Matrix measures, random moments and Gaussian ensembles

Jan Nagel
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: jan.nagel@rub.de

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: holger.dette@rub.de

April 15, 2009

Abstract

We consider the moment space $\mathcal{M}_n$ corresponding to $p \times p$ real or complex matrix measures defined on the interval $[0, 1]$. The asymptotic properties of the first k components of a uniformly distributed vector $(S_{1,n}, \dots, S_{n,n})^* \sim \mathcal{U}(\mathcal{M}_n)$ are studied if $n \to \infty$. In particular, it is shown that an appropriately centered and standardized version of the vector $(S_{1,n}, \dots, S_{k,n})^*$ converges weakly to a vector of k independent $p \times p$ Gaussian ensembles. For the proof of our results we use some new relations between ordinary moments and canonical moments of matrix measures which are of own interest. In particular, it is shown that the first k canonical moments corresponding to the uniform distribution on the real or complex moment space $\mathcal{M}_n$ are independent multivariate Beta distributed random variables and that each of these random variables converge in distribution (if the parameters converge to infinity) to the Gaussian orthogonal ensemble or to the Gaussian unitary ensemble, respectively.

Keyword and Phrases: Gaussian ensemble, matrix measures, canonical moments, multivariate Beta distribution, Jacobi ensemble, random matrix

AMS Subject Classification: 60F05, 15A52, 30E05

1 Introduction

A real (complex) matrix measure μ on the interval $[0, 1]$ is a $p \times p$ matrix $\mu = (\mu_{i,j})_{i,j=1}^p$ of signed real (complex) measures $\mu_{i,j}$, such that for each Borel set $A \subset [0, 1]$ the matrix $\mu(A) =$

$(\mu_{i,j}(A))_{i,j=1}^p$ is symmetric (hermitian) and nonnegative definite. Additionally, we require the matrix measure to be normalized, that is $\mu([0,1]) = I_p$, where I_p denotes the $p \times p$ identity matrix. In recent years considerable interest has been shown in generalizing many of the results on classical moment theory, orthogonal polynomials quadrature formulas etc. to the case of matrix measures. Among many others we refer to the early paper of Krein (1949) and to the more recent works of Geronimo (1982), Aptekarev and Nikishin (1983), Rodman (1990), Sinap and van Assche (1994), Duran and van Assche (1995), Duran (1995, 1996, 1999) and Duran and Lopez-Rodriguez (1996, 1997), Grünbaum (2003), Grünbaum et al. (2005) and Damanik et al. (2008) among many others.

The aim of the present paper is to explore the relations between moments of matrix measures and Gaussian ensembles, an important distribution in the area of random matrices [see Mehta (2004)]. Both fields have been investigated rather independently and in this paper we demonstrate that there exists a deep connection between random moments and Gaussian ensembles, if the "dimension" of the moment space converges to infinity. To be precise, consider the real case and recall that the moments of a real matrix measure μ on the interval $[0,1]$ are defined by

$$S_k = \int_0^1 x^k d\mu(x) \in \mathcal{S}_p(\mathbb{R}); \qquad k = 0, 1, 2, \dots \tag{1.1}$$

and the nth moment space is given by

$$\mathcal{M}_n(\mathbb{R}) = \left\{ (S_1, \dots, S_n)^T \,\middle|\, S_j = \int_0^1 x^j d\mu(x),\ j = 1, \dots, n \right\} \subset (\mathcal{S}_p(\mathbb{R}))^n , \tag{1.2}$$

where $\mathcal{S}_p(\mathbb{R})$ denotes the set of all real symmetric $p \times p$ matrices. In the scalar case $p = 1$ this space has been investigated by numerous authors [see Karlin and Shapeley (1953), Karlin and Studden (1966), Skibinsky (1967), Dette and Studden (1997)] and some of these results have been generalized to the matrix case [see Chen and Li (1999), Dette and Studden (2002) among others]. In order to understand the geometric properties of the moment space $\mathcal{M}_n(\mathbb{R})$ in the case $p = 1$, Chang et al. (1993) proposed to consider a uniform distribution on $\mathcal{M}_n(\mathbb{R})$ and studied the asymptotic properties of random moment vectors. In particular, these authors showed that an appropriately centered and standardized uniformly distributed vector on the set $\mathcal{M}_n(\mathbb{R})$ converges weakly to a multivariate normal distribution. This work was continued and substantially extended by Gamboa and Lozada-Chang (2004) and Lozada-Chang (2005), who derived a corresponding large deviation principle in the one-dimensional case.

In the present paper we will investigate related questions for the moment space (1.2) corresponding to the matrix measures on the interval $[0,1]$. More precisely, we consider a uniformly distributed vector $(S_{1,n}, \dots, S_{n,n})^T$ on $\mathcal{M}_n(\mathbb{R}) \subset (\mathcal{S}_p(\mathbb{R}))^n$ (for a precise definition see Section 2) and show that the vector of the first k matrices converges weakly after an appropriate standardization, that is

$$\sqrt{n}(A^{-1} \otimes I_p)(S_{1,n} - S_1^0, \dots, S_{k,n} - S_k^0)^T \xrightarrow[n\to\infty]{\mathcal{D}} (G_1, \dots, G_k)^T ,$$

where $A \in \mathbb{R}^{k\times k}$ is a matrix which will be specified in Section 2, $\otimes$ denotes the Kronecker product, $S_j^0 = s_j^0 I_p$,

$$s_j^0 = \int_0^1 \frac{x^j dx}{\pi\sqrt{x(1-x)}} = \frac{1}{2^{2j}}\binom{2j}{j}, \qquad j = 0, 1, 2, \ldots, \tag{1.3}$$

are the moments of the arcsine distribution and $G_1, \ldots, G_k$ are independent random $p\times p$ matrices, each distributed as the Gaussian orthogonal ensemble. The proof is based on the introduction of new "coordinates" for the moment space $\mathcal{M}_n(\mathbb{R})$. More precisely, we define a one to one map from the interior of $\mathcal{M}_n(\mathbb{R})$ onto the product space $(0_p, I_p)^n$, where 0_p is the $p \times p$ matrix with vanishing entries and $(0, I_p)$ denotes the set of all positive definite $p \times p$ matrices $C \in \mathcal{S}_p(\mathbb{R})$ for which $C < I_p$ with respect to the Loewner ordering, that is $I_p - C$ is positive definite. The new coordinates are called canonical moments [see Dette and Studden (2002)], and they are related to the Verblunsky coefficients, which have been discussed for matrix measures on the unit circle [see Damanik et al. (2008)]. We show that for a uniformly distributed vector on the nth moment space $\mathcal{M}_n(\mathbb{R})$ the corresponding canonical moments are independent and have multivariate $p\times p$ Beta distributions. Each canonical moment converges weakly (after centering and standardizing it appropriately) to the Gaussian orthogonal ensemble, and this result will be used to obtain a corresponding asymptotic result for the vector $\sqrt{n}(S_{1,n} - S_1^0, \ldots, S_{k,n} - S_k^0)^T$.

The remaining part of this paper is organized as follows. In Section 2 we introduce the basic notation, define a uniform distribution on the moment space $\mathcal{M}_n(\mathbb{R})$ and state our main result. We also determine the volume of $\mathcal{M}_n(\mathbb{R})$ defined by (1.2). In particular, it is shown that the volume behaves asymptotically as $2^{-n^2p(p+1)/2}$, which means that the moment space $\mathcal{M}_n(\mathbb{R})$ defines a very small part of $(\mathcal{S}_p(\mathbb{R}))^n$. Canonical moments of matrix measures on the interval $[0, 1]$ are introduced in Section 3. The proof of our main result is given in Section 4, which contains several results which are of own interest. In particular we prove the weak convergence of the (appropriately standardized) multivariate Beta distribution to the Gaussian ensemble. Section 5 extends these results to random moment sequences corresponding to matrix measures with complex entries. Roughly speaking, a corresponding weak convergence result is still available, where the Gaussian orthogonal ensemble has to be replaced by the Gaussian unitary ensemble. Finally, the proofs of some technical results are deferred to an Appendix in Section 6.

2 The uniform distribution on the moment space of matrix measures

Throughout this paper let $(\mathcal{S}_p(\mathbb{R}), \mathcal{B}(\mathcal{S}_p(\mathbb{R})))$ denote the measurable set of all symmetric $p \times p$ matrices with real entries, where $\mathcal{B}(\mathcal{S}_p(\mathbb{R}))$ is the Borel field corresponding to the Frobenius norm on $\mathcal{S}_p(\mathbb{R})$. In order to define a uniform distribution on the matrix moment space $\mathcal{M}_n(\mathbb{R})$

we consider on $\mathcal{S}_p(\mathbb{R})$ the integration operator

$$dX = \prod_{i \le j} dx_{i,j} \ , \tag{2.1}$$

the product Lebesgue measure with respect to the independent entries of a symmetric matrix. For an integrable function $f : \mathcal{S}_p(\mathbb{R}) \to \mathbb{R}$ the integral

$$\int f(X) dX \tag{2.2}$$

is thus the iterated integral with respect to each element $x_{i,j}, \ i \le j$ [see e.g. Muirhead (1982) or Gupta and Nagar (2000)]. We will repeatedly integrate functions $F : \mathcal{S}_p(\mathbb{R}) \to \mathcal{S}_p(\mathbb{R})$, in this case we define

$$\int F(X) dX = \left(\int (F(X))_{i,j} dX \right)_{i,j=1}^{p} \ . \tag{2.3}$$

It was shown in Dette and Studden (2002) that the moment space $\mathcal{M}_n(\mathbb{R})$ is compact and has non empty interior, say $\text{Int}(\mathcal{M}_n(\mathbb{R}))$, which enables us to define a uniform distribution on $\mathcal{M}_n(\mathbb{R})$. To be precise we introduce the matrix valued Hankel matrices

$$\underline{H}_{2m} = \begin{pmatrix} S_0 & \cdots & S_m \\ \vdots & & \vdots \\ S_m & \dots & S_{2m} \end{pmatrix} \qquad \overline{H}_{2m} = \begin{pmatrix} S_1 - S_2 & \cdots & S_m - S_{m+1} \\ \vdots & & \vdots \\ S_m - S_{m+1} & \dots & S_{2m-1} - S_{2m} \end{pmatrix} \tag{2.4}$$

and

$$\underline{H}_{2m+1} = \begin{pmatrix} S_1 & \cdots & S_{m+1} \\ \vdots & & \vdots \\ S_{m+1} & \dots & S_{2m+1} \end{pmatrix} \qquad \overline{H}_{2m+1} = \begin{pmatrix} S_0 - S_1 & \cdots & S_m - S_{m+1} \\ \vdots & & \vdots \\ S_m - S_{m+1} & \dots & S_{2m} - S_{2m+1} \end{pmatrix}. \tag{2.5}$$

Dette and Studden (2002) showed that the point $(S_1, \dots, S_n)^T$ is in the interior of the moment space $\mathcal{M}_n(\mathbb{R})$ if and only if the matrices $\underline{\mathrm{H}}_n$ and $\overline{H}_n$ are positive definite.
For a point $(S_1, \dots, S_n)^T \in \mathcal{M}_n(\mathbb{R})$ we define

$$\begin{aligned} \underline{h}_{2m}^T &= (S_{m+1}, \cdots, S_{2m}) \\ \underline{h}_{2m-1}^T &= (S_m, \cdots, S_{2m-1}) \\ \bar{h}_{2m}^T &= (S_m - S_{m+1}, \cdots, S_{2m-1} - S_{2m}) \\ \bar{h}_{2m-1}^T &= (S_m - S_{m+1}, \cdots, S_{2m-2} - S_{2m-1}) \end{aligned}$$

and consider the $p \times p$ matrices

$$S_{n+1}^- = \underline{h}_n^T \underline{H}_{n-1}^{-1} \underline{h}_n, \quad n \ge 1 \ , \tag{2.6}$$

$$S_{n+1}^+ = S_n - \bar{h}_n^T \bar{H}_{n-1}^{-1} \bar{h}_n, \quad n \ge 2 \ , \tag{2.7}$$

(for the sake of completeness we also define $S_1^- = 0$ and $S_1^+ = I_p$, $S_2^+ = S_1$). Note that S_{n+1}^- and S_{n+1}^+ are continuous functions of the moments $S_1, \dots, S_n$ and that $S_{n+1}^- < S_{n+1}^+$ if and only if $(S_1, \dots, S_n)^T \in \mathrm{Int}(\mathcal{M}_n(\mathbb{R}))$. Moreover it follows that

$$\mathcal{M}_n(\mathbb{R}) = \{(S_1, \dots, S_n)^T \mid S_1^- \le S_1 \le S_1^+, \dots, S_n^- \le S_n \le S_n^+\}, \tag{2.8}$$

$$\mathrm{Int}(\mathcal{M}_n(\mathbb{R})) = \{(S_1, \dots, S_n)^T \mid S_1^- < S_1 < S_1^+, \dots, S_n^- < S_n < S_n^+\} \neq \emptyset \tag{2.9}$$

[see Dette and Studden (2002) for more details]. Consequently the $\frac{1}{2}np(p+1)$-dimensional volume

$$\mathcal{V}(\mathcal{M}_n(\mathbb{R})) = \int_{\mathcal{M}_n(\mathbb{R})} dS_1 \dots dS_n \tag{2.10}$$

of the moment space is positive and a uniform distribution on $\mathcal{M}_n(\mathbb{R})$ is well defined by the density

$$f(S_1, \dots, S_n) = \frac{1}{\mathcal{V}(\mathcal{M}_n(\mathbb{R}))} I_{\mathcal{M}_n(\mathbb{R})}(S_1, \dots, S_n) \ . \tag{2.11}$$

For the sake of brevity we use the notation $(S_1, \dots, S_n)^T \sim \mathcal{U}(\mathcal{M}_n(\mathbb{R}))$ throughout this paper. The following result gives the volume of the moment space $\mathcal{M}_n(\mathbb{R})$. The proof will be given in Section 3 where more powerful tools have been developed for this purpose [see Remark 3.6].

Theorem 2.1. *For the real moment space $\mathcal{M}_n(\mathbb{R})$ defined in* (1.2) *we have*

$$\mathcal{V}(\mathcal{M}_n(\mathbb{R})) = \prod_{k=1}^{n} B_p\left(\tfrac{1}{2}k(p+1), \tfrac{1}{2}k(p+1)\right) \ , \tag{2.12}$$

where $B_p(a, b)$ denotes the multivariate Beta function

$$B_p(a, b) := \frac{\Gamma_p(a)\Gamma_p(b)}{\Gamma_p(a+b)} \qquad a, b > \tfrac{1}{2}(p-1) \ . \tag{2.13}$$

and $\Gamma_p(a)$ the multivariate Gamma function

$$\begin{aligned}\Gamma_p(a) :&= \int_{X>0} \det X^{a-(p+1)/2} e^{-\mathrm{tr}(X)} dX \\ &= \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma(a - \tfrac{1}{2}(i-1)) \ , \quad a > \frac{1}{2}(p-1).\end{aligned}$$

As a simple consequence of Theorem 2.1 we obtain by Stirling's formula the following approximation for the volume of the nth moment space

$$\lim_{n\to\infty} \frac{\log \mathcal{V}(\mathcal{M}_n(\mathbb{R}))}{-n^2 \frac{p(p+1)}{2} \log(2)} = 1 \ ,$$

which shows that $\mathcal{M}_n(\mathbb{R})$ consists only of a very small part of $(\mathcal{S}_p(\mathbb{R}))^n$. We will conclude this section with the main result of this paper, which gives the asymptotic distribution of the vector of the first k components of a uniform distribution on $\mathcal{M}_n(\mathbb{R})$. For this purpose recall that a random symmetric matrix X is governed by the Gaussian orthogonal ensemble (GOE), if its density is given by

$$f(X) = (2\pi)^{-p/2}\pi^{-p(p-1)/4}\, e^{-\frac{1}{2}\mathrm{tr}X^2}. \tag{2.14}$$

Theorem 2.2. *If $\mathbf{S_{n,n}} = (S_{1,n},\dots,S_{n,n})^T \sim \mathcal{U}(\mathcal{M}_n(\mathbb{R}))$, then an appropriate standardization of the vector $\mathbf{S_{k,n}} = (S_{1,n},\dots,S_{k,n})^T$ converges weakly to a vector of independent Gaussian orthogonal ensembles, that is*

$$\sqrt{4n(p+1)}(A^{-1}\otimes I_p)(\mathbf{S_{k,n}} - \mathbf{S_k^0}) \xrightarrow[n\to\infty]{\mathcal{D}} \mathbf{G}\ .$$

Here $\mathbf{S_k^0} = (s_1^0 I_p,\dots,s_k^0 I_p)^T$, s_m^0 denotes the mth moment of the arcsine distribution on the interval $[0,1]$ defined in (1.3), A is a $k\times k$ lower triangular matrix with elements $a_{i,j}$ defined by

$$a_{i,j} = 2^{-2i+2}\binom{2i}{i-j} \qquad j\le i\ , \tag{2.15}$$

and $\mathbf{G} = (G_1,\dots,G_k)^T \in (\mathcal{S}_p(\mathbb{R}))^k$ is a vector of k independent Gaussian orthogonal ensembles, i.e. $G_1,\dots,G_k$ i.i.d. $\sim$ GOE.

The proof of Theorem 2.2 is complicated and stated in Section 4, which contains also several results of own interest. It requires some explanation of the relation between the ordinary and canonical moments of a matrix measure, which will be presented in the following section.

3 Symmetric canonical moments of matrix measures

Let $(S_1,\dots,S_n)^T \in \mathrm{Int}(\mathcal{M}_n(\mathbb{R}))$ be a vector of moments of a real matrix measure on the interval $[0,1]$ and recall that the matrices S_k^+ and S_k^- given in (2.7) and (2.6), respectively, depend only on the moments $S_1,\dots S_{k-1}$. The corresponding canonical moments of the moment point $(S_1,\dots,S_n)^T$ are defined by

$$\bar{U}_k = (S_k^+ - S_k^-)^{-1}(S_k - S_k^-), \qquad k=1,\dots,n\ , \tag{3.1}$$

whenever $S_k^+ - S_k^- > 0_p$, otherwise they are left undefined [see Dette and Studden (2002)]. Note that in general the matrices $\bar{U}_1,\dots,\bar{U}_n$ are not symmetric and a symmetric version of canonical moments, say $(U_1,\dots,U_n)^T$, can easily be obtained by the transformation

$$U_k := (S_k^+ - S_k^-)^{1/2}\bar{U}_k(S_k^+ - S_k^-)^{-1/2} = (S_k^+ - S_k^-)^{-1/2}(S_k - S_k^-)(S_k^+ - S_k^-)^{-1/2}\ . \tag{3.2}$$

Throughout this paper we will work with both definitions of these matrices (note that the matrices $\bar{U}_k$ and U_k are similar). It is shown in Theorem 2.7 in Dette and Studden (2002) that the "range" of the kth moment can be expressed in terms of the canonical moments $\bar{U}_i$ and $\bar{V}_i = I_p - \bar{U}_i$ for $i = 1, \ldots, k-1$:

$$S_k^+ - S_k^- = \bar{U}_1 \bar{V}_1 \ldots \bar{U}_{k-1} \bar{V}_{k-1} \,, \qquad k = 2, \ldots, n \tag{3.3}$$

Furthermore, the moment space $\mathcal{M}_n(\mathbb{R})$ is convex and by the discussion in Section 2 it follows that $\mathcal{M}_n(\mathbb{R})$ has non empty interior. Consequently, any random variable $(S_1, \ldots, S_n)^T \sim \mathcal{U}(\mathcal{M}_n(\mathbb{R}))$ with density (2.11) satisfies $P((S_1, \ldots, S_n)^T \in \text{Int}(\mathcal{M}_n(\mathbb{R}))) = 1$ and the corresponding canonical moments $\bar{U}_1, \ldots, \bar{U}_n$ and $U_1, \ldots, U_n$ are well defined with probability 1. Moreover, it is easy to see that $S_k^- < S_k < S_k^+$ implies $0_p < U_k < I_p$ whenever $(S_1, \ldots, S_n)^T \in \text{Int}(\mathcal{M}_n(\mathbb{R}))$, since the Loewner ordering is not changed by the multiplication with positive definite matrices. Therefore equation (3.2) defines a one to one mapping

$$\varphi_p : \begin{cases} \text{Int}(\mathcal{M}_n(\mathbb{R})) & \longrightarrow \ (0_p, I_p)^n \\ (S_1, \ldots, S_n)^T & \mapsto \varphi_p(S_1, \ldots, S_n) = (U_1, \ldots, U_n)^T \,, \end{cases} \tag{3.4}$$

from the interior of the moment space onto the "cube" $(0_p, I_p)^n$, where the open interval is defined by

$$(0_p, I_p) = \{A \in \mathcal{S}_p(\mathbb{R}) \mid 0_p < A < I_p\}.$$

In the following Lemma we collect some interesting properties of the matrix valued canonical moments, which will be useful in the following discussion. The proof can be found in the Appendix.

Lemma 3.1.

(a) If μ is a matrix measure on the interval $[0, 1]$ with corresponding canonical moments U_n^μ and $\nu = \mu^\gamma$ is the measure induced on the interval $[a, b]$ by the transformation $\gamma(x) = (b-a)x + a$ $(a < b)$ with corresponding canonical moments U_n^ν, then

$$U_n^\nu = U_n^\mu,$$

whenever the canonical moments are defined. In other words: the canonical moments are invariant under linear transformations.

(b) If the matrix measure is symmetric, then

$$U_{2n-1} = \tfrac{1}{2} I_p,$$

whenever the canonical moments are defined.

(c) Let μ denote a matrix measure on the interval $[0,1]$ with canonical moments U_n^μ and define σ as the symmetric matrix measure on the interval $[-1,1]$ induced by the transformation

$$\sigma([-x,x]) = \mu([0,x^2]) \tag{3.5}$$

with corresponding canonical moments U_n^σ, then

$$U_{2n-1}^\sigma = \tfrac{1}{2} I_p \ , \qquad U_{2n}^\sigma = U_n^\mu, \tag{3.6}$$

whenever the canonical moments are defined.

The following result shows that the ordinary moments of a matrix measure can be calculated recursively from the canonical moments $\bar{U}_j$. A similar result in the scalar case was shown by Skibinsky (1968).

Theorem 3.2. *For a moment point $(S_1,\ldots,S_n)^T \in \mathrm{Int}(\mathcal{M}_n(\mathbb{R}))$ with corresponding canonical moments $\bar{U}_1,\ldots,\bar{U}_n$, define $\zeta_0 = 0_p$, $\zeta_1 = \bar{U}_1$ and*

$$\zeta_j = \bar{V}_{j-1}\bar{U}_j \ , \qquad j = 2,\ldots,n. \tag{3.7}$$

Then we have

$$S_n = G_{n,n},$$

where $\{G_{i,j},\ i,j \in (1,\ldots,n)\}$ denotes an array of $p \times p$ matrices defined by $G_{i,j} = 0_p$ if $i > j$, $G_{0,j} = I_p$ and recursively by

$$G_{i,j} = G_{i,j-1} + \zeta_{j-i+1} G_{i-1,j} \ , \tag{3.8}$$

whenever $j \geq i \geq 1$. In particular, we have

$$S_n = \sum_{i \in I} \zeta_{i_n} \cdot \ldots \cdot \zeta_{i_1},$$

where the index set is defined by $I = \{(i_1,\ldots,i_n) | i_k \in \{1,\ldots,n\},\ i_1 = 1,\ i_k \leq i_{k-1}+1\}$.

Proof: We consider the (infinite dimensional) block Hankel matrix

$$\mathbf{M} = (S_{i+j})_{i,j \geq 0} \ , \tag{3.9}$$

which contains the moments of the matrix measure μ. Let $\{P_n(x)\}_{n\geq 0}$ denote the sequence of monic (this means that $P_n(x)$ has leading term $x^n I_p$) orthogonal matrix polynomials with respect to μ, that is

$$\int P_n(x)d\mu(x)P_m^T(x) = \begin{cases} 0_p \in \mathbb{R}^{p\times p} & \text{if } n \neq m \\ D_n \in \mathbb{R}^{p\times p} & \text{if } n = m \end{cases} \tag{3.10}$$

It was shown by Sinap and van Assche (1994) that these polynomials satisfy a three term recurrence relation

$$\begin{aligned} &P_0(x) = I_p, \qquad P_1(x) = xI_p - A_1, \\ &xP_n(x) = P_{n+1}(x) + A_{n+1}P_n(x) + B_{n+1}P_{n-1}(x), \qquad n \geq 1 \ . \end{aligned}$$

We define $\mathbf{P}(x) = (P_0^T(x), P_1^T(x), P_2^T(x), \dots)^T$, and

$$\mathbf{J} = \begin{pmatrix} A_1 & I_p & 0_p & \cdots & \\ B_2 & A_2 & I_p & 0_p & \\ 0_p & B_3 & A_3 & I_p & \\ \vdots & & \ddots & \ddots & \ddots \\ & & & & \end{pmatrix}, \tag{3.11}$$

then the recursion can be rewritten in the form

$$\mathbf{JP}(x) = x\mathbf{P}(x) \ . \tag{3.12}$$

If $\mathbf{F}(x) = (I_p, xI_p, x^2I_p, \dots)^T$ denotes the vector of matrix valued monomials, then it follows that

$$\mathbf{P}(x) = \mathbf{LF}(x) \ , \tag{3.13}$$

where $\mathbf{L}$ is a lower triangular block matrix containing the coefficients of the matrix polynomials $P_n(x)$. The following Lemmata are proved in the Appendix. The first specifies the inverse of the matrix $\mathbf{L}$.

Lemma 3.3. *The matrix* $\mathbf{L}$ *defined in* (3.13) *is non singular and its inverse* $\mathbf{K} := \mathbf{L}^{-1}$ *is defined by*

$$\mathbf{RK} = \mathbf{KJ} \ , \tag{3.14}$$

where the matrix $\mathbf{R}$ *is given by*

$$\mathbf{R} = \begin{pmatrix} 0_p & I_p & 0_p & & \\ & 0_p & I_p & 0_p & \\ & & \ddots & \ddots & \ddots \\ & & & & \end{pmatrix}.$$

Moreover $\mathbf{K}$ *is a lower triangular block matrix with the matrices* I_p *on the diagonal. If* $\mathbf{D} = diag(D_0, D_1, \dots)$ *is the block diagonal matrix with entries* D_j *defined by* (3.10), *then the Hankel matrix* $\mathbf{M}$ *defined by* (3.9) *has the representation*

$$\mathbf{M} = \mathbf{KDK}^T \ . \tag{3.15}$$

Lemma 3.4. *If μ denotes a matrix measure on the interval $[0,1]$ and σ the corresponding symmetric measure on the interval $[-1,1]$ defined by (3.5), then the monic orthogonal matrix polynomials $\{P_n(x)\}_{n\geq 0}$ with respect to the matrix measure σ satisfy the recurrence relations*

$$P_0(x) = I_p, \qquad P_1(x) = xI_p,$$
$$xP_n(x) = P_{n+1}(x) + \zeta_n^T P_{n-1}(x), \qquad n \geq 1 \ , \tag{3.16}$$

where $\zeta_n = \bar{V}_{n-1}\bar{U}_n$ and $\bar{U}_n$ are the canonical moments of the measure μ.

By Lemma 3.4 the orthogonal polynomials $P_n(x)$ with respect to the matrix measure σ on the interval $[-1,1]$ are even (if n is even) or odd (if n is odd) functions. Consequently it follows from Lemma 3.3 and the representation (3.13) for the block $K_{i,j} \in \mathbb{R}^{p\times p}$ in the position (i,j) of the matrix $\mathbf{K} = \mathbf{L}^{-1}$ that $K_{i,j} = 0_p$ if $i+j$ is odd. Moreover, for the elements of the corresponding matrix $\mathbf{J}$ in (3.11) we have $A_{n+1} = 0_p$ and $B_{n+1} = \zeta_n^T$, where ζ_n corresponds to the matrix measure μ. Observing (3.14) we obtain the recursion

$$K_{i+2j,i} = K_{i+2j-1,i-1} + K_{i+2j-1,i+1}\zeta_{i+1}^T \ . \tag{3.17}$$

With the definiton

$$G_{m,n} = K_{n+m,n-m}^T \qquad \text{for } 1 \leq m \leq n,$$

(if $m > n$ we define $G_{m,n} = 0_p$) one easily sees that the matrices $G_{m,n}$ satisfy the recursion (3.8). Finally the representation (3.15) yields for the moments S_n^μ and S_n^σ of the matrix measures μ and σ the relation

$$S_n^\mu = S_{2n}^\sigma = M_{2n,0} = K_{2n,0}D_0K_{0,0} = K_{2n,0} = G_{n,n} \ ,$$

where $M_{i,j}$ denotes the $p\times p$ matrix in the position (i,j) of the matrix $\mathbf{M}$ corresponding to the matrix measure σ. This proves the first part of Theorem 3.2. The remaining statement is obvious. □

In the following we will study the distribution of the canonical moments corresponding to a random moment vector uniformly distributed on $\mathcal{M}_n(\mathbb{R})$.

Theorem 3.5. *If $(S_1,\ldots,S_n)^T \sim \mathcal{U}(\mathcal{M}_n(\mathbb{R}))$ is a random vector with a uniform distribution on the nth moment space $\mathcal{M}_n(\mathbb{R})$, then the distribution of the corresponding random vector of matrix valued canonical moments $(U_1,\ldots,U_n)^T$ is absolute continuous with respect to the Lebesgue measure and its density is given by*

$$f(U_1,\ldots,U_n) = \frac{1}{\mathcal{V}(\mathcal{M}_n(\mathbb{R}))}\prod_{k=1}^{n} \det((U_k(I_p - U_k))^{(p+1)(n-k)/2} I_{(0_p,I_p)}(U_k) \ , \tag{3.18}$$

where $\mathcal{V}(\mathcal{M}_n(\mathbb{R}))$ *is defined in* (2.10).

Note that Theorem 3.5 shows that the random variables $U_1, \dots, U_n$ corresponding to a random moment vector $(S_1, \dots, S_n)^T \sim \mathcal{U}(\mathcal{M}_n(\mathbb{R}))$ are independent and have a multivariate Beta distribution, that is

$$U_k \sim Beta_p(\tfrac{1}{2}(n-k+1)(p+1), \tfrac{1}{2}(n-k+1)(p+1)) \ ,$$

where the density of a random variable $X \sim Beta_p(a,b)$ with a matrix valued Beta distribution with parameters a and b is given by

$$f(X) = B_p(a,b)^{-1} (\det X)^{a-(p+1)/2} (\det(I-X))^{b-(p+1)/2} I_{(0_p, I_p)}(X) \tag{3.19}$$

[see Olkin and Rubin (1964) or Muirhead (1982)] and the normalizing constant $B_p(a,b)$ is defined in (2.13). Note that this definition requires $a, b > (p-1)/2$.

Proof of Theorem 3.5: We first calculate the Jacobi determinant of the mapping φ_p in (3.4) from the ordinary to the canonical moments, which we denote by $J(\varphi_p)$. By the transformation formula (3.2) we can write $\varphi_p(S_1, \dots, S_n) = (\varphi_p^{(1)}(S_1), \dots, \varphi_p^{(n)}(S_n))$, where

$$\varphi_p^{(k)} : \ (S_k^-, S_k^+) \longrightarrow (0_p, I_p) \ ,$$

$$\varphi_p^{(k)}(S_k) = (S_k^+ - S_k^-)^{-1/2} S_k (S_k^+ - S_k^-)^{-1/2} - (S_k^+ - S_k^-)^{-1/2} S_k^- (S_k^+ - S_k^-)^{-1/2} \ .$$

Note that the transformation $\varphi_p^{(k)}$ is one to one and depends only on the moments $S_1, \dots, S_{k-1}$. This implies that the Jacobian $J(\varphi_p)$ is the product of the Jacobians of the transformations $\varphi_p^{(k)}$. For fixed nonsingular matrices A and B with $B \in \mathcal{S}_p(\mathbb{R})$ the Jacobian of the transformation $X \mapsto AXA^T + B$ is equal to $(\det A)^{p+1}$, see Theorem 2.1.6 in Muirhead (1982). Consequently we obtain with the aid of equality (3.3)

$$\begin{aligned} J(\varphi_p) &= \prod_{k=1}^n J(\varphi_p^{(k)}) = \prod_{k=1}^n \det(S_k^+ - S_k^-)^{-(p+1)/2} \\ &= \prod_{k=2}^n \det(\bar{U}_1 \bar{V}_1 \dots \bar{U}_{k-1} \bar{V}_{k-1})^{-(p+1)/2} \\ &= \prod_{k=2}^n \det(U_1 V_1 \dots U_{k-1} V_{k-1})^{-(p+1)/2} = \prod_{k=1}^{n-1} \det(U_k V_k)^{-(n-k)(p+1)/2} \ . \end{aligned}$$

This gives for the density of the vector $(U_1, \dots U_n)^T$

$$f(U_1, \dots, U_n) = \frac{1}{\mathcal{V}(\mathcal{M}_n(\mathbb{R}))} I\{(U_1, \dots, U_n)^T \in \varphi_p(\mathrm{Int}(\mathcal{M}_n(\mathbb{R})))\} \prod_{k=1}^{n-1} \det(U_k V_k)^{(n-k)(p+1)/2} \ ,$$

where $\varphi_p(\mathrm{Int}(\mathcal{M}_n(\mathbb{R}))) = (0_p, I_p)^n$. □

Remark 3.6. Note that the proof of Theorem 3.5 provides also a proof of the formula for the volume of the nth moment space in Theorem 2.1, because

$$\begin{aligned}\mathcal{V}(\mathcal{M}_n(\mathbb{R})) &= \int_{(0_p,I_p)^n} \prod_{k=1}^{n-1} \det(U_k(I_p - U_k)^{(n-k)(p+1)/2} dU_1, \ldots dU_n \\ &= \prod_{k=1}^{n} B_p\left(\tfrac{1}{2}k(p+1), \tfrac{1}{2}k(p+1)\right) .\end{aligned}$$

4 The multivariate Beta distribution and a proof of Theorem 2.2

The proof of Theorem 2.2 is separated in two steps. First we investigate the asymptotic properties of the multivariate Beta distribution (Section 4.1). In particular, we show that a standardized version of the random matrix $X_n \sim Beta_p(a_n, a_n)$ converges in distribution to the GOE if the parameter a_n tends to infinity. From this result and Theorem 3.5 we obtain a weak convergence of the vector $\mathbf{U_{k,n}} = (U_{1,n}, \ldots, U_{k,n})^T$ of canonical moments corresponding to the first k components of a vector $\mathbf{S_{n,n}} = (S_{1,n}, \ldots, S_{n,n})^T \sim \mathcal{U}(\mathcal{M}_n(\mathbb{R}))$ if n tends to infinity.
Secondly, we use the relation between ordinary and canonical moments of matrix measures on the interval $[0,1]$ to prove a corresponding statement regarding the weak convergence of the vector $\mathbf{S_{k,n}} = (S_{1,n}, \ldots, S_{k,n})^T$ (Section 4.2).

4.1 Some properties of the multivariate Beta distribution

By Theorem 3.5 the multivariate Beta distribution will play a particular role in the analysis of random moment sequences of matrix measures on the interval $[0,1]$. This distribution on $\mathcal{S}_p(\mathbb{R})$ can easily be defined by its density (3.19). Since the density depends on X only through the determinant of X or $I_p - X$, the distribution of a multivariate Beta distributed random variable X is invariant under the transformation $X \mapsto OXO^T$ for any orthogonal matrix $O \in \mathcal{O}(p)$, where

$$\mathcal{O}(p) = \{O \in \mathbb{R}^{p\times p} |\ OO^T = I_p\}$$

denotes the orthogonal group. For some properties following from this invariance see Gupta and Nagar (2000), chapter 9.5. The eigenvalues of a multivariate Beta distributed random variable follow the law of the Jacobi ensemble. To be precise recall that the Jacobi ensemble is defined as the distribution of a vector $\lambda = (\lambda_1, \ldots, \lambda_p)^T$ with density

$$c_J \, |\Delta(\lambda)|^\beta \prod_{i=1}^{p} \lambda_i^{a-1}(1-\lambda_i)^{b-1} I_{(0,1)}(\lambda_i) \tag{4.1}$$

where $\Delta(\lambda) = \prod_{i<j}(\lambda_j - \lambda_i)$ is the Vandermonde determinant, $a, b, \beta > 0$ and the constant c_J is given by

$$c_J = \prod_{j=1}^{p} \frac{\Gamma(1+\frac{\beta}{2})\Gamma(a+b+\frac{\beta}{2}(p+j-2))}{\Gamma(1+\frac{\beta}{2}j)\Gamma(a+\frac{\beta}{2}(j-1))\Gamma(b+\frac{\beta}{2}(j-1))} \ , \tag{4.2}$$

see for example Dumitriu and Edelman (2002). For the sake of simplicity we write

$$\lambda \sim \mathcal{J}_\beta^{(a,b)} \tag{4.3}$$

if a random vector $\lambda = (\lambda_1, \ldots, \lambda_p)^T$ has density (4.1). Usually only the cases $\beta = 1, 2$ and 4 are considered corresponding to matrices with real, complex and quaternion entries, respectively [see Dyson (1962)]. A symmetric random variable $X \sim Beta_p(a, b)$ can be factorized as $X = Odiag(\lambda)O^T$, where $O \in \mathcal{O}(p)$ and $diag(\lambda)$ is a diagonal matrix containing the eigenvalues of X. Integration with respect to the orthogonal matrix O shows that the eigenvalues are distributed according to the Jacobi ensemble $\mathcal{J}_1^{(a-(p-1)/2, b-(p-1)/2)}$, for the calculation we refer to Muirhead (1982). We now make use of the invariance of the multivariate Beta distribution and the distribution of the eigenvalues and calculate the first moments

$$\mathrm{E}[X^k] \;=\; \int X^k f(X) dX \tag{4.4}$$

of a multivariate Beta distribution.

Lemma 4.1. *Suppose $X \sim Beta_p(a, b)$, then the moments defined by* (4.4) *satisfy*

$$\mathrm{E}[X^k] = c_k I_p \ , \tag{4.5}$$

where the constant $c_k \in \mathbb{R}$ depends on the parameters a, b and p. In particular, we obtain for the first two moments of X

$$\mathrm{E}[X] = \frac{a}{a+b} I_p \ , \tag{4.6}$$

$$\mathrm{E}[X^2] = \frac{a}{(a+b)(a+b+1)} \left(a + 1 + (p-1)\frac{b}{2a+2b-1} \right) I_p \ . \tag{4.7}$$

Proof: Because of the invariance of the multivariate Beta distribution we obtain for any orthogonal matrix U

$$\mathrm{E}[X^k]U = \mathrm{E}[UX^kU^T]U = U\mathrm{E}[X^k] \ . \tag{4.8}$$

Therefore $\mathrm{E}[X^k]$ commutes with all orthogonal matrices, which gives $\mathrm{E}[X^k] = c_k I_p$. The real constant c_k can be determined by

$$c_k = \frac{1}{p}\mathrm{tr}\mathrm{E}[X^k] = \frac{1}{p}\mathrm{E}[\mathrm{tr}X^k] = \frac{1}{p}\mathrm{E}[\lambda_1^k + \cdots + \lambda_p^k] \ , \tag{4.9}$$

where the distribution of the eigenvalues $\lambda_1, \ldots, \lambda_p$ is the Jacobi ensemble with paramters $a - \frac{1}{2}(p-1), b - \frac{1}{2}(p-1)$ and $\beta = 1$. Therefore the moment $\mathrm{E}[X^k]$ is given by

$$\mathrm{E}[\lambda_1^k] \cdot I_p = c_J \int_0^1 \ldots \int_0^1 \lambda_1^k |\Delta(\lambda)| \prod_{j=1}^p \lambda_j^{a-(p+1)/2} (1-\lambda_j)^{b-(p+1)/2} d\lambda \ \cdot I_p \ . \tag{4.10}$$

This integral is known as Aomoto's generalization of the Selberg-integral [see Aomoto (1988)]. Aomoto showed that the eigenvalues of the Jacobi ensemble $\mathcal{J}_{2\gamma}^{(\alpha,\beta)}$ satisfy

$$\mathrm{E}[\lambda_1 \cdot \ldots \cdot \lambda_m] = \prod_{i=1}^m \frac{\alpha + \gamma(p-i)}{\alpha + \beta + \gamma(2p-i-1)} \tag{4.11}$$

for $1 \leq m \leq p$. By a similar method Mehta (2004) gets the recursion

$$(\alpha + \beta + 1 + 2\gamma(p-1))\mathrm{E}[\lambda_1^2] = (\alpha + 1 + 2\gamma(p-1))\mathrm{E}[\lambda_1] - \gamma(p-1)\mathrm{E}[\lambda_1\lambda_2] \ . \tag{4.12}$$

We combine equation (4.11) and (4.12) and obtain

$$\mathrm{E}[\lambda_1^2] = \frac{\alpha + \gamma(p-1)}{(\alpha + \beta + 2\gamma(p-1))(\alpha + \beta + 1 + 2\gamma(p-1))} \tag{4.13}$$

$$\times \left((\alpha + 1 + \gamma(p-1)) + \gamma(p-1) \frac{\beta + \gamma(p-1)}{\alpha + \beta + 2\gamma(p-1) - \gamma} \right) \ . \tag{4.14}$$

This completes the proof of Lemma 4.1 if we set $\alpha = a - \frac{1}{2}(p-1)$, $\beta = b - \frac{1}{2}(p-1)$ and $\gamma = \frac{1}{2}$. □

As a next step we state a result concerning the asymptotic properties of the multivariate Beta distribution if the parameters tend to infinity.

Theorem 4.2. *Assume that $X_n \sim Beta_p(a_n, a_n)$ for a sequence $a_n/n \to \gamma \in \mathbb{R}^+$, then*

(i) $X_n \xrightarrow[n\to\infty]{L^2} \frac{1}{2} I_p$,

(ii) $\sqrt{8\gamma n}\left(X_n - \frac{1}{2}I_p\right) \xrightarrow[n\to\infty]{\mathcal{D}} G$,

where the random variable G is distributed according to the GOE.

Proof: Let $||\cdot||$ denote the Frobenius norm, then we obtain by Lemma 4.1

$$\begin{aligned} \mathrm{E}[||X_n - \tfrac{1}{2}I_p||^2] &= \mathrm{trE}[(X_n - \mathrm{E}[X_n])^2] = \mathrm{tr}(\mathrm{E}[X_n^2] - \mathrm{E}[X_n]^2) \\ &= \frac{p}{8a_n + 4}\left(1 + (p-1)\frac{2a_n}{4a_n - 1}\right) \xrightarrow[n\to\infty]{} 0 \ . \end{aligned}$$

The proof of (ii) is based on the convergence theorem of Scheffé (1947), by which it suffices to show that the density f_n of the standardized random variable $\sqrt{8\gamma n}(X_n - \frac{1}{2}I_p)$ converges pointwise to the density f of the GOE given in (2.14). The density f_n is given by

$$\begin{aligned} f_n(X) =& B_p^{-1}(a_n, a_n)(8\gamma n)^{-p(p+1)/4} \\ &\times \det\left(\frac{1}{\sqrt{8\gamma n}}X + \frac{1}{2}I_p\right)^{a_n-(p+1)/2} \det\left(\frac{1}{2}I_p - \frac{1}{\sqrt{8\gamma n}}X\right)^{a_n-(p+1)/2} I_{(-\sqrt{2\gamma n}I_p, \sqrt{2\gamma n}I_p)}(X) \\ =& B_p^{-1}(a_n, a_n)(8\gamma n)^{-p(p+1)/4}2^{-2pa_n+p(p+1)} \\ &\times \det\left(I_p - \frac{1}{2\gamma n}X^2\right)^{a_n-(p+1)/2} I_{(-\sqrt{2\gamma n}I_p, \sqrt{2\gamma n}I_p)}(X) \ . \end{aligned}$$

We can diagonalize a fixed matrix $X \in \mathcal{S}_p(\mathbb{R})$ as $X = Odiag(\lambda)O^T$, where $O \in \mathcal{O}(p)$ is an orthogonal matrix and $\lambda = (\lambda_1, \dots \lambda_p)^T$ are the eigenvalues of X. Therefore is easy to see that each factor in the last formula satisfies

$$\begin{aligned} &\det\left(I_p - \frac{1}{2\gamma n}X^2\right)^{a_n-(p+1)/2} I_{(-\sqrt{2\gamma n}I_p, \sqrt{2\gamma n}I_p)}(X) \\ =& \prod_{i=1}^p \left(1 - \frac{1}{2\gamma n}\lambda_i^2\right)^{a_n-(p+1)/2} I_{(-\sqrt{2\gamma n}, \sqrt{2\gamma n})}(\lambda_i) \\ &\xrightarrow[n\to\infty]{} \prod_{i=1}^p e^{-\frac{1}{2}\lambda_i^2} = e^{-\frac{1}{2}\mathrm{tr}X^2} \ . \end{aligned}$$

As n tends to infinity, we obtain by Stirling's formula

$$\begin{aligned} & B_p^{-1}(a_n, a_n)(8\gamma n)^{-p(p+1)/4}2^{-2pa_n+p(p+1)} \\ =& (8\gamma n)^{-p(p+1)/4}2^{-2pa_n+p(p+1)}\frac{\Gamma_p(2a_n)}{\Gamma_p(a_n)^2} \\ =& \pi^{-p(p-1)/4}(8\gamma n)^{-p(p+1)/4}2^{-2pa_n+p(p+1)}\prod_{i=1}^p \frac{\Gamma(2a_n - \frac{1}{2}(i-1))}{\Gamma(a_n - \frac{1}{2}(i-1))^2} \\ =& \pi^{-p(p-1)/4}(2\pi)^{-p/2}(1 + o(1)) \ . \end{aligned}$$

In other words, the normalization constant of the density f_n converges to the normalization constant of the GOE, which completes the proof. □

By Theorem 3.5 it follows that for a vector of matrix-valued moments $\mathbf{S_{n,n}} = (S_{1,n}, \ldots, S_{n,n})^T \sim \mathcal{U}(\mathcal{M}_n(\mathbb{R}))$ chosen uniformly from the moment space $\mathcal{M}_n(\mathbb{R})$ the corresponding canonical moments $U_{1,n}, \ldots, U_{n,n}$ are independent multivariate Beta distributed. As n tends to infinity the parameters of the Beta distributions behave as $\frac{n}{2}(p+1)$. The following Theorem is thus a direct consequence of Theorem 4.2 with $\gamma = \frac{1}{2}(p+1)$.

Theorem 4.3. *Assume that* $\mathbf{S_{n,n}} = (S_{1,n}, \ldots, S_{n,n})^T \sim \mathcal{U}(\mathcal{M}_n(\mathbb{R}))$ *and let* $\mathbf{U_{k,n}} = (U_{1,n}, \ldots, U_{k,n})^T$ *denote the vector of the first* k *canonical moments corresponding to the random variable* $\mathbf{S_{n,n}}$. *Then*

$$\sqrt{4(p+1)n}\left(\mathbf{U_{k,n}} - \mathbf{U_k^0}\right) \xrightarrow[n\to\infty]{\mathcal{D}} \mathbf{G_k} \ ,$$

where $\mathbf{U_k^0} = \frac{1}{2}(I_p, \ldots, I_p)^T$ *and* $\mathbf{G_k}$ *consists of* k *independent matrices of the* GOE.

It follows from calculations in the scalar case that the canonical moments of the arcsine distribution defined in (1.3) are all equal 1/2 [Skibinsky (1969)]. Therefore we obtain

$$\begin{aligned} \mathbf{U_n^0} &= \tfrac{1}{2}(I_p, \ldots, I_p)^T. \\ &= \varphi_p(((s_1^0 I_p, \ldots, s_n^0 I_p)^T) \\ &= \varphi_p(\mathbf{S_n^0}) \end{aligned}$$

In other words the vector $\mathbf{U_k^0}$ used in the centering of Theorem 4.3 contains the canonical moments corresponding to the matrix measure μ defined by

$$d\mu(x) = \frac{1}{\pi\sqrt{x(1-x)}} I_p dx \ . \qquad (4.15)$$

For this reason the sequence of moments $\mathbf{S_n^0}$ of the matrix measure defined by (4.15) can be viewed as the "center" of the moment space $\mathcal{M}_n(\mathbb{R})$.

4.2 Asymptotic properties of random moments

In this Section we will use the results of Section 4.1 to prove Theorem 2.2. The basic idea of the proof consists of two steps. First we show that the inverse of the mapping

$$\varphi_p : \mathrm{Int}(\mathcal{M}_k(\mathbb{R})) \longrightarrow (0_p, I_p)^k \qquad (4.16)$$

defined as in (3.4) is differentiable in a sense defined below, secondly we use this property and Theorem 4.3 to establish the weak convergence of the vector $\mathbf{S_{k,n}}$ of the first k components of $\mathbf{S_{n,n}} = (S_{1,n}, \ldots, S_{n,n})^T \sim \mathcal{U}(\mathcal{M}_n(\mathbb{R}))$. For this purpose recall that

$$\sqrt{n}(\mathbf{S_{k,n}} - \mathbf{S_k^0}) = \sqrt{n}(\varphi_p^{-1}(\mathbf{U_{k,n}}) - \varphi_p^{-1}(\mathbf{U_k^0})) \ . \qquad (4.17)$$

where $\mathbf{U_{k,n}} = (U_{1,n}, \ldots, U_{k,n})^T$ denotes the vector of canonical moments corresponding to $\mathbf{S_{k,n}}$ and $\mathbf{U^0_k} = \frac{1}{2}(I_p, \ldots, I_p)^T$. In the scalar case $p = 1$ the quantity in (4.17) can be reduced by differentiating the mapping φ_1^{-1}, that is

$$\begin{aligned}\sqrt{n}(\mathbf{s_{k,n}} - \mathbf{s^0_k}) &= \sqrt{n}(\varphi_1^{-1}(\mathbf{u_{k,n}}) - \varphi_1^{-1}(\mathbf{u^0_k})) \\ &= \sqrt{n}\frac{\partial \varphi_1^{-1}}{\partial \mathbf{u_{k,n}}}(\mathbf{u^0_k})(\mathbf{u_{k,n}} - \mathbf{u^0_k}) + \sqrt{n}\ o(||\mathbf{u_{k,n}} - \mathbf{u^0_k}) ,\end{aligned} \tag{4.18}$$

where for the sake of readability the lower capital symbols $\mathbf{s_{k,n}},\ \mathbf{s^0_k},\ \mathbf{u_{k,n}}$ and $\mathbf{u^0_k}$ denote the moment vectors $\mathbf{S_{k,n}},\ \mathbf{S^0_k},\ \mathbf{U_{k,n}}$ and $\mathbf{U^0_k}$ in the case $p = 1$. Note that

$$\frac{\partial \varphi_1^{-1}}{\partial \mathbf{u_{k,n}}}(\mathbf{u^0_k}) = A ,$$

where the elements of the matrix A were found by Chang et al. (1993) and are defined by (2.15). In order to study the general matrix case we introduce the following concept of differentiability.

Definition 4.4. *Assume that $\mathcal{S} \subset (\mathcal{S}_p(\mathbb{R}))^n$ is an open set. A mapping $\Phi : \mathcal{S} \to (\mathbb{R}^{p\times p})^m$ is called matrix differentiable in a point $\mathbf{M^0} \in \mathcal{S}$, if there exists a matrix $\mathbf{L} \in \mathbb{R}^{mp\times np}$ such that*

$$\Phi(\mathbf{M^0} + \mathbf{H}) - \Phi(\mathbf{M^0}) = \mathbf{LH} + o(||\mathbf{H}||) . \tag{4.19}$$

In this case the matrix derivative of Φ at the point $\mathbf{M^0}$ is defined by $\frac{\partial \Phi}{\partial \mathbf{M}}(\mathbf{M^0}) := \mathbf{L}$.

Note that matrix differentiability is a stronger concept than total differentiability and that a linear mapping $\Phi_1(M) = AM + B$ is matrix differentiable with $\frac{\partial \Phi_1}{\partial M} = A$. On the other hand the mapping $\Phi_2(M) = M^2$ is only matrix differentiable at the points $M^0 = mI_p$. It is easy to see that matrix differentiability has the usual properties and we note for later reference

$$\frac{\partial \Phi}{\partial \mathbf{M}} = \left(\frac{\partial \Phi_1}{\partial \mathbf{M}}^T, \ldots, \frac{\partial \Phi_m}{\partial \mathbf{M}}^T\right)^T , \tag{4.20}$$

if $\Phi = (\Phi_1^T, \ldots, \Phi_m^T)^T$ is matrix differentiable, and

$$\frac{\partial(\Phi \cdot \Psi)}{\partial \mathbf{M}}(\mathbf{M^0}) = \Psi(\mathbf{M^0})\frac{\partial \Phi}{\partial \mathbf{M}}(\mathbf{M^0}) + \Phi(\mathbf{M^0})\frac{\partial \Psi}{\partial \mathbf{M}}(\mathbf{M^0}) . \tag{4.21}$$

if $m = 1$, Φ and Ψ are matrix differentiable in $\mathbf{M^0}$ and $\Psi(\mathbf{M^0}) = cI_p$ with $c \in \mathbb{R}$. Our next result shows that the inverse of the mapping φ_p defined in (3.4) is matrix differentiable and gives the derivative. The proof is complicated and given at the end of this Section.

Theorem 4.5. *The mapping $\varphi_p^{-1} : (0_p, I_p)^k \to \mathrm{Int}(\mathcal{M}_k(\mathbb{R}))$ defined by (3.4) is matrix differentiable at the point $\mathbf{U^0} = \frac{1}{2}(I_p, \ldots, I_p)^T$ with*

$$\frac{\partial \varphi_p^{-1}}{\partial \mathbf{U}}(\mathbf{U^0}) = A \otimes I_p , \tag{4.22}$$

where A is the lower triangular matrix defined in (2.15)*.*

With the aid of Theorem 4.5 we are now in a position to complete the proof of Theorem 2.2. More precisely we obtain from (4.17) and (4.22)

$$\begin{aligned}\sqrt{n}(\mathbf{S_{k,n}}-\mathbf{S_k^0}) =& \sqrt{n}(\varphi_p^{-1}(\mathbf{U_{k,n}})-\varphi_p^{-1}(\mathbf{U_k^0}))\\ =& \sqrt{n}(A\otimes I_p)(\mathbf{U_{k,n}}-\mathbf{U_k^0})+\sqrt{n}\ o_P(||\mathbf{U_{k,n}}-\mathbf{U_k^0}||)\end{aligned}$$

and the assertion of Theorem 2.2 follows because (4.6) and (4.7) yield that the expectation of $n||\mathbf{U_{k,n}}-\mathbf{U_k^0}||^2$ converges to $pk/8$, which implies that $n||\mathbf{U_{k,n}}-\mathbf{U_k^0}||^2 = O_P(1)$. Also note that $A\otimes I_p$ is nonsingular and $(A\otimes I_p)^{-1}=A^{-1}\otimes I_p$. □

Proof of Theorem 4.5: We first study for $1\le m\le k$ the mapping

$$\psi:\begin{cases}(0_p,I_p)^k\to\mathbb{R}^{p\times p}\\ (U_1,\ldots,U_k)\mapsto \bar U_m\end{cases}\tag{4.23}$$

where $\mathbf{U}=(U_1,\ldots,U_k)^T\in(0_p,I_p)^k$ is a vector of (symmetric) canonical moments defined by (3.2) and $\bar U_m$ the mth non symmetric canonical moment defined by (3.1). Note that $\bar U_m = D_m^{-1/2}U_mD_m^{1/2}$ where

$$D_m = D_m(\mathbf{U}) = S_m^+ - S_m^-\ ,$$

and D_m satisfies the recursion $D_{m+1}=D_m^{1/2}U_mV_mD_m^{1/2}$, $D_1=I_p$ [see Theorem 2.7 in Dette and Studden (2002)]. Obviously D_m depends continuously on $U_1,\ldots,U_{m-1}$. At the point $\mathbf{U^0}$ we have $D_m(\mathbf{U^0})=D_m(\frac12 I_p,\ldots,\frac12 I_p)=(\frac12)^{2m-2}I_p$ and $\psi(\mathbf{U^0})=\frac12 I_p$. With the notation $\tilde D_m = D_m(\mathbf{U^0}+\mathbf{H})$ we obtain for $\mathbf{H}=(H_1,\ldots,H_k)^T\in\mathcal{S}_p(\mathbb{R})^k$

$$\begin{aligned}\psi(\mathbf{U^0}+\mathbf{H})-\psi(\mathbf{U^0}) &= \tilde D_m^{-1/2}(\tfrac12 I_p+H_m)\tilde D_m^{1/2}-\tfrac12 I_p\\ &= \tilde D_m^{-1/2}H_m\tilde D_m^{1/2}\\ &= I_pH_m+\left(\tilde D_m^{-1/2}H_m\tilde D_m^{1/2}-H_m\right).\end{aligned}$$

The remainder can be estimated as follows

$$\begin{aligned}\tilde D_m^{-1/2}H_m\tilde D_m^{1/2}-H_m &= \tilde D_m^{-1/2}H_m\tilde D_m^{1/2}-\tilde D_m^{-1/2}H_mD_m^{1/2}+\tilde D_m^{-1/2}H_mD_m^{1/2}-H_m\\ &= \tilde D_m^{-1/2}H_m(\tilde D_m^{1/2}-D_m^{1/2})+(\tilde D_m^{-1/2}D_m^{1/2}-I_p)H_m\\ &= o(||\mathbf{H}||)\ .\end{aligned}$$

This yields $\frac{\partial\psi}{\partial\mathbf{U}}(\mathbf{U^0})=e_m^T\otimes I_p$ and as a consequence

$$\frac{\partial\bar U_m}{\partial\mathbf{U}}(\mathbf{U^0})=\frac{\partial u_m}{\partial\mathbf{u}}(\mathbf{u^0})\otimes I_p\ ,$$

where as in (4.18) u_m denotes the mth component of the vector $\mathbf{u}$ of canonical moments in the case $p = 1$ and $\mathbf{u^0}$ is the vector of scalar canonical moments corresponding to the arcsine distribution in (1.3). A similar argument shows that

$$\frac{\partial \bar{V}_m}{\partial \mathbf{U}}(\mathbf{U^0}) = -e_m^T \otimes I_p = \frac{\partial v_m}{\partial \mathbf{u}}(\mathbf{u^0}) \otimes I_p \ ,$$

where $v_m = 1 - u_m$. By (4.21) products of canonical moments $\bar{U}_m$ and $\bar{V}_m$ are also matrix differentiable at the point $\mathbf{U^0}$ (for sums this statement is trivial) and the derivative is the Kronecker product of the derivative in the case $p = 1$ with the unit matrix. For example we can calculate for $m \neq 1$

$$\begin{aligned}\frac{\partial \zeta_m}{\partial \mathbf{U}}(\mathbf{U^0}) &= \frac{\partial \bar{V}_{m-1}\bar{U}_m}{\partial \mathbf{U}}(\mathbf{U^0}) = \tfrac{1}{2} I_p \frac{\partial \bar{V}_{m-1}}{\partial \mathbf{U}}(\mathbf{U^0}) + \tfrac{1}{2} I_p \frac{\partial \bar{U}_m}{\partial \mathbf{U}}(\mathbf{U^0}) \\ &= \tfrac{1}{2}(e_m - e_{m-1})^T \otimes I_p = \frac{\partial v_{m-1} u_m}{\partial \mathbf{u}}(\mathbf{u^0}) \otimes I_p\end{aligned}$$

and for $m = 1$

$$\frac{\partial \zeta_1}{\partial \mathbf{U}}(\mathbf{U^0}) = \frac{\partial \bar{U}_1}{\partial \mathbf{U}}(\mathbf{U^0}) = \frac{\partial u_1}{\partial \mathbf{u}}(\mathbf{u^0}) \otimes I_p \ .$$

Finally Theorem 3.2 shows that the mth moment S_m is equal to the sum over products of the matrices ζ_m. Therefore the matrix derivative of S_m with respect the canonical moments $\mathbf{U}$ is given by $\frac{\partial S_m}{\partial \mathbf{U}}(\mathbf{U^0}) = \frac{\partial s_m}{\partial \mathbf{u}}(\mathbf{u^0}) \otimes I_p$ and (4.20) yields

$$\frac{\partial \varphi_p^{-1}}{\partial \mathbf{U}}(\mathbf{U^0}) = \frac{\partial \varphi_1^{-1}}{\partial \mathbf{u}}(\mathbf{u^0}) \otimes I_p = A \otimes I_p \ .$$

This completes the proof (note that $A \otimes I_p$ is non singular). □

5 Complex random moments

To a large extend, the case of complex matrix measures can be treated analogously to the case of real matrix measures. For the sake of brevity we only state the results in this Section and omit the proofs. The kth moment of a complex matrix measure on the interval $[0, 1]$ is defined as

(5.1) $$S_k = \int_0^1 x^k d\mu(x) \in \mathcal{S}_p(\mathbb{C}); \qquad k = 0, 1, 2, \ldots$$

where $\mathcal{S}_p(\mathbb{C})$ denotes the space of $p \times p$ hermitian matrices. The complex nth moment space

(5.2) $$\mathcal{M}_n(\mathbb{C}) = \left\{ (S_1, \ldots, S_n)^* \ \middle| \ S_j = \int_0^1 x^j d\mu(x), \ j = 1, \ldots, n \right\} \subset (\mathcal{S}_p(\mathbb{C}))^n$$

is characterised by the equations (2.8) as well [see Dette and Studden (2002)]. Here $A^* = \bar{A}^T$ denotes the conjugate transpose of the matrix A. For a point $(S_1, \dots, S_n)^* \in \mathrm{Int}(\mathcal{M}_n(\mathbb{C}))$ the complex canonical moments $U_1, \dots, U_n$ are therefore well-defined, where as in the real case

$$U_k = (S_k^+ - S_k^-)^{-1/2}(S_k - S_k^-)(S_k^+ - S_k^-)^{-1/2} \tag{5.3}$$

and the hermitian matrices S_k^- and S_k^+ are defined as in (2.6) and (2.7), respectively. The integration operator changes on $\mathcal{S}_p(\mathbb{C})$ to

$$dX = \prod_{i=1}^{p} dx_{ii} \prod_{i<j} d\mathrm{Re}x_{ij} d\mathrm{Im}x_{ij} \ , \tag{5.4}$$

that is, we integrate with respect to the p^2 independent real entries of a hermitian matrix. Note that in this case for a nonsingular matrix $A \in \mathcal{S}_p(\mathbb{C})$ the Jacobian of the transformation $X \mapsto AXA$ is given by $(\det A)^{2p}$. The law of a random variable $X \in \mathcal{S}_p(\mathbb{C})$ is called complex multivariate Beta distribution with parameters $a, b > p - 1$ if its density is given by

$$f(X) = (B_p^{(2)}(a,b))^{-1} \det X^{a-p} \det(I_p - X)^{b-p} I_{(0_p, I_p)}(X) \tag{5.5}$$

and we denote this property by $X \sim Beta_p^{(2)}(a,b)$. The normalizing constant is the complex multivariate Beta function

$$B_p^{(2)}(a,b) = \frac{\Gamma_p^{(2)}(a)\Gamma_p^{(2)}(b)}{\Gamma_p^{(2)}(a+b)} \ ,$$

where $\Gamma_p^{(2)}(a) = \pi^{p(p-1)/2} \prod_{i=1}^p \Gamma(a - i + 1)$. For a more general discussion of the complex Beta distribution we refer to Khatri (1965) and Pillai and Jouris (1971). The eigenvalues of a $Beta_p^{(2)}(a,b)$-distributed random variable follow the law of the Jacobi ensemble $\mathcal{J}_2^{(a-p+1, b-p+1)}$ [see Pillai and Jouris (1971)] and similar arguments as given in the proof of Lemma 4.1 show that the first moments of a random variable $X \sim Beta_p^{(2)}(a,b)$ are given by

$$\begin{aligned} \mathrm{E}[X] &= \frac{a}{a+b} I_p \ , \\ \mathrm{E}[X^2] &= \frac{a}{(a+b)(a+b+1)} \left(a + 1 + (p-1)\frac{b}{a+b-1} \right) I_p \ . \end{aligned}$$

Proceeding as in Section 3, we get the following result for complex canonical moments.

Theorem 5.1. *Let* $\mathbf{S_{n,n}} = (S_{1,n}, \dots, S_{n,n})^*$ *be uniformly distributed on the complex moment space* $\mathcal{M}_n(\mathbb{C})$ *defined in* (5.2)*, then the corresponding canonical moments* $U_{1,n}, \dots, U_{n,n}$ *are independent and for* $k = 1, \dots, n$ $U_{k,n}$ *is complex multivariate Beta distributed with parameters* $(p(n-k+1), p(n-k+1))$.

For a sequence of complex random variables $X_n \sim Beta_p^{(2)}(a_n, a_n)$ an analoge of Theorem 4.3 holds, where in the limit the Gaussian orthogonal ensemble has to be replaced by the Gaussian unitary ensemble (GUE). Recall that a $p \times p$ hermitian matrix of the GUE is characterized by the density

$$f(X) = (2\pi)^{-p/2}\pi^{-p(p-1)/2}e^{-\frac{1}{2}\mathrm{tr}X^2} \ . \tag{5.6}$$

Theorem 5.2. *Assume that $X_n \sim Beta_p^{(2)}(a_n, a_n)$ for a sequence $a_n/n \to \gamma \in \mathbb{R}^+$, then*

(i) $X_n \xrightarrow[n\to\infty]{L^2} \frac{1}{2}I_p$,

(ii) $\sqrt{8\gamma n}\left(X_n - \frac{1}{2}I_p\right) \xrightarrow[n\to\infty]{\mathcal{D}} G$,

where the random variable G is distributed according to the GUE.

The remaining arguments in Section 4 remain essentially unchanged, which yields the following result on the weak convergence of random complex moments.

Theorem 5.3. *If $\mathbf{S_{n,n}} = (S_{1,n}, \dots, S_{n,n})^* \sim U(\mathcal{M}_n(\mathbb{C}))$, then the standardized vector of the first k moments $\mathbf{S_{k,n}} = (S_{1,n}, \dots, S_{k,n})^*$ converges weakly to a vector of independent Gaussian unitary ensembles, that is*

$$\sqrt{8np}(A^{-1} \otimes I_p)(\mathbf{S_{k,n}} - \mathbf{S_k^0}) \xrightarrow[n\to\infty]{\mathcal{D}} \mathbf{G} \ .$$

The matrix A and the vector $\mathbf{S_k^0}$ are defined as in Theorem 2.2 and $\mathbf{G} = (G_1, \dots, G_k)^*$, with $G_1, \dots, G_k$ i.i.d. $\sim$ GUE.

6 Appendix: Proof of auxiliary results

6.1 Proof of Lemma 3.1

(a) We denote by S_n and T_n the nth moment of the matrix measure μ and ν, respectively, then a straightforward calculation yields

$$T_n = \sum_{i=0}^{n-1} \binom{n}{i} a^{n-i}(b-a)^i S_i + (b-a)^n S_n \ . \tag{6.1}$$

Note that T_n^+ (T_n^-) is the unique maximal (minimal) matrix with respect to the Loewner ordering such that for fixed $T_0, \dots, T_{n-1}$ the vector $(T_0, \dots, T_n)$ is an element of the moment space of the matrix measure on the interval $[a, b]$. Therefore we obtain (note that the specification of $T_0, \ldots, T_{n-1}$ determines $S_0, \ldots, S_{n-1}$) that

$$T_n^{\pm} = \sum_{i=0}^{n-1} \binom{n}{i} a^{n-i}(b-a)^i S_i + (b-a)^n S_n^{\pm} \ .$$

This yields $T_n - T_n^{\pm} = (b-a)^n (S_n - S_n^{\pm})$, and the assertion (a) of Lemma 3.1 follows from the definition of the canonical moments in (3.2).

(b) We consider the transformation $\phi(x) = 1 - x$ and the measure $\nu = \mu^{\phi} = \mu$. The same arguments as in part (a) show

$$T_{2n} - T_{2n}^{\pm} = S_{2n} - S_{2n}^{\pm} \ , \qquad T_{2n-1} - T_{2n-1}^{\pm} = S_{2n-1}^{\mp} - S_{2n-1} \ ,$$

which implies for the corresponding canonical moments

$$U_{2n}^{\nu} = U_{2n}^{\mu} \ , \qquad U_{2n-1}^{\nu} = I_p - U_{2n-1}^{\mu} \ . \tag{6.2}$$

Because $\mu = \nu$ we obtain $U_{2n-1}^{\mu} = U_{2n-1}^{\nu} = I_p - U_{2n-1}^{\mu}$, which yields $\quad U_{2n-1}^{\mu} = \frac{1}{2} I_p$.

(c) We obtain for the moments S_k^{σ} of the matrix measure σ

$$\begin{aligned} S_{2n}^{\sigma} &= \int_{-1}^{1} t^{2n} d\sigma(t) = \int_0^1 t^n d\mu(t) = S_n \ , \\ S_{2n-1}^{\sigma} &= \int_{-1}^{1} t^{2n-1} d\sigma(t) = 0_p \ , \end{aligned} \tag{6.3}$$

where $S_1, S_2, \ldots$ denote the moments of μ. The measure σ is obviously symmetric and (b) yields $U_{2n-1}^{\sigma} = \frac{1}{2} I_p$. From (6.3) we have for the even moments

$$S_n^- \le S_{2n}^{\sigma} \le S_n^+ \ ,$$

which yields $S_{2n}^{\sigma -} = S_n^-$, $S_{2n}^{\sigma +} = S_n^+$. Consequently it follows

$$U_{2n-1}^{\sigma} = \tfrac{1}{2} I_p \ , \qquad U_{2n}^{\sigma} = U_n^{\mu} \ . \tag{6.4}$$

□

6.2 Proof of Lemma 3.3 and 3.4

Proof of Lemma 3.3: From (3.13) and (3.12) we obtain (observing that the matrix $\mathbf{R}$ acts as a shift operator) $\mathbf{LRF}(x) = x\mathbf{P}(x) = \mathbf{JP}(x) = \mathbf{JLF}(x)$, which yields

$$\mathbf{LR} = \mathbf{JL} \ . \tag{6.5}$$

It is easy to see that the matrix $\mathbf{L}$ is non singular and that the inverse matrix $\mathbf{K} := \mathbf{L}^{-1}$ is again a lower triangular block matrix with matrices I_p on the diagonal. From (6.5) we therefore obtain

$$\mathbf{RK} = \mathbf{KJ} \ .$$

On the other hand $\mathbf{F}(x) = \mathbf{KP}(x)$ and by the orthogonality relation (3.10) it follows

$$\mathbf{M} = \int \mathbf{F}(x) d\mu(x) \mathbf{F}^T(x) = \mathbf{K} \cdot \int \mathbf{P}(x) d\mu(x) \mathbf{P}^T(x) \cdot \mathbf{K}^T = \mathbf{KDK}^T \ ,$$

where the matrix $\mathbf{D} = diag(D_0, D_1, \dots)$ is defined by (3.10). □

Proof of Lemma 3.4: It follows from Favard's theorem [see Sinap and van Assche (1994) or Dette and Studden (2002)] that there exist matrices A_n, B_n such that the polynomials $\{P_n(x)\}_{n\geq 0}$ orthogonal with respect to the matrix measure σ satisfy a three term recurrence relation

$$\begin{aligned} &P_0(x) = I_p, \qquad P_1(x) = xI_p - A_1, \\ &xP_n(x) = P_{n+1}(x) + A_{n+1}P_n(x) + B_{n+1}P_{n-1}(x), \qquad n \geq 1 \ . \end{aligned}$$

We define $y = \frac{1}{2}(x+1)$ and obtain from Dette and Studden (2002) for the monic orthogonal polynomials $R_n(y) = 2^{-n}P_n(2y-1)$ with respect to the measure $\tilde{\sigma} = \sigma^{\frac{1}{2}(x+1)}$ on the interval $[0,1]$ the recursion

$$\begin{aligned} &R_0(y) = I_p, \qquad R_1(y) = yI_p - \zeta_1^{\tilde{\sigma}T}, \\ &yR_n(y) = R_{n+1}(y) + (\zeta_{2n+1}^{\tilde{\sigma}} + \zeta_{2n}^{\tilde{\sigma}})^T R_n(y) + (\zeta_{2n-1}^{\tilde{\sigma}}\zeta_{2n}^{\tilde{\sigma}})^T R_{n-1}(y), \qquad n \geq 1 \end{aligned} \tag{6.6}$$

where $\zeta_n^{\tilde{\sigma}} = \bar{V}_{n-1}^{\tilde{\sigma}}\bar{U}_n^{\tilde{\sigma}}$ for $n \geq 2$, $\zeta_1^{\tilde{\sigma}} = \bar{U}_1^{\tilde{\sigma}}$ and $\bar{U}_n^{\tilde{\sigma}}$ denote the canonical moments of the measure $\tilde{\sigma}$. Observing Lemma 3.1 (a) and (c) it follows that $\zeta_1^{\tilde{\sigma}} = \bar{U}_1^{\sigma} = \frac{1}{2}I_p$ and for $n \geq 1$

$$\zeta_{2n}^{\tilde{\sigma}} = \bar{V}_{2n-1}^{\tilde{\sigma}}\bar{U}_{2n}^{\tilde{\sigma}} = \bar{V}_{2n-1}^{\sigma}\bar{U}_{2n}^{\sigma} = \tfrac{1}{2}\bar{U}_n \ , \tag{6.7}$$

$$\zeta_{2n+1}^{\tilde{\sigma}} = \tfrac{1}{2}(I_p - \bar{U}_n) = \tfrac{1}{2}\bar{V}_n \ . \tag{6.8}$$

This yields for the polynomials $\{P_n(x)\}_{n\geq 0}$

$$P_0(x) = I_p, \qquad P_1(x) = xI_p$$

and (6.6) simplifies to

$$\tfrac{1}{2}(x+1)2^{-n}P_n(x) = 2^{-n-1}\left(P_{n+1}(x) + P_n(x) + \zeta_n^T P_{n-1}(x)\right), \qquad n \geq 1$$

which proves the assertion of Lemma 3.4. □

Acknowledgements. The authors are grateful to Martina Stein who typed parts of this paper with considerable technical expertise. The work of the authors was supported by the Sonderforschungsbereich Tr/12 (project C2, Fluctuations and universality of invariant random matrix ensembles) and in part by a grant of the Deutsche Forschungsgemeinschaft De 502/22-3.

References

Aomoto, K. (1988). Correlation functions of the selberg integral. In *Ramanujan Revisited*. Academic Press, San Diego.

Aptekarev, A. and Nikishin, E. (1983). The scattering problem for a discrete Sturm-Liouville operator. *Mat. Sb.*, 121(163):327–358.

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

Chen, G. N. and Li, Z. Q. (1999). The Nevanlinna-Pick interpolation problems and power moment problems for matrix-valued functions. *Lin. Alg. and Appl.*, 288:123–148.

Damanik, D., Pushnitski, A., and Simon, B. (2008). The analytic theory of matrix orthogonal polynomials, arxiv:0711.2703v2.

Dette, H. and Studden, W. J. (1997). *Canonical moments with applications in Statistics, Probability and Analysis*. Wiley and Sons, New York.

Dette, H. and Studden, W. J. (2002). Matrix measures, moment spaces and Favard's theorem on the interval $[0;1]$ and $[0;\infty)$. *Lin. Alg. and Appl.*, 345:169–193.

Dumitriu, I. and Edelman, A. (2002). Matrix models for beta-ensembles. *J. Math. Phys.*, 43:5830–5847.

Duran, A. J. (1995). On orthogonal polynomials with respect to a positive definite matrix of measures. *Can. J. Math.*, 47:88–112.

Duran, A. J. (1996). Markov's Theorem for orthogonal matrix polynomials. *Can. J. Math.*, 48:1180–1195.

Duran, A. J. (1999). Ratio asymptotics for orthogonal matrix polynomials. *Journal of Approximation Theory*, 100:304–344.

Duran, A. J. and Lopez-Rodriguez, P. (1996). Orthogonal matrix polynomials: zeros and Blumenthal's Theorem. *Journal of Approximation Theory*, 84:96–118.

Duran, A. J. and Lopez-Rodriguez, P. (1997). Density questions for the truncated matrix moment problem. *Can. J. Math.*, 49:708–721.

Duran, A. J. and van Assche, W. (1995). Orthogonal matrix polynomials and higher-order recurrence relations. *Lin. Alg. and Appl.*, 219:261–280.

Dyson, F. J. (1962). The threefold way. algebraic structure of symmetry groups and ensembles in quantum mechanics. *J. Math. Phys.*, 3:1199–1215.

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

Geronimo, J. S. (1982). Scattering theory and matrix orthogonal polynomials on the real line. *Circuits Systems Signal Process*, 1:471–495.

Grünbaum, F. A. (2003). Matrix valued Jacobi polynomials. *Bull. Sciences Math.*, 127:207–214.

Grünbaum, F. A., Pacharoni, I., and Tirao, J. A. (2005). Matrix valued orthogonal polynomials of the Jacobi type: The role of group representation theory. *Ann. Inst. Fourier (Grenoble)*, 55:2051–2068.

Gupta, A. K. and Nagar, D. (2000). *Matrix variate distributions*. Capman & Hall/CRC, Boca Raton.

Karlin, S. and Shapeley, L. S. (1953). Geometry of moment spaces. In *Amer. Math. Soc. Memoir No. 12*. American Mathematical Society, Providence, Rhode Island.

Karlin, S. and Studden, W. (1966). *Tchebysheff Systems: with Application in Analysis and Statistics*. Wiley, New York.

Khatri, C. G. (1965). Classical statistical analysis based on a certain multivariate complex gaussian distribution. *Annals of Mathematical Statistics*, 36:98–114.

Krein, M. G. (1949). Fundamental aspects of the representation theory of hermitian operators with deficiency index (m,m). *Ukrain. Mat. Zh.*, 1(2):3–66.

Lozada-Chang, L. V. (2005). Large deviations on moment spaces. *Electronic Journal of Probability*, 10:662–690.

Mehta, M. L. (2004). *Random matrices, Pure and Applied Mathematics*. Elsevier/Academic Press, Amsterdam.

Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. John Wiley & Sons, New York.

Olkin, I. and Rubin, H. (1964). Multivariate beta distributions and independence properties of the wishart distribution. *Ann. Math. Stat.*, 35:261–269.

Pillai, K. C. S. and Jouris, G. M. (1971). Some distribution problems in the multivariate complex gaussian case. *Annals of Mathematical Statistics*, 42:517–525.

Rodman, L. (1990). Orthogonal matrix polynomials. In Nevai, P., editor, *Orthogonal polynomials: theory and practice*, NATO ASI Series C., Volume 295. Kluwer, Dordrecht.

Scheffé, H. (1947). A useful convergence theorem for probability distributions. *Ann. Math. Stat.*, 18:434–438.

Sinap, A. and van Assche, W. (1994). Polynomial interpolation and Gaussian quadrature for matrix-valued functions. *Lin. Alg. and Appl.*, 207:71–114.

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

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

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