A note on random orthogonal polynomials
on a compact interval

Melanie Birke

Ruhr-Universität Bochum

Fakultät für Mathematik

44780 Bochum, Germany

e-mail: melanie.birke@ruhr-uni-bochum.de

Holger Dette

Ruhr-Universität Bochum

Fakultät für Mathematik

44780 Bochum, Germany

email: holger.dette@ruhr-uni-bochum.de

June 19, 2008

Abstract

We consider a uniform distribution on the set $\mathcal{M}_k$ of moments of order $k \in \mathbb{N}$ corresponding to probability measures on the interval $[0, 1]$. To each (random) vector of moments in $\mathcal{M}_{2n-1}$ we consider the corresponding uniquely determined monic (random) orthogonal polynomial of degree n and study the asymptotic properties of its roots if $n \to \infty$.

Keyword and phrases: moment space, random moment sequence, random orthogonal polynomial, arcsine distribution, Chebyshev polynomials, random matrices.
AMS Subject Classification: 60F15, 33C45, 44A60

1 Introduction

For a probability measure η on the interval $[0, 1]$ let

$$(1.1) \qquad c_k = c_k(\eta) = \int_0^1 x^k \eta(dx) \qquad k = 0, 1, 2, \ldots$$

denote the corresponding moments and consider the moment space (of order n) defined by

$$(1.2) \quad \mathcal{M}_n = \{(c_1(\eta), \ldots, c_n(\eta))^T \mid \quad \eta \text{ is a probability measure on the interval } [0, 1] \} .$$

The set $\mathcal{M}_n$ is a very small subset of $\mathbb{R}^n$ with volume proportional to 2^{-n^2} and has been studied extensively in the literature [see e.g. Karlin and Shapeley (1953), Skibinsky (1967, 1968, 1969),

Chang, Kemperman and Studden (1993) or Gamboa and Lozada-Chang (2004)]. In order to understand the structure of the moment space Chang, Kemperman and Studden (1993) assigned a uniform distribution over $\mathcal{M}_n$ and studied the asymptotic properties of the first k components of the random vector

(1.3) $$(C_{1,n}, \ldots, C_{n,n})^T \sim \mathcal{U}(\mathcal{M}_n)$$

if $n \to \infty$. In particular, they showed that an appropriately standardized version of the vector $(C_{1,n} \ldots, C_{k,n})^T$ converges weakly to a centered normal distribution, that is

(1.4) $$\sqrt{n}\, \{(C_{1,n}, \ldots, C_{k,n})^T - (c_1^0, \ldots, c_k^0)^T\} \to \mathcal{N}(0, \Sigma),$$

where c_k^0 denotes the kth moment of the arcsine distribution defined by

(1.5) $$c_k^0 = \frac{1}{\pi} \int_0^1 \frac{x^k}{\sqrt{x(1-x)}}\, dx = \frac{1}{2^{2k}} \binom{2k}{k}, \qquad k = 0, 1, \ldots ,$$

and the matrix $\Sigma \in \mathbb{R}^{k \times k}$ is given by

(1.6) $$\Sigma = \frac{1}{2} \left(c_{i+j}^0 - c_i^0 c_j^0\right)_{i,j=1}^k .$$

A large deviation principle for the random moment vector $(C_{1,n} \ldots, C_{k,n})^T$ was derived in Gamboa and Lozada-Chang (2004), while Dette and Gamboa (2007) investigated the asymptotic properties of a moment range process.

It is the purpose of the present paper to provide further insight in the probabilistic properties of quantities associated to random moment sequences. In particular, by the definition $\mathcal{L}(x^k) = C_{k,2n-1}$ $(k = 1, \ldots, 2n-1)$ each random vector $(C_{1,2n-1}, \ldots, C_{2n-1,2n-1})^T \sim \mathcal{U}(\mathcal{M}_{2n-1})$ defines a (random) moment functional and a corresponding sequence of monic random orthogonal polynomials $P_{0,n}(x), \ldots, P_{n,n}(x)$ satisfying

(1.7) $$\mathcal{L}(P_{k,n}\, P_{l,n}) = 0 \quad \text{if} \quad k \neq l\,;\; k, l \in \{0, \ldots, n\}$$

[see Chihara (1978)]. These polynomials are carefully introduced in Section 2, where we also state some non-standard results regarding moment theory. In Section 3 we derive the asymptotic properties of the (random) roots $X_{1,n}, \ldots, X_{n,n}$ of the polynomial $P_{n,n}(x)$ associated with the random moment sequence $(C_{1,2n-1}, \ldots, C_{2n-1,2n-1}) \sim \mathcal{U}(\mathcal{M}_{2n-1})$. In particular, it is shown that the empirical distribution function of the (random) roots $X_{1,n}, \ldots, X_{n,n}$ converges almost surely to the arcsine distribution if $n \to \infty$ and that an appropriate standardization of the vector $X_{1,n}, \ldots, X_{k,n}$ is asymptotically normal distributed, where the roots of the Chebyshev polynomial of the first kind are used for the centering, and the rate of convergence is $1/\sqrt{n}$.

2 Random coefficients in a three term recurrence relation

The set $\mathcal{M}_n$ defined by (1.2) is a very small compact subset of the unit cube $[0,1]^n$ with nonempty interior and volume

$$(2.1) \qquad \mathrm{Vol}(\mathcal{M}_n) = \prod_{k=1}^{2n} \frac{\Gamma(k)\Gamma(k)}{\Gamma(2k)} \approx c \cdot 2^{-n^2}$$

[see Karlin and Shapely (1953)]. The interior of $\mathcal{M}_n$ is denoted by $\mathcal{M}_n^0$ throughout this paper. It is well known that there exist an infinite number of probability measures on the interval $[0,1]$ with moments up to the order $2n-1$ given by $(c_1, \ldots, c_{2n-1})^T \in \mathcal{M}_{2n-1}^0$ [see Dette and Studden (1997)]. Moreover, there exists a unique measure η^- supported on exactly n points in the open interval $(0,1)$ such that $c_i(\eta^-) = \int_0' x^i d\eta^-(x) = c_i(i = 1, \ldots, 2n-1)$ and such that the point $(c_1(\eta^-), \ldots, c_{2n}(\eta^-))^T \in \partial\mathcal{M}_{2n}$ where $\partial\mathcal{M}_{2n}$ denotes the boundary of the set $\mathcal{M}_{2n}$. The measure η^- is called lower principal representation of the point $(c_1, \ldots, c_{2n-1})^T \in \mathcal{M}_{2n-1}^0$ [see Skibinsky (1967)]. A straightforward calculation shows [see e.g. Szegö (1975) or Chihara (1978)] that for a given vector

$$(c_1, \ldots c_{2n-1})^T = (c_1(\eta), \ldots, c_{2n-1}(\eta))^T \in \mathcal{M}_{2n-1}^0$$

the polynomials $P_{0,n}(x) = 1$,

$$(2.2) \qquad P_{m,n}(x) := \begin{vmatrix} c_0 & \cdots & c_{m-1} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ c_m & \cdots & c_{2m-1} & x^m \end{vmatrix} \Bigg/ \begin{vmatrix} c_0 & \cdots & c_{m-1} \\ \vdots & \ddots & \vdots \\ c_{m-1} & \cdots & c_{2m-2} \end{vmatrix} ; \quad m = 1, \ldots, n$$

have leading coefficient 1 and are orthogonal with respect to the measure η^-, that is

$$(2.3) \qquad \mathcal{L}(P_{k,n} \, P_{l,n}) = \int_0^1 P_{k,n}(x)P_{l,n}(x)d\eta^-(x) = 0 \qquad \text{if} \quad k \neq l,$$

where the moment functional $\mathcal{L}$ is defined by

$$(2.4) \qquad \mathcal{L}(x^k) = c_k = \int_0^1 x^k d\eta^-(x) \qquad (k = 0, \ldots, 2n-1) .$$

In other words: each vector $(c_1, \ldots, c_{2n-1})^T \in \mathcal{M}_{2n-1}^0$ uniquely determines monic orthogonal polynomials $P_{0,n}(x), \ldots, P_{n,n}(x)$ satisfying (2.3). In the following we consider a one-to-one mapping of the set $\mathcal{M}_{2n-1}^0$ on the cube $(0,1)^{2n-1}$, which was introduced by Skibinsky (1967). The new coordinates are called canonical moments and have been studied by numerous authors [see e.g. Dette and Studden (1997) for a detailed discussion]. To be precise, let $\mathcal{P}([0,1])$ denote the set of

all probability measures on the interval $[0,1]$, $\Phi_k(x) = (x, \ldots, x^k)$ the vector of all monomials of order k and define for a fixed vector $c = (c_1, \ldots, c_k)^T \in \mathcal{M}_k$

$$\mathcal{S}_k(c) := \left\{ \mu \in \mathcal{P}([0,1]) : \int_0^1 \Phi_k(x)\mu(dx) = c \right\}$$

as the set of all probability measures on the interval $[0,1]$ whose moments up to the order k coincide with $c = (c_1, \ldots, c_k)^T$. Note that this set is a singleton if and only if $c \in \partial \mathcal{M}_k$ [see Dette and Studden (1997)]. For $k = 2, 3, \ldots$ and for a given point $(c_1, \ldots, c_{k-1})^T \in \mathcal{M}_{k-1}$ we define $c_k^+ = c_k^+(c_1, \ldots, c_{k-1})$ and $c_k^- = c_k^-(c_1, \ldots, c_{k-1})$ as the largest and smallest value of c_k such that $(c_1, \ldots, c_k)^T \in \partial \mathcal{M}_k$, that is

$$c_k^- = \min \left\{ \int_0^1 x^k \mu(dx) \mid \mu \in S_{k-1}(c_1, \ldots, c_{k-1}) \right\},$$
$$c_k^+ = \max \left\{ \int_0^1 x^k \mu(dx) \mid \mu \in S_{k-1}(c_1, \ldots, c_{k-1}) \right\}.$$

Note that $c_k^- \leq c_k \leq c_k^+$ and that both inequalities are strict if and only if $(c_1, \ldots, c_{k-1})^T \in \mathcal{M}_{k-1}^0$ [see Dette and Studden (1997)]. For a moment point $c = (c_1, \ldots, c_n)^T$ in the interior of the moment space $\mathcal{M}_n$ the canonical moments or canonical coordinates of the vector c are defined by

$$(2.5) \qquad p_1 = c_1, \text{ and } p_k = \frac{c_k - c_k^-}{c_k^+ - c_k^-} \quad k = 2, \ldots, n .$$

Note that $0 < p_k < 1$ $(k = 1, \ldots, n)$ if $(c_1, \ldots, c_n)^T \in \mathcal{M}_n^0$, and that the definition (2.5) defines a one to one mapping between $\mathcal{M}_n^0$ and the open unit cube $(0,1)^n$. For more details regarding canonical moments we refer to the work of Skibinsky (1967,1968,1969) and to the monograph of Dette and Studden (1997). In particular it is shown in the lastnamed reference that the three term recurrence relation corresponding to the monic orthogonal polynomials defined by (2.2) can be represented in terms of canonical moments, that is $P_0(x) = 1; P_1(x) = x - \zeta_1$, and for $1 \leq m \leq n-1$

$$(2.6) \qquad P_{m+1}(x) = (x - \zeta_{2m} - \zeta_{2m+1})P_m(x) - \zeta_{2m-1}\zeta_{2m}P_{m-1}(x),$$

where the quantities ζ_k are given by $\zeta_0 = 0$, $\zeta_1 = p_1$, $\zeta_k = q_{k-1}p_k$ with $q_{k-1} = 1 - p_{k-1}$ if $k \geq 2$. Moreover, from the representation (2.5) it is easy to see that $\partial c_1/\partial p_1 = 1$,

$$(2.7) \qquad \frac{\partial c_k}{\partial p_j} = \begin{cases} 0 & \text{if } j > k \\ r_k(c_1, \ldots, c_{k-1}) & \text{if } j = k \end{cases}$$

$(k \geq 1)$ where

$$(2.8) \qquad r_{k+1}(c_1, \ldots, c_k) = r_{k+1} = c_{k+1}^+ - c_{k+1}^- = \prod_{j=1}^k p_j(1 - p_j).$$

4

denotes the range of the moment space $\mathcal{M}_k$ (with the convention $r_1 = 1$). In the following section we will use these results to study stochastic properties of the roots of random orthogonal polynomials associated with a uniform distribution on the moment space $\mathcal{M}_{2n-1}$.

3 Asymptotic zero distribution of random orthogonal polynomials

For each $n \in \mathbb{N}$ let $C_{2n-1} = (C_{1,2n-1}, \ldots, C_{2n-1,2n-1})^T \sim \mathcal{U}(\mathcal{M}_{2n-1})$ denote a uniformly distributed vector on the moment space $\mathcal{M}_{2n-1}$. From (2.1) it follows that

$$P\left(C_{2n-1} \in \mathcal{M}_{2n-1}^0\right) = 1 \ ,$$

and consequently the random canonical moments, say $P_{1,2n-1}, \ldots, P_{2n-1,2n-1}$, corresponding to the random vector C_{2n-1} are well defined with probability 1. Observing the representation (2.7) it follows that the density of the random vector $(P_{1,2n-1}, \ldots, P_{2n-1,2n-1})$ is given by

(3.1) $$\prod_{j=1}^{2n-2} \frac{\Gamma(4n-2j)}{(\Gamma(2n-j))^2} \prod_{j=1}^{2n-2} (p_i(1-p_i))^{2n-j-1},$$

which means that $\{(P_{j,2n-1})_{j=1}^{2n-1} \mid n \in \mathbb{N}\}$ is a triangular array of rowwise independent random variables, where $P_{j,2n-1}$ has a symmetric Beta-distribution on the interval $[0,1]$ with parameter $2n-j$, that is $P_{j,2n-1} \sim B(2n-j, 2n-j)$. In what follows, let $P_{m,n}(x)$ denote the mth random monic orthogonal polynomial associated with the random vector C_{2n-1} by equation (2.2) and $X_{1,n}, \ldots, X_{m,n}$ the corresponding roots which are real with probability 1 [see Szegö (1975)]. Our first result gives an explicit representation for the joint density of the random vector $(X_{1,n}, \ldots, X_{n,n})^T$.

Theorem 3.1. *The joint density of the roots $X_{1,n}, \ldots, X_{n,n}$ of the monic random orthogonal polynomial $P_{n,n}(x)$ corresponding to a random vector $C_{2n-1} = (C_{1,2n-1}, \ldots, C_{2n-1,2n-1})^T \sim \mathcal{U}(\mathcal{M}_{2n-1})$ by equation (2.2) is given by*

(3.2) $$f(x_1, \ldots, x_n) = c \cdot \prod_{1 \le i < j \le n} |x_i - x_j|^4 \prod_{i=1}^{n} I_{[0,1]}(x_i),$$

where the normalizing constant c is defined by

$$c = \frac{1}{\Gamma(n+1)} \prod_{r=0}^{n-1} \frac{\Gamma(2r+2n)}{\Gamma(2r+1)^2 \Gamma(2r+2)}.$$

Proof. Consider the random canonical moments $(P_{1,2n-1}, \ldots, P_{2n-1,2n-1})^T$ corresponding to the random vector $(C_{1,2n-1}, \ldots, C_{2n-1,2n-1})^T \sim \mathcal{U}(\mathcal{M}_{2n-1})$. Let $\Xi_1 = \Xi_{1,2n-1} = P_{1,2n-1}$ and for $j \geq 2$ $\Xi_j = \Xi_{j,2n-1} = (1 - P_{j-1,2n-1})P_{j,2n-1}$, then it follows from the recursive relation (2.6) that the random orthogonal polynomial $P_{n,n}(x)$ can be represented as the determinant of a symmetric tridiagonal matrix, that is

$$
(3.3) \qquad P_{n,n}(x) = \begin{vmatrix} x - \Xi_1 & -\sqrt{\Xi_1\Xi_2} \\ -\sqrt{\Xi_1\Xi_2} & x - \Xi_2 - \Xi_3 \\ & -\sqrt{\Xi_3\Xi_4} & \ddots \\ & & & \ddots \\ & & & \ddots & & -\sqrt{\Xi_{2n-3}\Xi_{2n-2}} \\ & & & & -\sqrt{\Xi_{2n-3}\Xi_{2n-2}} & x - \Xi_{2n-2} - \Xi_{2n-1} \end{vmatrix}
$$

Consequently, the roots $X_{1,n}, \ldots, X_{n,n}$ of the polynomial $P_{n,n}(x)$ are the eigenvalues of the random Jacobi matrix $J_{n,n}$ where for $m < n$

$$
(3.4) \qquad J_{m,n} = \begin{pmatrix} \Xi_1 & \sqrt{\Xi_1\Xi_2} \\ \sqrt{\Xi_1\Xi_2} & \Xi_2 + \Xi_3 \\ & \sqrt{\Xi_3\Xi_4} & \ddots \\ & & & \ddots \\ & & & \ddots & & \sqrt{\Xi_{2m-3}\Xi_{2m-2}} \\ & & & & \sqrt{\Xi_{2m-3}\Xi_{2m-2}} & \Xi_{2m-2} + \Xi_{2m-1} \end{pmatrix} \in \mathbb{R}^{m \times m} .
$$

Now define for $j = 0, \ldots, 2n - 1$ the quantities $\alpha_j = 2P_{j+1,2n-1} - 1$ then $\alpha_0, \ldots, \alpha_{2n-1}$ are independent random variables and α_j follows a symmetric Beta-distribution on the interval $[-1, 1]$ with parameter $2n - j - 1$. Moreover, a straightforward calculation shows that the matrix $\tilde{J}_n = 4J_{n,n} - I_n$ can be represented as

$$
(3.5) \qquad \tilde{J}_n := \begin{pmatrix} b_1 & a_1 \\ a_1 & b_2 & \ddots \\ & \ddots & \ddots & a_{n-1} \\ & & a_{n-1} & b_n \end{pmatrix} \in \mathbb{R}^{n \times n}
$$

with entries

$$
\begin{aligned} b_{k+1} &= (1 - \alpha_{2k-1})\alpha_{2k} - (1 + \alpha_{2k-1})\alpha_{2k-2} \\ a_{k+1} &= \left\{(1 - \alpha_{2k-1})(1 - \alpha_{2k}^2)(1 + \alpha_{2k+1})\right\}^{1/2} . \end{aligned}
$$

and $\alpha_{-1} = -1$. This matrix has been considered recently by Killip and Nenciu (2004) in a more general context. By the results of these authors we obtain that the density of the random

6

eigenvalues $\lambda_1, \ldots, \lambda_n$ of the matrix $\tilde{J}_n$ is given by $\tilde{c} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^4$, where $\tilde{c}$ is an appropriate normalizing constant. Transferring this result to the eigenvalues of the matrix $J_{n,n}$ it follows that the density of the roots of the random polynomial $P_{n,n}(x)$ is given by (3.2). □

The density defined by (3.2) is a special case of the Jacobi ensemble in the symplectic case [see e.g. Mehta (2004)] which has found considerable interest in the recent literature [see e.g. Collins (2005) or Johnstone (2008) among others]. The density of the general Jacobi β-ensemble is given by

$$c_{a_n, b_n}^n \prod_{1 \leq i \leq j \leq n} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^n \lambda_i^{a_n} (1 - \lambda_i)^{b_n} I_{[0,1]}(\lambda_i)$$

where $a_n, b_n \geq -1$, c_{a_n, b_n}^n is a normalizing constant and $\beta > 0$ [see Killip and Nenciu (2004)]. While most authors consider the case where $n \to \infty$, $a_n/n \to \alpha$; $b_n/n \to \beta$, much less attention has been paid to the case where a_n and b_n are fixed. We use the explicit representation for the density of the roots of $P_{n,n}(x)$ in terms of eigenvalues of the random matrix $J_{n,n}$ defined in (3.4) to derive asymptotic properties for the empirical distribution function of the random variables $X_{1,n}, \ldots, X_{n,n}$.

Theorem 3.2. *Let*

$$F_n(x) = \frac{1}{n} \sum_{j=1}^n I\{X_{j,n} \leq x\}$$

denote the empirical distribution function of the roots of the monic random orthogonal polynomial $P_{n,n}(x)$ associated with the random vector $C_{2n-1} = (C_{1,2n-1}, \ldots, C_{2n-1,2n-1})^T \sim \mathcal{U}(\mathcal{M}_{2n-1})$, then for all $x \in [0,1]$

$$\lim_{n \to \infty} F_n(x) = \frac{1}{\pi} \int_0^x \frac{dt}{\sqrt{t(1-t)}} \qquad a.s.$$

Proof. Define

(3.6)
$$D_{m,n} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2\sqrt{2}} & & & & \\ \frac{1}{2\sqrt{2}} & \frac{1}{2} & \frac{1}{4} & & & \\ & \frac{1}{4} & \frac{1}{2} & & & \\ & & & \ddots & \frac{1}{4} & \\ & & & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \in \mathbb{R}^{m \times m},$$

then a straightforward calculation shows that the characteristic polynomial of the matrix $D_{n,n}$ is given by the monic Chebyshev polynomial of the first kind on the interval $[0,1]$, that is

(3.7)
$$\overline{T}_n(x) := \det(xI_n - D_{n,n}) = 2^{-2n+1} \cos(n \arccos(2x - 1)),$$

7

which has roots

$$(3.8) \qquad x_{k,n} = \frac{\cos((2k-1)\frac{\pi}{2n}) + 1}{2} \; ; \qquad k = 1, \ldots, n.$$

Moreover, define $T_k(x) = 2^{2k-1} \, \overline{T}_k(x)$ for $k \geq 1$, $T_0(x) = 1/\sqrt{2}$ and $T(x) = (T_0(x), \ldots, T_{n-1}(x))^T$, then it follows from the relation $\cos((n+1)z) + \cos((n-1)z) = 2\cos(nz)\cos(z)$ by a straightforward calculation that

$$(3.9) \qquad x_{k,n} \, T(x_{k,n}) = D_{n,n} \, T(x_{k,n}) \; ; \qquad k = 1, \ldots, n .$$

This shows that the vectors

$$(3.10) \qquad t_{k,n} = T(x_{k,n}) = \left(1/\sqrt{2}, \cos(\frac{2k-1}{2n}\pi), \ldots, \cos((n-1)\frac{2k-1}{2n}\pi)\right)^T \; ; \; k = 1, \ldots, n$$

are the eigenvectors of the matrix $D_{n,n}$ corresponding to the eigenvalues $x_{1,n}, \ldots, x_{n,n}$, respectively. In the following discussion define

$$N_n(x) = \frac{1}{n} \sum_{j=1}^{n} I\{x_{j,n} \leq x\}$$

as the empirical distribution function of the roots of $\overline{T}_n(x)$ (or $T_n(x)$), then an elementary calculation shows that for all $x \in [0,1]$

$$(3.11) \qquad \lim_{n \to \infty} N_n(x) = \frac{1}{\pi} \int_0^x \frac{dt}{\sqrt{t(1-t)}}.$$

From Bai (1999) we have for the Levy distance between the empirical distribution functions F_n and N_n the estimate

$$L^3(F_n, N_n) \; \leq \; \frac{1}{n} \sum_{j=1}^{n} | \, X_{(j),n} - x_{(j),n} \, |^2 \leq \frac{1}{n} \, \text{tr} \, (J_{n,n} - D_{n,n})^2$$

$$(3.12) \qquad \qquad \leq \; \frac{1}{n} \sum_{i=1}^{n} (J_{ii} - D_{ii})^2 + \frac{2}{n} \sum_{i=1}^{n-1} (J_{ii+1} - D_{ii+1})^2,$$

where $X_{(1),n} \leq \cdots \leq X_{(n),n}$ and $x_{(1),n} \leq \cdots \leq x_{(n),n}$ denote the ordered roots of the polynomials $P_{n,n}(x)$ and $\overline{T}_n(x)$, respectively, and J_{ij} and D_{ij} are the elements of the tridiagonal matrices $J_{n,n}$ and $D_{n,n}$ defined in (3.4) and (3.6), respectively. Using the notation $Q_{j,2n-1} = 1 - P_{j,2n-1}$ $(j = 1, \ldots, 2n-1)$ it now follows by a straightforward calculation that for $2 \leq i \leq n$

$$(J_{ii} - D_{ii})^2 \; = \; \left(Q_{2i-3,2n-1}P_{2i-2,2n-1} + Q_{2i-2,2n-1}P_{2i-1,2n-1} - \frac{1}{2}\right)^2$$

$$\leq \; \left\{|Q_{2i-3,2n-1} - \frac{1}{2}| + \frac{1}{2}|P_{2i-2,2n-1} - \frac{1}{2}| + |Q_{2i-2,2n-1} - \frac{1}{2}| + \frac{1}{2}|P_{2i-1,2n-1} - \frac{1}{2}|\right\}^2$$

$$\leq \; |P_{2i-3,2n-1} - \frac{1}{2}|^2 + 9|P_{2i-2,2n-1} - \frac{1}{2}|^2 + |P_{2i-1,2n-1} - \frac{1}{2}|^2$$

8

and $(J_{11} - D_{11})^2 \leq |P_{1,2n-1} - \frac{1}{2}|^2$. Similarly we obtain for $2 \leq i \leq n-1$

$$\left|J_{ii+1} - D_{ii+1}\right|^2 \leq \left|Q_{2i-2,2n-1}P_{2i-1,2n-1}Q_{2i-1,2n-1}P_{2i,2n-1} - \frac{1}{16}\right|$$

$$\leq \frac{1}{8}\left\{2\left|P_{2i-2,2n-1} - \frac{1}{2}\right| + 3\left|P_{2i-1,2n-1} - \frac{1}{2}\right| + \left|P_{2i,2n-1} - \frac{1}{2}\right|\right\}$$

and $\left|J_{1,2} - D_{1,2}\right|^2 \leq \frac{5}{2}\left|P_{1,2n-1} - \frac{1}{2}\right| + \frac{1}{4}\left|P_{2,2n-1} - \frac{1}{2}\right|$. In the following discussion we will show that

(3.13)
$$\frac{1}{n}\sum_{i=1}^{n-1}\left|P_{2i,2n-1} - \frac{1}{2}\right|^2 \xrightarrow{a.s.} 0,$$

(3.14)
$$\frac{1}{n}\sum_{i=1}^{n}\left|P_{2i-1,2n-1} - \frac{1}{2}\right|^2 \xrightarrow{a.s.} 0,$$

which directly implies

$$\frac{1}{n}\sum_{i=1}^{n}\left(J_{ii} - D_{ii}\right)^2 \xrightarrow{a.s.} 0.$$

Moreover, for the remaining sum in (3.12) it follows that

$$\frac{1}{n}\sum_{i=1}^{n-1}(J_{ii+1} - D_{ii+1})^2 \leq \gamma\left\{\frac{1}{n}\sum_{i=1}^{n-1}\left|P_{2i,2n-1} - \frac{1}{2}\right|^2\right\}^{1/2} + \left\{\frac{1}{n}\sum_{i=1}^{n}\left|P_{2i-1,2n-1} - \frac{1}{2}\right|^2\right\}^{1/2}$$

$$\overset{a.s.}{=} o(1),$$

where the constant γ does not depend on n. The assertion is now a consequence of (3.11) and (3.12), which yields for the Levy-distance between the empirical distribution function F_n and distribution function F of the arcsine measure

$$L(F_n, F) \leq L(F_n, N_n) + L(N_n, F) = o(1) \quad \text{a.s.}$$

For a proof of the almost sure convergence in (3.13) and (3.14) we restrict ourselves to the statement (3.13), the remaining case is treated similarly. In order to prove (3.13) we will use a strong law of large numbers for arrays of rowwise independent random variables. To be precise, define

$$Z_{n,k} = (P_{2k,2n-1} - \frac{1}{2})^2 - \frac{1}{8(2n-2k)+4} \qquad k = 1, \ldots, n-1$$

then $E[Z_{n,k}] = 0$ and a tedious calculation shows that for $k = 1, \ldots, n-1$ and any $t > 0$

$$P(|Z_{n,k}| > t) \leq P(|Z_{n,n-1}| > t),$$

where the distribution of the random variable $Z_{n,n-1}$ does not depend on n [note that $P_{2n-2,2n-1} \sim B(2,2)$]. Consequently, we obtain from Theorem 2 in Hu, Móricz and Taylor (1989) that

$$Z_n = \frac{1}{n^{1/p}}\sum_{k=1}^{n-1}Z_{n,k} \xrightarrow{a.s.} 0$$

9

for any $p \in [1, 2)$. Observing that

$$\frac{1}{4} \sum_{k=1}^{n-1} \frac{1}{2(2n - 1 - 2k) + 1} = O(\log n)$$

it follows for any $p \in [1, 2)$ that

$$\frac{1}{n} \sum_{k=1}^{n-1} \left| P_{2k,2n-1} - \frac{1}{2} \right|^2 = \frac{1}{n^{1-1/p}} \left\{ Z_n + O\left(\frac{\log n}{n^{1/p}} \right) \right\} = o\left(\frac{1}{n^{1-1/p}} \right) \quad \text{a.s.},$$

which establishes (3.13) and completes the proof of Theorem 3.2. $\qquad \square$

Our final result refers to the asymptotic behaviour of the roots of the mth orthogonal polynomial associated with the random moment vector $C_{2n-1} = (C_{1,2n-1}, \ldots, C_{2n-1,2n-1})^T \sim \mathcal{U}(\mathcal{M}_{2n-1})$, where m is fixed. In this case, the limiting distribution is normal, where the roots of the Chebyshev polynomial $\overline{T}_n(x)$ defined in (3.7) are used for the centering.

Theorem 3.3. *Let $X_{1,n}, \ldots, X_{m,n}$ denote the roots of the m-th random monic orthogonal polynomial $P_{m,n}(x)$, which corresponds to the random vector $C_{2n-1} \sim \mathcal{U}(\mathcal{M}_{2n-1})$ by equation (2.2), then*

$$4\sqrt{n} \left\{ (X_{1,n}, \ldots, X_{m,n})^T - (x_{1,n}, \ldots, x_{m,n})^T \right\} \xrightarrow[n \to \infty]{\mathcal{D}} \mathcal{N}(0, \Gamma),$$

where $x_{1,m}, \ldots, x_{m,m}$ are the roots of the Chebyshev polynomial of the first kind defined by (3.8) and the matrix Γ is given by $\Gamma = \frac{2}{m}(\gamma_{k,l})_{k,l=1}^m$, where

$$
\begin{aligned}
\gamma_{k,l} \;=\; & \frac{1}{4} + \frac{1}{2} \sum_{j=2}^{m} T_{j-1}^2(x_{k,m}) \, T_{j-1}^2(x_{l,m}) - \frac{1}{4} \sum_{j=1}^{m-1} \left(T_{j-1}^2(x_{k,m}) \, T_j^2(x_{l,m}) + T_{j-1}^2(x_{l,m}) \, T_j^2(x_{k,m}) \right) \\
& + \frac{1}{4} \, T_1(x_{k,m}) \, T_1(x_{l,m}) + \frac{1}{2} \sum_{j=2}^{m-1} T_{j-1}(x_{k,m}) \, T_j(x_{k,m}) \, T_{j-1}(x_{l,m}) \, T_j(x_{l,m}) \\
& - \frac{1}{4} \sum_{i=1}^{m-2} \left(T_{j-1}(x_{k,m}) \, T_j(x_{k,m}) \, T_j(x_{l,m}) \, T_{j+1}(x_{l,m}) + T_{j-1}(x_{l,m}) \, T_j(x_{l,m}) \, T_j(x_{k,m}) \, T_{j+1}(x_{k,m}) \right),
\end{aligned}
$$

$T_j(x) = \cos(j \arccos(2x - 1))$ and $x_{k,m}$ is the kth zero of the polynomial $T_m(x)$ defined in (3.8).

Proof. It is easy to see that for fixed $m \in \mathbb{N}$ we have for $k = 1, \ldots, m$

(3.15) $$4\sqrt{n} \left(P_{k,2n-1} - \frac{1}{2} \right) \xrightarrow[n \to \infty]{\mathcal{D}} \mathcal{N}(0, 1),$$

which implies that the vector $4\sqrt{n}(P_{1,2n-1}, \ldots, P_{m,2n-1})^T$ is asymptotically multivariate normal distributed with mean 0 and covariance matrix I_m (note that $P_{1,2n-1}, \ldots, P_{m,2n-1}$ are independent random variables). A straightforward calculation now shows that

$$(3.16) \quad 4\sqrt{n}\left\{\left(\Xi_{1,2n-1}, \ldots, \Xi_{2m-1,2n-1}\right)^T - \left(\frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{4}\right)^T\right\} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_{2m-1}(0, A_{2m-1}),$$

where the matrix $A_{2m-1} \in \mathbb{R}^{2m-1 \times 2m-1}$ is tridiagonal and given by

$$(3.17) \qquad A_{2m-1} = \begin{pmatrix} 1 & -\frac{1}{2} & & & & & \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{4} & & & & \\ & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & & \ddots & -\frac{1}{4} & \\ & & & & -\frac{1}{4} & \frac{1}{2} \end{pmatrix}.$$

A further application of the delta-method finally yields

$$(3.18) \qquad 4\sqrt{n}\left\{ \begin{pmatrix} \Xi_1 \\ \Xi_2 + \Xi_3 \\ \vdots \\ \Xi_{2m-2} + \Xi_{2m-1} \\ \sqrt{\Xi_1 \Xi_2} \\ \sqrt{\Xi_3 \Xi_4} \\ \vdots \\ \sqrt{\Xi_{2m-3}\Xi_{2m-2}} \end{pmatrix} - \begin{pmatrix} 1/2 \\ 1/2 \\ \vdots \\ 1/2 \\ 1/(2\sqrt{2}) \\ 1/2 \\ \vdots \\ 1/2 \end{pmatrix} \right\} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, V)$$

where the covariance matrix $V = diag(V_{11}, V_{22}) \in \mathbb{R}^{2m-1 \times 2m-1}$ is block diagonal with blocks $V_{11} = A_m \in \mathbb{R}^{m \times m}$ and

$$V_{22} = \frac{1}{8} \begin{pmatrix} 1 & -\frac{1}{\sqrt{2}} & & & & \\ -\frac{1}{\sqrt{2}} & 1 & -\frac{1}{2} & & & \\ & -\frac{1}{2} & 1 & -\frac{1}{2} & & \\ & & \ddots & \ddots & \ddots & \\ & & & & \ddots & -\frac{1}{2} \\ & & & & -\frac{1}{2} & 1 \end{pmatrix} \in \mathbb{R}^{m-1 \times m-1}.$$

Consequently, the difference of the matrices $J_{m,n} - D_{m,n}$ defined by (3.4) and (3.6) converges also weakly, that is

$$\lim_{n\to\infty} 4\sqrt{n}\,(J_{m,n} - D_m) = \mathcal{S} \qquad \text{(weakly)},$$

where $\mathcal{S}$ denotes a triangular matrix given by

$$
\mathcal{S} = \begin{pmatrix}
M_1 & N_1 & & & & \\
N_1 & M_2 & N_2 & & & \\
& N_2 & M_3 & N_3 & & \\
& & \ddots & \ddots & & \ddots \\
& & & \ddots & & N_{m-1} \\
& & & & N_{m-1} & M_m
\end{pmatrix} ,
$$

and $\{M_1, \ldots, M_m\}$ and $\{N_1, \ldots, N_{m-1}\}$ define independent samples $M_1 \sim \mathcal{N}(0,1)$, $M_j \sim \mathcal{N}(0, 1/2)$ $(j = 2, \ldots, m)$ $N_j \sim \mathcal{N}(0, 1/8)$ $(j = 1, \ldots, m-1)$ such that

(3.19)
$$
\mathrm{Cov}\,(M_i, M_j) = \begin{cases} -\frac{1}{2} & \text{if} \quad i = 1, j = 2 \quad \text{or} \quad i = 2, j = 1 \\ -\frac{1}{4} & \text{if} \quad j = i+1 \quad \text{or} \quad j = i-1 \\ 0 & \text{else} \end{cases}
$$

(3.20)
$$
\mathrm{Cov}\,(N_i, N_j) = \begin{cases} -\frac{1}{8\sqrt{2}} & \text{if} \quad i = 1, j = 2 \quad \text{or} \quad i = 2, j = 1 \\ -\frac{1}{16} & \text{if} \quad j = i+1 \quad \text{or} \quad j = i-1 \\ 0 & \text{else} \end{cases}
$$

Using the same arguments as in Dimitriu and Edelman (2005) it now follows that the first order properties of the matrix $J_{m,n}$ are the same as the first order properties of the matrix $D_{m,n} + \frac{1}{4\sqrt{n}}\mathcal{S}$. In particular, using Lemma 2.1 in this reference we obtain for $i = 1, \ldots, m$ that the asymptotic distribution of the vector

$$
4\sqrt{n} \left\{ (X_{1,n}, \ldots, X_{m,n})^T - (x_{1,n}, \ldots x_{m,n})^T \right\}
$$

coincides with the distribution of the random vector

(3.21)
$$
G = \left(\frac{t_{1,m}^T \, \mathcal{S} \, t_{1,m}}{t_{1,m}^T t_{1,m}}, \ldots, \frac{t_{m,m}^T \, \mathcal{S} \, t_{m,m}}{t_{m,m}^T t_{m,m}} \right)^T
$$

where the vectors $t_{1,m}, \ldots, t_{m,m} \in \mathbb{R}^m$ are defined by (3.10). Obviously, G has a multivariate normal distribution with mean zero and it remains to calculate the corresponding covariance matrix. For this purpose we note that for $k = 1, \ldots, m$

$$
t_{k,m}^T \, t_{k,m} = \frac{1}{2} + \sum_{l=1}^{m-1} \cos^2(l \, \frac{2k-1}{2m}\pi) = \frac{m}{2} \ .
$$

Therefore it remains to calculate the covariance matrix of the vector $(t_{1,m}^T \, \mathcal{S} \, t_{1,m}, \ldots, t_{m,m}^T \, \mathcal{S} \, t_{m,m})$. For this purpose we denote by $(v_1, \ldots, v_m)$ and $(w_1, \ldots, w_m)$ the components of the vectors $t_{k,m}$

and $t_{l,m}$, respectively, and obtain

$$\text{Cov } (t_{k,m}^T \, \mathcal{S} \, t_{k,m}, \; t_{l,m}^T \, \mathcal{S} \, t_{l,m}) = \text{Cov}\Big(\sum_{i=1}^{m} v_i^2 M_i, \sum_{j=1}^{m} w_j^2 M_j \Big) + 4 \, \text{Cov}\Big(\sum_{i=1}^{m-1} v_i v_{i+1} N_i, \sum_{j=1}^{m-1} w_j w_{j+1} N_j \Big)$$

$$= v_1^2 w_1^2 + \frac{1}{2} \sum_{j=2}^{m} v_j^2 w_j^2 - \frac{1}{4} \Big\{ w_2^2 + v_2^2 + \sum_{i=2}^{m-1} (v_i^2 w_{i+1}^2 + v_{i+1}^2 w_i^2) \Big\}$$

$$+ \frac{1}{2} \sum_{i=1}^{m-1} v_i v_{i+1} w_i w_{i+1} - \frac{1}{4} \Big\{ v_2 w_2 w_3 + w_2 v_2 v_3 + \sum_{i=2}^{m-2} (v_i v_{i+1} w_{i+1} w_{i+2} + v_{i+1} v_{i+2} w_i w_{i+1}) \Big\},$$

where we have used (3.19) and (3.20) and $w_1 = v_1 = 1/\sqrt{2}$. This proves the assertion of the Theorem. $\qquad\square$.

References

Z. D. Bai (1999). Methodologies in spectral analysis of large dimensional random matrices, a review. Statistica Sinica 9, 611-677.

F. C. Chang, J. H. B. Kemperman and W. J. Studden (1993). A normal limit theorem for moment sequences. Ann. Probab. 21, 1295-1309.

T. Chihara (1978). An Introduction to Orthogonal Polynomials. Gordon and Breach, New York.

H. Dette, F. Gamboa (2007). Asymptotic properties of the algebraic moment range process. Acta Math. Hung. 116, 247-264.

H. Dette, W. J. Studden (1997). The theory of canonical moments with applications in statistics, probability, and analysis. John Wiley & Sons Inc., New York. ISBN 0-471-10991-6. A Wiley-Interscience Publication.

B. Collins (2005). Product of random projections, Jacobi ensembles and universality problems arising from free probability. Probability Theory and Related Fields 133, 315-344.

J. Dumitriu, A. Edelman (2005). Eigenvalues of Hermite and Laguerre ensembles: large beta asymptotics. Ann. I. Henri Poincare, Probabilités et Statistiques 41, 1083-1099.

F. Gamboa and L.V. Lozada-Chang (2004). Large deviations for random power moment problem. Ann. Probab. 32, 2819-2837.

T.-C. Hu, F. Móricz, R. L. Taylor (1989). Strong laws of large numbers for arrays of rowwise independent random variables. Acta Math. Hung. 54, 153-162.

I. M. Johnstone (2008). Multivariate analysis and Jacobi ensembles: largest eigenvalue, Tracy Widom limits and rates of convergence. Annals of Statistics, to appear.

S. Karlin, L. S. Shapeley (1953). Geometry of moment spaces. Amer. Math. Soc. Memoir No. 12, Amer. Math. Soc., Providence, Rhode Island.

R. Killip, I. Nenciu I (2004). Matrix models for circular ensembles. Int. Math. Res. Not. 50, 2665-701.

M. L. Mehta (2004). Random Matrices. Academic Press.

M. Skibinsky (1967). The range of the $(n + 1)$-th moment for distributions on $[0, 1]$. J. Appl. Probability 4, 543–552.

M. Skibinsky (1968). Extreme nth moments for distributions on $[0, 1]$ and the inverse of a moment space map. J. Appl. Probability 5, 693–701.

M. Skibinsky (1969). Some striking properties of binomial and beta moments. Ann. Math. Statist. 40, 1753–1764.

G. Szegö (1975). Orthogonal Polynomials. Amer. Math. Soc. Colloqu. Publ. 23, Providence, RI.