Journal of Multivariate Analysis 131 (2014) 17–31

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Distributions on matrix moment spaces

Holger Dette [a], Matthias Guhlich [a], Jan Nagel [b,*]

[a] Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany

[b] Technische Universität München, Fakultät für Mathematik, 85748 Garching, Germany

ARTICLE INFO

Article history:
Received 11 March 2013
Available online 27 June 2014

AMS subject classifications:
60F05
15A52
30E05

Keywords:
Moment spaces
Random matrices
Matrix measures
Random moments
Gaussian ensemble
Laguerre ensemble
Wigner's semicircle law
Marchenko–Pastur law

ABSTRACT

In this paper we define distributions on the moment spaces corresponding to $p \times p$ real or complex matrix measures on the real line with an unbounded support. For random vectors on the unbounded matricial moment spaces we prove the convergence in distribution to the Gaussian orthogonal ensemble or the Gaussian unitary ensemble, respectively.

1. Introduction

In recent years there has been considerable interest in generalizing many of the results on classical moment theory to the case of matrix measures. Wiener and Masani [26] introduced matrix measures on the unit circle in the study of multivariate stochastic processes and their spectral theory. Whittle [25] followed the same approach and established a connection to matrix polynomials, that is, polynomials with matricial coefficients. Already Karlin and McGregor [17] studied a random walk with a doubly infinite transition matrix with help of matrix polynomials, however without special consideration of the matricial structure. Delsarte et al. [3] orthogonalized polynomials with respect to matrix measures on the unit circle. Duran and van Assche [11], Duran [8,9] and Duran and Lopez-Rodriguez [10] were the first who investigated matrix orthogonal polynomials with respect to matrix measures on the real line and generalized many results from the scalar case to the matrix case. Typical examples include the three-term-recursion, quadrature formulas and ratio asymptotics. Applications to stochastic processes with two-dimensional state space were discussed by Dette et al. [6], who expressed transition probabilities and the recurrence of states in terms of matrix measures and matrix orthogonal polynomials.

In contrast to moment spaces corresponding to (probability) measures the structure of moment spaces corresponding to matrix measures is much richer and not very well understood. In the scalar case Chang et al. [2] investigated a uniform distribution on the moment space corresponding to measures on the interval [0, 1]. Their investigation was motivated by the consideration of a "typical" point in the moment space and they studied the asymptotic properties of random moment

* Corresponding author.
E-mail addresses: holger.dette@rub.de (H. Dette), jan.nagel@ma.tum.de, jan.nagel@tum.de (J. Nagel).

http://dx.doi.org/10.1016/j.jmva.2014.06.015

vectors with increasing dimension. Gamboa and Lozada-Chang [12] considered large deviation principles for random moment sequences on this space, while Lozada-Chang [18] investigated similar problems for moment spaces corresponding to more general functions defined on a bounded set. More recently, Gamboa and Rouault [14] discussed random spectral measures related to moment spaces of measures on the interval [0, 1] and moment spaces related to measures defined on the unit circle. Dette and Nagel [4] considered distributions on moment spaces corresponding to scalar measures on the real line with an unbounded support.

For matrix measures the corresponding moment of a matrix measure is given by a symmetric (Hermitian) matrix and Dette and Studden [7] obtained a characterization of the compact moment space corresponding to matrix measures on a compact interval. Dette and Nagel [5] used these results to investigate the asymptotic properties of random vectors with values in the moment space corresponding to matrix measures on the interval [0, 1]. The aim of the present paper is to get a better understanding of the properties in the non compact case. For this purpose we define probability distributions on matrix moment spaces corresponding to measures with an unbounded support and study their asymptotic behavior with an increasing dimension.

The remaining part of this paper is organized as follows. In Section 2 we introduce the basic notation, define distributions on the moment spaces corresponding to matrix measures on unbounded intervals and state our main results. In Section 3 we consider matrix orthogonal polynomials and their relation to moments of matrix measures. In Section 4 we use this relation to prove our main results. Finally in Section 5 we extend these results to matrix moment spaces corresponding to matrix measures with complex entries. Finally some technical details have been deferred to the Appendix.

2. Matrix moment spaces

Throughout this paper let $(\mathcal{S}_p(\mathbb{R}), \mathcal{B}(\mathcal{S}_p(\mathbb{R})))$ denote the measurable space of all $p \times p$ symmetric matrices with real entries, where $\mathcal{B}(\mathcal{S}_p(\mathbb{R}))$ is the Borel field corresponding to the Frobenius norm $\|A\| = \sqrt{\operatorname{tr}(A^2)}$ on $\mathcal{S}_p(\mathbb{R})$. For properties of this norm and general results in matrix theory we refer to the book of Horn and Johnson [16]. The set $\mathcal{S}_p^+(\mathbb{R}) \subset \mathcal{S}_p(\mathbb{R})$ denotes the subset of positive definite matrices and for a matrix $A \in \mathcal{S}_p(\mathbb{R})$, $|A|$ is the determinant of A. Let T be a subset of the real line with corresponding Borel field $\mathcal{B}(T)$. A ($\mathcal{S}_p(\mathbb{R})$-valued) matrix measure Σ on a measurable space $(T, \mathcal{B}(T))$ is a $p \times p$ matrix of signed measures on $(T, \mathcal{B}(T))$ such that for all Borel sets $A \subset T$ the matrix $\Sigma(A)$ is symmetric and nonnegative definite. Additionally we require the matrix measure to be normalized, that is $\Sigma(T) = I_p$, where I_p denotes the $p \times p$ identity matrix. We consider on $\mathcal{S}_p(\mathbb{R})$ the integration operator

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

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 the iterated integral with respect to each of the elements $x_{ij},\ i \le j$ (see Muirhead [20] or Gupta and Nagar [15]). The kth moment of a matrix measure is then defined as

$$M_k(\Sigma) := \int x^k d\Sigma(x) \tag{2.3}$$

for $k \ge 0$. The set of all $\mathbb{R}$-valued matrix measures on $(T, \mathcal{B}(T))$ for which all moments exist is denoted by $\mathcal{P}_p(T)$ and we define the nth moment space of matrix measures by

$$\mathcal{M}_{p,n}(T) := \left\{ (M_1(\Sigma), \dots, M_n(\Sigma))^T \,|\, \Sigma \in \mathcal{P}_p(T) \right\}. \tag{2.4}$$

Analogous to the compact case in Dette and Nagel [5] we obtain a characterization of the moment spaces $\mathcal{M}_{p,n}([0, \infty))$ and $\mathcal{M}_{p,n}(\mathbb{R})$ in terms of Hankel matrices, which are defined for matrices $M_k \in \mathcal{S}_p(\mathbb{R})$, $k \ge 0$ as

$$\underline{H}_{2m} = \begin{pmatrix} M_0 & \cdots & M_m \\ \vdots & & \vdots \\ M_m & \cdots & M_{2m} \end{pmatrix}, \qquad \overline{H}_{2m} = \begin{pmatrix} M_1 - M_2 & \cdots & M_m - M_{m+1} \\ \vdots & & \vdots \\ M_m - M_{m+1} & \cdots & M_{2m-1} - M_{2m} \end{pmatrix}, \tag{2.5}$$

and

$$\underline{H}_{2m+1} = \begin{pmatrix} M_1 & \cdots & M_{m+1} \\ \vdots & & \vdots \\ M_{m+1} & \cdots & M_{2m+1} \end{pmatrix}, \qquad \overline{H}_{2m+1} = \begin{pmatrix} M_0 - M_1 & \cdots & M_m - M_{m+1} \\ \vdots & & \vdots \\ M_m - M_{m+1} & \cdots & M_{2m} - M_{2m+1} \end{pmatrix}. \tag{2.6}$$

The following lemmas give a characterization of $\mathcal{M}_{p,n}([0, \infty))$ and $\mathcal{M}_{p,n}(\mathbb{R})$. The proof follows by similar arguments as in Dette and Studden [7] and is therefore omitted. Note that the authors consider non-normalized measures, but the arguments can be extended to matrix probability measures.

Lemma 2.1. *A vector of matrices $(M_1, \ldots, M_n)^T \in \mathcal{S}_p(\mathbb{R})^n$ is an element of the moment space $\mathcal{M}_{p,n}([0,\infty))$ if and only if for all $k \le n$ the Hankel matrices $\underline{H}_k$ are nonnegative definite.*

The vector $(M_1, \ldots, M_n)^T$ is an interior point of $\mathcal{M}_{p,n}([0,\infty))$ if and only if for all $k \le n$ the matrices $\underline{H}_k$ are positive definite.

We define the vectors of matrix moments

$$\underline{h}_{2m}^T = (M_{m+1}, \ldots, M_{2m}),$$
$$\underline{h}_{2m-1}^T = (M_m, \ldots, M_{2m-1}),$$
$$\bar{h}_{2m}^T = (M_m - M_{m+1}, \ldots, M_{2m-1} - M_{2m}),$$
$$\bar{h}_{2m-1}^T = (M_m - M_{m+1}, \ldots, M_{2m-2} - M_{2m-1}),$$

and the symmetric matrices

$$M_{n+1}^- := \underline{h}_n^T \underline{H}_{n-1}^{-1} \underline{h}_n, \quad n \ge 1, \tag{2.7}$$

$$M_{n+1}^+ := M_n - \bar{h}_n^T \bar{H}_{n-1}^{-1} \bar{h}_n, \quad n \ge 2, \tag{2.8}$$

where for the sake of completeness we set $M_1^- = 0_p$, $M_1^+ = M_0$ and $M_2^+ = M_1$. If $(M_1, \ldots, M_n)^T$ is in the interior of $\mathcal{M}_{p,n}([0,\infty))$, there exist in contrast to the compact case no upper bound for the $(n+1)$th moment. A vector of symmetric matrices $(M_1, \ldots, M_{n+1})^T$ is a moment vector in $\mathcal{M}_{p,n+1}([0,\infty))$ if and only if

$$M_{n+1}^- \le M_{n+1} \tag{2.9}$$

where the matrices are compared with respect to the Loewner ordering. That is, $A \le B$ if and only if $B - A$ is nonnegative definite and $A < B$ if and only if $B - A$ is positive definite. The inequality (2.9) follows because $M_{n+1} - M_{n+1}^-$ is the Schur complement of M_{n+1} in $\underline{H}_{n+1}$ and the matrix $\underline{H}_{n+1}$ is nonnegative definite if and only if the matrix $\underline{H}_{n-1}$ and the Schur complement $M_{n+1} - M_{n+1}^-$ are nonnegative definite. $(M_1, \ldots, M_{n+1})^T$ is an interior point of $\mathcal{M}_{p,n+1}([0,\infty))$ if and only if $M_{n+1}^- < M_{n+1}$ holds. For the remaining case of the whole real line, we obtain the following characterization of elements of the corresponding moment space.

Lemma 2.2. *The vector of matrices $(M_1, \ldots, M_n)^T \in \mathcal{S}_p(\mathbb{R})^n$ is a moment vector in $\mathcal{M}_{p,n}(\mathbb{R})$ if and only if for all k with $2k \le n$ the Hankel matrices $\underline{H}_{2k}$ are nonnegative definite.*

The vector $(M_1, \ldots, M_n)^T$ in the interior of $\mathcal{M}_{p,n}(\mathbb{R})$ if and only if for all k with $2k \le n$ the Hankel matrices $\underline{H}_{2k}$ are positive definite.

Consider the moment vector $(M_1, \ldots, M_{2n})^T \in \operatorname{Int} \mathcal{M}_{p,2n}(\mathbb{R})$. There exist no bounds (with respect to the Loewner ordering) for the next moment, that is for any $M_{2n+1} \in \mathcal{S}_p(\mathbb{R})$ the matrix vector $(M_1, \ldots, M_{2n+1})^T$ is an interior point in $\mathcal{M}_{p,2n+1}(\mathbb{R})$. For the next even moment we can define as in (2.7) the matrix M_{2n+2}^-. Then $(M_1, \ldots, M_{2n+2})^T$ is a moment vector if and only if $M_{2n+2}^- \le M_{2n+2}$ and it is in the interior of $\mathcal{M}_{p,2n+2}(\mathbb{R})$ if and only if $M_{2n+2}^- < M_{2n+2}$.

Now we define a density on the moment space $\mathcal{M}_{p,n}([0,\infty))$ by

$$g_{p,n}^{(\gamma,\delta)}(\boldsymbol{M}_n) = \prod_{k=1}^{n} c_{p,k} \left|(M_{k-1} - M_{k-1}^-)^{-1}(M_k - M_k^-)\right|^{\gamma_k} \cdot \exp\left(-\delta_k \operatorname{tr}(M_{k-1} - M_{k-1}^-)^{-1}(M_k - M_k^-)\right) \mathbb{1}_{\{M_k > M_k^-\}}, \tag{2.10}$$

where the parameters satisfy $\gamma_k > \frac{1}{2}(p-1)$ and $\delta_k > 0$ and we will show in Section 4 that the normalization constant is given by

$$c_{p,k} = \frac{\delta_k^{p\gamma_k + \frac{1}{2}p(p+1)(n-k+1)}}{\Gamma_p\left(\gamma_k + \frac{1}{2}(p+1)(n-k+1)\right)}. \tag{2.11}$$

Here Γ_p denotes the multivariate gamma function, which is defined by

$$\Gamma_p(z) := \int_{\mathcal{S}_p^+(\mathbb{R})} |X|^{z - \frac{1}{2}(p+1)} e^{-\operatorname{tr}(X)} dX$$

(see Muirhead [20]). The density $g_{p,n}^{(\gamma,\delta)}$ is the matrix analog of the density on the scalar moment space considered by Dette and Nagel [4], which in turn is the natural extension of "beta-type" densities in the compact case. Our next result gives the asymptotic distribution of the vector of the first k components of random matrix moments distributed according to the density $g_{p,n}^{(\gamma,\delta)}$. For this purpose recall that a random symmetric $p \times p$ matrix is governed by the Gaussian orthogonal ensemble (GOE_p), if its density is given by

$$f_G(X) = (2\pi)^{-p/2} \pi^{-p(p-1)/4} e^{-\frac{1}{2}\operatorname{tr} X^2}. \tag{2.12}$$

We define the matricial Marchenko–Pastur distribution η_p by

$$d\eta_p(x) := \frac{\sqrt{x(4-x)}}{2\pi x} \mathbb{1}_{\{0<x<4\}} dx \cdot I_p, \tag{2.13}$$

and obtain from the diagonal structure of this measure that the kth moment satisfies $M_k(\eta_p) = c_k I_p$, where c_k is the kth moment of the scalar Marchenko–Pastur distribution, that is, the kth Catalan number (see e.g. Tulino and Verdú [23]).

Theorem 2.3. *Assume that the vector of random moments $\boldsymbol{M}_n \in \mathcal{M}_{p,n}([0,\infty))$ has a distribution with density $g_{p,n}^{(\gamma,\delta)}$ where $\delta_i = n\frac{1}{2}(p+1)$ for all i. For $k \geq 1$ denote by $\boldsymbol{M}_k^{(n)}$ the projection of $\boldsymbol{M}_n$ onto the first k matrices and let $\boldsymbol{M}_k(\eta_p) = (c_1 I_p, \ldots, c_k I_p)^T$ contain the first k moments of the matricial Marchenko–Pastur distribution η_p. Then the convergence*

$$\sqrt{n\frac{1}{2}(p+1)}(C^{-1} \otimes I_p)(\boldsymbol{M}_k^{(n)} - \boldsymbol{M}_k(\eta_p)) \xrightarrow[n\to\infty]{\mathcal{D}} \boldsymbol{X}_k$$

holds, where $C \in \mathbb{R}^{k\times k}$ is a lower triangular matrix with entries $C_{1,1} = \cdots = C_{k,k} = 1$,

$$C_{i,j} = \binom{2i}{i-j} - \binom{2i}{i-j-1}, \quad j < i,$$

and $\boldsymbol{X}_k = (X_1, \ldots, X_k)^T$ is a vector of independent, GOE_p-distributed random matrices.

The proof of Theorems 2.3 and 2.4 below is referred to Section 4, which contains also several results of independent interest. It requires some more detailed explanation of the relation between the moment of a matrix measure and the coefficients of matrix orthogonal polynomials, which will be presented in the following section. For the definition of a density on the moment space $\mathcal{M}_{p,n}(\mathbb{R})$ we also need some basic facts about matrix polynomials. A matrix polynomial is a polynomial

$$P(x) = A_n x^n + A_{n-1}x^{n-1} + \cdots + A_0,$$

with coefficients $A_k \in \mathbb{R}^{p\times p}$. If $\Sigma \in \mathcal{P}_p(\mathbb{R})$ is a matrix measure, we define the (left) inner product on the space of matrix polynomials by

$$\prec P(x), Q(x) \succ \;:= \int P(x)d\Sigma(x)Q(x)^T = \left(\sum_{k,l=1}^{p} \int P(x)_{ik}Q(x)_{jl}d\mu_{kl}(x)\right)_{1\le i,j\le p}. \tag{2.14}$$

This inner product is matrix valued and $\mathbb{R}$-linear in both arguments. Since it is not real valued, it is not a scalar product but we nevertheless have the property that $\prec P(x), P(x) \succ = 0_p$ implies $P(x) = 0_p$. A scalar product may be defined as $\mathrm{tr} \prec \cdot, \cdot \succ$, however, we need to preserve the matrix structure of the product. The inner product (2.14) was considered by Duran [8] and several applications are discussed in Sinap and van Assche [21]. It is possible to construct matrix polynomials $P_k(x)$ orthogonal with respect to $\prec \cdot, \cdot \succ$, that is,

$$\prec P_n(x), P_m(x) \succ = 0_p \tag{2.15}$$

for $n \neq m$. The degree of a matrix polynomial $P_n(x) = \sum_{k=0}^n A_k x^k$ is n, if $A_n \neq 0_p$ and $P_n(x)$ is called monic, if $A_n = I_p$. In order to construct monic orthogonal polynomials up to degree N with respect to a matrix measure Σ, we assume that the Hankel matrix

$$\underline{H}_{2N-2} = \begin{pmatrix} M_0 & \cdots & M_{N-1} \\ \vdots & & \vdots \\ M_{N-1} & \cdots & M_{2N-2} \end{pmatrix}$$

of the moments of Σ is positive definite. By Lemma 2.2 this is equivalent to the fact that $(M_1, \ldots, M_{2N-2})^T \in \mathrm{Int}\, \mathcal{M}_{p,2N-2}(\mathbb{R})$. If $P_n(x) = \sum_{k=0}^n A_k x^k$ is a monic matrix polynomial of degree $n < N$, then for any $z \in \mathbb{R}^p \setminus \{0\}$,

$$\begin{aligned} z^T \prec P_n(x), P_n(x) \succ z &= z^T(A_0^T, \ldots, A_n^T)\underline{H}_{2n}(A_0^T, \ldots, A_n^T)^T z \\ &= (z^T A_0^T, \ldots, z^T A_n^T)\underline{H}_{2n}(z^T A_0^T, \ldots, z^T A_n^T)^T > 0. \end{aligned}$$

This shows that $\prec P_n(x), P_n(x) \succ$ is positive definite.[1] Then the Gram–Schmidt-procedure can be applied to the matrix monomials $I_p, xI_p, x^2I_p, \ldots$, which results in $P_0(x) = I_p$ and recursively

$$P_n(x) = x^n I_p - \prec x^n I_p, P_{n-1}(x) \succ \prec P_{n-1}(x), P_{n-1}(x) \succ^{-1} P_{n-1}(x) - \cdots - \prec x^n I_p, P_0(x) \succ \prec P_0(x), P_0(x) \succ^{-1} P_0(x) \tag{2.16}$$

[1] The assumption that the polynomials $P_n(x)$ is monic is indeed necessary, otherwise $\prec P_n(x), P_n(x) \succ$ would be singular for all vectors z which are in the kernel of all matrices $A_1, \ldots, A_n$.

for $1 \leq n \leq N$. The basic properties of the inner product yield (2.15) for $n \neq m$. Note that if $\underline{H}_{2N}$ is not positive definite, $\prec P_N(x), P_N(x) \succ$ may be singular and $P_{N+1}(x)$ cannot be defined. In the following discussion we suppose that $\underline{H}_{2N}$ is positive definite for all $N \geq 1$. As in the scalar case, the monic orthogonal matrix polynomials satisfy a three term recursion

$$xP_n(x) = P_{n+1}(x) + B_{n+1}P_n(x) + A_nP_{n-1}(x), \quad n \geq 1 \tag{2.17}$$

with matricial recursion coefficients $A_n, B_{n+1} \in \mathbb{R}^{p\times p}$. By an induction argument, the recursion coefficients can be recursively calculated from

$$\prec P_n(x), x^nI_p \succ = A_nA_{n-1}\dots A_0, \tag{2.18}$$

$$\prec P_n(x), x^{n+1}I_p \succ = B_{n+1}A_n\dots A_0 + A_nB_nA_{n-1}\dots A_0 + \cdots + A_n\dots A_1B_1A_0 \tag{2.19}$$

where $A_0 = \prec I_p, I_p \succ = I_p = M_0$ (see Wall [24] for the scalar versions). Eq. (2.18) gives the identity

$$A_k = \prec P_k(x), P_k(x) \succ \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1}$$

and from Eq. (2.19) we get a recursion for B_k. The mapping $\boldsymbol{M}_{2n-1} \mapsto (B_1, A_1, \dots, B_n)$ is even invertible, since Eqs. (2.18) and (2.19) allow a recursive calculation of the moments from the recursion coefficients. For our probabilistic analysis we introduce the symmetrized matrices

$$\mathcal{A}_k(M_{2n-1}) := \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} \prec P_k(x), P_k(x) \succ \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} \tag{2.20}$$

for $1 \leq k \leq n-1$ and for $1 \leq k \leq n$,

$$\mathcal{B}_k(M_{2n-1}) := \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} B_k \prec P_{k-1}(x), P_{k-1}(x) \succ^{1/2}. \tag{2.21}$$

Here we write $\mathcal{A}_k(\boldsymbol{M}_{2n-1})$ and $\mathcal{B}_k(\boldsymbol{M}_{2n-1})$ to emphasize the dependence of the recursion coefficients from the moments. Now we define a density on the moment space $\mathcal{M}_{p,2n-1}(\mathbb{R})$ by

$$h^{(\gamma,\delta)}_{2n-1}(\boldsymbol{M}_{2n-1}) = \prod_{k=1}^{n} c_{\mathcal{B}_k} \exp\left(-\delta_{2k-1}\mathrm{tr}\,\mathcal{B}_k(\boldsymbol{M}_{2n-1})^2\right) \cdot \prod_{k=1}^{n-1} c_{\mathcal{A}_k} |\mathcal{A}_k(\boldsymbol{M}_{2n-1})|^{\gamma_k} \exp\left(-\delta_{2k}\mathrm{tr}\,\mathcal{A}_k(\boldsymbol{M}_{2n-1})\right) \mathbb{1}_{\{\mathcal{A}_k > 0_p\}} \tag{2.22}$$

with parameters $\gamma_k > -1,\ \delta_k > 0$. The normalization constants $c_{\mathcal{B}_k}$ and $c_{\mathcal{A}_k}$ are given by

$$c_{\mathcal{B}_k} = \frac{1}{2^{p/2}}\left(\frac{2\delta_{2k-1}}{\pi}\right)^{p(p+1)/4}, \tag{2.23}$$

$$c_{\mathcal{A}_k} = \frac{\delta_{2k}^{p\gamma_k + p(p+1)(n-k)}}{\Gamma_p(\gamma_k + (p+1)(n-k))}. \tag{2.24}$$

Again, this density defines the natural analog of the distributions analyzed in the scalar case. The following result gives a central limit theorem for the corresponding vector of random moments. The centering constants are the moments of the matrix-semicircle distribution ρ_p defined by

$$d\rho_p(x) := \frac{1}{2\pi}\sqrt{4-x^2}\mathbb{1}_{\{-2<x<2\}}dx \cdot I_p. \tag{2.25}$$

Theorem 2.4. *Assume that the vector of random moments $\boldsymbol{M}_{2n-1} \in \mathcal{M}_{p,2n-1}(\mathbb{R})$ has a distribution with density $h^{(\gamma,\delta)}_{2n-1}$, where the parameters of the density satisfy $\delta_{2i} = n(p+1)$ and $\delta_{2i-1} = \frac{1}{2}n(p+1)$ for all $i \geq 1$. For k fixed denote by $\boldsymbol{M}^{(n)}_k = (M_1, \dots, M_k)$ the vector of the first k matrices of $\boldsymbol{M}_{2n-1}$. Then*

$$\sqrt{n(p+1)}(D^{-1} \otimes I_p)(\boldsymbol{M}^{(n)}_k - \boldsymbol{M}_k(\rho_p)) \xrightarrow[n\to\infty]{\mathcal{D}} \boldsymbol{X}_k,$$

where $\boldsymbol{M}_k(\rho_p)$ are the first moments of the matrix-semicircle distribution, the matrices in $\boldsymbol{X}_k = (X_1, \dots, X_k)^T$ are independent and GOE_p-distributed and D is a $k \times k$ lower triangular matrix with $D_{i,j} = 0$ if $i+j$ is odd, $D_{1,1} = \cdots = D_{k,k} = 1$ and the remaining entries are given by

$$D_{i,j} = \binom{i}{\frac{i-j}{2}} - \binom{i}{\frac{i-j}{2}-1}.$$

Remark 2.5. Dette and Nagel [5] proved a similar result for random matrix moments uniformly distributed on the space $\mathcal{M}_{p,n}([0,1])$. It is also possible to obtain results of this type for a more general class of densities

$$f^{(\gamma,\delta)}_{p,n}(\boldsymbol{M}_n) = \prod_{k=1}^{n} c_{p,k}|(M_k^+ - M_k^-)^{-1}(M_k - M_k^-)|^{\gamma_k}|(M_k^+ - M_k^-)^{-1}(M_k^+ - M_k)|^{\delta_k}\mathbb{1}_{\{M_k^- < M_k < M_k^+\}}$$

where $\gamma = (\gamma_k)_k$, $\delta = (\delta_k)_k$ are fixed sequences of parameters with $\gamma_k, \delta_k > \frac{1}{2}(p-1)$. If $\gamma_k = \delta_k = 0$ for all k, we obtain the uniform distribution on $\mathcal{M}_{p,n}([0,1])$ considered in this reference.

3. Matrix orthogonal polynomials

An important tool for the proof of the results in Section 2 is matrix polynomials. If the support of a matrix measure Σ is a subset of $[0, \infty)$, it follows by similar arguments as in Dette and Studden [7] that there exists a sequence of matrices $Z_n \in \mathbb{R}^{p\times p}$, such that for $n \geq 1$ the recursion coefficients in (2.17) of the monic matrix polynomials are given by

$$A_n^T = Z_{2n-1}Z_{2n}, \tag{3.1}$$

$$B_n^T = Z_{2n-2} + Z_{2n-1}, \tag{3.2}$$

where for the sake of completeness we define $Z_0 = 0_p$ (note that these authors define the inner product in a different way). Throughout this paper we call the variables Z_k (and similar quantities) canonical variables. Dette and Studden [7] also showed the representation

$$Z_n = (M_{n-1} - M_{n-1}^-)^{-1}(M_n - M_n^-) \tag{3.3}$$

for $n \geq 1$ where $M_0^- = 0_p$. Since we assume all Hankel matrices $\underline{H}_{2N}$ to be positive definite, we have $M_{n-1} > M_{n-1}^-$.

This section will provide a recursive method to calculate the moments from the recursion coefficients or from the canonical variables Z_n. A similar result in the scalar case was shown by Skibinsky [22]. We first need to make some definitions. Denote by $\boldsymbol{P}(x)^T = (P_0(x)^T, P_1(x)^T, P_2(x)^T, \ldots)$ the vector of monic orthogonal matrix polynomials and by $\boldsymbol{F}(x)^T = (I_p, xI_p, x^2I_p, \ldots)$ the vector of matricial monomials. If

$$\boldsymbol{J} = \begin{pmatrix} B_1 & I_p & & & \\ A_1 & B_2 & I_p & & \\ & A_2 & B_3 & I_p & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \tag{3.4}$$

is the block-tridiagonal matrix of the recursion coefficients, the recursion (2.17) can be written as

$$\boldsymbol{J}\boldsymbol{P}(x) = x\boldsymbol{P}(x). \tag{3.5}$$

The infinite Hankel matrix of the moments of Σ is denoted by

$$\boldsymbol{M} = (M_{i+j})_{i,j\geq 0}, \tag{3.6}$$

and let $\boldsymbol{L}$ be the lower block-triangular matrix containing the coefficients of the matrix polynomials, that is

$$\boldsymbol{P}(x) = \boldsymbol{L}\boldsymbol{F}(x). \tag{3.7}$$

Furthermore, $\boldsymbol{R}$ is the shift-operator

$$\boldsymbol{R} = \begin{pmatrix} 0_p & I_p & 0_p & \cdots & & \\ 0_p & 0_p & I_p & 0_p & & \\ \vdots & & \ddots & \ddots & \ddots & \end{pmatrix}.$$

Although we consider matrices of infinite size, a multiplication of two matrices involves only a finite sum and all products are well-defined. The following lemma is due to Dette and Nagel [5].

Lemma 3.1. *The matrix $\boldsymbol{L}$ defined by (3.7) is non-singular and the inverse $\boldsymbol{K} = \boldsymbol{L}^{-1}$ is recursively defined by*

$$\boldsymbol{R}\boldsymbol{K} = \boldsymbol{K}\boldsymbol{J}.$$

Let $\boldsymbol{D} = \mathrm{diag}(D_0, D_1, \ldots)$ be the block-diagonal matrix with diagonal entries $D_n = \prec P_n(x), P_n(x) \succ$ then the moment matrix $\boldsymbol{M}$ of Σ has the representation

$$\boldsymbol{M} = \boldsymbol{K}\boldsymbol{D}\boldsymbol{K}^T.$$

If we write out the block entries in $\boldsymbol{R}\boldsymbol{K} = \boldsymbol{K}\boldsymbol{J}$, we see that the blocks of the matrix $\boldsymbol{K}$ can be calculated with the recursion

$$K_{i+1,j} = K_{i,j-1} + K_{i,j}B_{j+1} + K_{i,j+1}A_{j+1}, \quad 0 \leq j \leq i$$

where $K_{i,j} = 0_p$ if $j < 0$ and $K_{i,i} = I_p$ (note that the numbering of the blocks in $\boldsymbol{K}$ starts at 0). By the last assertion in Lemma 3.1, the moment M_n is then given by the block in the position $n, 0$ in the matrix $\boldsymbol{K}\boldsymbol{D}\boldsymbol{K}^T$ and therefore

$$M_n = K_{n,0}D_{0,0}K_{0,0} = K_{n,0}M_0.$$

In order to obtain a recursion for the matrices Z_n, we need the following lemma to express the canonical variables as the recursion coefficients of another measure. The proof is given in the Appendix.

Lemma 3.2. *Let Σ be a matrix measure on $[0,\infty)$ and Σ_s the symmetric matrix measure on $\mathbb{R}$ defined by*

$$\Sigma_s([-x,x]) = \Sigma([0,x^2]) \tag{3.8}$$

for all $x \geq 0$. Then the monic matrix polynomials orthogonal with respect to Σ_s satisfy the recursion

$$xP_n(x) = P_{n+1}(x) + Z_n(\Sigma)^T P_{n-1}(x), \quad n \geq 0$$

where $Z_n(\Sigma)$ is the matrix in the recursion of the polynomials orthogonal with respect to Σ.

Lemma 3.1 together with Lemma 3.2 enables us to prove a recursive relation to compute the moments of Σ from the canonical variables Z_n.

Theorem 3.3. *Let $\Sigma \in \mathcal{P}_p([0,\infty))$ be a matrix probability measure with canonical variables Z_n defined in (3.3). Define the triangular array of matrices $G_{i,j}$, $i,j \geq 0$ by $G_{i,j} = 0$ if $i > j$, $G_{0,j} = I_p$ and*

$$G_{i,j} = G_{i,j-1} + Z_{j-i+1} G_{i-1,j}$$

for $j \geq i \geq 1$. Then $G_{n,n} = M_n(\Sigma)$.

Proof. For the matrix measure Σ on $[0,\infty)$ we define the symmetric measure Σ_s by (3.8). According to Lemma 3.2 the monic orthogonal polynomials $P_n(x)$ of Σ_s of even (odd) degree are even (odd) functions. This implies for the block matrix $\boldsymbol{K} = (K_{i,j})_{i,j\geq 0}$ in Lemma 3.1, which satisfies $\boldsymbol{F}(x) = \boldsymbol{KP}(x)$, that block $K_{i,j}$ is the matrix of zeros if $i+j$ is odd. Furthermore, the blocks in the matrix $\boldsymbol{J}$ of the recursion coefficients are $B_n = 0_p$ and $A_n = Z_n^T$. Here Z_n is calculated from the moments of Σ, not of Σ_s. The equation $\boldsymbol{RK} = \boldsymbol{KJ}$ gives the recursion

$$K_{i+2j,i} = K_{i+2j-1,i-1} + K_{i+2j-1,i+1} Z_{i+1}^T. \tag{3.9}$$

Now define

$$G_{i,j} = K_{i+j,j-i}^T \quad \text{for } 1 \leq i \leq j, \text{ else } G_{i,j} = 0_p,$$

then the matrices $G_{i,j}$ satisfy the asserted recursion. From the equation $\boldsymbol{M} = \boldsymbol{KDK}^T$ in Lemma 3.1 we obtain

$$M_n(\Sigma) = M_{2n}(\Sigma_s) = K_{0,0} D_0\, K_{2n,0}^T = K_{2n,0}^T = G_{n,n}.$$

Here we used that Σ is a probability measure and $D_0 = M_0 = I_p$. $\square$

4. Weak convergence of random matrix moments

4.1. Matrix measures on the half-line and proof of Theorem 2.3

4.1.1. Convergence of canonical variables

Dette and Nagel [5] use symmetrized canonical moments to study random matrix moment on the interval [0, 1]. The auxiliary variables in the study of the random moments on the half-line are symmetrized versions of the canonical variables Z_k in the recursion of matrix polynomials orthogonal on $[0,\infty)$. For moments $\boldsymbol{M}_n = (M_1,\ldots,M_n)^T$ in the interior of $\mathcal{M}_{p,n}([0,\infty))$ we define

$$\mathcal{Z}_k := (M_{k-1} - M_{k-1}^-)^{-1/2} (M_k - M_k^-)(M_{k-1} - M_{k-1}^-)^{-1/2} \in \mathcal{S}_p^+(\mathbb{R}) \tag{4.1}$$

for $k = 1,\ldots,n$. The symmetric matrix $\mathcal{Z}_k$ is positive definite and similar to Z_k. We obtain a mapping

$$\psi_{p,n} : \begin{cases} \operatorname{Int} \mathcal{M}_{p,n}^{\beta}([0,\infty)) \longrightarrow \mathcal{S}_p^+(\mathbb{R})^n \\ \boldsymbol{M}_n \longmapsto \boldsymbol{\mathcal{Z}}_n = (\mathcal{Z}_1,\ldots,\mathcal{Z}_n)^T, \end{cases} \tag{4.2}$$

that maps a moment vector to the vector of corresponding canonical variables. If $\mathcal{Z}_k$ and $M_1,\ldots,M_{k-1}$ are given, M_k can be calculated with Eq. (4.1). Recursively, all moments $M_1,\ldots,M_n$ can be recovered from canonical variables $\mathcal{Z}_1,\ldots,\mathcal{Z}_n$ and the mapping $\psi_{p,n}$ is one-to-one. The next theorem gives the distribution of the canonical variables $\mathcal{Z}_n$ of the random vector of moments with density $g_{p,n}^{(\gamma,\delta)}$ defined in (2.10). Recall that the distribution of a random variable $X \in \mathcal{S}_p(\mathbb{R})$ is given by the Laguerre orthogonal ensemble $LOE_p(\gamma, W)$ with parameter $\gamma > \frac{1}{2}(p-1)$ and scale matrix $W \in \mathcal{S}_p^+(\mathbb{R})$, if it has the density

$$f_L(X) = \frac{1}{\Gamma_p(\gamma)|W|^{\gamma}} |X|^{\gamma - \frac{1}{2}(p+1)} e^{-\operatorname{tr} W^{-1} X} \mathbb{1}_{\{X \in \mathcal{S}_p^+(\mathbb{R})\}} \tag{4.3}$$

(see e.g. Muirhead [20]). If $W = I_p$, we call this distribution central Laguerre orthogonal ensemble and write $X \sim LOE_p(\gamma)$.

Theorem 4.1. *If the vector of random moments $\boldsymbol{M}_n$ has a distribution with density $g_{p,n}^{(\gamma,\delta)}$ and $\boldsymbol{\mathcal{Z}}_n = (\mathcal{Z}_1, \ldots, \mathcal{Z}_n) = \psi_{p,n}(\boldsymbol{M}_n)$, then the random matrices $\mathcal{Z}_1, \ldots, \mathcal{Z}_n$ are independent and*

$$\mathcal{Z}_k \sim LOE_p\left(\gamma_k + \frac{1}{2}(p+1)(n-k+1), \delta_k^{-1} I_p\right)$$

for $k = 1, \ldots, n$.

Proof. By definition of $\psi_{p,n}$ the canonical variable $\mathcal{Z}_k$ depends only on $M_1, \ldots, M_k$, and consequently the Jacobian determinant is obtained as

$$\begin{aligned}\left|\frac{\partial \mathrm{vec}(\boldsymbol{\mathcal{Z}}_n)}{\partial \mathrm{vec}(\boldsymbol{M}_n)}\right| &= \prod_{k=1}^{n}\left|\frac{\partial \mathrm{vec}(\mathcal{Z}_k)}{\partial \mathrm{vec}(M_k)}\right| = \prod_{k=1}^{n}|M_{k-1} - M_{k-1}^{-}|^{-\frac{1}{2}(p+1)} \\ &= \prod_{k=2}^{n}|\mathcal{Z}_1 \ldots \mathcal{Z}_{k-1}|^{-\frac{1}{2}(p+1)} = \prod_{k=1}^{n}|\mathcal{Z}_k|^{-\frac{1}{2}(p+1)(n-k)}.\end{aligned}$$

Here we use the fact, that for $X, A \in \mathcal{S}_p(\mathbb{R})$ the determinant of the Jacobi matrix $J(\Phi_A)$ of the transformation $\Phi_A : X \mapsto AXA$ is given by

$$|J(\Phi_A)| = |A|^{p+1}$$

(see e.g. Mathai [19]). The joint density of the canonical variables $\mathcal{Z}_1, \ldots, \mathcal{Z}_n$ is given by

$$g_{\mathcal{Z}}^{(\gamma,\delta)}(\boldsymbol{\mathcal{Z}_n}) = \prod_{k=1}^{n} c_{p,k}|\mathcal{Z}_k|^{\gamma_k + \frac{1}{2}(p+1)(n-k)} \exp(-\delta_k \mathrm{tr}\,\mathcal{Z}_k)\mathbb{1}_{\{\mathcal{Z}_k > 0_p\}},$$

and we obtain the distribution according to definition (4.3). Note that this argument also gives the value of the normalization constant in (2.11). □

The following result provides the weak convergence of the Laguerre orthogonal ensemble if the parameter tends to infinity.

Theorem 4.2. *Suppose $X_n \sim LOE_p(\gamma_n)$ is a sequence of random matrices with parameters $\gamma_n \to \infty$. Then it holds that*

$$\frac{1}{\sqrt{\gamma_n}}\left(X_n - \gamma_n I_p\right) \xrightarrow[n\to\infty]{\mathcal{D}} GOE_p.$$

Proof. The moment generating function of X_n (i.e., the joint moment generating function of all real random variables in X_n) can be written as a function

$$E\left[e^{\mathrm{tr}\,(KX_n)}\right] \tag{4.4}$$

of a symmetric matrix $K \in \mathcal{S}_p(\mathbb{R})$ (compare Gupta and Nagar [15]). By definition (4.3),

$$\begin{aligned}E\left[e^{\mathrm{tr}\,(KX_n)}\right] &= \int_{\mathcal{S}_p^+(\mathbb{R})} \frac{1}{\Gamma_p(\gamma_n)}|X|^{\gamma_n - \frac{1}{2}(p+1)} e^{-\mathrm{tr}\,X} e^{\mathrm{tr}\,(KX)} dX \\ &= |I_p - K|^{-\gamma_n} \int_{\mathcal{S}_p^+(\mathbb{R})} \frac{|I_p - K|^{\gamma_n}}{\Gamma_p(\gamma_n)}|X|^{\gamma_n - \frac{1}{2}(p+1)} e^{-\mathrm{tr}\,(I_p - K)X} dX = |I_p - K|^{-\gamma_n}\end{aligned}$$

for $K < I_p$. The moment generating function of the standardized random variable is therefore given by

$$E\left[e^{\mathrm{tr}\left(K\frac{1}{\sqrt{\gamma_n}}(X_n - \gamma_n I_p)\right)}\right] = E\left[e^{\mathrm{tr}\left(\frac{1}{\sqrt{\gamma_n}}KX_n\right)}\right] e^{-\sqrt{\gamma_n}\mathrm{tr}\,K} = \left|I_p - \frac{1}{\sqrt{\gamma_n}}K\right|^{-\gamma_n} e^{-\sqrt{\gamma_n}\mathrm{tr}\,K}.$$

Let $\kappa_1, \ldots, \kappa_p$ denote the eigenvalues of the matrix K, then the moment generating function can be written as

$$\prod_{i=1}^{p}\left(1 - \frac{\kappa_i}{\sqrt{\gamma_n}}\right)^{-\gamma_n} e^{-\sqrt{\gamma_n}\kappa_i} = \prod_{i=1}^{p}\exp\left\{-\gamma_n \log\left(1 - \frac{\kappa_i}{\sqrt{\gamma_n}}\right) - \sqrt{\gamma_n}\kappa_i\right\}.$$

Now expanding the logarithm yields for $-I_p < K < I_p$

$$\begin{aligned}E\left[e^{\mathrm{tr}\left(K\frac{1}{\sqrt{\gamma_n}}(X_n - \gamma_n I_p)\right)}\right] &= \prod_{i=1}^{p}\exp\left\{\gamma_n \sum_{k=1}^{\infty}\frac{1}{k}\left(\frac{\kappa_i}{\sqrt{\gamma_n}}\right)^k - \sqrt{\gamma_n}\kappa_i\right\} \\ &= \prod_{i=1}^{p}\exp\left\{\frac{1}{2}\kappa_i^2 + \sum_{k=3}^{\infty}\frac{1}{k}\kappa_i^k\sqrt{\gamma_n}^{2-k}\right\} \xrightarrow[n\to\infty]{} \prod_{i=1}^{p}\exp\left\{\frac{1}{2}\kappa_i^2\right\} = \exp\left(\frac{1}{2}\mathrm{tr}\,K^2\right).\end{aligned}$$

Because $\exp(\frac{1}{2}\operatorname{tr} K^2)$ is the moment generating function of the Gaussian orthogonal ensemble the assertion of Theorem 2.3 follows. $\square$

The random canonical variables $\mathcal{Z}_k$ are distributed according to the Laguerre orthogonal ensemble and the weak convergence for $n \to \infty$ follows directly from Theorem 4.2. Note that $\mathcal{Z}_k \sim LOE_p(\gamma_k + \frac{1}{2}(p+1)(n-k+1), \delta_k^{-1} I_p)$ implies that

$$\delta_k \mathcal{Z}_k \sim LOE_p\left(\gamma_k + \frac{1}{2}(p+1)(n-k+1)\right)$$

is distributed according to the central Laguerre orthogonal ensemble. The appropriately standardized random matrix

$$\sqrt{\delta_k}(\mathcal{Z}_k - I_p) = \frac{1}{\sqrt{\delta_k}}(\delta_k \mathcal{Z}_k - \delta_k)$$

tends in distribution to the Gaussian orthogonal ensemble if δ_k depends on n and tends to infinity with the same rate as $\gamma_k + \frac{1}{2}(p+1)(n-k+1)$. This yields the following result.

Theorem 4.3. *Assume that the vector of random moments $\boldsymbol{M}_n$ has a distribution with density $g_{p,n}^{(\gamma,\delta)}$ defined in (2.10) and parameters $\delta_k = \frac{1}{2}(p+1)n$ for all k. Then for each component $\mathcal{Z}_k$ of $\boldsymbol{\mathcal{Z}}_n = (\mathcal{Z}_1, \dots, \mathcal{Z}_n) = \psi_{p,n}(\boldsymbol{M}_n)$ the weak convergence*

$$\sqrt{\frac{1}{2}(p+1)n}(\mathcal{Z}_k - I_p) \xrightarrow[n\to\infty]{\mathcal{D}} GOE_p$$

holds.

The non-standardized canonical variable $\mathcal{Z}_k$ converges under the assumptions of Theorem 4.3 in probability to the identity matrix. Recalling the definition of the matricial Marchenko–Pastur distribution (2.13) and definition (4.1) the diagonal structure is carried over to the canonical variables. Therefore we obtain from the scalar case (see Lemma 3.3 in Dette and Nagel [4])

$$\boldsymbol{\mathcal{Z}}_n^0 = \psi_{p,n}(\boldsymbol{M}_n(\eta_p)) = \psi_{1,n}(\boldsymbol{c}_n) \otimes I_p = (1, \dots, 1)^T \otimes I_p = (I_p, \dots, I_p)^T.$$

In other words, the random canonical variables $\mathcal{Z}_k$ converge in the situation of Theorem 4.3 in probability to the corresponding variables of the matricial Marchenko–Pastur distribution η_p. By the continuity of the mapping $\psi_{p,n}^{-1}$, the random moments M_k converge to the moments $c_k I_p$ of the matricial Marchenko–Pastur distribution.

4.1.2. Proof of Theorem 2.3

For a proof of Theorem 2.3 we need the following concept of differentiability on the level of matrices. It has been applied previously by Dette and Nagel [5] and Gamboa et al. [13] to prove weak convergence of matricial random variables in related situations.

Definition 4.4. Let $\mathcal{O}$ be an open subset of $\mathcal{S}_p(\mathbb{R})^n$. A mapping $F : \mathcal{O} \longrightarrow (\mathbb{R}^{p\times p})^m$ is called matrix differentiable at $\boldsymbol{X}^0$, if there exists a matrix $\boldsymbol{L} \in \mathbb{R}^{mp\times np}$, such that

$$F(\boldsymbol{X}^0 + \boldsymbol{H}) - F(\boldsymbol{X}^0) = \boldsymbol{L}\boldsymbol{H} + o(\|\boldsymbol{H}\|)$$

as $\|\boldsymbol{H}\| \to 0$. In this case we define the matrix derivative of F at $\boldsymbol{X}^0$ as

$$F'(\boldsymbol{X}^0) = \frac{\partial F}{\partial \boldsymbol{X}}(\boldsymbol{X}^0) = \boldsymbol{L}.$$

Matrix differentiability is a more restrictive form of Fréchet differentiability (for more details on Fréchet differentiability, we refer to Averbukh and Smolyanov [1]). In fact, we require the linear mapping which is the derivative in the Fréchet sense to be a left multiplication, so that we can apply a product rule. If $F, G : \mathcal{O} \to \mathbb{R}^{p\times p}$ are matrix differentiable at X^0, $G(X^0) = xI_p$, then $F \cdot G$ is matrix differentiable at X^0 and

$$(F\cdot G)'(\boldsymbol{X}^0) = G(\boldsymbol{X}^0)F'(\boldsymbol{X}^0) + F(\boldsymbol{X}^0)G'(\boldsymbol{X}^0).$$

Now the assertion of Theorem 2.3 is proven analogously to Theorem 2.2 in Dette and Nagel [5] and follows from Theorem 4.3 and Lemma 4.5 below, which shows that the inverse of the mapping $\psi_{p,k}$ is matrix differentiable. The matrix derivative is the Kronecker product of the derivative in the scalar case with the identity matrix.

Lemma 4.5. *The mapping $\psi_{p,k}^{-1} : \mathcal{S}_p^+(\mathbb{R})^k \longrightarrow \operatorname{Int}\mathcal{M}_{p,k}([0,\infty))$ with $\psi_{p,k}^{-1}(\boldsymbol{\mathcal{Z}}_k) = \boldsymbol{M}_k$ is matrix differentiable at $\boldsymbol{\mathcal{Z}}_k^0 = (I_p, \dots, I_p)^T$ with derivative*

$$(\psi_{p,k}^{-1})'(\boldsymbol{\mathcal{Z}}_k^0) = C \otimes I_p.$$

Proof. For $1 \le r \le k$ consider the mapping

$$F_r : \mathcal{S}_p^+(\mathbb{R})^k \to \mathbb{R}^{p\times p}, \qquad F_r(\boldsymbol{\mathcal{Z}}) = Z_r,$$

which maps the symmetric canonical variables onto the rth non-symmetric canonical variable. We have the relation

$$Z_r = (M_{r-1} - M_{r-1}^-)^{-1/2} \mathcal{Z}_r (M_{r-1} - M_{r-1}^-)^{1/2}.$$

Now arguments similar to those in the proof of Theorem 4.5 in Dette and Nagel [5] show

$$\frac{\partial \mathcal{Z}_r}{\partial \boldsymbol{\mathcal{Z}}}(\boldsymbol{\mathcal{Z}}^0) = e_r^T \otimes I_p = \frac{\partial z_r}{\partial \boldsymbol{z}}(\boldsymbol{z}^0) \otimes I_p.$$

This means that the matrix derivative is the Kronecker product of the derivative in the scalar case with the identity matrix. By Lemma 3.2, the moments are sums of products of canonical variables and so the product rule implies the same structure for the derivative of the mapping $\psi_{p,k}^{-1}$, which yields

$$\frac{\partial \psi_{p,k}^{-1}}{\partial \boldsymbol{\mathcal{Z}}}(\boldsymbol{\mathcal{Z}}_k^0) = \frac{\partial \boldsymbol{M}_k}{\partial \boldsymbol{\mathcal{Z}}}(\boldsymbol{\mathcal{Z}}_k^0) = \frac{\partial \boldsymbol{m}_k}{\partial \boldsymbol{z}}(\boldsymbol{z}_k^0) \otimes I_p,$$

and the assertion follows from the calculations in the scalar case (see Dette and Nagel [4]). □

4.2. *Matrix measures on $\mathbb{R}$ and proof of Theorem 2.4*

4.2.1. *Convergence of recursion coefficients*

In this section we prove the result for the moment space corresponding to matrix measures $\Sigma \in \mathcal{P}_p(\mathbb{R})$. In this case the auxiliary variables $\mathcal{A}_k$ and $\mathcal{B}_k$ are the symmetrized versions of the recursion coefficients of orthogonal matrix polynomials defined in Eqs. (2.20) and (2.21), respectively. Note that $\mathcal{A}_k \in \mathcal{S}_p(\mathbb{R})$ and $\mathcal{A}_k > 0_p$ which follows from the symmetry of the inner product in (2.20) and from the fact that $\prec P_k(x), P_k(x) \succ$ is positive definite. The following lemma is proven in the Appendix and shows the symmetry of the recursion variables $\mathcal{B}_k$ as well. Note that the matrices A_k and $\mathcal{A}_k$ and the matrices B_k and $\mathcal{B}_k$ are similar.

Lemma 4.6. *The matrix $\mathcal{B}_k$ defined in (2.21) is symmetric, that is $\mathcal{B}_k \in \mathcal{S}_p(\mathbb{R})$.*

With definitions (2.20) and (2.21) we construct the continuous mapping

$$\xi_{p,2n-1} : \begin{cases} \operatorname{Int} \mathcal{M}_{p,2n-1}(\mathbb{R}) \longrightarrow (\mathcal{S}_p(\mathbb{R}) \times \mathcal{S}_p^+(\mathbb{R}))^{n-1} \times \mathcal{S}_p(\mathbb{R}) \\ \boldsymbol{M}_{2n-1} \longmapsto \boldsymbol{\mathcal{R}}_{2n-1} = (\mathcal{B}_1, \mathcal{A}_1, \mathcal{B}_2, \dots, \mathcal{B}_n)^T. \end{cases} \tag{4.5}$$

The symmetric recursion coefficients are in a one-to-one correspondence with the non-symmetric coefficients and the mapping $\xi_{p,2n-1}$ is one-to-one. (Note that $A_1 = \mathcal{A}_1$ and $\prec P_n(x), P_n(x) \succ = A_1 \dots A_n$.)

Theorem 4.7. *Assume that the vector of random moments $\boldsymbol{M}_{2n-1} \in \operatorname{Int} \mathcal{M}_{p,2n-1}(\mathbb{R})$ has a distribution with density $h_{2n-1}^{(\gamma,\delta)}$. Then the random recursion coefficients $\mathcal{B}_1, \mathcal{A}_1, \dots, \mathcal{B}_n$ are independent and*

$$\sqrt{2\delta_{2k-1}}\mathcal{B}_k \sim GOE_p,$$
$$\delta_{2k}\mathcal{A}_k \sim LOE_p\left(\gamma_k + \frac{1}{2}(p+1)(2n-2k)\right).$$

Proof. By Eqs. (2.19) and (2.21) the recursion coefficient $\mathcal{B}_k$ depends only on the moments $M_1, \dots, M_{2k-1}$. The random variable $\mathcal{A}_k$ depends only on $M_1, \dots, M_{2k}$ by identity (2.20). The Jacobi matrix of $\xi_{p,2n-1}$ is therefore a lower block-triangular matrix with determinant

$$\left|\frac{\partial \operatorname{vec}(\xi_{p,2n-1}(\boldsymbol{M}_{2n-1}))}{\partial \operatorname{vec}(\boldsymbol{M}_{2n-1})}\right| = \left(\prod_{k=1}^{n} \left|\frac{\partial \operatorname{vec}(\mathcal{B}_k)}{\partial \operatorname{vec}(M_{2k-1})}\right|\right)\left(\prod_{k=1}^{n-1} \left|\frac{\partial \operatorname{vec}(\mathcal{A}_k)}{\partial \operatorname{vec}(M_{2k})}\right|\right).$$

We obtain from (2.20)

$$\mathcal{A}_k = \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} M_{2k} \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} + R_1,$$

where R_1 is independent from M_{2k} and therefore (see Mathai [19])

$$\left|\frac{\partial \operatorname{vec}(\mathcal{A}_k)}{\partial \operatorname{vec}(M_{2k})}\right| = |{\prec} P_{k-1}(x), P_{k-1}(x) {\succ}|^{-\frac{1}{2}(p+1)} = |A_{k-1} \dots A_1|^{-\frac{1}{2}(p+1)} = |\mathcal{A}_{k-1} \dots \mathcal{A}_1|^{-\frac{1}{2}(p+1)}.$$

Rearranging Eq. (2.19) and an application of (2.21) yields for the other recursion coefficient

$$\begin{aligned}\mathcal{B}_k &= \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} B_k \prec P_{k-1}(x), P_{k-1}(x) \succ^{1/2} \\ &= \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} M_{2k-1}(A_{k-1}\dots A_1)^{-1} \prec P_{k-1}(x), P_{k-1}(x) \succ^{1/2} + R_2 \\ &= \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} M_{2k-1} \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} + R_2,\end{aligned}$$

with a matrix R_2 not depending on M_{2k-1}, which gives

$$\left|\frac{\partial \mathrm{vec}(\mathcal{B}_k)}{\partial \mathrm{vec}(M_{2k-1})}\right| = | \prec P_{k-1}(x), P_{k-1}(x) \succ |^{-\frac{1}{2}(p+1)} = |\mathcal{A}_{k-1}\dots\mathcal{A}_1|^{-\frac{1}{2}(p+1)}.$$

This gives for the Jacobian determinant of $\xi_{p,2n-1}$

$$\begin{aligned}\left|\frac{\partial \mathrm{vec}(\xi_{p,2n-1}(\boldsymbol{M}_{2n-1}))}{\partial \mathrm{vec}(\boldsymbol{M}_{2n-1})}\right| &= \left(\prod_{k=2}^{n} |\mathcal{A}_{k-1}\dots\mathcal{A}_1|^{-\frac{1}{2}(p+1)}\right)\left(\prod_{k=2}^{n-1} |\mathcal{A}_{k-1}\dots\mathcal{A}_1|^{-\frac{1}{2}(p+1)}\right) \\ &= \left(\prod_{k=1}^{n-1} |\mathcal{A}_k|^{-\frac{1}{2}(p+1)(n-k)}\right)\left(\prod_{k=1}^{n-1} |\mathcal{A}_k|^{-\frac{1}{2}(p+1)(n-1-k)}\right) \\ &= \prod_{k=1}^{n-1} |\mathcal{A}_k|^{-\frac{1}{2}(p+1)(2n-2k-1)}.\end{aligned}$$

The recursion coefficients have the joint density

$$\begin{aligned}h_{\mathcal{R}}^{(\gamma,\delta)}(\mathcal{R}_{2n-1}) = &\prod_{k=1}^{n} c_{\mathcal{B}_k} \exp\left(-\delta_{2k-1}\mathrm{tr}\,\mathcal{B}_k^2\right) \\ &\cdot \prod_{k=1}^{n-1} c_{\mathcal{A}_k} |\mathcal{A}_k|^{\gamma_k+\frac{1}{2}(p+1)(2n-2k-1)} \exp\left(-\delta_{2k}\mathrm{tr}\,\mathcal{A}_k\right) \mathbb{1}_{\{\mathcal{A}_k>0_p\}}.\end{aligned}$$

The product structure of the density yields the independence of the recursion parameters and the density of $\sqrt{2\delta_{2k-1}}\mathcal{B}_k$ is given by

$$c_{\mathcal{B}_k}|\sqrt{2\delta_{2k-1}}I_p|^{-\frac{1}{2}(p+1)} \exp\left(-\frac{1}{2}\mathrm{tr}\,\mathcal{B}_k^2\right) = c_{\mathcal{B}_k}\sqrt{2\delta_{2k-1}}^{-\frac{1}{2}p(p+1)} \exp\left(-\frac{1}{2}\mathrm{tr}\,\mathcal{B}_k^2\right).$$

Consequently $\sqrt{2\delta_{2k-1}}\mathcal{B}_k \sim GOE_p$ and

$$c_{\mathcal{B}_k} = \sqrt{2\delta_{2k-1}}^{\frac{1}{2}p(p+1)}\sqrt{2\pi^{\frac{1}{2}(p+1)}}^{-p} = \frac{1}{2^{p/2}}\left(\frac{2\delta_{2k-1}}{\pi}\right)^{p(p+1)/4}.$$

The distribution of $\mathcal{A}_k$ is the non-central Laguerre orthogonal ensemble $LOE_p(\gamma_k + (p+1)(n-k), \delta_{2k}^{-1}I_p)$, which implies $\delta_{2k}\mathcal{A}_k \sim LOE_p(\gamma_k + (p+1)(n-k))$ and

$$\begin{aligned}c_{\mathcal{A}_k} &= \left(\Gamma_p(\gamma_k + (p+1)(n-k))|\delta_{2k}^{-1}I_p|^{\gamma_k+(p+1)(n-k)}\right)^{-1} \\ &= \left(\Gamma_p(\gamma_k + (p+1)(n-k))\right)^{-1}\delta_{2k}^{p\gamma_k+p(p+1)(n-k)}.\end{aligned}$$

This proves Theorem 4.7 and the identities (2.23) and (2.24). □

Theorem 4.8. *If in the situation of Theorem 4.7 the parameters satisfy $\delta_{2k} = (p+1)n$ and $\delta_{2k-1} = \frac{1}{2}(p+1)n$, then we have for the kth recursion coefficient $\mathcal{B}_k$*

$$\sqrt{(p+1)n}\mathcal{B}_k \sim GOE_p$$

and the kth rescaled recursion coefficient $\mathcal{A}_k$ converges in distribution to the Gaussian orthogonal ensemble, that is

$$\sqrt{(p+1)n}(\mathcal{A}_k - I_p) \xrightarrow[n\to\infty]{\mathcal{D}} GOE_p.$$

Theorem 4.8 is a direct consequence from the distribution in Theorem 4.7. The weak convergence of $\mathcal{A}_k$ follows from the convergence of the Laguerre orthogonal ensemble in Theorem 4.2. Under the assumptions of Theorem 4.8 the matrix vector $\mathcal{R}_k^{(n)}$, which contains the first k matrix entries of $\mathcal{R}_{2n-1} = (\mathcal{B}_1, \mathcal{A}_1, \dots, \mathcal{B}_n)^T$, converges for k fixed in probability to $\mathcal{R}_k^0$, which is the projection of

$$\mathcal{R}^0 = (0_p, I_p, 0_p, I_p, \dots)^T$$

onto the first k matrix entries. It follows directly from the scalar case in Dette and Nagel [4], that

$$\mathcal{R}_k^0 = \xi_{p,k}(\boldsymbol{M}_k^0),$$

where $\boldsymbol{M}_k^0$ is the projection of

$$\boldsymbol{M}^0 = (0_p, c_1 I_p, 0_p, c_2 I_p, \dots)^T$$

onto the first k matrices (the generalization of definition (4.5) to an even number of matricial arguments is straightforward). The vector $\boldsymbol{M}^0$ contains the moments of the semicircle distribution ρ multiplied with the identity matrix, therefore $\boldsymbol{M}^0 = \boldsymbol{M}(\rho_p)$ is the moment vector of the matrix-semicircle distribution.

4.2.2. Proof of Theorem 2.4

The proof of Theorem 2.4 now follows as the proof of Theorem 2.3 from the following lemma and the weak convergence of the random recursion coefficients in Theorem 4.8.

Lemma 4.9. *The mapping $\xi_{p,k}^{-1}$ with $\xi_{p,k}^{-1}(\mathcal{R}_k) = \boldsymbol{M}_k$ is matrix differentiable at $\mathcal{R}_k^0$ with derivative*

$$(\xi_{p,k}^{-1})'(\mathcal{R}_k^0) = D \otimes I_p.$$

Proof. A similar argumentation as in the proof of Lemma 4.5 shows

$$\frac{\partial A_i}{\partial \mathcal{R}}(\mathcal{R}_k^0) = e_{2i}^T \otimes I_p = \frac{\partial a_i}{\partial \boldsymbol{r}}(\boldsymbol{r}_k^0) \otimes I_p$$

for $2i \le k$ and

$$\frac{\partial B_i}{\partial \mathcal{R}}(\mathcal{R}_k^0) = e_{2i-1}^T \otimes I_p = \frac{\partial b_i}{\partial \boldsymbol{r}}(\boldsymbol{r}_k^0) \otimes I_p$$

for $2i - 1 \le k$. From the structure of these derivatives and Lemma 3.1 we conclude that the derivative of M_i with respect to the recursion coefficients is the Kronecker product of the derivative in the scalar case with the identity matrix and

$$\frac{\partial \xi_{p,k}^{-1}}{\partial \mathcal{R}}(\mathcal{R}_k^0) = \frac{\partial \boldsymbol{M}_k}{\partial \mathcal{R}}(\mathcal{R}_k^0) = \frac{\partial \boldsymbol{m}_k}{\partial \boldsymbol{r}}(\boldsymbol{r}_k^0) \otimes I_p = D \otimes I_p. \quad \square$$

5. Complex random moments

To a large extent, the case of complex matrix measures can be treated analogously to the case of real matrix measures. For the sake of brevity we state only the results and omit the proofs. The integration operator changes on the space of Hermitian $p \times p$ matrices $\mathcal{S}_p(\mathbb{C})$ to

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

that is, we integrate with respect to the p^2 independent real entries of a Hermitian matrix. The kth moment of a complex matrix measure on T is defined as

$$M_k(\Sigma) := \int_T x^k d\Sigma(x) \in \mathcal{S}_p(\mathbb{C}),$$

where $\mathcal{S}_p(\mathbb{C})$ denotes the space of $p \times p$ Hermitian matrices. The complex nth moment space is denoted by $\mathcal{M}_{p,n}^2(T)$.

5.1. Random moments of complex matrix measures on the half-line

We define the density on the complex moment space $\mathcal{M}_{p,n}^2([0,\infty))$ as in the real case by

$$g_{p,n}^{(\gamma,\delta)}(\boldsymbol{M}_n) = \prod_{k=1}^{n} c_{p,k} \left|(M_{k-1} - M_{k-1}^-)^{-1}(M_k - M_k^-)\right|^{\gamma_k} \cdot \exp\left(-\delta_k \mathrm{tr}\,(M_{k-1} - M_{k-1}^-)^{-1}(M_k - M_k^-)\right) \mathbb{1}_{\{M_k > M_k^-\}}, \tag{5.1}$$

where the parameters satisfy $\gamma_k > p - 1$ and $\delta_k > 0$ and the normalization constant is given by

$$c_{p,k} = \frac{\delta_k^{p\gamma_k + p^2(n-k+1)}}{\Gamma_p^{(2)}(\gamma_k + 2(n-k+1))}. \tag{5.2}$$

The canonical variables Z_k in (3.3) are well defined and as in the real case the symmetrized version is given by

$$\mathcal{Z}_k := (M_{k-1} - M_{k-1}^-)^{-1/2}(M_k - M_k^-)(M_{k-1} - M_{k-1}^-)^{-1/2} \in \mathcal{S}_p^+(\mathbb{C}) \tag{5.3}$$

for $k = 1, \ldots, n$. The Hermitian matrix $\mathcal{Z}_k$ is positive definite and similar to Z_k. We obtain a mapping

$$\psi_{p,n} : \begin{cases} \operatorname{Int} \mathcal{M}_{p,n}^2([0,\infty)) \longrightarrow \mathcal{S}_p^+(\mathbb{C})^n \\ \boldsymbol{M}_n \longmapsto \boldsymbol{\mathcal{Z}}_n = (\mathcal{Z}_1, \ldots, \mathcal{Z}_n)^*, \end{cases} \tag{5.4}$$

that maps a moment vector to the vector of corresponding canonical variables. Recall that the distribution of a random variable $X \in \mathcal{S}_p(\mathbb{C})$ is given by the Laguerre unitary ensemble $LUE_p(\gamma, W)$ with parameter $\gamma > p - 1$ and scale matrix $W \in \mathcal{S}_p^+(\mathbb{C})$, if it has the density

$$f_L(X) = \frac{1}{\Gamma_p^{(2)}(\gamma)|W|^\gamma} |X|^{\gamma-p} e^{-\operatorname{tr} W^{-1}X} \mathbb{1}_{\{X \in \mathcal{S}_p^+(\mathbb{C})\}}.$$

If $W = I_p$, we call this distribution central Laguerre unitary ensemble and write $X \sim LUE_p(\gamma)$. $\Gamma_p^{(2)}$ is the complex multivariate gamma function and given by

$$\Gamma_p^{(2)}(z) := \int_{\mathcal{S}_p^+(\mathbb{C})} |X|^{z-p} e^{-\operatorname{tr} X} dX.$$

Proceeding as in Section 4 gives the following result.

Theorem 5.1. *Assume that the vector of random moments $\boldsymbol{M}_n$ has a distribution with density $g_{p,n}^{(\gamma,\delta)}$ and $\boldsymbol{\mathcal{Z}}_n = (\mathcal{Z}_1, \ldots, \mathcal{Z}_n) = \psi_{p,n}(\boldsymbol{M}_n)$, then the random matrices $\mathcal{Z}_1, \ldots, \mathcal{Z}_n$ are independent and*

$$\mathcal{Z}_k \sim LUE_p(\gamma_k + p(n-k+1), \delta_k^{-1} I_p)$$

for $k = 1, \ldots, n$.

We proceed with asymptotic results for the random matricial moments in $\mathcal{M}_{p,n}^2([0,\infty))$ for $n \to \infty$. The following theorem provides the weak convergence of the Laguerre unitary ensemble if the parameter tends to infinity. For this purpose recall that a random variable X taking values in $\mathcal{S}_p(\mathbb{C})$ is distributed according to the Gaussian unitary ensemble GUE_p, if it has the density

$$f_G(X) = (2\pi^p)^{-p/2} \pi^{-p(p-1)/2} e^{-\frac{1}{2}\operatorname{tr} X^2}.$$

Theorem 5.2. *Suppose $X_n \sim LUE_p(\gamma_n)$ is a sequence of random matrices with parameters $\gamma_n \to \infty$. Then it holds that*

$$\frac{1}{\sqrt{\gamma_n}} \left(X_n - \gamma_n I_p \right) \xrightarrow[n\to\infty]{\mathcal{D}} GUE_p.$$

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

Theorem 5.3. *Assume that the vector of random moments $\boldsymbol{M}_n \in \mathcal{M}_{p,n}^2([0,\infty))$ has a distribution with density $g_{p,n}^{(\gamma,\delta)}$ and parameters $\delta_i = np$ for all i. For $k \geq 1$ fixed denote by $\boldsymbol{M}_k^{(n)}$ the projection of $\boldsymbol{M}_n$ onto the first k matrices and assume that $\boldsymbol{M}_k(\eta_p) = (c_1 I_p, \ldots, c_k I_p)^*$ contains the first k moments of the matricial Marchenko–Pastur distribution η_p. Then the convergence*

$$\sqrt{np}(C^{-1} \otimes I_p)(\boldsymbol{M}_k^{(n)} - \boldsymbol{M}_k(\eta_p)) \xrightarrow[n\to\infty]{\mathcal{D}} \boldsymbol{X}_k$$

holds, where $C \in \mathbb{R}^{k\times k}$ is the lower triangular matrix in Theorem 2.3 and $\boldsymbol{X}_k = (X_1, \ldots, X_k)^*$ is a vector of independent, GUE_p-distributed random matrices.

5.2. Random moments of complex matrix measures on $\mathbb{R}$

The density on $\mathcal{M}_{p,2n-1}^2(\mathbb{R})$ in the complex case is given by

$$\begin{aligned} h_{2n-1}^{(\gamma,\delta)}(\boldsymbol{M}_{2n-1}) &= \prod_{k=1}^{n} c_{\mathcal{B}_k} \exp\left(-\delta_{2k-1} \operatorname{tr} \mathcal{B}_k(\boldsymbol{M}_{2n-1})^2\right) \\ &\quad \cdot \prod_{k=1}^{n-1} c_{\mathcal{A}_k} |\mathcal{A}_k(\boldsymbol{M}_{2n-1})|^{\gamma_k} \exp\left(-\delta_{2k} \operatorname{tr} \mathcal{A}_k(\boldsymbol{M}_{2n-1})\right) \mathbb{1}_{\{\mathcal{A}_k > 0_p\}} \end{aligned} \tag{5.5}$$

with parameters $\gamma_k > -1,\ \delta_k > 0$ and the normalization constants $c_{\mathcal{B}_k}$ and $c_{\mathcal{A}_k}$ are given by

$$c_{\mathcal{B}_k} = \frac{1}{2^{p/2}}\left(\frac{2\delta_{2k-1}}{\pi}\right)^{p^2/2} \tag{5.6}$$

$$c_{\mathcal{A}_k} = \frac{\delta_{2k}^{p\gamma_k+p^2(2n-2k)}}{\Gamma_p^{(2)}(\gamma_k+p(2n-2k))}. \tag{5.7}$$

Theorem 5.4. *Assume that the vector of random moments $\boldsymbol{M}_{2n-1} \in \operatorname{Int} \mathcal{M}^2_{p,2n-1}(\mathbb{R})$ has a distribution with density $h^{(\gamma,\delta)}_{2n-1}$. Then the random recursion coefficients $\mathcal{B}_1, \mathcal{A}_1, \ldots, \mathcal{B}_n$ are independent and*

$$\sqrt{2\delta_{2k-1}}\mathcal{B}_k \sim GUE_p,$$
$$\delta_{2k}\mathcal{A}_k \sim LUE_p(\gamma_k + p(2n-2k)).$$

Theorem 5.5. *If in the situation of Theorem 5.4 the parameters satisfy $\delta_{2k} = 2np$ and $\delta_{2k-1} = np$, then we have for the recursion coefficient $\mathcal{B}_k$*

$$\sqrt{2np}\mathcal{B}_k \sim GUE_p$$

and the rescaled recursion coefficients $\mathcal{A}_k$ converge in distribution to the Gaussian unitary ensemble,

$$\sqrt{2np}(\mathcal{A}_k - I_p) \xrightarrow[n\to\infty]{\mathcal{D}} GUE_p.$$

The following result completes this section and gives a central limit theorem for the complex random moments.

Theorem 5.6. *Assume that the vector of random moments $\boldsymbol{M}_{2n-1} \in \mathcal{M}^2_{p,2n-1}(\mathbb{R})$ has a distribution with density $h^{(\gamma,\delta)}_{2n-1}$, where the parameters satisfy $\delta_{2i} = 2np$ and $\delta_{2i-1} = np$ for all $i \geq 1$. For k fixed denote by $\boldsymbol{M}_k^{(n)}$ the vector of the first k matrices in $\boldsymbol{M}_{2n-1}$. Then*

$$\sqrt{2np}(D^{-1} \otimes I_p)(\boldsymbol{M}_k^{(n)} - \boldsymbol{M}_k(\rho_p)) \xrightarrow[n\to\infty]{\mathcal{D}} \boldsymbol{X}_k,$$

where $\boldsymbol{M}_k(\rho_p)$ are the first moments of the matrix-semicircle distribution, $\boldsymbol{X}_k = (X_1, \ldots, X_k)^*$ is a vector of independent and GUE_p-distributed random variables and D is the $k \times k$ lower triangular matrix in Theorem 2.4.

Acknowledgments

The authors would also like to thank the anonymous referee for constructive comments on an earlier version of this paper. The work of the authors was supported by the Sonderforschungsbereich Tr/12 (project C2, Fluctuations and universality of invariant random matrix ensembles).

Appendix

A.1. Proof of Lemma 3.2

The monic polynomials orthogonal with respect to Σ_s satisfy a recursion as in (2.17). Eq. (2.16) and a simple induction argument yield that $P_{2n}(x)$ is an even function and $P_{2n-1}(x)$ is an odd function and consequently $B_n = 0_p$ for all n. For the other recursion coefficients we have

$$A_n^T = (M_{2n-2}(\Sigma_s) - M_{2n-2}(\Sigma_s)^-)^{-1}(M_{2n}(\Sigma_s) - M_{2n}(\Sigma_s)^-).$$

A suitable approximation of the identity by step functions shows that the even moments of Σ_s are given by

$$M_{2n}(\Sigma_s) = \int x^{2n} d\Sigma_s(x) = \int x^n d\Sigma(x) = M_n(\Sigma).$$

Therefore $M_n(\Sigma)^- \geq M_{2n}(\Sigma_s)^-$ and $M_{2n}(\Sigma_s)^- \geq M_n(\Sigma)^-$, which implies $M_{2n}(\Sigma_s)^- = M_n(\Sigma)^-$. We get for the recursion coefficient

$$A_n^T = (M_{n-1}(\Sigma) - M_{n-1}(\Sigma)^-)^{-1}(M_n(\Sigma) - M_n(\Sigma)^-) = Z_n(\Sigma). \quad \square$$

A.2. Proof of Lemma 4.6

Denote by $C^{(n)}_{n-1}$ the coefficient of x^{n-1} in $P_n(x)$, then by the recursion (2.17) we have

$$C^{(k)}_{k-1} = -B_k + C^{(k-1)}_{k-2}.$$

The construction of the matrix orthogonal polynomials in (2.16) yields

$$C_{k-1}^{(k)} = - \prec x^k I_p, P_{k-1}(x) \succ \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1}.$$

Another application of the recursion gives for the recursion coefficient

$$\begin{aligned}
B_k &= C_{k-2}^{(k-1)} - C_{k-1}^{(k)} \\
&= \prec x^k I_p, P_{k-1}(x) \succ \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1} \\
&\quad - \prec x^{k-1} I_p, P_{k-2}(x) \succ \prec P_{k-2}(x), P_{k-2}(x) \succ^{-1} \\
&= \prec x^{k-1} I_p, xP_{k-1}(x) \succ \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1} \\
&\quad - \prec x^{k-1} I_p, P_{k-2}(x) \succ \prec P_{k-2}(x), P_{k-2}(x) \succ^{-1} \\
&= \left(\prec x^{k-1} I_p, P_k(x) \succ + \prec x^{k-1} I_p, P_{k-1}(x) \succ B_k^T + \prec x^{k-1} I_p, P_{k-2}(x) \succ A_{k-1}^T \right) \\
&\quad \cdot \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1} - \prec x^{k-1} I_p, P_{k-2}(x) \succ \prec P_{k-2}(x), P_{k-2}(x) \succ^{-1} \\
&= \prec P_{k-1}(x), P_{k-1}(x) \succ B_k^T \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1}.
\end{aligned}$$

The last equation implies

$$\begin{aligned}
\mathcal{B}_k &= \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} B_k \prec P_{k-1}(x), P_{k-1}(x) \succ^{1/2} \\
&= \prec P_{k-1}(x), P_{k-1}(x) \succ^{1/2} B_k^T \prec P_{k-1}(x), P_{k-1}(x) \succ^{-1/2} \\
&= \mathcal{B}_k^T. \quad \square
\end{aligned}$$

References

[1] V. Averbukh, O. Smolyanov, The various definitions of the derivative in linear topological spaces, Russian Math. Surveys 23 (1968) 67–113.
[2] F.C. Chang, J.H.B. Kemperman, W.J. Studden, A normal limit theorem for moment sequences, Ann. Probab. 21 (1993) 1295–1309.
[3] P. Delsarte, Y.V. Genin, Y.G. Kamp, Orthogonal polynomial matrices on the unit circle, IEEE Trans. Circuits Syst. 25 (1978) 149–160.
[4] H. Dette, J. Nagel, Distributions on unbounded moment spaces and random moment sequences, Ann. Probab. 40 (2012) 2690–2704.
[5] H. Dette, J. Nagel, Matrix measures, random moments and Gaussian ensembles, J. Theoret. Probab. 25 (1) (2012) 25–49.
[6] H. Dette, B. Reuther, W.J. Studden, M. Zygmunt, Matrix measures and random walks with a block tridiagonal transition matrix, SIAM J. Matrix Anal. Appl. 29 (2006) 117–142.
[7] H. Dette, W.J. Studden, Matrix measures, moment spaces and Favard's theorem on the interval [0; 1] and [0; ∞), Linear Algebra Appl. 345 (2002) 169–193.
[8] A.J. Duran, On orthogonal polynomials with respect to a positive definite matrix of measures, Canad. J. Math. 47 (1995) 88–112.
[9] A.J. Duran, Ratio asymptotics for orthogonal matrix polynomials, J. Approx. Theory 100 (1999) 304–344.
[10] A.J. Duran, P. Lopez-Rodriguez, Orthogonal matrix polynomials: zeros and Blumenthal's theorem, J. Approx. Theory 84 (1996) 96–118.
[11] A.J. Duran, W. van Assche, Orthogonal matrix polynomials and higher-order recurrence relations, Linear Algebra Appl. 219 (1995) 261–280.
[12] F. Gamboa, L.V. Lozada-Chang, Large deviations for random power moment problem, Ann. Probab. 32 (2004) 2819–2837.
[13] F. Gamboa, J. Nagel, A. Rouault, J. Wagener, Large deviations for random matricial moment problems, J. Multivariate Anal. 106 (2012) 17–35.
[14] F. Gamboa, A. Rouault, Canonical moments and random spectral measures, J. Theoret. Probab. (2009) http://dx.doi.org/10.1007/s10959-009-0239-1.
[15] A.K. Gupta, D.K. Nagar, Matrix Variate Distributions, Capman & Hall/CRC, Boca Raton, 2000.
[16] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[17] S. Karlin, J. McGregor, Random walks, Illinois J. Math. 3 (1959) 66–81.
[18] L.V. Lozada-Chang, Asymptotic behavior of moment sequences, Electron. J. Probab. 10 (2005) 662–690.
[19] A. Mathai, Jacobians of Matrix Transformations and Functions of Matrix Argument, World Scientific Publ., Singapore, 1997.
[20] R.J. Muirhead, Aspects of Multivariate Statistical Theory, John Wiley & Sons, New York, 1982.
[21] A. Sinap, W. van Assche, Orthogonal matrix polynomials and applications, J. Comput. Appl. Math. 66 (1996) 27–52.
[22] M. Skibinsky, Extreme nth moments for distributions on [0, 1] and the inverse of a moment space map, J. Appl. Probab. 5 (1968) 693–701.
[23] A.M. Tulino, S. Verdú, Random Matrix Theory and Wireless Communications, in: Found. Trends Commun. Inf. Theory, Now Publishers Inc., Hanover, MA, 2004.
[24] H.S. Wall, Analytic Theory of Continued Fractions, D. van Nostrand Company, Inc., New York, 1948, XIII.
[25] P. Whittle, On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix, Biometrika 50 (1963) 129–134.
[26] N. Wiener, P. Masani, The prediction theory of multivariate stochastic processes, Acta Math. 98 (1957) 111–150.