Hankel determinants of random moment sequences

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

Dominik Tomecki
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany
e-mail: dominik.tomecki@rub.de

August 5, 2015

Abstract

For $t \in [0,1]$ let $\underline{H}_{2\lfloor nt \rfloor} = (m_{i+j})_{i,j=0}^{\lfloor nt \rfloor}$ denote the Hankel matrix of order $2\lfloor nt \rfloor$ of a random vector $(m_1, \ldots, m_{2n})$ on the moment space $\mathcal{M}_{2n}(T)$ of all moments (up to the order $2n$) of probability measures on the interval $T \subset \mathbb{R}$. In this paper we study the asymptotic properties of the stochastic process $\{\log \det \underline{H}_{2\lfloor nt \rfloor}\}_{t \in [0,1]}$ as $n \to \infty$. In particular weak convergence and corresponding large deviation principles are derived after appropriate standardization.

Keyword and Phrases: random Hankel determinant, random moment sequences, weak convergence, large deviation principle, canonical moments, arcsine distribution

AMS Subject Classification: 60F05, 60F10, 30E05, 15B52

1 Introduction

Hankel matrices are well studied objects in mathematics with applications in various fields such as orthogonal polynomials, random matrices or operator theory. Asymptotic properties of functions of non-random Hankel matrices such as the determinant, condition number or smallest eigenvalue have been studied by Hirschman Jr. (1966), Zamarashkin and Tyrtyshnikov (2001), Basor et al. (2001) or Berg and Szwarc (2011) among others. Recently, random Hankel matrices have also been considered in the literature with the main focus on matrices with independent entries. For example, Bryc et al. (2006) studied the limiting spectral measure of large Hankel (and Toeplitz) matrices, while some results regarding the operator norm can be found in Bose and Sen (2007). The present paper takes a different look at random Hankel matrices (more precisely, at their log-determinants) with not necessarily independent entries defined by a distribution on a moment space. More precisely, our investigations are motivated by the fact that Hankel matrices are usually used to characterize the solution of classical moment problems. To be precise, let $T \subset \mathbb{R}$ denote an interval and define $\mathcal{P}(T)$ as the set of all probability measures on the Borel field of T with existing moments. For a measure $\mu \in \mathcal{P}(T)$ we denote by

$$m_k = m_k(\mu) = \int_T x^k \mu(dx) \ ; \qquad k = 0, 1, 2, \ldots$$

the k-th moment and define

$$\mathcal{M}(T) = \big\{ \boldsymbol{m}(\mu) = (m_1(\mu), m_2(\mu), \ldots)^t | \ \mu \in \mathcal{P}(T) \big\} \subset \mathbb{R}^{\mathbb{N}} \tag{1.1}$$

as the set of all moment sequences. We denote by Π_n ($n \in \mathbb{N}$) the canonical projection onto the first n coordinates and call $\mathcal{M}_n(T) = \Pi_n(\mathcal{M}(T)) \subset \mathbb{R}^n$ the n-th moment space. The Hamburger moment problem is to decide if a given sequence $(m_n)_{n\in\mathbb{N}}$ is an element of $\mathcal{M}(\mathbb{R})$ and it is well known that this is the case if and only if the Hankel matrices $\underline{H}_{2k} = (m_{i+j})_{i,j=0}^k$ are nonnegative definite for all $k \in \mathbb{N}$ [see Shohat and Tamarkin (1943)]. Moreover the vector $\boldsymbol{m}_{2n} = (m_1, \ldots, m_{2n})$ is an element of the moment space $\mathcal{M}_{2n}(\mathbb{R})$ if and only if the Hankel matrix $\underline{H}_{2n} = (m_{i+j})_{i,j=0}^n$ is nonnegative definite. Similar characterization can be obtained for the Stieltjes and Hausdorff moment problem corresponding to measures on the half line $\mathbb{R}_0^+ = [0, \infty)$ and the interval $[0, 1]$, respectively.

Chang et al. (1993) considered the "classical" moment space corresponding to measures on the interval $[0, 1]$ [see Karlin and Shapeley (1953), Krein and Nudelman (1977), for some early references] and equipped $\mathcal{M}_n([0, 1])$ with a uniform distribution. They proved asymptotic normality of an appropriately standardized version of a projection $\Pi_k(\boldsymbol{m}_n)$ of a uniformly distributed vector $\boldsymbol{m}_n$ on $\mathcal{M}_n([0, 1])$ as $n \to \infty$. Gamboa and Lozada-Chang (2004) investigated corresponding large deviation principles, while Lozada-Chang (2005) studied similar problems for moment spaces corresponding to more general functions defined on a bounded set. More recently, some of these results have been generalized by Dette and Nagel (2012) to the moment spaces $\mathcal{M}_n([0, \infty))$ and $\mathcal{M}_n(\mathbb{R})$ corresponding to unbounded intervals.

The present paper is devoted to the asymptotic analysis of Hankel determinants of random moment vectors on $\mathcal{M}_{2n}(T)$. For example, if $\boldsymbol{m}_{2n} = (m_1, \ldots, m_{2n})$ denotes a random vector uniformly distributed on the $2n$th moment space $\mathcal{M}_{2n}([0, 1])$, then it is shown in this paper that an appropriately transformed and standardized version of the determinant of the random Hankel matrix $\underline{H}_{2n} = (m_{i+j})_{i,j=0}^n$ converges weakly, that is

$$\frac{2}{\sqrt{n}} \Big\{ \log\det \underline{H}_{2n} - \log\det \underline{H}_{2n}^0 + \frac{n}{2} \Big\} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \tag{1.2}$$

where $\underline{H}_{2n}^0 = (m_{i+j}^0)_{i,j=0}^n$ denotes the Hankel determinant of the moments of the arcsine distribution on the interval $[0, 1]$, that is $m_\ell = \binom{2\ell}{\ell} 2^{-2\ell}$. Moreover, the sequence

$$-\frac{1}{n} \Big\{ \log\det \underline{H}_{2n} - \log\det \underline{H}_{2n}^0 \Big\} \tag{1.3}$$

satisfies a large deviation principle with a good rate function. It will be demonstrated in Section 2 that the moments of the arcsine distribution used for the centering in (1.2) and (1.3) correspond to the center of the moment space $\mathcal{M}([0,1])$.

Similar results are available for the moment space $\mathcal{M}_{2n}([0,\infty))$ and $\mathcal{M}_{2n}(\mathbb{R})$, where the centering has to be performed by the logarithms of the determinants of the Hankel matrices corresponding to the Marcenko-Pastur law and Wigner's semi-circle law, respectively (in these cases the corresponding Hankel determinants $\underline{H}_{2n}$ have value 1). These measures play a very important role in the theory of random matrices, free probability and quantum probability, see the books of Hiai and Petz (2000) and Hora and Obata (2007) among others.

The remaining part of this paper is organized as follows. In Section 2 we present some facts on moment theory and introduce random moment sequences on the spaces $\mathcal{M}_n([0,1]), \mathcal{M}_n([0,\infty))$ and $\mathcal{M}_{2n-1}(\mathbb{R})$. We also state some basic properties of these random variables which will be useful in the following discussion. In Section 3 it is shown that for the canonical distributions on the moment space $\mathcal{M}_{2n}(T)$ an appropriately standardized version of the stochastic process

$$\{D_{2\lfloor nt \rfloor}\}_{t\in[0,1]} = \{\log \det \underline{H}_{2\lfloor nt \rfloor}\}_{t\in[0,1]} \tag{1.4}$$

converges weakly to a Gaussian process. The centering and scaling is different for the three moment spaces under consideration. We also study the asymptotic properties of the vector $(D_{n,2},\dots,D_{n,2k})^t$ for any fixed k. Large deviation principles are investigated in Section 4, while some technical results which are required for the proofs are provided in the Appendix.

2 Some basic facts about moment theory

Similar to cumulants, canonical moments provide a one-to-one transformation of the ordinary moments. They appear naturally in the continued fraction expansion of the Stieltjes transform of a probability measure but are less known than cumulants. Therefore, we state some basic facts in the following two paragraphs, where we distinguish between bounded and unbounded intervals.

2.1 Canonical moments

Canonical moments have been investigated in a series of papers by Skibinsky (1967, 1968, 1969) and roughly speaking define a one-to-one mapping from the set of moments $\mathcal{M}([0,1])$ (or more generally from $\mathcal{M}([a,b])$ for any finite interval $[a,b]\subset\mathbb{R}$) onto the set $[0,1]^{\mathbb{N}}$. They have implicitly been discussed before in the work of Verblunsky (1935, 1936), who mainly considered measures on the unit circle. In this section we briefly present some basic facts for the sake of a self contained presentation and discuss corresponding results for the set $\mathcal{M}([0,\infty))$ and $\mathcal{M}(\mathbb{R})$. For details we refer to the monographs of Dette and Studden (1997) and Wall (1948). For a given vector

$\boldsymbol{m}_{k-1} = (m_1, \ldots, m_{k-1})^T \in \mathcal{M}_{k-1}([0,1])$ of moments of a probability measure on the interval $[0,1]$ define

$$m_k^- = \min\left\{ m_k(\mu) \,\middle|\, \mu \in \mathcal{P}([0,1]) \text{ with } \int_0^1 t^i d\mu(t) = m_i \text{ for } i = 1, \ldots, k-1 \right\},$$
$$m_k^+ = \max\left\{ m_k(\mu) \,\middle|\, \mu \in \mathcal{P}([0,1]) \text{ with } \int_0^1 t^i d\mu(t) = m_i \text{ for } i = 1, \ldots, k-1 \right\}.$$

Throughout this paper let $\operatorname{Int} C$ denote the interior of a set C. It is shown in Dette and Studden (1997) that $\boldsymbol{m}_k = (m_1, \ldots, m_k)^T \in \operatorname{Int} \mathcal{M}_k([0,1])$ if and only if $m_k^- < m_k < m_k^+$. In this case the canonical moments of order $l = 1, \ldots, k$ are defined as

$$p_l = p_l(\boldsymbol{m}_k) = \frac{m_l - m_l^-}{m_l^+ - m_l^-}\ ; \qquad l = 1, \ldots, k. \tag{2.1}$$

Note that for $\boldsymbol{m}_k \in \operatorname{Int} \mathcal{M}_k([0,1])$ we have $p_l \in (0,1)$; $l = 1, \ldots, k$; and that p_k describes the relative position of the moment m_k in the set of all possible k-th moments with fixed moments $m_1, \ldots, m_{k-1}$. It can also be shown that the definition (2.1) defines a one-to one mapping from $\operatorname{Int} \mathcal{M}_n([0,1])$ onto the open cube $(0,1)^n$. As an example consider the arcsine distribution μ^0 on the interval $[0,1]$ with density $1/(\pi\sqrt{x(1-x)})$, then the corresponding canonical moments are given by $p_\ell = 1/2$ for all $\ell \in \mathbb{N}$ [see Dette and Studden (1997)]. Consequently, the sequence of moments of the arcsine distribution defines the center of the moment space $\mathcal{M}([0,1])$. Note however, that it is not the barycenter of the moment space.
The determinant of the Hankel matrix $\underline{H}_{2k} = (m_{i+j})_{i,j=0}^k$ of the moment vector $(m_1, \ldots, m_{2k})$ can easily be expressed in terms of the corresponding canonical moments, that is

$$\det \underline{H}_{2k} = \det(m_{i+j})_{i,j=0}^k = \big(p_1 q_1 p_2\big)^k \prod_{j=2}^k \big(q_{2k-2} p_{2k-1} q_{2k-1} p_{2k}\big)^{k-j+1}\ , \tag{2.2}$$

where $q_j = 1 - p_j$ [see Dette and Studden (1997), Theorem 1.4.10].
In the case $T = [0, \infty)$ the upper bound m_k^+ is in general not finite, but we can still define for a point $\boldsymbol{m}_{k-1} \in \operatorname{Int} \mathcal{M}_{k-1}([0,\infty))$ the lower bound

$$m_k^- = \min\left\{ m_k(\mu) \,\middle|\, \mu \in \mathcal{P}([0,\infty)) \text{ with } \int_0^\infty t^i d\mu(t) = m_i \text{ for } i = 1, \ldots, k-1 \right\},$$

where $\boldsymbol{m}_k = (m_1, \ldots, m_k)^T \in \operatorname{Int} \mathcal{M}_k([0,\infty))$ if and only if $m_k > m_k^-$. In this case, the analogues of the canonical moments are defined by the quantities

$$z_l \;=\; \frac{m_l - m_l^-}{m_{l-1} - m_{l-1}^-} \qquad l = 1, \ldots, k \tag{2.3}$$

(with $m_0^- = 0$). As in the case of a bounded interval the definition (2.3) provides a one to one mapping from $\operatorname{Int} \mathcal{M}_n([0,\infty))$ onto $(\mathbb{R}^+)^n$, and it can be shown using similar arguments as in Dette and Studden (1997) that

$$\det \underline{H}_k = (m_k - m_k^-) \det \underline{H}_{k-2}\ ,\ \ k \geq 2.$$

Consequently, the determinant of the Hankel matrix is given by

$$\det \underline{H}_{2k} = \det(m_{i+j})_{i,j=0}^{k} = \prod_{j=1}^{k} \left(z_{2k-1} z_{2k}\right)^{k-j+1} , \tag{2.4}$$

Finally, in the case $T = \mathbb{R}$ neither m_k^- nor m_k^+ are in general finite. Nevertheless, there exists an analogue of the quantities p_i and z_i defined in (2.2) and (2.4). To be precise, we define for a vector $\boldsymbol{m}_{2n-1} = (m_1, \ldots, m_{2n-1}) \in \mathcal{M}_{2n-1}(\mathbb{R})$ with $\underline{H}_{2n-2} > 0$ the polynomial

$$P_k(x) \;=\; \begin{vmatrix} m_0 & \cdots & m_{k-1} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ m_k & \cdots & m_{2k-1} & x^k \end{vmatrix} \Big/ \begin{vmatrix} m_0 & \cdots & m_{k-1} \\ \vdots & \ddots & \vdots \\ m_{k-1} & \cdots & m_{2k-2} \end{vmatrix} ; \quad k = 1, \ldots, n \tag{2.5}$$

[see Chihara (1978)]. We consider a one to one mapping

$$\xi_n : \begin{cases} \operatorname{Int} \mathcal{M}_{2n-1}(\mathbb{R}) \longrightarrow \left(\mathbb{R} \times \mathbb{R}^+\right)^{n-1} \times \mathbb{R} \\ \boldsymbol{m}_{2n-1} \mapsto (b_1, a_1, \ldots, a_{n-1}, b_n)^T \end{cases} \tag{2.6}$$

defined by

$$\int_{\mathbb{R}} x^k P_k(x) d\mu(x) \;=\; a_1 \ldots a_k; \qquad k = 1, \ldots, n-1, \tag{2.7}$$

$$\int_{\mathbb{R}} x^{k+1} P_k(x) d\mu(x) \;=\; a_1 \ldots a_k (b_1 + \cdots + b_{k+1}); \qquad k = 0, \ldots, n-1, \tag{2.8}$$

where μ is any measure with first $2n-1$ moments given by $(m_1, \ldots, m_{2n-1})$ [see for example Wall (1948)]. Note that $P_1(x), \ldots, P_n(x)$ are orthogonal polynomials with leading coefficient 1 with respect to the measure μ. It is now easy to see that the determinant of the Hankel matrix can be represented as

$$\det \underline{H}_{2k} = \det(m_{i+j})_{i,j=0}^{k} = \prod_{j=1}^{k} a_j^{k-j+1} . \tag{2.9}$$

In the following section we will equip these moment spaces with distributions. We begin with the moment space corresponding to measures on bounded intervals.

2.2 Distributions on moment spaces

Chang et al. (1993) considered a uniformly distributed vector on the set $\mathcal{M}_n([0,1])$ and showed that an appropriately standardized version of a projection $\Pi_k(\boldsymbol{m}_n)$ onto its first k components is asymptotically normal distributed, where the centering has to be performed with the moments of the arcsine distribution. A key ingredient in their proof is the following lemma, which shows that the canonical moments of a uniformly distributed vector $\boldsymbol{m}_n$ on $\mathcal{M}_n([0,1])$ are independent [for a proof of the following result see Dette and Studden (1997)]. For this and the following statements we will make the dependence of the canonical moments on the dimension of the moment space $\mathcal{M}_n([0,1])$ more explicit. More precisely, we use the notation $p_{n,\ell}(\boldsymbol{m}_n)$ instead of $p_\ell(\boldsymbol{m}_n)$, and the symbol $\beta(a,b)$ denotes a Beta-distribution on the interval $[0,1]$ with density

$$I_{[0,1]}(x)x^{a-1}(1-x)^{b-1}/B(a,b).$$

Lemma 2.1. *For a uniformly distributed random vector $\boldsymbol{m}_n$ on the nth moment space $\mathcal{M}_n([0,1])$ the canonical moments $p_{n,1}(\boldsymbol{m}_n),\dots,p_{n,n}(\boldsymbol{m}_n)$ defined by (2.1) are independent and Beta-distributed, that is*

$$p_{n,i}(\boldsymbol{m}_n) \sim \beta(n-i+1, n-i+1) \ .$$

Note that the mapping between the (regular) moments and the canonical moments has only been defined on the the interior of $\mathcal{M}_n([0,1])$. However, $\mathcal{M}_n([0,1])$ is a closed, convex set and therefore its boundary has Lebesgue measure 0. Since we endow this space with the uniform distribution, the random variables $p_{n,i}$ are a.s. well-defined. We also note that Dette and Nagel (2012) defined more general distributions on $\mathcal{M}_n([0,1])$, which contain the uniform distribution as a special case. In order to define an analogue of the uniform distribution on the unbounded moment space $\mathcal{M}_n([0,\infty))$ these authors use the relation (2.3). To be precise, consider a random vector $\boldsymbol{m}_n$ and denote the quantities in (2.3) by $z_{n,1}(\boldsymbol{m}_n),\dots,z_{n,n}(\boldsymbol{m}_n)$. A density on the moment space $\mathcal{M}_n([0,\infty))$ is then defined by

$$g_n^{(\gamma,\delta)}(\boldsymbol{m}_n) \;=\; c_n^{[0,\infty)} \prod_{k=1}^{n} z_{n,k}(\boldsymbol{m}_n)^{\gamma_{n,k}} \exp(-\delta_{n,k} z_{n,k}(\boldsymbol{m}_n)) \mathbb{1}_{\{z_{n,k}(\boldsymbol{m}_n)>0\}}, \tag{2.10}$$

where the constants satisfy $\gamma_{n,k} > -(n-k+1)$, $\delta_{n,k} > 0$ for $k=1,\dots,n$, and the normalizing constant is given by $c_n^{[0,\infty)} = \prod_{k=1}^n (\delta_{n,k}^{\gamma_k+n-k+1})/\Gamma(\gamma_{n,k}+n-k+1))$. The analogue of Lemma 2.1 is now provided by the following result, where the symbol $\gamma(a,b)$ denotes a Gamma distribution $(a,b>0)$ with density

$$\frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} \mathbb{1}_{[0,\infty)}(x) \ .$$

Lemma 2.2. *For a random vector $\boldsymbol{m}_n$ with density (2.10) on $\mathcal{M}_n([0,\infty))$ the canonical moments $z_{n,k}(\boldsymbol{m}_n)$ defined by (2.3) are independent and Gamma-distributed, that is*

$$z_{n,k}(\boldsymbol{m}_n) \sim \gamma(\gamma_{n,k}+n-k+1, \delta_{n,k}), \qquad k=1,\dots,n.$$

A proof of Lemma 2.2 can be found in Dette and Nagel (2012) and we conclude this section with the corresponding statements for the moment space $\mathcal{M}_{2n-1}(\mathbb{R})$. Following Dette and Nagel (2012) we define a distribution on $\mathcal{M}_{2n-1}(\mathbb{R})$ by

$$
\begin{aligned}
h_{2n-1}^{(\gamma,\delta)}(\boldsymbol{m}_{2n-1}) = & \prod_{k=1}^{n} \sqrt{\frac{\delta_{n,2k-1}}{\pi}} \exp\left(-\delta_{n,2k-1} b_{n,k}^2(\boldsymbol{m}_{2n-1})\right) \\
& \times \prod_{k=1}^{n-1} \frac{{\delta_{n,2k}}^{\gamma_{n,k}+2n-2k}}{\Gamma(\gamma_{n,k}+2n-2k)} a_{n,k}^{\gamma_{n,k}}(\boldsymbol{m}_{2n-1}) \exp\left(-\delta_{n,2k} a_{n,k}(\boldsymbol{m}_{2n-1})\right) \mathbb{1}_{\{a_{n,k}(\boldsymbol{m}_{2n-1})>0\}},
\end{aligned} \tag{2.11}
$$

where the constants satisfy $\gamma_{n,k} > -2(n-k)$ for $k = 1, \ldots, n-1$ and $\delta_{n,1}, \ldots, \delta_{1,2n-1} > 0$. The distribution of the corresponding quantities a_k and b_k defined by (2.7) and (2.8) is specified in the following result.

Lemma 2.3. *Let $\boldsymbol{m}_{2n-1} \in \mathcal{M}_{2n-1}(\mathbb{R})$ be a random vector with density $h_{2n-1}^{(\gamma,\delta)}$ defined in (2.11). Then the random coefficients $(b_{n,1}, a_{n,1} \ldots, a_{n,n-1}, b_{n,n})^T$ defined by (2.7) and (2.8) are independent and*

$$
b_{n,k} \sim \mathcal{N}(0, \tfrac{1}{2\delta_{n,2k-1}}), \;\; a_{n,k} \sim \gamma(\gamma_{n,k} + 2n - 2k, \delta_{n,2k}).
$$

Remark 2.4. There exists an interesting relation to random matrix theory in particular to the β-ensembles considered by Dumitriu and Edelman (2002); Edelman and Sutton (2008); Ramírez et al. (2011) among others. To be precise, consider exemplarily the moment space $\mathcal{M}_{2n-1}(\mathbb{R})$. It can be shown that for a point $\boldsymbol{m}_{2n-1} \in \operatorname{Int} \mathcal{M}_{2n-1}(\mathbb{R})$ the polynomials defined in (2.5) satisfy the three term recurrence relation

$$
xP_k(x) = P_{k+1}(x) + b_{k+1}P_k(x) + a_k P_{k-1}(x); \qquad k = 1, \ldots, n-1, \tag{2.12}
$$

$(P_0(x) = 1, P_1(x) = x - b_1)$, where the coefficients in the recursion are defined by (2.7) and (2.8). A straightforward calculation now shows that the polynomial $P_n(x)$ is the characteristic polynomial $\det(xI_n - A_n)$ of the matrix

$$
A_n \;=\; \begin{pmatrix} b_1 & \sqrt{a_1} & & & \\ \sqrt{a_1} & b_2 & \sqrt{a_2} & & \\ & \ddots & \ddots & \ddots & \\ & & \sqrt{a_{n-2}} & b_{n-1} & \sqrt{a_{n-1}} \\ & & & \sqrt{a_{n-1}} & b_n \end{pmatrix}. \tag{2.13}
$$

If $\boldsymbol{m}_{2n-1}$ is a random vector on $\mathcal{M}_{2n-1}(\mathbb{R})$ with density $h_{2n-1}^{(\gamma,\delta)}$ defined in (2.11), and $\delta_{n,2k-1} = 1/2$ $(k = 1, \ldots, n)$, $\delta_{n,2k} = 1$, $\gamma_{n,k} = (\frac{1}{2}\beta - 2)(n-k)$ $(k = 1, \ldots, n-1)$ for some $\beta > 0$, then it follows from Lemma 2.3 that the coefficients in this matrix are independent with distributions $b_i \sim \mathcal{N}(0, 1)$, $a_i \sim \frac{1}{2}\chi^2_{\beta(n-i)}$. This means that $P_n(x)$ is the characteristic polynomial of the random the matrix (2.13) corresponding to the β- Hermite ensemble as introduced by Dumitriu and Edelman

(2002). While the common matrix literature investigates spectral properties of this matrix, the random Hankel determinant corresponds to a product of L^2-norms of the (random) polynomials $P_1, \ldots, P_n$, that is

$$\prod_{i=1}^{n} \int_{\mathbb{R}} P_i^2(x)\mu(dx) \;=\; \prod_{i=1}^{n} a_i^{n-i+1},$$

where μ denotes a random measure whose first $2n-1$ moments are defined the random Jacobi matrix (2.13).

We also note that a similar interpretation is available for the random moment sequences on $\mathcal{M}_{2n-1}([0,1])$ and $\mathcal{M}_{2n-1}([0,\infty))$ observing the results of Killip and Nenciu (2004) and Dumitriu and Edelman (2002), respectively.

3 Weak convergence of Hankel determinant processes

Throughout this section we investigate the asymptotic properties of the stochastic process

$$\{D_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n})\}_{t\in[0,1]} = \{\log\det \underline{H}_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n})\}_{t\in[0,1]} \tag{3.1}$$

where $\underline{H}_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n}) = (m_{i+j}(\boldsymbol{m}_{2n}))_{i,j=0}^{\lfloor nt\rfloor}$ is the Hankel determinant of a random vector $\boldsymbol{m}_{2n}$ on the moment space $\mathcal{M}_{2n}(T)$. We also investigate the asymptotic properties of the vector $(D_{n,2}(\boldsymbol{m}_{2n}), \ldots, D_{n,2k}(\boldsymbol{m}_{2n}))$ for some fixed $k \in \mathbb{N}$. In the following discussion we treat the cases of a bounded and unbounded moment space separately.

3.1 Hankel determinants from $\mathcal{M}_{2n}([0,1])$

Throughout this paper the symbol $Y_n \xrightarrow{d} Y$ denotes weak convergence of a vector valued sequence of random variables $(Y_n)_{n\in\mathbb{N}}$. Moreover, let $l^\infty([0,1])$ denote the space of bounded real-valued functions on the interval $[0,1]$ with the topology induced by the uniform norm. We denote by $X_n \Longrightarrow X$ the weak convergence of a sequence $(X_n)_{n\in\mathbb{N}}$ of random variables in $l^\infty([0,1])$ to an $l^\infty([0,1])$-valued random variable X. We also use the convention $0\log(0) := 0$ and in the following $s\wedge t$ and $s\vee t$ denote the minimum and maximum of $s,t \in [0,1]$, respectively. Consider a uniformly distributed random vector $\boldsymbol{m}_{2n}$ on $\mathcal{M}_{2n}([0,1])$, that is $\boldsymbol{m}_{2n} \sim \mathcal{U}(\mathcal{M}_{2n})$. We first investigate the weak convergence of the vector

$$H_n^k = (D_{n,2}(\boldsymbol{m}_{2n}), \ldots, D_{n,2k}(\boldsymbol{m}_{2n}))^t \tag{3.2}$$

for a fixed $k \in \mathbb{N}$.

Theorem 3.1. *If $\boldsymbol{m}_{2n} \sim \mathcal{U}(\mathcal{M}_{2n})$, then the random vector*

$$\sqrt{4n}(H_n^k - D_{n,2k}^0) \xrightarrow{d} \mathcal{N}(0,\Sigma_k)\ ,$$

where $D^0_{n,2k}$ denotes the log-determinant of the Hankel matrix corresponding to the arcsine distribution, that is

$$D^0_{n,2k} = \log\det\Big(\binom{2(i+j)}{i+j}2^{-2(i+j)}\Big)_{i,j=0,\dots,k} \quad = \quad -k(2k+1)\log(2), \tag{3.3}$$

and the asymptotic covariance matrix is given by

$$\Sigma_k = (i\wedge j)^k_{i,j=1} \ . \tag{3.4}$$

Proof: In all proofs of this paper we do not reflect the dependence of the canonical moments on the vector of random moments and use the notation $p_{2n,i} = p_{2n,i}(\boldsymbol{m}_{2n})$. According to Lemma 2.1 the canonical moments $p_{2n,i}$ are independent and $\beta(2n-i+1, 2n-i+1)$ distributed $(i = 1, 2, \dots, 2n)$. Therefore it follows from Theorem 3.1 in Dette and Nagel (2012) and the Delta method that

$$\sqrt{4n}(\log(p_{2n,k}) - \log(\tfrac{1}{2})) \xrightarrow{d} \mathcal{N}(0,1) \tag{3.5}$$

$$\sqrt{4n}(\log(p_{2n,k}(1-p_{2n,k})) - \log(\tfrac{1}{4})) \xrightarrow{P} 0 \tag{3.6}$$

Because the canonical moments $p_{n,1}, \dots, p_{n,n}$ are independent we obtain that the random vector

$$B_n = \sqrt{4n}\big((\log(q_{2n,2}), \dots, \log(q_{2n,2k}))^t - (\log(\tfrac{1}{2}), \dots, \log(\tfrac{1}{2}))^t\big)$$

converges weakly to a standard k-dimensional Gaussian distribution.
Next, note that the representation (3.3) follows from (2.2) and the fact that the canonical moments of the arsine distribution are all given by $1/2$. Consequently, we can decompose the vector H^k_n as follows

$$\sqrt{4n}(H^k_n - D^0_{n,2k}) = S_n - T_n \ ,$$

where the components of the vectors S_n and T_n are given by

$$S_{ni} = \sum_{j=1}^{i} \sqrt{4n}(i-j+1)(\log(q_{2n,2j}p_{2n,2j}) + \log(q_{2n,2j-1}p_{2n,2j-1}) - 2\log(\tfrac{1}{4})) \ ,$$

$$T_{ni} = \sqrt{4n}\sum_{j=1}^{i}(\log(q_{2n,2j}) - \log(\tfrac{1}{2})) \ ,$$

respectively $(i = 1, \dots, k)$. Observing (3.6) we see that S_n converges in probability to 0. For a proof of weak convergence of the random variable T_n note that $T_n = A_k B_n$, where the matrix $A_k \in \mathbb{R}^{k\times k}$ is given by

$$A_k = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}$$

This yields the weak convergence $T_n \xrightarrow{d} \mathcal{N}(0, \Sigma_k)$ and the representation (3.4) for the matrix $\Sigma_k = A_k A_k^t$ follows by a straightforward calculation.

$\square$

While Theorem 3.1 holds for any fixed $k \in \mathbb{N}$, the following result provides a process version.

Theorem 3.2. *Let $\boldsymbol{m}_{2n}$ denote a uniformly distributed random vector on $\mathcal{M}_{2n}([0,1])$, then*

$$\big\{\mathcal{G}_n(t)\big\}_{t\in[0,1]} \;:=\; \frac{2}{\sqrt{n}}\Big\{D_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n}) - D^0_{n,2\lfloor nt\rfloor} + \frac{n}{2}r(t)\Big\}_{t\in[0,1]} \;\Longrightarrow\; \{\mathcal{G}^{[0,1]}(t)\}_{t\in[0,1]}\ ,$$

where $D_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n})$ is defined in (3.1),

$$r(t) = t + (1-t)\log(1-t) \tag{3.7}$$

and $\mathcal{G}^{[0,1]}$ is a centered continuous Gaussian process on the interval $[0,1]$ with covariance kernel

$$f(s,t) \;=\; \int_0^{s\wedge t} \tfrac{(t-x)(s-x)}{(1-x)^2}\,dx = (s\wedge t)(2 - s\vee t) - (s+t-2)\log(1-s\wedge t) \tag{3.8}$$

Proof: It is shown later (more precisely, in the proof of (3.10)) that the kernel f is in fact nonnegative definite, that is for all $k \in \mathbb{N}$, $s_1, \ldots, s_k, t_1, \ldots, t_k \in [0,1]$ the matrices $(f(s_i,t_j))_{i,j=1}^k$ are nonnegative definite. A simple calculation shows that $\mathbb{E}[(\mathcal{G}^{[0,1]}(t) - \mathcal{G}^{[0,1]}(s))^4] \le 48(t-s)^2$, and consequently the existence of the process $\mathcal{G}^{[0,1]} = \{\mathcal{G}^{[0,1]}(t)\}_{t\in[0,1]}$ follows from Theorem 3.23 in Kallenberg (2002). Moreover, since $\mathcal{G}^{[0,1]}$ is continuous and $l^\infty([0,1])$ is a complete space, Theorem 1.3.2 in van der Vaart and Wellner (1995) shows that $\mathcal{G}^{[0,1]}$ is tight. For the following discussion we define

$$\begin{aligned}\tilde{\xi}_{n,i} &= \log(q_{2n,2i}p_{2n,2i}) + \log(q_{2n,2i-1}p_{2n,2i-1}),\\ \xi_{n,i}(t) &= \frac{2}{\sqrt{n}}(\lfloor nt\rfloor - i + 1)\big(\tilde{\xi}_{n,i} - \mathbb{E}[\tilde{\xi}_{n,i}]\big),\end{aligned}$$

and obtain by (2.2) the decomposition

$$\mathcal{G}_n(t) \;= S_n(t) + 2R_n(t) - 2T_n(t) + 2U_n(t), \tag{3.9}$$

where the processes S_n, R_n, T_n and U_n are defined by

$$\begin{aligned}S_n(t) &= \sum_{i=1}^{\lfloor nt\rfloor-1} \xi_{n,i}(t),\\ T_n(t) &= \frac{1}{\sqrt{n}}\Big\{\lfloor nt\rfloor\log(2) - \log(p_{2n,2\lfloor nt\rfloor}) + \sum_{i=1}^{\lfloor nt\rfloor-1}\log(q_{2n,2i})\Big\},\end{aligned}$$

$$R_n(t) = \frac{I\{nt \geq 1\}}{\sqrt{n}} \log(q_{2n,2\lfloor nt\rfloor-1} p_{2n,2\lfloor nt\rfloor-1}),$$

$$U_n(t) = \frac{1}{\sqrt{n}} \Big\{ \sum_{i=1}^{\lfloor nt\rfloor-1} (\lfloor nt\rfloor - i + 1)\mathbb{E}[\tilde{\xi}_{n,i}] - D^0_{n,2\lfloor nt\rfloor} + \frac{n}{2} r(t) + \lfloor nt\rfloor \log(2) \Big\},$$

respectively. With these notations the proof of Theorem 3.2 follows from the assertions

$$S_n \Longrightarrow \mathcal{G}^{[0,1]}\ , \tag{3.10}$$

$$T_n \Longrightarrow 0\ , \tag{3.11}$$

$$R_n \Longrightarrow 0\ , \tag{3.12}$$

$$||U_n||_\infty \xrightarrow{n\to\infty} 0\ , \tag{3.13}$$

and a simple application of Slutsky's theorem.

Proof of (3.10). For each $k \in \mathbb{N}$ consider $0 = t_0 \leq t_1 \leq \ldots \leq t_k \leq 1$ and define the k-dimensional random variable $S^*_n := (S_n(t_1), \ldots, S_n(t_k))^t$. Let $c = (c_1, \ldots, c_k)^t \in \mathbb{R}^k$ be an arbitrary vector. then

$$\begin{aligned} c^t S^*_n &= \sum_{i=1}^k c_i S_n(t_i) = \sum_{i=1}^k c_i \sum_{j=1}^{\lfloor nt_i\rfloor-1} \xi_{n,j}(t_i) = \sum_{i=1}^k c_i \sum_{l=1}^{i} \sum_{j=\lfloor nt_{l-1}\rfloor\vee 1}^{\lfloor nt_l\rfloor-1} \xi_{n,j}(t_i) \\ &= \frac{2}{\sqrt{n}} \sum_{l=1}^k \sum_{j=\lfloor nt_{l-1}\rfloor\vee 1}^{\lfloor nt_l\rfloor-1} \big(\tilde{\xi}_{n,j} - \mathbb{E}[\tilde{\xi}_{n,j}]\big) \sum_{i=l}^k c_i(\lfloor nt_i\rfloor - j + 1) \end{aligned}$$

In order to calculate the variance of $c^t S^*_n$ we assume $0 \leq s \leq t \leq 1$, use the approximation (A.2) in the Appendix and obtain

$$\begin{aligned} \mathrm{cov}(S_n(s), S_n(t)) =& \frac{4}{n} \sum_{i=1}^{\lfloor ns\rfloor-1} (\lfloor nt\rfloor - i + 1)(\lfloor ns\rfloor - i + 1)\,\mathrm{Var}(\tilde{\xi}_{n,i}) \\ =& \frac{4}{n} \sum_{i=1}^{\lfloor ns\rfloor-1} \frac{(\lfloor nt\rfloor - i + 1)(\lfloor ns\rfloor - i + 1)}{4(n-i+1)^2} \\ &+ \frac{4}{n} \sum_{i=1}^{\lfloor ns\rfloor-1} (\lfloor nt\rfloor - i + 1)(\lfloor ns\rfloor - i + 1) O\left((n-i+1)^{-3}\right) \\ =& \frac{1}{n} \sum_{i=1}^{\lfloor ns\rfloor-1} \frac{\left(t - \frac{i-1+nt-\lfloor nt\rfloor}{n}\right)\left(s - \frac{i-1+ns-\lfloor ns\rfloor}{n}\right)}{(1-\frac{i-1}{n})^2} + \frac{4}{n} \sum_{i=1}^{\lfloor ns\rfloor-1} O((n-i+1)^{-1}). \end{aligned}$$

Interpreting the first term as Riemann-sum, we can calculate the limit

$$\lim_{n\to\infty} \mathrm{cov}(S_n(s), S_n(t)) = \int_0^s \frac{(t-x)(s-x)}{(1-x)^2}\, dx = f(s,t),$$

which gives

$$\lim_{n\to\infty} \mathrm{Var}(c^t S_n^*) = \lim_{n\to\infty} c^t \operatorname{cov}(S_n^*, S_n^*)c \to c^t \Sigma c \ ,$$

where the matrix Σ is given by $\Sigma = (f(t_i, t_j))_{i,j=1,\ldots,k}$ and the covariance kernel f is defined in (3.8). Consequently we obtain that this kernel is nonnegative definite.
We now prove the weak convergence of $c^t S_n^*$ by verifying the Lyapunov–condition. For this purpose we use the notation $c^* := \max\{|c_1|, \ldots, |c_n|\}$ and obtain

$$\begin{aligned}
&\frac{2^4}{n^2} \sum_{l=1}^{k} \sum_{j=\lfloor nt_{l-1}\rfloor \vee 1}^{\lfloor nt_l\rfloor -1} \mathbb{E}\big[(\tilde{\xi}_{n,j} - \mathbb{E}[\tilde{\xi}_{n,j}])^4\big] \Big(\sum_{i=l}^{k} c_i (\lfloor nt_i \rfloor - j + 1) \Big)^4 \\
&\le \frac{2^4}{n^2} \sum_{l=1}^{k} \sum_{j=\lfloor nt_{l-1}\rfloor \vee 1}^{\lfloor nt_l\rfloor -1} \mathbb{E}\big[(\tilde{\xi}_{n,j} - \mathbb{E}[\tilde{\xi}_{n,j}]^4\big] (kc^*)^4 (n-j+1)^4 \ \le \ \frac{(2kc^*)^4 C^2}{n} \to 0 \ ,
\end{aligned}$$

where we have used the estimate (A.5) in Appendix A for the moments. Consequently, Lyapunov‘s Theorem implies convergence of the finite dimensional distributions, that is

$$S_n^* = (S_n(t_1), \ldots, S_n(t_k))^t \xrightarrow{d} \mathcal{N}(0, \Sigma).$$

We finally prove that S_n is asymptotically tight, that is

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\omega_n\left(\tfrac{1}{m}\right) > \epsilon\right) \ = \ 0 \ , \tag{3.14}$$

where $\omega_n(a) = \sup\{|S_n(t) - S_n(s)| \mid 0 \le t - s \le a\}$ denotes the modulus of continuity of the process S_n. The statement (3.10) then follows from Theorem 1.5.4 in van der Vaart and Wellner (1995). For a proof of (3.14) we introduce the notation

$$d_{n,i} = \frac{2}{\sqrt{n}} \begin{cases} \lfloor nt \rfloor - \lfloor ns \rfloor & i \le \lfloor ns \rfloor - 1 \\ \lfloor nt \rfloor - i + 1 & \lfloor ns \rfloor - 1 < i \le \lfloor nt \rfloor - 1 \\ 0 & \text{else} \end{cases} \ ,$$

and obtain the following representation

$$S_n(t) - S_n(s) = \sum_{i=1}^{n} d_{n,i} \big(\tilde{\xi}_{n,i} - \mathbb{E}[\tilde{\xi}_{n,i}] \big).$$

The inequalities (A.4) and (A.5) in the Appendix then yield

$$\begin{aligned}
\mathbb{E}\left[(S_n(t) - S_n(s))^4\right] &\le C^2 \Big(\sum_{i=1}^{n} d_{n,i}^2 \frac{1}{(n-i+1)^2} \Big)^2 \\
&\le (2C)^2 \Big(\frac{1}{n} \sum_{i=1}^{\lfloor ns \rfloor -1} \frac{(t - s + \frac{1}{n})^2}{(1 - \frac{i}{n} + \frac{1}{n})^2} + \frac{1}{n} \sum_{i=\lfloor ns \rfloor \vee 1}^{\lfloor nt \rfloor -1} \frac{(t - \frac{i-1}{n})^2}{(1 - \frac{i-1}{n})^2} \Big)^2
\end{aligned}$$

$$
\begin{aligned}
&\le (2C)^2\Big(I\left\{s \ge \tfrac{2}{n}\right\}\left(t-s+\tfrac{1}{n}\right)^2 \int_{\frac{1}{n}}^{s} \frac{1}{\left(1-x+\frac{1}{n}\right)^2}\,dx + \frac{(\lfloor nt\rfloor - \lfloor ns\rfloor)}{n}\Big)^2 \\
&\le (2C)^2\Big(I\left\{s \ge \tfrac{2}{n}\right\}\left(t-s+\tfrac{1}{n}\right)^2\Big(\tfrac{1}{1-s+1/n}-1\Big) + \left(t-s+\tfrac{1}{n}\right)\Big)^2 \\
&\le (4C)^2\left(t-s+\tfrac{1}{n}\right)^2. \qquad (3.15)
\end{aligned}
$$

Consequently, we obtain

$$
\limsup_{n\to\infty}\sum_{k=1}^{m} E\left[(S_n(\tfrac{k}{m}) - S_n(\tfrac{k-1}{m}))^4\right] \le \limsup_{n\to\infty}\sum_{k=1}^{m}(4C)^2\left(\tfrac{1}{m}+\tfrac{1}{n}\right)^2 = \frac{(4C)^2}{m}. \qquad (3.16)
$$

Now assume that $0 \le r \le s \le t \le 1$. If $t - r \ge \frac{1}{n}$, Hölders's inequality and (3.15) yield the estimate

$$
\begin{aligned}
\mathbb{E}\left[(S_n(s)-S_n(r))^2(S_n(t)-S_n(s))^2\right] &\le (4C)^2\left(s-r+\tfrac{1}{n}\right)\left(t-s+\tfrac{1}{n}\right) \\
&\le (4C)^2\left(\tfrac{t-r}{2}+\tfrac{1}{n}\right)^2 \le (6C)^2(t-r)^2,
\end{aligned}
$$

which also holds if $t - r < \frac{1}{n}$ (because we have $S_n(r) = S_n(s)$ or $S_n(s) = S_n(t)$ in this case). Therefore Lemma 3.1 in Shorack and Wellner (1986) and (3.16) show that

$$
\begin{aligned}
\limsup_{n\to\infty}\mathbb{P}\left(\omega_n\left(\tfrac{1}{m}\right)\ge\epsilon\right) &\le \limsup_{n\to\infty}\frac{1}{\epsilon^4}\Big\{\sum_{k=1}^{m}\mathbb{E}\big[\big(S_n(\tfrac{k}{m})-S_n(\tfrac{k-1}{m})\big)^4\big] + \frac{K(6C)^2}{m}\Big\} \\
&= \frac{(4C)^2 + K(6C)^2}{\epsilon^4 m}
\end{aligned}
$$

for an absolute constant K. This proves (3.14) and completes the proof of (3.10).

Proof of (3.11) *and* (3.12)*:* These statements follow by similar arguments as given in the proof of assertion (3.10) using the estimates (A.1) - (A.7) in Appendix A. The details are omitted for the sake of brevity.

Proof of (3.13)*:* By (3.3) we have

$$
D^0_{n,2\lfloor nt\rfloor} = -(2\lfloor nt\rfloor+1)\lfloor nt\rfloor\log(2)\ ,
$$

and the estimate (A.1) from Section A yields the approximation

$$
\begin{aligned}
\sum_{i=1}^{\lfloor nt\rfloor-1}(\lfloor nt\rfloor-i+1)\mathbb{E}[\tilde{\xi}_{n,i}] &= \sum_{i=1}^{\lfloor nt\rfloor-1}(\lfloor nt\rfloor-i+1)\big(-4\log(2) - \tfrac{1}{2(n-i+1)} + O\big(\tfrac{1}{(2n-2i+1)^2}\big)\big) \\
&= -2\log(2)\big(\lfloor nt\rfloor^2+2\lfloor nt\rfloor\big) + \log(16) - \tfrac{\lfloor nt\rfloor}{2} + \tfrac{n-\lfloor nt\rfloor}{2}(G_n - G_{n-\lfloor nt\rfloor+1}) \\
&\quad + O(\log(n))
\end{aligned}
$$

(uniformly with respect to $t \in [0,1]$), where $G_n = \sum_{i=1}^n \frac{1}{i}$ is the nth partial sum of the harmonic series. Therefore

$$U_n(t) = \frac{1}{\sqrt{n}}\Big(\tfrac{nt-\lfloor nt \rfloor}{2} + \tfrac{n}{2}(1-t)\log(1-t) + \tfrac{n-\lfloor nt \rfloor}{2}(G_n - G_{n-\lfloor nt \rfloor+1})\Big) + O\Big(\frac{\log(n)}{\sqrt{n}}\Big).$$

Using the approximation $G_n = \log(n) + \gamma + O(\frac{1}{n})$, where γ is the Euler-Mascheroni constant, we can easily see that

$$\begin{aligned} U_n(t) &= \frac{1}{\sqrt{n}}\Big\{\tfrac{n}{2}(1-t)\log(1-t) - \tfrac{n-nt}{2}\log\Big(1-t+\tfrac{nt-\lfloor nt \rfloor+1}{n}\Big)\Big\} + O\Big(\frac{\log(n)}{\sqrt{n}}\Big) \\ &= \frac{-n(1-t)}{2\sqrt{n}}\log\Big(1+\tfrac{nt-\lfloor nt \rfloor+1}{n(1-t)}\Big) + O\Big(\frac{\log(n)}{\sqrt{n}}\Big) \ = \ o(1)\ , \end{aligned}$$

uniformly with respect $t \in [0,1]$, which completes the proof of Theorem 3.2. □

Remark 3.3. Similar results as stated in Theorem 3.1 and 3.2 can be obtained for the Hankel matrices $\underline{H}_{2n+1} = (m_{i+j+1})_{i,j=0}^n$, $\overline{H}_{2n} = (m_{i+j+1} - m_{i+j+2})_{i,j=0}^{n-1}$ and $\overline{H}_{2n+1} = (m_{i+j} - m_{i+j+1})_{i,j=0}^n$, which are commonly used to characterize Hausdorff moment sequences [see Karlin and Studden (1966)]. The details are omitted for the sake of brevity.

3.2 Hankel determinants from $\mathcal{M}_{2n}([0,\infty))$ and $\mathcal{M}_{2n}(\mathbb{R})$

In this section we will derive analogues of Theorem 3.2 for random moment sequences on unbounded moment spaces, where the corresponding distributions are defined by (2.10) and (2.11), respectively. For the sake of brevity we omit the discussion of $D_{n,2k}(\boldsymbol{m}_{2n})$ for fixed k (corresponding results can be easily obtained using similar arguments as given in the proof of Theorem 3.1) and concentrate on the stochastic process $\{D_{n,2\lfloor nt \rfloor}(\boldsymbol{m}_{2n})\}_{t\in[0,1]}$.

Theorem 3.4. *Let $\boldsymbol{m}_{2n}$ denote a random vector on $\mathcal{M}_{2n}([0,\infty))$ with density $g_n^{(\gamma,\delta)}$ defined in (2.10), where $\gamma_{2n,1}, \ldots \gamma_{2n,2n} > 0$ are bounded by a constant which does not depend on n and $\delta_{2n,i} = 2n - i + 1 + \gamma_{2n,i}$, then*

$$\big\{\mathcal{G}_n(t)\big\}_{t\in[0,1]} \ := \ \Big\{\frac{1}{n}D_{n,2\lfloor nt \rfloor}(\boldsymbol{m}_{2n})\Big\}_{t\in[0,1]} \ \Longrightarrow \ \{\mathcal{G}^{[0,\infty)}(t)\}_{t\in[0,1]}$$

where $D_{n,2\lfloor nt \rfloor}(\boldsymbol{m}_{2n})$ is defined in (3.1), and $\mathcal{G}^{[0,\infty)}$ is a continuous Gaussian process on the interval $[0,1]$ with mean $-r(t)/2$ and covariance kernel

$$\begin{aligned} g(s,t) &= \int_0^{s\wedge t} \frac{(t-x)(s-x)}{1-x}\,dx \\ &= \frac{1}{2}(s\wedge t)(s+t-2+s\vee t) + (s-1)(t-1)\log(1-s\wedge t) \end{aligned} \tag{3.17}$$

Proof: We will use the decomposition $\frac{1}{n}D_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n}) = S_n^{[0,\infty)}(t) + R_n^{[0,\infty)}(t)$, where the processes $S_n^{[0,\infty)}$ and $R_n^{[0,\infty)}$ are defined by

$$S_n^{[0,\infty)}(t) = \frac{1}{n}\left(D_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n}) - \mathbb{E}[D_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n})]\right)$$
$$R_n^{[0,\infty)}(t) = \frac{1}{n}\mathbb{E}[D_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n})]$$

Observing the fact that $bX \sim \gamma(a,1)$, whenever $X \sim \gamma(a,b)$, and using the approximations (A.8) - (A.10) from the Appendix it can be shown by similar arguments as given in the proof of Theorem 3.2 that $S_n^{[0,\infty)}$ converges weakly to a centered continuous Gaussian process on the interval $[0,1]$ with covariance kernel (3.17). For the remaining term $R_n^{[0,\infty)}(t)$ we use (2.4), Lemma 2.2 and the approximation

$$\begin{aligned}\mathbb{E}(\log(z_{2n,i})) &= \mathbb{E}\left[\log\left(z_{2n,i}\cdot(2n-i+1+\gamma_{2n,i})\right)\right] - \log(2n-i+1+\gamma_{2n,i})\\ &= -\frac{1}{2(2n-i+1+\gamma_{2n,i})} + O((2n-i+1+\gamma_{2n,i})^{-2}).\end{aligned}$$

This yields (uniformly with respect to $t \in [0,1]$)

$$R_n^{[0,\infty)}(t) = -\frac{1}{4n}\sum_{i=1}^{\lfloor nt\rfloor}\left(\frac{t-\frac{i}{n}+\frac{\lfloor nt\rfloor-nt+1}{n}}{1-\frac{i}{n}+\frac{\gamma_{2n,2i}+1/2}{n}} + \frac{t-\frac{i}{n}+\frac{\lfloor nt\rfloor-nt+1}{n}}{1-\frac{i}{n}+\frac{\gamma_{2n,2i-1}+1}{n}}\right) + O\left(\frac{\log(n)}{n}\right),$$

and a careful calculation shows that this term converges uniformly to $-r(t)/2$, where $r(t)$ is defined in (3.7). This yields the assertion. □

We conclude this section with a corresponding result for the moment space $\mathcal{M}_{2n}(\mathbb{R})$. The proof is similar to that of Theorem 3.2 and therefore omitted.

Theorem 3.5. *Let $\boldsymbol{m}_{2n-1}$ denote a random vector on $\mathcal{M}_{2n-1}(\mathbb{R})$ with density $h_n^{(\gamma,\delta)}$ defined in (2.10), where $\gamma_{n,1},\dots\gamma_{n,n}$ are bounded by a constant which does not depend on n and $\delta_{n,2i} = 2n-2i+\gamma_{n,i}$, then*

$$\left\{\frac{1}{n}D_{n,2\lfloor(n-1)t\rfloor}(\boldsymbol{m}_{2n-1})\right\}_{t\in[0,1]} \implies \{\mathcal{G}^{\mathbb{R}}(t)\}_{t\in[0,1]}$$

where $D_{n,2\lfloor(n-1)t\rfloor}(\boldsymbol{m}_{2n})$ is defined in (3.1), and $\mathcal{G}^{\mathbb{R}}$ is a continuous Gaussian process on the interval $[0,1]$ with mean $-r(t)/4$ and covariance kernel $g(s,t)/2$, defined by (3.7) and (3.8), respectively.

4 Large deviations

Throughout this section we consider large deviation principles (LDP) for the moment space $\mathcal{M}_{2n}([0,1])$. Similar results can be obtained for moment spaces corresponding to unbounded intervals. For fixed k the sequence $(H_n^k)_{n\in\mathbb{N}}$ defined in (3.2) for a uniformly distributed vector $\boldsymbol{m}_{2n}$ on the moment space $\mathcal{M}_{2n}([0,1])$ satisfies an LDP with a good rate function. To see this, observe that the sequence of canonical moments $(Y_n)_{n\in\mathbb{N}} = ((p_{2n,1},\dots,p_{2n,2k}))_{n\in\mathbb{N}}$ satisfies a large deviation principle with good rate function

$$I(x) = 2\sum_{i=1}^{2k}\big(-\log(x_i - x_i^2) - \log(4)\big)$$

(c.f. Gamboa and Lozada-Chang (2004)). As the function that maps the canonical moments to the logarithms of the Hankel-determinants is obviously continuous, the contraction principle [Theorem 4.2.1 in Dembo and Zeitouni (1998))] shows that $(H_n^k)_{n\in\mathbb{N}}$ satisfies an LDP with a good rate function. However, due to the complicated form of this map it is not possible to explicitly represent the corresponding rate function in terms of standard functions.
The investigation of LDP-properties of the logarithm of the lower Hankel determinant with increasing dimension turns out to be substantially more complicated, and we consider again the process $\{D_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n})\}_{t\in[0,1]}$, which has to be normalized differently, that is

$$Z_n(t) = -\frac{1}{n}\big(D_{n,2\lfloor nt\rfloor}(\boldsymbol{m}_{2n}) - D^0_{n,2\lfloor nt\rfloor}\big)\ ,$$

where $q_{n,0}=1$. Let $\mathcal{S}([0,1])$ denote the space of all signed regular Borel measures on the interval $[0,1]$ endowed with the weak-$*$-topology (with $(C([0,1]),||\cdot||_\infty)$ as predual). Then its (topological) dual space is the space $C([0,1])$ of all continuous functions on the interval $[0,1]$. In the following we interpret the process Z_n as the distribution function of a random measure $\nu_n\in\mathcal{S}([0,1])$. To be precise note that the process Z_n is piecewise constant with jumps at the points $\frac{1}{n},\dots,\frac{n}{n}$. Therefore ν_n is a linear combination of Dirac-measures and a simple calculation shows that

$$\nu_n = -\frac{1}{n}\sum_{i=1}^{n}\Big\{\sum_{j=1}^{i}\log(4q_{2n,2j-1}p_{2n,2j-1}) + \sum_{j=1}^{i-1}\log(4q_{2n,2j}p_{2n,2j}) + \log(2p_{2n,2i})\Big\}\delta_{\frac{i}{n}}\ , \tag{4.1}$$

where δ_t denotes the Dirac measure at the point $t\in[0,1]$. In order to investigate the large deviation properties of the sequence of random measures $\{\nu_n\}_{n\in\mathbb{N}}$ we first derive the limit of the (normalized) logarithmic moment generating function.

Theorem 4.1. *Let ν_n denote the random measure defined in (4.1). For any Riemann-integrable function $f\in l^\infty([0,1])$ we have*

$$\Lambda(f) = \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}[e^{n\nu_n(f)}] = \begin{cases} -\int\limits_0^1 \log\big(1-\frac{G(x)}{2(1-x)}\big)\,dx & K<2 \\ \infty & K>2\end{cases}\ ,$$

where

$$G(x) = \int_x^1 f(t)\,dt\ ; \qquad K = \sup_{x\in[0,1)}\frac{G(x)}{1-x}\ .$$

It is in general unknown what happens in the case $K = 2$.

Proof: Interpreting the sequences

$$x_{n,j} = -\sum_{i=j}^{n} f\left(\tfrac{i}{n}\right) \text{ and } y_{n,j} = -\sum_{i=j+1}^{n} f\left(\tfrac{i}{n}\right)$$

as Riemann-sums, we get the approximations

$$\sup_{j=1,\dots,n} \left|G\left(\tfrac{j}{n}\right) + \tfrac{x_{n,j}}{n}\right| \xrightarrow{n\to\infty} 0 \ , \ \sup_{j=1,\dots,n} \left|G\left(\tfrac{j}{n}\right) + \tfrac{y_{n,j}}{n}\right| \xrightarrow{n\to\infty} 0 \ . \tag{4.2}$$

This yields for the logarithm of the moment generating function

$$\begin{aligned}&\frac{1}{n}\log\left(\mathbb{E}[\exp(n\nu_n(f))]\right)\\ =&\frac{1}{n}\sum_{j=1}^{n}\left(\log\left(\mathbb{E}\left[q_{2n,2j-1}^{x_{n,j}} p_{2n,2j-1}^{x_{n,j}}\right]\right) + \log\left(\mathbb{E}\left[q_{2n,2j}^{y_{n,j}} p_{2n,2j}^{x_{n,j}}\right]\right) + (3x_{n,j} + y_{n,j})\log(2)\right) .\end{aligned}$$

For the determination of the limit we now consider the two cases $K > 2$ and $K < 2$ separately.
(1) In the case $K > 2$ we choose constants $\delta, C > 0$ such that for all sufficiently large n there exists a $j_n \in \{1, \dots, n\}$ with $1 - \frac{j_n}{n} > C$ and $G(\frac{j_n}{n}) \geq (2 + \delta)(1 - \frac{j_n}{n})$ (this is possible since the function G is continuous). Choosing another constant $0 < \epsilon < \delta C$ and considering (4.2), we get the following approximation for sufficiently large n:

$$\tfrac{x_{n,j_n}}{n} \leq \left|\tfrac{x_{n,j_n}}{n} + G\left(\tfrac{j}{n}\right)\right| - G\left(\tfrac{j_n}{n}\right) \leq \epsilon - (2+\delta)(1 - \tfrac{j_n}{n}) \ .$$

Therefore $2n - 2j_n + 1 + x_{n,j_n} \leq 1 + n(\epsilon - \delta C) < -1$, which yields $\mathbb{E}[q_{n,2j_n-1}^{x_{n,j_n}} p_{n,2j_n-1}^{x_{n,j_n}}] = \infty$ and the assertion follows.
(2) In the case $K < 2$ we use the formula

$$\log(\Gamma(x)) = (x - \tfrac{1}{2})\log(x) - x + \frac{\log(2\pi)}{2} + 2\phi_0(x) \ ,$$

where

$$\phi_0(x) = \int_0^\infty \frac{\arctan\left(\frac{t}{x}\right)}{\exp(2\pi t) - 1}\, dt \tag{4.3}$$

[cf. (4.3) in Dette and Gamboa (2007)]. Using the representation (4.1) we can show that

$$\begin{aligned}\frac{1}{n}\mathbb{E}[\exp(n\nu_n(f))] &= B_{n,1} + B_{n,2} + B_{n,3} + B_{n,4}\\ &\quad + R(2n - 2j + 2, x_{n,j}, x_{n,j}) + R(2n - 2j + 1, x_{n,j}, y_{n,j}) \ ,\end{aligned} \tag{4.4}$$

where

$$B_{n,1} = -\frac{1}{2n}\sum_{j=1}^{n} \log\left(1 + \tfrac{x_{n,j}}{2n-2j+2}\right) \ ,$$

$$B_{n,2} = -\frac{1}{2n}\sum_{j=1}^{n} \log\left(1 + \tfrac{x_{n,j}}{2n-2j+1}\right) \ ,$$

$$B_{n,3} = \frac{1}{n}\sum_{j=1}^{n} (2n-2j+1+x_{n,j}) \log\left(1 + \tfrac{-f(j/n)}{2(2n-2j+1)+x_{n,j}+y_{n,j}}\right) \ ,$$

$$B_{n,4} = \frac{1}{n}\sum_{j=1}^{n} (2n-2j+1+y_{n,j}-\tfrac{1}{2}) \log\left(1 + \tfrac{f(j/n)}{2(2n-2j+1)+x_{n,j}+y_{n,j}}\right) \ ,$$

and the remaining two terms are defined by

$$R(a,x,y) = 2(\phi_0(a+x) - \phi_0(a) + \phi_0(a+y) - \phi_0(a)) - 4(\phi_0(2a+x+y) - \phi_0(2a)) \ .$$

We now investigate the terms in this decomposition separately. The first term $B_{n,1}$ can be interpreted as Riemann-sum, using (4.2), that is

$$B_{n,1} = -\frac{1}{2n}\sum_{j=1}^{n} \log\left(1 - \frac{G\left(\frac{j}{n}\right)+o(1)}{2\left(1-\frac{j-1}{n}\right)+o(1)}\right) \xrightarrow{n\to\infty} -\frac{1}{2}\int_0^1 \log\left(1 - \frac{G(x)}{2(1-x)}\right) dx \ . \tag{4.5}$$

Analogously, the second term converges to the same limit, i.e.

$$B_{n,2} \xrightarrow{n\to\infty} -\frac{1}{2}\int_0^1 \log\left(1 - \frac{G(x)}{2(1-x)}\right) dx \ . \tag{4.6}$$

For the the terms $B_{n,3}$ and $B_{n,4}$ we use the Taylor-approximation $\log(1+x) = x + O(x^2)$ $(x \to 0)$ and obtain

$$B_{n,3} = -\frac{1}{2n}\sum_{j=1}^{n} \frac{2\left(1-\frac{j}{n}\right) - G\left(\frac{j}{n}\right) + o(1)}{2\left(1-\frac{j}{n}\right) - G\left(\frac{j}{n}\right) + o(1)} f\left(\tfrac{j}{n}\right) + O\Big(\frac{1}{n}\Big) \xrightarrow{n\to\infty} -\frac{G(0)}{2} \ , \tag{4.7}$$

$$B_{n,4} = \xrightarrow{n\to\infty} \frac{G(0)}{2} \ , \tag{4.8}$$

and it remains to show that the last two terms in (4.4) are asymptotically negligible.
For this purpose we note that the following inequality holds for the function ϕ_0 defined in (4.3) [cf. formula (4.10) in Dette and Gamboa (2007)]

$$|\phi_0(a+x) - \phi_0(a)| \leq C\frac{|x|}{(a \wedge (a+x))^2} \quad \text{with} \quad C = \int_0^\infty \frac{t}{\exp(2\pi t)-1}\, dt \ ,$$

where $a > 0, x > -a$. This gives

$$|R(a,x,y)| \le 2C\Big(\frac{|x|}{(a\wedge(a+x))^2} + \frac{|y|}{(a\wedge(a+y))^2}\Big) + 4C\Big(\frac{|x+y|}{(2a\wedge(2a+x+y))^2}\Big) ,$$

and using this inequality to estimate the terms $R(2n-2j+2, x_{n,j}, x_{n,j})$ and $R(2n-2j+1, x_{n,j}, y_{n,j})$ in (4.4) yields six terms, which have a similar form. For the sake of brevity we will only show exemplarily the convergence

$$D_n := \frac{1}{n}\sum_{j=1}^{n} \frac{|x_{n,j}|}{((2n-2j+2)\wedge(2n-2j+2+x_{n,j}))^2} \xrightarrow{n\to\infty} 0 .$$

The other five sums can be approximated in a similar way and the details are omitted. For sufficiently small $\epsilon > 0$ and sufficiently large n, we obtain by similar arguments as in the case $K > 2$:

$$x_{n,j} \ge -(\epsilon + K)(n-j) - \epsilon j ,$$
$$2n - 2j + x_{n,j} \ge (2 - K - 2\epsilon)(n-j) .$$

Choosing $\delta = \min\{2 - K - 2\epsilon, 2\} > 0$, we get the inequalities

$$(2n-2j+2)\wedge(2n-2j+2+x_{n,j}) \ge \delta(n-j+1) ,$$
$$|x_{n,j}| \le n(|G(\tfrac{j}{n})| + \epsilon) \le ||f||_\infty(n-j+1) + n\epsilon .$$

This yields

$$\limsup_{n\to\infty} D_n \ \le\ \limsup_{n\to\infty}\Big\{\frac{||f||_\infty}{\delta^2 n}\sum_{j=1}^{n}\frac{1}{n-j+1} + \frac{\epsilon}{\delta^2}\sum_{j=1}^{n}\frac{1}{(n-j+1)^2}\Big\} \ =\ \frac{\epsilon}{\delta^2}\sum_{j=1}^{\infty}\frac{1}{j^2} . \tag{4.9}$$

Considering the limit $\epsilon \searrow 0$ on the right hand side of (4.9) we obtain that the last two terms in (4.4) converge to 0, and the assertion follows from (4.5) - (4.8).

□

Lemma 4.2. *The sequence $(\nu_n)_{n\in\mathbb{N}}$ of random measures defined by (4.1) is exponentially tight.*

Proof: By the Banach-Alaoglu theorem the set

$$K_\alpha = \Big\{\mu \in \mathcal{M}([0,1]) \ \Big|\ \sup_{\substack{f\in C([0,1])\\ ||f||\le 1}} \mu(f) \le \alpha\Big\}$$

is compact (note that we endowed $\mathcal{S}([0,1])$ with the weak-$*$-topology). We define the modified measure

$$\nu_n' = -\frac{1}{n}\sum_{i=1}^{n}\Big\{\sum_{j=1}^{i}\log(4q_{2n,2j-1}p_{2n,2j-1}) + \sum_{j=1}^{i-1}\log(4q_{2n,2j}p_{2n,2j}) + \log(p_{2n,2i})\Big\}\delta_{\frac{i}{n}} .$$

Observing $\nu_n = \nu_n' - \frac{\log(2)}{n} \sum_{i=1}^{n} \delta_{\frac{i}{n}}$ we can see $\nu_n(f) \le \nu_n'(f) + \log(2)$ for all $f \in C([0,1])$ with $||f||_\infty \le 1$. Since ν_n' is a positive measure, we get by Markov's inequality

$$\begin{aligned}\frac{1}{n}\log \mathbb{P}(\nu_n \in K_\alpha^c) &\le \frac{1}{n}\log \mathbb{P}\Big(\sup_{\substack{f\in C([0,1])\\ ||f||\le 1}} \nu_n'(f) > \alpha - \log(2)\Big) \le \frac{1}{n}\log \mathbb{E}[\exp(n\nu_n'(1))] - \alpha + \log(2)\\ &= \frac{1}{n}\log \mathbb{E}[\exp(n(\nu_n(1)+\log(2))] - \alpha + \log(2) \xrightarrow{n\to\infty} \Lambda(1) - \alpha + 2\log(2)\ ,\end{aligned}$$

which yields the assertion.

□

Theorem 4.3. *Let Λ^* be the Fenchel-Legendre transform of Λ and let E denote the set of all exposed points of Λ^* which have an exposing hyperplane λ that satisfies*

$$\lim_{n\to\infty} \frac{1}{n}\log \mathbb{E}[\exp(\nu_n(n\lambda)] \text{ exists and } \Lambda(\gamma\lambda) < \infty \text{ for some } \gamma > 1.$$

Then

$$-\inf_{x\in E\cap\Gamma^\circ} \Lambda^*(x) \le \liminf_{n\to\infty} \frac{1}{n}\log \mathbb{P}(\nu_n \in \Gamma) \le \limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(\nu_n \in \Gamma) \le -\inf_{x\in\overline{\Gamma}} \Lambda^*(x)$$

for all measurable sets $\Gamma \subset \mathcal{S}([0,1])$.

Proof: This follows directly from Baldi's theorem [c.f. Theorem 4.5.20 in Dembo and Zeitouni (1998)].

□

The main difficulty in proving an LDP for the process $\{Z_n(t)\}_{t\in[0,1]}$ consists in the fact that an explicit representation of the Fenchel-Legendre transform Λ^* is not available. This makes it difficult to eliminate the set E in the lower bound in Theorem 4.3. On the other hand - in contrast to the LDP for the process $\{Z_n(t)\}_{t\in[0,1]}$ - the LDP for the random variable $Z_n(t)$ with a fixed t can be established.

Theorem 4.4. *For a fixed $t \in (0,1]$ the sequence $(Z_n(t))_{n\in\mathbb{N}}$ satisfies a large deviation principle with good rate function*

$$\Lambda^*(x) = \sup_{\lambda < \frac{2}{t}} \Big\{ \lambda x + \int_0^t \log\Big(1 - \tfrac{\lambda(t-x)}{2(1-x)}\Big)\ dx \Big\}$$

Proof: We will again apply Baldi's theorem. To calculate the normalized cumulant generating function of $Z_n(t)$, note that

$$\mathbb{E}[\exp(\lambda Z_n(t))] = \mathbb{E}[\nu_n(\lambda I\{\cdot \le t\})]\ ,$$

and Theorem 4.1 yields

$$\Lambda_t(\lambda) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[\exp(n\lambda Z_n(t))]$$
$$= \Lambda(\lambda I\{\cdot \le t\}) = \begin{cases} -\int_0^t \log\left(1 - \frac{\lambda(t-x)}{2(1-x)}\right)\, dx & \lambda < \frac{2}{t} \\ \infty & \lambda > \frac{2}{t} \end{cases}.$$

It now follows by similar arguments as given in the proof of Lemma 4.2 that the sequence $(Z_n(t))_{n\in\mathbb{N}}$ is exponentially tight (note that we can use the euclidean topology on $\mathbb{R}$ because the interval $[0, \alpha]$ is compact) and Baldi's theorem yields an analogue of the inequality in Theorem 4.3, where the set E has to be replaced by an analogue set E_t. It remains to prove that the lower bound remains correct if one removes the set E_t.

In order to see this, we define the new function

$$\tilde{\Lambda}_t : \begin{cases} \mathbb{R} & \to (-\infty, \infty] \\ \lambda & \mapsto \begin{cases} \Lambda_t(\lambda) & \text{if } \lambda \ne \frac{2}{t} \\ \lim\limits_{\epsilon \searrow 0} \Lambda_t(\lambda - \epsilon) & \text{if } \lambda = \frac{2}{t} \end{cases} \end{cases}$$

Λ_t and $\tilde{\Lambda}_t$ have the same Fenchel-Legendre transform and it is therefore sufficient to prove

$$\inf_{x\in E_t\cap F} \tilde{\Lambda}_t^*(x) = \inf_{x\in F} \tilde{\Lambda}_t^*(x) \qquad (4.10)$$

for all open sets $F \subset \mathbb{R}$. It is easy to see that $\tilde{\Lambda}_t$ is an essentially smooth function and the identity (4.10) follows by an adaptation of the arguments in the proof of the Gärtner-Ellis theorem [Theorem 2.3.6 in Dembo and Zeitouni (1998)]. By Lemma 1.2.18 in the same reference the rate function Λ^* is a good rate function, which yields the assertion.

□

Our final result specializes Theorem 4.4 to the case $t = 1$, where the rate function can be determined explicitly. The proof follows by a straightforward calculation of $\Lambda_1(\lambda)$ and its convex conjugate.

Corollary 4.5. *The sequence $(Z_n(1))_{n\in\mathbb{N}}$ satisfies an LDP with good rate function*

$$I(x) = \begin{cases} 2x - 1 - \log(2x) & x > 0 \\ \infty & \textit{else} \end{cases}.$$

A Auxiliary results

In the proof of the results we make frequent use of the following approximations, which can can be derived from the approximations given in Dette and Gamboa (2007). Throughout this section

C denotes a positive constant.

$$\left|\mathbb{E}\left[\tilde{\xi}_{n,i}\right] + 4\log(2) + \frac{1}{2(n-i+1)}\right| \leq \frac{C}{(2n-2i+1)^2} \tag{A.1}$$

$$\left|\operatorname{Var}\left(\tilde{\xi}_{n,i}\right) - \frac{1}{4(n-i+1)^2}\right| \leq \frac{C}{(2n-2i+1)^3} \tag{A.2}$$

$$|\mathbb{E}[\log(q_{n,i})] + \log(2)| \leq \frac{C}{n-i+1} \tag{A.3}$$

$$\operatorname{Var}\left(\tilde{\xi}_{n,i}\right) \leq \frac{C}{(n-i+1)^2} \tag{A.4}$$

$$\mathbb{E}\left[\left(\tilde{\xi}_{n,i} - \mathbb{E}[\tilde{\xi}_{n,i}]\right)^4\right] \leq \frac{C^2}{(n-i+1)^4} \tag{A.5}$$

Also, a direct estimate of the occurring integrals show for $i < 2n$:

$$\mathbb{E}\left[|\log(q_{2n,i})|^k\right] \leq 3 \sup_{x\in[0,1]} |\log(x)|^k x < \infty \tag{A.6}$$

Lastly, one can prove by differentiation under the integral that for a random variable $X \sim \beta(a,b)$

$$\operatorname{Var}(\log(X)) = \psi_1(a) - \psi_1(a+b) \tag{A.7}$$

where $\psi_1(x) = \frac{d^2}{dx^2}\log(\Gamma(x)) = x^{-1} + O(x^{-2}) \quad (x \to \infty)$ denotes the trigamma function.

We also need to approximate the moments of gamma-distributed random variables. Using the notation $d_i = \frac{\Gamma^{(i)}(k)}{\Gamma(k)}$ we can see that

$$\begin{aligned}
\frac{d}{dk}\log(\Gamma(k)) = d_1 &= \log(k) - \frac{1}{2k} + O(k^{-2}) \\
\frac{d^2}{dk^2}\log(\Gamma(k)) = d_2 - d_1^2 &= \frac{1}{k} + \frac{1}{2k^2} + O(k^{-3}) \\
\frac{d^4}{dk^4}\log(\Gamma(k)) = d_4 - 3d_2^2 - 6d_1^4 - 4d_3a_1 + 12d_1^2d_2 &= O(k^{-3})
\end{aligned}$$

where the first part of the equations follows from formally differentiating the term, while the second part follows from the approximations of the polygamma functions in Abramowitz and Stegun (1964). If $X \sim \gamma(k,1)$, then for $Y = \log(X)$ the following equations hold

$$\mathbb{E}[Y] = d_1 = \log(k) - \frac{1}{2k} + O(k^{-2}) \tag{A.8}$$

$$\operatorname{Var}(Y) = d_2 - d_1^2 = \frac{1}{k} + \frac{1}{2k^2} + O(k^{-3}) \tag{A.9}$$

$$\begin{aligned}
\mathbb{E}[(Y - \mathbb{E}[Y])^4] &= d_4 - 4d_1d_3 + 6d_1^2d_2 - 3d_1^4 = 3(d_2 - d_1^2)^2 + O(k^{-3}) \\
&= \frac{3}{k^2} + O(k^{-3})
\end{aligned} \tag{A.10}$$

Acknowledgements. The authors would like to thank M. Stein who typed this manuscript with considerable technical expertise. The work of H. Dette was partially supported by the Deutsche Forschungsgemeinschaft (DFG Research Unit 1735).

References

Abramowitz and Stegun (1964). *Handbook of Mathematical Functions.* Dover Publications.

Barthe, F., Gamboa, F., Lozada Chang, L. V., and Rouault, A. (2010). Generalized Dirichlet distributions on the ball and moments. *Alea.*, 7:319–340.

Basor, E. L., Chen, Y., and Widom, H. (2001). Determinants of Hankel matrices. *Journal of Functional Analysis*, 179(1):214–234.

Berg, C. and Szwarc, R. (2011). The smallest eigenvalue of Hankel matrices. *Constructive Approximation*, 34(1):107–133.

Bose, A. and Sen, A. (2007). Spectral norm of random large dimensional noncentral Toeplitz and Hankel matrices. *Electronic Communications in Probability*, 12:21–27.

Bryc, W., Dembo, A., and Jiang, T. (2006). Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Annals of Probability*, 34(1):1–38.

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

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

Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications.* Springer, 2. edition edition.

Dette, H. and Gamboa, F. (2007). Asymptotic properties of the algebraic moment range process. *Acta Math. Hungar.*, 116:247–264.

Dette, H. and Nagel, J. (2012). Distributions on unbounded moment spaces and random moment sequences. *The Annals of Probability*, 40(6):2690–2704.

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

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

Edelman, A. and Sutton, B. D. (2008). The beta-jacobi matrix model, the cs decomposition, and generalized singular value problems. *Found. Comput. Math.*, 8:259–285.

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

Girko, V. L. (1993). *Theory of Random Determinants*. Kluwer Academic Publishers, Dordrecht.

Hiai, F. and Petz, D. (2000). *The Semicircle Law, Free Random Variables and Entropy*. American Mathematical Society, R.I.

Hirschman Jr., I. I. (1966). The strong Szegö limit theorem for Toeplitz determinants. *American Journal of Mathematics*, 88(3):577–614.

Hora, A. and Obata, N. (2007). *Quantum Probability and Spectral Analysis of Graphs*. Springer, Berlin.

Kallenberg, O. (2002). *Foundations of Modern Probability*. Springer.

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). *Tchebycheff systems: with applications in analysis and statistics*. Interscience Publishers.

Killip, R. and Nenciu, I. (2004). Matrix models for circular ensembles. *Int. Math. Res. Not.*, 50:2665–2701.

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.

Krein, M. G. and Nudelman, A. A. (1977). *The Markov Moment Problem and Extremal Problems.* American Mathematical Society., Providence, RI.

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

Ramírez, J., Rider, B., and Virág, B. (2011). Beta ensembles, stochastic airy spectrum, and a diffusion. *J. Amer. Math. Soc.*, 24:919–944.

Shohat, J. A. and Tamarkin, J. D. (1943). *The Problem of Moments.* American Mathematical Society, New York.

Shorack, G. and Wellner, J. (1986). *Empirical Processes With Applications to Statistics*. Society for Industrial and Applied Mathematics.

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.

van der Vaart, A. and Wellner, J. (1995). *Weak Convergence and Empirical Processes*. Springer.

Verblunsky, S. (1935). On positive harmonic functions: A contribution to the algebra of Fourier series. *Proc. London Math. Soc.*, 38:125–157.

Verblunsky, S. (1936). On positive harmonic functions (second paper). *Proc. London Math. Soc.*, 40:290–320.

Wall, H. S. (1948). *Analytic Theory of Continued Fractions.* D. van Nostrand Company, Inc. XIII, New York.

Widom, H. (1966). Hankel matrices. *Transactions of the American Mathematical Society*, 121:1–35.

Zamarashkin, N. L. and Tyrtyshnikov, E. E. (2001). Eigenvalue estimates for Hankel matrices. *Sbornik: Mathematics*, 192(4):537.