Universality of limiting spectral distribution under projective criteria.

Florence Merlevède and Magda Peligrad

Université Paris-Est, LAMA and CNRS UMR 8050, Email: florence.merlevede@u-pem.fr
Department of Mathematical Sciences, University of Cincinnati, PO Box 210025, Cincinnati, Oh 45221-0025, USA. email: peligrm@ucmail.uc.edu

Abstract

This paper has double scope. In the first part we study the limiting empirical spectral distribution of a $n \times n$ symmetric matrix with dependent entries. For a class of generalized martingales we show that the asymptotic behavior of the empirical spectral distribution depends only on the covariance structure. Applications are given to strongly mixing random fields. The technique is based on a blend of blocking procedure, martingale techniques and multivariate Lindeberg's method. This means that, for this class, the study of the limiting spectral distribution is reduced to the Gaussian case. The second part of the paper contains a survey of several old and new asymptotic results for the empirical spectral distribution for Gaussian processes, which can be combined with our universality results.

1 Introduction.

The distribution of the eigenvalues of random matrices is useful in many fields of science such as statistics, physics and engineering. The celebrated paper by Wigner (1958) deals with symmetric matrices having i.i.d entries below the diagonal. Wigner proved a global universality result, showing that, asymptotically and with probability one, the empirical distribution of eigenvalues is distributed according to the semicircle law (see Chapter 2 in Bai and Silverstein (2010) for more details). The only parameter of this law is the variance of an entry. This result was expanded in various directions. The first generalization was to decrease the degree of stationarity by replacing the condition of equal variance by weaker assumptions of the Lindeberg's type. Another direction of generalization deals with weakening the hypotheses of independence by considering various notions of weak dependence. For symmetric Gaussian matrices with correlated entries, works of Khorunzhy and Pastur (1994), Boutet de Monvel et al. (1996), Boutet de Monvel and Khorunzhy (1999), Chakrabarty et al. (2016), Peligrad and Peligrad (2016) showed that the limiting spectral distribution of the symmetric matrix depends only on the covariance structure of the underlying Gaussian process. The limiting spectral distribution is rather complicated and the best way to describe it is by specifying an equation satisfied by its Stieltjes transform.

A way to symmetrize a matrix is to multiply it with its transpose. These matrices, known under the name of Gram matrices or sample covariance matrices, play an important role in statistical studies of large data sets. The spectral analysis of large-dimensional sample covariance matrices has been actively studied starting with the seminal work of Marchenko and Pastur (1967) who considered independent random samples from an independent multidimensional vector. A big step forward was the study of the dependent case represented in numerous papers. Basically, the entries of the matrix were allowed to be linear combinations of an independent sequence. The first paper where such a model was considered is by Yin and Krishnaiah (1983) followed by important contributions by Yin (1986), Silverstein (1995), Silverstein and Bai (1995), Hachem et al. (2005), Pfaffel and Schlemm (2011), Yao (2012), Pan et al. (2014), Davis et al. (2014), among many others. Another type of models was considered by Bai and Zhou (2008) based on independent columns. The dependence type-condition imposed to the colums is in particular satisfied for isotropic vectors with log-concave distribution (see Pajor and Pastur (2009)) but may be hard to verify for non linear time series (such that ARCH models) or requires rate of convergence of mixing coefficients. Let us also mention the recent papers by Yaskov (2016-a, 2016-b) where a weaker version of the Bai-Zhou's dependence type condition has been introduced.

In two recent papers Banna et al. (2015) and Merlevède-Peligrad (2016), have shown that, for two situations, namely for symmetric matrices whose entries are functions of independent and identically distributed random fields or for large sample covariance matrices generated by random matrices with independent rows, the limiting spectral distribution of eigenvalues counting measure always exists and can be described via an equation satisfied by its Stieltjes transform.

Even if many models encountered in time series analysis can be rewritten as functions of an i.i.d. sequence, this assumption is not completely satisfactory since many stationary processes, even with trivial left sigma field, cannot be in general represented as a function of an i.i.d sequence, as shown for instance in Rosenblatt (2009). Moreover, the assumption of independence of the rows or of the columns generating the large sample covariance matrices may be too restrictive.

The main goal of our paper is then to continue the study of the asymptotic behavior of the empirical eigenvalues distribution of symmetric matrices and large sample covariance matrices associated with random fields when the variables are not necessarily functions of an i.i.d. sequence or when the rows (or columns) are not necessarily independent. In the first part of the paper we shall show that the universality results hold for both symmetric and symmetrized random matrices when the dependence is controlled by projective type coefficients. These coefficients are easy to estimate in terms of strong mixing coefficients. By "universality" we mean that the limiting distribution of the eigenvalues counting measure depends only of the process' covariance structure. Therefore our result reduces the study of the limiting spectral distribution (LSD) to the case where the entries of the underlying matrix are observations of a Gaussian random field with the same covariance structure. In the second part of the paper we survey old and new results for the Gaussian case, which one can combine with the universality theorems, for obtaining the existence and the characterization of LSD.

Our paper is organized as follows. Section 2 contains the notations and the universality results. In Section 3 we apply our results to classes of strongly mixing random fields. Then, Section 4 is dedicated to LSD results for symmetrized matrices when the entries are observations of a Gaussian random field. All the proofs are postponed to Section 5. Several auxiliary results needed in the proofs are given in Section 6.

Here are some notations used all along the paper. The notation $[x]$ is used to denote the integer part of a real x. The notation $\mathbf{0}_{p,q}$ means a matrix of size $p \times q$, $(p,q) \in \mathbb{N}^2$ with entries 0. For a matrix A, we denote by A^T its transpose matrix, by $\mathrm{Tr}(A)$ its trace. We shall also use the notation $\|X\|_r$ for the $\mathbb{L}^r$-norm ($r \geq 1$) of a real valued random variable X. For two sequences of positive numbers (a_n) and (b_n) the notation $a_n \ll b_n$ means that there is a constant C such that $a_n \leq C b_n$ for all $n \in \mathbb{N}$. We use bold small letters to denote an element of $\mathbb{Z}^2$, hence $\mathbf{u} = (u_1, u_2) \in \mathbb{Z}^2$. For $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ in $\mathbb{Z}^2$, the following notations will be used: $|\mathbf{u} - \mathbf{v}| = \max(|u_1 - v_1|, |u_2 - v_2|)$ and $\mathbf{u} \wedge \mathbf{v} = (u_1 \wedge u_2, v_1 \wedge v_2)$ (where $u_1 \wedge u_2 = \min(u_1, u_2)$).

2 Results

Let $(X_{\mathbf{u}})_{\mathbf{u} \in \mathbb{N}^2}$ be a real-valued random field defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We consider the symmetric $n \times n$ random matrix $\mathbf{X}_n$ such that, for any i and j in $\{1, \ldots, n\}$

$$(\mathbf{X}_n)_{ij} = X_{ij} \ \text{ for } i \geq j \ \text{ and } \qquad (1)$$
$$(\mathbf{X}_n)_{ij} = X_{ji} \ \text{ for } i < j \, .$$

Denote by $\lambda_1^n \leq \cdots \leq \lambda_n^n$ the eigenvalues of

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

and define its spectral distribution function by

$$\mathbf{F}^{\mathbb{X}_n}(t) = \frac{1}{n} \sum_{1 \leq k \leq n} I(\lambda_k \leq t) \,,$$

where $I(A)$ denotes the indicator of an event A. The Stieltjes transform of $\mathbb{X}_n$ is given by

$$S^{\mathbb{X}_n}(z) = \int \frac{1}{x-z} dF^{\mathbb{X}_n}(x) = \frac{1}{n} \mathrm{Tr}(\mathbb{X}_n - z\mathbf{I}_n)^{-1} \,, \tag{3}$$

where $z = u + iv \in \mathbb{C}^+$ (the set of complex numbers with positive imaginary part), and $\mathbf{I}_n$ is the identity matrix of order n. In particular, if the random field is an array of i.i.d. random variables with variance $\sigma^2 > 0$, then Wigner (1958) proved that, with probability one, and for any $z \in \mathbb{C}^+$, $S^{\mathbb{X}_n}(z)$ converges to $S(z)$, which satisfies the equation $\sigma^2 S^2 + S + z^{-1} = 0$. Its solution

$$S(z) = -(z - \sqrt{z^2 - 4\sigma^2})(2\sigma^2)^{-1} \tag{4}$$

is the well-known Stieltjes transform of the semicircle law, which has the density

$$g(x) = \frac{1}{2\pi\sigma^2} \sqrt{4\sigma^2 - x^2} I(|x| \leq 2\sigma) \,.$$

(See for instance Theorem 2.5 in Bai-Silverstein (2010)). Note that it is not necessary for the random variables to have the same law for this result to hold. Indeed, if the random field $(X_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$ is an array of independent centered random variables with common positive variance σ^2, which satisfies the Lindeberg's condition given in Condition 1 below, then for all $z \in \mathbb{C}^+$, $S^{\mathbb{X}_n}(z)$ converges almost surely to the Stieltjes transform of the semicircle law with parameter σ^2 (see for instance Theorem 2.9 in Bai and Silverstein (2010)). Note that the necessity of the Lindeberg's condition has been stated in Girko's book (1990).

Another way to state the Wigner's result is to say that the Lévy distance between the distribution function $\mathbf{F}^{\mathbb{X}_n}$ and G, defined by $G(x) = \int_{-\infty}^x g(u)du$, converges to zero almost surely. Recall that the Lévy metric d between two distribution functions F and G, defined by

$$d(F,G) = \inf\{\varepsilon > 0 \; : \; F(x-\varepsilon) - \varepsilon \leq G(x) \leq F(x+\varepsilon) + \varepsilon \,, \, \forall x \in \mathbb{R}\} \,.$$

The aim of this paper is to specify a class of random fields for which the limiting behavior of $\mathbf{F}^{\mathbb{X}_n}(t)$ depends only on the covariances of the random variables $(X_{\mathbf{u}})_{\mathbf{u}\in\mathbb{N}^2}$ and not on the structural dependence structure. In other words, we shall show that the limiting spectral distribution of $\mathbf{F}^{\mathbb{X}_n}(t)$ can be deduced from that one of $\mathbf{F}^{\mathbb{Y}_n}(t)$ where $\mathbb{Y}_n$ is a Gaussian matrix with the same covariance structure as $\mathbb{X}_n$. Since the estimate of the Lévy distance between $\mathbf{F}^{\mathbb{X}_n}$ and $\mathbf{F}^{\mathbb{Y}_n}$ can be given in terms of their Stieltjes transforms (see, for instance, Theorem B.12 and Lemma B.18 in Bai and Silverstein (2010) or Proposition 2.1 in Bobkov et al. (2010)), we shall compare their Stieltjes transforms.

Our first result compares the Stieltjes transform of a matrix satisfying martingale-like projective conditions with the Stieltjes transform of an independent matrix whose entries are observations of a Gaussian random field with the same covariance structure. We shall assume that $\mathbb{X}_n$ is defined by (2), and satisfies the Lindeberg's condition below:

Condition 1. *(i) The variables $(X_{ij})_{i,j}$ are centered at expectations.*
(2) There exists a positive constant C such that, for any positive integer n,

$$\frac{1}{n^2} \sum_{n \geq i \geq j \geq 1} \mathbb{E}(X_{ij}^2) < C.$$

(iii) For every $\varepsilon > 0$,

$$L_n(\varepsilon) = \frac{1}{n^2} \sum_{n \geq i \geq j \geq 1} \mathbb{E}(X_{ij}^2 I(|X_{ij}| > \varepsilon n^{1/2})) \to 0.$$

Clearly the items (ii) and (iii) of this condition are satisfied as soon as the family (X_{ij}^2) is uniformly integrable or the random field is stationary and in $\mathbb{L}^2$ (recall that a random field $(X_{\mathbf{u}})_{\mathbf{u} \in \mathbb{Z}^2}$ is said to be (strictly) stationary if the law of $(X_{\mathbf{u}+\mathbf{v}})_{\mathbf{u} \in \mathbb{Z}^2}$ does not depend on $\mathbf{v} \in \mathbb{Z}^2$).

To introduce our martingale-like projective conditions (6) and (7) below as well as our regularity-type condition (8), we need to introduce the filtrations we shall consider:

For any non-negative integer a, let us introduce the following filtrations:

$$\begin{aligned}
\mathcal{F}^a_{i,\infty} &= \sigma(X_{uv} : 1 \leq u \leq i - a \,,\, v \geq 1) \text{ if } i > a \text{ and } \mathcal{F}^a_{i,\infty} = \{\Omega, \emptyset\} \text{ otherwise} \\
\mathcal{F}^a_{\infty,j} &= \sigma(X_{uv} : u \geq 1 \,,\, 1 \leq v \leq j - a) \text{ if } j > a \text{ and } \mathcal{F}^a_{\infty,j} = \{\Omega, \emptyset\} \text{ otherwise} \\
\mathcal{F}^a_{ij} &= \mathcal{F}^a_{i,\infty} \cup \mathcal{F}^a_{\infty,j} \,.
\end{aligned} \tag{5}$$

Note that X_{ij} is adapted to $\mathcal{F}^0_{ij}$. We are now in position to state our first result.

Theorem 2. *Assume that $\mathbb{X}_n$ satisfies Condition 1 and, as $n \to \infty$,*

$$\sup_{i \geq j} \|\mathbb{E}(X_{ij} | \mathcal{F}^n_{ij})\|_2 \to 0 \tag{6}$$

and

$$n^2 \sup \|\mathbb{E}(X_{ij} X_{ab} | \mathcal{F}^n_{i \wedge a, j \wedge b}) - \mathbb{E}(X_{ij} X_{ab})\|_1 \to 0, \tag{7}$$

where the supremum is taken over all pairs $(i,j) \neq (a,b)$ with $i \geq j$ and $a \geq b$. In addition assume that

$$\sup_{i \geq j} \|\mathbb{E}(X_{ij}^2 | \mathcal{F}^n_{ij}) - \mathbb{E}(X_{ij}^2)\|_1 \to 0 \text{ as } n \to \infty. \tag{8}$$

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

$$S^{\mathbb{X}_n}(z) - S^{\mathbb{Y}_n}(z) \to 0 \text{ in probability as } n \to \infty, \tag{9}$$

where $\mathbf{Y}_n$ is a Gaussian matrix of centered random variables with the same covariance structure as $\mathbf{X}_n$ and independent of $\mathbf{X}_n$ and $\mathbb{Y}_n = \mathbf{Y}_n / \sqrt{n}$.

Comment 3. (i) Conditions (6) and (7) can be viewed as a generalization of the martingale condition given in Basu and Dorea (1979) which is $\mathbb{E}(X_{ij} | \mathcal{F}^1_{i,j}) = 0$ a.s. for any $i \geq j \geq 1$. Both conditions (6) and (7) are obviously satisfied for this type of martingale random field, and then, the conditions of Theorem 2 are reduced just to Condition 1 and (8). Results for other type of martingale random fields based on the lexicographic order can be found in Merlevède et al. (2015).

(ii) Note also that Condition (8) is a regularity condition. For instance, in case where $\mathcal{F}^\infty_{ij} = \bigcap_{n \geq 0} \mathcal{F}^n_{ij}$ is the trivial σ-field, then this condition is automatically satisfied. Let us also mention that the conditions (6), (7) and (8) are natural extensions of projective criteria used for obtaining various limit theorems for sequences of random variables. As in the case of random sequences, the conditions (6), (7) and (8) can be handled either with the help of "physical measure of dependence" as developed in El Machkouri et al. (2013) for functions of i.i.d. random fields or by using mixing coefficients (see Section 3.1).

(iii) We should also mention that we can allow for dependence of n of the variables in $\mathbb{X}_n$.

The conditions in the theorem below have to be then generalized in a natural way. For instance, conditions (6) and (7) should become

$$\lim_{m\to\infty}\sup_{n\geq 1}\sup_{i\geq j}\|\mathbb{E}(X_{ij,n}|\mathcal{F}^m_{ij,n})\|_2=0$$

and

$$\lim_{m\to\infty} m^2\sup_{n\geq 1}\sup_{(i,j)\neq(a,b)}\|\mathbb{E}(X_{ij,n}X_{ab,n}|\mathcal{F}^m_{i\wedge a,j\wedge b,n})-\mathbb{E}(X_{ij,n}X_{ab,n})\|_1=0\,.$$

Based on the above theorem we shall treat two special cases of symmetric random matrices, namely $\mathbf{X}+\mathbf{X}^T$ and the covariance matrix given in definition (17).

We consider first the symmetric $n\times n$ matrix $\mathbf{Z}_n=[Z_{ij}]_{i,j=1}^n$ with $Z_{ij}=X_{ij}+X_{ji}$ and we set

$$\mathbb{Z}_n:=\frac{1}{\sqrt{2n}}\mathbf{Z}_n\,. \tag{10}$$

This type of symmetrization is important since it leads to a symmetric covariance structure. Indeed, if $(X_{ij})_{(i,j)\in\mathbb{Z}^2}$ is $\mathbb{L}^2$-stationary meaning that, for any $(i,j)\in\mathbb{Z}^2$, $\mathbb{E}(Y_{ij})=m$ and

$$\operatorname{cov}(X_{u,v},X_{k+u,\ell+v})=\operatorname{cov}(X_{0,0},X_{k,\ell})=c_{k,\ell}\,,$$

for any integers u,v,k,ℓ, we get that $(Z_{ij})_{(i,j)\in\mathbb{Z}^2}$ is also a $\mathbb{L}^2$-stationary random field satisfying

$$\operatorname{cov}(Z_{i,j},Z_{k,\ell})=b(k-i,\ell-j)+b(k-j,\ell-i)\text{ with }b(u,v)=\gamma_{u,v}+\gamma_{v,u}\,.$$

Notice then that $b(u,v)=b(v,u)$. This symmetry condition on the covariances is used for instance in Khorunzhy and Pastur (1994, Theorem 2) to derive the limiting spectral distribution of symmetric matrices associated with a stationary Gaussian random field when its associated series of covariances is absolutely summable.

Our next Theorem 4 shows that a similar conclusion as in Theorem 2 holds for $\mathbb{Z}_n$ defined above. However, due to the structure of each of the entries, the sequence (X_{ij}) has to satisfy the conditions of Theorem 2 but with the conditional expectations taken with respect to a larger filtration. Roughly speaking the filtrations in Theorem 2 are the union of two half planes, whereas in Theorem 4 they are defined as the sigma-algebras generated by all the variables outside the union of two squares. More precisely these latter filtrations are defined as follows: for any non-negative integer a,

$$\widetilde{\mathcal{F}}^a_{ij}=\sigma\big(X_{uv}\ :\ (u,v)\in\mathbb{Z}^2\text{ such that }\max(|i-u|,|j-v|)\geq a\big)\,. \tag{11}$$

Note that X_{ij} is adapted to $\widetilde{\mathcal{F}}^0_{ij}$.

Theorem 4. *Assume that $\mathbb{Z}_n$ is defined by (10) where the variables X_{ij} satisfy condition 1. In addition assume that*

$$\sup_{i\geq j}\|\mathbb{E}(X_{ij}|\widetilde{\mathcal{F}}^n_{ij})\|_2\to 0\ \textit{as}\ n\to\infty, \tag{12}$$

$$n^2\sup\|\mathbb{E}(X_{ij}X_{ab}|\widetilde{\mathcal{F}}^n_{ij}\cap\widetilde{\mathcal{F}}^n_{ab})-\mathbb{E}(X_{ij}X_{ab})\|_1\to 0\ \textit{as}\ n\to\infty, \tag{13}$$

where the supremum is taken over all pairs $(i,j)\neq(a,b)$. In addition assume that

$$\sup_{(i,j)}\|\mathbb{E}(X^2_{ij}|\widetilde{\mathcal{F}}^n_{ij})-\mathbb{E}(X^2_{ij})\|_1\to 0\ \textit{as}\ n\to\infty. \tag{14}$$

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

$$S^{\mathbb{Z}_n}(z)-S^{\mathbb{W}_n}(z)\to 0\ \textit{in probability as}\ n\to\infty, \tag{15}$$

where $\mathbf{W}_n=[W_{ij}]_{i,j=1}^n$ with $W_{ij}=Y_{ij}+Y_{ji}$, (Y_{ij}) being a real-valued Gaussian centered random field with the same covariance structure as $\mathbf{X}_n$ and independent of $\mathbf{X}_n$, and $\mathbb{W}_n=\mathbf{W}_n/\sqrt{2n}$.

Let $(X_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$ be a random field of real-valued square integrable variables and $(Y_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$ be a real-valued Gaussian random field with the same covariances, and independent of $(X_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$. Let N and p be two positive integers and consider the $N\times p$ matrices

$$\mathcal{X}_{N,p} = \big(X_{ij}\big)_{1\le i\le N, 1\le j\le p}\,,\ \Gamma_{N,p} = \big(Y_{ij}\big)_{1\le i\le N, 1\le j\le p}\,. \tag{16}$$

Define now the symmetric matrices $\mathbb{B}_N$ and $\mathbb{G}_N$ of order N by

$$\mathbb{B}_N = \frac{1}{N}\mathcal{X}_{N,p}\mathcal{X}_{N,p}^T\,,\ \mathbb{G}_N = \frac{1}{N}\Gamma_{N,p}\Gamma_{N,p}^T. \tag{17}$$

The matrix $\mathbb{B}_N$ is usually referred to as the sample covariance matrix associated with the process $(X_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$. It is also known under the name of Gram random matrix. In particular, if the random field $(X_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$ is an array of i.i.d. random variables with zero mean and variance σ^2, then the famous Marchenko and Pastur (1967) theorem states that, if $p/N\to c\in(0,\infty)$, then, for all $z\in\mathbb{C}^+$, $S^{\mathbb{B}_N}(z)$ converges almost surely to $S(z)=S$, which is the unique solution with $\mathrm{Im}S(z)\ge 0$ of the quadratic equation: for any $z\in\mathbb{C}^+$,

$$z\sigma^2S^2 + (z - c\sigma^2 + \sigma^2)S + 1 = 0\,. \tag{18}$$

This means that $\mathbb{P}(d(F^{\mathbb{G}_N}, F_c)\to 0) = 1$, where F_c is a probability distribution function of the so-called Marchenko-Pastur distribution with parameter $c>0$. That is F_c has density

$$g_c(x) = \frac{1}{2\pi x\sigma^2}\sqrt{(x-a)(b-x)}I(a\le x\le b)$$

and a point mass $1-c$ at the origin if $c<1$, where $a=\sigma^2(1-\sqrt{c})^2$ and $b=\sigma^2(1+\sqrt{c})^2$. Note that this result still holds if the random field $(X_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$ is an array of independent centered random variables with common positive variance σ^2, which satisfies the Lindeberg's condition 1 (see Pastur (1972)). Moreover, in this situation, the Lindeberg's condition is necessary as shown in Girko (1995, Theorem 4.1, Chapter 3) (see also Corollary 2.3 in Yaskov (2016-b)).

When we relax the independence assumption of the entries, the following result holds.

Theorem 5. *Assume $(X_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$ is as in Theorem 2. Then, if $p/N\to c\in(0,\infty)$, for all $z\in\mathbb{C}^+$,*

$$S^{\mathbb{B}_N}(z) - S^{\mathbb{G}_N}(z)\to 0 \text{ in probability, as } N\to\infty.$$

All our results can be easily reformulated for random matrices with entries from a stationary random field. For some applications it is interesting to formulate sufficient conditions in terms of the conditional expectation of a single random variable. For this case it is natural to work with the extended filtrations.

Let now $(X_{ij})_{(i,j)\in\mathbb{Z}^2}$ be a stationary real-valued random field. For any non-negative integer a, let us introduce the following filtrations:

$$\mathcal{F}^a_{i,\infty} = \sigma(X_{uv} : u\le i-a,\ v\in\mathbb{Z});$$
$$\mathcal{F}^a_{\infty,j} = \sigma(X_{uv} : v\le j-a,\ u\in\mathbb{Z});\ \mathcal{F}^a_{ij} = \mathcal{F}^a_{i,\infty}\cup\mathcal{F}^a_{\infty,j}\,.$$

We call the random field regular if for any $\mathbf{u}\in\mathbb{Z}^2$, $\mathbb{E}(X_{\mathbf{0}}X_{\mathbf{u}}|\mathcal{F}^{\infty}_{\mathbf{u}\wedge\mathbf{0}}) = \mathbb{E}(X_{\mathbf{0}}X_{\mathbf{u}})$ a.s.

Theorem 6. *Assume that $\mathbb{X}_n$ is defined by (2) where (X_{ij}) is a stationary, centered and regular random field. Assume the couple of conditions*

$$\sum_{\mathbf{u}\in V_0}\mathbb{E}|X_{\mathbf{u}}\mathbb{E}(X_{\mathbf{0}}|\mathcal{F}^{|u|}_{\mathbf{0}})| < \infty \tag{19}$$

and

$$p^2\sup_{\mathbf{u}\in V_0:|\mathbf{u}|>p}\mathbb{E}|X_{\mathbf{u}}\mathbb{E}(X_{\mathbf{0}}|\mathcal{F}^p_{\mathbf{0}})|\to 0,\ \text{as } p\to\infty,$$

where $V_0=\{\mathbf{u}=(u_1,u_2)\in\mathbb{Z}^2 : u_1\le 0 \text{ or } u_2\le 0\}$. Then the conclusions of Theorems 2 and 5 hold.

Condition (19) implies that $\sum_{\mathbf{u}\in\mathbb{Z}^2} |\mathrm{cov}(X_{\mathbf{0}}, X_{\mathbf{u}})| < \infty$ and is in the spirit of condition (2.3) given in Dedecker (1998) to derive a central limit theorem for stationary random fields. As we shall see in Subsection 3.1, when applied to stationary strongly mixing random fields, the conditions of Theorem 6 require a rate of convergence of the strong mixing coefficients with only one point in the future whereas the conditions of Theorems 2 and 5 require a rate of convergence of the strong mixing coefficients with two points in the future.

Combining Theorem 6 with Theorem 11 concerning Gaussian covariance matrices, the following corollary holds:

Corollary 7. *Let $\mathbb{B}_N$ be defined by* (17)*. Under the assumptions of Theorem 6 and if $p/N \to c \in (0,\infty)$, $d(F^{\mathbb{B}_N}, F) \to 0$ in probability where F is a nonrandom distribution function whose Stieltjes transform $S(z)$, $z \in \mathbb{C}^+$, is uniquely defined by the spectral density of (X_{ij}) and satisfies the equation stated in Theorem 11.*

3 Examples

3.1 Strongly mixing random fields

Let us first recall the definition of the strong mixing coefficient of Rosenblatt (1956): For any two σ-algebras $\mathcal{A}$ and $\mathcal{B}$, the strong mixing coefficient $\alpha(\mathcal{A},\mathcal{B})$ is defined by:

$$\alpha(\mathcal{A},\mathcal{B}) = \sup\{|\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)|; A \in \mathcal{A} \text{ and } B \in \mathcal{B}\}.$$

An equivalent definition is:

$$2\alpha(\mathcal{A},\mathcal{B}) = \sup\{|\mathbb{E}(|\mathbb{P}(B|\mathcal{A}) - \mathbb{P}(B)|) : B \in \mathcal{B}\},$$

and, according to Bradley (2007), Theorem 4.4, item (a2), one also has

$$4\alpha(\mathcal{A},\mathcal{B}) = \sup\{\|\mathbb{E}(Y|\mathcal{A})\|_1 : \ Y \ \mathcal{B}\text{-measurable}, \ \|Y\|_\infty = 1 \text{ and } \mathbb{E}(Y) = 0\}. \tag{20}$$

For a random field $\mathbf{X} = (X_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$, let

$$\alpha_{1,\mathbf{X}}(n) = \sup_{i,j} \alpha(\mathcal{F}^n_{ij}, \sigma(X_{ij})) \text{ and } \alpha_{2,\mathbf{X}}(n) = \sup_{(i,j)\neq(a,b)} \alpha(\mathcal{F}^n_{i\wedge a, j\wedge b}, \sigma(X_{ij}, X_{ab})).$$

Note that $\alpha_{1,\mathbf{X}}(n) \le \alpha_{2,\mathbf{X}}(n)$. For a bounded centered random field, the mixing condition required by Theorem 2 (or by Theorem 5) is

$$n^2 \alpha_{2,\mathbf{X}}(n) \to 0,$$

while for Theorem 6, provided the random field is stationary, we need the couple of conditions:

$$\alpha_{2,\mathbf{X}}(n) \to 0 \text{ and } \sum_{n\ge1} n\alpha_{1,\mathbf{X}}(n) < \infty.$$

If for some $\delta > 0$ we have $\sup_{\mathbf{u}} \|X_{\mathbf{u}}\|_{2+\delta} < \infty$ and the random field is centered then, by the properties of the mixing coefficients, applying, for instance, Lemma 4 in Merlevède and Peligrad (2006) (see also Bradley (2007) and Annex C in Rio (2017)), we infer that the conclusions of Theorems 2 and 5 are implied by

$$n^2 (\alpha_{2,\mathbf{X}}(n))^{\delta/(2+\delta)} \to 0.$$

Moreover, if we assume stationarity of the random field, Theorem 6 requires the couple of conditions:

$$\alpha_{2,\mathbf{X}}(n) \to 0 \text{ and } \sum_{n\ge1} n^{1+4/\delta} \alpha_{1,\mathbf{X}}(n) < \infty.$$

Slightly more general results can be given in terms of the quantile function of $|X_{\mathbf{u}}|$ (see Rio (2017) or the computations in the proof of Theorem 6.40 in Merlevède et al. (2019) where similar projective quantities as those involved in Theorems 2 or 6 are handled).

We refer to the monograph by Doukhan (1994) for examples of strong mixing random fields. Let us also mention the paper by Dombry and Eyi-Minko (2012) where, for max-infinitely divisible random fields on $\mathbb{Z}^d$, upper bounds of the strong mixing coefficients are given with the help of the extremal coefficient function (examples such as the Brown-Resnick process and the moving maxima process are considered). Strong mixing coefficients can also be controlled in the case of bounded spin systems. For instance, in case where the family of Gibbs specifications satisfies the weak mixing condition introduced by Dobrushin and Shlosman (1985), the coefficient $\alpha_2(n)$ decreases exponentially fast. This is then the case for Ising models with external fields in the regions where the temperature is strictly larger than the critical one (we refer to Dedecker (2001, Section 2.3) and to Laroche (1995) for more details).

Below, is another example of a random field for which the strong mixing coefficients can be handled.

Example: Functions of two independent strong mixing random fields. Let us consider two real-valued independent processes $\mathbf{U} = (U_{ij}, i, j \in \mathbb{Z})$ and $\mathbf{V} = (V_{ij}, i, j \in \mathbb{Z})$ such that, setting $\mathbf{U}^{(j)} = (U_{ij}, i \in \mathbb{Z})$, the processes $\mathbf{U}^{(j)}$, $j \in \mathbb{Z}$, are mutually independent and have the same law as $(U_i, i \in \mathbb{Z})$ and, setting $\mathbf{V}_{(i)} = (V_{ij}, j \in \mathbb{Z})$, the processes $\mathbf{V}_{(i)}$, $i \in \mathbb{Z}$ are also mutually independent and have the same law as $(V_j, j \in \mathbb{Z})$. For any measurable function h from $\mathbb{R}^2$ to $\mathbb{R}$, let

$$X_{ij} = h(U_{ij}, V_{ij}) - \mathbb{E}(h(U_{ij}, V_{ij})), \tag{21}$$

provided the expectation exists. Note that the random field $\mathbf{X} = (X_{ij}, i, j \in \mathbb{Z})$ does not have independent entries across the rows nor the columns (except if we have that for any j fixed, the r.v.'s U_{ij}, $i \in \mathbb{Z}$ are mutually independent as well as the r.v.'s V_{ij}, $j \in \mathbb{Z}$, for any i fixed). Hence, the results in Merlevède and Peligrad (2016) do not apply. Let $\mathcal{F}_k^{\mathbf{U}} = \sigma(U_\ell, \ell \le k)$ and $\mathcal{F}_k^{\mathbf{V}} = \sigma(V_\ell, \ell \le k)$, and define

$$\alpha_{1,\mathbf{U}}(n) = \sup_i \alpha(\mathcal{F}_{i-n}^{\mathbf{U}}, \sigma(U_i)) \ , \ \alpha_{2,\mathbf{U}}(n) = \sup_{i,j\,:\,j>i} \alpha(\mathcal{F}_{i-n}^{\mathbf{U}}, \sigma(U_i, U_j))$$

and

$$\alpha_{1,\mathbf{V}}(n) = \sup_i \alpha(\mathcal{F}_{i-n}^{\mathbf{V}}, \sigma(V_i)) \ , \ \alpha_{2,\mathbf{V}}(n) = \sup_{i,j\,:\,j>i} \alpha(\mathcal{F}_{i-n}^{\mathbf{V}}, \sigma(V_i, V_j)) \, .$$

Due to the definition of the strong mixing coefficients, it follows that

$$\alpha_{1,\mathbf{X}}(n) \le \alpha_{1,\mathbf{U}}(n) + \alpha_{1,\mathbf{V}}(n) \text{ and } \alpha_{2,\mathbf{X}}(n) \le \alpha_{2,\mathbf{U}}(n) + \alpha_{2,\mathbf{V}}(n) \, .$$

(See for instance Theorem 6.2 in Bradley (2007)). So, if we assume for instance that the function h is bounded and that $n^2(\alpha_{2,\mathbf{U}}(n) + \alpha_{2,\mathbf{V}}(n)) \to 0$, then Theorem 2 applies. Moreover if we assume in addition that the sequences $(U_{ij}, i \in \mathbb{Z})$ and $(V_{ij}, j \in \mathbb{Z})$ are stationary and that

$$\sum_{n \ge 1} n \big(\alpha_{2,\mathbf{U}}(n) + \alpha_{2,\mathbf{V}}(n)\big) < \infty \, ,$$

then, according to Corollary 7, we derive that, if $p/N \to c \in (0, \infty)$, for all $z \in \mathbb{C}^+$,

$$S^{\mathbb{B}_N}(z) \to S(z) \text{ in probability as } N \to \infty \, ,$$

where $\mathbb{B}_N$ is the Gram random matrix defined by (17) and S is defined in Theorem 11.

3.2 A convolution example

Let $\mathbf{U} = (U_{ij}, i, j \in \mathbb{Z})$ be a stationary centered regular martingale difference random field in $\mathbb{L}^2$, meaning that $\sup_{i,j} \|U_{ij}\|_2 < \infty$ and that, setting $\mathcal{G}_{ij}^a = \sigma(V_{k\ell}, k \leq i - a \text{ or } \ell \leq j - a)$,

$$\mathbb{E}(U_{ij}|\mathcal{G}_{ij}^1) = 0 \text{ a.s. and } \|\mathbb{E}(U_{\mathbf{0}}^2|\mathcal{G}_{\mathbf{0}}^n) - \mathbb{E}(U_{\mathbf{0}}^2)\|_1 \to 0 \text{ as } n \to \infty \,.$$

Let $\varepsilon = (\varepsilon_{ij}, i, j \in \mathbb{Z})$ be an iid centered random field in $\mathbb{L}^\infty$, independent of $\mathbf{U}$ and $(a_{k\ell}, k, \ell \in \mathbb{N})$ be a double indexed sequence of real numbers such that $\sum_{k,\ell \in \mathbb{N}} (k^2 + \ell^2)|a_{k,\ell}| < \infty$. Set $V_{ij} = \sum_{k,\ell \in \mathbb{Z}} a_{k,\ell} \varepsilon_{i-k,j-\ell}$ and define the stationary centered random field $\mathbf{X} = (X_{\mathbf{u}})_{\mathbf{u} \in \mathbb{Z}^2}$ in $\mathbb{L}^2$ by setting $X_{ij} = U_{ij} + V_{ij}$. It is easy to see that $\mathbf{X}$ satisfies the conditions of Theorem 6.

4 LSD for stationary Gaussian random fields

In this section we survey several old and new results for stationary Gaussian random fields that could be combined with our universality results in order to decide that the LSD exists and to characterize it. Relevant to this part is the notion of spectral density. We consider a centered stationary Gaussian random field $(Y_{ij})_{(i,j) \in \mathbb{Z}^2}$, meaning that for any $(i, j) \in \mathbb{Z}^2$, $\mathbb{E}(Y_{ij}) = 0$ and

$$\operatorname{cov}(Y_{u,v}, Y_{k+u,\ell+v}) = \operatorname{cov}(Y_{0,0}, Y_{k,\ell}) = \gamma_{k,\ell} \,,$$

for any integers u, v, k, ℓ. According to the Bochner-Herglotz representation (see for instance Theorem 1.7.4 in Sasvári (2013)), since the covariance function is positive definite, there exists a unique spectral measure such that

$$\operatorname{cov}(Y_{0,0}, Y_{k,\ell}) = \int_{[0,1]^2} \mathrm{e}^{2\pi \mathrm{i}(ku+\ell v)} F(\mathrm{d}u, \mathrm{d}v), \quad \text{for all } k, \ell \in \mathbb{Z} \,.$$

If F is absolutely continuous with respect to the Lebesgue measure $\lambda \otimes \lambda$, we have

$$\gamma_{k,\ell} := \operatorname{cov}(Y_{0,0}, Y_{k,\ell}) = \int_{[0,1]^2} \mathrm{e}^{2\pi \mathrm{i}(ku+\ell v)} f(u, v) \mathrm{d}u \mathrm{d}v, \quad \text{for all } k, \ell \in \mathbb{Z} \,. \tag{22}$$

Khorunzhy and Pastur (1994) and Boutet de Monvel and Khorunzhy (1999) treated a class of Gaussian fields with absolutely summable covariances,

$$\sum_{k,\ell \in \mathbb{Z}} |\gamma_{k,\ell}| < \infty \,, \tag{23}$$

and a certain symmetry condition. They described the limiting spectral distribution via an equation satisfied by the Stieltjes transform of the limiting distribution. Since the covariance structure is determined by the spectral density, this limiting spectral distribution can be expressed in terms of spectral density which generates the covariance structure. More precisely, if we consider the $n \times n$ random matrix $\mathbf{Y}_n$ with entries $Y_{k,\ell}$ and the symmetric matrix

$$\mathbb{W}_n = \frac{1}{\sqrt{2n}} (\mathbf{Y}_n + \mathbf{Y}_n^T) \,, \tag{24}$$

Theorem 2 in Khorunzhy and Pastur (1994) (see also in Theorem 17.2.1. in Pastur and Shcherbina (2011)) gives the following:

Theorem 8. *Let $(Y_{k,\ell})_{(k,\ell) \in \mathbb{Z}^2}$ be a centered stationary Gaussian random field with spectral density $f(x, y)$. Denote $b(x, y) = 2^{-1}(f(x, y) + f(y, x))$. Assume that* (23) *holds. Let $\mathbb{W}_n$ be defined by* (24)*. Then $\mathbb{P}(d(F^{\mathbb{W}_n}, F) \to 0) = 1$, where F is a nonrandom distribution function whose Stieltjes transform $S(z)$ is uniquely determined by the relations:*

$$S(z) = \int_0^1 g(x, z) dx \,, \; z \in \mathbb{C}^+ \,, \tag{25}$$

$$g(x,z) = -\Big(z + \int_0^1 b(x,y)g(y,z)dy\Big)^{-1}, \tag{26}$$

where for any $z \in \mathbb{C}^+$ and any $x \in [0,1)$, $g(x,z)$ is analytic in z and

$$\operatorname{Im} g(x,z) \cdot \operatorname{Im} z > 0\,,\ |g(x,z)| \le (\operatorname{Im} z)^{-1}\,,$$

and is periodic and continuous in x.

For the symmetric matrix $\mathbb{W}_n$ defined by (24) and constructed from a stationary Gaussian random field, Chakrabarty, Hazram and Sarkar (2016) proved the existence of its limiting spectral distribution provided that the spectral density of the Gaussian process exists. Their result goes then beyond the condition (23) requiring that the covariances are absolutely summable. It was completed recently by C. Peligrad and M. Peligrad (2016) who obtained a characterization of the limiting empirical spectral distribution for symmetric matrices with entries selected from a stationary Gaussian field under the sole condition that its spectral density exists. Their Theorem 2 is the following:

Theorem 9. *Let $(Y_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ be a centered stationary Gaussian random field with spectral density $f(x,y)$. Let $\mathbb{W}_n$ be defined by (24). Then, $\mathbb{P}(d(F^{\mathbb{W}_n}, F) \to 0) = 1$, where the Stieltjes transform $S(z)$ of F is uniquely defined by the relation (25) where for almost all x, $g(x,z)$ is a solution of the equation (26).*

If the spectral density has the structure $f(x,y) = u(x)u(y)$, the equation (25) simplifies as

$$S(z) = -\frac{1}{z}(1 + v^2(z))\,,\ z \in \mathbb{C}^+\,,$$

where $v(z)$ is solution of the equation

$$v(z) = -\int_0^1 \frac{u(y)\mathrm{d}y}{z + u(y)v(z)}\,,\ z \in \mathbb{C}^+\,,$$

with $v(z)$ analytic, $\operatorname{Im} v(z) > 0$ and $|v(z)| \le (\operatorname{Im} z)^{-1}\|Y_{0,0}\|_2$ (see the proof of Remark 3 in Peligrad-Peligrad (2016)).

In particular, if the random field is an array of i.i.d. random variables with zero mean and variance σ^2, then $u(x)$ is constant and $S(z)$ satisfies the equation (4).

The following new result gives the existence of LSD for large covariance matrices associated with a stationary Gaussian random field. Its proof is based on the method of proof in Chakrabarty *et al.* (2016).

Proposition 10. *Let $(Y_{ij})_{(i,j)\in\mathbb{Z}^2}$ be a stationary real-valued Gaussian process with mean zero. Assume that this process has a spectral density on $[0,1]^2$ denoted by f. Let N and p be two positive integers and consider $\Gamma_{N,p}$ the $N \times p$ matrix defined by $\Gamma_{N,p} = \big(Y_{ij}\big)_{1\le i\le N, 1\le j\le p}$. Let also $\mathbb{G}_N = \frac{1}{N}\Gamma_{N,p}\Gamma_{N,p}^T$. Then, when $p/N \to c \in (0,\infty)$, there exists a deterministic probability measure μ_f determined solely by c and the spectral density f, and such that the spectral empirical measure $\mu_{\mathbb{G}_N}$ converges weakly in probability to μ_f.*

For the case where the covariances are absolutely summable, we cite the following result which is Theorem 2.1 in Boutet de Monvel *et al.* (1996). It allows to characterize the LSD μ_f of $\mathbb{G}_N$ via an equation satisfied by its Stieltjes transform.

Theorem 11. *Assume that the assumptions of Proposition 10 and that condition (23) hold. Then, when $p/N \to c \in (0,\infty)$, $\mathbb{P}(d(F^{\mathbb{G}_N}, F) \to 0) = 1$ 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 of the equation

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

with $f(x,y)$ the spectral density given in (22).

When we assume that the entries of $\Gamma_{N,p}$ is a sequence of i.i.d. random variables with mean zero and variance σ^2, then $S(z)$ satisfies the equation (18) of the Marchenko-Pastur distribution. In view of Proposition 10 and of Theorem 11, it is still an open question if, without imposing the summability condition (23) on the covariances, one could still characterize the LSD of $\mathbb{G}_N$.

5 Proofs

The notation $V_n^1 = \{(i,j); i \geq j \text{ with } i \text{ and } j \text{ in } \{1,\dots,n\}\}$ will be often used along the proofs.

5.1 Proof of Theorem 2

The proof is based on a Bernstein-type blocking procedure for random fields and the Lindeberg's method. The blocking argument, originally introduced by Bernstein (1927) in order to prove an extension of the central limit theorem to r.v.'s satisfying dependent conditions, consists of making "big blocks" interlaced by "small blocks" which have a negligible behavior compared to the one of the "big blocks". In the context of random fields, this blocking argument can also be used (see for instance Tone (2011), where the asymptotic normality of the normalized partial sum of a Hilbert-space valued stationary and mixing random field is proved with the help of a blocking procedure). In our context, the "big" blocks, called $\mathbf{B}_{i,j}$ in the figure (28) below, are of size p (with p such that $p/n \to 0$) and the "small" blocks will consist of bands of width K with entries which are zero and with K negligible with respect to p. As we shall see below, this blocking procedure can be efficiently done because, roughly speaking, the limiting spectral density distribution is not affected by changing a number of $o(n^2)$ variables.

Now the Lindeberg's method will consist of replacing one by one each of the "big" blocks with blocks of the same size but whose entries are those of a Gaussian random field having the same covariance structure as the initial process.

The blocking procedure combined with the Lindeberg's method does not seem very classical in the context of random matrices. It has been however recently used in Banna et al. (2015) and in Merlevède and Peligrad (2016), but in the context where the entries of the matrices are functions of an i.i.d. random field in the first mentioned paper, or in the context where the rows or the columns of the matrix are independent, in the second one. These conditions are not assumed in the context of the present paper. This makes the situation more delicate. Indeed, concentration inequalities for the Stieltjes transform around its mean are not available, hence we cannot restrict the study to the difference between the expectations of the two Stieltjes transforms. However, as we shall see, this issue can be bypassed by approximating the random matrix with "big" blocks $\mathbb{B}(\mathbf{X}_n)$ defined in (28), by another one where the "big" blocks will have a certain martingale difference property. Hence, in particular, they are uncorrelated. This new uncorrelated block matrix will be called $\mathbb{B}(\mathbf{X}'_n)$ in the proof below. A similar treatment will be done to the matrices with the Gaussian field entries, having a suitable covariance structure.

We turn now to the details of the proof of Theorem 2, and first, to our blocking procedure, which involves several steps. We then start by some preliminary considerations.

Let (K), (c_K) and (p_K) be sequences of integers converging to ∞ such that $p_K = c_K K$. Assume that

$$c_K^2 K^2 \sup_{(i,j)\neq(a,b); i\geq j, a\geq b} \|\mathbb{E}(X_{ij}X_{ab}|\mathcal{F}^K_{i\wedge a, j\wedge b}) - \mathbb{E}(X_{ij}X_{ab})\|_1 \to 0 \text{ as } K\to\infty. \tag{27}$$

This selection of (c_K) is possible by (7). In what follows, to soothe the notation, we suppress the subscript K to denote the sequences (c_K) and (p_K).

We describe now the blocking procedure which is based on the known fact that the limiting spectral density distribution is not affected by changing a number of $o(n^2)$ variables. We first notice that, without restricting the generality we may and do assume that $n = q(p+K)+p$ where $q := q_n$ is a sequence of positive integers depending on n, K and p. Indeed, if $(n-p)/(p+K)$ is not an integer, we set $n' = q(p+K)+p$ where $q = [(n-p)/(p+K)]$ and we notice that, by the Cauchy's interlacing law (see for instance Relation (2.96) in Tao's monograph (2012) for more details),

$$|S^{\mathbb{X}_n} - S^{\mathbb{X}_{n'}}| \ll \frac{n-n'}{n} \ll \frac{p+K}{n} \to 0 \text{ as } n \to \infty.$$

Therefore, we shall assume, from now on, that $n = q(p+K)+p$.

To introduce the big blocks, for a given symmetric matrix $\mathbf{Z}_n = \{Z_{ij}\}_{i,j=1}^n$ we shall associate the following checkerboard structure

$$\mathbb{B}(\mathbf{Z}_n) = \frac{1}{\sqrt{n}} \begin{pmatrix} 0_{p,p} & & & & & & & & \\ 0_{K,p} & 0_{K,K} & & & & & & & \\ \mathbf{B}_{1,1} & 0_{p,K} & 0_{p,p} & & & & & & \\ 0_{K,p} & 0_{K,K} & 0_{K,p} & 0_{K,K} & & & & & \\ \mathbf{B}_{2,1} & 0_{p,K} & \mathbf{B}_{2,2} & 0_{p,K} & 0_{p,p} & & & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \\ \mathbf{B}_{q-1,1} & 0_{p,K} & \mathbf{B}_{q-1,2} & 0_{p,K} & \mathbf{B}_{q-1,3} & \dots & 0_{p,p} & & \\ 0_{K,p} & 0_{K,K} & 0_{K,p} & 0_{K,K} & 0_{K,p} & \dots & 0_{K,p} & 0_{K,K} & \\ \mathbf{B}_{q,1} & 0_{p,K} & \mathbf{B}_{q,2} & 0_{p,K} & \mathbf{B}_{q,3} & \dots & \mathbf{B}_{q.q} & 0_{p,K} & 0_{p,p} \end{pmatrix}. \tag{28}$$

The rest of the matrix is completed by symmetry. Here each $\mathbf{B}_{k,\ell}$ denotes a block matrix $p \times p$ indexed by a set of indexes in $\mathcal{E}_{k,\ell}$ defined below, and whose entries are identical to the matrix $\mathbf{Z}_n$. To be more precise, we define

$$\mathcal{E}_{k,\ell} = \{(u,v) \in E_k \times E_{\ell-1}\} \text{ where } E_\ell = \big[\ell(p+K)+1\,,\, \ell(p+K)+p\big] \cap \mathbb{N}\,. \tag{29}$$

We shall order the blocks in the lexicographic order starting with the top of the matrix. To soothe the notations all along the paper, we shall use the following convention: for any $k = 1, \dots, q_n$ and any $\ell = 1, \dots, k$,

$$\mathbf{B}_u = \mathbf{B}_{k,\ell} \text{ and } \mathbf{I}_u = \mathcal{E}_{k,\ell} \text{ where } u = k(k-1)/2 + \ell. \tag{30}$$

To avoid confusion, when block matrices are constructed from different symmetric matrices, we shall also use the notation $\mathbf{B}_i = \mathbf{B}_i(\mathbf{Z}_n)$ to identify the variables in the block matrix. Note that the "big" blocks $\mathbf{B}_i$ are separated by bands of K rows and K columns. Variables in two different blocks are separated by at least either K rows or K columns.

Using Lemma 14, Theorem A.43 in Bai and Silverstein (2010) and taking into account Condition 1, straightforward computations lead to

$$\mathbb{E}|S^{\mathbb{X}_n} - S^{\mathbb{B}(\mathbf{X}_n)}|^2 \to 0 \text{ as } n \to \infty.$$

Similarly, we define $\mathbb{B}(\mathbf{Y}_n)$, and one can prove that

$$\mathbb{E}|S^{\mathbb{Y}_n} - S^{\mathbb{B}(\mathbf{Y}_n)}|^2 \to 0 \text{ as } n \to \infty.$$

We introduce a filtration

$$\mathcal{B}_u = \sigma(\mathbf{B}_1(\mathbf{X}_n), \mathbf{B}_2(\mathbf{X}_n), \dots, \mathbf{B}_u(\mathbf{X}_n)) \text{ for } u \geq 1 \text{ and } \mathcal{B}_0 = \{\emptyset, \Omega\}. \tag{31}$$

To introduce martingale structure, for $1 \le u \le q(q+1)/2$ and $j \in \mathbf{I}_u$ define the variables

$$X'_{\mathbf{j}} = X_{\mathbf{j}} - \mathbb{E}(X_{\mathbf{j}}|\mathcal{B}_{u-1}).$$

Then we define a new block matrix, say $\mathbb{B}(\mathbf{X}'_n)$, with blocks $\mathbf{B}'_i = \mathbf{B}_i(\mathbf{X}'_n)$ having a similar structure as $\mathbb{B}(\mathbf{X}_n)$ where the entries in these big blocks are $X'_{\mathbf{j}}$, $j \in \mathbf{I}_u$. Note that by Lemma 14

$$\mathbb{E}|S^{\mathbb{B}(\mathbf{X}_n)} - S^{\mathbb{B}(\mathbf{X}'_n)}|^2 \ll \frac{1}{n^2} \sum_{u \ge 1} \sum_{j \in \mathbf{I}_u} \mathbb{E}|\mathbb{E}(X_{\mathbf{j}}|\mathcal{B}_{u-1})|^2 \le \sup_{i \ge j} \mathbb{E}|\mathbb{E}(X_{ij}|\mathcal{F}_{ij}^K)|^2,$$

which converges to 0 uniformly in n when $K \to \infty$ by (6). Here and in the sequel we shall keep in mind that the range for the index u is from 1 to $q(q+1)/2$. For simplicity, we shall denote the sum from $u = 1$ to $u = q(q+1)/2$ by a sum over $u \ge 1$.

We proceed similarly for the matrix $\mathbb{B}(\mathbf{Y}_n)$. We introduce the filtration

$$\mathcal{H}_u = \sigma(\mathbf{B}_1(\mathbf{Y}_n), \mathbf{B}_2(\mathbf{Y}_n), \dots, \mathbf{B}_u(\mathbf{Y}_n)) \text{ for } u \ge 1 \text{ and } \mathcal{H}_0 = \{\emptyset, \Omega\}, \tag{32}$$

and for any $j \in \mathbf{I}_u$ define the variables

$$Y'_{\mathbf{j}} = Y_{\mathbf{j}} - \mathbb{E}(Y_{\mathbf{j}}|\mathcal{H}_{u-1}).$$

Notice that $(Y'_{\mathbf{j}}, 1 \le u \le q(q+1)/2, \mathbf{j} \in I_u)$ is also a Gaussian vector. In addition, by using the properties of conditional expectation we can easily notice that the random vectors $(Y'_{\mathbf{j}}, \mathbf{j} \in I_u)_u$ are orthogonal. Therefore $(Y'_{\mathbf{j}}, \mathbf{j} \in I_u)_u$ are mutually independent. We shall also prove that for $\mathbf{j} \in I_u$

$$\|\mathbb{E}(Y_{\mathbf{j}}|\mathcal{H}_{u-1})\|_2 \le \|\mathbb{E}(X_{\mathbf{j}}|\mathcal{B}_{u-1})\|_2. \tag{33}$$

To prove the inequality above, it suffices to notice the following facts. Let

$$\mathcal{V}_u = \overline{\text{span}}(1, (Y_{\mathbf{j}}\,,\, 1 \le v \le u\,,\, \mathbf{j} \in \mathbf{I}_v))$$

and

$$\mathcal{V}^*_u = \overline{\text{span}}(1, (X_{\mathbf{j}}\,,\, 1 \le v \le u\,,\, \mathbf{j} \in \mathbf{I}_v)),$$

where the closure is taken in $\mathbb{L}^2$. Denote by $\Pi_{\mathcal{V}_u}(\cdot)$ the orthogonal projection on $\mathcal{V}_u$ and by $\Pi_{\mathcal{V}^*_u}(\cdot)$ the orthogonal projection on $\mathcal{V}^*_u$. Since $(Y'_{\mathbf{j}}, 1 \le u \le q(q+1)/2, \mathbf{j} \in I_u)$ is a Gaussian process,

$$\mathbb{E}(Y_{\mathbf{j}}|\mathcal{H}_{u-1}) = \Pi_{\mathcal{V}_{u-1}}\big(Y_{\mathbf{j}}\big) \text{ a.s. and in } \mathbb{L}^2\,.$$

Since $(Y_{k\ell})_{1 \le \ell \le k \le n}$ has the same covariance structure as $(X_{k\ell})_{1 \le \ell \le k \le n}$, we observe that

$$\|\Pi_{\mathcal{V}_{u-1}}\big(Y_{\mathbf{j}}\big)\|_2 = \|\Pi_{\mathcal{V}^*_{u-1}}\big(X_{\mathbf{j}}\big)\|_2\,.$$

But,

$$\|\Pi_{\mathcal{V}^*_{u-1}}\big(X_{\mathbf{j}}\big)\|_2 \le \|\mathbb{E}(X_{\mathbf{j}}|\mathcal{B}_{u-1})\|_2\,,$$

which proves (33). Then we define a new block matrix, say $\mathbb{B}(\mathbf{Y}'_n)$, with blocks $\mathbf{\Gamma}'_i = \mathbf{B}_i(\mathbf{Y}'_n)$ having a similar structure as $\mathbb{B}(\mathbf{Y}_n)$ where the entries in these big blocks are $Y'_{\mathbf{j}}$. Therefore, by Lemma 14 and (33),

$$\begin{aligned}\mathbb{E}|S^{\mathbb{B}(\mathbf{Y}_n)} - S^{\mathbb{B}(\mathbf{Y}'_n)}|^2 &\ll \frac{1}{n^2} \sum_{u \ge 1} \sum_{j \in \mathbf{I}_u} \mathbb{E}|\mathbb{E}(Y_{\mathbf{j}}|\mathcal{H}_{u-1})|^2 \le \frac{1}{n^2} \sum_{u \ge 1} \sum_{\mathbf{j} \in \mathbf{I}_u} \mathbb{E}|\mathbb{E}(X_{\mathbf{j}}|\mathcal{B}_{u-1})|^2 \\ &\le \sup_{i \ge j} \mathbb{E}|\mathbb{E}(X_{ij}|\mathcal{F}_{ij}^K)|^2,\end{aligned}$$

which converges to 0 as $K \to \infty$ by (6), uniformly in n. The proof is reduced to showing that

$$\mathbb{E}|S^{\mathbb{B}(\mathbf{X}'_n)} - S^{\mathbb{B}(\mathbf{Y}'_n)}| \to 0 \text{ as } n \to \infty, \tag{34}$$

which we shall achieve at the end of several steps.

We write $S^{\mathbb{B}(\mathbf{X}'_n)}$ and $S^{\mathbb{B}(\mathbf{Y}'_n)}$ as function of the entries. So

$$S^{\mathbb{B}(\mathbf{X}'_n)} = s(\mathbf{B}'_1, \dots, \mathbf{B}'_{q(q+1)/2}) \text{ and } S^{\mathbb{B}(\mathbf{Y}'_n)} = s(\mathbf{\Gamma}'_1, \dots, \mathbf{\Gamma}'_{q(q+1)/2})\,. \tag{35}$$

We note that in our proofs the order in which we treat the blocks is critically important for using the power of the martingale structure, but the location of a variable in a block is not going to matter. With our functional notation we use the following decomposition:

$$S^{\mathbb{B}(\mathbf{X}'_n)} - S^{\mathbb{B}(\mathbf{Y}'_n)} = s(\mathbf{B}'_1, \dots, \mathbf{B}'_{q(q+1)/2}) - s(\mathbf{\Gamma}'_1, \dots, \mathbf{\Gamma}'_{q(q+1)/2}) \tag{36}$$
$$= \sum\nolimits_{u=1}^{q(q+1)/2} \Big(s(\mathbf{B}'_1, \dots, \mathbf{B}'_u, \mathbf{\Gamma}'_{u+1}, \dots, \mathbf{\Gamma}'_{q(q+1)/2}) - s(\mathbf{B}'_1, \dots, \mathbf{B}'_{u-1}, \mathbf{\Gamma}'_u, \dots, \mathbf{\Gamma}'_{q(q+1)/2}) \Big)\,.$$

We also denote

$$\mathbf{C}_u = (\mathbf{B}'_1, \dots, \mathbf{B}'_{u-1}, \mathbf{0}_u, \mathbf{\Gamma}'_{u+1}, \dots, \mathbf{\Gamma}'_{q(q+1)/2}),$$

where $\mathbf{0}_u$ is a null vector with p^2 entries 0. We shall use the Taylor expansion in Lemma 12, applied for a fixed index u, to the function

$$s(\mathbf{B}'_1, \dots, \mathbf{B}'_{u-1}, \mathbf{B}_u(\mathbf{X}'_n), \mathbf{\Gamma}'_{u+1}, \dots, \mathbf{\Gamma}'_{q(q+1)/2}),$$

where s is defined in (35). We can view this function simply as a function of a vector $\mathbf{x} = (x_{\mathbf{i}},\ u \in \{1, \dots, q(q+1)/2\}$ and $\mathbf{i} \in \mathbf{I}_u)$. By using (36), Lemma 12 with $A = 4\varepsilon n^{1/2}$ and (60), we obtain

$$S^{\mathbb{B}(\mathbf{X}'_n)} - S^{\mathbb{B}(\mathbf{Y}'_n)} = R'_1 + R'_2 + R'_3, \tag{37}$$

where

$$R'_1 = \sum_{u \ge 1} \sum_{\mathbf{j} \in \mathbf{I}_u} (X'_{\mathbf{j}} - Y'_{\mathbf{j}}) \partial_{\mathbf{j}} s(\mathbf{C}_u),$$

$$R'_2 = \frac{1}{2} \sum_{u \ge 1} \Big(\Big(\sum_{\mathbf{j} \in \mathbf{I}_u} X'_{\mathbf{j}} \partial_{\mathbf{j}} \Big)^2 - \Big(\sum_{\mathbf{j} \in \mathbf{I}_u} Y'_{\mathbf{j}} \partial_{\mathbf{j}} \Big)^2 \Big) s(\mathbf{C}_u)$$

and

$$|R'_3| \le \sum_{u \ge 1} |R_{u3}| + \sum_{u \ge 1} |R'_{u3}|,$$

with

$$|R_{u3}| \ll \frac{1}{n^2} p^2 \sum_{\mathbf{j} \in \mathbf{I}_u} (X'_{\mathbf{j}})^2 I(|X'_{\mathbf{j}}| > 4\varepsilon n^{1/2}) + \varepsilon n^{1/2} \frac{1}{n^{5/2}} p^4 \sum_{\mathbf{j} \in \mathbf{I}_u} (X'_{\mathbf{j}})^2$$

and

$$|R'_{u3}| \ll \frac{1}{n^2} p^2 \sum_{\mathbf{j} \in \mathbf{I}_u} (Y'_{\mathbf{j}})^2 I(|Y'_{\mathbf{j}}| > 4\varepsilon n^{1/2}) + \varepsilon n^{1/2} \frac{1}{n^{5/2}} p^4 \sum_{\mathbf{j} \in \mathbf{I}_u} (Y'_{\mathbf{j}})^2\,.$$

We treat first the term $|R'_3|$. By taking the expected value and considering condition 1, we obtain

$$\sum_{u \ge 1} \mathbb{E}|R_{u3}| \ll p^2 \frac{1}{n^2} \sum_{(i,j) \in V_n^1} \mathbb{E}(|X'_{ij}|^2 I(|X'_{ij}| > 4\varepsilon n^{1/2})) + \varepsilon p^4.$$

Notice now the following fact: If U is a real-valued random variable and $\mathcal{F}$ is a sigma-field, then setting $V = U - \mathbb{E}(U|\mathcal{F})$, the following inequality holds: for any $m \ge 1$ and any $a > 0$,

$$\mathbb{E}\big(|V|^m I(|V| > 4a)\big) \le 3 \times 2^m \mathbb{E}\big(|U|^m I(|U| > a)\big)\,. \tag{38}$$

This implies that

$$\mathbb{E}(|X'_{ij}|^2 I(|X'_{ij}| > 4\varepsilon n^{1/2})) \ll \mathbb{E}(X_{ij}^2 I(|X_{ij}| > \varepsilon n^{1/2})).$$

Therefore

$$\sum_{u\geq 1}\mathbb{E}|R_{u3}| \ll p^2\frac{1}{n^2}\sum_{(i,j)\in V_n^1}\mathbb{E}(X_{ij}^2 I(|X_{ij}|>\varepsilon n^{1/2})) + \varepsilon p^4$$
$$= p^2 L_n(\varepsilon) + \varepsilon p^4.$$

We let first $n\to\infty$ and take into account Condition 1 and then we let $\varepsilon\to 0$. It follows that

$$\lim_{n\to\infty}\sum_{u\geq 1}\mathbb{E}|R_{u3}| = 0. \tag{39}$$

We handle now the quantity $\sum_{u\geq 1}\mathbb{E}|R'_{u3}|$. Taking into account that $\mathbb{E}(Y_{\mathbf{u}}^2)=\mathbb{E}(X_{\mathbf{u}}^2)$, condition 1 and inequality (38), note first that

$$\sum_{u\geq 1}\mathbb{E}|R'_{u3}| \ll p^2\frac{1}{n^2}\sum_{(i,j)\in V_n^1}\mathbb{E}(Y_{ij}^2 I(|Y_{ij}|>\varepsilon n^{1/2})) + \varepsilon p^4.$$

To treat the first term in the right-hand side, some computations are needed. Note first that if N is a centered Gaussian random variable with variance σ^2,

$$\mathbb{E}(N^2 I(|N|>\varepsilon n^{1/2})) = \frac{2\sigma}{\sqrt{2\pi}}\varepsilon\sqrt{n}e^{-\varepsilon^2 n/(2\sigma^2)} + \sigma^2\mathbb{P}(|N|>\varepsilon\sqrt{n}). \tag{40}$$

Let now $\sigma_{ij}^2=\mathbb{E}(X_{ij}^2)$. For any $\eta>0$, we then have

$$\sigma_{ij}^2\mathbb{P}(|Y_{ij}|>\varepsilon\sqrt{n}) \leq \eta^2\varepsilon^2 n\mathbb{P}(|Y_{ij}|>\varepsilon\sqrt{n}) + \mathbb{E}(X_{ij}^2 I(|X_{ij}|>\eta\varepsilon n^{1/2}))$$
$$\leq \eta^2\mathbb{E}(Y_{ij}^2) + \mathbb{E}(X_{ij}^2 I(|X_{ij}|>\eta\varepsilon n^{1/2})) = \eta^2\mathbb{E}(X_{ij}^2) + \mathbb{E}(X_{ij}^2 I(|X_{ij}|>\eta\varepsilon n^{1/2})).$$

On another hand, let A be a positive real. Observe that, for any $\eta>0$,

$$\sigma_{ij}\varepsilon\sqrt{n}e^{-\varepsilon^2 n/(2\sigma_{ij}^2)}I(\sigma_{ij}>\varepsilon\sqrt{n}/A) \leq A\sigma_{ij}^2 I(\sigma_{ij}>\varepsilon\sqrt{n}/A)$$
$$\leq A\mathbb{E}(X_{ij}^2 I(|X_{ij}|>\eta n^{1/2})) + An\eta^2 I(\sigma_{ij}>\varepsilon\sqrt{n}/A)$$
$$\leq A\mathbb{E}(X_{ij}^2 I(|X_{ij}|>\eta n^{1/2})) + \eta^2 A^3\sigma_{ij}^2/\varepsilon^2.$$

Moreover

$$\sigma_{ij}\varepsilon\sqrt{n}e^{-\varepsilon^2 n/(2\sigma_{ij}^2)}I(\sigma_{ij}\leq\varepsilon\sqrt{n}/A) = \sqrt{2}\sigma_{ij}^2\frac{\varepsilon\sqrt{n}}{\sqrt{2}\sigma_{ij}}e^{-\varepsilon^2 n/(2\sigma_{ij}^2)}I(\sigma_{ij}\leq\varepsilon\sqrt{n}/A) \leq 2\sigma_{ij}^2 e^{-A^2/4}.$$

So, overall, taking into account the above considerations and (40), it follows that, for any $\varepsilon>0$, any $\eta>0$ and any positive real A,

$$\mathbb{E}(Y_{ij}^2 I(|Y_{ij}|>\varepsilon n^{1/2})) \ll \eta^2\mathbb{E}(X_{ij}^2) + \mathbb{E}(X_{ij}^2 I(|X_{ij}|>\eta\varepsilon n^{1/2}))$$
$$+ A\mathbb{E}(X_{ij}^2 I(|X_{ij}|>\eta n^{1/2})) + \eta^2 A^3\sigma_{ij}^2/\varepsilon^2 + \mathbb{E}(X_{ij}^2)e^{-A^2/4}.$$

Hence, taking into account condition 1, it follows that for any $\varepsilon>0$, any $\eta>0$ and any $A>0$,

$$\sum_{u\geq 1}\mathbb{E}|R'_{u3}| \ll \varepsilon p^4 + p^2\eta^2 + p^2 L_n(\eta\varepsilon) + Ap^2 L_n(\eta) + p^2\eta^2 A^3/\varepsilon^2 + p^2 e^{-A^2/4}.$$

Letting $n\to\infty$, then $\eta\to 0$ and finally $A\to\infty$, it follows that

$$\limsup_{n\to\infty}\sum_{u\geq 1}\mathbb{E}|R'_{u3}| \ll \varepsilon p^4.$$

Letting then $\varepsilon \to 0$ and taking into account (39), it follows that $\mathbb{E}|R_3'| \to 0$, as $n \to \infty$.

We treat now the term R_1', and we shall compute $\mathbb{E}|R_1'|^2$. Let

$$D_u = \sum_{\mathbf{j} \in \mathbf{I}_u} X_{\mathbf{j}}' \partial_{\mathbf{j}} s(\mathbf{C}_u) \quad \text{and} \quad \widetilde{D}_u = \sum_{\mathbf{j} \in \mathbf{I}_u} Y_{\mathbf{j}}' \partial_{\mathbf{j}} s(\mathbf{C}_u),$$

and note that

$$\mathbb{E}|R_1'|^2 \leq 2\mathbb{E}\Big|\sum_{u \geq 1} D_u\Big|^2 + 2\mathbb{E}\Big|\sum_{u \geq 1} \widetilde{D}_u\Big|^2.$$

By definition of the $X_{\mathbf{j}}'$ for $\mathbf{j}$ in $\mathbf{I}_u$, the random variables $(D_u)_{u \geq 1}$ are orthogonal. Moreover, since the random vectors $(Y_{\mathbf{j}}', \mathbf{j} \in I_u)_u$ are mutually independent, the variables $(\widetilde{D}_u)_{u \geq 1}$ are also orthogonal. Hence we get

$$\mathbb{E}|R_1'|^2 \leq 2\sum_{u \geq 1} \mathbb{E}|D_u|^2 + 2\sum_{u \geq 1} \mathbb{E}|\widetilde{D}_u|^2.$$

Therefore, by using Cauchy-Schwarz's inequality, taking into account (60), the fact that $\mathbb{E}(Y_{\mathbf{u}}^2) = \mathbb{E}(X_{\mathbf{u}}^2)$ and condition 1, it follows that

$$\mathbb{E}|R_1'|^2 \ll \frac{p^2}{n^3} \sum_{(i,j) \in V_n^1} \mathbb{E}(X_{ij}^2) \ll \frac{p^2}{n},$$

which converges to 0 when $n \to \infty$.

Now we treat the term $R_2' = 2^{-1} \sum_{u \geq 1} R_{u2}$ where

$$R_{u2} = \Big(\Big(\sum_{\mathbf{j} \in \mathbf{I}_u} X_{\mathbf{j}}' \partial_{\mathbf{j}}\Big)^2 - \Big(\sum_{\mathbf{j} \in \mathbf{I}_u} Y_{\mathbf{j}}' \partial_{\mathbf{j}}\Big)^2\Big) s(\mathbf{C}_u).$$

We write R_{u2} as a sum of differences of the type $\big(X_{\mathbf{j}}' X_{\mathbf{i}}' - Y_{\mathbf{j}}' Y_{\mathbf{i}}'\big) \partial_{\mathbf{j}} \partial_{\mathbf{i}} s(\mathbf{C}_u)$ were $\mathbf{i}, \mathbf{j} \in \mathbf{I}_u$. To introduce martingale structure we add and substract some terms. Hence we write

$$\begin{aligned}
\big(X_{\mathbf{j}}' X_{\mathbf{i}}' - Y_{\mathbf{j}}' Y_{\mathbf{i}}'\big) \partial_{\mathbf{j}} \partial_{\mathbf{i}} s(\mathbf{C}_u) &= \big(X_{\mathbf{j}}' X_{\mathbf{i}}' - \mathbb{E}(X_{\mathbf{j}}' X_{\mathbf{i}}' | \mathcal{B}_{u-1})\big) \partial_{\mathbf{j}} \partial_{\mathbf{i}} s(\mathbf{C}_u) \\
&\quad + \big(\mathbb{E}(X_{\mathbf{j}}' X_{\mathbf{i}}' | \mathcal{B}_{u-1}) - \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}})\big) \partial_{\mathbf{j}} \partial_{\mathbf{i}} s(\mathbf{C}_u) + \big(\mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}}) - Y_{\mathbf{j}}' Y_{\mathbf{i}}'\big) \partial_{\mathbf{j}} \partial_{\mathbf{i}} s(\mathbf{C}_u) \\
&:= I_{u\mathbf{ij}}^{(1)} + I_{u\mathbf{ij}}^{(2)} + I_{u\mathbf{ij}}^{(3)} . \qquad (41)
\end{aligned}$$

Taking into account the properties of the conditional expectation, we obtain

$$\mathbb{E}(X_{\mathbf{j}}' X_{\mathbf{i}}' | \mathcal{B}_{u-1}) = \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}} | \mathcal{B}_{u-1}) - \mathbb{E}(X_{\mathbf{j}} | \mathcal{B}_{u-1}) \mathbb{E}(X_{\mathbf{i}} | \mathcal{B}_{u-1}).$$

Therefore

$$\begin{aligned}
\Big|\sum_{\mathbf{i},\mathbf{j} \in \mathbf{I}_u} I_{u\mathbf{ij}}^{(2)}\Big| &\leq \Big|\sum_{\mathbf{i},\mathbf{j} \in \mathbf{I}_u} \big(\mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}} | \mathcal{B}_{u-1}) - \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}})\big) \partial_{\mathbf{j}} \partial_{\mathbf{i}} s(\mathbf{C}_u)\Big| \\
&\quad + \Big|\sum_{\mathbf{i},\mathbf{j} \in \mathbf{I}_u} \mathbb{E}(X_{\mathbf{j}} | \mathcal{B}_{u-1}) \mathbb{E}(X_{\mathbf{i}} | \mathcal{B}_{u-1}) \partial_{\mathbf{j}} \partial_{\mathbf{i}} s(\mathbf{C}_u)\Big| \\
&:= I_{1,u} + I_{2,u} . \qquad (42)
\end{aligned}$$

Let us handle the term $I_{2,u}$. By Lemma 13,

$$I_{2,u} \leq \frac{c_4}{n^2} \sum_{\mathbf{i} \in \mathbf{I}_u} |\mathbb{E}(X_{\mathbf{i}} | \mathcal{B}_{u-1})|^2 ,$$

and then

$$\sum_{u\geq 1}\mathbb{E}(I_{2,u}) \leq \frac{c_4}{n^2}\sum_{u\geq 1}\sum_{\mathbf{i}\in\mathbf{I}_u}\|\mathbb{E}(X_{\mathbf{i}}|\mathcal{B}_{u-1})\|_2^2\,. \tag{43}$$

Therefore, using the contractivity of conditional expectation, we get

$$\sum_{u=1}^{q(q+1)/2}\mathbb{E}(I_{2,u}) \ll \frac{1}{n^2}\Big(\frac{n}{p}\Big)^2 p^2 \sup_{i\geq j}\mathbb{E}|\mathbb{E}(X_{ij}|\mathcal{F}_{ij}^K)|^2 \ll \sup_{i\geq j}\mathbb{E}|\mathbb{E}(X_{ij}|\mathcal{F}_{ij}^K)|^2\,.$$

Hence, by condition (6),

$$\lim_{K\to\infty}\limsup_{n\to\infty}\sum_{u=1}^{q(q+1)/2}\mathbb{E}(I_{2,u}) = 0.$$

We handle now the term $I_{1,u}$ in (42). Using (60), we first write

$$\mathbb{E}(I_{1,u}) \ll \frac{1}{n^2}\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u}\|\mathbb{E}(X_{\mathbf{j}}X_{\mathbf{i}}|\mathcal{B}_{u-1}) - \mathbb{E}(X_{\mathbf{j}}X_{\mathbf{i}})\|_1.$$

By using the contractivity of conditional expectation, we then get

$$\begin{aligned}\sum_{u\geq 1}\mathbb{E}(I_{1,u}) &\ll \frac{1}{n^2}\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u}\mathbb{E}|\mathbb{E}(X_{\mathbf{j}}X_{\mathbf{i}}|\mathcal{B}_{u-1}) - \mathbb{E}(X_{\mathbf{j}}X_{\mathbf{i}})| \qquad (44)\\
&\ll \frac{1}{n^2}\Big(\frac{n}{p}\Big)^2 p^4 \sup_{(i,j)\neq(a,b);i\geq j,a\geq b}\mathbb{E}|\mathbb{E}(X_{ij}X_{ab}|\mathcal{F}_{i\wedge a,j\wedge b}^K) - \mathbb{E}(X_{ij}X_{ab})|\\
&\quad + \frac{1}{n^2}\Big(\frac{n}{p}\Big)^2 p^2 \sup_{i\geq j}\mathbb{E}|\mathbb{E}(X_{ij}^2|\mathcal{F}_{ij}^K) - \mathbb{E}(X_{ij}^2)|\\
&\ll p^2 \sup_{(i,j)\neq(a,b);i\geq j,a\geq b}\mathbb{E}|\mathbb{E}(X_{ij}X_{ab}|\mathcal{F}_{i\wedge a,j\wedge b}^K) - \mathbb{E}(X_{ij}X_{ab})|\\
&\quad + \sup_{i\geq j}\mathbb{E}|\mathbb{E}(X_{ij}^2|\mathcal{F}_{ij}^K) - \mathbb{E}(X_{ij}^2)|.\end{aligned}$$

The first term converges to 0 by (27) and the second term is convergent to 0 by (8). Hence,

$$\lim_{K\to\infty}\limsup_{n\to\infty}\sum_{u\geq 1}\mathbb{E}(I_{1,u}) = 0.$$

Overall, starting from (42) and taking into account the above considerations, we get

$$\sum_{u\geq 1}\mathbb{E}\Big|\sum_{\mathbf{i}\in\mathbf{I}_u}\sum_{\mathbf{j}\in\mathbf{I}_u} I_{u\mathbf{i}\mathbf{j}}^{(2)}\Big| = 0.$$

We treat now the negligibility of the term $\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u} I_{u\mathbf{i}\mathbf{j}}^{(1)}$ in the following way. First we truncate

$$\bar{X}_{\mathbf{i}}' = X_{\mathbf{i}}' I(|X_{\mathbf{i}}'| \leq 4n^{1/2}) \text{ and } \tilde{X}_{\mathbf{i}}' = X_{\mathbf{i}}' I(|X_{\mathbf{i}}'| > 4n^{1/2})$$

and write

$$X_{\mathbf{j}}'X_{\mathbf{i}}' - \mathbb{E}(X_{\mathbf{j}}'X_{\mathbf{i}}'|\mathcal{B}_{u-1}) = X_{\mathbf{j}}'\bar{X}_{\mathbf{i}}' - \mathbb{E}(X_{\mathbf{j}}'\bar{X}_{\mathbf{i}}'|\mathcal{B}_{u-1}) + X_{\mathbf{j}}'\tilde{X}_{\mathbf{i}}' - \mathbb{E}(X_{\mathbf{j}}'\tilde{X}_{\mathbf{i}}'|\mathcal{B}_{u-1})\,.$$

Therefore, by the triangle inequality, the Cauchy-Schwarz inequality, the Minkowski's inequalities and the properties of conditional expectation, we easily obtain

$$\begin{aligned}\mathbb{E}\Big|\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u} I_{u\mathbf{i}\mathbf{j}}^{(1)}\Big| &\ll \frac{1}{n^2}\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u}\|X_{\mathbf{j}}'\|_2\|\tilde{X}_{\mathbf{i}}'\|_2\\
&+\Big\|\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u}[X_{\mathbf{j}}'\bar{X}_{\mathbf{i}}' - \mathbb{E}(X_{\mathbf{j}}'\bar{X}_{\mathbf{i}}'|\mathcal{B}_{u-1})]\partial_{\mathbf{j}}\partial_{\mathbf{i}}s(\mathbf{C}_u)\Big\|_2\\
&:= A_n + \Big\|\sum_{u\geq 1}D_u'\Big\|_2\,.\end{aligned}$$

By the fact that that the terms D'_u are orthogonal, by (60), the level of truncation and condition 1, we have

$$\begin{aligned}\mathbb{E}\Big|\sum_{u\geq 1} D'_u\Big|^2 &\ll \frac{1}{n^4}\sum_{u\geq 1}\mathbb{E}\Big(\sum_{i,j\in\mathbf{I}_u}|X'_{\mathbf{j}}\bar{X}'_{\mathbf{i}}-\mathbb{E}(X'_{\mathbf{j}}\bar{X}'_{\mathbf{i}}|\mathcal{B}_{u-1})|\Big)^2\\ &\ll \frac{1}{n^4}\sum_{u\geq 1}\Big(\sum_{i,j\in\mathbf{I}_u}\|X'_{\mathbf{j}}\bar{X}'_{\mathbf{i}}\|_2\Big)^2 \ll \frac{p^4}{n^4}\sum_{u\geq 1}\sum_{i,j\in\mathbf{I}_u}\|X'_{\mathbf{j}}\bar{X}'_{\mathbf{i}}\|_2^2\\ &\ll \frac{p^6}{n^3}\sum_{u\geq 1}\sum_{j\in\mathbf{I}_u}\mathbb{E}(X_{\mathbf{j}}^2) \ll \frac{p^6}{n}\to 0 \text{ as } n\to\infty.\end{aligned}$$

Also, by the Cauchy-Schwarz's inequality and condition 1,

$$\begin{aligned}A_n = \frac{1}{n^2}\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u}\|X'_{\mathbf{j}}\|_2\|\tilde{X}'_{\mathbf{i}}\|_2 &\leq p^2\Big(\frac{1}{n^2}\sum_{u\geq 1}\sum_{\mathbf{j}\in\mathbf{I}_u}\|X'_{\mathbf{j}}\|_2^2\Big)^{1/2}\Big(\frac{1}{n^2}\sum_{u\geq 1}\sum_{\mathbf{i}\in\mathbf{I}_u}\|\tilde{X}'_{\mathbf{i}}\|_2^2\Big)^{1/2}\\ &\leq p^2\Big(\frac{1}{n^2}\sum_{\mathbf{i}\in V_n^1}\mathbb{E}(X'^2_{\mathbf{i}}I(|X'_{\mathbf{i}}|>4n^{1/2}))\Big)^{1/2}.\end{aligned}$$

Using (38), we derive that $A_n \ll\ p^2\ L_n^{1/2}(1)$, which converges to 0 for any p fixed as $n\to\infty$. Overall, it follows that

$$\mathbb{E}\Big|\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u} I^{(1)}_{u\mathbf{i}\mathbf{j}}\Big| \to 0\,,\ \text{as } n\to\infty.$$

To end the proof, it remains only to treat the term containing the Gaussian random variables. With this aim, we write $I^{(3)}_{u\mathbf{i}\mathbf{j}} = A_{u\mathbf{i}\mathbf{j}} + B_{u\mathbf{i}\mathbf{j}}$, where

$$A_{u\mathbf{i}\mathbf{j}} := (\mathbb{E}(X_{\mathbf{j}}X_{\mathbf{i}}) - \mathbb{E}(Y'_{\mathbf{j}}Y'_{\mathbf{i}}))\partial_{\mathbf{j}}\partial_{\mathbf{i}}s(\mathbf{C}_u)$$

and

$$B_{u\mathbf{i}\mathbf{j}} := (\mathbb{E}(Y'_{\mathbf{j}}Y'_{\mathbf{i}}) - Y'_{\mathbf{j}}Y'_{\mathbf{i}})\partial_{\mathbf{j}}\partial_{\mathbf{i}}s(\mathbf{C}_u)\,.$$

We use the orthogonality of $\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u} B_{u\mathbf{i}\mathbf{j}}$ and (60). This leads to

$$\begin{aligned}\mathbb{E}\Big|\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u} B_{u\mathbf{i}\mathbf{j}}\Big|^2 &\ll \frac{1}{n^4}\sum_{u\geq 1}\Big\|\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u}|\mathbb{E}(Y'_{\mathbf{j}}Y'_{\mathbf{i}}) - Y'_{\mathbf{j}}Y'_{\mathbf{i}}|\Big\|_2^2\\ \leq \frac{1}{n^4}\sum_{u\geq 1}\Big(\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u}\|\mathbb{E}(Y'_{\mathbf{j}}Y'_{\mathbf{i}}) - Y'_{\mathbf{j}}Y'_{\mathbf{i}}\|_2\Big)^2 &\leq \frac{4}{n^4}\sum_{u\geq 1}\Big(\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u}\|Y'_{\mathbf{j}}\|_4\|Y'_{\mathbf{i}}\|_4\Big)^2\\ &\leq \frac{4^3}{n^4}\sum_{u\geq 1}\Big(\sum_{\mathbf{i}\in\mathbf{I}_u}\|Y_{\mathbf{i}}\|_4\Big)^4 \leq \frac{4^3p^6}{n^4}\sum_{u\geq 1}\sum_{\mathbf{i}\in\mathbf{I}_u}\|Y_{\mathbf{i}}\|_4^4\,.\end{aligned}$$

Since the r.v.'s $Y_{\mathbf{i}}$ are Gaussian, $\|Y_{\mathbf{i}}\|_4^4 = 3\|Y_{\mathbf{i}}\|_2^4 = 3\|X_{\mathbf{i}}\|_2^4$. Therefore, for any $\varepsilon>0$,

$$\begin{aligned}\mathbb{E}\Big|\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in\mathbf{I}_u} B_{u\mathbf{i}\mathbf{j}}\Big|^2 &\ll \frac{p^6}{n^4}\sum_{u\geq 1}\sum_{\mathbf{i}\in\mathbf{I}_u}\|X_{\mathbf{i}}\|_2^4\\ &\ll \frac{p^6\varepsilon^2}{n^3}\sum_{u\geq 1}\sum_{\mathbf{i}\in\mathbf{I}_u}\|X_{\mathbf{i}}\|_2^2 + \frac{p^6}{n^4}\sum_{u\geq 1}\sum_{\mathbf{i}\in\mathbf{I}_u}\|X_{\mathbf{i}}^2 I(|X_{\mathbf{i}}|>\varepsilon\sqrt{n})\|_1^2\\ &\ll \frac{p^6\varepsilon^2}{n^3}\sum_{(i,j)\in V_n^1}\|X_{ij}\|_2^2 + p^6\Big(\frac{1}{n^2}\sum_{(i,j)\in V_n^1}\|X_{ij}^2 I(|X_{ij}|>\varepsilon\sqrt{n})\|_1\Big)^2.\end{aligned}$$

Letting $n \to \infty$ and after $\varepsilon \to 0$, and taking into account condition 1, it follows that

$$\mathbb{E}\Big|\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in \mathbf{I}_u} B_{u\mathbf{ij}}\Big|^2 \to 0 \ \text{ as } n \to \infty.$$

On the other hand, since

$$\mathbb{E}(Y'_{\mathbf{j}}Y'_{\mathbf{i}}) = \mathbb{E}(Y_{\mathbf{j}}Y_{\mathbf{i}}) - \mathbb{E}\big(\mathbb{E}(Y_{\mathbf{j}}|\mathcal{H}_{u-1})\mathbb{E}(Y_{\mathbf{i}}|\mathcal{H}_{u-1})\big)$$

and $\mathbb{E}(Y_{\mathbf{j}}Y_{\mathbf{i}}) = \mathbb{E}(X_{\mathbf{j}}X_{\mathbf{i}})$, we get, by the same arguments as those leading to (43),

$$\mathbb{E}\Big|\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in \mathbf{I}_u} A_{u\mathbf{ij}}\Big| \ll \frac{1}{n^2}\sum_{u\geq 1}\sum_{\mathbf{i}\in \mathbf{I}_u} \|\mathbb{E}(Y_{\mathbf{i}}|\mathcal{H}_{u-1})\|_2^2\,.$$

Hence, by (33) and the contractivity of conditional expectation,

$$\mathbb{E}\Big|\sum_{u\geq 1}\sum_{\mathbf{i},\mathbf{j}\in \mathbf{I}_u} A_{u\mathbf{ij}}\Big| \ll \frac{1}{n^2}\sum_{u\geq 1}\sum_{\mathbf{i}\in \mathbf{I}_u} \|\mathbb{E}(X_{\mathbf{i}}|\mathcal{B}_{u-1})\|_2^2 \ll \sup_{i\geq j}\|\mathbb{E}(X_{ij}|\mathcal{F}^K_{ij})\|_2^2\,,$$

which converges to 0, as $K \to \infty$, by condition (6). This completes the proof of the theorem. $\Diamond$

5.2 Proof of Theorem 4

The proof follows the lines of the proof of Theorem 2 with Z_{ij} instead of X_{ij} and with W_{ij} instead of Y_{ij}. We point here the differences. The filtrations $\mathcal{B}_u$ and $\mathcal{H}_u$ respectively defined in (31) and (32) have to be defined as follows. If $u = k(k-1)/2 + \ell$ with $1 \leq \ell \leq k$ and $1 \leq k \leq q$, then

$$\mathcal{B}_u = \mathcal{B}'_u \vee \mathcal{B}''_u \text{ for } u \geq 1 \text{ and } \mathcal{B}_0 = \{\emptyset, \Omega\} \tag{45}$$

and

$$\mathcal{H}_u = \mathcal{H}'_u \vee \mathcal{H}''_u \text{ for } u \geq 1 \text{ and } \mathcal{H}_0 = \{\emptyset, \Omega\}, \tag{46}$$

where

$$\mathcal{B}'_u = \mathcal{B}'_{k,\ell} = \sigma\big(X_{ab}\,, (a,b) \in \cup_{j=1}^{\ell}\mathcal{E}_{kj} \text{ or } (a,b) \in \cup_{i=1}^{k-1}\cup_{j=1}^{i}\mathcal{E}_{ij}\big)\,,$$
$$\mathcal{B}''_u = \mathcal{B}''_{k,\ell} = \sigma\big(X_{ba}\,, (a,b) \in \cup_{j=1}^{\ell}\mathcal{E}_{kj} \text{ or } (a,b) \in \cup_{i=1}^{k-1}\cup_{j=1}^{i}\mathcal{E}_{ij}\big)\,,$$

and $\mathcal{H}'_u$ and $\mathcal{H}''_u$ defined, respectively, as $\mathcal{B}''_u$ and $\mathcal{B}''_u$ with the X_{ab} (resp. X_{ba}) replaced by Y_{ab} (resp. Y_{ba}). According to the proof of Theorem 2, the proof will be achieved if we can show that if $u = k(k-1)/2 + \ell$ with $1 \leq \ell \leq k$ and $1 \leq k \leq q$, then for any (i,j) and (a,b) in $\mathcal{E}_{k,\ell}$

$$\|\mathbb{E}(Z_{ij}|\mathcal{B}_{u-1})\|_2 \leq \|\mathbb{E}(X_{ij}|\widetilde{\mathcal{F}}^K_{ij})\|_2 + \|\mathbb{E}(X_{ji}|\widetilde{\mathcal{F}}^K_{ji})\|_2 \tag{47}$$

and

$$\begin{aligned}\|\mathbb{E}(Z_{ij}Z_{ab}|\mathcal{B}_{u-1}) - \mathbb{E}(Z_{ij}Z_{ab})\|_1 &\leq \|\mathbb{E}(X_{ij}X_{ab}|\widetilde{\mathcal{F}}^K_{ij}\cap\widetilde{\mathcal{F}}^K_{ab}) - \mathbb{E}(X_{ij}X_{ab})\|_1 \\ &+ \|\mathbb{E}(X_{ij}X_{ba}|\widetilde{\mathcal{F}}^K_{ij}\cap\widetilde{\mathcal{F}}^K_{ba}) - \mathbb{E}(X_{ij}X_{ba})\|_1 + \|\mathbb{E}(X_{ji}X_{ab}|\widetilde{\mathcal{F}}^K_{ji}\cap\widetilde{\mathcal{F}}^K_{ab}) - \mathbb{E}(X_{ji}X_{ab})\|_1 \\ &+ \|\mathbb{E}(X_{ji}X_{ba}|\widetilde{\mathcal{F}}^K_{ji}\cap\widetilde{\mathcal{F}}^K_{ba}) - \mathbb{E}(X_{ji}X_{ba})\|_1. \end{aligned} \tag{48}$$

To prove the inequalities above, we fix, all along the rest of the proof k and ℓ such that $1 \leq k \leq q$ and $1 \leq \ell \leq k$ and also a (i,j) in $\mathcal{E}_{k\ell}$. We notice that if (u,v) belongs to $\cup_{m=1}^{\ell-1}\mathcal{E}_{km}$ then $j - v \geq K$, and if (a,b) belongs to $\cup_{r=1}^{k-1}\cup_{m=1}^{r}\mathcal{E}_{rm}$ then $i - a \geq K$. So $\mathcal{H}'_{u-1} \subseteq \widetilde{\mathcal{F}}^K_{ij}$. In addition, if (a,b) belongs to $\cup_{m=1}^{\ell-1}\mathcal{E}_{km}$ then $i - b \geq p + 2K$ and if (a,b) belongs to $\cup_{r=1}^{k-1}\cup_{m=1}^{r}\mathcal{E}_{rm}$ then $i - b \geq 2K + p$. Therefore the distance between (i,j) and all the points (v,u) such that (a,b) belongs either to

$\cup_{m=1}^{\ell-1}\mathcal{E}_{km}$ or to $\cup_{r=1}^{k-1}\cup_{m=1}^{r}\mathcal{E}_{rm}$ is larger than K. This shows that $\mathcal{B}''_{u-1}\subseteq\widetilde{\mathcal{F}}_{ij}^{K}$. The two latter inclusions prove that $\mathcal{B}_{u-1}\subseteq\widetilde{\mathcal{F}}_{ij}^{K}$. Let us prove now that $\mathcal{H}_{u-1}\subseteq\widetilde{\mathcal{F}}_{ji}^{K}$. If (a,b) belongs to $\cup_{m=1}^{\ell-1}\mathcal{E}_{km}$ then $a-j\geq K$, and if (a,b) belongs to $\cup_{r=1}^{k-1}\cup_{m=1}^{r}\mathcal{E}_{rm}$ then $i-b\geq 2K+p$. So $\mathcal{H}'_{u-1}\subseteq\widetilde{\mathcal{F}}_{ji}^{K}$. In addition, if (a,b) belongs to $\cup_{m=1}^{\ell-1}\mathcal{E}_{km}$ then $j-b\geq K$ and if (a,b) belongs to $\cup_{r=1}^{k-1}\cup_{m=1}^{r}\mathcal{E}_{rm}$ then $i-a\geq K$. Therefore the distance between (j,i) and all the points (b,a) such that (a,b) belongs to $\cup_{m=1}^{\ell-1}\mathcal{E}_{km}$ or to $\cup_{r=1}^{k-1}\cup_{m=1}^{r}\mathcal{E}_{rm}$ is larger than K. Hence $\mathcal{B}''_{u-1}\subseteq\widetilde{\mathcal{F}}_{ji}^{K}$. This ends the proof of $\mathcal{B}_{u-1}\subseteq\widetilde{\mathcal{F}}_{ji}^{K}$. Since $\mathcal{B}_{u-1}\subseteq\widetilde{\mathcal{F}}_{ji}^{K}$ and $\mathcal{B}_{u-1}\subseteq\widetilde{\mathcal{F}}_{ji}^{K}$, (47) follows by applying the tower lemma and using contraction. To prove (48), we use again the tower lemma together with the contractivity of the norm for the conditional expectation and the fact that the above inclusions imply that for any (i,j) and (a,b) belonging to $\mathcal{E}_{k,\ell}$, $\mathcal{B}_{u-1}\subseteq\widetilde{\mathcal{F}}_{ij}^{K}\cap\widetilde{\mathcal{F}}_{ab}^{K}$, $\mathcal{B}_{u-1}\subseteq\widetilde{\mathcal{F}}_{ij}^{K}\cap\widetilde{\mathcal{F}}_{ba}^{K}$, $\mathcal{B}_{u-1}\subseteq\widetilde{\mathcal{F}}_{ji}^{K}\cap\widetilde{\mathcal{F}}_{ab}^{K}$ and $\mathcal{B}_{u-1}\subseteq\widetilde{\mathcal{F}}_{ji}^{K}\cap\widetilde{\mathcal{F}}_{ba}^{K}$. $\Diamond$

5.3 Proof of Theorem 5

The proof is very similar to the proof of Theorem 9 from Merlevède and Peligrad (2016). We give it here for completeness.

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

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

Notice that the eigenvalues of $\mathbb{X}_n^2$ are the eigenvalues of $N^{-1}\mathcal{X}_{N,p}\mathcal{X}_{N,p}^{T}$ together with the eigenvalues of $N^{-1}\mathcal{X}_{N,p}^{T}\mathcal{X}_{N,p}$. Assuming that $N\leq p$ (otherwise exchange the role of $\mathcal{X}_{N,p}$ and $\mathcal{X}_{N,p}^{T}$ everywhere), the following relation holds: for any $z\in\mathbb{C}^+$

$$S^{\mathbb{B}_N}(z)=z^{-1/2}\frac{n}{2N}S^{\mathbb{X}_n}(z^{1/2})+\frac{N-p}{2Nz}. \tag{49}$$

(See, for instance, page 549 in Rashidi Far *et al.* (2008) for additional arguments leading to the relation above). Consider now a real-valued centered Gaussian random field $(Y_{k\ell})_{(k,\ell)\in\mathbb{Z}^2}$ independent of $(X_{k\ell})_{(k,\ell)\in\mathbb{Z}^2}$ and with covariance function given by:

$$\mathbb{E}(Y_{k\ell}Y_{ij})=\mathbb{E}(X_{k\ell}X_{ij})\text{ for any }(k,\ell)\text{ and }(i,j)\text{ in }\mathbb{Z}^2, \tag{50}$$

and define the $N\times p$ matrix

$$\Gamma_{N,p}=\big(Y_{ij}\big)_{1\leq i\leq N,1\leq j\leq p}.$$

Let $\mathbb{G}_N=\frac{1}{N}\Gamma_{N,p}\Gamma_{N,p}^{T}$ and $\mathbb{H}_n$ be the symmetric matrix of order n defined by

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

Assuming that $N\leq p$, the following relation holds: for any $z\in\mathbb{C}^+$

$$S^{\mathbb{G}_N}(z)=z^{-1/2}\frac{n}{2N}S^{\mathbb{H}_n}(z^{1/2})+\frac{N-p}{2Nz}. \tag{51}$$

In view of relations (49) and (51), to prove that for any $z\in\mathbb{C}^+$,

$$\big|S^{\mathbb{B}_N}(z)-S^{\mathbb{G}_N}(z)\big|\to 0\ \text{ in probability} \tag{52}$$

it suffices to prove that, for any $z\in\mathbb{C}^+$,

$$\big|S^{\mathbb{X}_n}(z)-S^{\mathbb{H}_n}(z)\big|\to 0\ \text{ in probability} \tag{53}$$

(since $n/N \to 1+c$). Clearly (53) follows from the proof of Theorem 2 together with Comment 3 (iii), by noticing the following facts. The entries $x_{i,j}$ and $g_{i,j}$ of the matrices $n^{1/2}\mathbb{X}_n$ and $n^{1/2}\mathbb{H}_n$ respectively, satisfy

$$x_{i,j} = \alpha_{i,j}^{(n)} X_{ji} \ , \ g_{i,j} = \alpha_{i,j}^{(n)} Y_{ji} \text{ if } 1 \le j \le i \le n \text{ and } x_{i,j} = x_{j,i} \ , \ g_{i,j} = g_{j,i} \text{ if } 1 \le j \le i \le n \, ,$$

where $(\alpha_{i,j}^{(n)})$ is a sequence of positive numbers defined by: $\alpha_{i,j}^{(n)} = \frac{n^{1/2}}{N^{1/2}} \mathbf{1}_{N+1 \le i \le n} \mathbf{1}_{1 \le j \le N}$. Hence $\mathbb{E}(g_{k,\ell} g_{i,j}) = \alpha_{k,\ell}^{(n)} \alpha_{i,j}^{(n)} \mathbb{E}(X_{k\ell} X_{ij})$, $\max_{1 \le j \le i \le n} \alpha_{i,j} = \frac{n^{1/2}}{N^{1/2}} := \alpha^{(n)}$ and $\lim_{n\to\infty} \alpha^{(n)} = \sqrt{1+c}$. $\lozenge$

5.4 Proof of Theorem 6

The proof of this theorem follows all the steps of the proof of Theorem 2 (with the same notations) with the exception of the treatment of terms which appear in (44). By stationarity

$$\frac{1}{n^2} \sum_{u=1}^{q(q+1)/2} \sum_{\mathbf{i},\mathbf{j} \in \mathbf{I}_u} \mathbb{E}|\mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}} | \mathcal{B}_{u-1}) - \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}})| \le \frac{1}{p^2} \sum_{\mathbf{i},\mathbf{j} \in \mathcal{E}_p} \mathbb{E}|\mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}} | \mathcal{F}_{\mathbf{0}}^K) - \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}})| \, ,$$

where above and below $\mathcal{E}_p = [1,p]^2 \cap \mathbb{N}^2$. For $\mathbf{i}$ fixed in $\mathcal{E}_p$, we shall divide the last sum in three parts according to $\mathbf{j} \in \mathcal{E}_p$, with $|\mathbf{i}-\mathbf{j}| \le d$ or $d \le |\mathbf{i}-\mathbf{j}| \le K$ or $|\mathbf{i}-\mathbf{j}| > K$, where d is a positive integer less than K. Since for this case $\mathcal{F}_{\mathbf{0}}^K \subset \mathcal{F}_{\mathbf{i}\wedge\mathbf{j}}^K$, by the properties of conditional expectations and stationarity we have

$$\begin{aligned} \sum_{\mathbf{j} \in \mathcal{E}_p, |\mathbf{i}-\mathbf{j}| \le d} \mathbb{E}|\mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}} | \mathcal{F}_{\mathbf{0}}^K) - \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}})| &\le \sum_{\mathbf{j} \in \mathcal{E}_p, |\mathbf{i}-\mathbf{j}| \le d} \mathbb{E}|\mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}} | \mathcal{F}_{\mathbf{i}\wedge\mathbf{j}}^K) - \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}})| \\ &= \sum_{\mathbf{j} \in \mathcal{E}_p, |\mathbf{i}-\mathbf{j}| \le d} \mathbb{E}|\mathbb{E}(X_{\mathbf{0}} X_{\mathbf{j}-\mathbf{i}} | \mathcal{F}_{(\mathbf{j}-\mathbf{i})\wedge\mathbf{0}}^K) - \mathbb{E}(X_{\mathbf{0}} X_{\mathbf{j}-\mathbf{i}})| \, . \end{aligned}$$

Therefore

$$\frac{1}{p^2} \sum_{\mathbf{i},\mathbf{j} \in \mathcal{E}_p, |\mathbf{i}-\mathbf{j}| \le d} \mathbb{E}|\mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}} | \mathcal{F}_{\mathbf{0}}^K) - \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}})| \le \sum_{\mathbf{u}, |\mathbf{u}| \le d} \mathbb{E}|\mathbb{E}(X_{\mathbf{0}} X_{\mathbf{u}} | \mathcal{F}_{\mathbf{u}\wedge\mathbf{0}}^K) - \mathbb{E}(X_{\mathbf{0}} X_{\mathbf{u}})|,$$

which converges to 0 for d fixed as $K \to \infty$ by the regularity condition of the random field.

Now we treat the part of the sum where $\mathbf{j} \in \mathcal{E}_p$, with $d < |\mathbf{i}-\mathbf{j}| \le K$. For this case we note that $\mathcal{F}_{\mathbf{0}}^K \subset \mathcal{F}_{\mathbf{i}}^{|\mathbf{i}-\mathbf{j}|}$ and $\mathcal{F}_{\mathbf{0}}^K \subset \mathcal{F}_{\mathbf{j}}^{|\mathbf{i}-\mathbf{j}|}$. By the properties of conditional expectation, stationarity and some computations we infer that

$$\sum_{\mathbf{j} \in \mathcal{E}_p, d < |\mathbf{i}-\mathbf{j}| \le K} \mathbb{E}|\mathbb{E}(X_{\mathbf{i}} X_{\mathbf{j}} | \mathcal{F}_{\mathbf{0}}^K) - \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}})| \le 2 \sum_{\mathbf{u} \in V_0, |\mathbf{u}| > d} \mathbb{E}|X_{\mathbf{u}} \mathbb{E}(X_{\mathbf{0}} | \mathcal{F}_{\mathbf{0}}^{|\mathbf{u}|})| \, ,$$

where we recall that $V_0 = \{\mathbf{u} = (u_1, u_2) \in \mathbb{Z}^2 \, : \, u_1 \le 0 \text{ or } u_2 \le 0\}$. It follows that

$$\frac{1}{p^2} \sum_{\mathbf{i},\mathbf{j} \in \mathcal{E}_p, d < |\mathbf{i}-\mathbf{j}| \le K} \mathbb{E}|\mathbb{E}(X_{\mathbf{i}} X_{\mathbf{j}} | \mathcal{F}_{\mathbf{0}}^K) - \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}})| \le 2 \sum_{\mathbf{u} \in V_0, |\mathbf{u}| > d} \mathbb{E}|X_{\mathbf{u}} \mathbb{E}(X_{\mathbf{0}} | \mathcal{F}_{\mathbf{0}}^{|\mathbf{u}|})| \, ,$$

which converges to 0 as $d \to \infty$ uniformly in p and K by (19).

Finally, for the third sum where $\mathbf{i},\mathbf{j} \in \mathcal{E}_p$, $|\mathbf{i}-\mathbf{j}| \ge K$ we either have $\sigma(X_{\mathbf{i}}) \subset \mathcal{F}_{\mathbf{j}}^K$ or $\sigma(X_{\mathbf{j}}) \subset \mathcal{F}_{\mathbf{i}}^K$. Moreover $\mathcal{F}_{\mathbf{0}}^K \subset \mathcal{F}_{\mathbf{i}}^K$ and $\mathcal{F}_{\mathbf{0}}^K \subset \mathcal{F}_{\mathbf{j}}^K$. By the properties of conditional expectation, when $\sigma(X_{\mathbf{i}}) \subset \mathcal{F}_{\mathbf{j}}^K$ we have

$$\mathbb{E}|\mathbb{E}(X_{\mathbf{i}} X_{\mathbf{j}} | \mathcal{F}_{\mathbf{0}}^K) - \mathbb{E}(X_{\mathbf{j}} X_{\mathbf{i}})| \le 2\mathbb{E}(|\mathbb{E}(X_{\mathbf{i}} X_{\mathbf{j}} | \mathcal{F}_{\mathbf{j}}^K)|) = 2\mathbb{E}(|X_{\mathbf{i}-\mathbf{j}} \mathbb{E}(X_{\mathbf{0}} | \mathcal{F}_{\mathbf{0}}^K)|) \, .$$

When $\sigma(X_{\mathbf{j}}) \subset \mathcal{F}_{\mathbf{i}}^K$, similarly, we have

$$\mathbb{E}|\mathbb{E}(X_{\mathbf{i}}X_{\mathbf{j}}|\mathcal{F}_{\mathbf{0}}^K) - \mathbb{E}(X_{\mathbf{j}}X_{\mathbf{i}})| \leq 2\mathbb{E}(|X_{\mathbf{j}-\mathbf{i}}\mathbb{E}(X_{\mathbf{0}}|\mathcal{F}_{\mathbf{0}}^K)|)\,.$$

Therefore

$$\frac{1}{p^2}\sum_{\mathbf{i},\mathbf{j}\in\mathcal{E}_p,|\mathbf{i}-\mathbf{j}|\geq K}\mathbb{E}|\mathbb{E}(X_{\mathbf{i}}X_{\mathbf{j}}|\mathcal{F}_{\mathbf{0}}^K) - \mathbb{E}(X_{\mathbf{j}}X_{\mathbf{i}})| \leq 2p^2 \sup_{\mathbf{u}\in V_0:|\mathbf{u}|>K}\mathbb{E}|X_{\mathbf{u}}\mathbb{E}(X_{\mathbf{0}}|\mathcal{F}_{\mathbf{0}}^K)|\,.$$

Since we can take K close to p, the result follows by letting first $p \to \infty$ followed by $d \to \infty$. $\diamondsuit$

5.5 Proof of Proposition 10

The proof uses similar arguments as those given in the proof of Theorem 2.1 in Chakrabarty *et al.* (2016).

For any integers k and ℓ define

$$c_{k\ell} = \int_{[0,1]^2} \mathrm{e}^{-2\pi\mathrm{i}(kx+\ell y)}\sqrt{f(x,y)}dxdy\,.$$

There are real numbers and satisfy $\sum_{k,\ell\in\mathbb{Z}} c_{k\ell}^2 < \infty$. Let now $(U_{ij})_{(i,j)\in\mathbb{Z}^2}$ be i.i.d. real-valued random variables with law $\mathcal{N}(0,1)$. Then, without restricting the generality, we can write

$$Y_{ij} = \sum_{k,\ell\in\mathbb{Z}} c_{k\ell}U_{i-k,j-\ell}\,. \tag{54}$$

(See Fact 4.1 in Chakrabarty *et al.* (2016)).

The result will follow if we can prove that when N, p tend to infinity such that $p/N \to c$, then there exists a deterministic probability measure μ_f depending only on c and f, and such that for any $\varepsilon > 0$,

$$\mathbb{P}\big(d(\mu_{\mathbb{G}_N}, \mu_f) > \varepsilon\big) \to 0 \text{ as } N \to \infty. \tag{55}$$

Clearly the identity (54) holds in distribution, hence to prove (55), without loss of generality, we may and do assume from now on, that Y_{ij} is given by (54). To prove (55), we shall use Fact 4.3 in Chakrabarty *et al.* (2016)) and first truncate the series (54). Hence we fix a positive integer m and we define

$$Y_{ij}^{(m)} = \sum_{k=-m}^{m}\sum_{\ell=-m}^{m} c_{k\ell}U_{i-k,j-\ell}\,.$$

Let $\Gamma_{N,p}^{(m)} = \big(Y_{ij}^{(m)}\big)_{1\leq i\leq N, 1\leq j\leq p}$. Define also $\mathbb{G}_N^{(m)} = \frac{1}{N}\Gamma_{N,p}^{(m)}(\Gamma_{N,p}^{(m)})^T$. By Theorem 2.1 in Boutet de Monvel *et al.* (1996), we have that for any positive integer m, there exists a deterministic probability measure μ_m depending only on c and on the complex-valued function $\chi^{(m)}$ defined on $[0,1]^2$ by $\chi^{(m)} = \sum_{k,\ell\in\mathbb{Z}}\mathbb{E}(Y_{00}^{(m)}Y_{k\ell}^{(m)})\mathrm{e}^{-2\pi\mathrm{i}(kx+\ell y)}$, and such that for any $\varepsilon > 0$,

$$\mathbb{P}\big(d(\mu_{\mathbb{G}_N^{(m)}}, \mu_m) > \varepsilon\big) \to 0 \text{ as } N \to \infty. \tag{56}$$

Notice that

$$\mathbb{E}(Y_{00}^{(m)}Y_{k\ell}^{(m)}) = \sum_{u=\max(k-m,m)}^{\min(k+m,m)}\ \sum_{v=\max(\ell-m,m)}^{\min(\ell+m,m)} c_{uv}c_{k-u,\ell-v}\,.$$

Since the $c_{k\ell}$'s depend only on f, it follows that $\chi^{(m)}$ depends only on m and f. Therefore μ_m can be rewritten as $\mu_{m,f}$. Notice now that by Corollary A.42 in Bai and Silverstein (2010), for any $\varepsilon > 0$,

$$\mathbb{P}\big(d(\mu_{\mathbb{G}_N}, \mu_{\mathbb{G}_N^{(m)}}) > \varepsilon\big) \le \frac{1}{\varepsilon^2}\mathbb{E}(d^2(\mu_{\mathbb{G}_N}, \mu_{\mathbb{G}_N^{(m)}})$$

$$\le \frac{\sqrt{2}}{p\sqrt{N}\varepsilon^2}\big\|\mathrm{Tr}^{1/2}\big(\mathbb{G}_N^{(m)} + \mathbb{G}_N\big)\mathrm{Tr}^{1/2}\big((\Gamma_{N,p}^{(m)} - \Gamma_{N,p})(\Gamma_{N,p}^{(m)} - \Gamma_{N,p})^T\big)\big\|_1 .$$

Therefore, by the Cauchy-Schwarz's inequality and simple algebra,

$$\mathbb{P}\big(d(\mu_{\mathbb{G}_N}, \mu_{\mathbb{G}_N^{(m)}}) > \varepsilon\big) \le \frac{\sqrt{2}}{p\sqrt{N}\varepsilon^2}\big\|\mathrm{Tr}\big(\mathbb{G}_N^{(m)} + \mathbb{G}_N\big)\big\|_1^{1/2}\big\|\mathrm{Tr}^{1/2}\big((\Gamma_{N,p}^{(m)} - \Gamma_{N,p})(\Gamma_{N,p}^{(m)} - \Gamma_{N,p})^T\big)\big\|_1^{1/2}$$

$$\ll \Big(\sum_{k,\ell\in\mathbb{Z}\,:\,|k|\vee|\ell|>m} c_{k\ell}^2\Big)^{1/2} .$$

This proves that, for any $\varepsilon > 0$,

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}\big(d(\mu_{\mathbb{G}_N}, \mu_{\mathbb{G}_N^{(m)}}) > \varepsilon\big) = 0 . \tag{57}$$

Taking into account (56) and (57), Fact 4.3 in Chakrabarty *et al.* (2016) and the fact that the space of probability measures on $\mathbb{R}$ is a complete metric space when equipped with the Lévy distance, (55) follows. $\diamondsuit$

6 Useful technical lemmas

Below we give a Taylor expansion of a more convenient type for using Lindeberg's method:

Lemma 12. *Let $f(\cdot)$ be a function from $\mathbb{R}^\ell$ to $\mathbb{C}$, three times differentiable, with continuous third partial derivatives and such that*

$$|\partial_i\partial_j f(\mathbf{x})| \le L_2 \text{ and } |\partial_i\partial_j\partial_k f(\mathbf{x})| \le L_3 \text{ for all } i,j,k \in \{1,\dots,\ell\} \text{ and } \mathbf{x}\in\mathbb{R}^\ell .$$

Then, for any $\mathbf{a} = (a_1,\dots,a_\ell)$ and $\mathbf{b} = (b_1,\dots,b_\ell)$ in $\mathbb{R}^\ell$,

$$f(\mathbf{a}) - f(\mathbf{b}) = \sum_{k=1}^{2}\frac{1}{k!}[(\sum_{j=1}^{\ell} a_j\partial_j)^k - (\sum_{j=1}^{\ell} b_j\partial_j)^k]f(0,\dots,0) + R_3$$

where $|R_3| \le R(\mathbf{a}) + R(\mathbf{b})$, with

$$R(\mathbf{c}) \le 4\ell L_2\sum_{j=1}^{\ell} c_j^2 I(|a_j| > A) + 2AL_3\ell^2\Big(\sum_{j=1}^{\ell} c_j^2\Big),$$

where $\mathbf{c}$ equals $\mathbf{a}$ or $\mathbf{b}$.

This Lemma can be applied in conjunction with Stieltjes transform. Let $A(\mathbf{x})$ be the matrix defined by

$$(A(\mathbf{x}))_{ij} = \begin{cases} \frac{1}{\sqrt{n}}x_{ij} & i \ge j \\ \frac{1}{\sqrt{n}}x_{ji} & i < j \end{cases} \tag{58}$$

Let $z \in \mathbb{C}^+$ and $s := s_z$ be the function defined from $\mathbb{R}^N$ to $\mathbb{C}$ by

$$s(\mathbf{x}) = \frac{1}{n}\mathrm{Tr}(A(\mathbf{x}) - z\mathbf{I}_n)^{-1} , \tag{59}$$

where $\mathbf{I}_n$ is the identity matrix of order n.

The function s, as defined above, admits partial derivatives of all orders. Next we give a lemma concerning the derivatives of $s(\mathbf{x})$ which is easily obtained by using the computations in Chatterjee (2006) (see the proof of Lemma 12 in Merlevède and Peligrad (2016) for a complete proof of its last inequality).

Lemma 13. *Let $z = u + iv \in \mathbb{C}^+$ and let $(a_{ij})_{1\le j\le i\le n}$ and $(b_{ij})_{1\le j\le i\le n}$ be real numbers. There exist universal positive constants c_1, c_2 and c_3 depending only on the imaginary part of z such that*

$$|\partial_{\mathbf{u}} s| \le \frac{c_1}{n^{3/2}}\,,\ |\partial_{\mathbf{u}}\partial_{\mathbf{v}} s| \le \frac{c_2}{n^2} \ \text{ and } \ |\partial_{\mathbf{u}}\partial_{\mathbf{v}}\partial_{\mathbf{w}} s| \le \frac{c_3}{n^{5/2}}\,. \tag{60}$$

Furthermore there exists an universal positive constant c_4 depending only on the imaginary part of z such that for any subset $\mathcal{I}_n$ of $\{(i,j)\}_{1\le j\le i\le n}$ and any $\mathbf{x}$,

$$\Big|\sum_{\mathbf{u}\in\mathcal{I}_n}\sum_{\mathbf{v}\in\mathcal{I}_n} a_{\mathbf{u}} b_{\mathbf{v}} \partial_{\mathbf{u}}\partial_{\mathbf{v}} s_n(\mathbf{x})\Big| \le \frac{c_4}{n^2}\Big(\sum_{\mathbf{u}\in\mathcal{I}_n} a_{\mathbf{u}}^2 \sum_{\mathbf{v}\in\mathcal{I}_n} b_{\mathbf{v}}^2\Big)^{1/2}.$$

The following lemma is Lemma 2.1 in Götze *et al.* (2012).

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

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

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

Acknowledgements. This paper was supported in part by a Charles Phelps Taft Memorial Fund grant, and the NSF grant DMS-1811373. We are grateful to the referee for carefully reading our manuscript and for helpful comments that improved its presentation.

References

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

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

[3] Banna, M., Merlevède, F. and Peligrad, M. (2015). On the limiting spectral distribution for a large class of symmetric random matrices with correlated entries. *Stochastic Process. Appl.* **125**, no. 7, 2700–2726.

[4] Basu, A.K. and Dorea, C.C.Y. (1979). On functional central limit theorem for stationary martingale random fields. *Acta Math. Acad. Sci. Hungar.* **33** 307-316.

[5] Bernstein, S. (1927). Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Math. Ann.* **97** , no. 1, 1–59.

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

[7] 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.

[8] Bradley R. (2007) *Introduction to Strong Mixing Conditions*, Vol 1,2,3 Kendrick Press.

[9] Bobkov, S.G., Götze, F. and Tikhomirov, A.N. (2010). On concentration of empirical measures and convergence to the semi-circle law. *J. Theor. Probab.* **23** 792–823.

[10] Chakrabarty, A., Hazra, R.S. and Sarkar, D. (2016). From random matrices to long range dependence. *Random Matrices Theory Appl.* **5** no. 2, 52 pp.

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

[12] Davis, R., Pfaffel, O. and Stelzer, R. (2014). Limit theory for the largest eigenvalues of sample covariance matrices with heavy-tails. *Stochastic Process. Appl.* **124**, no. 1, 18–50.

[13] Dedecker, J. (1998). A central limit theorem for stationary random fields. *Probab. Theory Related Fields* **1**10, 397–426.

[14] Dedecker, J. (2001). Exponential inequalities and functional central limit theorem for random fields. *ESAIM Probability and Statistics* **5**, 77–104.

[15] Dobrushin, R.L. and Shlosman, S.B. (1985). Constructive criterion for the uniqueness of Gibbs field. *Statistical physics and dynamical systems*, (Köszeg, 1984), 347–370, *Progr. Phys.*, **10**, Birkhäuser Boston, Boston, MA.

[16] Dombry, C. and Eyi-Minko, F. (2012). Strong mixing properties of max-infinitely divisible random fields. *Stochastic Process. Appl.* **122**, no. 11, 3790–3811.

[17] Doukhan, P. (1984) Mixing: Properties and Examples. Springer-Verlag.

[18] El Machkouri, M., Volný, D. and Wu, W.B. (2013). A central limit theorem for stationary random fields. *Stochastic Process. Appl.* **123**, no. 1, 1–14.

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

[20] Girko, V.L. (1995). *Statistical analysis of observations of increasing dimension.* Vol. **28.** Springer Science & Business Media.

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

[22] 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.

[23] 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.

[24] Laroche, E. (1995). Hypercontractivité pour des systèmes de spins de portée infinie. *Probab. Theory Relat. Fields.* **101**, 89–132.

[25] Marchenko, V. and Pastur, L. (1967). Distribution of eigenvalues for some sets of random matrices. *Mat. Sb.* **72**, 507-536.

[26] Merlevède, F., Peligrad, C. and Peligrad, M. (2015). On the universality of spectral limit for random matrices with martingale differences entries. *Random Matrices Theory Appl.* **4**, no. 1, 33 pp.

[27] Merlevède, F. and Peligrad, M. (2006). On the weak invariance principle for stationary sequences under projective criteria. *J. Theor. Probab.* **19**, no. 3, 647–689,

[28] Merlevède, F. and Peligrad, M. (2016). On the empirical spectral distribution for matrices with long memory and independent rows. *Stochastic Process. Appl.* **126**, no. 9, 2734–2760.

[29] Merlevède, F., Peligrad, M. and Utev, S. (2019). *Functional Gaussian Approximation for Dependent Structures.* Oxford Studies in Probability. **6**, Oxford University Press.

[30] Pajor, A. and Pastur, L. (2009). On the limiting empirical measure of eigenvalues of the sum of rank one matrices with log-concave distribution. *Studia Math.* **195**, no. 1, 11–29.

[31] Pan, G., Gao, J. and Yang, Y. (2014). Testing independence among a large number of high-dimensional random vectors. *J. Amer. Statist. Assoc.* **109**, 600–612.

[32] Pastur, L. (1972). On the spectrum of random matrices. *Teor. Mat. Fiz.* **10**, 102–112.

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

[34] Peligrad, C. and Peligrad, M. (2016). The limiting spectral distribution in terms of spectral density. *Random Matrices Theory Appl.* **5**, no 1, 19pp.

[35] Pfaffel, O. and Schlemm, E. (2011). Eigenvalue distribution of large sample covariance matrices of linear processes. *Probab. Math. Statist.* **31**, 313–329

[36] 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.

[37] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition, *Proc. Nat. Acad. Sci. U. S. A.* **42**, 43–47.

[38] Rosenblatt, M. (2009). A comment on a conjecture of N. Wiener. *Stat. and Prob. Letters.* **79**, 347–348

[39] Rio, E. (2017). *Asymptotic theory of weakly dependent random processes.* Probability Theory and Stochastic Modelling, **80**. Springer.

[40] Sasvári, Z. (2013). *Multivariate characteristic and correlation functions.* De Gruyter Studies in Mathematics, **50**. Walter de Gruyter & Co., Berlin, x+366 pp.

[41] Silverstein, J. W. (1995). Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. (English summary) *J. Multivariate Anal.* **55**, 331-339.

[42] Silverstein, J. W. and Bai. Z. D. (1995). On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *J. Multivariate Anal.* **54**, 175–192.

[43] Tao, T. (2012). *Topics in Random Matrix Theory*, Graduate Studies in Mathematics **132**, Amer. Math.

[44] Tone, C. (2011). Central limit theorems for Hilbert-space valued random fields satisfying a strong mixing condition. *ALEA Lat. Am. J. Probab. Math. Stat.* **8**, 77–94

[45] Wigner, E.P. (1958). On the distribution of the roots of certain symmetric matrices. *Ann. of Math.* **67**, 325-327.

[46] Yao, J. (2012). A note on a Mar Ì†cenko-Pastur type theorem for time series. *Statist. Probab. Lett.* **82**, 22–28.

[47] Yaskov, P. (2016-a). A short proof of the Marchenko-Pastur theorem. *C. R. Math. Acad. Sci. Paris.* **354**, no. 3, 319–322.

[48] Yaskov, P. (2016-b). Necessary and sufficient conditions for the Marchenko-Pastur theorem. *Electron. Commun. Probab.* **21**, no. 73, 1–8.

[49] Yin, Y. Q. (1986). Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.* **20**, 50–68.

[50] Yin, Y. Q. and Krishnaiah, P. R. (1983). A limit theorem for the eigenvalues of product of two random matrices. *J. Multivariate Anal.* **13**, no. 4, 489–507.