Truncations of Haar unitary matrices and bivariate tied-down Brownian bridge

A. Rouault (Versailles-Saint-Quentin),
joint work with C. Donati-Martin (UPMC)

12 octobre 2010
Workshop on Large Random Matrices
Telecom-Paris

Sketch of talk

- Introduction and main result
- Idea of Proof
- Related questions

Outline

Motivation

In computational biology, an important question is to measure the similarity between two genomic (long) sequences. If the sequences σ and τ are assumed to be random elements of $\mathfrak{S}_n$, the set of permutations of $[\![n]\!]$, biologists are interested in

$$O_p(\sigma, \tau) = \#\{i \leq p : \sigma \circ \tau^{-1}(i) \leq p\} \ , p = 1, \cdots , n \, .$$

Motivation

In computational biology, an important question is to measure the similarity between two genomic (long) sequences. If the sequences σ and τ are assumed to be random elements of $\mathfrak{S}_n$, the set of permutations of $[[n]]$, biologists are interested in

$$O_p(\sigma, \tau) = \#\{i \leq p : \sigma \circ \tau^{-1}(i) \leq p\}\ , p = 1, \cdots, n\,.$$

More generally, G. Chapuy (2007) introduced the discrepancy process

$$T^n_{p,q}(\sigma) = \#\{i \leq p : \sigma(i) \leq q\}\ , p, q = 1, \cdots, n\,,$$

Motivation

In computational biology, an important question is to measure the similarity between two genomic (long) sequences. If the sequences σ and τ are assumed to be random elements of $\mathfrak{S}_n$, the set of permutations of $[[n]]$, biologists are interested in

$$O_p(\sigma,\tau) = \#\{i \le p : \sigma \circ \tau^{-1}(i) \le p\} \ , p = 1, \cdots , n \, .$$

More generally, G. Chapuy (2007) introduced the discrepancy process

$$T^n_{p,q}(\sigma) = \#\{i \le p : \sigma(i) \le q\} \ , p, q = 1, \cdots , n \, ,$$

and proved that the normalized "discrepancy" process

$$n^{-1/2}\left(T^n_{\lfloor ns\rfloor,\lfloor nt\rfloor}(\sigma) - stn\right) \ , \ s,t \in [0,1]$$

converges in distribution to the tied down bivariate Brownian bridge, of covariance $(s \wedge s' - ss')(t \wedge t' - tt')$.

If σ is represented by the matrix $U(\sigma)$, the integer $Y^n_{p,q}(\sigma)$ is the sum of all elements of the upper-left $p \times q$ submatrix of $U(\sigma)$, i.e.

$$T^n_{p,q}(\sigma) = \operatorname{Tr} D_1 U(\sigma) D_2 U(\sigma)^*$$

where $D_1 = I_p$, $D_2 = I_q$ and $I_k = \mathrm{diag}(1, \cdots, 1, 0, \cdots, 0)$ (with k times 1) .

If σ is represented by the matrix $U(\sigma)$, the integer $Y^n_{p,q}(\sigma)$ is the sum of all elements of the upper-left $p \times q$ submatrix of $U(\sigma)$, i.e.

$$T^n_{p,q}(\sigma) = \operatorname{Tr} D_1 U(\sigma) D_2 U(\sigma)^*$$

where $D_1 = I_p$, $D_2 = I_q$ and $I_k = \operatorname{diag}(1, \cdots, 1, 0, \cdots, 0)$ (with k times 1) .

Instead of picking randomly σ in the group $\mathfrak{S}_n$, we propose to pick a random element U in the group $\mathbb{U}(n)$ and to study

$$T^n_{p,q} = \operatorname{Tr} D_1 U D_2 U^* = \sum_{i \le p, j \le q} |U_{ij}|^2 .$$

Outline

1. Introduction
   - Motivation
   - Main result
   - Previous related results
2. Sketch of the proof
   - Preliminary remarks
   - Combinatorics of the unitary group
   - Fidi convergence
   - Tightness
3. Complementary remarks
   - The marginals
   - Orthogonal case (in progress)
   - Conjectured universality

Main result

Theorem (CDM,AR, 2010)

The process

$$W^{(n)} = \left\{ T^{(n)}_{\lfloor ns \rfloor, \lfloor nt \rfloor} - \mathbb{E}\, T^{(n)}_{\lfloor ns \rfloor, \lfloor nt \rfloor} \;, \; s, t \in [0,1] \right\}$$

converges in distribution in $D([0,1]^2)$ *to the bivariate tied down Brownian bridge, i.e. the Gaussian process* $W^{(\infty)}$ *with covariance*

$$\mathbb{E}\left[W^{(\infty)}(s,t) W^{(\infty)}(s',t') \right] = (s \wedge s' - ss')(t \wedge t' - tt')\,.$$

Main result

Theorem (CDM,AR, 2010)

The process

$$W^{(n)} = \left\{ T^{(n)}_{\lfloor ns \rfloor, \lfloor nt \rfloor} - \mathbb{E}\, T^{(n)}_{\lfloor ns \rfloor, \lfloor nt \rfloor} \,,\; s, t \in [0,1] \right\}$$

converges in distribution in $D([0,1]^2)$ to the bivariate tied down Brownian bridge, i.e. the Gaussian process $W^{(\infty)}$ with covariance

$$\mathbb{E}\left[W^{(\infty)}(s,t) W^{(\infty)}(s',t') \right] = (s \wedge s' - ss')(t \wedge t' - tt')\,.$$

No normalization here !

Main result

Theorem (CDM,AR, 2010)

The process

$$W^{(n)} = \left\{ T^{(n)}_{\lfloor ns\rfloor,\lfloor nt\rfloor} - \mathbb{E}\, T^{(n)}_{\lfloor ns\rfloor,\lfloor nt\rfloor}\ ,\ \ s,t \in [0,1] \right\}$$

converges in distribution in $D([0,1]^2)$ to the bivariate tied down Brownian bridge, i.e. the Gaussian process $W^{(\infty)}$ with covariance

$$\mathbb{E}\left[W^{(\infty)}(s,t) W^{(\infty)}(s',t') \right] = (s \wedge s' - ss')(t \wedge t' - tt')\,.$$

No normalization here !

- If $\sigma \in \mathfrak{S}_n$, then $|U_{ij}|^2(\sigma) = U_{ij}(\sigma)$ and if σ is Haar distributed $\mathrm{Var}(|U_{ij}|^2) = n^{-1}(1-n^{-1})$
- If U is Haar distributed in $\mathbb{U}(n)$, then $\mathrm{Var}(|U_{ij}|^2) = n^{-2}$.

Outline

1. Introduction
   - Motivation
   - Main result
   - Previous related results
2. Sketch of the proof
   - Preliminary remarks
   - Combinatorics of the unitary group
   - Fidi convergence
   - Tightness
3. Complementary remarks
   - The marginals
   - Orthogonal case (in progress)
   - Conjectured universality

Previous related results

- If q is fixed, the vector $(U_{i,q})_{i=1}^{n}$ is uniformly distributed on the n dimensional complex sphere. It is well known (Silverstein 1981) that the process

$$n^{1/2}\left(\sum_{i=1}^{\lfloor ns\rfloor}|U_{iq}|^2 - s\right) \quad , \quad s \in [0,1]$$

converges in distribution to the Brownian bridge.

Previous related results

- If q is fixed, the vector $(U_{i,q})_{i=1}^n$ is uniformly distributed on the n dimensional complex sphere. It is well known (Silverstein 1981) that the process

$$n^{1/2}\left(\sum_{i=1}^{\lfloor ns\rfloor}|U_{iq}|^2 - s\right) \quad , \quad s\in[0,1]$$

converges in distribution to the Brownian bridge.

- If $p=q$, Diaconis and d'Aristotile (99, 06) were interested by partial traces and proved that $\{\sum_{i=1}^{\lfloor ns\rfloor} U_{ii} \ , \ s\in[0,1]\}$ converges without normalization to the Brownian motion.

As usual, the proof is divided in two parts : convergence of the fi.di. distributions of $W^{(n)}$ and tightness. The main tool is the computation of cumulants and their asymptotics. We state a formula for the cumulants of variables of the form $X = Tr(AUBU^*)$ for deterministic matrices A, B of size n, and we apply it to the computation of the second and fourth cumulant of $T_{p,q}$. This formula relies on the notion of second order freeness introduced by Mingo, Sniady and Speicher (06-07). Roughly speaking, whereas the freeness, introduced by Voiculescu, provides the asymptotic behavior of expectation of traces of random matrices, the second order freeness describes the leading order of the fluctuations of these traces.

Outline

Preliminary remarks : Some moments

Elementary computations give

$$\mathbb{E}|U_{ij}|^{2k} = \frac{(n-1)!k!}{(n-1+k)!}$$

$$\mathbb{E}\left(|U_{i,j}|^2|U_{i,k}|^2\right) = \frac{1}{n(n+1)} \; , \; \mathbb{E}\left(|U_{i,j}|^2|U_{k,\ell}|^2\right) = \frac{1}{n^2-1} \, .$$

From these relations, we can compute the first moments of $T_{p,q}$.

$$\mathbb{E}\, T_{p,q} = \frac{pq}{n} \quad , \quad \lim_n \frac{1}{n}\mathbb{E}\, T_{p,q} = st \, .$$

$$\mathrm{Var}\; T_{p,q} = pq\frac{(n-p)(n-q)}{n^2} \quad , \quad \lim_n \mathrm{Var}\; T_{p,q} = st(1-s)(1-t) \, .$$

Outline

Combinatorics of the unitary group

The expectations of products of entries of U can be described by a special function, called the Weingarten function (see [5]) defined as follows :

$$\mathrm{Wg}(N,\pi) = \mathbb{E}(U_{11}\dots U_{pp}\bar{U}_{1\pi(1)}\dots\bar{U}_{p\pi(p)})$$

where $\pi \in \mathcal{S}_p$. Then,

$$\begin{aligned} &\mathbb{E}(U_{i'_1 j'_1}\dots U_{i'_p j'_p}\bar{U}_{i_1 j_1}\dots\bar{U}_{i_p j_p}) \\ &\quad = \sum_{\alpha,\beta\in\mathcal{S}_p} \delta_{i_1 i'_{\alpha(1)}}\dots\delta_{i_p i'_{\alpha(p)}}\delta_{j_1 j_{\beta(1)}}\dots\delta_{j_p i'_{\beta(p)}}\,\mathrm{Wg}(N,\beta\alpha^{-1})\,. \end{aligned} \tag{1}$$

The Weingarten functions for $p = 1, 2$ are given by :

$$\begin{aligned} \mathrm{Wg}(n,(1)) &= \frac{1}{n} \\ \mathrm{Wg}(n,(1)(2)) &= \frac{1}{n^2-1}\ ,\ \mathrm{Wg}(n,(12)) = -\frac{1}{n(n^2-1)} \end{aligned} \tag{2}$$

Cumulants of random variables

$$\kappa_r(a_1, \cdots, a_r) = \sum_{C \in \mathcal{P}(r)} \text{Möb}(C, 1_r) \mathbb{E}_C(a_1, \cdots, a_r)$$

where

- $\mathcal{P}(r)$ is the set of partitions of $[[r]]$
- If $C = \{C_1, \cdots, C_k\}$ is the decomposition of C in blocks, then

$$\text{Möb}(C, 1-r) = (-1)^{k-1}(k-1)! \ , \ \mathbb{E}_C(a_1, \ldots a_r) = \prod_{i=1}^{k} \mathbb{E}(\prod_{j \in C_i} a_j).$$

Cumulants of random matrices

Cumulants of random matrices If $X_1, \dots X_{2l}$ are random matrices, for $\pi = \pi_1 \times \cdots \times \pi_r \in \mathfrak{S}_{2l}$ with $\pi_i = (\pi_{i,1}, \dots, \pi_{i,\ell(i)})$ let

$$\kappa_\pi(X_1, \dots, X_{2l}) := \kappa_r \left(\mathrm{Tr}(X_{\pi_{1,1}} \cdots X_{\pi_{1,\ell(1)}}), \dots, \mathrm{Tr}(X_{\pi_{r,1}} \cdots X_{\pi_{r,\ell(r)}}) \right)$$

Cumulants of random matrices If $X_1, \dots X_{2l}$ are random matrices, for $\pi = \pi_1 \times \cdots \times \pi_r \in \mathfrak{S}_{2l}$ with $\pi_i = (\pi_{i,1}, \dots, \pi_{i,\ell(i)})$ let

$$\kappa_\pi(X_1, \dots, X_{2l}) := \kappa_r \left(\mathrm{Tr}(X_{\pi_{1,1}} \cdots X_{\pi_{1,\ell(1)}}), \dots, \mathrm{Tr}(X_{\pi_{r,1}} \cdots X_{\pi_{r,\ell(r)}}) \right)$$

For $A = \{A_1, \dots, A_k\}$ a σ-invariant partition of $[[2l]]$ let $\sigma_i = \sigma|_{A_i}$ and

$$\kappa_{\sigma,A}(X_1, \dots, X_{2l}) := \kappa_{\sigma_1}(X_1, \dots, X_{2l}) \cdots \kappa_{\sigma_k}(X_1, \dots, X_{2l}) \,.$$

Cumulants of random matrices If $X_1, \dots X_{2l}$ are random matrices, for $\pi = \pi_1 \times \cdots \times \pi_r \in \mathfrak{S}_{2l}$ with $\pi_i = (\pi_{i,1}, \dots, \pi_{i,\ell(i)})$ let

$$\kappa_\pi(X_1, \dots, X_{2l}) := \kappa_r \left(\mathrm{Tr}(X_{\pi_{1,1}} \cdots X_{\pi_{1,\ell(1)}}), \dots, \mathrm{Tr}(X_{\pi_{r,1}} \cdots X_{\pi_{r,\ell(r)}}) \right)$$

For $A = \{A_1, \dots, A_k\}$ a σ-invariant partition of $[[2l]]$ let $\sigma_i = \sigma|_{A_i}$ and

$$\kappa_{\sigma,A}(X_1, \dots, X_{2l}) := \kappa_{\sigma_1}(X_1, \dots, X_{2l}) \cdots \kappa_{\sigma_k}(X_1, \dots, X_{2l}) .$$

A sequence $\{B_1, \dots, B_s\}_n$ a of $n \times n$ deterministic matrices is said to have a limit distribution if there exists a non commutative probability space $(\mathcal{A}, \varphi)$ and $b_1, \dots b_s \in \mathcal{A}$ such that for any polynomial p in s non commuting variables,

$$\lim_{n \to \infty} n^{-1} \mathrm{Tr}(p(B_1, \dots, B_s)) = \varphi p(b_1, \dots, b_s).$$

Proposition (From Mingo, Sniady, Speicher)

Let $U_n \in \mathbb{U}(n)$ Haar distributed and $\{B_1, \dots, B_s\}_n$ a sequence with a limit distribution. Let $r > 1$ and $\epsilon_1, \dots \epsilon_{2r} \in \{-1, 1\}$ such that $\sum \epsilon_i = 0$. Consider $p_1, \dots p_{2r}$ polynomials in s non commuting variables. Let

$$D_i = p_i(B_1, \dots, B_s) \quad , \quad X_j = \mathrm{Tr}(D_{2j-1} U^{\epsilon(2j-1)} D_{2j} U^{\epsilon(2j)}) \, ,$$

$(i \le 2r, 1 \le j \le r)$. Then,

$$\kappa_r(X_1, \dots, X_r) = \sum_{\pi \in \mathcal{S}_{2r}^{(\epsilon)}} \sum_{A,B} C_{\tilde{\pi}, \tilde{A}} \, \kappa_{\gamma\pi^{-1}, B}(D_1, \dots, D_{2r}) \qquad (3)$$

Moreover, for $r \ge 3$,

$$\lim_{n \to \infty} \kappa_r(X_1, \dots, X_r) = 0.$$

Above, the second sum is taken over pairs of partitions of $[[2r]]$ such that A is π invariant, B is $\gamma\pi^{-1}$ invariant and $A \vee B = 1_{[[2r]]}$ the one block partition. γ is given by the product of transpositions $\prod_{i \leq r}(2i-1, 2i)$ and $C_{\pi,A}$ are relative cumulants :

$$C_{\pi,A} = \sum_{C \in [\pi,A], C=\{V_1,\dots V_k\}} \text{Möb}(C,A)\,\text{Wg}(\pi|_{V_1})\dots\text{Wg}(\pi|_{V_k}) \quad (4)$$

for A π invariant. The other expressions are too complicated to be exposed here.

Outline

1. Introduction
   - Motivation
   - Main result
   - Previous related results
2. Sketch of the proof
   - Preliminary remarks
   - Combinatorics of the unitary group
   - Fidi convergence
   - Tightness
3. Complementary remarks
   - The marginals
   - Orthogonal case (in progress)
   - Conjectured universality

Fi.di. convergence

Let $(a_i)_{i\leq k} \in \mathbb{R}$, $(s_i, t_i)_{i\leq k} \in [0,1]^2$. We must prove the convergence in distribution of $X^{(n)} = \sum_{i=1}^k a_i Y^{(n)}_{p_i,q_i}$ with $p_i = \lfloor ns_i \rfloor$, $q_i = \lfloor nt_i \rfloor$ to a Gaussian distribution.
We have

$$X^{(n)} = \sum_{i=1}^{k} a_i[\mathrm{Tr}(D_{2i-1}UD_{2i}U^\star) - \mathbb{E}(\mathrm{Tr}(D_{2i-1}UD_{2i}U^\star))]$$

where $D_{2i-1} = I_{p_i}$, $D_{2i} = I_{q_i}$. Now, $\{D_{2i-1}, D_{2i}, i = 1, \dots k\}$ are commuting projectors with a limit distribution $\{q_{2i-1}, q_{2i}, i = 1, \dots k\}$ on a probability space $(\mathcal{A}, \phi)$ with $\phi(q_{2i-1}) = s_i$, $\phi(q_{2i}) = t_i$ and $q_i q_j = q_i$ if $u_i \leq u_j$ (and $= q_j$ otherwise) where $u_i = s_i$ for i odd and $u_i = t_i$ for i even.

Let $r \geq 3$, then

$$\begin{aligned}\kappa_r(X^{(n)},\ldots,X^{(n)}) &= \sum_{i_1,\ldots,i_r=1}^{k} a_{i_1}\ldots a_{i_r}\kappa_r(Y^{(n)}_{p_{i_1},q_{i_1}},\ldots,Y^{(n)}_{p_{i_r},q_{i_r}}) \\ &= \sum_{i_1,\ldots,i_r=1}^{k} a_{i_1}\ldots a_{i_r}\kappa_r(X_{i_1},\ldots,X_{i_r})\end{aligned}$$

where $X_{i_p} = \mathrm{Tr}(D_{2i_p-1}UD_{2i_p}U^\star)$. From Proposition 2.1

$$\lim_{n\to\infty}\kappa_r(X_{i_r},\ldots,X_{i_r}) = 0. \tag{5}$$

Now, the second cumulant is given by

$$\kappa_2(X^{(n)},X^{(n)}) = \sum_{i,j=1}^{k} a_i a_j \kappa_2(\mathrm{Tr}(D_{2i-1}UD_{2i}U^\star),\mathrm{Tr}(D_{2j-1}UD_{2j}U^\star)).$$

$$\kappa_2(\mathrm{Tr}(D_{2i-1}UD_{2i}U^\star),\mathrm{Tr}(D_{2j-1}UD_{2j}U^\star)) = \frac{(p_i\wedge p_j)(q_i\wedge q_j)}{n^2-1} - \frac{(p_i\wedge p_j)q_iq_j}{n(n^2-1)} - \frac{p_ip_j(q_i\wedge q_j)}{n(n^2-1)} + \frac{p_ip_jq_iq_j}{n^2(n^2-1)}.$$

In the limit, we get

$$\begin{aligned}\lim_n \kappa_2(\mathrm{Tr}(D_{2i-1}UD_{2i}U^\star), \mathrm{Tr}(D_{2j-1}UD_{2j}U^\star)) \\ = \quad (s_i \wedge s_j - s_i s_j)(t_i \wedge t_j - t_i t_j).\end{aligned}$$

Thus, we get the convergence of $X^{(n)}$ to a centered Gaussian distribution with variance

$$\sum_{i,j=1}^{k} a_i a_j (s_i \wedge s_j - s_i s_j)(t_i \wedge t_j - t_i t_j).$$

Outline

1. Introduction
   - Motivation
   - Main result
   - Previous related results
2. Sketch of the proof
   - Preliminary remarks
   - Combinatorics of the unitary group
   - Fidi convergence
   - Tightness
3. Complementary remarks
   - The marginals
   - Orthogonal case (in progress)
   - Conjectured universality

Tightness

Let $p \leq p' \leq n$ and $q \leq q' \leq n$

$$\begin{aligned}
\Delta^{(n)}_{p,q}(p',q') &= Y^{(n)}_{p',q'} - Y^{(n)}_{p',q} - Y^{(n)}_{p,q'} + Y^{(n)}_{p,q} \\
&= \sum_{p+1 \leq i \leq p'} \sum_{q+1 \leq i \leq q'} |U_{i,j}|^2 - \mathbb{E}|U_{i,j}|^2 \\
&\overset{\text{(d)}}{=} Y^{(n)}_{p'-p,q'-q} \,.
\end{aligned}$$

A criterion adapted from Bickel and Wichura says that it is enough to prove

$$\mathbb{E}\left[\left(Y^{(n)}_{p,q}\right)^4\right] = O(p^2 q^2 n^{-4})\,. \tag{6}$$

We now give an estimate for $\kappa_4(T_{p,q})$ which, helped by the above estimates, will be sufficient. From (3),

$$\kappa_4 = \sum_{\pi \in S_8^{(\epsilon)}} \sum_{A,B} C_{\tilde{\pi},\tilde{A}}\, \kappa_{\gamma\pi^{-1},B}(D_1, \ldots, D_8)$$

where $\mathcal{S}_8^{(\epsilon)}$ is the subset of $\mathcal{S}_8$ which sends $\{1,3,5,7\}$ onto $\{2,4,6,8\}$ and reversely, $\gamma = (12)(34)(56)(78) \in \mathcal{S}_8$, A and B are partitions of $[[8]]$ such that A is π-invariant, B is $\gamma\pi^{-1}$-invariant, $A \vee B = 1_{[[8]]}$, and finally

$$D_1 = D_3 = D_5 = D_7 = I_p\ ,\ \ D_2 = D_4 = D_6 = D_8 = I_q\,.$$

Outline

Asymptotics of the marginal

Let us recall the notation

$$A_{p,q} = D_1 U D_2 U^* = V_{p,q} V_{p,q}^*$$

where $V_{p,q}$ is the upper-left submatrix of U. As proved by Collins (2005) $A_{p,q}$ belongs to the Jacobi unitary ensemble (JUE) and

$$T_{p,q} = \mathrm{Tr}\, A_{p,q} = p \int x d\mu^{(p)}(x) \, ,$$

where $\mu^{(p)}$ is the empirical spectral distribution

$$\mu^{(p)} = \frac{1}{p} \sum_{k=1}^{p} \delta_{\lambda_k^{(p)}} \, ,$$

and the $\lambda_k^{(p)}$'s are the eigenvalues of $A_{p,q}$.

For the JUE, the equilibrium measure is the Kesten-McKay distribution of density

$$C_{u_-,u_+}\frac{\sqrt{(x-u_-)(u_+-x)}}{2\pi(4-x^2)}\mathbf{1}_{(u_-,u_+)}(x) \tag{7}$$

where $-2 \le u_- < u_+ \le 2$ ($u_\pm$ depending on s, t). By continuity, we recover

$$\lim_n \frac{1}{n} T_{\lfloor ns\rfloor,\lfloor nt\rfloor} = s\int x\pi_{u_-,u_+}(x)dx = st\,,$$

in probability. It could also be possible to recover the fluctuation result for the marginal distribution, i.e. $T_{\lfloor ns\rfloor,\lfloor nt\rfloor} - \mathbb{E}\,T_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ converges in distribution to $\mathcal{N}(0, s(1-s)t(1-t))$ from the known results on the fluctuations of linear statistics of $\mu^{(p)}$.

Outline

Orthogonal case (in progress)

$\mathbb{U}(n) \to \mathbb{O}(n)$ $T_{p,q} = \sum_{i \le p, j \le q} U_{ij}^2$

In multivariate (real) analysis of variance, $T_{p,q}$ is known as the Bartlett-Nanda-Pillai statistics., used to test equalities of covariances matrices from Gaussian populations. Asymptotic studies :

1. p, q fixed, $n \to \infty$ (large sample framework),
2. q fixed, $n, p \to \infty$ and $p/n \to s < 1$ fixed (high-dimensional framework, see Fujikoshi et al.).
3. $p/n \to s, q/n \to t$ with s, t fixed. This case is considered in the Bai and Silverstein's book, and a CLT for $T_{p,q}$ was proved by Bai, Jiang, Yao, Zhang (2009).

Outline

One conjecture of D. Chafai

If M is a $n \times n$ matrix with i.i.d. entries with the same four first moments as the Gaussian standard then the matrix U of the eigenvectors of MM^* satisfy the same asymptotic result as in our theorem.

Bibliography I

Z. Bai and J.W. Silverstein.
Springer Series in Statistics, Springer, New York, 2010.

Z. Bai, D. Jiang, J-F. Yao, and S. Zheng.
Ann. Statist., 37, 3822–3840, 2009.

P.J. Bickel and M.J. Wichura.
Ann. Math. Statist., 42(5) :1656–1670, 1971.

G. Chapuy.
Discrete Math. Theor. Comput. Sci. Proc. AH 415–426, 2007.

B. Collins.
Int. Math. Res. Not., 17, 953–982, 2003.

B. Collins.
Probab. Theory Related Fields, 133, 315–344, 2005.

Bibliography II

B. Collins and P. Sniady.
Comm. Math. Phys., 264, 773–795, 2006.

Y Fujikoshi, T Himeno and H. Wakaki.
J. Statist. Plann. Inference, 138 3457–3466, 2008.

F. Hiai and D. Petz.
Amer. Math. Soc., Providence, 2000.

K. Johansson.
Duke Math. J., 91, 151–204, 1998.

J.A. Mingo and R. Speicher.
J. Funct. Anal., 235 (1), 226–270, 2006.

J.A. Mingo, P. Śniady, and R. Speicher.
Adv. Math., 209(1), 212–240, 2007.

Bibliography III

R.J. Muirhead.
Aspects of multivariate statistical theory, John Wiley, 1982.

J.W. Silverstein.
SIAM Journal on Math. Anal., 12, 1981

J.W. Silverstein.
Annals of Probab., 18(3), 1174–1194, 1990.

D. Voiculescu.
Invent. Math., 104(1), 201–220, 1991.