Determinants of Random Block Hankel Matrices

Dominik Tomecki

Ruhr-Universität Bochum

Fakultät für Mathematik

44780 Bochum, Germany

email: dominik.tomecki@rub.de

Holger Dette

Ruhr-Universität Bochum

Fakultät für Mathematik

44780 Bochum, Germany

email: holger.dette@rub.de

June 27, 2017

Abstract

We consider the moment space $\mathcal{M}_{2n+1}^{p_n}$ of moments up to the order $2n + 1$ of $p_n \times p_n$ real matrix measures defined on the interval $[0, 1]$. The asymptotic properties of the Hankel determinant

$$\big\{ \log \det(M_{i+j}^{p_n})_{i,j=0,\ldots,\lfloor nt \rfloor} \big\}_{t \in [0,1]}$$

of a uniformly distributed vector $(M_1, \ldots, M_{2n+1})^t \sim \mathcal{U}(\mathcal{M}_{2n+1})$ are studied when the dimension n of the moment space and the size of the matrices p_n converge to infinity. In particular weak convergence of an appropriately centered and standardized version of this process is established. Mod-Gaussian convergence is shown and several large and moderate deviation principles are derived. Our results are based on some new relations between determinants of subblocks of the Jacobi-beta-ensemble, which are of their own interest and generalize Bartlett decomposition-type results for the Jacobi-beta-ensemble from the literature.

Keyword and Phrases: mod-ϕ-convergence, moment spaces, matrix measures, large deviations, Jacobi-beta-ensemble, Bartlett-decomposition
AMS Subject Classification: 60F05, 60F10, 44A60, 47B35

1 Introduction

A $p \times p$ matrix $\mu = (\mu_{i,j})_{i,j=1}^p$ of signed real measures $\mu_{i,j}$, such that for each Borel set $A \subset [0, 1]$ the matrix $\mu(A) = (\mu_{i,j}(A))_{i,j=1}^p$ is symmetric and nonnegative definite, is called matrix measure on the interval $[0, 1]$. Matrix measures have been studied extensively in the literature generalising many classical results in the context of moment theory, orthogonal polynomials, quadrature formulas [see Krein (1949) for an early reference and Duran and van Assche (1995), Duran (1999) and Duran and Lopez-Rodriguez (1997), Grünbaum (2003), Grünbaum et al. (2005), Damanik et al. (2008), Gamboa et al. (2012) and Gamboa et al. (2016) for some more recent references among many others].
In a recent paper Dette and Nagel (2012b) defined $\mathcal{P}_p$ as the set of all matrix measures on the interval $[0, 1]$ satisfying the condition $\int_0^1 d\mu(x) = I_p$ [here and throughout this paper

I_p denotes the $p \times p$ identity matrix] and studied a uniform distribution on the moment space

$$\mathcal{M}_n^p = \left\{ (M_1, \ldots, M_n)^t \mid M_j = \int_0^1 x^j d\mu(x), \ j = 1, \ldots, n \ , \ \mu \in \mathcal{P}_p \right\} \qquad (1.1)$$

of all (matrix) moments

$$M_k = \int_0^1 x^k d\mu(x) = \left(\int_0^1 x^k d\mu_{i,j}(x) \right)_{1 \le i,j \le p} ; \quad k = 0, 1, 2, \ldots \qquad (1.2)$$

up to the order n corresponding to matrix measures on the interval $[0,1]$. Note that $\mathcal{M}_n^p \subset (\mathcal{S}_p)^n$, where $\mathcal{S}_p$ denotes the set of all nonnegative definite matrices of size $p \times p$. If $(M_{1,n}^p, \ldots, M_{n,n}^p)^t$ is a uniformly distributed vector on $\mathcal{M}_n^p$, Dette and Nagel (2012b) showed that an appropriately centered and standardized version of the vector $(M_{1,n}^p, \ldots, M_{k,n}^p)^t$ converges weakly to a vector of k independent $p \times p$ Gaussian ensembles, thus generalising the meanwhile classical results of Chang et al. (1993) for the case $p = 1$.

The one-dimensional case has also been studied intensively with respect to other properties of random moment sequences and we refer to Gamboa and Lozada-Chang (2004), Lozada-Chang (2005) for large deviation results and to Dette and Nagel (2012a) for some results on more general moment spaces. Recently Dette and Tomecki (2016) examined the asymptotic properties of a stochastic process of Hankel determinants

$$\left\{ \log \det(M_{i+j,2n}^p)_{i,j=0,\ldots,\lfloor nt \rfloor} \right\}_{t \in [0,1]}$$

of a uniformly distributed moment vector $(M_{1,2n}^p, \ldots, M_{2n,2n}^p)^t$ on the moment space $\mathcal{M}_{2n}^p$ in the case $p = 1$ and derived weak convergence and large deviation principles for this process.

In the present paper we will investigate properties of a stochastic process corresponding to the determinant of matrix valued random Hankel matrices, where the dimension of the moment space and the dimension p_n of the matrix measures converge to infinity. To be precise, consider a uniformly random vector $(M_{1,2n+1}^{p_n}, \ldots, M_{2n+1,2n+1}^{p_n})^t$ on the moment space $\mathcal{M}_{2n+1}^{p_n}$, where p_n is a sequence of integers converging to infinity as $n \to \infty$ and define the stochastic process

$$H_n(t) = \left\{ \log \det(M_{i+j,2n+1}^{p_n})_{i,j=0,\ldots,\lfloor nt \rfloor} \right\}_{t \in [0,1]}. \qquad (1.3)$$

We establish weak convergence of the process $\{H_n(t)\}_{t \in [0,1]}$ with a Gaussian limit, Mod-Gaussian convergence (for fixed $t \in [0,1]$) and derive several moderate and large deviation principles. In Section 2 we will present some basic facts about matrix-valued moment spaces. Section 3 is devoted to the investigation of distributional properties of determinants corresponding to subblocks of the Jacobi-beta-ensemble. These results are of own interest and provide a new point on classical results about a Bartlett- type decomposition [see Bartlett (1933)] for the Jacobi-beta-ensemble [see for example Kshirsagar (1961)]. In Section 4 we prove weak convergence of the process (1.3). Finally, in Section 5 we examine mod-ϕ-convergence, as well as moderate and large deviations. Our results are based on several delicate estimates of the cumulants of logarithms of beta-distributed random variables, stated in the Appendix, which also contains proofs of the more technical results and some inequalities about polygamma functions.

2 Moment spaces of matrix-valued meausres

We begin recalling some basic facts about the moment space $\mathcal{M}_n^p$ defined in (1.1) [see Dette and Studden (2002) for a detailed discussion]. In the following we compare matrices with respect to the Löwner (partial) ordering. Thus for two $p \times p$ symmetric matrices A, B we use the notation $A < B$ $(A \leq B)$ if and only if the difference $B - A$ is positive definite (positive semi-definite). We denote by $\mathcal{S}_p$ the set of nonnegative definite (symmetric) $p \times p$ matrices. Let $\mathcal{B}([0, 1])$ denote the Borel field on the interval $[0, 1]$. A map

$$\mu = (\mu_{i,j})_{1 \leq i,j \leq p} : \mathcal{B}([0, 1]) \to \mathbb{R}^{p \times p}$$

is called a *matrix-valued measure*, if $\mu_{i,j}$ is a signed measure for all $1 \leq i, j \leq p$ and $\mu(A) \in \mathcal{S}_p$ for every Borel set $A \subset [0, 1]$. We denote by $\mathcal{P}_p$ the set of all $p \times p$ matrix-valued measures on the interval $[0, 1]$ satisfying $\mu([0, 1]) = I_p$ and consider the nth moment space $\mathcal{M}_n^p$ defined in (1.1), which is a subset of $(\mathcal{S}_p)^n$. Note that in the one dimensional case $\mathcal{P}_1$ is the set of all probability measures on the interval $[0, 1]$.

Dette and Studden (2002) introduced new "coordinates" for the moment space $\mathcal{M}_n^p$ defining a one to one map from the interior of $\mathcal{M}_n^p$ onto the product space $(\mathcal{E}_p)^n$, where $\mathcal{E}_p$ denotes the "cube"

$$\mathcal{E}_p = \{A \in \mathcal{S}_p \mid 0_p < A < I_p\}.$$

Here and throughout the paper, 0_p denotes the $p \times p$ matrix with all elements equal to zero and I_p denotes the $p \times p$ identity matrix.

The new coordinates are called canonical moments [see Dette and Studden (2002)], and they are related to the Verblunsky coefficients, which have been discussed for matrix measures on the unit circle [see Damanik et al. (2008) and Simon (2005a,b)]. They turn out to be extremely useful in analyzing the asymptotic properties of the stochastic process defined in (1.3). The definition of matrix valued canonical moments relies on the introduction of Block-Hankel-matrices:

$$\underline{H}_{2k} = \begin{pmatrix} M_0 & \cdots & M_k \\ \vdots & \ddots & \vdots \\ M_k & \cdots & M_{2k} \end{pmatrix} \qquad \overline{H}_{2k} = \begin{pmatrix} M_1 - M_2 & \cdots & M_k - M_{k+1} \\ \vdots & \ddots & \vdots \\ M_k - M_{k+1} & \cdots & M_{2k-1} - M_{2k} \end{pmatrix}$$

$$\underline{H}_{2k+1} = \begin{pmatrix} M_1 & \cdots & M_{k+1} \\ \vdots & \ddots & \vdots \\ M_{k+1} & \cdots & M_{2k+1} \end{pmatrix} \qquad \overline{H}_{2k+1} = \begin{pmatrix} M_0 - M_1 & \cdots & M_k - M_{k+1} \\ \vdots & \ddots & \vdots \\ M_k - M_{k+1} & \cdots & M_{2k} - M_{2k+1} \end{pmatrix}$$

which - as in the one dimensional case - can be used to characterize elements of the moment space $\mathcal{M}_n^p$. More precisely, the vector of matrices $M = (M_1, \ldots, M_n)^t \in (\mathcal{S}_p)^n$ satisfies

$$M \in \mathcal{M}_p^n \quad \text{if and only if} \quad \underline{H}_k \geq 0_k, \overline{H}_k \geq 0_k \quad \text{for all} \quad k \leq n$$
$$M \in \text{Int}\left(\mathcal{M}_p^n\right) \quad \text{if and only if} \quad \underline{H}_k > 0_k, \overline{H}_k > 0_k \quad \text{for all} \quad k \leq n \ ,$$

where $\text{Int}\left(\mathcal{M}_p^n\right)$ denotes the interior of the set $\mathcal{M}_p^n$. We now introduce the vectors of matrices

$$\underline{h}_{2k}^t = \left(M_{k+1}, \ldots, M_{2k}\right), \ \underline{h}_{2k-1}^t = \left(M_k, \ldots, M_{2k-1}\right)$$
$$\overline{h}_{2k}^t = \left(M_k - M_{k+1}, \ldots, M_{2k-1} - M_{2k}\right), \ \overline{h}_{2k-1}^t = \left(M_k - M_{k+1}, \ldots, M_{2k-2} - M_{2k-1}\right)$$

and define the "extremal" matrices $M_1^- = 0_p$, $M_1^+ = I_p$ and $M_2^+ = M_1$ (the phrase extremal will shortly become clear). If M is an element of the interior of the moment space $\mathcal{M}_p^n$, the extremal moments of larger order are defined as

$$M_n^- = \underline{h}_{n-1}^t \underline{H}_{n-2}^{-1} \underline{h}_{n-1} \quad \text{for } n \geq 2$$
$$M_n^+ = M_{n-1} - \overline{h}_{n-1}^t \overline{H}_{n-2}^{-1} \overline{h}_{n-1} \quad \text{for } n \geq 3.$$

Dette and Studden (2002) showed that the extremal moments provide a convenient tool to characterise the moment space $\mathcal{M}_p^n$. In particular by considering Schur complements of $\underline{H}_k$ and $\overline{H}_k$, they showed that

$$(M_1, \ldots, M_n)^t \in \text{Int}\left(\mathcal{M}_p^n\right) \quad \text{if and only if} \quad M_k^- < M_k < M_k^+ \quad \text{for all} \quad k \leq n .$$

This property is then used to define for a point $(M_1, \ldots, M_n)^t \in \text{Int}\left(\mathcal{M}_p^n\right)$ matrix valued canonical moments as follows

$$\mathcal{U}_i = (M_i^+ - M_i^-)^{-1/2}(M_i - M_i^-)(M_i^+ - M_i^-)^{-1/2} , \quad i = 1, \ldots, n. \tag{2.1}$$

Given $M_1, \ldots M_{k-1}$, the moment M_k can be calculated from the canonical moment U_k and therefore equation (2.1) defines a one to one mapping

$$\varphi_p : \left\{ \begin{array}{ll} \text{Int}\left(\mathcal{M}_n^p\right) & \to \mathcal{E}_p^n \\ (M_{1,n}^p, \ldots, M_{n,n}^p)^t & \mapsto \varphi_p(M_{1,n}^p, \ldots, M_{n,n}^p) = (U_{1,n}^p, \ldots, U_{n,n}^p)^t , \end{array} \right.$$

from the interior of the moment space onto $\mathcal{E}_p^n$.

We conclude this section with a very interesting and useful relation between canonical moments and determinants of Hankel matrices $\underline{H}_{2n} = \det(M_{i+j,2n}^p)_{i,j=0,\ldots,n}$ corresponding to a point $(M_{1,2n}^p, \ldots, M_{2n,2n}^p)^t \in \text{Int}(\mathcal{M}_{2n}^p)$, i.e

$$\det \underline{H}_{2n} = \prod_{i=1}^n \left(\det(I_p - \mathcal{U}_{2i-2,2n}^p) \det(\mathcal{U}_{2i-1,2n}^p) \det(I_p - \mathcal{U}_{2i-1,2n}^p) \det(\mathcal{U}_{2i,2n}^p) \right)^{n-i+1} \tag{2.2}$$

$$= \prod_{i=1}^n \left\{ \det(\mathcal{U}_{2i-1,2n}^p) \det(I_p - \mathcal{U}_{2i-1,2n}^p) \det(\mathcal{U}_{2i,2n}^p) \right\}^{n-i+1} (\det(I_p - \mathcal{U}_{2i,2n}^p))^{n-i},$$

where $U_0 = 0_p$ [see Dette and Studden (2005) for a proof].

3 The distribution of random Hankel block matrices

By identifying a symmetric matrix $M = (m_{i,j})_{i,j=1}^p \in \mathcal{S}_p$ with the vector containing the entries $m_{i,j}$ for $1 \leq i \leq j \leq p$, we get a subset of $\mathbb{R}^{p(p+1)/2}$ with non-empty interior. This identification allows us to integrate on $\mathcal{S}_p$ via the usual Lebesgue-measure and to define distributions on $\mathcal{S}_p$ by specifying their (Lebesgue-)densities. We are particularly interested in matrix valued Beta distributions [see Olkin and Rubin (1964) or Muirhead (1982)] supported on the set $\mathcal{E}_p$ with a density proportional to

$$f_{\gamma,\delta}(U) = (\det U)^{\gamma - e_1}(\det(I_p - U))^{\delta - e_1} I\{0_p < U < I_p\}. \tag{3.1}$$

where $e_1 = \frac{p+1}{2}$ and the parameters γ, δ satisfy $\gamma, \delta > e_1 - 1$. These distributions are a special case of the Jacobi-beta-ensemble $J\beta E_p(\gamma, \delta)$, which defines a density to be proportional to (3.1), where the constant e_β is given by $e_\beta = 1 + \frac{\beta}{2}(p-1)$. The parameter β

varies when the entries of the matrix are are real ($\beta = 1$), complex ($\beta = 2$) or quaternions ($\beta = 4$) [see for example Arashi et al. (2011)]. In the present paper, we will only consider the real case.

Consider a uniform distribution on the n-th moment space $\mathcal{M}_n^p$ defined in (1.2) and denote by $\mathcal{U}_{i,n}^p$ the i-th canonical moment in (2.1), then it is shown in Theorem 3.5 of Dette and Nagel (2012b) that $\mathcal{U}_{i,n}^p, \ldots, \mathcal{U}_{i,n}^p$ are independent and distributed according to a Jacobi-beta-ensemble, that is

$$\mathcal{U}_{i,n}^p \sim J\beta E_p(e_1(n-i+1), e_1(n-i+1)). \tag{3.2}$$

¿From this result and formula (2.2) it is obvious that the distribution of determinants of random variables governed by the Jacobi-Ensemble will be essential for the following analysis of the process. Our first main result, which is of independent interest, provides an important tool to determine the distribution of the process $\{H_n(t)\}_{t\in[0,1]}$ defined in (1.3). Throughout this paper we will use the notation $M^{[k]} \in \mathbb{R}^{k \times k}$ for the upper left $k \times k$ subblock of the matrix $M \in \mathbb{R}^{p \times p}$ and $\beta(\gamma, \delta)$ denotes a Beta distribution on the interval $[0, 1]$.

Theorem 3.1 (Subblocks of Jacobi-Ensembles). *Assume $p > 1$ and that $U \sim J\beta E_p(\gamma, \delta)$. If we denote by $V = U^{[p-1]}$ the upper left $(p-1) \times (p-1)$ subblock of the random matrix U, then*

$$\left(V, \frac{\det U}{\det V}, \frac{\det(I_p - U)}{\det(I_{p-1} - V)}\right) \stackrel{d}{=} J\beta E_{p-1}(\gamma, \delta) \otimes (p_{p-1,1}, (1 - p_{p-1,1})p_{p-1,2}),$$

where the random variables $p_{i,1} \sim \beta(\gamma - i/2, \delta)$ and $p_{i,2} \sim \beta(\delta - i/2, i/2)$ are beta-distributed and independent.

Proof: We consider the Cholesky decomposition of the matrix

$$U = \begin{pmatrix} V & B \\ B^t & c \end{pmatrix} = \begin{pmatrix} T^t & 0 \\ t^t & \sqrt{t_p} \end{pmatrix} \begin{pmatrix} T & t \\ 0 & \sqrt{t_p} \end{pmatrix} = \begin{pmatrix} T^t T & T^t t \\ t^t T & t^t t + t_p \end{pmatrix}, \tag{3.3}$$

where $V = U^{[p-1]}$ and T is an upper triangular $(p-1) \times (p-1)$ matrix with strictly positive entries on the diagonal. The matrix U satisfies almost surely the inequality $0_p < U < I_p$. As a matrix is positive definite if and only if its main subblock and the correpsonding Schur complement are positve definite, we obtain for the terms V, B and c in (3.3):

$$0_{p-1} < V \text{ and } c - B^t V^{-1} B > 0 \tag{3.4}$$
$$V < I_{p-1} \text{ and } 1 - c - B^t(I_{p-1} - V)^{-1}B > 0 \tag{3.5}$$

This implies $c < 1$ and we conclude $t_p = c - t^t t < 1$. Therefore the random variable

$$v = (1 - t_p)^{-1/2}(I_{p-1} + T(I_{p-1} - V)^{-1}T^t)^{1/2}t.$$

is well defined.

We will now determine the joint density of the random variables V, v and t_p (up to a constant). For this purpose note that the equation $c = t^t t + t_p$ yields the Schur complement of $I_{p-1} - V$ in $I_p - U$ as

$$\begin{aligned} 1 - c - B^t(I_{p-1} - V)^{-1}B &= 1 - t^t t - t_p - t^t T(I_{p-1} - V)^{-1}T^t t \\ &= (1 - t_p)(1 - t^t(I_{p-1} + T(I_{p-1} - V)^{-1}T^t)t(1 - t_p)^{-1}) \\ &= (1 - t_p)(1 - v^t v), \end{aligned}$$

5

which yields combined with the well-known formula for the determinant of a Schur-complement

$$\det(I_p - U) = \det(I_{p-1} - V) \cdot (1 - t_p)(1 - v^t v). \tag{3.6}$$

Similarly, observing (3.3)

$$\det(V)(c - B^t V^{-1} B) = \det U = (\det(T) \sqrt{t_p})^2 = \det(T^t T) t_p = \det(V) t_p, \tag{3.7}$$

which also implies

$$c - B^t V^{-1} B = t_p. \tag{3.8}$$

Using (3.4), (3.5), (3.6) and (3.7) it follows that the density of V, v and t_p is proportional to the function

$$g_{\gamma,\delta}(V) = (\det V)^{\gamma - e_1} t_p^{\gamma - e_1} (\det(I_{p-1} - V))^{\delta - e_1} ((1 - t_p)(1 - v^t v))^{\delta - e_1} \det D^{-1} \tag{3.9}$$
$$\times I\{0_{p-1} < V < I_{p-1}\} I\{0 < t_p < 1\} I\{v^t v < 1\},$$

where $\det D^{-1}$ is the Jacobi-determinant of the corresponding transformation from U to V, v and t_p. As this transformation leaves the matrix V unchanged we obtain $\det D^{-1} = \det M^{-1}$, where the matrix M is given by

$$M = \begin{pmatrix} \frac{\partial v}{\partial B} & \frac{\partial v}{\partial c} \\ \frac{\partial t_p}{\partial B} & \frac{\partial t_p}{\partial c} \end{pmatrix}$$

and

$$\frac{\partial t_p}{\partial c} = 1 \; , \; \frac{\partial t_p}{\partial B} = -2 B^t V^{-1}.$$

In order to calculate the remaining elements of the matrix M, we simplify the representation of v using the formula

$$I_{p-1} + T(I_{p-1} - V)^{-1} T^t = I_{p-1} + (V^{-1} - I_{p-1})^{-1} = (I_{p-1} - V)^{-1},$$

where the second equality stems from an application of the principal axis transform. From this equation, (3.3) and (3.8) we can rewrite the vector v as

$$v = (1 - c + B^t V^{-1} B)^{-1/2} (I_{p-1} - V)^{-1/2} (T^t)^{-1} B.$$

Standard calculus, observing this representation and (3.8) now gives

$$\frac{\partial v}{\partial c} = \frac{1}{2(1 - t_p)} v$$
$$\frac{\partial v}{\partial B} = -\frac{1}{1 - t_p} v B^t V^{-1} + (1 - t_p)^{-1/2} (I_{p-1} - V)^{-1/2} (T^t)^{-1} \; ,$$

and it follows

$$\det M = \det \left(\frac{\partial v}{\partial B} - \frac{\partial v}{\partial c} \cdot \frac{\partial t_p}{\partial B} \right) = (1 - t_p)^{-(p-1)/2} (\det(I_{p-1} - V))^{-(p-1)/2} (\det V)^{-1/2}.$$

6

From (3.9) we obtain that the joint density of V, v and t_p is proportional to the function

$$
\begin{aligned}
g_{\gamma,\delta}(V) &= (\det V)^{\gamma-e_1+1/2}(\det(I_{p-1}-V))^{\delta-e_1+1/2} \cdot t_p^{\gamma-e_1}(1-t_p)^{\delta-e_1+(p-1)/2} \cdot (1-v^t v)^{\delta-e_1} \\
&\quad \times I\{0_{p-1}<V<I_{p-1}\}I\{0<t_p<1\}I\{v^t v<1\} \\
&= (\det V)^{\gamma-p/2}(\det(I_{p-1}-V)^{\delta-p/2} \cdot t_p^{\gamma-(p-1)/2-1}(1-t_p)^{\delta-1} \cdot (1-v^t v)^{\delta-(p+1)/2} \\
&\quad \times I\{0_{p-1}<V<I_{p-1}\}I\{0<t_p<1\}I\{v^t v<1\}
\end{aligned}
$$

Therefore the random variables V, v and t_p are independent with $V \sim J\beta E_{p-1}(\gamma,\delta)$ and $t_p \sim \beta(\gamma-(p-1)/2,\delta)$. Obviously,

$$
\frac{\det U}{\det V} = t_p , \quad \frac{\det(I_p-U)}{\det(I_{p-1}-V)} = (1-t_p)(1-v^t v),
$$

and the assertion now follows if we can prove $1-v^t v \sim \beta(\delta-(p-1)/2,(p-1)/2)$, or equivalently $v^t v \sim \beta((p-1)/2,\delta-(p-1)/2)$.

To see this, we will apply Lemma 2.1 in Song and Gupta (1997). Since v has a density proportional to $g(\|v\|_2^2)$ with the function $g(x)=(1-x)^{\delta-(p+1)/2}I\{x<1\}$, the density of $\|v\|_2$ is proportional to

$$
x^{p-2}g(x^2)I\{0<x\} = x^{p-2}(1-x^2)^{\delta-(p+1)/2}I\{0<x<1\}
$$

Using a simple substitution, the density of $\|v\|_2^2$ is therefore proportional to

$$
x^{p/2-1}(1-x)^{\delta-(p+1)/2}I\{0<x<1\}x^{-1/2} = x^{(p-1)/2-1}(1-x)^{\delta-(p-1)/2-1}I\{0<x<1\},
$$

i.e. $v^t v = \|v\|_2^2 \sim \beta((p-1)/2,\delta-(p-1)/2)$. This concludes the proof of Theorem 3.1.
$\square$

Theorem 3.2. *Assume $M \sim J\beta E_p(\gamma,\delta)$, and let $M^{[j]}$ denote the $j \times j$ upper left subblock of the matrix M. There exist independent random variables $p_{0,1},\ldots,p_{p-1,1}$ and $p_{1,2},\ldots,p_{p-1,2}$ with*

$$
\begin{aligned}
p_{i,1} &\sim \beta(\gamma-i/2,\delta) , \quad i=0,\ldots,p-1, &\quad (3.10) \\
p_{i,2} &\sim \beta(\delta-i/2,i/2) , \quad i=1,\ldots,p-1, &\quad (3.11)
\end{aligned}
$$

such that for all real a,b the identity

$$
\begin{aligned}
&\log\left((\det M^{[j]})^a \cdot (\det(I_j-M^{[j]}))^b\right) \\
&= \log(p_{0,1}^a(1-p_{0,1})^b) + \sum_{i=1}^{j-1}\left\{\log(p_{i,1}^a(1-p_{i,1})^b) + \log(p_{i,2}^b)\right\} \quad (3.12)
\end{aligned}
$$

holds for all $1 \le j \le p$ simultaneously. In particular

$$
\det M = \prod_{i=0}^{p-1} p_{i,1}.
$$

Proof: For $1 \le i \le p-1$ choose

$$
p_{0,1} = \det M^{[1]} , \quad p_{i,1} = \frac{\det M^{[i+1]}}{\det M^{[i]}} , \quad p_{i,2} = \frac{\det(I_{i+1}-M^{[i+1]})}{(1-p_{i,1})\det(I_i-M^{[i]})} ,
$$

7

then the identity (3.12) obviously holds. The statements (3.10) and (3.11) are now proved by induction with respect to the parameter p. The claim is obviously correct for $p = 1$, since

$$p_{0,1} = \det M^{[1]} = M^{[1]} \sim J\beta E_1(\gamma, \delta) = \beta(\gamma, \delta) \ .$$

holds.

Now assume that (3.10) and (3.11) are satisfied for $1, \dots, p - 1$, then an application of Theorem 3.1 yields

$$M^{[p-1]} \sim J\beta E_{p-1}(\gamma, \delta),$$
$$p_{p-1,1} \sim \beta(\gamma - (p-1)/2, \delta),$$
$$p_{p-1,2} \sim \beta(\gamma - (p-1)/2, (p-1)/2),$$

and these three random variables are independent. Since $M^{[p-1]} \sim J\beta E_{p-1}(\gamma, \delta)$ and

$$(M^{[p]})^{[i]} = (M^{[p-1]})^{[i]} \quad (i \leq p - 1)$$

it follows from the induction hypothesis that

$$p_{i,1} \sim \beta(\gamma - i/2, \delta) \ , \ i = 0 \dots p - 2,$$
$$p_{i,2} \sim \beta(\delta - i/2, i/2) \ , \ i = 1 \dots p - 2,$$

and these random variables are independent. It remains to show that $(p_{i,1})_{i=0}^{p-2}$ and $(p_{i,2})_{i=1}^{p-2}$ are also (jointly) independent of $(p_{p-1,1}, p_{p-1,2})$. This follows directly from the fact that $(p_{i,1})_{i=0}^{p-2}$ and $(p_{i,2})_{i=1}^{p-2}$ can be written as functions of $M^{[p-1]}$, which is in turn independent of $(p_{p-1,1}, p_{p-1,2})$.

$\square$

Remark 3.3. A well-known result in random matrix theory is the Bartlett decomposition [see Bartlett (1933)], which states that in the Cholesky decomposition of a Wishart-distributed random matrix the entries are independent and normal resp. χ^2-distributed. A corresponding result for the Jacobi-beta-ensemble was derived by Kshirsagar (1961) and reads as follows. If $X \sim J\beta E_p(\gamma, \delta)$ has the (random) Cholesky decomposition $X = T^t T$ for some upper triangular matrix T, then the diagonal entries $t_{11}, \dots, t_{pp}$ of T are independent and their squares are beta-distributed, that is $t_{ii}^2 \sim \beta(\gamma - (i-1)/2, \delta)$. This result is a special case of Theorem 3.2 that can be obtained for $a = 1$, $b = 0$. To see this, denote by K_i the $p \times i$-matrix with $(K_i)_{j,k} = \delta_{jk}$. Then the equation

$$X^{[i]} = K_i^t X K_i = K_i^t T^t K_i K_i^t T K_i = (K_i^t T K_i)^t K_i^t T K_i$$

holds. Noting $\det(K_i^t T K_i) = t_{11} \cdot \dots \cdot t_{ii}$ we can conclude that

$$t_{ii}^2 = \frac{\det(K_i^t T K_i)^2}{\det(K_{i-1}^t T K_{i-1})^2} = \frac{\det(X^{[i]})}{\det(X^{[i-1]})} = p_{i-1,1}$$

are independent random variables with $t_{ii}^2 \sim \beta(\gamma - (i-1)/2, \delta)$.

4 Random Hankel determinant processes

Let $(M^p_{1,2n+1}, \ldots, M^p_{2n+1,2n+1})$ denote a uniformly distributed random vector on the $(2n+1)$-th moment space $\mathcal{M}^{p_n}_{2n+1}$ and recall the definition of the stochastic process $\{H_n(t)\}_{t\in[0,1]}$ in (1.3). From (2.2), (3.2) and Theorem 3.2 we obtain the following representation

$$H_n(t) = \sum_{i=1}^{\lfloor nt \rfloor} \Big\{ \sum_{j=1}^{p_n-1} \Big(\log(r^{\lfloor nt \rfloor - i}_{2n+1,2i,j}) + \log(r^{\lfloor nt \rfloor - i+1}_{2n+1,2i-1,j}) \Big)$$

$$+ \sum_{j=0}^{p_n-1} \log(p^{\lfloor nt \rfloor - i+1}_{2n+1,2i,j}(1 - p_{2n+1,2i,j})^{\lfloor nt \rfloor - i})$$

$$+ \sum_{j=0}^{p_n-1} \log(p^{\lfloor nt \rfloor - i+1}_{2n+1,2i-1,j}(1 - p_{2n+1,2i-1,j})^{\lfloor nt \rfloor - i+1}) \Big\} \, ,$$

where the random variables $p_{2n+1,i,j}$ and $r_{2n+1,i,j}$ are independent and beta-distributed, that is

$$p_{2n+1,i,j} \sim \beta\Big(\frac{p_n+1}{2}(2n-i+2) - j/2, \frac{p_n+1}{2}(2n-i+2)\Big) \tag{4.1}$$

$$r_{2n+1,i,j} \sim \beta\Big(\frac{p_n+1}{2}(2n-i+2) - j/2, j/2\Big). \tag{4.2}$$

In the following discussion, we will consider a more general process. When viewing the Hankel-determinant as a function of the canonical moment matrices $\mathcal{U}^{p_n}_{i,2n}$ [see equation (2.2)], we can not only vary the dimension n of the Hankel-matrix, but also the size p_n of the canonical moment matrices in (2.2). To this extent we introduce a new parameter $s \in (0,1]$ and consider the upper left $\lfloor p_n s \rfloor \times \lfloor p_n s \rfloor$ subblocks of the canonical moment matrices. According to Theorem 3.2, the distribution of the logarithm of the corresponding Hankel-determinant can be written as

$$H_n(s,t) = \sum_{i=1}^{\lfloor nt \rfloor} \Big\{ \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \Big(\log(r^{\lfloor nt \rfloor - i}_{2n+1,2i,j}) + \log(r^{\lfloor nt \rfloor - i+1}_{2n+1,2i-1,j}) \Big)$$

$$+ \sum_{j=0}^{\lfloor p_n s \rfloor - 1} \log(p^{\lfloor nt \rfloor - i+1}_{2n+1,2i,j}(1 - p_{2n+1,2i,j})^{\lfloor nt \rfloor - i}) \tag{4.3}$$

$$+ \sum_{j=0}^{\lfloor p_n s \rfloor - 1} \log(p^{\lfloor nt \rfloor - i+1}_{2n+1,2i-1,j}(1 - p_{2n+1,2i-1,j})^{\lfloor nt \rfloor - i+1}) \Big\} \, ,$$

where we use the convention $H_n(0,t) = 0$ and the process in (1.3) is obtained as $H_n(t) = H_n(1,t)$. In the following discussion we will investigate the weak convergence of the process $\{H_n(s,t)\}_{s,t\in[0,1]}$ as $n, p_n \to \infty$.

For this purpose we state a general result with sufficient conditions for the weak convergence of a process of the form (4.3), which might be of independent interest. The proof can be found in the Appendix A.1.

Theorem 4.1. *For each $n \in \mathbb{N}$ let T_n be a finite set, let $\{X_n(i) \mid i \in T_n\}$ be real valued random variables and let $g_n : T_n \times [0,1]^k \to \mathbb{R}$ be a real-valued function. Consider a process of the form*

$$Z_n(t_1, \ldots, t_k) = \sum_{i \in T_n} g_n(i, t_1, \ldots, t_k) X_n(i)$$

9

and suppose that the following assumptions are satisfied

(C1) *The random variables* $(X_n(i))_{i \in T_n}$ *are independent.*

(C2) g_n *is right-continuous in each of the last k components.*

(C3) $\mathbb{E}[X_n(i)] = 0.$

(C4) $\mathbb{E}\left[X_n^{2k+2}(i)\right] \leq C \mathbb{E}\left[X_n^2(i)\right]^{k+1} < \infty$ *for some universal constant $C > 1$.*

(C5) $\sup\limits_{i \in T_n} g_n^2(i, t) \operatorname{Var}(X_n(i)) \xrightarrow{n \to \infty} 0$ *for all fixed $t \in [0,1]^k$.*

(C6) *There exists a function $f : [0,1]^{2k} \to \mathbb{R}$ such that for all $s, t \in [0,1]^k$*

$$\operatorname{cov}(Z_n(s), Z_n(t)) = \sum_{i \in T_n} g_n(i,s)g_n(i,t) \operatorname{Var}(X_n(i)) \xrightarrow{n \to \infty} f(s,t).$$

(C7) *There are sequences $h_n^{(j)} \xrightarrow{n \to \infty} 0$ for $1 \leq j \leq k$ such that for any two vectors $T = (t_1, \ldots, t_k)$ and $T' = (t_1, \ldots, t_j', \ldots, t_k)$ the inequality*

$$\operatorname{Var}(Z_n(T) - Z_n(T')) \leq C(|t_j - t_j'| + h_n^{(j)})$$

holds.

(C8) *For any $t_j \leq t_j' \leq t_j'' \leq t_j + h_n^{(j)}$ at least one of the equations*

$$g_n(i, t_1, \ldots, t_j, \ldots, t_k) = g_n(i, t_1, \ldots, t_j', \ldots, t_k)$$

or

$$g_n(i, t_1, \ldots, t_j', \ldots, t_k) = g_n(i, t_1, \ldots, t_j'', \ldots, t_k)$$

holds.

Then the process Z_n converges weakly in $l^\infty([0,1]^k)$ to a centered, continuous Gaussian process with covariance kernel f.

Theorem 4.2. *If $p_n \xrightarrow{n \to \infty} \infty$, then*

$$\{\widetilde{H}_n(s,t)\}_{s,t \in [0,1]} := \frac{1}{\sqrt{n}}\{H_n(s,t) - \mathbb{E}[H_n(s,t)]\}_{s,t \in [0,1]} \implies \{\mathcal{G}(s,t)\}_{s,t \in [0,1]}$$

in $l^\infty([0,1]^2)$, where $\{H_n(s,t)\}_{s,t \in [0,1]}$ is the process defined in (4.3) and $\mathcal{G}$ is a centered continuous Gaussian process with covariance kernel

$$\operatorname{cov}(\mathcal{G}(s_1, t_1), \mathcal{G}(s_2, t_2)) = \frac{(s_1 \wedge s_2)^2 c(t_1, t_2)}{2}.$$

The function c is given by

$$c(t_1, t_2) = \int_0^{t_1 \wedge t_2} \frac{(t_1 - x)(t_2 - x)}{(1-x)^2} dx.$$

Proof: We use the decomposition

$$\widetilde{H}_n = A_n + B_n + C_n + D_n + E_n$$

of the process $\widetilde{H}_n$, where the processes on the right-hand side are defined by

$$A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} (\lfloor nt \rfloor - i) \left(\log(r_{2n+1,2i,j}) - \mathbb{E}\left[\log(r_{2n+1,2i,j}) \right] \right)$$

$$B_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} (\lfloor nt \rfloor - i + 1) \left(\log(r_{2n+1,2i-1,j}) - \mathbb{E}\left[\log(r_{2n+1,2i-1,j}) \right] \right)$$

$$C_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=0}^{\lfloor p_n s \rfloor - 1} (\lfloor nt \rfloor - i) \Big\{ \log(p_{2n+1,2i,j}(1 - p_{2n+1,2i,j}))$$

$$- \mathbb{E}\left[\log(p_{2n+1,2i,j}(1 - p_{2n+1,2i,j})) \right] \Big\}$$

$$D_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=0}^{\lfloor p_n s \rfloor - 1} (\lfloor nt \rfloor - i + 1) \Big\{ \log(p_{2n+1,2i-1,j}(1 - p_{2n+1,2i-1,j}))$$

$$- \mathbb{E}\left[\log(p_{2n+1,2i-1,j}(1 - p_{2n+1,2i-1,j})) \right] \Big\}$$

$$E_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=0}^{\lfloor p_n s \rfloor - 1} \left(\log(p_{2n+1,2i,j}) - \mathbb{E}\left[\log(p_{2n+1,2i,j}) \right] \right)$$

and the random variables $p_{2n+1,i,j}, r_{2n+1,i,j}$ are independent and beta-distributed, as specified in (4.1) and (4.2) respectively. We will now apply Theorem 4.1 to each of these processes to prove

$$A_n \implies \mathcal{G}' \tag{4.4}$$
$$B_n \implies \mathcal{G}' \tag{4.5}$$
$$C_n \implies 0 \tag{4.6}$$
$$D_n \implies 0 \tag{4.7}$$
$$E_n \implies 0, \tag{4.8}$$

where $\mathcal{G}'$ is a continuous centered Gaussian process with covariance kernel

$$\mathrm{cov}\left(\mathcal{G}'(s_1, t_1), \mathcal{G}'(s_2, t_2) \right) = \frac{(s_1 \wedge s_2)^2 c(t_1, t_2)}{4}.$$

Since A_n and B_n are independent it follows that $(A_n, B_n) \implies (\mathcal{G}', \mathcal{G}'')$, where $\mathcal{G}''$ is an independent copy of $\mathcal{G}'$ (c.f. Example 1.4.6 in van der Vaart and Wellner (1996)), and the continuous mapping theorem implies $A_n + B_n \implies \mathcal{G}' + \mathcal{G}'' \stackrel{d}{=} \mathcal{G}$. The assertion of Theorem 4.2 now follows from Slutsky's lemma. We will omit the proof of (4.5) and (4.7) because the arguments are similar as in the proof of (4.4) and (4.6), respectively.

Proof of (4.4): We can represent the process A_n in a form such that Theorem 4.1 is aplicable, that is

$$A_n(s, t) = \sum_{(i,j) \in T_n} g_n((i,j), s, t) X_n((i,j)) , \tag{4.9}$$

where

$$T_n = \{1, \ldots, n-1\} \times \{1, \ldots, p_n - 1\}$$

$$g_n((i,j), s, t) = \frac{1}{\sqrt{n}}(\lfloor nt \rfloor - i)I\{i \le \lfloor nt \rfloor - 1\}I\{j \le \lfloor p_n s \rfloor - 1\}$$

$$X_n((i,j)) = \log(r_{2n+1,2i,j}) - \mathbb{E}\left[\log(r_{2n+1,2i,j})\right] \tag{4.10}$$

It is obvious that A_n satisfies the conditions **(C1)**, **(C2)** and **(C3)**. Condition **(C4)** is proved in Theorem A.8 in Appendix A.2 (note that the parameters of the distribution of $r_{2n+1,i,j}$ are bounded from below by $\frac{1}{2}$).

By (A.9) from Appendix A.2 the variance of the logarithm of a beta distributed random variable $X \sim \beta(a, b)$ can be calculated as

$$\mathrm{Var}\,(\log X) = \psi_1(a) - \psi_1(a+b),$$

where

$$\psi_k(x) = \frac{d^{k+1}}{dx^{k+1}} \log \Gamma(x) \tag{4.11}$$

denotes the Polygamma function. An application of formula (A.23) from Appendix A.4 shows that

$$
\begin{aligned}
&g_n^2((i,j), s, t)\,\mathrm{Var}\,(X_n(i,j)) \\
&\le \frac{3(n-i)^2 j/2}{n\left(\frac{p_n+1}{2}(2n-2i+2) - j/2\right)\frac{p_n+1}{2}(2n-2i+2)} \\
&\le \frac{3(n-i)^2 j/2}{n\left((p_n+1)(n-i+1/2) + (p_n+1-j)/2\right)(p_n+1)(n-i+1)} \\
&\le \frac{3j/2}{n\,(p_n+1)^2} \le \frac{3}{n\,(p_n+1)} \xrightarrow{n\to\infty} 0\;,
\end{aligned}
$$

and therefore condition **(C5)** is also satisfied. To see that condition **(C6)** holds, define for $s_1, s_2, t_1, t_2 \in [0,1]$ the minima $s = s_1 \wedge s_2$ and $t = t_1 \wedge t_2$. By (A.25) we have a decomposition

$$\mathrm{cov}\,(A_n(s_1, t_1), A_n(s_2, t_2)) = S_n + R_n,$$

where

$$S_n = \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{1}{n} \frac{(\lfloor nt_1 \rfloor - i)(\lfloor nt_2 \rfloor - i)j/2}{\left(\frac{p_n+1}{2}(2n-2i+2) - j/2\right)\frac{p_n+1}{2}(2n-2i+2)},$$

and the remainder R_n satisfies the inequality

$$|R_n| \le \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{4}{n} \frac{(\lfloor nt_2 \rfloor - i)(\lfloor nt_2 \rfloor - i)}{\left(\frac{p_n+1}{2}(2n-2i+2) - j/2\right)^2}.$$

12

Now note that

$$S_n = \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{1}{n} \frac{(\lfloor nt_1 \rfloor - i)(\lfloor nt_2 \rfloor - i)j/2}{\left(\frac{p_n+1}{2}(2n - 2i + 2) - j/2\right)\frac{p_n+1}{2}(2n - 2i + 2)}$$

$$\geq \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{1}{n} \frac{(nt_1 - i - 1)(nt_2 - i - 1)j/2}{\left(\frac{p_n+1}{2}(2n - 2i + 2)\right)^2}$$

$$= \frac{1}{n} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{j}{2(p_n + 1)^2} \sum_{i=1}^{\lfloor nt \rfloor - 1} \frac{(nt_1 - i - 1)(nt_2 - i - 1)}{(n - i + 1)^2}$$

$$= \frac{1}{4}(s^2 + O(p_n^{-1}))\left(c(t_1, t_2) + O(n^{-1})\right) \xrightarrow{n \to \infty} \frac{(s_1 \wedge s_2)^2 c(t_1, t_2)}{4}.$$

Moreover,

$$S_n = \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{1}{n} \frac{(\lfloor nt_1 \rfloor - i)(\lfloor nt_2 \rfloor - i)j/2}{\left(\frac{p_n+1}{2}(2n - 2i + 2) - j/2\right)\frac{p_n+1}{2}(2n - 2i + 2)}$$

$$\leq \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{1}{n} \frac{(nt_1 - i)(nt_2 - i)j/2}{\left(\frac{p_n+1}{2}(2n - 2i)\right)^2}$$

$$= \left(\frac{s^2}{4} + O(p_n^{-1})\right)\left(c(t_1, t_2) + O(n^{-1})\right) \xrightarrow{n \to \infty} \frac{(s_1 \wedge s_2)^2 c(t_1, t_2)}{4},$$

which gives

$$S_n \xrightarrow{n \to \infty} \frac{(s_1 \wedge s_2)^2 c(t_1, t_2)}{4}.$$

Finally, the remainder R_n vanishes asymptotically, that is

$$|R_n| \leq \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{4}{n} \frac{(\lfloor nt_2 \rfloor - i)(\lfloor nt_2 \rfloor - i)}{\left(\frac{p_n+1}{2}(2n - 2i + 2) - j/2\right)^2}$$

$$\leq \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{4}{n} \frac{(n - i)^2}{((p_n + 1)(n - i + 1/2))^2} \leq \frac{4}{p_n} \xrightarrow{n \to \infty} 0,$$

which proves condition (C6). To show (C7) note that

$$\mathrm{Var}\left(\log r_{2n,2i,j}\right) \leq 5 \frac{j/2}{\left(\frac{p_n+1}{2}(2n - 2i + 2) - j/2\right)^2} \leq \frac{5j}{((p_n + 1)(n - i))^2} \qquad (4.12)$$

holds by the upper bound (A.23) in Appendix A.4. Assume now that $0 \leq t_1 < t_2 \leq 1$

and $0 \leq s \leq 1$ are real numbers. Then for $X = (s, t_1)$ and $Y = (s, t_2)$ we have

$$\mathrm{Var}\left(Z_n(Y) - Z_n(X)\right) \leq \sum_{i=1}^{\lfloor nt_1 \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{5}{n} (\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor)^2 \frac{j}{(p_n + 1)^2 (n - i)^2}$$

$$+ \sum_{i=\lfloor nt_1 \rfloor}^{\lfloor nt_2 \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{5}{n} (\lfloor nt_2 \rfloor - i)^2 \frac{j}{(p_n + 1)^2 (n - i)^2}$$

$$\leq \sum_{i=1}^{\lfloor nt_1 \rfloor - 1} \frac{5}{n} (\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor)^2 \frac{1}{(n - i)^2} + \sum_{i=\lfloor nt_1 \rfloor}^{\lfloor nt_2 \rfloor - 1} \frac{5}{n} (\lfloor nt_2 \rfloor - i)^2 \frac{1}{(n - i)^2}$$

$$\leq \frac{(\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor)^2}{n^2} \sum_{i=1}^{\lfloor nt_1 \rfloor - 1} \frac{5}{n} \frac{1}{(1 - i/n)^2} + \sum_{i=\lfloor nt_1 \rfloor}^{\lfloor nt_2 \rfloor - 1} \frac{5}{n}$$

$$\leq 10 \left(t_2 - t_1 + \frac{1}{n} \right).$$

For the increments in the second coordinate, let $0 \leq s_1 < s_2 \leq 1$ and $0 \leq t \leq 1$ be real numbers and set $X = (s_1, t)$, $Y = (s_2, t)$, then

$$\mathrm{Var}\left(Z_n(Y) - Z_n(X)\right) = \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=\lfloor p_n s_1 \rfloor}^{\lfloor p_n s_2 \rfloor - 1} \frac{1}{n} (\lfloor nt \rfloor - i)^2 \, \mathrm{Var}\left(\log r_{2n, 2i, j}\right)$$

$$\leq \sum_{i=1}^{\lfloor nt \rfloor - i} \sum_{j=\lfloor p_n s_1 \rfloor}^{\lfloor p_n s_2 \rfloor - 1} \frac{5}{n} \frac{j}{(p_n + 1)^2} \leq \frac{5}{p_n + 1} \sum_{j=\lfloor p_n s_1 \rfloor}^{\lfloor p_n s_2 \rfloor - 1} 1$$

$$\leq 5 \left(s_2 - s_1 + \frac{1}{p_n} \right).$$

Therefore condition **(C7)** is also satisfied with $h_n^{(1)} = \frac{1}{p_n}$ and $h_n^{(2)} = \frac{1}{n}$. It is obvious that **(C8)** holds for these sequences and the assertion (4.4) follows from Theorem 4.1.

Proof of (4.6): The process C_n can be decomposed in a similar way as A_n in (4.9), where the random variables $X_n((i, j))$ in (4.10) are now defined by

$$X_n((i, j)) = \log\left(p_{2n+1, 2i, j}(1 - p_{2n+1, 2i, j})\right) - \mathbb{E}\left[\log\left(p_{2n+1, 2i, j}(1 - p_{2n+1, 2i, j})\right)\right].$$

It is again obvious that **(C1)**, **(C2)** and **(C3)** hold. Condition **(C4)** is a consequence of Theorem A.9 in Appendix A.2.
For a proof of **(C5)** note that by formula (A.11) in Appendix A.2 the variance of $\log(X(1 - X))$ for a random variable $X \sim \beta(a, b)$ can be calculated as

$$\mathrm{Var}\left(\log(X(1 - X))\right) = \psi_1(a) + \psi_1(b) - 4\psi_1(a + b).$$

An application of (A.26) now gives

$$\operatorname{Var}\left(\log(p_{2n+1,2i,j}(1-p_{2n+1,2i,j}))\right) \le \left(6 + \frac{(j/2)^2}{\left(\frac{p_n+1}{2}(2n-2i+2)-\frac{j}{2}\right)}\right)\frac{1}{\left(\frac{p_n+1}{2}(2n-2i+2)\right)^2}$$

$$\le \left(6 + \frac{j^2}{(p_n+1)(n-i)}\right)\frac{1}{((p_n+1)(n-i))^2}$$

$$= \frac{6}{((p_n+1)(n-i))^2} + \frac{j^2}{((p_n+1)(n-i))^3}$$

$$\le \frac{6(j+1)}{((p_n+1)(n-i))^2}, \tag{4.13}$$

and condition **(C5)** follows from the inequality

$$\frac{1}{n}(\lfloor nt \rfloor - i)^2 \operatorname{Var}\left(\log(p_{2n+1,2i,j}(1-p_{2n+1,2i,j}))\right) \le \frac{1}{n}(\lfloor nt \rfloor - i)^2 \frac{6(j+1)}{((p_n+1)(n-i))^2} \le \frac{6}{np_n}.$$

For a proof of assumption **(C6)** note that

$$\operatorname{Var}\left(C_n(s,t)\right) \le \sum_{i=1}^{n-1}\sum_{j=0}^{p_n-1}\frac{1}{n}\left(6 + \frac{j^2}{((p_n+1)(n-i))^2}\right)\frac{1}{(p_n+1)^2}$$

$$\le \frac{6}{p_n+1} + \sum_{i=1}^{n-1}\sum_{j=0}^{p_n-1}\frac{1}{n}\frac{1}{(p_n+1)i}\frac{j^2}{(p_n+1)^2}$$

$$\le \frac{6}{p_n+1} + \frac{\log(n)+1}{n} \xrightarrow{n\to\infty} 0.$$

By the Cauchy-Schwarz-inequality this implies

$$\operatorname{cov}\left(C_n(s_1,t_1), C_n(s_2,t_2)\right) \xrightarrow{n\to\infty} 0.$$

Assumption **(C7)** follows by similar calculations as given for the proof of (4.4) (note the similarity between (4.12) and (4.13)). More specifically, the inequalities

$$\operatorname{Var}\left(C_n(s,t_1) - C_n(s,t_2)\right) \le 24\left(|t_2 - t_1| + \frac{1}{n}\right)$$

$$\operatorname{Var}\left(C_n(s_1,t) - C_n(s_2,t)\right) \le 12\left(|s_2 - s_1| + \frac{1}{p_n}\right)$$

hold for any real numbers $s,t \in [0,1]$, $0 \le s_1 < s_2 \le 1$ and $0 \le t_1 \le t_2 \le 1$. Finally condition **(C8)** is obvious with $h_n^{(1)} = \frac{1}{p_n}$, $h_n^{(2)} = \frac{1}{n}$ and the assertion (4.6) follows from Theorem 4.1.

Proof of (4.8): We use again Theorem 4.1 to prove the assertion. Conditions **(C1)** – **(C4)** hold by similar arguments as in the proof of A_n. Observing the inequality

$$\operatorname{Var}\left(\log(p_{2n+1,2i,j})\right) \le 5\frac{\frac{p_n+1}{2}(2n-2i+2)}{\left(\frac{p_n+1}{2}(2n-2i+2)-j/2\right)\left((p_n+1)(2n-2i+2)-j/2\right)}$$

$$\le 5\frac{(p_n+1)(n-i+1)}{((p_n+1)(n-i+1/2))\left((p_n+1)(2n-2i+1)\right)}$$

$$\le 5\frac{n-i+1}{(p_n+1)(n-i+1/2)^2} \le \frac{5}{(p_n+1)(n-i+1/2)} \tag{4.14}$$

15

(which follows from (A.23)) condition **(C5)** is obviously satisfied. To prove **(C6)**, note that it follows from (4.14)

$$\text{Var}\left(E_n(s,t)\right) \leq \frac{1}{n}\sum_{i=1}^{n}\sum_{j=0}^{p_n-1}\frac{5}{(p_n+1)(n-i+1/2)} = O\left(\frac{\log(n)}{n}\right) \xrightarrow{n\to\infty} 0.$$

Again, by the Cauchy-Schwarz-inequality, this implies

$$\text{cov}\left(E_n(s_1,t_1), E_n(s_2,t_2)\right) \xrightarrow{n\to\infty} 0.$$

For a proof of **(C7)** let $0 \leq s_1 < s_2 \leq 1$ and $0 \leq t \leq 1$, then we obtain for $X = (s_1, t)$ and $Y = (s_2, t)$ from (4.14) the estimate

$$\text{Var}\left(E_n(X) - E_n(Y)\right) \leq \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor}\sum_{j=\lfloor p_n s_1 \rfloor}^{\lfloor p_n s_2 \rfloor - 1}\frac{5}{(p_n+1)(n-i+1/2)} \leq 10\left(s_2 - s_1 + \frac{1}{p_n}\right),$$

and similarly for $0 \leq t_1 < t_2 \leq 1$, $0 \leq s \leq 1$, $X = (s, t_1)$ and $Y = (s, t_2)$

$$\text{Var}\left(E_n(X) - E_n(Y)\right) \leq \frac{1}{n}\sum_{i=\lfloor nt_1 \rfloor + 1}^{\lfloor nt_2 \rfloor}\sum_{j=0}^{\lfloor p_n s \rfloor - 1}\frac{5}{(p_n+1)(n-i+1/2)} \leq 10\left(t_2 - t_1 + \frac{1}{n}\right).$$

Finally assumption **(C8)** is obvious with $h_n^{(1)} = \frac{1}{p_n}$ and $h_n^{(2)} = \frac{1}{n}$. $\qquad\square$

Corollary 4.3. *If* $p_n \xrightarrow{n\to\infty} \infty$, *then*

$$\{\widetilde{H}_n(t)\}_{t\in[0,1]} := \frac{1}{\sqrt{n}}\{H_n(t) - \mathbb{E}[H_n(t)]\}_{t\in[0,1]} \implies \{\mathcal{G}(t)\}_{t\in[0,1]},$$

in $l^\infty([0,1])$ *where* $\{H_n(t)\}_{t\in[0,1]}$ *is the process defined in (1.3) and* $\{\mathcal{G}(t)\}_{t\in[0,1]}$ *is a centered continuous Gaussian process with covariance kernel*

$$\text{cov}\left(\mathcal{G}(t_1), \mathcal{G}(t_2)\right) = \frac{1}{2}\int_0^{t_1\wedge t_2}\frac{(t_1-x)(t_2-x)}{(1-x)^2}dx.$$

We conclude with a Glivenko-Cantelli type Theorem.

Theorem 4.4. *If* $p_n \xrightarrow{n\to\infty} \infty$, *we have for any* $\varepsilon > 0$

$$n^{-\varepsilon}\|\widetilde{H}_n\| \xrightarrow{a.s.} 0.$$

Proof: We will prove a more general result, namely that for a process Z_n satisfying assumptions **(C1)** – **(C8)** we have $a_n\|Z_n\|_\infty \xrightarrow{a.s.} 0$ for all sequences a_n that satisfy $a_n = O(n^{-(1+\delta)/(2k+2)})$ for some $\delta > 0$. In order to prove this, note that by (A.6) and (A.7) we can apply Theorem A.3 from Appendix A.1 to conclude that

$$\mathbb{P}\left(\sup_{s,k\in[0,1]^k}|Z_n(s) - Z_n(t)| > \lambda\right) \leq C''\lambda^{-(2k+2)}$$

16

holds for all $\lambda > 0$, where C'' is some constant independent of λ. Combined with **(C4)** and **(C6)** this yields

$$\sum_{n=1}^{\infty} \mathbb{P}\left(a_n \|Z_n\|_{\infty} > \lambda\right)$$

$$\leq \sum_{n=1}^{\infty} \mathbb{P}\left(\sup_{s,t \in [0,1]^k} |Z_n(s) - Z_n(t)| > a_n^{-1}\lambda/2\right) + \mathbb{P}\left(|Z_n(0)| > a_n^{-1}\lambda/2\right)$$

$$\leq \sum_{n=1}^{\infty} \left(C'' + C(f(0,0) + o(1))^{k+1}\right) a_n^{2k+2}(\lambda/2)^{-(2k+2)}$$

$$= \left(C'' + C(f(0,0) + o(1))^{k+1}\right)(\lambda/2)^{-(2k+2)} \sum_{n=1}^{\infty} O(n^{-(1+\delta)}) < \infty.$$

By the Borel-Cantelli Lemma, this implies $a_n \|Z_n\|_{\infty} \xrightarrow{a.s.} 0$.
To prove the statement of Theorem 4.4, set $k = \lceil \varepsilon^{-1} + 1 \rceil$, $\delta = 1$ and $a_n = n^{-1/(k+1)}$. We can now define a k-dimensional partial sum process $\widehat{H}_n$ via

$$\widehat{H}_n(s_1, \ldots, s_k) = \widetilde{H}_n(s_1, s_2).$$

Similar to the proof of Theorem 4.2, we can decompose $\widehat{H}_n$ into a sum of processes that satisfy the conditions **(C1)** – **(C8)** and from the previous result we can conclude

$$n^{-\varepsilon} \|\widetilde{H}_n\|_{\infty} \leq a_n \|\widehat{H}_n\|_{\infty} \xrightarrow{a.s.} 0.$$

$\square$

Corollary 4.5 (Law of large numbers). *If $p_n \xrightarrow{n \to \infty} \infty$, we have*

$$\frac{H_n(s,t)}{np_n} \xrightarrow[\text{uniformly}]{a.s.} -\frac{s^2}{2}(t + (1-t)\log(1-t))$$

for $n \to \infty$.

Proof: By Theorem 4.4 we know

$$\frac{\|H_n - \mathbb{E}[H_n]\|_{\infty}}{np_n} = p_n^{-1} n^{-1/2} \|\tilde{H}_n\|_{\infty} \xrightarrow{a.s.} 0$$

The assertion now follows from Lemma 5.1 and Lemma 5.2 of the following section, which yield

$$\frac{\mathbb{E}[H_n]}{np_n} \xrightarrow[\text{uniformly}]{n \to \infty} -\frac{s^2}{2} \int_0^t \frac{t-x}{1-x} \, dx = -\frac{s^2}{2}(t + (1-t)\log(1-t)).$$

$\square$

5 Mod-ϕ-convergence, moderate and large deviations

In this section we study further stochastic properties of the random variables $H_n(s,t)$ defined in (4.3). We are particularly interested in the recently introduced concept of mod-ϕ-convergence [see Feray et al. (2016)] and large deviation properties. The different limiting results for the sequence $a_n^{-1}\big(H_n(s,t) - \mathbb{E}\left[H_n(s,t)\right]\big)$ obtained in this section and in Section 4 are summarised in Figure 1.

To be precise consider the strip

$$S_{(a,b)} = \{z \in \mathbb{C} \mid a < \Re(z) < b\},$$

where $a < 0 < b$ and let ϕ be a non-constant infinitely divisible distribution with moment generating function $\exp(\eta(z))$. Let $(t_n)_{n\in\mathbb{N}}$ denote a real valued sequence and $(X_n)_{n\in\mathbb{N}}$ be a sequence of real-valued random variables with existing moment-generating functions on $S_{(a,b)}$ such that $t_n \to \infty$ and that for some non-vanishing analytic function ψ on $S_{(a,b)}$

$$\exp(-t_n\eta(z))\mathbb{E}\left[\exp(zX_n)\right] \to \psi(z)$$

holds locally uniform on $S_{(a,b)}$. Following Feray et al. (2016) the sequence $(X_n)_{n\in\mathbb{N}}$ is said to converge mod-ϕ on $S_{(a,b)}$ with speed $(t_n)_{n\in\mathbb{N}}$.

Mod-ϕ-convergence is a very strong mode of convergence that implies the asymptotic behaviour of $(X_n)_{n\in\mathbb{N}}$ at different scales. Most prominently, Berry-Esseen bounds and large deviation results can be derived from mod-ϕ-convergence. Particularly the large deviation results are stronger than the results that are usually obtained by a large deviation principle. The former gives an asymptotic equivalent for the probability $\mathbb{P}\left(X_n \geq t_n x\right)$, while the latter only yields the limiting behaviour for the logarithm of the probability.

The core idea behind mod-ϕ-convergence is that the distribution of X_n is close to the distribution of the sum of t_n i.i.d. ϕ-distributed random variables. The function $\psi(z)$ measures the error made in this approximation and yields further refinement in the asymptotic formulas.
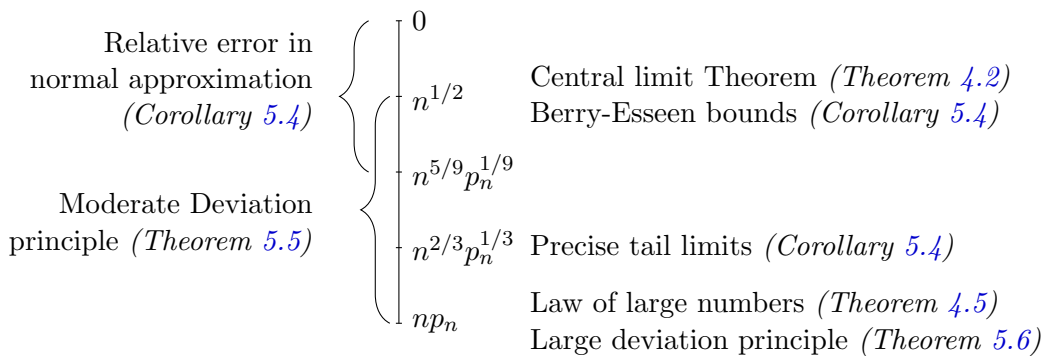

Figure 1: *Overview of all the limiting results for the sequence $a_n^{-1}\big(H_n(s,t) - \mathbb{E}\left[H_n(s,t)\right]\big)$, depending on the order of a_n.*

In the following we will first establish mod-ϕ-convergence of the sequence $(H_n(s,t))_{n\in\mathbb{N}}$ defined in (4.3). Large deviation principles are discussed in the second part of this Section. As the moment generating function is closely related to cumulants, we first provide estimates for these objects in Lemma 5.1 and Lemma 5.2, which are proved in Appendix A.3.

Lemma 5.1. *Let $r_{2n+1,i,j} \sim \beta\left(\frac{p_n+1}{2}(2n-i+2)-j/2, j/2\right)$ denote independent beta distributed random variables, then the cumulants of the random variables*

$$S_n = \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} (\lfloor nt \rfloor - i) \log(r_{2n+1,2i,j}),$$

$$S_n' = \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} (\lfloor nt \rfloor - i + 1) \log(r_{2n+1,2i-1,j})$$

satisfy the inequalities

$$\int_0^{t-2/n} \left(\frac{t-2/n-x}{1-x}\right)^m dx \le \frac{(-1)^m \kappa_m(S_n)}{n \frac{\lfloor p_n s - 1 \rfloor \lfloor p_n s \rfloor}{(p_n+1)^m} \frac{(m-1)!}{4}} \le \left(1 + \frac{m}{p_n}\right) \int_0^t \left(\frac{t-x}{1-x}\right)^m dx$$

$$\int_0^{t-2/n} \left(\frac{t-1/n-x}{1+1/n-x}\right)^m dx \le \frac{(-1)^m \kappa_m^n(S_n')}{n \frac{\lfloor p_n s - 1 \rfloor \lfloor p_n s \rfloor}{(p_n+1)^m} \frac{(m-1)!}{4}} \le \left(1 + \frac{m}{p_n}\right)\left(\int_0^t \left(\frac{t-x}{1-x}\right)^m dx + \frac{1}{n}\right)$$

for all $m \ge 1$. In particular

$$\max\{|\kappa_m(S_n)|, |\kappa_m(S_n')|\} \le 2 \cdot (m+1)! n p_n^{2-m}.$$

Lemma 5.2. *Let $p_{2n+1,i,j} \sim \beta\left(\frac{p_n+1}{2}(2n-i+2)-j/2, \frac{p_n+1}{2}(2n-i+2)\right)$ denote independent beta distributed random variables, then the cumulants of the random variables*

$$T_n = \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=0}^{\lfloor p_n s \rfloor - 1} \log(p_{2n+1,2i,j}^{\lfloor nt \rfloor - i + 1}(1 - p_{2n+1,2i,j})^{\lfloor nt \rfloor - i}),$$

$$T_n' = \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=0}^{\lfloor p_n s \rfloor - 1} \log(p_{2n+1,2i-1,j}^{\lfloor nt \rfloor - i + 1}(1 - p_{2n+1,2i-1,j})^{\lfloor nt \rfloor - i + 1})$$

satisfy the inequalities

$$\max\{|\kappa_m(T_n)|, |\kappa_m(T_n')|\} \le 6 \cdot 4^m (m+1)! p_n^{-m}\left(np_n + (\log(n)+1)p_n^2\right)$$
$$\le 12 \cdot 4^m (m+1)! n p_n^{2-m}$$

for all $m \ge 1$.

Theorem 5.3 (Mod-Gaussian-convergence)**.** *For any fixed $s, t \in (0, 1]$ the sequence*

$$\left(\frac{p_n}{n}\right)^{1/3} \left(H_n(s,t) - \mathbb{E}\left[H_n(s,t)\right]\right)$$

converges mod-Gaussian on any strip $S_{(a,b)}$ $(-\infty < a < b < \infty)$ with speed

$$t_n = \left(\frac{p_n}{n}\right)^{2/3} \kappa_2(H_n(s,t)) \sim n^{1/3} p_n^{2/3} \frac{s^2}{2} \int_0^t \left(\frac{t-x}{1-x}\right)^2 dx \to \infty \qquad (5.1)$$

and limiting function

$$\psi(z) = \exp\left(-z^3 \frac{s^2}{6} \int_0^t \left(\frac{t-x}{1-x}\right)^3 dx\right).$$

Proof: Recalling the definition of $H_n(s,t)$ in (4.3), it follows by an application of Lemma 5.1 and 5.2 that $t_n \to \infty$. More precisely

$$t_n = 2\frac{p_n^{2/3}}{n^{2/3}} n \frac{(\lfloor p_n s \rfloor - 1)\lfloor p_n s \rfloor}{(p_n + 1)^2} \frac{1}{4}\left(\int_0^t \left(\frac{t-x}{1-x}\right)^2 dx + o(1)\right) + \frac{p_n^{2/3}}{n^{2/3}} O(np_n^{-1} + \log(n))$$

$$\sim n^{1/3} p_n^{2/3} \frac{s^2}{2}\int_0^t \left(\frac{t-x}{1-x}\right)^2 dx \to \infty,$$

which proves (5.1). By the definition of the cumulant-generating function the equation

$$\log\left\{\exp(-t_n z^2/2)\mathbb{E}\left[\exp(zp_n^{1/3}n^{-1/3}(H_n(s,t)) - \mathbb{E}\left[H_n(s,t)\right])\right]\right\}$$

$$= \sum_{j=3}^{\infty} \frac{z^j}{j!}\kappa_j(p_n^{1/3}n^{-1/3}H_n(s,t))$$

$$= \frac{z^3}{6}p_n n^{-1}\kappa_3(H_n(s,t)) + \sum_{j=4}^{\infty} \frac{z^j}{j!}p_n^{j/3}n^{-j/3}\kappa_j(H_n(s,t)).$$

holds. As before we get by an application of Lemma 5.1 and 5.2

$$\frac{z^3}{6}p_n n^{-1}\kappa_3(H_n(s,t)) = -\frac{z^3}{3}\left(\int_0^t \left(\frac{t-x}{1-x}\right)^3 dx + o(1)\right)\frac{p_n\lfloor p_n s - 1\rfloor\lfloor p_n s\rfloor}{(p_n+1)^3}\frac{1}{2}$$

$$+ z^3 O(p_n n^{-1}p_n^{-3}(np_n + (\log(n) + 1)p_n^2))$$

$$= -\frac{z^3}{6}s^2\int_0^t \left(\frac{t-x}{1-x}\right)^3 dx + o(1)$$

The remainder converges locally uniform to 0, which follows from the inequality

$$\left|\sum_{j=4}^{\infty} \frac{z^j}{j!}p_n^{j/3}n^{-j/3}\kappa_j(H_n(s,t))\right| \leq 24\sum_{j=4}^{\infty} \frac{|z|^j}{j!}p_n^{j/3}n^{-j/3}(j+1)!4^j np_n^{2-j}$$

$$= 9216|z|^4 n^{-1/3}p_n^{-2/3}\sum_{j=0}^{\infty}(j+4)\left(\frac{4|z|}{n^{1/3}p_n^{2/3}}\right)^j.$$

$\square$

From the mod-ϕ-converge established in Theorem 5.3 we can derive several conclusions. Exemplarily we mention the following results, which are obtained by an application of Proposition 4.1.1, Theorem 4.3.1 and Theorem 4.2.1 in Feray et al. (2016)

Corollary 5.4.

(1) *Berry-Esseen bound*

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{H_n(s,t) - \mathbb{E}\left[H_n(s,t)\right]}{\sqrt{\mathrm{Var}\left(H_n(s,t)\right)}} \leq x\right) - \Phi(x)\right| = o\left(\frac{1}{\sqrt{t_n}}\right) = o\left(n^{-1/6}p_n^{-1/3}\right).$$

(2) *Relative error in normal approximation: For any sequence x_n satisfying $x_n = o(t_n^{1/6}) = o(n^{1/18}p_n^{1/9})$ we have*

$$\mathbb{P}\left(\frac{H_n(s,t) - \mathbb{E}\left[H_n(s,t)\right]}{\sqrt{\mathrm{Var}\left(H_n(s,t)\right)}} \leq x_n\right) = \Phi(x_n)(1 + o(1)).$$

(3) Precise tail limits: For $x > 0$ we have

$$\mathbb{P}\Big(H_n(s,t) - \mathbb{E}\left[H_n(s,t)\right] \geq \frac{n^{1/3}t_n}{p_n^{1/3}}x\Big)$$

$$= \frac{1}{x\sqrt{2\pi t_n}}\exp\Big(-t_n x^2/2 - x^3\frac{s^2}{6}\int_0^t \Big(\frac{t-y}{1-y}\Big)^3 dy\Big)(1+o(1)).$$

For $x < 0$ we have

$$\mathbb{P}\Big(H_n(s,t) - \mathbb{E}\left[H_n(s,t)\right] \leq \frac{n^{1/3}t_n}{p_n^{1/3}}x\Big)$$

$$= \frac{1}{-x\sqrt{2\pi t_n}}\exp\Big(-t_n x^2/2 - x^3\frac{s^2}{6}\int_0^t \Big(\frac{t-y}{1-y}\Big)^3 dy\Big)(1+o(1)).$$

In Part (3) of Corollary 5.4, the speed of the large deviation bound is $t_n \sim Cn^{1/3}p_n^{2/3}$ (c.f. formula (5.1)) and $H_n(s,t)$ is rescaled by $\frac{n^{1/3}t_n}{p_n^{1/3}} \sim Cn^{2/3}p_n^{1/3}$. This rescaling correponds to a moderate deviation principle. Using the Gärtner-Ellis theorem we can strengthen the large deviation principle to a speed of np_n^2 and a rescaling with np_n [see Theorem 5.6 below]. Before we do this, we will provide a more general version of a moderate deviation principle for the sequence $H_n(s,t)$. This result contains part (3) of Corollary 5.4 as a special case (using $a_n = n^{2/3}p_n^{1/3}$). However, the latter provides more precise information about the limit.

Theorem 5.5 (Moderate deviations). *Let $s,t \in (0,1]$ be fixed and let a_n be a sequence with $\sqrt{n} = o(a_n)$ and $a_n = o(np_n)$. Then*

$$X_n = \frac{H_n(s,t) - \mathbb{E}\left[H_n(s,t)\right]}{a_n}$$

satisfies a large deviation principle with speed $b_n = \frac{a_n^2}{n}$ and good rate function

$$I(x) = x^2\Big(s^2\int_0^t \Big(\frac{t-y}{1-y}\Big)^2 dy\Big)^{-1}$$

Proof: We will use the Gärtner-Ellis theorem [see Theorem 2.3.6 in Dembo and Zeitouni (2009) and the subsequent remark]. From Lemma 5.1 and Lemma 5.2 it follows

$$\frac{1}{b_n}\log\mathbb{E}\left[\exp(b_n x X_n/a_n)\right] = \sum_{j=2}^{\infty}\frac{x^j b_n^{j-1}\kappa_j(X_n)}{j!a_n^j} = \frac{x^2}{2n}\kappa_2(X_j) + O\Big(\sum_{j=3}^{\infty}\frac{x^j a_n^{j-2}}{j!}(j+1)!\frac{n}{p_n^{j-2}}\Big)$$

$$= \frac{x^2}{2n}\Big(\int_0^t \Big(\frac{t-y}{1-y}\Big)^2 dy + o(1)\Big)n\frac{(\lfloor p_n s\rfloor - 1)\lfloor p_n s\rfloor}{(p_n+1)^2}\frac{1}{2}$$

$$+ \frac{x^2}{2n}O\Big(\frac{n}{p_n} + (\log(n)+1)\Big) + O\Big(\sum_{j=3}^{\infty}x^j(j+1)\Big(\frac{a_n}{np_n}\Big)^{j-2}\Big)$$

$$= x^2\frac{s^2}{4}\int_0^t \Big(\frac{t-y}{1-y}\Big)^2 dy + o(1).$$

Since $I(x)$ is the Fenchel-Legendre transform of the limit on the right hand side of the last equation, the Gärtner-Ellis theorem yields the desired result.

$\square$

21

The large deviation principle in Theorem 5.5 is called a moderate deviation principle, because the scale at which the deviations occur is between the scaling of a central limit theorem and the scale of a law of large numbers. Indeed, at scale $\sqrt{n}$ the sequence $H_n(s,t) - \mathbb{E}[H_n(s,t)]$ is asymptotically Gaussian by Theorem 4.2 and at scale np_n the sequence $H_n(s,t)$ satisfies a law of large numbers by Theorem 4.5.

Theorem 5.6 (Large deviations). *The sequence $\frac{H_n(s,t)}{np_n}$ satisfies a large deviation principle with speed np_n^2 and good rate function*

$$\Lambda^*(x) = \sup_{\lambda > -\frac{1}{t}} \left(\lambda x + \frac{s^2}{2} \int_0^t \log\left(1 + \lambda \frac{t-y}{1-y} \right) dy \right).$$

In particular the sequence $\frac{H_n(s,1)}{np_n}$ satisfies a large deviation principle with speed np_n^2 and good rate function

$$\Lambda^*(x) = \begin{cases} -(x + s^2/2) + s^2 \log(s) + \frac{s^2}{2} \log(-2x) & x < 0 \\ \infty & x \geq 0 \end{cases}.$$

Proof: The Gärtner-Ellis theorem, Lemma 5.1 and Lemma 5.2 yield

$$\Lambda(x) = \limsup_{n \to \infty} \frac{1}{np_n^2} \log \mathbb{E}\left[\exp(p_n x H_n(s,t)) \right] = \limsup_{n \to \infty} \sum_{j=1}^{\infty} \frac{x^j}{j!} p_n^{j-2} n^{-1} \kappa_j(H_n(s,t))$$

$$= \limsup_{n \to \infty} \sum_{j=1}^{\infty} \left\{ \frac{1}{2} \frac{(-x)^j}{j} (s^2 + o(1)) \left(\int_0^t \left(\frac{t-y}{1-y} \right)^j dy + o(1) \right) \right.$$

$$\left. + O\left(4^j x^j (j+1) \left(\frac{1}{p_n} + \frac{\log(n)+1}{n} \right) \right) \right\}$$

$$= \frac{s^2}{2} \int_0^t \sum_{j=1}^{\infty} \left(-x \frac{t-y}{1-y} \right)^j \frac{1}{j} \, dy = -\frac{s^2}{2} \int_0^t \log\left(1 + x \frac{t-y}{1-y} \right) dy.$$

The remaining part follows by a straightforward calculation of

$$\Lambda^*(x) = \sup_{\lambda > -1} \left\{ \lambda x + \frac{s^2}{2} \log(1 + \lambda) \right\}.$$

$\square$

Remark 5.7. Note that the mod-Gaussian convergence provides moderate deviation principles while the application of the Gärtner-Ellis theorem yields the full large deviation principle. This is most likely due to the fact that we used a rather simple approximation when proving Theorem 5.3, based on upper bounds for the cumulants of order larger than three. In contrast to this, Theorem 5.6 uses all the cumulants in calculating the limiting function. This situation is similar to Example 2.1.3 in Feray et al. (2016), where a sequence is proven to converge mod-Poisson as well as mod-Gaussian, and the results obtained from the mod-Poisson convergence are much stronger than the ones obtained from mod-Gaussian convergence.

For $t = 1$ the limit of the cumulant generating function $\Lambda(x)$ in Theorem 5.6 is $\Lambda(x) = -\frac{s^2}{2} \log(1+x)$. This is the cumulant-generating function of a negative $\gamma(\frac{s^2}{2}, 1)$-distribution. A comparison with Theorem 4.2.1 from Feray et al. (2016) suggests the conjecture that $C \cdot (p_n + O(1)) H_n(s,t)$ converges mod-γ with a speed $t_n \sim C n p_n^2$, as this would imply the large deviation principle in Theorem 5.6.

References

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York, ninth dover printing, tenth gpo printing edition. 34, 38

Alzer, H. (2001). Mean-value inequalities for the polygamma functions. *aequationes mathematicae*, 61(1):151–161. 39

Arashi, M., Nagar, D. K., and Far, Z. F. (2011). Properties of the complex bimatrix variate beta distribution. *Linear Algebra and its Applications*, 434(9):2018–2029. 5

Bartlett, M. (1933). On the theory of statistical regression. *Proc. R. Soc. Edinburgh*, 53:260–283. 2, 8

Billingsley, P. (1971). *Weak Convergence of Measures.* Society for Industrial and Applied Mathematics. 27

Chang, F. C., Kemperman, J. H. B., and Studden, W. J. (1993). A normal limit theorem for moment sequences. *Ann. Probab.*, 21:1295–1309. 2

Damanik, D., Pushnitski, A., and Simon, B. (2008). The analytic theory of matrix orthogonal polynomials, arxiv:0711.2703v2. 1, 3

Dembo, A. and Zeitouni, O. (2009). *Large Deviations Techniques and Applications.* Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg. 21

Dette, H. and Nagel, J. (2012a). Distributions on unbounded moment spaces and random moment sequences. *The Annals of Probability*, 40(6):2690–2704. 2

Dette, H. and Nagel, J. (2012b). Matrix measures, random moments, and gaussian ensembles. *Journal of Theoretical Probability*, 25(1):25–49. 1, 2, 5

Dette, H. and Studden, W. J. (2002). Matrix measures, moment spaces and favard's theorem for the interval $[0, 1]$ and $[0, \infty)$. *Linear Algebra and its Applications*, 345(1–3):169 – 193. 3, 4

Dette, H. and Studden, W. J. (2005). A note on the maximization of matrix valued hankel determinants with applications. *Journal of Computational and Applied Mathematics*, 177(1):129 – 140. 4

Dette, H. and Tomecki, D. (2016). Hankel determinants of random moment sequences. *Journal of Theoretical Probability*, pages 1–26. 2

Duran, A. J. (1999). Ratio asymptotics for orthogonal matrix polynomials. *Journal of Approximation Theory*, 100:304–344. 1

Duran, A. J. and Lopez-Rodriguez, P. (1997). Density questions for the truncated matrix moment problem. *Can. J. Math.*, 49:708–721. 1

Duran, A. J. and van Assche, W. (1995). Orthogonal matrix polynomials and higher-order recurrence relations. *Lin. Alg. and Appl.*, 219:261–280. 1

Feray, V., Meliot, P., and Nikeghbali, A. (2016). *Mod-ϕ Convergence: Normality Zones and Precise Deviations*. SpringerBriefs in Probability and Mathematical Statistics. Springer International Publishing. 18, 20, 22

Gamboa, F. and Lozada-Chang, L. V. (2004). Large deviations for random power moment problem. *Ann. Probab.*, 32:2819–2837. 2

Gamboa, F., Nagel, J., and Rouault, A. (2016). Sum rules and large deviations for spectral matrix measures. Preprint available at https://arxiv.org/abs/1601.08135. 1

Gamboa, F., Nagel, J., Rouault, A., and Wagener, J. (2012). Large deviations for random matricial moment problems. *Journal of Multivariate Analysis*, 106:17–35. 1

Grünbaum, F. A. (2003). Matrix valued Jacobi polynomials. *Bull. Sciences Math.*, 127:207–214. 1

Grünbaum, F. A., Pacharoni, I., and Tirao, J. A. (2005). Matrix valued orthogonal polynomials of the Jacobi type: The role of group representation theory. *Ann. Inst. Fourier (Grenoble)*, 55:2051–2068. 1

Kallenberg, O. (2002). *Foundations of Modern Probability*. Probability and Its Applications. Springer New York. 25, 27

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. 1

Kshirsagar, A. M. (1961). The non-central multivariate beta distribution. *Ann. Math. Statist.*, 32(1):104–111. 2, 8

Lozada-Chang, L. V. (2005). Large deviations on moment spaces. *Electronic Journal of Probability*, 10:662–690. 2

Muirhead, R. (1982). *Aspects of Multivariate Statistical Theory*. Wiley Series in Probability and Statistics. Wiley. 4

Olkin, I. and Rubin, H. (1964). Multivariate beta distributions and independence properties of the wishart distribution. *Ann. Math. Stat.*, 35:261–269. 4

Rosenthal, H. P. (1970). On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables. *Israel Journal of Mathematics*, 8(3):273–303. 30

Shorack, G. and Wellner, J. (2009). *Empirical Processes with Applications to Statistics*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM). 27

Simon, B. (2005a). *Orthogonal Polynomials on the Unit Circle, Part 1: Classical theory*. AMS Colloquium publications. American Mathematical Society. 3

Simon, B. (2005b). *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory.* AMS Colloquium publications. American Mathematical Suciety. 3

Smith, P. J. (1995). A recursive formulation of the old problem of obtaining moments from cumulants and vice versa. *The American Statistician*, 49(2):217–218. 32

Song, D. and Gupta, A. K. (1997). L_p-norm uniform distribution. *Proc. Amer. Math. Soc.*, 125(2):595–601. 7

van der Vaart, A. W. and Wellner, J. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics.* Springer Series in Statistics. Springer. 11, 25, 32

A   Technical Details

A.1   Proof of Theorem 4.1

We begin with some general remarks regarding the assumptions made in Theorem 4.1. Conditions **(C1)**, **(C3)**, **(C5)** and **(C6)** are rather common and necessary to apply the Lindeberg central limit theorem. **(C4)** is used to prove the Lindeberg condition and in combination with **(C7)** it also allows us to bound the increments of the process via the Markov inequality. The latter condition also guarantees that a continuous version of the limiting process exists. Condition **(C8)** essentially means that the process Z_n is constant on all rectangular sets of the form

$$\prod_{j=1}^{k}[a_j h_n^{(j)}, (a_j + 1)h_n^{(j)}),$$

where a_j are integers satisfying $0 \leq a_j \leq \frac{1}{h_n^{(j)}}$.

The proof is split into several parts. First, we prove in Lemma A.1 that the finite-dimensional distributions of Z_n converge weakly to a centered Gaussian distribution with covariance kernel f. This also implies that f is nonnegative definite, and therefore we can conclude that a centered Gaussian process with covariance kernel f exists. We then show in Theorem A.2 that this process can be chosen to have continuous paths. Finally, we show in Theorem A.6 that the process Z_n is asymptotically tight. The assertion now follows from Theorem 1.5.4 in van der Vaart and Wellner (1996).

Lemma A.1 (Convergence of finite-dimensional distributions). *For any $t_1, \ldots, t_m \in [0,1]^k$ we have*

$$Z_n^{\star} = (Z_n(t_1), \ldots, Z_n(t_m)) \xrightarrow{d} \mathcal{N}(0, \Sigma),$$

where the elements of $\Sigma \in \mathbb{R}^{m \times m}$ are given by $\Sigma_{i,j} = f(t_i, t_j)$.

Proof: By the Cramér-Wold theorem it is sufficient to prove the weak convergence

$$c^t Z_n^{\star} \xrightarrow{d} \mathcal{N}(0, c^t \Sigma c)$$

for any arbitrary vector $c = (c_1, \ldots, c_m)^t \in \mathbb{R}^m$. For this purpose we use the Lindeberg central limit theorem in the the form of Theorem 5.12 of Kallenberg (2002). The first step is showing that

$$\mathrm{Var}\left(c^t Z_n^{\star}\right) \to c^t \Sigma c,$$

which is equivalent to proving $\text{cov}\left(Z_n(s), Z_n(t)\right) \to f(s, t)$ for all $s, t \in [0, 1]^k$. This follows directly from assumptions **(C1)** and **(C6)** and the following calculation

$$
\begin{aligned}
\text{cov}\left(Z_n(s), Z_n(t)\right) &= \sum_{i,j \in T_n} g_n(i, s) g_n(j, t) \, \text{cov}\left(X_n(i), X_n(j)\right) \\
&= \sum_{i \in T_n} g_n(i, s) g_n(i, t) \, \text{Var}\left(X_n(i)\right) \to f(s, t).
\end{aligned}
$$

To prove the Lindeberg-condition, observe the inequality

$$
\mathbb{E}\left[X^2 I\{|X| > \varepsilon\}\right] \leq \mathbb{E}\left[\frac{X^4}{\varepsilon^2} I\{|X| > \varepsilon\}\right] \leq \mathbb{E}\left[X^4\right] \varepsilon^{-2},
$$

and note that

$$
\mathbb{E}\left[X_n^4(i)\right] \leq \mathbb{E}\left[X_n^{2k+2}(i)\right]^{\frac{2}{k+1}} \leq C^{\frac{2}{k+1}} \mathbb{E}\left[X_n^2(i)\right]^2 \leq C \mathbb{E}\left[X_n^2(i)\right]^2
$$

by assumption **(C4)** and Jensen's inequality. Define $c^\star = \max\{|c_1|, \ldots, |c_m|\}$ and observe that

$$
c^t Z_n^\star = \sum_{j=1}^m c_j Z_n(t_j) = \sum_{i \in T_n} \left(\sum_{j=1}^m c_j g_n(i, t_j)\right) X_n(i).
$$

A further application of Jensen's inequality and **(C5)** yield

$$
\begin{aligned}
&\sum_{i \in T_n} \left(\sum_{j=1}^m c_j g_n(i, t_j)\right)^4 \mathbb{E}\left[X_n^4(i)\right] \leq \sum_{i \in T_n} m^3 \sum_{j=1}^m c_j^4 g_n^4(i, t_j) \mathbb{E}\left[X_n^4(i)\right] \\
&\leq C m^3 (c^\star)^4 \sum_{j=1}^m \left\{\left(\sup_{i \in T_n} g_n^2(i, t_j) \, \text{Var}\left(X_n(i)\right)\right)\left(\sum_{i \in T_n} g_n^2(i, t_j) \, \text{Var}\left(X_n(i)\right)\right)\right\} \\
&\leq C m^3 (c^\star)^4 \sum_{j=1}^m \left(\sup_{i \in T_n} g_n^2(i, t_j) \, \text{Var}\left(X_n(i)\right)\right)\left(f(t_j, t_j) + o(1)\right) \xrightarrow{n \to \infty} 0,
\end{aligned}
$$

which proves the Lindeberg-condition.

$\square$

Theorem A.2 (Continuity of the limit process). *There exists a continuous, centered Gaussian process with covariance kernel f.*

Proof: Let G be a centered Gaussian process with covariance function f. For arbitrary vectors $s = (s_1, \ldots, s_k), t = (t_1, \ldots, t_k) \in [0, 1]^k$ and $1 \leq i \leq k + 1$ define

$$
M_i = (s_1, \ldots, s_{i-1}, t_i, \ldots, t_k).
$$

From assumptions **(C6)** and **(C7)** we can conclude

$$
\text{Var}\left(G(M_i) - G(M_{i+1})\right) = \lim_{n \to \infty} \text{Var}\left(Z_n(M_i) - Z_n(M_{i+1})\right) \leq C|t_i - s_i|.
$$

This yields

$$\mathbb{E}\left[(G(t) - G(s))^{2k+2}\right] \leq k^{2k+1} \sum_{i=1}^{k} \mathbb{E}\left[(G(M_i) - G(M_{i+1}))^{2k+2}\right]$$

$$= k^{2k+1} \sum_{i=1}^{k} (2k+1)!! \operatorname{Var}\left(G(M_i) - G(M_{i+1})\right)^{k+1}$$

$$\leq k^{2k+1} \sum_{i=1}^{k} (2k+1)!!(C|t_i - s_i|)^{k+1}$$

$$\leq (Ck^2)^{k+1}(2k+1)!! \cdot \|t - s\|_{\infty}^{k+1},$$

and Theorem 3.23 in Kallenberg (2002) implies the existence of a continuous version of the process G.

$\square$

The proof of asymptotic tightness of Z_n requires some preparations. Typically, the asymptotic tightness of a one-dimensional random process $H_n \in l^{\infty}([0,1])$ is proven by showing a bound of the form

$$\mathbb{E}\left[(H_n(s) - H_n(t))^b\right] \leq C|t - s|^a,$$

where $a > 1$ and $b > 0$ are some parameters. Theorem 6 on p. 51 in Shorack and Wellner (2009) then yields the asymptotic tightness of H_n. However, since we are mostly interested in partial sum processes, such an inequality cannot hold. This is due to the discontinuity of partial sum processes at fixed points $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}$. We will therefore use a similar, but slightly more delicate argument generalizing Theorem 6.2 in Billingsley (1971) to more than one dimension. Informally speaking, we show that an increment $|Z(t) - Z(s)|$ is small with high probability (w.h.p.), if t and s are close to each other and only differ in one coordinate. From this we deduce that the increments of Z are "simultaneously" small in the sense that $\sup_{s,t} |Z(t) - Z(s)|$ is small w.h.p..

In order to achieve this, we use a chaining type argument and define a "dyadic lattice" on the cube $[0,1]^k$. Starting with the 2^k vertices $\{(i_1, \ldots, i_k) \mid i_1, \ldots, i_k \in \{0,1\}\}$ we subdivide the lattice in each step, to gain the lattice $\{(i_1, \ldots, i_k)/2^n \mid i_1, \ldots, i_k \in \{0, 1, \ldots, 2^n\}\}$ after n steps. Then an induction argument shows that the increments to the nearest neighbors *within* a lattice of length 2^{-n} are small w.h.p.. Using the assumptions on the increments of Z, we can find for all points s, t in the n-th lattice points S, T in the $(n-1)$-th lattice, so that $|Z(s) - Z(S)|$ and $|Z(t) - Z(T)|$ are small w.h.p.. Summing up all "small increments", we can see that all increments $|Z(s) - Z(t)|$ are "simultaneously small" w.h.p. for all dyadic rationals s, t. Using the right-continuity of Z this can be strengthened to hold for all real numbers s, t.

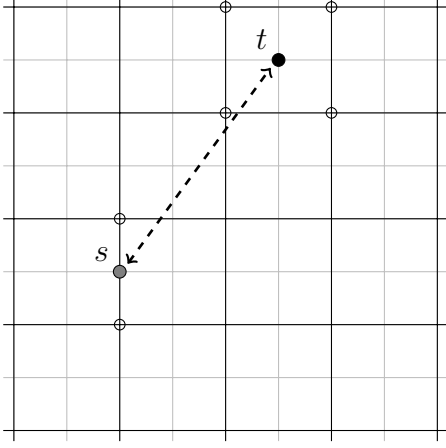This argument is visualized in the following two graphics for the two-dimensional $(k = 2)$ case.

Figure 2: *Points s and t on the n-th lattice with feasible points S, T on the $(n-1)$-th lattice marked.*
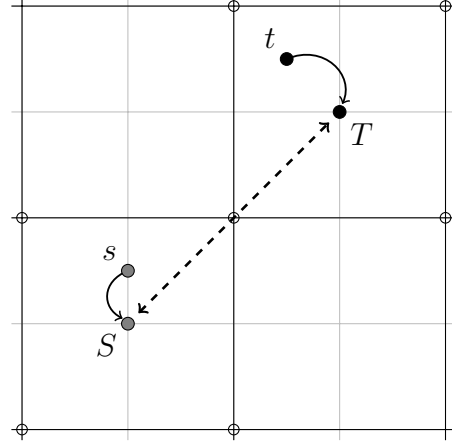

Figure 3: *Feasible points S and T were chosen. New feasible points for S and T on the $(n-2)$-th lattice are marked.*

While the assumptions in the following theorem may seem technical, they can typically be proven by a simple application of the Markov-inequality and estimates on the moments of Z.

Theorem A.3 (Global increments). *Let $Z : [0,1]^k \to \mathbb{R}$ be a process, which is right-continuous in every coordinate. Define for $1 \le i \le k$ and $s = (s_1, \ldots, s_k)$ the increment in the i-th coordinate by:*

$$m^i(s_1, \ldots, s_k, r, t) = \min\{|Z(s_1, \ldots, s_{i-1}, t, s_{i+1}, \ldots, s_k) - Z(s)|,$$
$$|Z(s_1, \ldots, s_{i-1}, r, s_{i+1}, \ldots, s_k) - Z(s)|\}.$$

Assume that there exist constants $\gamma > 0$, $\delta > k$ such that for all $1 \le i \le k$ and $r \le s_i \le t$ the inequality

$$\mathbb{P}\left(m^i(s_1, \ldots, s_k, r, t) > \lambda\right) \le C\lambda^{-\gamma}|t - r|^\delta \tag{A.1}$$

holds with a universal constant C. Further assume that there exists a function η such that the inequality

$$\mathbb{P}\left(|Z(s) - Z(t)| > \varepsilon\right) \le \eta(\|t - s\|_\infty, \varepsilon) \tag{A.2}$$

is satisfied. Then we have

$$\mathbb{P}\left(\sup_{s,t \in [0,1]^k} |Z(s) - Z(t)| > 4k\lambda\right) \le CC'\lambda^{-\gamma} + 4^k\eta(1, \lambda),$$

where C' is a universal constant that only depends on γ, δ and k.

Proof: Let $\theta_0, \theta_1, \theta_2, \ldots$ be arbitrary positive numbers and consider the event

$$M = \bigcap_{m=1}^\infty \bigcap_{i_1,\ldots,i_k=1}^{2^m-1} \bigcap_{j=1}^k \left\{m^j\left(\frac{i}{2^m}, \frac{i_j-1}{2^m}, \frac{i_j+1}{2^m}\right) \le \lambda\theta_m\right\} \bigcap \left\{\max_{s,t \in \{0,1\}^k} |Z(t) - Z(s)| \le \lambda\theta_0\right\}.$$

By assumptions (A.1) and (A.2) the complimentary event has a probability of at most

$$C\lambda^{-\gamma}k2^\delta \sum_{m=1}^\infty 2^{m(k-\delta)}\theta_m^{-\gamma} + 4^k\eta(1, \lambda\theta_0).$$

28

On M the following inequality

$$\left| Z\left(\frac{s}{2^m}\right) - Z\left(\frac{t}{2^m}\right) \right| \leq 2k \sum_{i=0}^{m} \lambda\theta_i \tag{A.3}$$

holds for all $s, t \in \{0, \ldots, 2^m\}^k$, which is obtained by an induction argument with respect to m. If the inequality holds for $m-1$ we successively choose $s'_1, \ldots, s'_k$ as follows. Assume we have already chosen $s'_1, \ldots, s'_j$ and set $S_j = \left(\frac{s'_1}{2^{m-1}}, \ldots, \frac{s'_j}{2^{m-1}}, \frac{s_{j+1}}{2^m}, \ldots, \frac{s_k}{2^m}\right)^t$. Then

$$s'_{j+1} = \begin{cases} \frac{s_{j+1}}{2} & \text{if } s_{j+1} \text{ is even} \\ \displaystyle\operatorname*{argmin}_{0 \leq p \leq 2^{m-1}} \left| Z\left(\frac{s_1}{2^{m-1}}, \ldots, \frac{s_{i-1}}{2^{m-1}}, \frac{p}{2^{m-1}}, \frac{s_{i+1}}{2^m}, \ldots, \frac{s_k}{2^m}\right) - Z(S_j)\right| & \text{if } s_{j+1} \text{ is odd} \end{cases} \cdot$$

The sequence S_j has the following three properties

1. $S_0 = \frac{s}{2^m}$.

2. $|Z(S_j) - Z(S_{j+1})| < \lambda\theta_m$ for $0 \leq j < k$.

3. $S_k \in \left\{(i_1, \ldots, i_k)/2^{m-1} \mid i_1, \ldots, i_k \in \{0, \ldots, 2^{m-1}\}\right\}$.

a similar constructions yield a sequence T_j from t_j, which proves (A.3):

$$\left| Z\left(\frac{s}{2^m}\right) - Z\left(\frac{t}{2^m}\right) \right| \leq |Z(S_k) - Z(T_k)| + \sum_{j=0}^{k-1} \left(|Z(S_j) - Z(S_{j+1})| + |Z(T_j) - Z(T_{j+1})|\right)$$

$$\leq \left(2k \sum_{i=0}^{m-1} \lambda\theta_i\right) + 2k\lambda\theta_m = 2k \sum_{i=0}^{m} \lambda\theta_i.$$

We now choose $\theta_i = (i+1)^{-2}$. Then the inequality

$$|Z(s) - Z(t)| \leq 2k\lambda \sum_{i=0}^{\infty} (i+1)^{-2} = k\lambda\frac{\pi^2}{3} \leq 4k\lambda$$

holds on M for all dyadic rational points $s, t \in [0,1]^k$. Since the paths of Z are right-continuous in every coordinate, the inequality is also satisfied for all real vectors $s, t \in [0,1]^k$. The theorem now follows by choosing

$$C' = k2^\delta \sum_{m=1}^{\infty} 2^{m(k-\delta)}\theta_m^{-\gamma} = k2^\delta \sum_{m=1}^{\infty} 2^{m(k-\delta)}(m+1)^{-\gamma}.$$

$\square$

Corollary A.4 (local increments). *Assume the process Z satisfies the assumptions of Theorem A.3. Then for all $\varepsilon > 0$ and $r \in [0,1]^k$*

$$\mathbb{P}\left(\sup_{\substack{\|s\|_\infty, \|t\|_\infty < \varepsilon \\ r+s, r+t \in [0,1]^k}} |Z(r+s) - Z(r+t)| > 4k\lambda \right) \leq \varepsilon^\delta C C' \lambda^{-\gamma} + 4^k \eta(\varepsilon, \lambda).$$

Proof: Consider the modified process $Z' : [0,1]^k \to \mathbb{R}, Z'(s) = Z((r + \varepsilon s) \wedge 1)$. This process satisfies the requirements of Theorem A.3, if η is replaced by $\eta(\epsilon \cdot, \cdot)$ and C is eplaced by $C\varepsilon^\delta$.

$\square$

Lemma A.5 (a Rosenthal-type inequality). *Let $X_1, \ldots, X_n$ be independent centered random variables and $p > 2$. Then the inequality*

$$\mathbb{E}\Big[\Big| \sum_{i=1}^{n} X_i \Big|^p \Big] \leq R(p) \Big(\sum_{i=1}^{n} \mathbb{E}\left[|X_i|^p\right]^{2/p} \Big)^{p/2}$$

holds, where $R(p)$ is a universal constant only depending on p.

Proof: By Jensen's inequality the inequalities

$$\sum_{i=1}^{n} \mathbb{E}\left[|X_i|^p\right] = \Big(n^{2/p} \Big(\sum_{i=1}^{n} \frac{1}{n} \mathbb{E}\left[|X_i|^p\right] \Big)^{2/p} \Big)^{p/2} \leq \Big(n^{2/p-1} \sum_{i=1}^{n} \mathbb{E}\left[|X_i|^p\right]^{2/p} \Big)^{p/2}$$

$$\leq \Big(\sum_{i=1}^{n} \mathbb{E}\left[|X_i|^p\right]^{2/p} \Big)^{p/2}$$

$$\Big(\sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] \Big)^{p/2} \leq \Big(\sum_{i=1}^{n} \mathbb{E}\left[|X_i|^p\right]^{2/p} \Big)^{p/2}$$

hold. The assertion now follows from Rosenthal's inequality [c.f. Theorem 3 in Rosenthal (1970)]:

$$\mathbb{E}\Big[\Big| \sum_{i=1}^{n} X_i \Big|^p \Big] \leq R(p) \max\Big\{ \sum_{i=1}^{n} \mathbb{E}\left[|X_i|^p\right], \Big(\sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] \Big)^{p/2} \Big\}.$$

$\square$

Theorem A.6. *The process Z_n is asymptotically tight.*

Proof: We will prove that Z_n satisfies the conditions of Theorem A.3 resp. Corollary A.4 with a function η that depends on n. Let $1 \leq j \leq k$ and $s_1, \ldots, s_k \in [0,1]$ be arbitrary. For $r \leq s_j \leq t$ define

$$R = (s_1, \ldots, r, \ldots, s_k)^t,$$
$$S = (s_1, \ldots, s_j, \ldots, s_k)^t,$$
$$T = (s_1, \ldots, t, \ldots, s_k)^t,$$

then an application of the Markov- resp. Hölder-inequality yields

$$\begin{aligned}
&\mathbb{P}\left(m^j(s_1, \ldots, s_k, r, t) > \lambda \right) \\
&\leq \mathbb{P}\left(|Z_n(T) - Z_n(S)||Z_n(S) - Z_n(R)| > \lambda^2 \right) \\
&\leq \lambda^{-2k-2} \mathbb{E}\left[|Z_n(T) - Z_n(S)|^{k+1} |Z_n(S) - Z_n(R)|^{k+1} \right] \\
&\leq \lambda^{-2k-2} \left(\mathbb{E}\left[|Z_n(T) - Z_n(S)|^{2k+2} \right] \mathbb{E}\left[|Z_n(S) - Z_n(R)|^{2k+2} \right] \right)^{\frac{1}{2}}.
\end{aligned} \tag{A.4}$$

By condition **(C4)**, **(C7)** and Lemma A.5 from Appendix A.1 it follows that

$$
\begin{aligned}
\mathbb{E}\left[|Z(T) - Z(S)|^{2k+2}\right] &\leq R(2k+2)\Big(\sum_{i \in T_n}(g_n(i,T) - g_n(i,S))^2 \mathbb{E}\left[|X_n(i)|^{2k+2}\right]^{2/(2k+2)}\Big)^{k+1} \\
&\leq R(2k+2)C^{2k+2}\Big(\sum_{i \in T_n}(g_n(i,T) - g_n(i,S))^2 \mathbb{E}\left[|X_n(i)|^2\right]\Big)^{k+1} \\
&= R(2k+2)C^{2k+2}\operatorname{Var}\left(Z_n(T) - Z_n(S)\right)^{k+1} \\
&\leq R(2k+2)C^{3k+3}\left(t - s_j + h_n^{(j)}\right)^{k+1},
\end{aligned} \tag{A.5}
$$

and a similar argument yields the inequality

$$
\mathbb{E}\left[|Z(S) - Z(R)|^{2k+2}\right] \leq R(2k+2)C^{3k+3}\left(s_j - r + h_n^{(j)}\right)^{k+1}.
$$

For $t - r \geq h_n^{(j)}$ this implies

$$
\begin{aligned}
\mathbb{E}\left[|Z(T) - Z(S)|^{2k+2}\right]\mathbb{E}\left[|Z(S) - Z(R)|^{2k+2}\right] &\leq c_{r,k}\left((t - s_j + h_n^{(j)})(s_j - r + h_n^{(j)})\right)^{k+1} \\
&\leq c_{r,k}\left(\frac{t - r + 2h_n^{(j)}}{2}\right)^{k+1} \\
&\leq c_{r,k}2^{k+1}\left(t - r\right)^{k+1},
\end{aligned}
$$

where $c_{r,k} = R(2k+2)^2 C^{6k+6}$ This inequality is also correct for $t - r < h_n^{(j)}$, since in this case $Z_n(T) = Z_n(S)$ or $Z_n(S) = Z_n(R)$ holds by **(C8)**. Plugging this inequality into (A.4) yields the estimate

$$
\mathbb{P}\left(m^j(s_1, \ldots, s_k, r, t) > \lambda\right) \leq \lambda^{-2k-2}R(2k+2)C^{3k+3}2^{k+1}(t-r)^{k+1}. \tag{A.6}
$$

Let $s, t \in [0,1]^k$ be arbitrary, set $S_i = (s_1, \ldots, s_{i-1}, t_i, \ldots, t_k)^t$ and note that $Z(t) - Z(s) = \sum_{i=1}^{k}(Z(S_i) - Z(S_{i+1}))$. From (A.5) we can conclude

$$
\begin{aligned}
&\mathbb{P}\left(|Z(t) - Z(s)| > \lambda\right) \\
&\leq \sum_{i=1}^{k}\mathbb{P}\left(|Z(S_i) - Z(S_{i+1})| > \lambda/k\right) \leq \sum_{i=1}^{k}k^{2k+2}\lambda^{-2k-2}\mathbb{E}\left[|Z(S_i) - Z(S_{i+1})|^{2k+2}\right] \\
&\leq \sum_{i=1}^{k}k^{2k+2}\lambda^{-2k-2}R(2k+2)C^{3k+3}(|t_i - s_i| + h_n^{(i)})^{k+1} \\
&\leq k^{2k+3}\lambda^{-2k-2}R(2k+2)C^{3k+3}(\|t - s\|_\infty + h_n)^{k+1}. \tag{A.7}
\end{aligned}
$$

where $h_n = \max_{i=1}^{k}h_n^{(i)}$. Let m be a positive integer and define for $j \in \{1, \ldots, m\}^k$ the set $K_j = \prod_{i=1}^{k}\left[\frac{j_i-1}{m}, \frac{j_i}{m}\right]$. The inequalities (A.6) and (A.7) allow us to apply Corollary A.4 with $\delta = k+1$, $\gamma = 2k+2$ and $\varepsilon = \frac{1}{m}$, which gives the inequality

$$
\begin{aligned}
\mathbb{P}\left(\sup_{s,t \in K_j}|Z(t) - Z(s)| > \lambda\right) &\leq D(m^{-k-1}\lambda^{-2k-2} + \lambda^{-2k-2}(m^{-1} + h_n)^{k+1}) \\
&\leq 2D\lambda^{-2k-2}(m^{-1} + h_n)^{k+1}
\end{aligned}
$$

for some constant D that depends only on k and C. This yields

$$\limsup_{n\to\infty}\mathbb{P}\Big(\sup_{j\in\{1,\ldots,m\}^k}\sup_{s,t\in K_j}|Z(t)-Z(s)|>\lambda\Big)$$

$$\leq\limsup_{n\to\infty}\sum_{j\in\{1,\ldots,m\}^k}\mathbb{P}\Big(\sup_{s,t\in K_j}|Z(t)-Z(s)|>\lambda\Big)$$

$$\leq\limsup_{n\to\infty}m^k 2D\lambda^{-2k-2}(m^{-1}+h_n)^{k+1}=\frac{2D\lambda^{-2k-2}}{m}\xrightarrow{m\to\infty}0.$$

Since the finite-dimensional distributions of $Z_n(t)$ converge weakly by Theorem A.1, Theorem 1.5.6 in van der Vaart and Wellner (1996) yields the asymptotic tightness of Z_n.

$\square$

A.2 Moments of logarithms of beta-distributed random variables

For the application of Theorem 4.1 we need precise estimates of the central moments of the log-beta distribution, which are given in this Section.

Lemma A.7. *Let $n\geq 2$ and Y be a random variable with finite n-th moment. Denote by μ_n resp. κ_n the n-th central moment resp. the n-th cumulant of Y. If the inequality*

$$|\kappa_m|\leq C\kappa_2^{m/2}$$

holds for some constant $C\geq 1$ and all $2\leq m\leq n$, then the inequality

$$|\mu_m|\leq (C+m)^m\mu_2^{m/2}$$

holds for all $0\leq m\leq n$.

Proof: We will show this theorem with an induction argument. For $n=0,1,2$ the inequality holds trivially. For $n\geq 3$ we obtain from the recursion

$$\mu_n=\kappa_n+\sum_{m=2}^{n-2}\binom{n-1}{m-1}\kappa_m\mu_{n-m}$$

[see for example Smith (1995)]

$$|\mu_m|\leq C\kappa_2^{m/2}+\sum_{i=2}^{m-2}\binom{m-1}{i-1}C\kappa_2^{i/2}(C+m-i)^{m-i}\kappa_2^{(m-i)/2}$$

$$=C\Big\{1+\sum_{i=2}^{m-2}\binom{m-1}{i}(C+i)^i\Big\}\kappa_2^{m/2}\leq Cv\Big\{\sum_{i=0}^{m-1}\binom{m-1}{i}(C+m-1)^i\Big\}\kappa_2^{m/2}$$

$$=C(C+m)^{m-1}\kappa_2^{n/2}\leq (C+m)^m\mu_2^{m/2}.$$

$\square$

Theorem A.8. *If $a, b \geq M > 0$, $X \sim \beta(a, b)$ and $Y = \log(X)$, then the inequality*

$$\left| \mathbb{E}\left[(Y - \mathbb{E}[Y])^n \right] \right| \leq \left(n! 2^{n/2} (M \wedge 1)^{-(n-1)/2} \left(1 + \frac{1}{M} \right) + n \right)^n \operatorname{Var}(Y)^{n/2}$$

holds.

Proof: In the following we will show that the cumulants κ_n of Y satisfy

$$|\kappa_m| \leq n! 2^{n/2} (M \wedge 1)^{-(n-1)/2} \left(1 + \frac{1}{M} \right) \kappa_2^{m/2} \tag{A.8}$$

for all $2 \leq m \leq n$. An application of Lemma A.7 then yields the desired result. For a proof of (A.8) we denote by

$$K(t) = \log \mathbb{E}\left[\exp(tY) \right] = \log \mathbb{E}\left[X^t \right] = \log \left(\frac{B(a + t, b)}{B(a, b)} \right) = \log \left(\frac{\Gamma(a + t)\Gamma(b)\Gamma(a + b)}{\Gamma(a + b + t)\Gamma(a)\Gamma(b)} \right)$$

the cumulant generating function of the random variable Y. For $m \geq 1$ the m-th derivative of K can be calculated as

$$K^{(m)}(t) = \psi_{m-1}(a + t) - \psi_{m-1}(a + b + t),$$

where $\psi_k(x) = \frac{d^{k+1}}{dx^{k+1}} \log \Gamma(x)$ denotes the polygamma function of order k. This yields

$$\kappa_m = K^{(m)}(0) = \psi_{m-1}(a) - \psi_{m-1}(a + b) \tag{A.9}$$

for $m \geq 1$. Applying formula (A.23) and (A.24) from Appendix A.4 yields

$$\begin{aligned}
\frac{|\kappa_m|}{|\kappa_2|^{m/2}} &\leq m! \min(a, b) a^{-m} \left(1 + a^{-1} \right) \left(\frac{a(a + b)}{b} \right)^{m/2} = m! \left(\frac{1}{a} + \frac{1}{b} \right)^{m/2} \min(a, b)(1 + a^{-1}) \\
&\leq m! 2^{m/2} \max(a^{-1}, b^{-1})^{m/2} \max(a^{-1}, b^{-1})^{-1} (1 + a^{-1}) \\
&\leq m! 2^{m/2} M^{-(m-1)/2} \left(1 + \frac{1}{M} \right) \leq n! 2^{n/2} (M \wedge 1)^{-(n-1)/2} \left(1 + \frac{1}{M} \right). \tag{A.10}
\end{aligned}$$

$\square$

Theorem A.9. *If $a, b \geq M > 0$, $X \sim \beta(a, b)$ and $Y = \log(X(1 - X))$, then there exists a constant $C_n(M)$ depending only on n and M such that the inequality*

$$\left| \mathbb{E}\left[(Y - \mathbb{E}[Y])^n \right] \right| \leq C_n(M) \operatorname{Var}(Y)^{n/2}$$

holds.

Proof: We will show that for $n \geq 2$ the quotient $|\kappa_n|/\kappa_2^{n/2}$ is bounded by a constant depending only on n and M. The assertion then follows from the same arguments as used in the proof of Theorem A.8. The only difference is that the bound has a more complex structure and we will therefore omit an explicit representation of $C_n(M)$.
The cumulant-generating function of Y is given by

$$K(t) = \log \mathbb{E}\left[X^t (1 - X)^t \right] = \log \left(\frac{B(a + t, b + t)}{B(a, b)} \right) = \log \left(\frac{\Gamma(a + t)\Gamma(b + t)\Gamma(a + b)}{\Gamma(a + b + 2t)\Gamma(a)\Gamma(b)} \right).$$

For $n \geq 1$ the n-th derivative of the cumulant-generating function can be written as

$$K^{(n)}(t) = \psi_{n-1}(a+t) + \psi_{n-1}(b+t) - 2^n \psi_{n-1}(a+b+2t),$$

and this yields the representation

$$\kappa_n = K^{(n)}(0) = \psi_{n-1}(a) + \psi_{n-1}(b) - 2^n \psi_{n-1}(a+b) \tag{A.11}$$

$$= \left\{ \psi_{n-1}(a) + \psi_{n-1}(b) - 2\psi_{n-1}\left(\frac{a+b}{2}\right) \right\} + \left\{ \psi_{n-1}\left(\frac{a+b}{2}\right) - \psi_{n-1}\left(\frac{a+b+1}{2}\right) \right\}$$

for the n-th cumulant of Y, where we have used formula (6.4.8) in Abramowitz and Stegun (1964). In the following, we will multiple times use the fact that $(-1)^n \psi_{n-1}$ is a nonnegative decreasing function, which is apparent from formula (A.21) in Appendix A.4. By the mean-value theorem there exists a $\xi \in \left(\frac{a+b}{2}, \frac{a+b+1}{2}\right)$ such that the inequality

$$\psi_1\left(\frac{a+b}{2}\right) - \psi_1\left(\frac{a+b+1}{2}\right) = -\frac{1}{2}\psi_2(\xi) \geq \frac{1}{2}\left|\psi_2\left(\frac{a+b+1}{2}\right)\right|$$

$$\geq 2(a+b+1)^{-2} \geq \frac{1}{\left(1+\frac{1}{M}\right)^2 (a+b)^2} \tag{A.12}$$

is satisfied, where the lower bound for $|\psi_2|$ follows from (A.22) in Appendix A.4. From (A.22), (A.23) and (A.27) we know

$$\psi_1(a) + \psi_1(b) - 2\psi_1\left(\frac{a+b}{2}\right) \geq \left(\sqrt{\psi_1(a)} - \sqrt{\psi_1(b)}\right)^2$$

$$= \left(\frac{\psi_1(a) - \psi_1(b)}{\sqrt{\psi_1(a)} + \sqrt{\psi_1(b)}}\right)^2 \geq \left(\frac{b-a}{ab(a^{-1/2} + b^{-1/2})}\right)^2 \frac{M}{M+1} \tag{A.13}$$

$$\geq \left(\frac{b-a}{ab\sqrt{2}(a^{-1} + b^{-1})^{1/2}}\right)^2 \frac{M}{M+1} = \frac{(b-a)^2}{2ab(a+b)} \frac{M}{M+1}.$$

Now (A.11) and (A.13) yield

$$\frac{|\kappa_n|}{\kappa_2^{n/2}} \leq \frac{\left|\psi_{n-1}(a) + \psi_{n-1}(b) - 2\psi_{n-1}\left(\frac{a+b}{2}\right)\right|}{\kappa_2^{n/2}} + \frac{\left|\psi_{n-1}\left(\frac{a+b}{2}\right) - \psi_{n-1}\left(\frac{a+b+1}{2}\right)\right|}{\left|\psi_1\left(\frac{a+b}{2}\right) - \psi_1\left(\frac{a+b+1}{2}\right)\right|^{n/2}}.$$

The formulas (A.9) and (A.10) show that the second term is bounded and it only remains to prove that the first term is bounded. Since the term is zero for $a = b$, we will assume $a \neq b$ in the following.

For this purpose let $0 < c < d$ be positive numbers and set $x = \frac{c+d}{2}$, $h = \frac{d-c}{2}$. Then $x + h = d$ and $x - h = c$. By the generalized mean-value theorem there are numbers $0 < \xi' < \xi < h$ such that

$$\frac{\psi_{n-1}(x+h) + \psi_{n-1}(x-h) - 2\psi_{n-1}(x)}{h^2} = \frac{\psi_n(x+\xi) - \psi_n(x-\xi)}{2\xi}$$

$$= \frac{\psi_{n+1}(x+\xi') + \psi_{n+1}(x-\xi')}{2}$$

holds. Note that $x + \xi \in \left(\frac{c+d}{2}, d\right)$ and $x - \xi \in \left(c, \frac{c+d}{2}\right)$. Applying this to $c = \min(a,b)$ and $d = \max(a,b)$ and using the monotonicity of $|\psi_{n+1}|$ yields

$$\left|\psi_{n-1}(a) + \psi_{n-1}(b) - 2\psi_{n-1}\left(\frac{a+b}{2}\right)\right| \leq |\psi_{n+1}(\min(a,b))|\frac{(b-a)^2}{4} \tag{A.14}$$

$$\leq (n+1)! \min(a,b)^{-(n+1)}\left(1 + \frac{1}{M}\right)(b-a)^2,$$

where the upper bound for $|\psi_{n+1}|$ stems from (A.22). We will now consider three separate cases.

Case 1: $\frac{1}{2} \leq \frac{\min(a,b)}{\max(a,b)}$ and $\frac{(ab(a+b))^{n/2}}{\min(a,b)^{n+1}} \leq |b-a|^{n-2}$. An application of (A.13) and (A.14) yields the desired result:

$$f_n := \frac{\left|\psi_{n-1}(a) + \psi_{n-1}(b) - 2\psi_{n-1}\left(\frac{a+b}{2}\right)\right|}{\kappa_2^{n/2}} \leq (n+1)!\left(1 + \frac{1}{M}\right)\frac{(b-a)^2}{\min(a,b)^{n+1}\left(\frac{(b-a)^2}{2ab(a+b)}\right)^{n/2}}$$

$$= (n+1)!\left(1 + \frac{1}{M}\right)\frac{(2ab(a+b))^{n/2}}{\min(a,b)^{n+1}|b-a|^{n-2}} \leq (n+1)!\left(1 + \frac{1}{M}\right)2^{n/2}.$$

Case 2: $\frac{1}{2} \leq \frac{\min(a,b)}{\max(a,b)}$ and $\frac{(ab(a+b))^{n/2}}{\min(a,b)^{n+1}} > |b-a|^{n-2}$. An application of (A.12) and (A.14) yields the desired result:

$$f_n \leq (n+1)!\left(1 + \frac{1}{M}\right)\frac{(b-a)^2}{\min(a,b)^{n+1}}\left(\left(1 + \frac{1}{M}\right)(a+b)\right)^n$$

$$\leq (n+1)!\left(1 + \frac{1}{M}\right)^{n+1}\left(\frac{(ab(a+b))^{n/2}}{\min(a,b)^{n+1}}\right)^{2/(n-2)}\min(a,b)^{-(n+1)}(a+b)^n$$

$$= (n+1)!\left(1 + \frac{1}{M}\right)^{n+1}\frac{a^{n/(n-2)}b^{n/(n-2)}(a+b)^{n+n/(n-2)}}{\min(a,b)^{n+1+2(n+1)/(n-2)}}$$

$$= (n+1)!\left(1 + \frac{1}{M}\right)^{n+1}\left(\frac{\max(a,b)}{\min(a,b)}\right)^{n/(n-2)}\left(1 + \frac{\max(a,b)}{\min(a,b)}\right)^{n+n/(n-2)}$$

$$\leq (n+1)!\left(1 + \frac{1}{M}\right)^{n+1}3^{n+2n/(n-2)}.$$

Case 3: $\frac{\min(a,b)}{\max(a,b)} < \frac{1}{2}$. The estimate (A.22) yields

$$\left|\psi_{n-1}(a) + \psi_{n-1}(b) - 2\psi_{n-1}\left(\frac{a+b}{2}\right)\right| \leq |\psi_{n-1}(a)| + |\psi_{n-1}(b)| + 2\left|\psi_{n-1}\left(\frac{a+b}{2}\right)\right|$$

$$\leq 4|\psi_{n-1}(\min(a,b))| \leq 4n!\min(a,b)^{-(n-1)}\left(1 + \frac{1}{M}\right).$$

Jointly with (A.13) this implies

$$f_n \leq 4n!\left(1 + \frac{1}{M}\right)^{1+n/2}\min(a,b)^{-(n-1)}\frac{(2ab(a+b))^{n/2}}{|b-a|^n}$$

$$\leq 2^{2+n}n!\left(1 + \frac{1}{M}\right)^{1+n/2}\frac{\max(a,b)^n\min(a,b)^{n/2}}{\max(a,b)^n2^{-n}\min(a,b)^{n-1}}$$

$$= 2^{2+2n}n!\left(1 + \frac{1}{M}\right)^{1+n/2}\min(a,b)^{1-n/2} \leq 2^{2+2n}n!\left(1 + \frac{1}{M}\right)^{1+n/2}M^{1-n/2}$$

$$\leq 2^{2+2n}n!\left(1 + \frac{1}{M}\right)^n,$$

where we used the inequality

$$|b-a| = \max(a,b) - \min(a,b) > \max(a,b) - \frac{1}{2}\max(a,b) = \frac{\max(a,b)}{2}$$

in the third line.

$\square$

A.3   Proofs of Lemma 5.1 and Lemma 5.2

Proof of Lemma 5.1 If $X \sim \beta(a,b)$, then we obtain from formula (A.9) in Appendix A.2

$$(-1)^m \kappa_m(\log(X)) = (-1)^{m+1}(\psi_{m-1}(a+b) - \psi_{m-1}(a)) = \int_a^{a+b} (-1)^{m+1} \psi_m(t)\, dt.$$

As $(-1)^{m+1}\psi_m(t) \geq 0$ (see formula (A.21) in Appendix A.4) it follows from (A.22) that

$$(-1)^m \kappa_m(\log(X)) \geq \int_a^{a+b} (m-1)! t^{-m}\, dt \geq (m-1)! b(a+b)^{-m}, \qquad \text{(A.15)}$$

$$(-1)^m \kappa_m(\log(X)) \leq \int_a^{a+b} (m-1)! t^{-m} \left(1 + \frac{m}{t}\right) dt \leq (m-1)! b a^{-m} \left(1 + \frac{m}{a}\right). \quad \text{(A.16)}$$

Applying (A.15) yields the lower bound

$$(-1)^m \kappa_m(S_n) \geq \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} (\lfloor nt \rfloor - i)^m \frac{j(m-1)!}{2(p_n+1)^m(n-i+1)^m}$$

$$\geq (m-1)! \frac{\lfloor p_n s - 1 \rfloor \lfloor p_n s \rfloor}{4(p_n+1)^m} \sum_{i=1}^{\lfloor nt \rfloor - 1} \frac{(nt-i-1)^m}{(n-i+1)^m}$$

$$= (m-1)! \frac{\lfloor p_n s - 1 \rfloor \lfloor p_n s \rfloor}{4(p_n+1)^m} \sum_{i=0}^{\lfloor nt \rfloor - 2} \frac{(t-2/n-i/n)^m}{(1-i/n)^m}$$

$$\geq n \frac{\lfloor p_n s - 1 \rfloor \lfloor p_n s \rfloor}{(p_n+1)^m} \frac{(m-1)!}{4} \int_0^{t-2/n} \left(\frac{t-2/n-x}{1-x}\right)^m dx \ ,$$

while the upper bound follows in a similar manner from (A.16), i.e.

$$(-1)^m \kappa_m(S_n) = (-1)^m \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} (\lfloor nt \rfloor - i)^m \kappa_m(\log(r_{2n+1,2i,j}))$$

$$\leq \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{(\lfloor nt \rfloor - i)^m j(m-1)!}{2((p_n+1)(n-i+1) - j/2)^m} \left(1 + \frac{m}{(p_n+1)(n-i+1) - j/2}\right)$$

$$\leq \sum_{i=1}^{\lfloor nt \rfloor - 1} \sum_{j=1}^{\lfloor p_n s \rfloor - 1} \frac{(nt-i)^m j(m-1)!}{2((p_n+1)^m(n-i)^m)} \left(1 + \frac{m}{(p_n+1)(n-i)}\right)$$

$$\leq \frac{\lfloor p_n s - 1 \rfloor \lfloor p_n s \rfloor}{(p_n+1)^m} \left(1 + \frac{m}{p_n}\right) \frac{(m-1)!}{4} \sum_{i=1}^{\lfloor nt \rfloor - 1} \frac{(t-i/n)^m}{(1-i/n)^m}$$

$$\leq n \frac{\lfloor p_n s - 1 \rfloor \lfloor p_n s \rfloor}{(p_n+1)^m} \left(1 + \frac{m}{p_n}\right) \frac{(m-1)!}{4} \int_0^t \left(\frac{t-x}{1-x}\right)^m dx.$$

The bounds for $\kappa_m(S_n')$ are proven in essentially the same way.

Proof of Lemma 5.2. For a beta distributed random variable $X \sim \beta(a, b)$ a similar calculation as given in (A.11) of Appendix A.2 shows

$$
\begin{aligned}
\kappa_m(\log(X^{d+1}(1-X)^d)) &= (d+1)^m \psi_{m-1}(a) + d^m \psi_{m-1}(b) - (2d+1)^m \psi_{m-1}(a+b) \\
&= d^m \big\{ \psi_{m-1}(a) + \psi_{m-1}(b) - 2^m \psi_{m-1}(a+b) \big\} \\
&\quad + ((d+1)^m - d^m)\psi_{m-1}(a) - ((2d+1)^m - (2d)^m)\psi_{m-1}(a+b) \\
&= d^m \big\{ \psi_{m-1}(a) + \psi_{m-1}(b) - 2^m \psi_{m-1}(a+b) \big\} \\
&\quad + \sum_{k=0}^{m-2} \binom{m}{k} d^k (\psi_{m-1}(a) - 2^k \psi_{m-1}(a+b)) \\
&\quad + md^{m-1}(\psi_{m-1}(a) - 2^{m-1}\psi_{m-1}(a+b)). \tag{A.17}
\end{aligned}
$$

Using the same calculations as in (A.11) we get

$$
\begin{aligned}
\psi_{m-1}(a) + \psi_{m-1}(b) - 2^m \psi_{m-1}(a+b) &= \Big\{ \psi_{m-1}(a) + \psi_{m-1}(b) - 2\psi_{m-1}\Big(\frac{a+b}{2}\Big) \Big\} \\
&\quad + \Big\{ \psi_{m-1}\Big(\frac{a+b}{2}\Big) - \psi_{m-1}\Big(\frac{a+b+1}{2}\Big) \Big\}
\end{aligned}
$$

Applying (A.14), the mean-value theorem and the estimate (A.22) from Appendix A.4 yields

$$
\begin{aligned}
|\psi_{m-1}(a) + \psi_{m-1}(b) - 2^m \psi_{m-1}(a+b)| &\leq \Big(1 + \frac{1}{\min(a,b)}\Big) \frac{(m+1)!(b-a)^2}{\min(a,b)^{(m+1)}} \\
&\quad + m!2^m (a+b)^{-m}\Big(1 + \frac{2}{a+b}\Big). \\
&\leq \frac{(m+1)!2^m}{\min(a,b)^m}\Big(1 + \frac{2}{\min(a,b)}\Big)\Big(\frac{(b-a)^2}{\min(a,b)} + 1\Big).
\end{aligned}
\tag{A.18}
$$

The second part of the sum can be approximated using (A.22)

$$
\begin{aligned}
\Big|\sum_{k=0}^{m-2} \binom{m}{k} d^k (\psi_{m-1}(a) - 2^k \psi_{m-1}(a+b))\Big| &\leq 2|\psi_{m-1}(a)| \sum_{k=0}^{m-2} \binom{m}{k}(2d)^k \\
&\leq 2|\psi_{m-1}(a)|(2d+1)^{m-2} \\
&\leq (m-1)!2^{m-1}\frac{(d+1)^{m-2}}{a^{m-1}}\Big(1 + \frac{1}{a}\Big).
\end{aligned}
\tag{A.19}
$$

For the last summand, we again use the same formula as in (A.11), together with the bound (A.10)

$$
\begin{aligned}
md^{m-1}|\psi_{m-1}(a) - 2^{m-1}\psi_{m-1}(a+b)| &\leq md^{m-1}\Big|\psi_{m-1}(a) - \psi_{m-1}\Big(\frac{a+b}{2}\Big)\Big| \\
&\quad + \frac{md^{m-1}}{2}\Big|\psi_{m-1}\Big(\frac{a+b}{2}\Big) - \psi_{m-1}\Big(\frac{a+b+1}{2}\Big)\Big| \\
&\leq md^{m-1}\frac{|b-a|+1}{2}|\psi_m(\min(a,b))| \\
&\leq (m+1)!\frac{|b-a|+1}{2}\frac{d^{m-1}}{\min(a,b)^m}\Big(1 + \frac{1}{\min(a,b)}\Big).
\end{aligned}
\tag{A.20}
$$

Finally, note that $\min\left\{\frac{p_n+1}{2}(2n - 2i + 2) - \frac{j}{2}, \frac{p_n+1}{2}(2n - 2i + 2)\right\} \geq 1$. Combining this inequality and plugging the inequalities (A.18), (A.19) and (A.20) into (A.17) shows the desired result

$$
\begin{aligned}
|\kappa_m(T_n)| &\leq \sum_{i=1}^{\lfloor nt\rfloor}\sum_{j=0}^{\lfloor p_ns\rfloor-1}\left\{\frac{3(m+1)!2^m(\lfloor nt\rfloor - i)^m}{((p_n+1)(n-i+1/2))^m}\cdot\left(\frac{j^2}{4(p_n+1)(n-i+1/2)}+1\right)\right. \\
&\quad + (m-1)!2^m\frac{(\lfloor nt\rfloor - i + 1)^{m-2}}{((p_n+1)(n-i+1/2))^{m-1}} \\
&\quad \left. + (m+1)!(j+1)\frac{(j+1)(\lfloor nt\rfloor - i)^{m-1}}{((p_n+1)(n-i+1/2))^m}\right\} \\
&\leq 3(m+1)!\sum_{i=1}^{n}\sum_{j=0}^{p_n-1}\left\{\frac{2^m}{p_n^m}\left(\frac{j^2}{(p_n+1)(2n-2i+1)}+1\right)+\frac{p_n^{1-m}4^m}{n-i+1/2}\right\} \\
&\leq 6\cdot 4^m(m+1)!\sum_{i=1}^{n}\sum_{j=0}^{p_n-1}\left\{p_n^{1-m}\frac{1}{n-i+1}+p_n^{-m}\right\} \\
&\leq 6\cdot 4^m(m+1)!p_n^{-m}(np_n+(\log(n)+1)p_n^2)\leq 12\cdot 4^m(m+1)!np_n^{2-m}.
\end{aligned}
$$

The inequality for $\kappa_m(T_n')$ is proven in essentially the same way. Since the exponents of $p_{2n+1,2i-1,j}$ and $1 - p_{2n+1,2i-1,j}$ in the definition of T_n' are the same, we do not need (A.17) and can apply (A.18) directly to the formula (A.11) of the cumulant. This makes the proof for T_n' much easier and the details are omitted for the sake of brevity.

A.4 On the Polygamma functions

Throughout this section a, b, t and z will be positive real numbers.

Recall the definition of the Polygamma function in (4.11) and note that formula 6.4.1 from Abramowitz and Stegun (1964) states

$$
\psi_k(z) = (-1)^{k+1}\int_0^\infty \frac{t^k e^{-tz}}{1-e^{-t}}dt. \tag{A.21}
$$

In particular, this implies that ψ_k is positive and decreasing for odd k and negative and increasing for even k. Observing the estimate

$$
|\psi_n(z)| = \int_0^\infty \frac{t^n e^{-zt}}{1-e^{-t}}dt \geq \int_0^\infty t^{n-1}e^{-zt}dt = \int_0^\infty s^{n-1}e^{-s}z^{-n}ds = (n-1)!z^{-n}
$$

we obtain the inequality

$$
(n-1)!z^{-n} \leq |\psi_n(z)| \leq (n-1)!z^{-n} + n!z^{-(n+1)} \leq n!z^{-n}\left(1+\frac{1}{z}\right), \tag{A.22}
$$

where we have used the inequality $\frac{1}{1-e^{-t}} \leq 1+\frac{1}{t}$ for the upper bound. From (A.22) we obtain the inequalities

$$
\frac{b}{a(a+b)} \leq \psi_1(a) - \psi_1(a+b) = \int_a^{a+b}|\psi_2(x)|dx \leq \left(1+\frac{2}{a}\right)\frac{b}{a(a+b)}. \tag{A.23}
$$

This method also yields upper bounds for differences of higher-order polygamma functions. Note that $|\psi_k|$ is decreasing, which yields together with (A.22)

$$|\psi_k(a) - \psi_k(a + b)| = \int_a^{a+b} |\psi_{k+1}(t)|dt \le b|\psi_{k+1}(a)| \le b(k+1)!a^{-(k+1)}\left(1 + \frac{1}{a}\right).$$

From the monotonicity of ψ_k and (A.22) we can furthermore deduce

$$|\psi_k(a) - \psi_k(a + b)| \le |\psi_k(a)| \le k!a^{-k}\left(1 + \frac{1}{a}\right).$$

Combining these inequalities proves

$$|\psi_k(a) - \psi_k(a + b)| \le (k+1)!\min(a, b)a^{-(k+1)}\left(1 + \frac{1}{a}\right). \tag{A.24}$$

Inequality (A.22) also yields

$$\left|\psi_1(a) - \psi_1(a + b) - \frac{b}{a(a+b)}\right| = \left|\psi_1(a) - \psi_1(a + b) + \left(\frac{1}{a} - \frac{1}{a+b}\right)\right| \le \frac{2}{a^2} + \frac{2}{(a+b)^2} \le \frac{4}{a^2} \tag{A.25}$$

and

$$\left|\psi_1(a) + \psi_1(b) - 4\psi_1(a + b) - \frac{(a-b)^2}{ab(a+b)}\right|$$
$$\le \left|\psi_1(a) - \frac{1}{a}\right| + \left|\psi_1(b) - \frac{1}{b}\right| + 4\left|\psi_1(a + b) - \frac{1}{a+b}\right|$$
$$\le \frac{2}{a^2} + \frac{2}{b^2} + \frac{8}{(a+b)^2} \le \frac{6}{(a \wedge b)^2},$$

where the latter implies

$$|\psi_1(a) + \psi_1(b) - 4\psi_1(a + b)| \le \left(6 + \frac{(a-b)^2}{(a \wedge b)}\right)\frac{1}{(a \wedge b)^2}. \tag{A.26}$$

Finally, $|\psi_n|$ is log-convex by formula (1.4) from Alzer (2001), i.e.

$$\left|\psi_n\left(\frac{a+b}{2}\right)\right| \le \sqrt{|\psi_n(a)\psi_n(b)|} \le \frac{|\psi_n(a)| + |\psi_n(b)|}{2}. \tag{A.27}$$