



Eigenvectors of Sample Covariance Matrices : Universality of global fluctuations

Ali Bouferroum

► To cite this version:

Ali Bouferroum. Eigenvectors of Sample Covariance Matrices : Universality of global fluctuations. 2013. hal-00835504

HAL Id: hal-00835504

<https://hal.science/hal-00835504>

Preprint submitted on 18 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EIGENVECTORS OF SAMPLE COVARIANCE MATRICES: UNIVERSALITY OF GLOBAL FLUCTUATIONS

ALI BOUFERROUM

Université Paris-Est

ABSTRACT. In this paper, we prove a universality result of convergence for a bivariate random process defined by the eigenvectors of a sample covariance matrix. Let $V_n = (v_{ij})_{i \leq n, j \leq m}$ be a $n \times m$ random matrix, where $(n/m) \rightarrow y > 0$ as $n \rightarrow \infty$, and let $X_n = (1/m)V_n V_n^*$ be the sample covariance matrix associated to V_n . Consider the spectral decomposition of X_n given by $U_n D_n U_n^*$, where $U_n = (u_{ij})_{n \times n}$ is an eigenmatrix of X_n . We prove, under some moments conditions, that the bivariate random process

$$\left(B_{s,t}^n = \sum_{\substack{1 \leq i \leq \lfloor ns \rfloor \\ 1 \leq j \leq \lfloor nt \rfloor}} \left(|u_{i,j}|^2 - \frac{1}{n} \right) \right)_{(s,t) \in [0,1]^2}$$

converges in distribution to a bivariate Brownian bridge. This type of result has been already proved for Wishart matrices (LOE/LUE) and Wigner matrices. This supports the intuition that the eigenmatrix of a sample covariance matrix is in a way "asymptotically Haar distributed". Our analysis follows closely the one of Benaych-Georges for Wigner matrices, itself inspired by Silverstein works on the eigenvectors of sample covariance matrices.

KEYWORDS AND PHRASES. Random matrices; Sample covariance matrices; Haar measure; Eigenvectors; Delocalization; Brownian bridge; Spectral decomposition; Method of moments.

1. INTRODUCTION

The eigenvalues of random matrices attracted considerable attention in the recent years [1, 6, 28]. Less is known for the eigenvectors. Therefore, recent research on the limiting behavior of eigenvectors has attracted considerable interest among mathematicians and statisticians, see among others, Silverstein [23, 24, 25], Bai-Pan [3], Bai-Miao-Pan [2], Ledoit-Péché [19], Benaych-Georges [7], Pillai-Yin [20]. The recent progress on the study of eigenvectors refers to a delocalization property shown for the eigenvectors of some types of random matrices, see Erdős-Schlein-Yau [15, 14], Bordenave-Guionnet [11], Schenker [22], Cacciapuoti-Maltsev-Schlein [12], Rudelson-Vershynin [21] and Vu-Wang [29]. For Wigner matrices, a universal properties of eigenvector coefficients were given recently, see Knowles-Yin [18] and Tao-Vu [27].

In practical applications, the eigenvectors of large random matrices play a role as important as that played by the eigenvalues. For example, in multivariate analysis, the Principal Component Analysis is based on eigenvectors of sample covariance matrices. The directions of the principal

Date: Compiled June 18, 2013.

2010 Mathematics Subject Classification. 15B52; 60F05.

components are of particular interest, however, the exact distribution of the eigenmatrix (matrix of eigenvectors) of this type of matrices cannot be computed and few works had been devoted to this subject until quite recently. One of the reasons is that while the eigenvalues of an Hermitian matrix admit variational characterizations as extrema of certain functions, the eigenvectors can be characterized as the argmax of these functions, hence are more sensitive to perturbations of the entries of the matrix.

Recently, it was proved in [17] that the entries of the first $o(n/\log n)$ columns of a Haar distributed matrix can be approximated simultaneously by independent standard normals. Based on these evidentiary support and motivated by the fact that the eigenmatrix of Wishart matrix is Haar (uniformly) distributed, we believe that the eigenmatrix of a sample covariance matrix X_n is "asymptotically Haar distributed" over the unitary group $\mathcal{U}(n)$ of $n \times n$ unitary matrices for the complex case; or over the orthogonal group $\mathcal{O}(n)$ of $n \times n$ orthogonal matrices for the real case. A question asked here is how to formulate the wording of "asymptotically Haar distributed"? Silverstein discussed this terminology in details in [23].

Let $V_n = (v_{ij})_{n \times m}$, $i = 1, \dots, n$; $j = 1, \dots, m = m(n)$ where $(n/m) \rightarrow y > 0$ as $n \rightarrow \infty$, be an observation matrix of i.i.d. real or complex random variables $\{v_{ij}\}_{i,j=1,2,\dots}$ such that

$$(1) \quad \mathbb{E}(v_{11}) = 0, \quad \mathbb{E}(|v_{11}|^2) = 1,$$

and $V_j = (v_{1j}, \dots, v_{nj})'$ be the j^{th} column of V_n . In this paper, we will consider a simplified version of sample covariance matrices with large dimension n and sample size m

$$X_n = (1/m)V_n V_n^*,$$

where V_n^* denotes the conjugate transpose of the matrix V_n . Let us define the cumulative distribution of the eigenvalues of X_n , for each $u \in \mathbb{R}$, as

$$F_n^X(u) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{\lambda_i \leq u\}},$$

this function describes the global behavior of the spectrum of X_n . Recall that for a matrix X_n defined as above, the previous empirical cumulative function F_n^X converges almost surely for every $u \geq 0$, as $n \rightarrow \infty$, to a non-random distribution function F_y^{MP} which has the Marchenko-Pastur density

$$f_y^{MP}(u) = \left(1 - \frac{1}{y}\right)_+ \delta_0 + \frac{\sqrt{(b-u)(u-a)}}{2\pi y u} \mathbf{1}_{[a,b]}(u),$$

where $a = (1 - \sqrt{y})^2$ and $b = (1 + \sqrt{y})^2$ (atom $1 - 1/y$ at the origin if and only if $y > 1$), see [30] and [6, Theorems 3.6, 3.7].

Let $U_n D_n U_n^*$ denote the spectral decomposition of the sample covariance matrix X_n , where $D_n = \text{diag}(\lambda_1, \dots, \lambda_n)$, and the λ_i 's are the eigenvalues of X_n arranged along the diagonal of D_n in non-decreasing order, and $U_n = \{u_{ij}\}$ is the associated eigenmatrix for X_n . Let us define a bivariate random process $B_{s,t}^n$ by

$$(2) \quad \left(B_{s,t}^n = \sqrt{\frac{\beta}{2}} \sum_{\substack{1 \leq i \leq \lfloor ns \rfloor \\ 1 \leq j \leq \lfloor nt \rfloor}} \left(|u_{i,j}|^2 - \frac{1}{n} \right) \right)_{(s,t) \in [0,1]^2}$$

where $\beta = 2$ in the complex case and $\beta = 1$ in the real case, and $\lfloor a \rfloor$ denotes the greatest integer less than or equal to a . It is well known [13] that if U_n is Haar distributed over the group $\mathcal{U}(n)$ or the group $\mathcal{O}(n)$, then $B_{s,t}^n$ weakly converges (i.e. in the sense of convergence of all finite-dimensional marginals) to a Brownian bridge $(B_{s,t})$ as n tends to infinity, i.e: the centered continuous Gaussian process $(B_{s,t})_{(s,t) \in [0,1]^2}$ with covariance

$$(3) \quad \mathbb{E} (B_{s,t} B_{s',t'}) = (\min\{s, s'\} - ss') (\min\{t, t'\} - tt').$$

Conversely, if $B_{s,t}^n$ weakly converges to a Brownian bridge, it then reveals some evidence supporting the conjecture that the eigenmatrix U_n is asymptotically Haar distributed.

In this paper, we will prove that for a sample covariance matrix X_n defined as above, $B_{s,t}^n$ has a limit in a weaker sense if v_{11} has moments of all orders, and that this weak limit is the bivariate Brownian bridge if and only if v_{11} has the same fourth moment as in the case of LOE/LUE matrix (Wishart-Laguerre orthogonal/unitary ensembles). This work is inspired by the work of Benaych-Georges for Wigner matrices [7], itself is inspired by Silverstein's works in [24, 25] for a univariate process defined by the eigenmatrix of a sample covariance matrix as

$$Y_t^n = \sqrt{\frac{n}{2}} \sum_{1 \leq i \leq \lfloor nt \rfloor} \left(|y_i|^2 - \frac{1}{n} \right), \quad \text{where } y = X_n x_n \text{ for some vector } x_n.$$

The rest of the paper is organized as follows. The main theorem is presented in section 2 with some remarks. The proof of this theorem is mainly contained in Sections 3 and 4: Essentially, the problem is transformed into showing convergence on an appropriate space $D[0,1] \times D[0,+\infty[$ instead of the space $D[0,1]^2$. After that, the proof will consist of studying the moments of a weighted spectral law of X_n according to the process B_n . In section 5, we finish by giving a version of tightness and convergence in the Skorokhod topology of the process $B_{s,t}^n$ under some additional hypotheses on the atom distribution.

Acknowledgment: I would like to express my sincere appreciation and gratitude to my advisor Djamil Chafaï for his academic guidance and enthusiastic encouragement. I also would like to thank Florent Benaych-Georges for pointing out some references and for his encouragement.

2. MAIN RESULT

Let us consider a matrix $V_n = (v_{ij}^{(n)})_{n \times m}$, $i = 1, \dots, n$; $j = 1, \dots, m = m(n)$ where $(n/m) \rightarrow y > 0$ as n tends to infinity. Let $X_n = (1/m)V_n V_n^*$ be its associated sample covariance matrix of dimension n and sample size m . Let $U_n D_n U_n^*$ denote the spectral decomposition of the sample covariance matrix X_n , where $D_n = \text{diag}(\lambda_1, \dots, \lambda_n)$, and λ_i 's are the eigenvalues of X_n arranged along the diagonal of D_n with a non-decreasing order, and $U_n = \{u_{ij}\}_{1 \leq i,j \leq n}$ is the associated

eigenmatrix of X_n . Note that U_n is not uniquely defined, however, one can choose it in any measurable way. We consider the bivariate càd-làg process $(B_{s,t}^n)_{(s,t) \in [0,1]^2}$ defined as:

$$\left(B_{s,t}^n = \sqrt{\frac{\beta}{2}} \sum_{\substack{1 \leq i \leq \lfloor ns \rfloor \\ 1 \leq j \leq \lfloor nt \rfloor}} \left(|u_{i,j}|^2 - \frac{1}{n} \right) \right)_{(s,t) \in [0,1]^2},$$

where $\beta = 2$ in the complex case and $\beta = 1$ in the real case.

Theorem 2.1 (Main result). *Suppose in the definition above of the sample covariance matrix $X_n = (1/m)V_n V_n^*$ that*

$$(4) \quad \{v_{ij}^{(n)}\}_{i,j=1,2,\dots} \text{ are i.i.d. complex or real random variables,}$$

with

$$(5) \quad \mathbb{E}(v_{11}^{(n)}) = 0, \quad \mathbb{E}|v_{11}^{(n)}|^2 = 1,$$

and

$$(6) \quad \forall k \geq 0, \quad \sup_n \mathbb{E}|v_{1,1}^{(n)}|^k < \infty,$$

Then the sequence

$$(\text{distribution}(B^n))_{n \geq 1}$$

has a unique possible accumulation point supported by $C([0,1]^2)$ in the sense of convergence of all finite-dimensional marginals. This accumulation point is the distribution of a centered Gaussian process which depends on the distribution of $v_{11}^{(n)}$ only through $\lim_{n \rightarrow \infty} \mathbb{E}(|v_{11}^{(n)}|^4)$, and which is the bivariate Brownian bridge if and only if

$$(7) \quad \lim_{n \rightarrow \infty} \mathbb{E}(|v_{11}^{(n)}|^4) = 4 - \beta.$$

Remark 2.2 (Dependence of entries on n). *The distribution of the entries $v_{i,j} = v_{i,j}^{(n)}$ are allowed to depend on n . For brevity of notations, we write $v_{i,j}$ instead of $v_{i,j}^{(n)}$.*

Remark 2.3 (Matching with LOE/LUE). *Note that the unique possible accumulation point supported by $C([0,1]^2)$ of our sequence $(\text{distribution}(B^n))_{n \geq 1}$ which is a centered Gaussian process depends on the distribution of the $v_{i,j}$'s only through $\lim_{n \rightarrow \infty} \mathbb{E}(|v_{11}|^4)$, and this limiting distribution is the bivariate Brownian bridge if and only if $\lim_{n \rightarrow \infty} \mathbb{E}(|v_{11}|^4)$ is the same as for a LOE or LUE matrix, i.e. equal to $4 - \beta$.*

SIMULATION

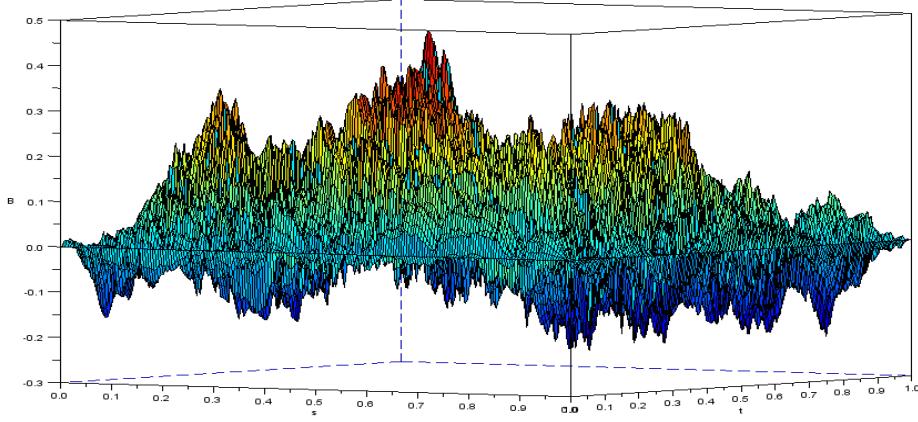


FIGURE 1. Simulation of the bivariate random process $B^n(s, t)$ for a sample covariance matrix with standard real normal atom distribution (Wishart matrix). the matrix V_n is $n \times m$ with $n = m = 500$.

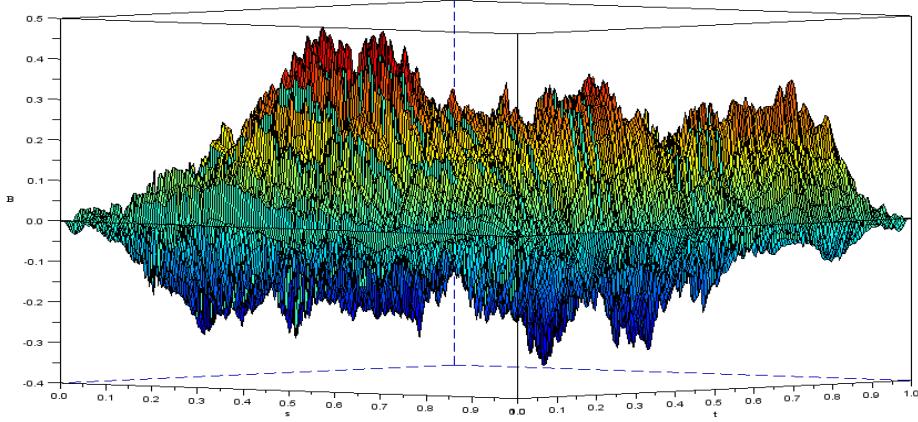


FIGURE 2. Simulation of the bivariate random process $B^n(s, t)$ for a sample covariance matrix with atom distribution: $v_{1,1} = CB - \mathbb{E}(CB)$ with $CB \stackrel{d}{=} \text{Binomial}(6, (1/2 - \sqrt{1/12}))$. The matrix V_n is $n \times m$ with $n = m = 500$.

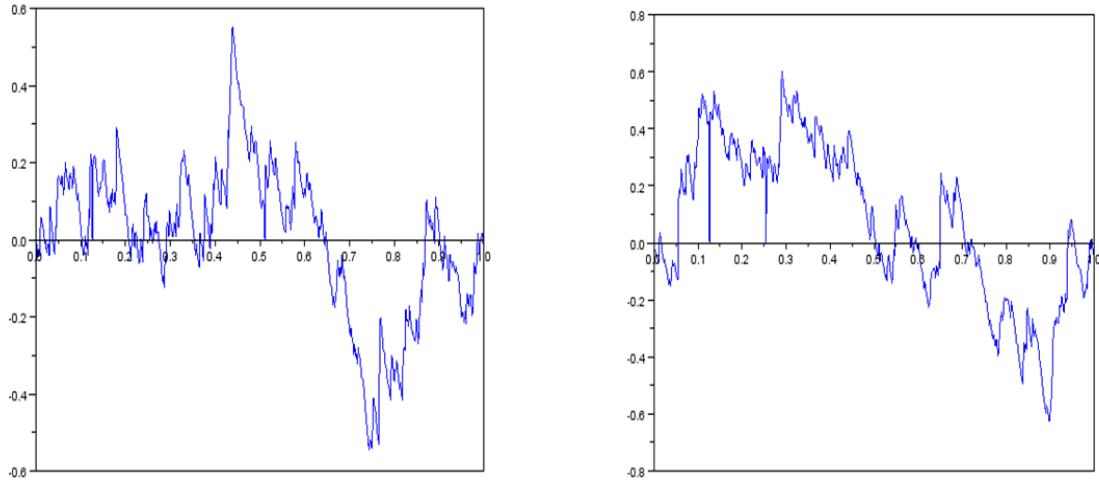


FIGURE 3. Simulation of the univariate random process $B^n(s, 1)$ for two different choices of atom distributions: **Left:** Wishart matrix. **Right:** Sample covariance matrix with atom distribution $v_{1,1} = CB - \mathbb{E}(CB)$ with $CB \stackrel{d}{=} \text{Binomial}(6, (1/2 - \sqrt{1/12}))$. For both pictures, the matrix V_n is $n \times m$ with $n = m = 500$.

3. OUTLINE OF THE PROOF OF THEOREM 2.1

In this paper we denote:

- $C([0, 1]^2)$ (resp. $C_c([0, 1] \times [0, +\infty))$) the space of real valued continuous functions on $[0, 1]^2$ (resp. of real valued compactly supported continuous functions on $[0, 1] \times [0, +\infty]$), endowed with the uniform convergence topology.
- $D_c(\mathbb{R}, [0, 1])$ the set of compactly supported càd-làg functions on \mathbb{R} taking values in $[0, 1]$, endowed with the topology defined by the fact that $f_n \rightarrow f$ if and only if the bounds of the support of f_n tend to those of the support of f and for all $M > 0$, after restriction to $[-M, M]$, $f_n \rightarrow f$ with the topology of $D[-M, M]$ being deduced from the Skorokhod topology of $D([0, 1])$ defined in [10, Chapter 3].
- $D([0, 1]^2)$ (resp. $D_c([0, 1] \times [0, +\infty])$) the space of real valued functions $f : [0, 1]^2 \rightarrow \mathbb{R}$ (resp. of compactly supported functions $f : [0, 1] \times [0, +\infty] \rightarrow \mathbb{R}$) admitting limits in all "orthants", more precisely such that for each s_0, t_0 , for each pair of symbols $\circ, \bullet \in \{<, \geq\}$,

$$\lim_{\substack{s \rightarrow \circ s_0 \\ t \rightarrow \bullet t_0}} f(s, t)$$

exists, and is equal to $f(s_0, t_0)$ if both \circ and \bullet are \geq . The space $D([0, 1]^2)$ is endowed with the Skorokhod topology defined in [8] and the space $D_c([0, 1] \times [0, +\infty])$ is endowed with the topology defined by: $f_n \rightarrow f$ if and only if for all $M > 0$, after restriction to $[0, 1] \times [0, M]$, $f_n \rightarrow f$ in the sense of the space $D([0, 1]^2)$.

- $D_0([0, 1]^2)$ the set of functions in $D([0, 1]^2)$ vanishing at the border of $[0, 1]^2$, endowed with the induced topology.

3.1. From $D([0, 1]^2)$ to $D([0, 1] \times [0, +\infty[)$. As we have seen in the introduction, the cumulative distribution function F_n^X of X_n converges almost surely, as $n \rightarrow \infty$, to a non-random distribution function F_y^{MP} (Marchenko-Pastur law). The proof of Theorem 2.1 can be reduced to the following remark, inspired by some ideas of Silverstein [24, 25] and of Benaych-Georges [7]: even though we do not have any "direct access" to the eigenvectors of X_n , we have access to the process $(B_{s, F_n^X(u)}^n)_{s \in [0, 1], u \in [0, +\infty[}$, for $F_n^X(u) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{\lambda_i \leq u\}}$. Indeed,

$$B_{s, F_n^X(u)}^n = \sqrt{\frac{\beta}{2}} \sum_{1 \leq i \leq \lfloor ns \rfloor} \sum_{1 \leq j \leq n: \lambda_j \leq u} \left(|u_{i,j}|^2 - \frac{1}{n} \right),$$

hence, for all fixed $s \in [0, 1]$, the function $u \in \mathbb{R} \mapsto B_{s, F_n^X(u)}^n$ is the cumulative distribution function of the signed measure

$$(8) \quad \sqrt{\frac{\beta}{2}} \sum_{1 \leq i \leq \lfloor ns \rfloor} \sum_{1 \leq j \leq n} \left(|u_{i,j}|^2 - \frac{1}{n} \right) \delta_{\lambda_j},$$

which can be considered as a difference between two random probability measures: $\sum_{1 \leq j \leq n} |u_{i,j}|^2 \delta_{\lambda_j}$ (weighted spectral law of X_n) and $\frac{1}{n} \sum_{1 \leq j \leq n} \delta_{\lambda_j}$ (empirical spectral law of X_n). The law (8) can be studied via its moments

$$\sum_{1 \leq i \leq \lfloor ns \rfloor} \left(e_i^* X_n^k e_i - \frac{1}{n} \text{Tr} X_n^k \right)$$

for $k \geq 1$, the e_i 's being the vectors of the canonical basis. From the asymptotic behavior of the moments of the signed measure (8), one can then find out the asymptotic behavior of its cumulative distribution function.

Once the asymptotic distribution of the process $(B_{s, F_n^X(u)}^n)_{s \in [0, 1], u \in [0, +\infty[}$ is identified, one can obtain the asymptotic distribution of the process $(B_{s,t}^n)_{(s,t) \in [0,1]^2}$, because

$$F_n^X(u) \xrightarrow[n \rightarrow \infty]{} F_y^{MP}(u), \quad \text{almost surely for every } u \geq 0.$$

The following proposition is the key of the proof, since it allows transferring our problem from the eigenvectors to some more accessible objects: the weighted spectral distributions of the sample covariance matrix X_n .

Proposition 3.1 (From the process B_n to a weighted spectral process). *To prove Theorem 2.1, it suffices to prove that each finite-dimensional marginal distribution of the process*

$$\left(\sum_{1 \leq i \leq \lfloor ns \rfloor} \left(e_i^* X_n^k e_i - \frac{1}{n} \text{Tr} X_n^k \right) \right)_{s \in [0,1], k \geq 1}$$

converges to a centered Gaussian process and that the covariance of the limiting process depends on the distribution of the $v_{i,j}$'s only through $\lim_{n \rightarrow \infty} \mathbb{E}(|v_{11}|^4)$, and that this covariance is the one of the bivariate Brownian bridge if and only if $\lim_{n \rightarrow \infty} \mathbb{E}(|v_{11}|^4) = 4 - \beta$.

Proof. It is known [30, 6] that the cumulative distribution function $F_n^X(u) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{\lambda_i \leq u\}}$ of the matrix X_n converges almost surely, as n tends to infinity, to a non-random distribution function F_y^{MP} defined by means of the Marchenko-Pastur law. Since the limit is continuous and compactly supported on $(0, +\infty)$, this convergence is uniform

$$\sup_{u \in [0, +\infty[} |F_n^X(u) - F_y^{MP}(u)| \xrightarrow{n \rightarrow \infty} 0 \quad \text{almost surely.}$$

Hence, it follows that

$$F_n^X \xrightarrow{n \rightarrow \infty} F_y^{MP} \quad \text{in } D_c([0, +\infty), [0, 1]),$$

see [6, section 10.1.2]. Moreover, the map

$$D_0([0, 1]^2) \times D_c([0, +\infty), [0, 1]) \rightarrow D_c([0, 1] \times [0, +\infty))$$

$$((G_{s,t})_{s,t \in [0,1]}, (g(u))_{u \in [0,+\infty)}) \mapsto (G_{s,g(u)})_{(s,u) \in [0,1] \times [0,+\infty)}$$

is continuous at any pair of continuous functions. Hence for any continuous process $(B_{s,t})_{s,t \in [0,1]}$ whose distribution is an accumulation point of the sequence $(\text{distribution}(B^n))_{n \geq 1}$ for the Skorokhod topology in $D([0, 1]^2)$, the process

$$(B_{s,F_n^X(u)}^n)_{s,u \in [0,1] \times [0,+\infty)}$$

converges in distribution (up to the extraction of a subsequence) to the process

$$(B_{s,F_y^{MP}(u)})_{s,u \in [0,1] \times [0,+\infty)}.$$

This assertion relies on two results which can be found in [10, Theorem 4.4 and Corollary 1 of Theorem 5.1 in Chapter 1]. Now, note that $F_y^{MP} : [0, +\infty) \rightarrow [0, 1]$ admits a right inverse, so the distribution of the process $(B_{s,t})_{s,t \in [0,1]}$ is entirely determined by that of the process $(B_{s,F_y^{MP}(u)})_{s,u \in [0,1] \times [0,+\infty)}$. Therefore, to prove Theorem 2.1, it suffices to prove that the sequence

$$(9) \quad (\text{distribution}(B_{s,F_n^X(u)}^n)_{s,u \in [0,1] \times [0,+\infty)})_{n \geq 1}$$

converges to a centered Gaussian process which depends on the distribution of the $v_{i,j}$'s only through $\lim_{n \rightarrow \infty} \mathbb{E}(|v_{11}|^4)$, and which is the bivariate Brownian bridge if and only if

$$\lim_{n \rightarrow \infty} \mathbb{E}(|v_{11}|^4) = 4 - \beta.$$

Now, let us prove that any (random) function $f \in C_c([0, 1] \times [0, +\infty))$ is entirely determined by the collection of real numbers $(\int_{u \in [0,+\infty)} u^k f(s, u) du)_{s \in [0,1], k \geq 0}$.

Lemma 3.2 (Technical characterization). *Let f be a random variable in $C_c([0, 1] \times [0, +\infty))$ such that with probability one, $f(s, u) = 0$ when $u > r, r > 0$. Then the distribution of f is entirely determined by the finite-dimensional marginals of the process*

$$(10) \quad \left(\int_{u \in [0,+\infty)} u^k f(s, u) du \right)_{s \in [0,1], k \geq 0}.$$

Moreover, in the case where the finite-dimensional marginals of the process of (10) are Gaussian and centered, then so are those of f .

Proof. Let us fix $(s, u_0) \in [0, 1] \times [0, r]$ and let, for each $p \geq 1$, $(P_{p,q})_{q \geq 1}$ be a sequence of polynomials that is uniformly bounded on $[0, r+1]$ and that converges pointwise to $1_{[u_0, u_0+1/p]}$ on $[0; r+1]$. Then one has, with probability one,

$$f(s, u_0) = \lim_{p \rightarrow \infty} p \int_{u_0}^{u_0+1/p} f(s, u) du = \lim_{p \rightarrow \infty} \lim_{q \rightarrow \infty} p \int_{u \in [0, +\infty)} P_{p,q}(u) f(s, u) du.$$

This proves the lemma, because any almost sure limit of a sequence of variables belonging to a space of centered Gaussian variables is Gaussian and centered. \square

Since the fourth moment of the entries of V_n is finite, we know that the largest eigenvalue of the sample covariance matrix X_n converges, almost surely, to $b = (1 + \sqrt{y})^2$ (see [4, 5]). Hence, for any random variable f taking values in $C_c([0, 1] \times [0, +\infty))$ such that the distribution of f is a limiting point of the sequence of (9), we have $f(s, u) = 0$, almost surely, when $u > b + \epsilon$.

As a consequence, it follows from the previous lemma and from what precedes that in order to prove Theorem 2.1, it suffices to prove that each finite-dimensional marginal distribution of the process

$$\int_{u \in \mathbb{R}} u^k B_{s, F_n^X(u)}^n du$$

converges to a centered Gaussian measure and that the covariance of the limit process depends on the distribution of the $v_{i,j}$'s only through $\lim_{n \rightarrow \infty} \mathbb{E}(|v_{11}|^4)$, and that this covariance is the one of the bivariate Brownian bridge if and only if $\lim_{n \rightarrow \infty} \mathbb{E}(|v_{11}|^4) = 4 - \beta$.

Recall that :

$$\begin{aligned} B_{s, F_n^X(u)}^n &= \sqrt{\frac{\beta}{2}} \sum_{1 \leq i \leq \lfloor ns \rfloor} \sum_{1 \leq j \leq n: \lambda_j \leq u} \left(|u_{i,j}|^2 - \frac{1}{n} \right) \\ &= \sqrt{\frac{\beta}{2}} \sum_{1 \leq i \leq \lfloor ns \rfloor} F_{\mu_{X_n, e_i} - \mu_{X_n}}, \end{aligned}$$

where

- μ_{X_n} is the empirical spectral law of X_n .
- μ_{X_n, e_i} is the weighted spectral law of X_n , defined by $\mu_{X_n, e_i} = \sum_{j=1}^n |u_{i,j}|^2 \delta_{\lambda_j}$.
- $F_{\mu_{X_n, e_i} - \mu_{X_n}}$ is the cumulative distribution function of the null-mass signed measure $\mu_{X_n, e_i} - \mu_{X_n}$.

Now, let us give the following lemma to complete the proof of Proposition 3.1.

Lemma 3.3 (Moment's calculation rule). *Let μ be a compactly supported null-mass signed measure and set $F_\mu(u) = \mu((-\infty, u])$. Then for all $k \geq 0$,*

$$\int_{u \in \mathbb{R}} u^k F_\mu(u) du = - \int_{x \in \mathbb{R}} \frac{x^{k+1}}{k+1} d\mu(x).$$

Proof. Let $a < b$ be such that the support of μ is contained in the open interval (a, b) . F_μ is null outside (a, b) and satisfies $F_\mu(u) = -\mu((u, b))$, so by Fubini's Theorem,

$$\begin{aligned} \int_{u \in \mathbb{R}} u^k F_\mu(u) du &= \int_{x=a}^b \int_{u=a}^x -u^k du d\mu(x) \\ &= \int_{x=a}^b \frac{-x^{k+1}}{k+1} d\mu(x) + \frac{a^{k+1}}{k+1} \mu((a, b)) \\ &= \int_{x=a}^b \frac{-x^{k+1}}{k+1} d\mu(x). \end{aligned}$$

□

It follows from this lemma that for all $s \in [0, 1]$, $k \geq 0$,

$$\begin{aligned} \int_{u \in \mathbb{R}} u^k B_{s, F_n^X}^n(u) du &= - \int_{x \in \mathbb{R}} \frac{-x^{k+1}}{k+1} d(\mu_{X_n, e_i} - \mu_{X_n})(x) \\ &= \frac{-1}{k+1} \sqrt{\frac{\beta}{2}} \sum_{1 \leq i \leq \lfloor ns \rfloor} \sum_{1 \leq j \leq n} \left(|u_{i,j}|^2 - \frac{1}{n} \right) \lambda_j^{k+1} \\ &= \frac{-1}{k+1} \sqrt{\frac{\beta}{2}} \sum_{1 \leq i \leq \lfloor ns \rfloor} \left(e_i^* X_n^{k+1} e_i - \frac{1}{n} \text{Tr} X_n^{k+1} \right), \end{aligned}$$

which completes the proof of Proposition 3.1. □

It follows from all what precedes that Theorem 2.1 is a direct consequence of the following proposition, whose proof is in Section 4.

Proposition 3.4 (Convergence of Moments). *Under Assumptions (4), (5) and (6) in Theorem 2.1, each finite-dimensional marginal distribution of the process*

$$\left(\sum_{1 \leq i \leq \lfloor ns \rfloor} \left(e_i^* X_n^k e_i - \frac{1}{n} \text{Tr} X_n^k \right) \right)_{s \in [0, 1], k \geq 1}$$

converges to a centered Gaussian measure. The covariance of the limit distribution, denoted by

$$(\text{Cov}_{s_1, s_2}(k_1, k_2))_{s_1, s_2 \in [0, 1], k_1, k_2 \geq 1},$$

depends on the distribution of the $v_{i,j}$'s only through $\lim_{n \rightarrow \infty} \mathbb{E}[|v_{1,1}|^4]$. Moreover, we have

$$\text{Cov}_{s_1, s_2}(1, 1) = (\lim_{n \rightarrow \infty} \mathbb{E}[|x_{1,2}|^4] - 1)(\min\{s_1, s_2\} - s_1 s_2).$$

4. PROOF OF PROPOSITION 3.4

Note that the expectation of the weighted spectral law $\mu_{X_n, e_i} = \sum_{j=1}^n |u_{i,j}|^2 \delta_{\lambda_j}$ does not depend on i . So for all $s \in [0, 1]$, $k \geq 1$,

$$(11) \quad \sum_{1 \leq i \leq ns} (e_i^* X_n^k e_i - \frac{1}{n} \text{Tr}(X_n^k)) = \sum_{1 \leq i \leq ns} (e_i^* X_n^k e_i - \mathbb{E}[e_i^* X_n^k e_i]) - \frac{\lfloor ns \rfloor}{n} \sum_{1 \leq i \leq n} (e_i^* X_n^k e_i - \mathbb{E}[e_i^* X_n^k e_i]).$$

Therefore, we are led to study the limit, as $n \rightarrow \infty$, of the finite-dimensional marginal distributions of the process

$$(12) \quad \left(\sum_{1 \leq i \leq ns} (e_i^* X_n^k e_i - \mathbb{E}[e_i^* X_n^k e_i]) \right)_{s \in [0,1], k \geq 1}.$$

Let us fix $p \geq 1$, $s_1, \dots, s_p \in [0, 1]$ and $k_1, \dots, k_p \geq 1$. We shall study the limit, as n tends to infinity, of

$$(13) \quad \mathbb{E} \left[\prod_{l=1}^p \sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \right]$$

We introduce the set

$$(14) \quad \mathcal{E} := \{\bar{0}^1, \dots, \bar{k}_1^1\} \cup \dots \cup \{\bar{0}^p, \dots, \bar{k}_p^p\},$$

where the sets $\{\bar{0}^1, \bar{1}^1, \dots\}$, $\{\bar{0}^2, \bar{1}^2, \dots\}$, \dots , $\{\bar{0}^p, \bar{1}^p, \dots\}$ are p disjoint copies of the set of non-negative integers. The set \mathcal{E} is ordered as presented in (14). In the rest of this paper, we denote $(1, \dots, n)^k$ the set of k -tuples of a set $\{1, 2, \dots, n\}$.

The expectation (13) can be expanded and expressed as a sum on the set $(1, \dots, n)^K$ indexed by the set \mathcal{E} introduced above, where $K = k_1 + \dots + k_p$. We get

$$(15) \quad \begin{aligned} & \mathbb{E} \left[\prod_{l=1}^p \sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \right] \\ &= \sum_{\pi \in (1, \dots, n)^K} \mathbb{E} \left[\prod_{l=1}^p \left(x_{\pi(\bar{0}^l), \pi(\bar{1}^l)} \cdots x_{\pi(\bar{k}_l-1^l), \pi(\bar{k}_l^l)} - \mathbb{E}(x_{\pi(\bar{0}^l), \pi(\bar{1}^l)} \cdots x_{\pi(\bar{k}_l-1^l), \pi(\bar{k}_l^l)}) \right) \right] \mathbf{1}_{\{1 \leq \pi(\bar{0}^l) = \pi(\bar{k}_l^l) \leq ns_l\}}, \end{aligned}$$

As we have not sufficient information on the laws of the $x_{i,j}$'s, we need to write the previous expression in terms of elements of the matrix V_n . Let us consider

$$\Pr = x_{\pi(\bar{0}^l), \pi(\bar{1}^l)} x_{\pi(\bar{1}^l), \pi(\bar{2}^l)} \cdots x_{\pi(\bar{k}_l-1^l), \pi(\bar{k}_l^l)}.$$

We have

$$\begin{aligned} \Pr &= \frac{1}{m^{k_l}} \left[\sum_{j=1}^m v_{\pi(\bar{0}^l), j} v_{j, \pi(\bar{1}^l)}^* \right] \left[\sum_{j=1}^m v_{\pi(\bar{1}^l), j} v_{j, \pi(\bar{2}^l)}^* \right] \cdots \left[\sum_{j=1}^m v_{\pi(\bar{k}_l-1^l), j} v_{j, \pi(\bar{k}_l^l)}^* \right] \\ &= \frac{1}{m^{k_l}} \sum_{\gamma \in (1, \dots, m)^{k_l}} v_{\pi(\bar{0}^l), \gamma(\bar{1}^l)} v_{\gamma(\bar{1}^l), \pi(\bar{1}^l)}^* v_{\pi(\bar{1}^l), \gamma(\bar{2}^l)} v_{\gamma(\bar{2}^l), \pi(\bar{2}^l)}^* \cdots v_{\pi(\bar{k}_l-1^l), \gamma(\bar{k}_l^l)} v_{\gamma(\bar{k}_l^l), \pi(\bar{k}_l^l)}^*. \end{aligned}$$

Thus, we obtain that

$$(16) \quad \begin{aligned} & \mathbb{E} \left[\prod_{l=1}^p \sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \right] \\ &= m^{-K} \sum_{\pi \in (1, \dots, n)^K} \mathbb{E} \left[\prod_{l=1}^p \sum_{\gamma \in (1, \dots, m)^{k_l}} (V_{l, \pi, \gamma}^{k_l} - \mathbb{E} V_{l, \pi, \gamma}^{k_l}) \right] \mathbf{1}_{\{1 \leq \pi(\bar{0}^l) = \pi(\bar{k}_l^l) \leq ns_l\}}, \end{aligned}$$

where

$$V_{l, \pi, \gamma}^{k_l} = v_{\pi(\bar{0}^l), \gamma(\bar{1}^l)} v_{\gamma(\bar{1}^l), \pi(\bar{1}^l)}^* v_{\pi(\bar{1}^l), \gamma(\bar{2}^l)} v_{\gamma(\bar{2}^l), \pi(\bar{2}^l)}^* \cdots v_{\pi(\bar{k}_l-1^l), \gamma(\bar{k}_l^l)} v_{\gamma(\bar{k}_l^l), \pi(\bar{k}_l^l)}^*.$$

The product and the sum in the expectation of (16) can be expressed and developed as the following:

$$\begin{aligned}
& \prod_{l=1}^p \sum_{\gamma \in (1, \dots, m)^{k_l}} v_{\pi(\bar{0}^l), \gamma(\bar{1}^l)} v_{\gamma(\bar{1}^l), \pi(\bar{1}^l)}^* v_{\pi(\bar{1}^l), \gamma(\bar{2}^l)} v_{\gamma(\bar{2}^l), \pi(\bar{2}^l)}^* \cdots v_{\pi(\bar{k_l-1}^l), \gamma(\bar{k_l}^l)} v_{\gamma(\bar{k_l}^l), \pi(\bar{k_l}^l)}^* \\
& = \sum_{\gamma \in (1, \dots, m)^K} (v_{\pi(\bar{0}^1), \gamma(\bar{1}^1)} v_{\gamma(\bar{1}^1), \pi(\bar{1}^1)}^* \cdots v_{\pi(\bar{k_1-1}^1), \gamma(\bar{k_1}^1)} v_{\gamma(\bar{k_1}^1), \pi(\bar{k_1}^1)}^*) \cdots (v_{\pi(\bar{0}^p), \gamma(\bar{1}^p)} v_{\gamma(\bar{1}^p), \pi(\bar{1}^p)}^* \cdots v_{\pi(\bar{k_p-1}^p), \gamma(\bar{k_p}^p)} v_{\gamma(\bar{k_p}^p), \pi(\bar{k_p}^p)}^*) \\
& = \sum_{\gamma \in (1, \dots, m)^K} \prod_{l=1}^p \left(v_{\pi(\bar{0}^l), \gamma(\bar{1}^l)} v_{\gamma(\bar{1}^l), \pi(\bar{1}^l)}^* \cdots v_{\pi(\bar{k_l-1}^l), \gamma(\bar{k_l}^l)} v_{\gamma(\bar{k_l}^l), \pi(\bar{k_l}^l)}^* \right).
\end{aligned}$$

Therefore, we find that the quantity (15) is equal to:

$$(17) \quad m^{-K} \sum_{\pi \in (1, \dots, n)^K} \sum_{\gamma \in (1, \dots, m)^K} \mathbb{E} \left[\prod_{l=1}^p (V_{l,\pi,\gamma} - \mathbb{E} V_{l,\pi,\gamma}) \right] \mathbf{1}_{\{1 \leq \pi(\bar{0}^l) = \pi(\bar{k_l}^l) \leq ns_l\}},$$

where

$$V_{l,\pi,\gamma} = v_{\pi(\bar{0}^l), \gamma(\bar{1}^l)} v_{\gamma(\bar{1}^l), \pi(\bar{1}^l)}^* \cdots v_{\pi(\bar{k_l-1}^l), \gamma(\bar{k_l}^l)} v_{\gamma(\bar{k_l}^l), \pi(\bar{k_l}^l)}^*.$$

Now, as we work directly with the elements $v_{i,j}$ of the starting matrix V_n , we can use the assumptions (4, 5, 6) of Theorem 2.1. Note that the fact that the variables $v_{i,j}$'s are i.i.d. allows us to group all the combinations which behave in the same way in the product of (17). Let $\text{Part}(\mathcal{E})$ denote the set of all partitions of \mathcal{E} , and set

$$\mathcal{E}_\gamma = \{\bar{1}^1, \dots, \bar{k_1}^1\} \cup \dots \cup \{\bar{1}^p, \dots, \bar{k_p}^p\}.$$

For each partition α in $\text{Part}(\mathcal{E})$, for each $x \in \mathcal{E}$, we denote by $\alpha(x)$ the index of the class of x , after having ordered the classes according to the order of their first element (for example, $\alpha(\bar{1}^1) = 1$; $\alpha(\bar{2}^1) = 1$ if $\bar{1}^1 \approx \bar{2}^1$ and $\alpha(\bar{2}^1) = 2$ if $\bar{1}^1 \not\approx \bar{2}^1$). Therefore, we can write (17) as two sums on the sets $\text{Part}(\mathcal{E})$, $\text{Part}(\mathcal{E}_\gamma)$ introduced above. We get

$$\begin{aligned}
& \mathbb{E} \left[\prod_{l=1}^p \sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \right] \\
(18) \quad & = m^{-K} \sum_{\pi \in \text{Part}(\mathcal{E})} \sum_{\gamma \in \text{Part}(\mathcal{E}_\gamma)} A(n, \pi) B(m, \gamma) \mathbb{E} \left[\prod_{l=1}^p (V_{l,\pi,\gamma} - \mathbb{E} V_{l,\pi,\gamma}) \right],
\end{aligned}$$

where:

- $V_{l,\pi,\gamma}$ is defined this time with two functions π and γ as shown in the previous paragraph for the general definition of α ,

$$V_{l,\pi,\gamma} = v_{\pi(\bar{0}^l), \gamma(\bar{1}^l)} v_{\gamma(\bar{1}^l), \pi(\bar{1}^l)}^* \cdots v_{\pi(\bar{k_l-1}^l), \gamma(\bar{k_l}^l)} v_{\gamma(\bar{k_l}^l), \pi(\bar{k_l}^l)}^*.$$

- For each $\pi \in \text{Part}(\mathcal{E})$, $A(n, \pi)$ is the number of families of indices of

$$(i_{\bar{0}^1}, \dots, i_{\bar{k}_1^1}, i_{\bar{0}^2}, \dots, i_{\bar{k}_2^2}, \dots, \dots, i_{\bar{0}^p}, \dots, i_{\bar{k}_p^p}) \in (1, \dots, n)^K$$

whose level sets partition is π and that satisfies, for each $l = 1, \dots, p$

$$(19) \quad 1 \leq \pi(\bar{0}^l) = \pi(\bar{k}_l^l) \leq ns_l.$$

- For each $\gamma \in \text{Part}(\mathcal{E}_\gamma)$, $B(m, \gamma)$ is the number of families of indices of

$$(i_{\bar{1}^1}, \dots, i_{\bar{k}_1^1}, i_{\bar{1}^2}, \dots, i_{\bar{k}_2^2}, \dots, \dots, i_{\bar{1}^p}, \dots, i_{\bar{k}_p^p}) \in (1, \dots, m)^K$$

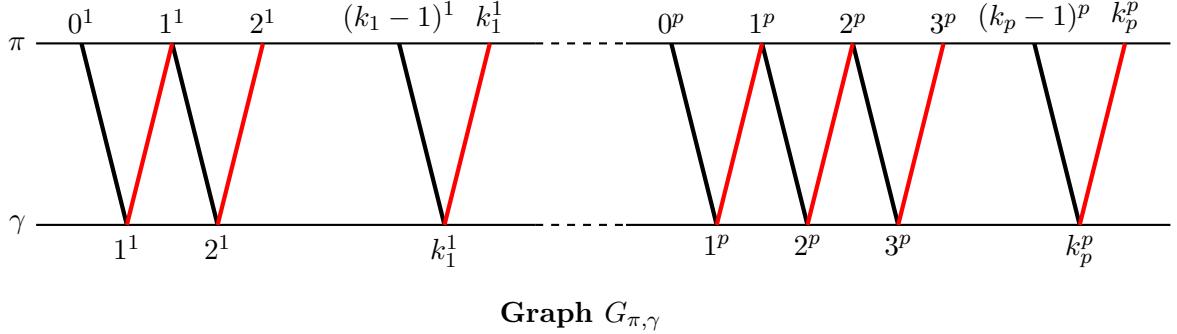
whose level sets partition is γ .

For any partitions $\pi \in \text{Part}(\mathcal{E})$ and $\gamma \in \text{Part}(\mathcal{E}_\gamma)$, let us define $G_{\pi, \gamma}$ to be the graph with vertex set

$$V_{\pi, \gamma} = \{\pi(x), \gamma(x); \quad x \in \mathcal{E} \text{ for } \pi, \quad x \in \mathcal{E}_\gamma \text{ for } \gamma\},$$

and edge set

$$E_{\pi, \gamma} = \left\{ \{\pi(\bar{m}^l), \gamma(\bar{m}^l)\}, \{\gamma(\bar{m}^l), \pi(\bar{m}^l)\} \quad ; \quad 1 \leq l \leq p, \quad m \in \{1, \dots, k_l\} \right\}.$$



Graph $G_{\pi, \gamma}$

For the term associated to a $(\pi, \gamma) \in (\text{Part}(\mathcal{E}) \times \text{Part}(\mathcal{E}_\gamma))$ in (18) to be non zero, we need to have:

- for each $l = 1, \dots, p$, $\pi(\bar{0}^l) = \pi(\bar{k}_l^l)$,
- each edge of $G_{\pi, \gamma}$ is visited at least twice by the union of the p paths $(\pi(\bar{0}^l), \gamma(\bar{1}^l), \pi(\bar{1}^l), \dots, \gamma(\bar{k}_l^l), \pi(\bar{k}_l^l))$, $l = 1, \dots, p$,
- for each $l = 1, \dots, p$, there exists $l' \neq l$ such that at least one edge of $G_{\pi, \gamma}$ is visited by both paths $(\pi(\bar{0}^l), \gamma(\bar{1}^l), \pi(\bar{1}^l), \dots, \gamma(\bar{k}_l^l), \pi(\bar{k}_l^l))$ and $(\pi(\bar{0}^{l'}), \gamma(\bar{1}^{l'}), \pi(\bar{1}^{l'}), \dots, \gamma(\bar{k}_{l'}^{l'}), \pi(\bar{k}_{l'}^{l'}))$.

Indeed, (i) is due to (19), (ii) is due to the fact that $v_{i,j}$'s are independent and centered and (iii) is due to the fact that the $v_{i,j}$'s are independent and that the variables $V_{l, \pi, \gamma} - \mathbb{E} V_{l, \pi, \gamma}$ are centered.

Now, let us define a function $s(\cdot)$ on the set \mathcal{E} in order to control the condition (19) in the following way: for each $l = 1, \dots, p$ and each $h = 0, \dots, k_l$, set

$$s(\bar{h}^l) = \begin{cases} s_l & \text{if } h = 0 \text{ or } h = k_l \\ 1 & \text{otherwise,} \end{cases}$$

and

$$(20) \quad s_\pi = \prod_{B \text{ bloc of } \pi} \min_{x \in B} s(x).$$

Then one can easily see that, as n tends to infinity,

$$A(n, \pi) \sim s_\pi n^{|\pi|},$$

and

$$B(m, \gamma) \sim m^{|\gamma|},$$

where $|\pi|$ denotes the number of vertices indexed by π in the graph $G_{\pi, \gamma}$, and $|\gamma|$ denotes the number of vertices indexed by γ in $G_{\pi, \gamma}$.

Therefore, for (π, γ) to have a non zero asymptotic contribution to (18), we need the following condition, in addition to (i), (ii) and (iii):

$$(iv) \quad K \leq |\pi| + |\gamma|.$$

Now, let us introduce this lemma which is the analogue of [1, Lemma 2.1.34]. Its proof goes along the same lines as the proof of the former (see also [7, Lemma 4.1]).

Lemma 4.1 (Combinatorics). *Let $(\pi, \gamma) \in (\text{Part}(\mathcal{E}) \times \text{Part}(\mathcal{E}_\gamma))$ satisfy (i), (ii) and (iii). Then the number C_G of connected components of $G_{\pi, \gamma}$ is such that $C_G \leq p/2$ and*

$$|V_{\pi, \gamma}| = |\pi| + |\gamma| \leq C_G - \frac{p}{2} + K.$$

As a consequence, if (π, γ) also satisfies (iv), we have

- (a) $C_G = p/2$,
- (b) p is even,
- (c) $|V_{\pi, \gamma}| = |\pi| + |\gamma| = K$.

Also note that by (ii), we have

$$(d) \quad |E_{\pi, \gamma}| \leq K,$$

where $|E_{\pi, \gamma}|$ denotes the number of edges of the graph $G_{\pi, \gamma}$. Therefore, by (18) and (c), we get

$$(21) \quad \lim_{n \rightarrow \infty} \mathbb{E} \left[\prod_{l=1}^p \sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \right] = \\ \sum_{\pi \in \text{Part}(\mathcal{E})} \sum_{\gamma \in \text{Part}(\mathcal{E}_\gamma)} s_\pi \lim_{n \rightarrow \infty} \mathbb{E} \left[\prod_{l=1}^p (V_{l, \pi, \gamma} - \mathbb{E} V_{l, \pi, \gamma}) \right],$$

where the sum is taken over the partitions (π, γ) which satisfy (i), (ii), (iii) and (iv) above, and such partitions also do satisfy (a), (b), (c) and (d) above.

Case where p is odd: By (b), we know that when p is odd, there is no couple of partitions (π, γ) satisfying the above conditions, hence

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\prod_{l=1}^p \sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \right] = 0.$$

Case where $p = 2$: In this case, by (a) we know that for each couple of partitions (π, γ) satisfying (i), (ii), (iii) and (iv) above, the graph $G_{\pi, \gamma}$ is connected. so that $|V_{\pi, \gamma}| - 1 \leq |E_{\pi, \gamma}|$. Therefore, by (c) and (d) $|E_{\pi, \gamma}|$ is either equal to K or $K - 1$:

- $|E_{\pi, \gamma}| = K - 1$: In this case, the graph $G_{\pi, \gamma}$ has exactly one more vertex than edges, hence it is a tree. As a consequence, the paths $(\pi(\bar{0}^1), \gamma(\bar{1}^1), \pi(\bar{1}^1), \gamma(\bar{k}_1^1), \pi(\bar{k}_1^1))$ and $(\pi(\bar{0}^2), \gamma(\bar{1}^2), \pi(\bar{1}^2), \dots, \gamma(\bar{k}_2^2), \pi(\bar{k}_2^2))$, which have the same beginning and ending vertices, satisfy the property that each visited edge is visited an even number of times. By an obvious cardinality argument, only one edge is visited more than twice, and it is visited four times (twice in each sense). The other edges are visited once in each sense. It follows that the expectation associated to a couple (π, γ) in (21) is equal to $\mathbb{E}[|v_{1,1}|^4] - 1$.
- $|E_{\pi, \gamma}| = K$: In this case, the graph $G_{\pi, \gamma}$ has exactly the same number of vertices as edges, hence it is a bracelet. Therefore, by a cardinality argument again, the paths $(\pi(\bar{0}^1), \gamma(\bar{1}^1), \pi(\bar{1}^1), \dots, \gamma(\bar{k}_1^1), \pi(\bar{k}_1^1))$ and $(\pi(\bar{0}^2), \gamma(\bar{1}^2), \pi(\bar{1}^2), \dots, \gamma(\bar{k}_2^2), \pi(\bar{k}_2^2))$ satisfy the property that they visit exactly twice of times each edge they visit (once in each sense). It follows that the expectation associated to a couple (π, γ) in (21) is equal to 1.

As a consequence, as n tends to infinity,

$$\mathbb{E} \left[\sum_{1 \leq i \leq ns_1} (e_i^* X_n^{k_1} e_i - \mathbb{E}[e_i^* X_n^{k_1} e_i]) \times \sum_{1 \leq i \leq ns_2} (e_i^* X_n^{k_2} e_i - \mathbb{E}[e_i^* X_n^{k_2} e_i]) \right]$$

converges to a number that we shall denote by

$$(22) \quad \text{Cov}_{s_1, s_2}(k_1, k_2),$$

which depends on the distribution of the $v_{i,j}$'s only through $\lim_{n \rightarrow \infty} \mathbb{E}[|v_{1,1}|^4]$.

Case where p is > 2 and even: By (a) above, for each couple of partitions (π, γ) satisfying (i), (ii), (iii) and (iv), $G_{\pi, \gamma}$ has exactly $p/2$ connected components. By (iii), each one of them contains the support of exactly two of the p paths

$$(\pi(\bar{0}^l), \gamma(\bar{1}^l), \pi(\bar{1}^l), \dots, \gamma(\bar{k}_l^l), \pi(\bar{k}_l^l)) \quad (l = 1, \dots, p).$$

To join every two paths having the same support in the expectation of (21), let us define $\sigma_{\pi, \gamma}$ to be the matching (i.e. a permutation all of whose cycles have length two) of $\{1, \dots, p\}$ such that for all $l = 1, \dots, p$, the paths with indices l and $\sigma_{\pi, \gamma}(l)$ are supported by the same connected component of $G_{\pi, \gamma}$.

We shall now partition the sum of (21) according to the value of the matching $\sigma_{\pi,\gamma}$ defined by (π, γ) . We get

$$(23) \quad \lim_{n \rightarrow \infty} \mathbb{E} \left[\prod_{l=1}^p \sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \right] = \\ \sum_{\sigma} \sum_{\pi, \gamma} s_{\pi} \lim_{n \rightarrow \infty} \mathbb{E} \left[\prod_{l=1}^p (V_{l, \pi, \gamma} - \mathbb{E} V_{l, \pi, \gamma}) \right] \mathbf{1}_{\{\sigma_{\pi, \gamma} = \sigma\}},$$

where the first sum is over the matchings σ of $\{1, \dots, p\}$ and the second sum is over the couples of partitions (π, γ) satisfying (i), (ii), (iii) and (iv).

Note that for each matching σ of $\{1, \dots, p\}$, the set of couples of partitions $(\pi, \gamma) \in (\text{Part}(\mathcal{E}) \times \text{Part}(\mathcal{E}_{\gamma}))$ such that $\sigma_{\pi, \gamma} = \sigma$ can be identified with the Cartesian product, indexed by the set of cycles $\{l, l'\}$ of σ , of the set of couples of partitions $(\bar{\pi}, \bar{\gamma}) \in (\text{Part}(\mathcal{E}^{l, l'}) \times \text{Part}(\mathcal{E}_{\gamma}^{l, l'}))$ such that

$$\mathcal{E}^{l, l'} = \{\bar{0}^l, \dots, \bar{k}_l^l\} \cup \{\bar{0}^{l'}, \dots, \bar{k}_l^{l'}\} \quad (\text{subset of } \mathcal{E})$$

and

$$\mathcal{E}_{\gamma}^{l, l'} = \{\bar{1}^l, \dots, \bar{k}_l^l\} \cup \{\bar{1}^{l'}, \dots, \bar{k}_l^{l'}\} \quad (\text{subset of } \mathcal{E}_{\gamma})$$

satisfying the following conditions

- (i') $\bar{\pi}(\bar{0}^l) = \bar{\pi}(\bar{k}_l^l)$ and $\bar{\pi}(\bar{0}^{l'}) = \bar{\pi}(\bar{k}_l^{l'})$,
- (ii') each edge of the graph $G_{\bar{\pi}, \bar{\gamma}}$ is visited at least twice by the union of its two paths indexed by the corresponding l and l' .
- (iii') at least one edge of $G_{\bar{\pi}, \bar{\gamma}}$ is visited by both previous paths,
- (iv') $k_l + k_{l'} \leq |\bar{\pi}| + |\bar{\gamma}|$.

Moreover, one can see that the factor s_{π} factorizes along the connected components of $G_{\pi, \gamma}$, and by the independence of the random variables $v_{i,j}$'s, the expectation

$$\mathbb{E} \left[\prod_{l=1}^p \sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \right]$$

also factorizes along the connected components of $G_{\pi, \gamma}$. Therefore, we get

$$(24) \quad \lim_{n \rightarrow \infty} \mathbb{E} \left[\prod_{l=1}^p \sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \right] = \\ \sum_{\sigma} \prod_{l, l'} \lim_{n \rightarrow \infty} \mathbb{E} \left[\sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \times \sum_{1 \leq i \leq ns_{l'}} (e_i^* X_n^{k_{l'}} e_i - \mathbb{E}[e_i^* X_n^{k_{l'}} e_i]) \right]$$

where the sum is over the matchings σ of $\{1, \dots, p\}$ and for each such σ , the product is over the cycles $\{l, l'\}$ of σ .

By the previous definition of $\text{Cov}_{s_1, s_2}(k_1, k_2)$ in (22), we get

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\prod_{l=1}^p \sum_{1 \leq i \leq ns_l} (e_i^* X_n^{k_l} e_i - \mathbb{E}[e_i^* X_n^{k_l} e_i]) \right] = \sum_{\sigma \text{ matching}} \prod_{\{l, l'\} \text{cycle of } \sigma} \text{Cov}_{s_l, s_{l'}}(k_l, k_{l'}).$$

By Wick's formula and Equation (11), we have proved the first part of Proposition (3.4).

We finish the proof by this last step.

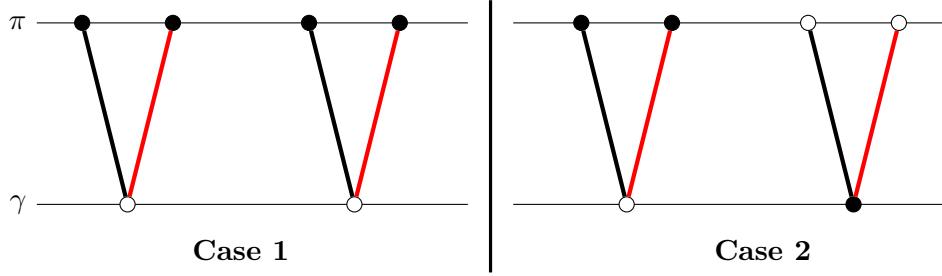
Computation of $\text{Cov}_{s_1, s_2}(1, 1)$: In this case, we have $p = 2$, and $K = k_1 + k_2 = 2$. Therefore, by (a), (c) and (d),

- $G_{\pi, \gamma}$ is connected,
- $V_{\pi, \gamma} = K = 2$,
- $E_{\pi, \gamma} \leq 2$.

With two vertices, there is exactly one tree and zero bracelet. Thus, we have by the paragraph devoted to the case $p = 2$,

$$\text{Cov}_{s_1, s_2}(1, 1) = \sum_{(\pi, \gamma), G_{\pi, \gamma} \text{ is a tree}} s_{\pi} \left(\lim_{n \rightarrow \infty} \mathbb{E}[|v_{1,1}|^4] - 1 \right)$$

For this tree, there are two associated couples of partitions (π, γ) :



Case (1): The partition π is defined by

$$\pi(\mathcal{E}) = \{\{\bar{0}^1, \bar{1}^1, \bar{0}^2, \bar{1}^2\}\},$$

hence by (20), $s_{\pi} = \min\{s_1, s_2\}$.

Case (2): In this case, the partition π is defined by

$$\pi(\mathcal{E}) = \{\{\bar{0}^1, \bar{1}^1\}, \{\bar{0}^2, \bar{1}^2\}\},$$

hence, $s_{\pi} = s_1 s_2$.

As a consequence,

$$\text{Cov}_{s_1, s_2}(1, 1) = \left(\lim_{n \rightarrow \infty} \mathbb{E}[|v_{1,1}|^4] - 1 \right) (\min\{s_1, s_2\} + s_1 s_2).$$

Now, by Equation (11), we get

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E} \left[\sum_{1 \leq i \leq ns_1} (e_i^* X_n e_i - \frac{1}{n} \text{Tr}(X_n)) \times \sum_{1 \leq i \leq ns_2} (e_i^* X_n e_i - \frac{1}{n} \text{Tr}(X_n)) \right] \\ &= \text{Cov}_{s_1, s_2}(1, 1) - s_2 \text{Cov}_{s_1, 1}(1, 1) - s_1 \text{Cov}_{1, s_2}(1, 1) + s_1 s_2 \text{Cov}_{1, 1}(1, 1) \\ &= \left(\lim_{n \rightarrow \infty} \mathbb{E}[|v_{1,1}|^4] - 1 \right) [(\min\{s_1, s_2\} + s_1 s_2) - (2s_1 s_2) - (2s_1 s_2) + (2s_1 s_2)] \\ &= \left(\lim_{n \rightarrow \infty} \mathbb{E}[|v_{1,1}|^4] - 1 \right) (\min\{s_1, s_2\} - s_1 s_2), \end{aligned}$$

which concludes the proof of Proposition 3.4. \square

5. TIGHTNESS AND CONVERGENCE IN THE SKOROKHOD TOPOLOGY

For Wigner matrices, Benaych-Georges proved that the bivariate process B^n converges in distribution, for the Skorokhod topology in $D[0, 1]^2$, to the bivariate Brownian bridge under several assumptions on the atom distribution: absolute continuity, moments of all orders and matching with a GUE/GOE matrix up to order 10 on the diagonal and up to order 12 off the diagonal. In order to prove this convergence, he used some ideas developed by Tao and Vu in [27], especially, the "Four Moment Theorem for eigenvectors of Wigner matrices", see [27, theorem 8].

To our knowledge, such a theorem is not yet available for the case of sample covariance matrices. We formulate the statement in Hypothesis 5.2 below. If this is indeed the case, convergence of the process B^n for the Skorokhod topology in $D[0, 1]^2$ will be also verified in our case. For proving this, we will follow closely the strategy of Benaych-Georges.

Definition 5.1 (Matching moments). *Let $k \geq 1$. Two random matrices $V_n = (v_{i,j})_{i \leq n, j \leq m}$, $V'_n = (v'_{i,j})_{i \leq n, j \leq m}$ are said to match up to order k , if one has*

$$\mathbb{E} Re(v_{i,j})^a Im(v_{i,j})^b = \mathbb{E} Re(v'_{i,j})^a Im(v'_{i,j})^b$$

whenever $a, b \geq 0$, $1 \leq i \leq n$ and $1 \leq j \leq m$ are integers such that $a + b \leq k$.

Before stating the theorem of convergence of our process B^n for the Skorokhod topology in $D[0, 1]^2$, let us give the following hypothesis that we will need:

Hypothesis 5.2 (Matching theorem for eigenvectors). *We suppose that if the matrix V_n matches a $(n \times m)$ -Gaussian matrix M_n (i.e. matrix whose elements are independent standard Gaussian variables) to order $l \geq 4$, then, for any fixed positive integer k and polynomial function G on \mathbb{C}^k , there exists a certain constant C independent of n such that*

$$(25) \quad |\mathbb{E}[G(n u_{p_1,i_1} \bar{u}_{q_1,i_1}, \dots, n u_{p_k,i_k} \bar{u}_{q_k,i_k})] - \mathbb{E}[G(n u'_{p_1,i_1} \bar{u}'_{q_1,i_1}, \dots, n u'_{p_k,i_k} \bar{u}'_{q_k,i_k})]| \leq Cn^{2-\frac{l}{2}}.$$

whenever $(i_1, p_1, q_1), \dots, (i_k, p_k, q_k)$ is a collection of indices in $\{1, \dots, n\}^3$, $U_n = (u_{i,j})_{1 \leq i,j \leq n}$ is the eigenmatrix of the sample covariance matrix $X_n = (1/m)V_n V_n^*$ and $U'_n = (u'_{i,j})_{1 \leq i,j \leq n}$ is the eigenmatrix of the Laguerre matrix $L_n = \frac{1}{m}M_n M_n^*$.

Theorem 5.3 (Convergence in the Skorokhod topology). *For the sample covariance matrix $X_n = (1/m)V_n V_n^*$ defined as in Section 2, suppose that*

- (i) *The distribution of the entries of (V_n) are absolutely continuous with respect to the Lebesgue measure.*
- (ii) $\forall k \geq 0$, $\sup_n \mathbb{E}|v_{1,1}^{(n)}|^k < \infty$.
- (iii) *(V_n) matches a $(n \times m)$ -Gaussian matrix M_n up to order l , and that Hypothesis 5.2 is satisfied.*

Then, for $l = 12$, the bivariate process B^n converges in distribution, for the Skorokhod topology in $D([0, 1]^2)$, to the bivariate Brownian bridge.

Remark 5.4 (Comments on the assumptions of Theorem 5.3). *These assumptions might not to be optimal, especially the continuity one and matching up to order 12. We hope to prove this theorem under Assumption (iii) for $l=4$ instead of $l=12$.*

Proving Theorem 5.3 consists to prove the following lemma of tightness and uniqueness of the accumulation point argument.

Lemma 5.5 (Tightness argument). *Under Assumptions of Theorem 5.3, the sequence (distribution $(B^n)_{n \geq 1}$) is C-tight, i.e. is tight and has only one $C([0, 1]^2)$ -supported accumulation point.*

Proof of Theorem 5.3. Note that Theorem 5.3 allows us to show that for all $0 \leq s < s' \leq 1$ and $0 \leq t < t' \leq 1$, the sequence of random variables

$$\frac{1}{\sqrt{(s' - s)(t' - t)}} \sum_{\substack{ns < i \leq ns' \\ nt < j \leq nt'}} (|u_{i,j}|^2 - 1/n)$$

admits a limit in distribution as $n \rightarrow \infty$, hence is bounded in probability (in the sense of [26, Def. 1.1: $\lim_{C \rightarrow \infty} \liminf_{n \rightarrow \infty} \mathbb{P}(|X_n| \leq C) = 1$]). In the next proposition, we improve these assertions by making them uniform on s, s', t, t', i, j and upgrading them to the L^2 and L^4 levels. This proposition is almost sufficient to apply the tightness argument of the (distribution $(B^n)_{n \geq 1}$).

Proposition 5.6 (Control of Jumps). *Suppose that Assumptions (i), (ii) and (iii) for $l = 4$ (resp. $l = 8$) are satisfied. Then as $n \rightarrow \infty$, the sequence*

$$(26) \quad n|u_{i,j}|^2 - 1 \quad (\text{resp.} \quad \frac{1}{\sqrt{(s' - s)(t' - t)}} \sum_{\substack{ns < i \leq ns' \\ nt < j \leq nt'}} (|u_{i,j}|^2 - 1/n))$$

is bounded for the L^4 (resp. L^2) norm, uniformly in i, j (resp. $s < s', t < t'$).

The proof of Proposition 5.6 goes along the same lines as the proof given by Benaych-Georges in [7, section 4.4]. Indeed, Hypothesis 5.2 "matching with Gaussian matrix" allows us to work with the entries of a Haar-distributed matrix instead of the entries of the eigenmatrix of the sample covariance matrix X_n . Note only that, if the second term of (26) have been bounded for $L^{2+\epsilon}$ instead of L^2 , Assumption (iii) for $l = 8$ would have been enough to prove the convergence of B^n in distribution, for the Skorokhod topology in $D([0, 1]^2)$, to the bivariate Brownian bridge.

Now, to prove Lemma 5.5, we give the following proposition, which is the obvious multidimensional generalization of Proposition 3.26 of [16, Chapter VI]:

For $f \in D([0, 1]^2)$ and $(s_0, t_0) \in [0, 1]^2$, we define $\Delta_{s_0, t_0} f$ to be the "maximal jump" of f at (s_0, t_0) , i.e.

$$\Delta_{s_0, t_0} f := \max_{\diamond, \diamond' \in \{\langle, \rangle, \geq\}} \left| f(s_0, t_0) - \lim_{\substack{s \rightarrow \diamond s_0 \\ t \rightarrow \diamond' t_0}} f(s, t) \right|.$$

Proposition 5.7 (C-Tightness). *If the sequence (distribution $(B^n)_{n \geq 1}$) is tight and satisfies*

$$(27) \quad \forall \epsilon > 0, \quad \mathbb{P}\left(\sup_{(s,t) \in [0,1]^2} \Delta_{s,t} B^n > \epsilon\right) \rightarrow_{n \rightarrow \infty} 0,$$

then the sequence (distribution $(B^n)_{n \geq 1}$) is C-tight, i.e. is tight and has only one $C([0, 1]^2)$ -supported accumulation point.

So to prove Lemma 5.5, let us first prove that the sequence (distribution $(B^n)_{n \geq 1}$) is tight. For this, we follow closely the proof of Benaych-Georges.

Note that the process B^n vanishes at the border of $[0, 1]^2$. So according to [9, Th. 3] and to Cauchy-Schwartz inequality, it suffices to prove that there exists $C < \infty$ such that for n large enough, for all $s < s', t < t' \in [0, 1]$,

$$\mathbb{E}[\{\sum_{ns < i \leq ns'} \sum_{nt < j \leq nt'} (|u_{i,j}|^2 - 1/n)\}^4] \leq C(s' - s)^2(t' - t)^2.$$

As in the proof of Proposition 5.6, one can suppose that the $u_{i,j}$'s are the entries of a Haar-distributed matrix. But in this case, the job has already been done in [13]: the unitary case is treated in Section 3.4.1 (see specifically Equation (3.25)) and the orthogonal case is treated, more elliptically, in Section 4.5.

Let us now prove (27). Note that $\sup_{(s,t) \in [0,1]^2} \Delta_{s,t} B^n = \max_{1 \leq i,j \leq n} |u_{i,j}|^2 - 1/n$. As a consequence, by the union bound, it suffices to prove that for each $\epsilon > 0$, there exists $C < \infty$ independent of i, j and n such that for all i, j ,

$$\mathbb{P}(|u_{i,j}|^2 - 1/n > \epsilon) \leq Cn^{-4},$$

which follows from Chebyshev's inequality and Proposition 5.6. \square

REFERENCES

- [1] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
- [2] Z. D. Bai, B. Q. Miao, and G. M. Pan. On asymptotics of eigenvectors of large sample covariance matrix. *Ann. Probab.*, 35(4):1532–1572, 2007.
- [3] Z. D. Bai and G. M. Pan. Limiting behavior of eigenvectors of large Wigner matrices. *J. Stat. Phys.*, 146(3):519–549, 2012.
- [4] Z. D. Bai, J. W. Silverstein, and Y. Q. Yin. A note on the largest eigenvalue of a large-dimensional sample covariance matrix. *J. Multivariate Anal.*, 26(2):166–168, 1988.
- [5] Z. D. Bai and Y. Q. Yin. Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. *Ann. Probab.*, 21(3):1275–1294, 1993.
- [6] Zhidong Bai and Jack W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics. Springer, New York, second edition, 2010.
- [7] F. Benaych-Georges. A universality result for the global fluctuations of the eigenvectors of wigner matrices. *Random Matrices Theory Appl.*, 01(04):1250011, 2012.
- [8] P. J. Bickel and M. J. Wichura. Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.*, 42:1656–1670, 1971.
- [9] P. J. Bickel and M. J. Wichura. Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.*, 42:1656–1670, 1971.
- [10] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.
- [11] C. Bordenave and A. Guionnet. Localization and delocalization of eigenvectors for heavy-tailed random matrices. *ArXiv:1201.1862*, January 2012.
- [12] C. Cacciapuoti, A. Maltsev, and B. Schlein. Local marchenko-pastur law at the hard edge of sample covariance matrices. *Journal of Mathematical Physics*, 54(4):043302, April 2013.
- [13] C. Donati-Martin and A. Rouault. Truncations of Haar distributed matrices, traces and bivariate Brownian bridges. *Random Matrices Theory Appl.*, 1(1):1150007, 24, 2012.
- [14] L. Erdős, B. Schlein, and H.-T. Yau. Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.*, 287(2):641–655, 2009.
- [15] L. Erdős, B. Schlein, and H.-T. Yau. Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.*, 37(3):815–852, 2009.

- [16] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1987.
- [17] T. Jiang. How many entries of a typical orthogonal matrix can be approximated by independent normals? *Ann. Probab.*, 34(4):1497–1529, 2006.
- [18] A. Knowles and J. Yin. Eigenvector distribution of Wigner matrices. *Probab. Theory Related Fields*, 155(3–4):543–582, 2013.
- [19] O. Ledoit and S. Péché. Eigenvectors of some large sample covariance matrix ensembles. *Probab. Theory Related Fields*, 151(1-2):233–264, 2011.
- [20] N. S. Pillai and J. Yin. Universality of Covariance Matrices. *ArXiv:1110.2501*, October 2011.
- [21] M. Rudelson and R. Vershynin. Delocalization of eigenvectors of random matrices with independent entries. *ArXiv:1306.2887*, June 2013.
- [22] J. Schenker. Eigenvector localization for random band matrices with power law band width. *Comm. Math. Phys.*, 290(3):1065–1097, 2009.
- [23] J. W. Silverstein. Describing the behavior of eigenvectors of random matrices using sequences of measures on orthogonal groups. *SIAM J. Math. Anal.*, 12(2):274–281, 1981.
- [24] J. W. Silverstein. Some limit theorems on the eigenvectors of large-dimensional sample covariance matrices. *J. Multivariate Anal.*, 15(3):295–324, 1984.
- [25] J. W. Silverstein. Weak convergence of random functions defined by the eigenvectors of sample covariance matrices. *Ann. Probab.*, 18(3):1174–1194, 1990.
- [26] T. Tao and V. Vu. Random matrices: universality of ESDs and the circular law. *Ann. Probab.*, 38(5):2023–2065, 2010. With an appendix by Manjunath Krishnapur.
- [27] T. Tao and V. Vu. Random matrices: universal properties of eigenvectors. *Random Matrices Theory Appl.*, 1(1):1150001, 27, 2012.
- [28] Terence Tao. *Topics in random matrix theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.
- [29] V. Vu and K. Wang. Random weighted projections, random quadratic forms and random eigenvectors. *ArXiv:1306.3099*, June 2013.
- [30] Y. Q. Yin. Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.*, 20(1):50–68, 1986.

ALI BOUFERROUM, LABORATOIRE D’ANALYSE ET DE MATHÉMATIQUES APPLIQUÉES, CNRS UMR8050,
UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE, FRANCE.

E-mail address: ali.bouferroum@univ-paris-est.fr