Lithuanian Mathematical Journal, Vol. 50, No. 2, 2010, pp. 121–132

ASYMPTOTIC DISTRIBUTION OF SINGULAR VALUES OF POWERS OF RANDOM MATRICES*

N. Alexeev [a], **F. Götze** [b], **and A. Tikhomirov** [c]

[a] Faculty of Mathematics and Mechanics, St. Petersburg State University, St. Petersburg, Russia
[b] Faculty of Mathematics, University of Bielefeld, Germany
[c] Department of Mathematics, Komi Research Center of Ural Branch of RAS, Syktyvkar State University, Syktyvkar, Russia
(e-mail: nikita.v.alexeev@gmail.com; goetze@math.uni-bielefeld.de; tichomir@math.uni-bielefeld.de)

Received November 24, 2009

Abstract. Let x be a complex random variable such that $\mathbf{E}x = 0$, $\mathbf{E}|x|^2 = 1$, and $\mathbf{E}|x|^4 < \infty$. Let x_{ij}, $i, j \in \{1, 2, \dots\}$, be independent copies of x. Let $\mathbf{X} = (N^{-1/2} x_{ij})$, $1 \leqslant i, j \leqslant N$, be a random matrix. Writing $\mathbf{X}^*$ for the adjoint matrix of $\mathbf{X}$, consider the product $\mathbf{X}^m \mathbf{X}^{*m}$ with some $m \in \{1, 2, \dots\}$. The matrix $\mathbf{X}^m \mathbf{X}^{*m}$ is Hermitian positive semidefinite. Let $\lambda_1, \lambda_2, \dots, \lambda_N$ be eigenvalues of $\mathbf{X}^m \mathbf{X}^{*m}$ (or squared singular values of the matrix $\mathbf{X}^m$). In this paper, we find the asymptotic distribution function $G^{(m)}(x) = \lim_{N\to\infty} \mathbf{E} F_N^{(m)}(x)$ of the empirical distribution function $F_N^{(m)}(x) = N^{-1} \sum_{k=1}^N \mathbb{I}\{\lambda_k \leqslant x\}$, where $\mathbb{I}\{A\}$ stands for the indicator function of an event A. With $m = 1$, our result turns to a well-known result of Marchenko and Pastur [V. Marchenko and L. Pastur, The eigenvalue distribution in some ensembles of random matrices, *Math. USSR Sb.*, 1:457–483, 1967].

MSC: 60F05, 15B52

Keywords: random matrices, Fuss–Catalan numbers, semi-circular law, Marchenko–Pastur distribution

1 INTRODUCTION

Let $\mathbf{X} = (N^{-1/2} x_{ij}^{(N)}), 1 \leqslant i, j \leqslant N$, be a random matrix. We assume that $x_{ij} \equiv x_{ij}^{(N)}$ are independent complex random variables such that

$$\mathbf{E}x_{ij} = 0, \qquad \mathbf{E}|x_{ij}|^2 = 1, \qquad \mathbf{E}|x_{ij}|^4 \leqslant B, \tag{1.1}$$

with some $B < \infty$ independent of N. We assume additionally that

$$L_N(\alpha) = N^{-2} \sum_{1\leqslant i,j\leqslant N} \mathbf{E}|x_{ij}|^4 \mathbb{I}\{|x_{ij}| > \alpha\sqrt{N}\} \to 0 \quad \text{as } N \to \infty \tag{1.2}$$

* Partially supported by the RF grant of the Leading Scientific Schools (No. NSH-4472.2010.1), RFBR (grant No. 09-01-12180), RFBR–DFG (grant No. 09-01-91331), and CRC 701 “Spectral Structures and Topological Methods in Mathematics”, Bielefeld.

for all $\alpha > 0$. Note that $x_{ij} \equiv x_{ij}^{(N)}$ and $\mathbf{X} \equiv \mathbf{X}^{(N)}$ may depend on N, which is not reflected in our further notation.

Writing $\mathbf{X}^*$ for the adjoint matrix of $\mathbf{X}$, we consider the product

$$\mathbf{W}^{(m)} = \mathbf{X}^m \mathbf{X}^{*m}$$

with some $m \in \{1, 2, \dots\}$. The matrix $\mathbf{W}^{(m)}$ is Hermitian positive semidefinite. Let $\lambda_1, \lambda_2, \dots, \lambda_N$ be eigenvalues of $\mathbf{W}^{(m)}$ (or squared singular values of the matrix $\mathbf{X}^m$). In this paper, we find the asymptotic distribution function

$$G^{(m)}(x) = \lim_{N\to\infty} \mathbf{E} F_N^{(m)}(x)$$

of the empirical distribution function

$$F_N^{(m)}(x) = N^{-1} \sum_{k=1}^{N} \mathbb{I}\{\lambda_k \leqslant x\},$$

where $\mathbb{I}\{A\}$ stands for the indicator function of an event A.

Theorem 1. *Assume that* (1.1) *and* (1.2) *hold. Then the limit* $G^{(m)}(x) = \lim_{N\to\infty} \mathbf{E} F_N^{(m)}(x)$ *exists. The function* $G^{(m)}(x)$ *is a distribution function and has the moments*

$$M_p^{(m)} = \int_{\mathbb{R}} x^p \, \mathrm{d}G^{(m)}(x) = \frac{1}{mp+1} \binom{pm+p}{p}. \tag{1.3}$$

Corollary 1. Let x_{ij} be independent copies of a random variable, say x, such that

$$\mathbf{E}x = 0, \qquad \mathbf{E}|x|^2 = 1, \qquad \mathbf{E}|x|^4 < \infty.$$

Let $\mathbf{X} = (N^{-1/2} x_{ij})$, $1 \leqslant i, j \leqslant N$. Then the limit $\lim_{N\to\infty} \mathbf{E} F_N^{(m)}(x)$ exists and is equal to $G^{(m)}(x)$.

Gessel and Xin [4] showed that, for any natural m, the sequence $M_1^{(m)}, M_2^{(m)}, \dots$ is a sequence of moments of some probability measure. Hence, $G^{(m)}$ is a probability distribution for any natural m. Since $M_p^{(m)} \leqslant c_m^p$ with some $c_m < \infty$, by Carleman's Theorem in [3] the measure $G^{(m)}$ is uniquely determined by its moments. The support of the measure $G^{(m)}$ is the interval $[0, m^{-m}(m+1)^{m+1}]$.

With $m = 1$, Theorem 1 turns to a well-known result of Marchenko and Pastur [6]. Namely, the asymptotic distribution $G^{(1)}$ of eigenvalues of the matrices $\mathbf{X}\mathbf{X}^*$ has the moments $M_p^{(1)} = \frac{1}{p+1}\binom{2p}{p}$. Note that, in the case $m = 1$, our fourth moment assumption is stronger than assumptions in Theorems 2.5 and 2.8 in [1]. The question of the weakest sufficient conditions in the case $m > 1$ remains an open problem.

In Free Probability Theory, $M_p^{(m)}$ are known as Fuss–Catalan numbers. Combinatorial properties of this sequence have been studied by Nica and Speicher [8]. Mlotkowski [7] investigated a family of distributions, say $G^{(m,r)}$, with real $m \geqslant 0$ and $0 \leqslant r \leqslant m$, such that $G^{(m,r)}$ has the moments $\frac{r}{mp+p+r}\binom{mp+p+r}{p}$. It is easy to check that $G^{(m,1)} = G^{(m)}$. Oravecz [9] proved that the powers of Voiculescu's circular element have the distribution $G^{(m)}$. This distribution belongs to the class of Free Bessel Laws (see [2]).

Let $\mathcal{M}_m(x) = \sum_{p=0}^{\infty} M_p^{(m)} x^p$ be the generating function of the sequence $M_p^{(m)}$. It satisfies the following functional equation (see Eq. (7.68) on p. 347 in [5]):

$$\mathcal{M}_m(x) = 1 + x\mathcal{M}_m^{m+1}(x). \tag{1.4}$$

Equation (1.4) allows us to describe $G^{(m)}$ in the framework of Free Probability Theory. In Free Probability Theory, the free multiplicative convolution $\xi \boxtimes \eta$ is defined for any positive random variables ξ and η (see [8], p. 287). The S-transform is a homomorphism with respect to free multiplicative convolution, i.e., if ξ and η are free independent positive variables, then $S_{\xi\boxtimes\eta}(z) = S_\xi(z)S_\eta(z)$. Recall that the S-transform, say $S(z)$, of a distribution μ is defined as follows. Let

$$M_p = \int\limits_{\mathbb{R}} x^p \,\mathrm{d}\mu(x), \qquad u(z) = \sum_{p=1}^{\infty} M_p z^p.$$

Then

$$S(z) = \frac{z+1}{z} u^{-1}(z), \tag{1.5}$$

where u^{-1} denotes the inverse function of u.

Equation (1.4) allows us to calculate the S-transform, say $S^{(m)}(z)$, of $G^{(m)}$:

$$S^{(m)}(z) = \frac{1}{(1+z)^m}. \tag{1.6}$$

This means that the family $G^{(m)}$ has the following property: if a random variable ξ has the distribution $G^{(m)}$, then the rth power of the S-transform of ξ is equal to the S-transform of the multiplicative free power $\xi^{\boxtimes r}$. This property holds for this family of distributions only.

To prove Theorem 1, we use truncation and the method of moments. Truncation means that we can replace (see Section 2.1 for details) $\mathbf{X}$ by the matrix $\widetilde{\mathbf{X}} = (\widetilde{X}_{ij})$ with truncated entries (here and below, $X_{ij} = N^{-1/2} x_{ij}$ denote the entries of matrix $\mathbf{X}$):

$$\widetilde{X}_{ij} = X_{ij}\mathbb{I}\{|X_{ij}| < \alpha_N\}, \tag{1.7}$$

where α_N is some sequence of positive numbers such that $\alpha_N \to 0$ as $N \to \infty$. Lemma 1 (see Section 2.1) reduces the proof of Theorem 1 to the proof of the following proposition.

Proposition 1. *Assume that $\alpha_N \to 0$ and $\beta_N \to 0$. Then Theorem 1 holds if*

$$\bigl|X_{ij}^{(N)}\bigr| \leqslant \alpha_N, \qquad \max_{1\leqslant i,j\leqslant N}\bigl|\mathbf{E}X_{ij}^{(N)}\bigr| \leqslant \beta_N N^{-3/2}, \qquad \bigl|\mathbf{E}\bigl|X_{ij}^{(N)}\bigr|^2 - 1/N\bigr| \leqslant \beta_N N^{-3/2}. \tag{1.8}$$

Let us explain our proof of Proposition 1. Denote by $\xi_m(N)$ a random variable with distribution $\mathbf{E}F_N^{(m)}$. We show that the moments $\mathbf{E}\xi_m^p(N)$ converge to $M_p^{(m)}$. In order to simplify the notation, assume for a while that X_{ij} are real random variables. Then one can represent $\mathbf{E}\xi_m^p(N)$ as

$$\mathbf{E}\xi_m^p(N) = \sum\nolimits^{(2mp)} N^{-1}\mathbf{E}\prod_{j=0}^{2mp-1} X_{i_j i_{j+1}}^{\varepsilon(j)},$$

where the sum $\sum^{(2mp)}$ is taken over $i_0, \dots, i_{2mp} \in \{1, \dots, N\}$ such that $i_{2mp} = i_0$. The notation $X_{i_j i_{j+1}}^{\varepsilon(j)}$ means that $X_{i_j i_{j+1}}^{+} := X_{i_j i_{j+1}}$ in the case $\varepsilon(j) = +$ and $X_{i_j i_{j+1}}^{-} := X_{i_{j+1} i_j}$ in the case $\varepsilon(j) = -$ (see Section 2.2 for a precise definition of the spin variable $\varepsilon(j)$). We investigate properties of the paths $(i_0, \dots, i_{2mp})$ by combinatorial methods. The moment $\mathbf{E}\xi_m^p(N)$ converges to the number of paths of a special type. Namely, one can describe such paths as follows: the cardinality of $\{i_0, \dots, i_{2mp}\}$ is equal to $mp+1$, and each factor $X_{i_j i_{j+1}}$ appears in the product $\prod_{j=0}^{2mp-1} X_{i_j i_{j+1}}^{\varepsilon}$ twice. In Section 2.4, we count the number of these paths.

2 THE PROOF OF THE MAIN RESULT

2.1 Truncation

Recalling that $X_{ij} = N^{-1/2}x_{ij}$, we can rewrite $L_N(\alpha)$ as

$$L_N(\alpha) = \sum_{1\leqslant i,j\leqslant N} \mathbf{E}|X_{i,j}|^4\mathbb{I}\{|X_{i,j}| > \alpha\}.$$

Since, for all $\alpha > 0$, the ratio $L_N(\alpha)/\alpha^4$ tends to 0, one can find a sequence $\alpha_N \downarrow 0$ such that $L_N(\alpha_N)/\alpha_N^4 \to 0$ and $N^\delta \alpha_N^{-1} \to \infty$ for any $\delta > 0$ as $N \to \infty$. Let $\widetilde{F}_N^{(m)}(t)$ denote the empirical spectral distribution function of the matrix $\widetilde{\mathbf{X}}^m\widetilde{\mathbf{X}}^{*m}$.

Lemma 1. *The limit behaviors of* $\mathbf{E}\widetilde{F}_N^{(m)}(t)$ *and* $\mathbf{E}F_N^{(m)}(t)$ *are the same, that is,*

$$\sup_{t\in\mathbb{R}}\bigl|\mathbf{E}\widetilde{F}_N^{(m)}(t) - \mathbf{E}F_N^{(m)}(t)\bigr| \to 0.$$

Proof. Since by definition $|\widetilde{F}_N^{(m)}(t) - F_N^{(m)}(t)| \neq 0$ only if there exist $i, j \in \{1,\dots,N\}$ such that $|X_{ij}| \geqslant \alpha_N$, we have

$$\bigl|\mathbf{E}\widetilde{F}_N^{(m)}(t) - \mathbf{E}F_N^{(m)}(t)\bigr| \leqslant \sum_{1\leqslant i,j\leqslant N} \mathbf{P}\bigl(|X_{ij}| \geqslant \alpha_N\bigr). \tag{2.1}$$

Estimating $\mathbf{P}(|X_{ij}| \geqslant \alpha_N) \leqslant \alpha_N^{-4}\mathbf{E}|X_{ij}|^4\mathbb{I}\{|X_{ij}| > \alpha_N\}$ and using inequality (2.1), we obtain

$$\sup_{t\in\mathbb{R}}\bigl|\mathbf{E}\widetilde{F}_N^{(m)}(t) - \mathbf{E}F_N^{(m)}(t)\bigr| \leqslant \alpha_N^{-4}\sum_{i,j=1}^{N} \mathbf{E}|X_{ij}|^4\mathbb{I}\bigl\{|X_{ij}| > \alpha_N\bigr\} = \alpha_N^{-4}L_N(\alpha_N) \to 0. \qquad \square \tag{2.2}$$

Note, that the lower-order moments of the truncated variables are asymptotically equal to the moments of the original variables. Writing for a while $X = X_{ij}$, we have, for $k \leqslant 3$,

$$\bigl|\mathbf{E}\widetilde{X}^k - \mathbf{E}X^k\bigr| \leqslant \mathbf{E}|X|^k\mathbb{I}\bigl\{|X| > \alpha_N\bigr\}. \tag{2.3}$$

The right-hand side of (2.3) can be estimated as

$$\mathbf{E}|X|^k\mathbb{I}\bigl\{|X| > \alpha_N\bigr\} \leqslant \alpha_N^{k-4}\mathbf{E}|X|^4 \leqslant \beta_N N^{-3/2}, \tag{2.4}$$

where $\beta_N = B\alpha_N^{k-4}N^{-1/2} \to 0$ as $N \to \infty$.

Lemma 1 shows that the limit behaviors of $\widetilde{F}_N^{(m)}(t)$ and $F_N^{(m)}(t)$ are the same. Thus, we may replace $\mathbf{X}$ by $\widetilde{\mathbf{X}}$ in the following arguments and assume that $\mathbf{X}$ is truncated, that is, that the entries of $\mathbf{X}$ satisfy assumption (1.8).

2.2 Moments of the spectral distribution

We apply the method of moments. Recall that $\lambda_1, \lambda_2, \dots, \lambda_N$ denote the eigenvalues of $\mathbf{X}^m\mathbf{X}^{*m}$. We can write

$$\mathbf{E}\xi_m^p(N) = N^{-1}\mathbf{E}\sum_{j=1}^{N}\lambda_j^p = N^{-1}\mathbf{E}\,\mathrm{Tr}\bigl(\mathbf{X}^m\mathbf{X}^{*m}\bigr)^p. \tag{2.5}$$

We assume that m and p are fixed and study the asymptotics of $\mathbf{E}\xi_m^p(N)$ as $N \to \infty$. In order to simplify the notation, henceforth we assume that X_{ij} are real random variables.

In the Hermitian case, the trace of $\mathbf{X}^{2k}$ can be rewritten in terms of the entries of $\mathbf{X}$ via

$$\mathbf{E}\,\mathrm{Tr}\,\mathbf{X}^{2k} = \sum\nolimits^{(2k)} \mathbf{E} \prod_{j=0}^{2k-1} X_{i_j i_{j+1}}, \tag{2.6}$$

where the sum $\sum^{(s)}$ is taken over $i_0, \dots, i_s \in \{1, \dots, N\}$ such that $i_s = i_0$.

In the non-Hermitian case, $\mathbf{E}\,\mathrm{Tr}(\mathbf{X}^m \mathbf{X}^{*m})^p$ has a similar representation. An entry of $\mathbf{X}^m \mathbf{X}^{*m}$ is given by

$$\left[\mathbf{X}^m \mathbf{X}^{*m}\right]_{ik} = \sum_{1 \leqslant i_j \leqslant N} X_{i i_1} X_{i_1 i_2} \cdots X_{i_{m-1} i_m} X_{i_{m+1} i_m} \cdots X_{k i_{2m-1}}. \tag{2.7}$$

We write $X^+_{i_j i_{j+1}} := X_{i_j i_{j+1}}$ and $X^-_{i_j i_{j+1}} := X_{i_{j+1} i_j}$. Then the right-hand side of (2.7) takes the form

$$\left[\mathbf{X}^m \mathbf{X}^{*m}\right]_{ik} = \sum_{1 \leqslant i_j \leqslant N} \prod_{j=0}^{2m-1} X^{\varepsilon(j)}_{i_j i_{j+1}}, \tag{2.8}$$

where $i_0 = i$, $i_{2m} = k$, and the "spin" variable $\varepsilon(j)$ takes values $\varepsilon(j) = +$ for $j < m$ and $\varepsilon(j) = -$ for $j \geqslant m$. Since $(\mathbf{X}^m \mathbf{X}^{*m})^p = \mathbf{X}^m \mathbf{X}^{*m} \cdots \mathbf{X}^m \mathbf{X}^{*m}$ (p times), one needs to change the order of indices in $X^{\varepsilon}_{i_j i_{j+1}}$ if the spin $\varepsilon = -$ and

$$\varepsilon(j) = \begin{cases} + & \text{if } j \ (\mathrm{mod}\ 2m) \in \{0, \dots, m-1\}, \\ - & \text{if } j \ (\mathrm{mod}\ 2m) \in \{m, \dots, 2m-1\}. \end{cases} \tag{2.9}$$

Using these notions, (2.5) takes the form

$$\mathbf{E}\xi_m^p(N) = N^{-1}\mathbf{E}\,\mathrm{Tr}\big(\mathbf{X}^m \mathbf{X}^{*m}\big)^p = \sum\nolimits^{(2mp)} N^{-1} \mathbf{E} \prod_{j=0}^{2mp-1} X^{\varepsilon(j)}_{i_j i_{j+1}}. \tag{2.10}$$

A crucial notion in the proof is that of "paths" of indices of type $(i_0, i_1, \dots, i_{2mp-1})$.

2.3 Description of paths

We consider a path $\mathbf{i} = (i_0, \dots, i_{2mp-1})$ which corresponds to a product $\prod_{j=0}^{2mp-1} X^{\varepsilon}_{i_j i_{j+1}}$. Let $\mathcal{P}$ be the set of pairs $\{(j, j+1)^{\varepsilon(j)}\}_{j=0}^{2m-2} \cup \{(2mp-1, 0)^-\}$, where $(j, j+1)^+ := (j, j+1)$, $(j, j+1)^- := (j+1, j)$, and $\varepsilon(j)$ is given by (2.9). We call pairs $(j, j+1)^{\varepsilon(j)}$ and $(k, k+1)^{\varepsilon(k)}$ equivalent (denoted by $(j, j+1)^{\varepsilon(j)} \sim (k, k+1)^{\varepsilon(k)}$) iff $X^{\varepsilon(j)}_{i_j i_{j+1}} \equiv X^{\varepsilon(k)}_{i_k i_{k+1}}$. We also call $(j, j+1)^{\varepsilon(j)}$ an edge of the path $\mathbf{i}$. We construct a directed graph $\mathcal{G}_{\mathbf{i}}$ as follows. A vertex $\mathcal{V}$ of $\mathcal{G}_{\mathbf{i}}$ is a subset of $\{0, 1, \dots, 2mp-1\}$ such that $j \in \mathcal{V}$ and $k \in \mathcal{V}$ if and only if $i_j = i_k$. There exists an edge $(\mathcal{V}, \mathcal{U})$ if and only if there exist $l \in \mathcal{V}$ and $r \in \mathcal{U}$ such that $(l, r) \in \mathcal{P}$ (note that $|l - r| = 1$). Denote by V the total number of vertices of the graph $\mathcal{G}_{\mathbf{i}}$ and by E the total number of its edges. Since the graph $\mathcal{G}$ is connected, $E \geqslant V - 1$. It is clear that V is the cardinality of $\{i_0, i_1, \dots, i_{2m-1}\}$ and E is the cardinality of the quotient set $\mathcal{P}/\sim$. Denote by k_r ($r = 1, \dots, E$) the cardinality of each equivalence class in $\mathcal{P}$. Note that $k_1 + k_2 + \cdots + k_E = 2mp$.

Remark 1. Consider paths $\mathbf{i} = (i_0, \ldots, i_{2mp-1})$ and $\mathbf{k} = (k_0, k_1, \ldots, k_{2m-1})$ such that $\mathcal{G}_{\mathbf{i}} = \mathcal{G}_{\mathbf{k}}$. It is clear that if x_{ij} are identically distributed, then

$$\mathbf{E} \prod_{j=0}^{2mp-1} X_{i_j i_{j+1}}^{\varepsilon(j)} = \mathbf{E} \prod_{j=0}^{2mp-1} X_{k_j k_{j+1}}^{\varepsilon(j)}.$$

We will show that, assuming our conditions, the asymptotic products corresponding to equivalent paths are equal as well.

DEFINITION 1. We define the contribution of a graph $\mathcal{G}$ to (2.10) as

$$Cont(\mathcal{G}) = \sum_{\mathbf{i}:\ \mathcal{G}_{\mathbf{i}} = \mathcal{G}} N^{-1} \mathbf{E} \prod_{j=0}^{2mp-1} X_{i_j i_{j+1}}^{\varepsilon(j)}.$$

Lemma 2. *Using this notation, we have that the contribution of the path $\mathcal{G}$ is asymptotically given by*

$$Cont(\mathcal{G}) \sim N^{V-1} \prod_{r=1}^{E} \mathbf{E} X_{i_s i_t}^{k_r} \tag{2.11}$$

as N tends to infinity.

Proof. Since X_{ij} are independent, we have

$$\mathbf{E} \prod_{j=0}^{2mp-1} X_{i_j i_{j+1}}^{\varepsilon} = \prod_{r=1}^{E} \mathbf{E} X_{i_s i_t}^{k_r}.$$

Furthermore, for any vertex $\mathcal{V}$, the number of possible values of corresponding indices (indices i_j such that $j \in \mathcal{V}$) lies between N and $N - 2mp \sim N$. The lower bound $N - 2mp$ is due to the fact that indices corresponding to this vertex should not coincide with indices corresponding to other vertices and that there are at most $2mp$ different indices. This yields the multiplicity N^V. Together with the factor N^{-1}, this finally leads to formula (2.11). $\square$

DEFINITION 2. We call a graph $\mathcal{G}_{\mathbf{i}}$ an (m,p)-regular graph if it has at least $mp + 1$ vertices and if $k_r \geqslant 2$ for all $r \in \{1, 2, \ldots, E\}$. We call the path $\mathbf{i}$ an (m,p)-regular path.

Lemma 3. *$Cont(\mathcal{G}_{\mathbf{i}})$ does not converge to zero if and only if $\mathcal{G}_{\mathbf{i}}$ is a regular path.*

Proof. Since the variables X_{ij} satisfy conditions (1.8), we have

$$\left|\mathbf{E} X_{ij}^k\right| \leqslant \mathbf{E} X_{ij}^2 |X_{ij}|^{k-2} \leqslant N^{-1} \alpha_N^{k-2}. \tag{2.12}$$

Of course, this estimation also holds for $k = 1$. First, we suppose that one of k_r is equal to 1 (without loss of generality, $k_1 = 1$). Then we have

$$\begin{aligned} \left| \prod_{r=1}^{E} \mathbf{E} X_{i(r)j(r)}^{k_r} \right| &= \left| \mathbf{E} X_{i(1)j(1)} \prod_{r=2}^{E} \mathbf{E} X_{i(r)j(r)}^{k_r} \right| \leqslant \beta_N N^{-3/2} N^{-E+1} \alpha_N^{\sum_r (k_r - 2)} \\ &= \beta_N N^{-3/2} N^{-E+1} \alpha_N^{2mp-1-2(E-1)} \leqslant N^{-E-1/2}, \end{aligned} \tag{2.13}$$

and the contribution of such a graph is bounded by

$$\left|Cont(\mathcal{G}_{\mathbf{i}})\right| \leqslant N^{V-1}N^{-E-1/2} = N^{V-E-1}N^{-1/2}. \tag{2.14}$$

Note that $V - E - 1 \leqslant 0$ since the graph $\mathcal{G}$ is connected, and hence, $N^{V-E-1}N^{-1/2}$ tends to 0.

Furthermore, we consider the case $V < mp + 1$. Note that $k_r \geqslant 2$ for any r and $E \leqslant 2mp/2 = mp$. Our truncation leads to

$$\left|\prod_{r=1}^{E} \mathbf{E}X_{i(r)j(r)}^{k_r}\right| \leqslant N^{-E}\alpha_N^{\sum_r (k_r-2)} = N^{-E}\alpha_N^{2mp-2E}. \tag{2.15}$$

Using inequality (2.15) to estimate the terms in (2.11), we obtain, for such a product,

$$N^{V-1}\left|\prod_{r=1}^{E} \mathbf{E}X_{i(r)j(r)}^{k_r}\right| \leqslant N^{V-E-1}\alpha_N^{2mp-2E}. \tag{2.16}$$

Note that $E \geqslant V - 1$ and $2mp - 2E \geqslant 0$. It follows that the right-hand side of (2.16) does not converge to 0 only if $2mp - 2E = 0$ and $V - E - 1 = 0$, i.e., $V = mp + 1$, and the graph $\mathcal{G}$ is a regular graph. □

Furthermore, we obtain the following:

Lemma 4. *A regular graph is a tree and has exactly $V = mp + 1$ vertices and exactly $E = mp$ edges (each representing an equivalence class of size $k_r = 2$).*

Remark 2. Due to the fact that $\mathbf{E}X_{ij}^2 \sim 1/N$ and by the remarks above, we can write the contribution of a regular graph $\mathcal{G}_{\mathrm{reg}}$ as

$$Cont(\mathcal{G}_{\mathrm{reg}}) \sim 1. \tag{2.17}$$

We now show the connection between the moments of the spectral distribution $\mathbf{E}F_m^{(N)}$ and the number of regular graphs. Indeed, $\xi_m(N)$ has distribution $\mathbf{E}F_m^{(N)}$. Denote by $T_{m,p}$ the set of all possible graphs of view $\mathcal{G}_{\mathbf{i}}$ and by $T_{m,p}^{\mathrm{reg}}$ the set of all (m,p)-regular graphs. Then

$$\mathbf{E}\xi_m^p(N) = \sum_{\mathcal{S}\in T_{m,p}} Cont(\mathcal{S}) \sim \sum_{\mathcal{S}\in T_{m,p}^{\mathrm{reg}}} 1 = \#T_{m,p}^{\mathrm{reg}}. \tag{2.18}$$

We can reformulate (2.18) as follows.

Lemma 5. $\lim_{N\to\infty} \mathbf{E}\xi_N^p$ *is equal to the number of (m,p)-regular graphs.*

2.4 Counting of the number of regular graphs

Lemma 6. *The number of all (m,p)-regular graphs is $\#T_{m,p}^{\mathrm{reg}} = M_p^{(m)}$.*

Proof. The numbers $M_p^{(m)} = \frac{1}{mp+1}\binom{mp+p}{p}$ satisfy the recurrence (see [5])

$$M_p^{(m)} = \sum_{p-1} \prod_{i=0}^{m-1} M_{p_i}^{(m)}, \qquad M_1^{(m)} = 1, \tag{2.19}$$

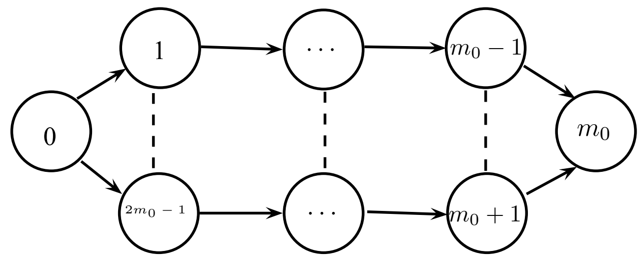

Figure 1. The path **i**. Vertices that correspond to equal indices are connected via dashed lines.

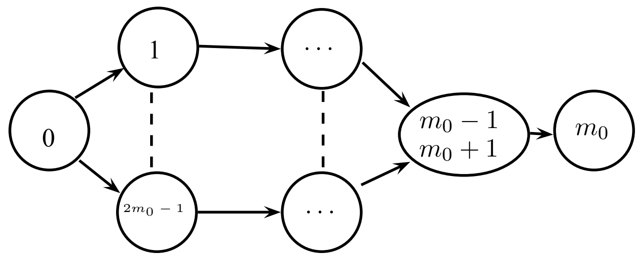

Figure 2. The path **j** is an $(m_0-1,1)$-path.

where the sum $\sum_{p-1}$ is taken over all $p_0+p_1+\cdots+p_m = p-1$. We will show that there is a one-to-one correspondence between collections of (m,p_k)-regular graphs $(\mathcal{G}_{m,p_0},\ldots,\mathcal{G}_{m,p_m})$: $\sum_{i=0}^m p_i = p-1$ and (m,p)-regular graphs $\mathcal{G}_{m,p}$. It follows that

$$\#T^{\text{reg}}_{m,p} = \#\bigcup_{p-1}\bigotimes_{i:=0}^{m} T^{\text{reg}}_{m,p_i} = \sum_{p-1}\prod_{i:=0}^{m}\#T^{\text{reg}}_{m,p_i} \tag{2.20}$$

and that the sequence $\#T^{\text{reg}}_{m,p}$ satisfies both the same recurrence and initial conditions as the sequence of Fuss–Catalan numbers $M_p^{(m)}$, and, therefore, these two sequences are equal.

Proposition 2. *The number* $\#T^{\text{reg}}_{m,1} = 1$ *for all* m. *If* $\mathcal{G}_{\mathbf{i}}$ *is an* $(m,1)$*-regular graph, then indices* i_k *and* i_l *are equal iff* $(k+l) = 2m$.

Proof. By induction. Consider $m=1$. In this case, it is clear that there is only one regular graph $0\to 1$ and Proposition 2 holds. Assume that Proposition 2 holds for all $m < m_0$. Consider the path $\mathbf{i}$ and the corresponding graph $\mathcal{G}_{\mathbf{i}}$ (see Fig. 1). This path has m_0+1 distinct indices, and it has $2m_0$ in total. It follows that there exist at least two one-element vertices of $\mathcal{G}_{\mathbf{i}}$. Let these one-element vertices be $\{s\}$ and $\{t\}$. Consider the pair $(i_{t-1},i_t)^{\varepsilon}(t)$. It must have an equal pair, but i_t is not equal to any other index. This means that $(i_{t-1},i_t)^{\varepsilon(t-1)} = (i_t,i_{t+1})^{\varepsilon(t)}$. It follows that $\varepsilon(t-1)\neq\varepsilon(t)$. There are exactly two possibilities for this: $t=m_0$ or $t=0$. Assume without loss of generality that $s=0$ and $t=m_0$. Therefore, $i_{m_0-1} = i_{m_0+1}$ (notice that $(m_0-1)+(m_0+1) = 2m_0$). Define the $(m_0-1,1)$-path $\mathbf{j}$ as follows: $j_k := i_k$ if $k\in\{0,\ldots,m_0-2\}$, $j_{m_0-1} := i_{m_0-1} = i_{m_0+1}$, $j_k := i_{k+2}$ if $k\in\{m_0,\ldots,2(m_0-1)-1\}$ (see Fig. 2). The path $\mathbf{j}$ is an $(m_0-1,1)$-regular path. There is only one such path by the induction hypothesis, and $(i_k = i_l)\Leftrightarrow(j_k = j_{l-2})\Leftrightarrow(k+(l-2) = 2(m_0-1))\Leftrightarrow(k+l = 2m_0)$. $\square$

DEFINITION 3. Notice that a vertex of a regular graph has two outgoing edges iff the corresponding index has the form i_{2mk} (because it should be $(i_j,i_{j+1})^{\varepsilon(j)} = (i_j,i_{j+1})$ and $(i_{j-1},i_j)^{\varepsilon(j-1)} = (i_j,i_{j-1})$, and this

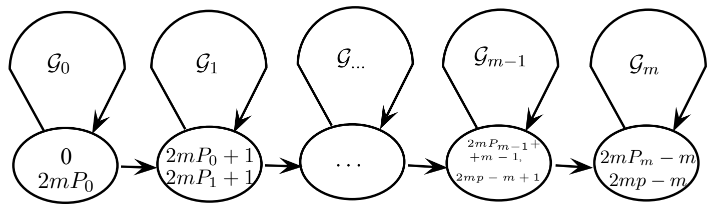

Figure 3. The regular graph $\mathcal{G}_{\mathbf{j}}$ obtained from the collection of (m, p_i)-regular graphs $(\mathcal{G}_0, \mathcal{G}_1, \dots, \mathcal{G}_m)$.

happens if and only if $j = 2mk$). The distance between such a vertex and vertex $\mathcal{V}$ is called a *type of vertex* $\mathcal{V}$. The *type of index* i_j is the type of a vertex $\mathcal{V}$ such that $j \in \mathcal{V}$. It is clear that index i_j has type $j \pmod{2m}$ if $j \pmod{2m} \in \{0, \dots, m-1\}$ or type $-j \pmod{2m}$ in the other case. There are $m+1$ types of vertices. Note that only indices of the same type can be equal (this is proved in the case $p = 1$ in Proposition 2 and will be proved for other cases below).

Consider a collection of (m, p_k)-regular paths $(\mathbf{i}_0, \mathbf{i}_1, \dots, \mathbf{i}_m)$ (such that $\sum_{k=0}^{m} p_k = p - 1$) and collection of the corresponding regular graphs $(\mathcal{G}_0, \mathcal{G}_1, \dots, \mathcal{G}_m)$. The sum of the lengths of paths is $2m(p-1)$. We indicate the recipe how to obtain the (m, p)-regular graph from these collections. We take an $(m, 1)$-regular graph and attach to its vertices the graphs from the collection in the following way: the graph $\mathcal{G}_0$ is attached to vertex of type 0, the graph $\mathcal{G}_1$ is attached to vertex of type $1, \dots,$ and the graph $\mathcal{G}_m$ to the vertex of type m. For a more detailed argument, we denote $\sum_{i=0}^{k} p_i$ by P_k ($P_m = p - 1$) and the indices of the kth path $\mathbf{i}_k$ by $i_j^{(k)}$. The resulting (m, p)-regular graph is denoted by $\mathcal{G}_{\mathbf{j}}$. Define the map $\Delta : \Delta(\mathbf{i}_0, \mathbf{i}_1, \dots, \mathbf{i}_m) = \mathbf{j}$ as follows:

$$
\begin{gathered}
j_0 := i_0^{(0)}, \quad j_1 = i_1^{(0)}, \quad \dots, \quad j_{2mP_0-1} := i_{2mp_0-1}^{(0)}, \quad j_{2mP_0} := j_0; \\
j_{2mP_0+1} := i_1^{(1)}, \quad \dots, \quad j_{2mP_1-1} := i_{2mp_1-1}^{(1)}, \quad j_{2mP_1} := i_0^{(1)}, \quad j_{2mP_1+1} := j_{2mP_0+1}; \\
j_{2mP_1+2} := i_2^{(2)}, \quad \dots, \quad j_{2mP_2-1} := i_{2mp_2-1}^{(2)}, \quad j_{2mP_2} := i_0^{(2)}, \quad j_{2mP_2+1} := i_1^{(2)}, \quad j_{2mP_2+2} := j_{2mP_1+2}; \\
\vdots \\
j_{2mP_{m-1}+m} := i_m^{(m)}, \quad \dots, \quad j_{2m(p-1)-1} := i_{2mp_m-1}^{(m)}, \quad j_{2m(p-1)} := i_0^{(m)}, \\
j_{2mp-m+1} := j_{2mP_{m-1}+m-1}, \quad \dots, \quad j_{2mp-k} := j_{2mP_k+k}, \quad \dots, \quad j_{2mp-1} := j_{2mP_1+1}.
\end{gathered}
\tag{2.21}
$$

Let $\widetilde{\Delta}$ be the corresponding map $\widetilde{\Delta}(\mathcal{G}_0, \mathcal{G}_1, \dots, \mathcal{G}_m) = \mathcal{G}_{\mathbf{j}}$. Graphically, the construction (2.21) looks as shown in Fig. 3.

Example. For example, we consider for $m = 2$ the collection of $(2, p_k)$-regular graphs $(\mathcal{G}_{2,2}, \mathcal{G}_{2,0}, \mathcal{G}_{2,1})$ (see Fig. 4) and obtain from it a $(2, 4)$-regular graph $\mathcal{G}_{2,4}$ (see Fig. 5).

Proposition 3. *Using the above construction, we get an* (m, p)*-regular graph.*

Proof. Indeed, the graph $\mathcal{G}_{\mathbf{j}}$ has exactly mp edges, and $k_r = 2$ for all $r = 1, 2, \dots, mp$. Furthermore, there are exactly $mp + 1$ vertices (there is no newly introduced vertex, and there are exactly $\sum_{i=0}^{m} (mp_i + 1) = m(p-1) + m + 1 = mp + 1$ vertices of graphs $\mathcal{G}_k$). $\square$

Note that the map Δ is an injection.

Now we consider an arbitrary (m, p)-regular path $\mathbf{i}$ and try to construct the inverse map for Δ. We denote

$$
\begin{gathered}
J_0 := \{j \colon i_j = i_0\}, \\
J_k := \{j \colon j \neq 2mp - k,\ i_j = i_{2mp-k}\}, \quad k \in \{1, \dots, m-1\},
\end{gathered}
$$

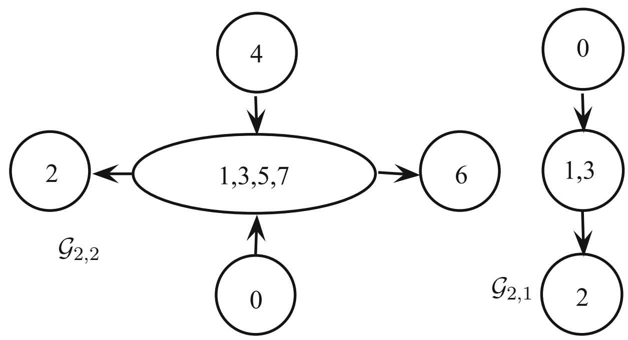

Figure 4. Graphs $\mathcal{G}_{2,2}$ and $\mathcal{G}_{2,1}$. Graph $\mathcal{G}_{m,0}$ is empty.

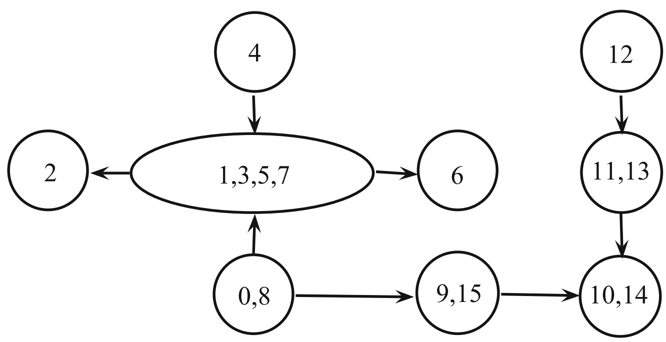

Figure 5. The resulting graph $\mathcal{G}_{2,4}$.

$$J_m := \{j\colon i_j = i_{2mp-m}\},$$
$$\overline{J_k} := \max(J_k), \qquad \underline{J_k} := \min(J_k). \tag{2.22}$$

We will prove that the sets J_k have some remarkable properties, and after that it will be clear how to obtain a collection of regular paths from one (large) regular path.

Proposition 4. *J_k is nonempty.*

Proof. Indeed, there is $0 \in J_0$ and $2mp - m \in J_m$. If J_k is void with $k \in \{1, \dots, m-1\}$, then the index i_{2mp-k} has no equal indices in the path $\mathbf{i}$. However, in this case, the pair $(i_{2mp-k-1}, i_{2mp-k})^-$ has no equivalent for the following reason. The index i_{2mp-k} appears in $(2mp-k, 2mp-k+1)^-$ and in $(2mp-k-1, 2mp-k)^-$ only, and they are not equivalent. However, each edge in a regular path has an equivalent one, a contradiction. Therefore, the initial assumption that J_k is void must be false. $\square$

Proposition 5. *J_k $(0 \leqslant k \leqslant m)$ are pairwise disjoint, and if $k < l$, then $J_k < J_l$ (for all $j \in J_k$ and all $i \in J_l$, the inequality $j < i$ holds).*

Proof. Indeed, if $J_k \cap J_l \neq \emptyset$, then $i_{2mp-k} = i_{2mp-l}$. The edges of the path $\mathbf{i}$ have the same orientation on the section $(2mp-k, 2mp-k-1)^-, \dots, (2mp-l+1, 2mp-l)$, and, therefore, the graph $G_{\mathbf{i}}$ has a cycle. However, a regular graph is a tree, a contradiction. Thus, $J_k \cap J_l = \emptyset$. We prove the second part of Proposition 5 for the case $l = k+1$ only (which is sufficient). Consider the edge $(2mp-(k+1), 2mp-k)^-$. It must be equivalent to an edge $(t, t+1)^+$ with some $t \in J_k$ and $t+1 \in J_{k+1}$. If there exists $s \in J_k$ such that $s > t$, then $s > t+1$ $(J_k \cap J_{k+1} = \emptyset)$. The edge $(t, t+1)$ is not equivalent to any edge in the section

$(t+1, t+2), \ldots, (s-1, s)$, because it has only one equivalent edge $(2mp-(k+1), 2mp-k)^-$. It follows that there are two different paths in the graph $G_{\mathbf{i}}$ that connect vertex $\mathcal{U}$ (such that $t+1 \in \mathcal{U}$) and vertex $\mathcal{V}$ (such that $s \in \mathcal{V}$ and $t \in \mathcal{V}$), that is, there is a cycle in the the graph $G_{\mathbf{i}}$, and, hence, we have a contradiction. Therefore, $t = \max J_k = \overline{J_k}$. Similarly, $t+1 = \underline{J_{k+1}}$. It follows that $\overline{J_k} + 1 = \underline{J_{k+1}}$ and, for all $j \in J_k$ and all $i \in J_{k+1}$, the inequality $j < i$ holds. $\square$

Proposition 6. *For all k, the difference $(\overline{J_k} - \underline{J_k})$ is divisible by $2m$.*

Proof. Denote $(\overline{J_k} - \underline{J_k}) \pmod{2m}$ by d_k. Notice that $(\overline{J_k} - \underline{J_k})$ is the number of edges in the path's section $(\underline{J_k}, \underline{J_k}+1), \ldots, (\overline{J_k}-1, \overline{J_k})$. Notice that the orientation of edges changes after every m steps. Edges of the form $(i_{\overline{J_k}}, i_{\underline{J_{k+1}}})^+$ have the same orientation. It follows that $d_0 \leqslant m-1$, $d_1 \leqslant m-1$, $(d_0+1+d_1) \leqslant m-1 \pmod{2m}$ (and so $(d_0+1+d_1) \leqslant m-1$, because $0 \leqslant d_0+1+d_1 \leqslant 2m-1$), $\ldots$, $0 \leqslant d_0+1+d_1+1+\cdots+d_{m-2}+1+d_{m-1} \leqslant m-1$ (similarly), i.e., $0 \leqslant \sum_{k=0}^{m-1} d_k + m - 1 \leqslant m-1$. Therefore, $d_k = 0$ for all $k = 0, 1, \ldots, m-1$. Consider all edges of the path $\mathbf{i}$. There are m edges of the form $(\overline{J_k}, \underline{J_{k+1}})$, m edges of the form $(2mp-k, 2mp-k+1)^-$ with some $k = 1, 2, \ldots, m$, and all the remaining ones are in sections of the form $(\underline{J_k}, \underline{J_k}+1), \ldots, (\overline{J_k}-1, \overline{J_k})$. There are $2mp$ edges in total. Therefore, $\sum_{k=0}^{m}(\overline{J_k} - \underline{J_k}) + m + m = 2mp$, and, hence, $\sum_{k=0}^{m} d_k = 0 \pmod{2m}$. It follows that also $d_m = 0$. $\square$

Proposition 7. *If $\underline{J_k} < t < \overline{J_k}$ and $\underline{J_l} < s < \overline{J_l}$, then $i_t \neq i_s$. In other words, the sections of the path $\mathbf{i}$ of the form $(\underline{J_k}, \underline{J_k}+1), \ldots, (\overline{J_k}-1, \overline{J_k})$ with $k = 0, 1, \ldots, m$ are disjoint.*

Proof. Without loss of generality, we consider $l > k$. Assume that $i_t = i_s$. In this case, the section $(\underline{J_k}, \underline{J_k}+1), \ldots, (s-1, s)$ contains the edge $(\overline{J_k}, \underline{J_{k+1}})$, and the section $(t, t+1), \ldots, (\overline{J_k}-1, \overline{J_k})$ does not contain it or its equivalent $(2mp-k-1, 2mp-k)^-$. Thus, there are two nonequal paths in the regular graph $G_{\mathbf{i}}$ that connect vertex $\mathcal{U}$ (such that $\overline{J_k} \in \mathcal{U}$) and vertex $\mathcal{V}$ (such that $s \in \mathcal{V}$ and $t \in \mathcal{V}$), that is, there is a cycle in the the graph $G_{\mathbf{i}}$. Therefore, the initial assumption must be false. $\square$

Now we can describe the inverse map for Δ. Let $p_k := (\overline{J_k} - \underline{J_k})/2m$ (p_k is a nonnegative integer by Proposition 6). Furthermore, we have that the sum $\sum_{k=0}^{m-1} p_k = p-1$ (see the proof of Proposition 6). Denote by $\mathbf{j}^{(k)}$ the kth resulting path (it has the length $2mp_k$, and if $p_k = 0$, then $\mathbf{j}_k$ is empty). Let

$$j_t^{(k)} := i_{\underline{J_k}+((t-k) \bmod 2mp_k)}, \quad t \in \{0, \ldots, 2mp_k - 1\},\ k \in \{0, \ldots, m\}. \tag{2.23}$$

Now one obtains the collection $(\mathcal{G}_{2,2}, \mathcal{G}_{2,0}, \mathcal{G}_{2,1})$ (see Fig. 4) from the graph $\mathcal{G}_{2,4}$ (see Fig. 5) in the way described in (2.23).

Proposition 8. *The collection of paths $(\mathbf{j}^{(0)}, \mathbf{j}^{(1)}, \ldots, \mathbf{j}^{(m)})$ (defined by (2.23)) is a collection of regular paths.*

Proof. In fact, the path $\mathbf{j}^{(k)}$ is almost the same as the section $(\underline{J_k}, \underline{J_k}+1), \ldots, (\overline{J_k}-1, \overline{J_k})$ of the regular path $\mathbf{i}$. This section contains $2mp_k$ edges. Each of these edges has an equivalent one in the same section by Proposition 7. Therefore, this section contains exactly $mp_k + 1$ distinct indices because of connectivity and noncyclicity. Hence, the path $\mathbf{j}^{(k)}$ is a regular one. $\square$

Thus, $\widetilde{\Delta}$ is a bijection between $T_{m,p}^{\mathrm{reg}}$ and $\bigcup T_{m,p_0}^{\mathrm{reg}} \times T_{m,p_1}^{\mathrm{reg}} \times \cdots \times T_{m,p_m}^{\mathrm{reg}}$, where the union is taken over all $p_0 + p_1 + \cdots + p_m = p-1$. Hence, $\# T_{m,p}^{\mathrm{reg}} = M_p^{(m)}$, and Lemma 6 is proved. $\square$

Lemmas 5 and 6 show that the moments of the spectral distribution converge to $M_p^{(m)}$. Thus, Theorem 1 is proved.

REFERENCES

1. Z.D. Bai, Methodologies in spectral analysis of large-dimensional random matrices, a review, *Stat. Sin.*, **9**(3):611–677, 1999.
2. T. Banica, S. Belinschi, M. Capitaine, and B. Collins, Free Bessel laws, preprint, `arXiv:0710.5931`.
3. T. Carleman, *Les Fonctions Quasi-analytiques*, Paris, 1926.
4. I.M. Gessel and G. Xin, The generating function of ternary trees and continued fractions, *Electron. J. Comb.*, **13**(1), 2006.
5. R.L. Graham, D.E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science,* 2nd edition, Addison-Wesley Professional, 1994.
6. V. Marchenko and L. Pastur, The eigenvalue distribution in some ensembles of random matrices, *Math. USSR Sb.*, **1**:457–483, 1967.
7. W. Mlotkowski, Fuss–Catalan numbers in noncommutative probability, preprint.
8. A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, Cambridge Univ. Press, 2006.
9. F. Oravecz, On the powers of Voiculescu's circular element, *Stud. Math.*, **145**:85–95, 2001.