{49}$$

where $A(\mathbf{x})$ is the matrix defined in (42) and $\mathbf{I}_n$ is the identity matrix of order n. The function f admits partial derivatives of all orders. In particular, denoting for any $\mathbf{u} \in \{(i,j)\}_{1\le j\le i\le n}$, $\partial_{\mathbf{u}} f$ for $\partial f_n / \partial x_{\mathbf{u}}$, the following upper bounds hold: for any $z = x + iy \in \mathbb{C}^+$ and any $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\{(i,j)\}_{1\le j\le i\le n}$,

$$|\partial_{\mathbf{u}} f| \le \frac{2}{y^2 n^{3/2}}\,,\ |\partial_{\mathbf{u}}\partial_{\mathbf{v}} f| \le \frac{4}{y^3 n^2} \text{ and } |\partial_{\mathbf{u}}\partial_{\mathbf{v}}\partial_{\mathbf{w}}| \le \frac{3 \times 2^{5/2}}{y^4 n^{5/2}}\,. \tag{50}$$

(See the equalities (20) and (21) in [8] together with the computations on pages 2074-2075). In addition, the following lemma has been proved in Merlevède and Peligrad (2014):

Lemma 15 *Let $z = x + iy \in \mathbb{C}^+$ and $f_n := f_{n,z}$ be defined by (49). Let $(a_{ij})_{1\le j\le i\le n}$ and $(b_{ij})_{1\le j\le i\le n}$ be real numbers. Then, for any subset $\mathcal{I}_n$ of $\{(i,j)\}_{1\le j\le i\le n}$ and any element $\mathbf{x}$ of $\mathbb{R}^N$,*

$$\Big|\sum_{\mathbf{u}\in\mathcal{I}_n}\sum_{\mathbf{v}\in\mathcal{I}_n} a_{\mathbf{u}}b_{\mathbf{v}}\partial_{\mathbf{u}}\partial_{\mathbf{v}}f(\mathbf{x})\Big| \le \frac{2}{y^3n^2}\Big(\sum_{\mathbf{u}\in\mathcal{I}_n} a_{\mathbf{u}}^2\sum_{\mathbf{v}\in\mathcal{I}_n} b_{\mathbf{v}}^2\Big)^{1/2}.$$

Next lemma is a consequence of the well-known Gaussian interpolation trick.

Lemma 16 *Let $(Y_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ and $(Z_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ be two centered real-valued Gaussian processes. Let $\mathbf{Y}_n$ be the symmetric random matrix of order n defined by $\big(\mathbf{Y}_n\big)_{i,j} = Y_{i,j}$ if $1\le j\le i\le n$ and $\big(\mathbf{Y}_n\big)_{i,j} = Y_{j,i}$ if $1\le i<j\le n$. Denote $\mathbb{Y}_n = \frac{1}{\sqrt{n}}\mathbf{Y}_n$. Define similarly $\mathbb{Z}_n$. Then, for any $z = x + \mathrm{i}y \in \mathbb{C}^+$,*

$$\begin{aligned}&\mathbb{E}\big(S_{\mathbb{Y}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{Z}_n}(z)\big)\\ &\qquad = \frac{1}{2}\sum_{1\le\ell\le k\le n}\sum_{1\le j\le i\le n}\int_0^1 \big(\mathbb{E}(Y_{k,\ell}Y_{i,j}) - \mathbb{E}(Z_{k,\ell}Z_{i,j})\big)\mathbb{E}\big(\partial_{k\ell}\partial_{ij}f(\mathbf{u}(t))\big)dt \qquad (51)\end{aligned}$$

where, for $t\in[0,1]$, $\mathbf{u}(t) = (\sqrt{t}Y_{k,\ell} + \sqrt{1-t}Z_{k,\ell})_{1\le\ell\le k\le n}$ and

$$\big|\mathbb{E}\big(S_{\mathbb{Y}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{Z}_n}(z)\big)\big| \le \frac{2}{n^2y^3}\sum_{1\le\ell\le k\le n}\sum_{1\le j\le i\le n}|\mathbb{E}(Y_{k,\ell}Y_{i,j}) - \mathbb{E}(Z_{k,\ell}Z_{i,j})|. \qquad (52)$$

Proof. Using the definition of f, we first write

$$\mathbb{E}\big(S_{\mathbb{Y}_n}(z)\big) = \mathbb{E}f\big((Y_{k,\ell})_{1\le\ell\le k\le n}\big) \text{ and } \mathbb{E}\big(S_{\mathbb{Z}_n}(z)\big) = \mathbb{E}f\big((Z_{k,\ell})_{1\le\ell\le k\le n}\big).$$

Equality (51) then follows from the usual interpolation trick (for an easy reference we cite Talagrand [17] Section 1.3, Lemma 1.3.1.). To obtain the upper bound (52), it suffices then to take into account (50). □

Below we give a Taylor expansion for functions of random variables of a convenient type for Lindeberg's method.

Lemma 17 *Let $f(\cdot)$ be a function from $\mathbb{R}^{d+m}$ to $\mathbb{C}$, three times differentiable, with continuous and bounded third partial derivatives, i.e. there is a constant L_3 such that*

$$|\partial_i\partial_j\partial_k f(\mathbf{x})| \le L_3 \text{ for all } i,j,k \text{ and } \mathbf{x}.$$

Let $\mathbf{X},\mathbf{Y},\mathbf{Z}$ be random vectors defined on a probability space $(\Omega,\mathcal{K},\mathbb{P})$, such that $\mathbf{X}$ and $\mathbf{Y}$ take their values in $\mathbb{R}^d$, and $\mathbf{Z}$ takes its values in $\mathbb{R}^m$. Assume in addition that $\mathbf{X}$ and $\mathbf{Y}$ are independent of $\mathbf{Z}$, and that $\mathbf{X}$ and $\mathbf{Y}$ are in $\mathbb{L}^3(\mathbb{R}^d)$, centered at expectation and have the same covariance structure. Then, for any permutation $\pi : \mathbb{R}^{d+m}\to\mathbb{R}^{d+m}$, we have

$$|\mathbb{E}f(\pi(\mathbf{X},\mathbf{Z})) - \mathbb{E}f(\pi(\mathbf{Y},\mathbf{Z}))| \le \frac{L_3d^2}{3}\Big(\sum_{j=1}^d\mathbb{E}(|X_j|^3) + \sum_{j=1}^d\mathbb{E}(|Y_j|^3)\Big).$$

The proof of this lemma is based on the following Taylor expansion for functions of several variables.

Lemma 18 *Let $g(\cdot)$ be a function from $\mathbb{R}^p$ to $\mathbb{R}$, three times differentiable, with continuous third partial derivatives and such that*

$$|\partial_i\partial_j\partial_k g(\mathbf{x})| \le L_3 \text{ for all } i,j,k \text{ and } \mathbf{x}.$$

Then, for any $\mathbf{a} = (a_1, \dots, a_p)$ *and* $\mathbf{b} = (b_1, \dots, b_p)$ *in* $\mathbb{R}^p$,

$$g(\mathbf{b}) - g(\mathbf{a}) = \sum_{k=1}^{p} (b_j - a_j)\partial_j g(\mathbf{0}) + \frac{1}{2}\sum_{j,k=1}^{p} (b_j b_k - a_j a_k)\partial_j \partial_k g(\mathbf{0}) + R_3(\mathbf{a}, \mathbf{b}).$$

$$\textit{with } |R_3(\mathbf{a}, \mathbf{b})| \le \frac{L_3}{6}\Big(\big(\sum_{j=1}^{p} |a_j|\big)^3 + \big(\sum_{j=1}^{p} |b_j|\big)^3 \Big) \le \frac{L_3 p^2}{6}\Big(\sum_{j=1}^{p} |a_j|^3 + |b_j|^3 \Big).$$

Proof of Lemma 18. We use Taylor expansion of second order for functions with bounded partial derivatives of order three. It is well-known that

$$g(\mathbf{a}) - g(\mathbf{0}_p) = \sum_{j=1}^{p} a_j \partial_j g(\mathbf{0}_p) + \frac{1}{2}\sum_{j,k=1}^{p} a_j a_k \partial_j \partial_k g(\mathbf{0}_p) + R_3(\mathbf{a}),$$

$$\text{where } |R_3(\mathbf{a})| \le \frac{L_3}{6}\big(\sum_{j=1}^{p} |a_j|\big)^3 \le \frac{L_3 p^2}{6}\sum_{j=1}^{p} |a_j|^3.$$

By writing a similar expression for $g(\mathbf{b}) - g(\mathbf{0}_p)$ and substracting them the result follows. $\square$

Proof of Lemma 17. For simplicity of the notation we shall prove it first for $f((\mathbf{X}, \mathbf{Z})) - f((\mathbf{Y}, \mathbf{Z}))$. We start by applying Lemma 18 to real and imaginary part of f and obtain

$$f(\mathbf{X}, \mathbf{Z}) - f(\mathbf{Y}, \mathbf{Z}) = \sum_{j=1}^{d} (X_j - Y_j)\partial_j f(\mathbf{0}_d, \mathbf{Z}) + \frac{1}{2}\sum_{j,k=1}^{d} (X_k X_j - Y_k Y_j)\partial_j \partial_k f(\mathbf{0}_d, \mathbf{Z}) + R_3,$$

$$\text{with } |R_3| \le \frac{L_3 d^2}{3}\Big(\sum_{j=1}^{d} |X_j|^3 + \sum_{j=1}^{d} |Y_j|^3 \Big).$$

By taking the expected value and taking into account the hypothesis of independence and the fact that $\mathbf{X}$ and $\mathbf{Y}$ are centered at expectations and have the same covariance structure, we obtain, for all $1 \le j \le d$

$$\mathbb{E}((X_j - Y_j)\partial_j f(\mathbf{0}_d, \mathbf{Z})) = (\mathbb{E}X_j - \mathbb{E}Y_j)\mathbb{E}\partial_j f(\mathbf{0}_d, \mathbf{Z}) = 0$$

and, for all $1 \le k, j \le d$,

$$\mathbb{E}(X_k X_j - Y_k Y_j)\partial_j \partial_k f(\mathbf{0}_d, \mathbf{Z}) = (\mathbb{E}(X_k X_j) - \mathbb{E}(Y_k Y_j))\mathbb{E}\partial_j \partial_k f(\mathbf{0}_d, \mathbf{Z}) = 0.$$

It follows that

$$\mathbb{E}f(\mathbf{X}, \mathbf{Z}) - \mathbb{E}f(\mathbf{Y}, \mathbf{Z}) = R_3,$$

$$\text{with } |R_3| \le \frac{L_3 d^2}{3}\Big(\sum_{j=1}^{d} |X_j|^3 + \sum_{j=1}^{p} |Y_j|^3 \Big).$$

It remains to note that the result remains valid for any permutation of variables $(\mathbf{X}, \mathbf{Z})$. The variables in $\mathbf{X}, \mathbf{Z}$ can hold any positions among the variables in function f since we just need all the derivatives of order three to be uniformly bounded. The difference in the proof consists only in re-denoting the partial derivatives; for instance instead of ∂_j we shall use ∂_{k_j} where k_j, $1 \le k_j \le d + m$ denotes the index of the variable X_j in $f(x_1, x_2, ..., x_{d+m})$. $\square$

We provide next a technical lemma on the behavior of the expected value of Stieltjes transform of symmetric matrices with Gaussian entries. In Lemma 19 and Proposition 20 below, we consider a stationary real-valued centered Gaussian random field $(G_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ with covariance function given by: for any $(k,\ell)\in\mathbb{Z}^2$ and any $(i,j)\in\mathbb{Z}^2$,

$$\mathbb{E}(G_{k,\ell}G_{i,j}) = \gamma_{k-i,\ell-j}\,,$$

satisfying (6) and (7). We define then two symmetric matrices of order n, $\mathbb{G}_n = n^{-1/2}[g_{k,\ell}]_{k,\ell=1}^n$ and $\mathbb{W}_n = n^{-1/2}[W_{k,\ell}]_{k,\ell=1}^n$ where the entries $g_{k,\ell}$ and $W_{k,\ell}$ are defined respectively by

$$g_{k,\ell} = G_{\max(k,\ell),\min(k,\ell)} \text{ and } W_{k,\ell} = \frac{1}{\sqrt{2}}\big(G_{k,\ell}+G_{\ell,k}\big)\,.$$

Lemma 19 *For any $z\in\mathbb{C}\backslash\mathbb{R}$ the following convergence holds:*

$$\lim_{n\to\infty}\big|\mathbb{E}(S_{\mathbb{G}_n}(z)) - \mathbb{E}(S_{\mathbb{W}_n}(z))\big| = 0\,.$$

As a consequence of this lemma and Theorem 2 in [13], we obtain the following result concerning the limiting spectral distribution of both $\mathbb{G}_n$ and $\mathbb{W}_n$.

Proposition 20 *For any $z\in\mathbb{C}\backslash\mathbb{R}$, $S_{\mathbb{G}_n}(z)$ and $S_{\mathbb{W}_n}(z)$ have almost surely the same limit, $S(z)$, defined by the relations* (8) *and* (9)*.*

Proof of Lemma 19. According to Lemma 16, for any $z\in\mathbb{C}\backslash\mathbb{R}$,

$$\big|\mathbb{E}\big(S_{\mathbb{G}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{W}_n}(z)\big)\big| \le \frac{2}{n^2|\operatorname{Im}(z)|^3}\sum_{1\le\ell\le k\le n}\sum_{1\le j\le i\le n}|\operatorname{Cov}(G_{k,\ell},G_{i,j}) - \operatorname{Cov}(W_{k,\ell},W_{i,j})|\,.$$

Taking into account (7), we get

$$\mathbb{E}(W_{k,\ell}W_{i,j}) = \gamma_{k-i,\ell-j} + \gamma_{k-j,\ell-i}\,. \tag{53}$$

Hence,

$$\big|\mathbb{E}\big(S_{\mathbb{G}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{W}_n}(z)\big)\big| \le \frac{2}{n^2|\operatorname{Im}(z)|^3}\sum_{1\le\ell\le k\le n}\sum_{1\le j\le i\le n}|\gamma_{k-j,\ell-i}|\,.$$

Using (7) and noticing that by stationarity $\gamma_{u,v}=\gamma_{-u,-v}$ for any $(u,v)\in\mathbb{Z}^2$, we get

$$\sum_{1\le\ell\le k\le n}\sum_{1\le j\le i\le n}|\gamma_{k-j,\ell-i}| \le 2\sum_{1\le\ell\le k\le n}\sum_{1\le j\le i\le k}|\gamma_{k-j,\ell-i}|\,.$$

By simple algebra, we infer that, for any positive integer m_n less than n,

$$\big|\mathbb{E}\big(S_{\mathbb{G}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{W}_n}(z)\big)\big| \le \frac{4}{|\operatorname{Im}(z)|^3}\Big(\frac{2m_n}{n}\sum_{p=0}^{n}\sum_{q=-n}^{n}|\gamma_{p,q}| + \sum_{p\ge m_n}\sum_{q\in\mathbb{Z}}|\gamma_{p,q}|\Big)\,,$$

for any $z\in\mathbb{C}\backslash\mathbb{R}$. The lemma then follows by taking into account (6) and by selecting m_n such that $m_n\to\infty$ and $m_n/n\to 0$. $\square$

Proof of Proposition 20. The Borel-Cantelli lemma together with Theorem 17.1.1 in [15] imply that, for any $z\in\mathbb{C}\backslash\mathbb{R}$,

$$\lim_{n\to\infty}|S_{\mathbb{G}_n}(z) - \mathbb{E}\big(S_{\mathbb{G}_n}(z)\big)| = 0 \text{ and } \lim_{n\to\infty}|S_{\mathbb{W}_n}(z) - \mathbb{E}\big(S_{\mathbb{W}_n}(z)\big)| = 0 \text{ a.s.}$$

Therefore, the proposition follows by Lemma 19 combined with Theorem 2 in [13] applied to $\mathbb{E}\big(S_{\mathbb{W}_n}(z)\big)$. Indeed the entries $(W_{k,\ell})_{1\le k,\ell\le n}$ of the matrix $n^{1/2}\mathbb{W}_n$ form a *symmetric* real-valued centered Gaussian random field whose covariance function satisfies (53). Hence relation (2.8) in [13] holds. In addition, by (6) , condition (2.9) in [13] is also satisfied. At this step, the reader should notice that Theorem 2 in [13] also requires additional conditions on the covariance function $\gamma_{k,\ell}$ (this function is denoted by $B(k,\ell)$ in this latter paper), namely $\gamma_{k,\ell} = \gamma_{\ell,k} = \gamma_{\ell,-k}$. In our case, the first holds (this is (7)) but not necessarily $\gamma_{\ell,k} = \gamma_{\ell,-k}$ since by stationarity we only have $\gamma_{\ell,k} = \gamma_{-\ell,-k}$. However a careful analysis of the proof of Theorem 2 in [13] (and in particular of their auxiliary lemmas) or of the proof of Theorem 17.2.1 in [15], shows that the only condition required on the covariance function to derive the limiting equation of the Stieljes transform is the absolute summability condition (2.9) in [13]. It is noteworthy to indicate that, in Theorem 2 of [13], the symmetry conditions on the covariance function $\gamma_{k,\ell}$ must only translate the fact that the entries of the matrix form a stationary *symmetric* real-valued centered Gaussian random field, so $\gamma_{k,\ell}$ has only to satisfy $\gamma_{k,\ell} = \gamma_{\ell,k} = \gamma_{-\ell,-k}$ for any $(k,\ell) \in \mathbb{Z}^2$. □

6 Appendix: proof of Theorem 11

By using inequality (41) together with the Borel-Cantelli lemma, it follows that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty} \left|S_{\mathbb{X}_n}(z) - \mathbb{E}\big(S_{\mathbb{X}_n}(z)\big)\right| = 0 \text{ almost surely.}$$

To prove the almost sure convergence (24), it suffices then to prove that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty} \left|\mathbb{E}\big(S_{\mathbb{X}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{G}_n}(z)\big)\right| = 0\,. \tag{54}$$

We start by truncating the entries of the matrix $\mathbb{X}_n$. Since $\mathbf{A_2}$ holds, we can consider a decreasing sequence of positive numbers τ_n such that, as $n \to \infty$,

$$\tau_n \to 0\ ,\ L_n(\tau_n) \to 0 \ \text{ and } \ \tau_n\sqrt{n} \to \infty\,. \tag{55}$$

Let $\bar{\mathbf{X}}_n = [\bar{X}^{(n)}_{k,\ell}]^n_{k,\ell=1}$ be the symmetric matrix of order n whose entries are given by:

$$\bar{X}^{(n)}_{k,\ell} = X^{(n)}_{k,\ell}\mathbf{1}_{|X^{(n)}_{k,\ell}|\le\tau_n\sqrt{n}} - \mathbb{E}\big(X^{(n)}_{k,\ell}\mathbf{1}_{|X^{(n)}_{k,\ell}|\le\tau_n\sqrt{n}}\big)\,.$$

Define $\bar{\mathbb{X}}_n := n^{-1/2}\bar{\mathbf{X}}_n$. Using (55), it has been proved in Section 2.1 of [10] that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty} \left|\mathbb{E}\big(S_{\mathbb{X}_n}(z)\big) - \mathbb{E}\big(S_{\bar{\mathbb{X}}_n}(z)\big)\right| = 0\,.$$

Therefore, to prove (54) (and then the theorem), it suffices to show that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty} \left|\mathbb{E}\big(S_{\bar{\mathbb{X}}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{G}_n}(z)\big)\right| = 0\,. \tag{56}$$

The proof of (56) is then divided in three steps. The first step consists of replacing in (56), the matrix $\mathbb{G}_n$ by a symmetric matrix $\bar{\mathbb{G}}_n$ of order n whose entries are real-valued Gaussian random variables with the same covariance structure as the entries of $\bar{\mathbb{X}}_n$. The second step consists of "approximating" $\bar{\mathbb{X}}_n$ and $\bar{\mathbb{G}}_n$ by matrices with "big square independent blocks" containing the entries spaced by "small blocks" around them containing only zeros as entries. Due to the assumption $\mathbf{A_3}$, the random variables contained in two different big blocks will be independent. The third and last step consists of proving the mean convergence (56) but with $\bar{\mathbb{X}}_n$ and $\mathbb{G}_n$ replaced by their approximating matrices with independent blocks. This step will be achieved with the help of the Lindeberg method.

Step 1. Let $\bar{\mathbf{G}}_n = [\bar{G}^{(n)}_{k,\ell}]^n_{k,\ell=1}$ be the symmetric matrix of order n whose entries $(\bar{G}^{(n)}_{k,\ell}\,, 1 \le \ell \le k \le n)$ are real-valued centered Gaussian random variables with the following covariance structure: for any $1 \le \ell \le k \le n$ and any $1 \le j \le i \le n$,

$$\mathbb{E}(\bar{G}^{(n)}_{k,\ell}\bar{G}^{(n)}_{i,j}) = \mathbb{E}(\bar{X}^{(n)}_{k,\ell}\bar{X}^{(n)}_{i,j})\,. \tag{57}$$

There is no loss of generality by assuming in the rest of the proof that the σ-fields $\sigma(\bar{G}^{(n)}_{k,\ell}\,, 1 \leq \ell \leq k \leq n)$ and $\sigma(X^{(n)}_{k,\ell}\,, 1 \leq \ell \leq k \leq n)$ are independent.

Denote $\bar{\mathbb{G}}_n = \frac{1}{\sqrt{n}} \bar{\mathbf{G}}_n$. We shall prove that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty} \big| \mathbb{E}\big(S_{\bar{\mathbb{G}}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{G}_n}(z)\big) \big| = 0\,. \tag{58}$$

Applying Lemma 16, we get

$$\big| \mathbb{E}\big(S_{\bar{\mathbb{G}}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{G}_n}(z)\big) \big| \leq \frac{2}{v^3 n^2} \sum_{1\leq \ell\leq k\leq n} \sum_{1\leq j\leq i\leq n} |\mathbb{E}(G^{(n)}_{k,\ell} G^{(n)}_{i,j}) - \mathbb{E}(\bar{G}^{(n)}_{k,\ell} \bar{G}^{(n)}_{i,j})|\,, \tag{59}$$

where $v = \mathrm{Im}(z)$. Recall now that $\mathbb{E}(G^{(n)}_{k,\ell} G^{(n)}_{i,j}) = \mathbb{E}(X^{(n)}_{k,\ell} X^{(n)}_{i,j})$ and $\mathbb{E}(\bar{G}^{(n)}_{k,\ell} \bar{G}^{(n)}_{i,j}) = \mathbb{E}(\bar{X}^{(n)}_{k,\ell} \bar{X}^{(n)}_{i,j})$. Hence, setting $b_n = \tau_n \sqrt{n}$, we have

$$\mathbb{E}(G^{(n)}_{k,\ell} G^{(n)}_{i,j}) - \mathbb{E}(\bar{G}^{(n)}_{k,\ell} \bar{G}^{(n)}_{i,j}) = \mathrm{Cov}\big(X^{(n)}_{k,\ell}, X^{(n)}_{i,j}\big) - \mathrm{Cov}\big(X^{(n)}_{k,\ell} \mathbf{1}_{|X^{(n)}_{k,\ell}|\leq b_n}, X^{(n)}_{i,j} \mathbf{1}_{|X^{(n)}_{i,j}|\leq b_n}\big)\,.$$

Note that

$$\begin{aligned}
&\mathrm{Cov}\big(X^{(n)}_{k,\ell}, X^{(n)}_{i,j}\big) - \mathrm{Cov}\big(X^{(n)}_{k,\ell} \mathbf{1}_{|X^{(n)}_{k,\ell}|\leq b_n}, X^{(n)}_{i,j} \mathbf{1}_{|X^{(n)}_{i,j}|\leq b_n}\big) \\
&= \mathrm{Cov}\big(X^{(n)}_{k,\ell} \mathbf{1}_{|X^{(n)}_{k,\ell}|\leq b_n}, X^{(n)}_{i,j} \mathbf{1}_{|X^{(n)}_{i,j}|> b_n}\big) + \mathrm{Cov}\big(X^{(n)}_{k,\ell} \mathbf{1}_{|X^{(n)}_{k,\ell}|> b_n}, X^{(n)}_{i,j} \mathbf{1}_{|X^{(n)}_{i,j}|> b_n}\big) \\
&\quad + \mathrm{Cov}\big(X^{(n)}_{k,\ell} \mathbf{1}_{|X^{(n)}_{k,\ell}|> b_n}, X^{(n)}_{i,j} \mathbf{1}_{|X^{(n)}_{i,j}|\leq b_n}\big)
\end{aligned}$$

implying, by Cauchy-Schwarz's inequality, that

$$\begin{aligned}
&\big|\mathrm{Cov}\big(X^{(n)}_{k,\ell}, X^{(n)}_{i,j}\big) - \mathrm{Cov}\big(X^{(n)}_{k,\ell} \mathbf{1}_{|X^{(n)}_{k,\ell}|\leq b_n}, X^{(n)}_{i,j} \mathbf{1}_{|X^{(n)}_{i,j}|\leq b_n}\big)\big| \\
&\leq 2 b_n \mathbb{E}\big(|X^{(n)}_{i,j}| \mathbf{1}_{|X^{(n)}_{i,j}|> b_n}\big) + 2 b_n \mathbb{E}\big(|X^{(n)}_{k,\ell}| \mathbf{1}_{|X^{(n)}_{k,\ell}|> b_n}\big) + 2 \|X^{(n)}_{i,j} \mathbf{1}_{|X^{(n)}_{i,j}|> b_n}\|_2 \|X^{(n)}_{k,\ell} \mathbf{1}_{|X^{(n)}_{k,\ell}|> b_n}\|_2 \\
&\leq 3\,\mathbb{E}\big(|X^{(n)}_{i,j}|^2 \mathbf{1}_{|X^{(n)}_{i,j}|> b_n}\big) + 3\,\mathbb{E}\big(|X^{(n)}_{k,\ell}|^2 \mathbf{1}_{|X^{(n)}_{k,\ell}|> b_n}\big)\,.
\end{aligned}$$

Note also that, by assumption $\mathbf{A}_3$,

$$\begin{aligned}
&\big|\mathrm{Cov}\big(X^{(n)}_{k,\ell} \mathbf{1}_{|X^{(n)}_{k,\ell}|\leq b_n}, X^{(n)}_{i,j} \mathbf{1}_{|X^{(n)}_{i,j}|\leq b_n}\big) - \mathrm{Cov}\big(X^{(n)}_{k,\ell}, X^{(n)}_{i,j}\big)\big| \\
&\quad = \mathbf{1}_{i\in[k-K,k+K]} \mathbf{1}_{j\in[\ell-K,\ell+K]} \big|\mathrm{Cov}\big(X^{(n)}_{k,\ell} \mathbf{1}_{|X^{(n)}_{k,\ell}|\leq b_n}, X^{(n)}_{i,j} \mathbf{1}_{|X^{(n)}_{i,j}|\leq b_n}\big) - \mathrm{Cov}\big(X^{(n)}_{k,\ell}, X^{(n)}_{i,j}\big)\big|\,.
\end{aligned}$$

So, overall,

$$\begin{aligned}
\big|\mathbb{E}(G^{(n)}_{k,\ell} G^{(n)}_{i,j}) - \mathbb{E}(\bar{G}^{(n)}_{k,\ell} \bar{G}^{(n)}_{i,j})\big| &\leq 3\,\mathbb{E}\big(|X^{(n)}_{i,j}|^2 \mathbf{1}_{|X^{(n)}_{i,j}|> b_n}\big) \mathbf{1}_{k\in[i-K,i+K]} \mathbf{1}_{\ell\in[j-K,j+K]} \\
&\quad + 3\,\mathbb{E}\big(|X^{(n)}_{k,\ell}|^2 \mathbf{1}_{|X^{(n)}_{k,\ell}|> b_n}\big) \mathbf{1}_{i\in[k-K,k+K]} \mathbf{1}_{j\in[\ell-K,\ell+K]}\,.
\end{aligned}$$

Hence, starting from (59) and taking into account the above inequality, we derive that

$$\big| \mathbb{E}\big(S_{\bar{\mathbb{G}}_n}(z)\big) - \mathbb{E}\big(S_{\mathbb{G}_n}(z)\big) \big| \leq \frac{12}{n^2 v^3} (2K+1)^2 \sum_{k=1}^{n} \sum_{\ell=1}^{k} \mathbb{E}\big(|X^{(n)}_{k,\ell}|^2 \mathbf{1}_{|X^{(n)}_{k,\ell}|> b_n}\big)\,.$$

which converges to zero as n tends to infinity, by assumption $\mathbf{A}_2$. This ends the proof of (58).

Step 2: Reduction to matrices with independent blocks. Let $p := p_n$ such that $p_n \to \infty$, $p_n/n \to 0$, and $\tau_n p_n^4 \to 0$. Clearly we can take $p_n > K$, $p_n + K \le n/3$, and set

$$q = q_n = \left[\frac{n}{p+K}\right] - 1\,. \tag{60}$$

Let

$$I_\ell = \big[\ell(p+K)+1\,,\ \ell(p+K)+p\big] \cap \mathbb{N} \ \text{ and } \mathcal{E}_{k,\ell} = \{(u,v) \in I_k \times I_{\ell-1}\}\,. \tag{61}$$

For any $k = 1, \ldots, q_n$ and any $\ell = 1, \ldots, k$, we define now real matrices $B_{k,\ell}$ of size $p \times p$ whose entries consist of all the $\bar{X}^{(n)}_{u,v}$ for $(u,v) \in \mathcal{E}_{k,\ell}$. More precisely,

$$B_{k,\ell} := \big\{b_{k,\ell}(i,j)\big\}_{i,j=1}^{p} \ \text{ where } b_{k,\ell}(i,j) = \bar{X}^{(n)}_{k(p+K)+i,(\ell-1)(p+K)+j}\,. \tag{62}$$

Similarly, we define real matrices $B^*_{k,\ell}$ of size $p \times p$ whose entries consist of all the $\bar{G}^{(n)}_{u,v}$ for $(u,v) \in \mathcal{E}_{k,\ell}$. Therefore,

$$B^*_{k,\ell} := \big\{b^*_{k,\ell}(i,j)\big\}_{i,j=1}^{p} \ \text{ where } b^*_{k,\ell}(i,j) = \bar{G}^{(n)}_{k(p+K)+i,(\ell-1)(p+K)+j}\,. \tag{63}$$

Using the blocks $B_{k,\ell}$ we construct now a $n \times n$ matrix, $\widetilde{\mathbf{X}}_n$, by inserting 0's. Actually we start from the matrix $\overline{\mathbf{X}}_n$, keep the blocks $B_{k,\ell}$ and $B^T_{k,\ell}$ and replace all the other variables by 0's. For the sake of clarity we describe the south-western part of the matrix $\widetilde{\mathbf{X}}_n$ below, the other part being constructed by symmetry.

$$\widetilde{\mathbf{X}}_n := \begin{pmatrix}
\mathbf{0}_{p,p} & & & & & \cdots & & & \\
\mathbf{0}_{K,p} & \mathbf{0}_{K,K} & & & & \cdots & & & \\
B_{1,1} & \mathbf{0}_{p,K} & \mathbf{0}_{p,p} & & & \cdots & & & \\
\mathbf{0}_{K,p} & \mathbf{0}_{K,K} & \mathbf{0}_{K,p} & \mathbf{0}_{K,K} & & \cdots & & & \\
B_{2,1} & \mathbf{0}_{p,K} & B_{2,2} & \mathbf{0}_{p,K} & \mathbf{0}_{p,p} & \cdots & & & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
B_{q-1,1} & \mathbf{0}_{p,K} & B_{q-1,2} & \mathbf{0}_{p,K} & B_{q-1,3} & \cdots & & & \\
\mathbf{0}_{K,p} & \mathbf{0}_{K,K} & \mathbf{0}_{K,p} & \mathbf{0}_{K,K} & \mathbf{0}_{K,p} & \cdots & \mathbf{0}_{K,K} & & \\
B_{q,1} & \mathbf{0}_{p,K} & B_{q,2} & \mathbf{0}_{p,K} & B_{q,3} & \cdots & B_{q,q} & \mathbf{0}_{p,p} & \\
\mathbf{0}_{m,p} & \mathbf{0}_{m,K} & \mathbf{0}_{m,p} & \mathbf{0}_{m,K} & \mathbf{0}_{m,p} & \cdots & \mathbf{0}_{m,K} & \mathbf{0}_{m,p} & \mathbf{0}_{m,m}
\end{pmatrix}, \tag{64}$$

where $m = m_n = n - q(p+K) - p$. $\widetilde{\mathbf{G}}_n$ is constructed as $\widetilde{\mathbf{X}}_n$ with the $B^*_{k,\ell}$ in place of the $B_{k,\ell}$.

In what follows, we shall prove that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty} \big|\mathbb{E}\big(S_{\bar{\mathbb{X}}_n}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n}(z)\big)\big| = 0 \ \text{ and } \ \lim_{n\to\infty} \big|\mathbb{E}\big(S_{\bar{\mathbb{G}}_n}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{G}}_n}(z)\big)\big| = 0\,, \tag{65}$$

with $\widetilde{\mathbb{X}}_n := n^{-1/2}\widetilde{\mathbf{X}}_n$, $\widetilde{\mathbb{G}}_n := n^{-1/2}\widetilde{\mathbf{G}}_n$.

To prove it we first introduce two other symmetric $n \times n$ matrices $\widehat{\mathbf{X}}_n = [\widehat{X}^{(n)}_{k,\ell}]_{k,\ell=1}^n$ and $\widehat{\mathbf{G}}_n = [\widehat{G}^{(n)}_{k,\ell}]_{k,\ell=1}^n$ constructed from $\bar{\mathbf{X}}_n$ and $\bar{\mathbf{G}}_n$ respectively, by replacing the entries by zeros in square blocks of size p around the diagonal. More precisely, for any $1 \le i, j \le n$,

$$\widehat{X}^{(n)}_{i,j} = 0 \ \text{ if } (i,j) \in \cup_{\ell=0}^{q_n} \mathcal{E}_{\ell,\ell+1} \ \text{ and } \ \widehat{X}^{(n)}_{i,j} = \bar{X}^{(n)}_{i,j} \ \text{ otherwise}$$

and

$$\widehat{G}^{(n)}_{i,j} = 0 \ \text{ if } (i,j) \in \cup_{\ell=0}^{q_n} \mathcal{E}_{\ell,\ell+1} \ \text{ and } \ \widehat{G}^{(n)}_{i,j} = \bar{G}^{(n)}_{i,j} \ \text{ otherwise}$$

where we recall that the sets $\mathcal{E}_{\ell,\ell+1}$ have been defined in (61). Denote now $\widehat{\mathbb{X}}_n = \frac{1}{\sqrt{n}}\widehat{\mathbf{X}}_n$ and $\widehat{\mathbb{G}}_n = \frac{1}{\sqrt{n}}\widehat{\mathbf{G}}_n$.

By Lemma 14, we get, for any $z = u + \mathrm{i}v \in \mathbb{C}^+$, that

$$\big|\mathbb{E}\big(S_{\bar{\mathbb{X}}_n}(z)\big) - \mathbb{E}\big(S_{\widehat{\mathbb{X}}_n}(z)\big)\big|^2 \le \mathbb{E}\big(\big|S_{\bar{\mathbb{X}}_n}(z) - S_{\widehat{\mathbb{X}}_n}(z)\big|^2\big) \le \frac{1}{n^2 v^4}\mathbb{E}\big(\mathrm{Tr}\big((\bar{\mathbf{X}}_n - \widehat{\mathbf{X}}_n)^2\big)\big)\,.$$

Therefore,

$$\big|\mathbb{E}\big(S_{\bar{\mathbb{X}}_n}(z)\big) - \mathbb{E}\big(S_{\widehat{\mathbb{X}}_n}(z)\big)\big|^2 \le \frac{1}{n^2 v^4}\sum_{\ell=0}^{q_n}\sum_{(i,j)\in\mathcal{E}_{\ell,\ell+1}}\mathbb{E}(|\bar{X}_{i,j}^{(n)}|^2)\,.$$

But $\|\bar{X}_{i,j}^{(n)}\|_\infty \le 2\tau_n\sqrt{n}$. Hence,

$$\big|\mathbb{E}\big(S_{\bar{\mathbb{X}}_n}(z)\big) - \mathbb{E}\big(S_{\widehat{\mathbb{X}}_n}(z)\big)\big|^2 \le \frac{4}{n^2 v^4}(q_n+1)p_n^2\tau_n^2 n \le \frac{4}{v^4}\tau_n^2 p_n\,.$$

By our selection of p_n, we obviously have that $\tau_n^2 p_n \to 0$ as $n \to \infty$. It follows that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty}\big|\mathbb{E}\big(S_{\bar{\mathbb{X}}_n}(z)\big) - \mathbb{E}\big(S_{\widehat{\mathbb{X}}_n}(z)\big)\big| = 0\,. \tag{66}$$

With similar arguments, we get that, for any $z = u + \mathrm{i}v \in \mathbb{C}^+$,

$$\big|\mathbb{E}\big(S_{\bar{\mathbb{G}}_n}(z)\big) - \mathbb{E}\big(S_{\widehat{\mathbb{G}}_n}(z)\big)\big|^2 \le \frac{1}{n^2 v^4}\sum_{\ell=0}^{q_n}\sum_{(i,j)\in\mathcal{E}_{\ell,\ell}}\mathbb{E}(|\bar{G}_{i,j}^{(n)}|^2)\,.$$

But $\|\bar{G}_{i,j}^{(n)}\|_2 = \|\bar{X}_{i,j}^{(n)}\|_2$. So, as before, we derive that for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty}\big|\mathbb{E}\big(S_{\bar{\mathbb{G}}_n}(z)\big) - \mathbb{E}\big(S_{\widehat{\mathbb{G}}_n}(z)\big)\big| = 0\,. \tag{67}$$

From (66) and (67), the mean convergence (65) follows if we prove that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty}\big|\mathbb{E}\big(S_{\widehat{\mathbb{X}}_n}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n}(z)\big)\big| = 0 \ \text{ and } \ \lim_{n\to\infty}\big|\mathbb{E}\big(S_{\widehat{\mathbb{G}}_n}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{G}}_n}(z)\big)\big| = 0\,, \tag{68}$$

For proving it, we shall use rank inequalities. Indeed, notice first that, for any $z = u + \mathrm{i}v \in \mathbb{C}^+$,

$$\begin{aligned}\big|S_{\widehat{\mathbb{X}}_n}(z) - S_{\widetilde{\mathbb{X}}_n}(z)\big| &= \Big|\int \frac{1}{x-z}dF^{\widehat{\mathbb{X}}_n}(x) - \int \frac{1}{x-z}dF^{\widetilde{\mathbb{X}}_n}(x)\Big| \\ &\le \Big|\int \frac{F^{\widehat{\mathbb{X}}_n}(x) - F^{\widetilde{\mathbb{X}}_n}}{(x-z)^2}dx\Big| \le \frac{\pi\,\big\|F^{\widehat{\mathbb{X}}_n} - F^{\widetilde{\mathbb{X}}_n}\big\|_\infty}{v}\,.\end{aligned}$$

Hence, by Theorem A.43 in Bai and Silverstein [2],

$$\big|S_{\widehat{\mathbb{X}}_n}(z) - S_{\widetilde{\mathbb{X}}_n}(z)\big| \le \frac{\pi}{vn}\mathrm{rank}\big(\widehat{\mathbb{X}}_n - \widetilde{\mathbb{X}}_n\big)\,.$$

But, by counting the numbers of rows and of columns with entries that can be different from zero, we infer that

$$\mathrm{rank}\big(\widehat{\mathbb{X}}_n - \widetilde{\mathbb{X}}_n\big) \le 2(q_n K + m_n) \le 2(np^{-1}K + p + 2K)\,.$$

Therefore,

$$\big|S_{\widehat{\mathbb{X}}_n}(z) - S_{\widetilde{\mathbb{X}}_n}(z)\big| \le \frac{2\pi}{v}(Kp^{-1} + pn^{-1} + 2Kn^{-1})\,.$$

With similar arguments, we get

$$\big|S_{\widehat{\mathbb{G}}_n}(z) - S_{\widetilde{\mathbb{G}}_n}(z)\big| \le \frac{2\pi}{v}(Kp^{-1} + pn^{-1} + 2Kn^{-1})\,.$$

Since $p = p_n \to \infty$ and $p_n/n \to 0$, as $n \to \infty$, (68) (and then (65)) follows from the two above inequalities. Therefore, to prove that the mean convergence (56) holds, it suffices to prove that, for any $z \in \mathbb{C}^+$,

$$\lim_{n\to\infty} \left|\mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{G}}_n}(z)\big)\right| = 0\,. \tag{69}$$

This is done in the next step.

Step 3: Lindeberg method. To prove (69), we shall use the Lindeberg method. Recall that the σ-fields $\sigma(\bar{G}^{(n)}_{k,\ell}\,,\, 1 \le \ell \le k \le n)$ and $\sigma(X^{(n)}_{k,\ell}\,,\, 1 \le \ell \le k \le n)$ are assumed to be independent. Furthermore, by the hypothesis $\mathbf{A}_3$, all the blocks $(B_{k,\ell})$ and $(B^*_{k,\ell})$ $1 \le \ell \le k \le q$ are independent.

The Lindeberg method consists of writing the difference of expectations as telescoping sums and using Taylor expansions. This method can be used in the context of random matrices since the function f, defined in (49), admits partial derivatives of all orders (see the equality (17) in Chatterjee [8]). In the traditional Lindeberg method, the telescoping sums consist of replacing one by one the random variables, involved in a partial sum, by a Gaussian random variable. Here, we shall replace one by one the blocks $B_{k,\ell}$ by the "Gaussian" ones $B^*_{k,\ell}$ with the same covariance structure. So, starting from the matrix $\widetilde{\mathbf{X}}_n = \widetilde{\mathbf{X}}_n(0)$, the first step is to replace its block B_{q_n,q_n} by $B^*_{q_n,q_n}$, this gives a new matrix. Note that, at the same time, $B^T_{q_n,q_n}$ will also be replaced by $(B^*_{q_n,q_n})^T$. We denote this matrix by $\widetilde{\mathbf{X}}_n(1)$ and re-denote the block replaced by $B(1)$ and the new one by $B^*(1)$. At the second step, we replace, in the new matrix $\widetilde{\mathbf{X}}_n(1)$, the block $B(2) := B_{q_n,q_n-1}$ by $B^*(2) := B^*_{q_n,q_n-1}$, and call the new matrix $\widetilde{\mathbf{X}}_n(2)$ and so on. Therefore, after the q_n-th step, in the matrix $\widetilde{\mathbf{X}}_n$ we have replaced the blocks $B(q_n - \ell + 1) = B_{q_n,\ell}\,,\ \ell = 1, \dots, q_n$ (and their transposed) by the blocks $B^*(q_n - \ell + 1) = B^*_{q_n,\ell}\,,\ \ell = 1, \dots, q_n$ (and their transposed) respectively. This matrix is denoted by $\widetilde{\mathbf{X}}_n(q_n)$. Next, the $q_n + 1$-th step will consist of replacing the block $B(q_n + 1) = B_{q_n-1,q_n-1}$ by $B^*_{q_n-1,q_n-1}$ and obtain the matrix $\widetilde{\mathbf{X}}_n(q_n + 1)$. So finally after $q_n(q_n + 1)/2$ steps, we have replaced all the blocks $B_{k,\ell}$ and $B^T_{k,\ell}$ of the matrix $\widetilde{\mathbb{X}}_n$ to obtain at the end the matrix $\widetilde{\mathbf{X}}_n(q_n(q_n + 1)/2) = \widetilde{\mathbb{G}}_n$.

Therefore we have

$$\mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{G}}_n}(z)\big) = \sum_{k=1}^{k_n} \Big(\mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n(k-1)}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n(k)}(z)\big)\Big)\,. \tag{70}$$

where $k_n = q_n(q_n + 1)/2$.

Let k in $\{1, \dots, k_n\}$. Observe that $\widetilde{\mathbb{X}}_n(k-1)$ and $\widetilde{\mathbb{X}}_n(k)$ differ only by the variables in the block $B(k)$ replaced at the step k. Define then the vector $\mathbf{X}$ of $\mathbb{R}^{p^2}$ consisting of all the entries of $B(k)$, the vector $\mathbf{Y}$ of $\mathbb{R}^{p^2}$ consisting of all the entries of $B^*(k)$ (in the same order we have defined the coordinates of $\mathbf{X}$). Denote by $\mathbf{Z}$ the vector of $\mathbb{R}^{N-p^2}$ (where $N = n(n+1)/2$) consisting of all the entries on and below the diagonal of $\widetilde{\mathbf{X}}_n(k-1)$ except the ones that are in the block matrix $B(k)$. More precisely if (u, v) are such that $B(k) = B_{u,v}$, then

$$\mathbf{X} = \big((b_{u,v}(i,j))_{j=1,\dots,p}\,,\, i = 1, \dots, p\big) \quad \text{and} \quad \mathbf{Y} = \big((b^*_{u,v}(i,j))_{j=1,\dots,p}\,,\, i = 1, \dots, p\big)$$

where $b_{u,v}(i,j)$ and $b^*_{u,v}(i,j)$ are defined in (62) and (63) respectively. In addition,

$$\mathbf{Z} = \big((\widetilde{\mathbf{X}}_n(k-1))_{i,j} \,:\, 1 \le j \le i \le n\,,\, (i,j) \notin \mathcal{E}_{u,v}\big)\,,$$

where $\mathcal{E}_{u,v}$ is defined in (61). The notations above allow to write

$$\mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n(k-1)}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n(k)}(z)\big) = \mathbb{E}f(\pi(\mathbf{X}, \mathbf{Z})) - \mathbb{E}f(\pi(\mathbf{Y}, \mathbf{Z}))\,,$$

where f is the function from $\mathbb{R}^N$ to $\mathbb{C}$ defined by (49) and $\pi : \mathbb{R}^N \to \mathbb{R}^N$ is a certain permutation. Note that, by our hypothesis $\mathbf{A}_3$ and our construction, the vectors $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are independent. Moreover $\mathbf{X}$ and $\mathbf{Y}$ are centered at expectation, have the same covariance structure and finite moments of order 3. Applying then Lemma 17 from Section 5 and taking into account (50), we derive that, for a constant C depending only on $\mathrm{Im}(z)$,

$$\left|\mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n(k-1)}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n(k)}\big)\right| \leq \frac{Cp^4}{n^{5/2}} \sum_{(i,j)\in\mathcal{E}_{u,v}} \Big(\mathbb{E}(|\bar{X}^{(n)}_{i,j}|^3) + \mathbb{E}(|\bar{G}^{(n)}_{i,j}|^3)\Big) \,.$$

So, overall,

$$\sum_{k=1}^{k_n} \left|\mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n(k-1)}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n(k)}\big)\right| \leq \frac{Cp^4}{n^{5/2}} \sum_{1\leq \ell\leq k\leq q}\ \sum_{(i,j)\in\mathcal{E}_{k,\ell}} \Big(\mathbb{E}(|\bar{X}^{(n)}_{i,j}|^3) + \mathbb{E}(|\bar{G}^{(n)}_{i,j}|^3)\Big) \,. \tag{71}$$

By taking into account that

$$\mathbb{E}(|\bar{X}^{(n)}_{i,j}|^3) \leq 2\tau_n\sqrt{n}\,\mathbb{E}(|X^{(n)}_{i,j}|^2)$$

and also

$$\mathbb{E}(|\bar{G}^{(n)}_{i,j}|^3) \leq 2\left(\mathbb{E}(|\bar{G}^{(n)}_{i,j}|^2)\right)^{3/2} = 2\left(\mathbb{E}(|\bar{X}^{(n)}_{i,j}|^2)\right)^{3/2} \leq 4\tau_n\sqrt{n}\,\mathbb{E}(|X^{(n)}_{i,j}|^2)\,,$$

it follows from (70) and (71) that, for a constant C' depending only on $\mathrm{Im}(z)$,

$$\left|\mathbb{E}\big(S_{\widetilde{\mathbb{X}}_n}(z)\big) - \mathbb{E}\big(S_{\widetilde{\mathbb{G}}_n}(z)\big)\right| \leq \frac{C'p^4}{n^2}\tau_n \sum_{1\leq j\leq i\leq n} \mathbb{E}(|X^{(n)}_{i,j}|^2)$$

which converges to 0 by $\mathbf{A}_1$ and the selection of p_n. This ends the proof of (69) and then of the theorem. $\square$

Acknowledgements. The authors would like to thank the referee for carefully reading the manuscript and J. Najim for helpful discussions.

References

[1] Anderson, G. and Zeitouni, O. (2008). A law of large numbers for finite-range dependent random matrices *Comm. Pure Appl. Math.* **61** 1118-1154.

[2] Bai, Z. and Silverstein, J.W. (2010). *Spectral analysis of large dimensional random matrices.* Springer, New York, second edition.

[3] Bai, Z. and Zhou, W. (2008). Large sample covariance matrices without independence structures in columns. *Statist. Sinica* **18** 425-442.

[4] Banna, M. and Merlevède, F. (2013). Limiting spectral distribution of large sample covariance matrices associated with a class of stationary processes. To appear in *J. Theoret. Probab.* (DOI: 10.1007/s10959-013-0508-x)

[5] Boutet de Monvel, A. and Khorunzhy, A. (1999). On the Norm and Eigenvalue Distribution of Large Random Matrices. *Ann. Probab.* **27** 913-944.

[6] Boutet de Monvel, A. Khorunzhy, A. and Vasilchuk, V. (1996). Limiting eigenvalue distribution of random matrices with correlated entries. *Markov Process. Related Fields* **2** 607-636.

[7] Chakrabarty A., Hazra R.S. and Sarkar D. (2014). From random matrices to long range dependence. *arXiv:math/1401.0780.*

[8] Chatterjee, S. (2006). A generalization of the Lindeberg principle. *Ann. Probab.* **34** 2061-2076.

[9] Girko, V. L. (1990). Theory of Random Determinants. Translated from the Russian. *Mathematics and Its Applications (Soviet Series)* **45**. Kluwer Academic Publishers Group, Dordrecht.

[10] Götze, F., Naumov, A. and A. Tikhomirov (2012). Semicircle law for a class of random matrixes with dependent entries. arXiv:math/0702386v1.

[11] Guntuboyina, A. and Leeb, H. (2009). Concentration of the spectral measure of large Wishart matrices with dependent entries. *Electron. Commun. Probab.* **14** 334-342.

[12] Hachem, W., Loubaton, P. and J. Najim (2005). The empirical eigenvalue distribution of a Gram matrix: from independence to stationarity, *Markov Process. Related Fields* **11** 629–648.

[13] Khorunzhy, A. and Pastur, L. (1994). On the eigenvalue distribution of the deformed Wigner ensemble of random matrices. In: V. A. Marchenko (ed.), *Spectral Operator Theory and Related Topics*, Adv. Soviet Math. **19**, Amer. Math. Soc., Providence, RI, 97-127.

[14] Merlevède, F. and Peligrad, M. (2014). On the empirical spectral distribution for matrices with long memory and independent rows. *arXiv: 1406.1216*

[15] Pastur, L. and Shcherbina, M. (2011). Eigenvalue distribution of large random matrices. *Mathematical Surveys and Monographs*, **171**. American Mathematical Society, Providence, RI.

[16] Rashidi Far, R., Oraby T., Bryc, W. and Speicher, R. (2008). On slow-fading MIMO systems with nonseparable correlation. *IEEE Trans. Inform. Theory* **54** 544-553.

[17] Talagrand M. (2010). *Mean Field Models for Spin Glasses. Vol 1. Basic Examples.* Springer.

[18] Yao, J. (2012). A note on a Marčenko-Pastur type theorem for time series. *Statist. Probab. Lett.* **82** 22-28.