Electronic Journal of Probability

Vol. 10 (2005), Paper no. 19, pages 662-690.

Journal URL
http://www.math.washington.edu/∼ejpecp/

Large Deviations on Moment Spaces

Li-Vang Lozada-Chang

Departamento Teoría de Funciones
Facultad de Matemáticas y Computación
Universidad de la Habana
Cuba and Laboratoire de Statistique et Probabilités
Université Paul Sabatier, France.
Email: livang@matcom.uh.cu

Abstract

In this paper we study asymptotic behavior of some moment spaces. We consider two different settings. In the first one, we work with ordinary multi-dimensional moments on the standard m-simplex. In the second one, we deal with the trigonometric moments on the unit circle of the complex plane. We state large and moderate deviation principles for moment sequences based on uniform distributions. In both cases the rate function of the large deviation principle is related to the reversed Kullback information with respect to the uniform measure on the integration domain.

Keywords: Random Moment Problem, Large Deviations, Multidimensional Moment, Kullback information, reversed Kullback information

AMS 2000 Mathematics Subject Classification: Primary: 60F10, 44A60 Secondary: 42A70

Submitted to EJP on May 5, 2004. Final version accepted on August 27, 2004.

1. Introduction

In this paper we present a *random* viewpoint on moment theory. We study the asymptotic behavior of the projection of random moment vectors.

To illustrate the idea, consider the following framework involving the ordinary moments on $[0,1]$. Let $\mathcal{P}([0,1])$ be the set of all probability measures on $[0,1]$. For $\mu \in \mathcal{P}([0,1])$, the moments are defined as

$$m_n(\mu) = \int_0^1 t^n \, \mathrm{d}\mu(t), \ n \in \mathbb{N}^*. \tag{1.1}$$

Among others, the "classic" moment problems are recovering a measure μ from its given moments and, given a sequence of real numbers $(c_n : n \in \mathbb{N})$, determining the existence of a measure μ on $[0,1]$ verifying that $c_n = m_n(\mu)$ for all $n \in \mathbb{N}^*$. In other words, it consists in characterizing the sequences $(m_n(\mu) : n \in \mathbb{N}^*)$, $\mu \in \mathcal{P}([0,1])$.

The work developed here is inspired by [CKS93]. In order to get better insight on structure and shape of the *moment spaces*

$$M_n = \Big\{ \big(m_1(\mu), m_2(\mu), ..., m_n(\mu)\big) : \mu \in \mathcal{P}([0,1]) \Big\}, \ n \in \mathbb{N}^*,$$

the authors construct a process of random k-dimensional moment vectors and establish its convergence and a normal central limit. More precisely, let X_n be a random vector of M_n having uniform distribution and define X_n^k as its projection on M_k (k a fixed positive integer). They show that $(X_n^k : n \in \mathbb{N})$ converges to $\mathbf{c}_0 = (c_{01}, c_{02}, ..., c_{0k})$, the k-dimensional moment vector of the arc-sine law on $[0,1]$, i.e.

$$c_{0j} = \int_0^1 \frac{x^j}{\pi\sqrt{x(1-x)}} \, \mathrm{d}x, \quad j = 1, 2, ..., k.$$

Moreover, they obtain that the limit distribution of the fluctuations is normal. Namely,

$$\sqrt{n}(X_n^k - \mathbf{c}_0) \xrightarrow[n\to\infty]{Law} \mathcal{N}_k(0, \Sigma),$$

where

$$\Sigma_{ij} = c_{0,i+j} - c_{0i}c_{0j}, \ i, j = 1, 2, ..., k.$$

In the same framework the large and moderate deviations behavior are studied in [GLC04]. There, the main result is that, for any borelian set A of $[0,1]$,

$$\mathbb{P}\big(X_n^k \in A\big) \approx \exp\left(-n \inf\{I(x) : x \in A\}\right)$$

where I is a convex function related to the Kullback information with respect to arc-sine law (see Section 2.2 for the definition of this functional). Here, $\approx$ stands for large deviation equivalence (see Section 2.1 for a precise formulation). The large deviation principle (LDP) is also stated for the sequence of (random) upper representation probability measures of the random vectors X_n, $n \in \mathbb{N}^*$.

In this paper we present similar asymptotic behavior analysis of two different moment space settings. We consider multi-dimensional power moments on the m-dimensional standard simplex and trigonometric moments on the complex unit circle.

Let us begin introducing a general setting. Let $\mathcal{L}$ be an infinite set of index and let $\{\phi_t : t \in \mathcal{L}\}$ be a family of continuous complex valued functions defined on a bounded set $\mathbf{K} \subset \Omega$ (with $\Omega = \mathbb{C}$ or $\mathbb{R}^d$, $d \in \mathbb{N}^*$). Let $(\mathcal{L}_n : n \in \mathbb{N}^*)$ be an increasing sequence of finite subsets of $\mathcal{L}$ such that $\bigcup_n \mathcal{L}_n = \mathcal{L}$. For $n \in \mathbb{N}^*$, the n-th *moment space* is the set

$$M_n := \left\{ \int_{\mathbf{K}} \Phi_n \,\mathrm{d}\mu : \mu \in \mathcal{P}(\mathbf{K}) \right\} \subset \mathbb{C}^{\,|\mathcal{L}_n|}$$

where $\Phi_n := (\phi_t : t \in \mathcal{L}_n)$, $|A|$ stands for the cardinality of the set A and $\mathcal{P}(\mathbf{K})$ denotes the set of all probability measures with support included in $\mathbf{K}$. It is well know that M_n is the convex hull of the curve $\{\Phi_n(x) : x \in \mathbf{K}\}$ [KN77, Theorem I.3.5]. The set $\mathcal{P}(\mathbf{K})$ will be endowed with the weak topology [Bil99].

For $n, k \in \mathbb{N}^*$ such that $n \geq k$, let $\Pi_{n,k} : \mathbb{C}^{\,|\mathcal{L}_n|} \to \mathbb{C}^{\,|\mathcal{L}_k|}$ be the natural projection map. Let $(\widetilde{M}_n : n \in \mathbb{N})$ be a sequence of sets verifying

$$M_n \subset \widetilde{M}_n \subset \mathbb{C}^{\,|\mathcal{L}_n|}, \quad n \in \mathbb{N}^*, \tag{1.2a}$$

$$\Pi_{n,k}\widetilde{M}_n \subset \Pi_{n+1,k}\widetilde{M}_{n+1}, \quad n, k \in \mathbb{N}^*,\ k \leq n \tag{1.2b}$$

$$\bigcap_{n \geq k} \Pi_{n,k}\widetilde{M}_n = M_k \quad k \in \mathbb{N}^*. \tag{1.2c}$$

The sets $\widetilde{M}_n$ could be seen as *extended* moment spaces. By (1.2c), these sets provide good approximation of moment spaces. In general, is not possible to have a tractable description of the moment spaces. Nevertheless, by means of the approximating sets $\widetilde{M}_n$ (easily describable) we will be able to deal with the asymptotics of moment spaces.

Assuming that it is feasible, we endow $\widetilde{M}_n$ with the uniform probability measure U_n, i.e. with the corresponding normalized Lebesgue measure. Let k be a fixed integer, we focus on the asymptotic behavior of the first k-dimensional marginal probability of U_n. More precisely, let X_n be a random vector of $\widetilde{M}_n$ with distribution U_n. We aim to find convergence rates (in the sense of large deviations) for the sequence of random vectors

$$X_n^k := \Pi_{n,k} X_n.$$

The two different settings are described as follows. The first one concerns the m-dimensional power moments on $\mathcal{S}$ (with m positive integer). The standard m-simplex is the set

$$\mathcal{S} := \left\{ (x_1, x_2, ..., x_m) \in \mathbb{R}_+^m : \sum_{i=1}^{m} x_i \leq 1 \right\}.$$

For $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_m) \in \mathbb{N}^m$, define

$$\phi_{\boldsymbol{\beta}}(x) = x_1^{\beta_1} x_2^{\beta_2} \cdots x_m^{\beta_m}. \tag{1.3}$$

Let

$$\mathcal{L}_n := \left\{ (\beta_1, \beta_2, ..., \beta_m) \in \mathbb{N}^m : 0 < \sum_{i=1}^{n} \beta_i \leq n \right\}, \quad n \in \mathbb{N}$$

and

$$\Phi_n^{\mathcal{S}} = \Big(\phi_{\boldsymbol{\beta}} : \boldsymbol{\beta} \in \mathcal{L}_n \Big).$$

The n-th moment space (on $\mathcal{S}$) is

$$M_n^{\mathcal{S}} := \left\{ \int_{\mathcal{S}} \Phi_n^{\mathcal{S}} \, \mathrm{d}\mu : \mu \in \mathcal{P}(\mathcal{S}) \right\}.$$

The second setting involves the trigonometric moments on $\mathbb{T}$, the unit circle of $\mathbb{C}$. The moment spaces are

$$M_n^{\mathbb{T}} := \left\{ \left(\int_{\mathbb{T}} z^j \, \mathrm{d}\mu(z) : \, 1 \le j \le n \right) : \mu \in \mathcal{P}(\mathbb{T}) \right\}, \quad n \in \mathbb{N}.$$

Theorem 2.2 of [Gup99b] state that an infinite multi-sequence is a moment multi-sequence if and only if it is *completely* monotone. Taking for all $n \in \mathbb{N}^*$, $\widetilde{M_n}$ as the set of all *partially* monotone multi-sequence of order n, denoted by $\mathcal{G}_n$, the conditions (1.2) are fulfilled (see Definition (3.1) and Theorem (3.2) for details).

In [Gup00] a normal central limit for $\left(X_n^k : n \in \mathbb{N}\right)$ is shown. However, this result presents a mistake. Here, we formulate the correct proposition and the corresponding proof. The central role played by the arc-sine law in [CKS93] and [GLC04], is in this case played by $\nu^{\mathcal{S}}$, the uniform probability measure on $\mathcal{S}$. In Section 4, in the context of the trigonometric moment, we obtain the convergence of the random moment vector sequence (X_n^k) to the moment vector of $\nu^{\mathbb{T}}$, the uniform probability measure on $\mathbb{T}$, and its limit normal distribution (after normalization).

Our main results is the statement of the large deviation principle (LDP) for $(X_n^k : n \in \mathbb{N})$ in both settings. Using the symbol $\mathbf{K}$ to represent $\mathcal{S}$ or $\mathbb{T}$ our result can be written as follows. For sake of simplicity we drop the index $\mathbf{K}$ of X_n^k and $\mathcal{L}_n$.

$$\mathbb{P}\left(X_n^k \in A\right) \approx \exp\left(-|\mathcal{L}_n| I_{\mathbf{K}}(A)\right) \quad (A \text{ measurable set})$$

where

$$I_{\mathbf{K}}(A) = \inf_{\mu \in \mathcal{P}(\mathbf{K}, A)} \left\{ \int_{\mathbf{K}} \ln\left(\frac{\mathrm{d}\nu^{\mathbf{K}}}{\mathrm{d}\mu} \right) \mathrm{d}\nu^{\mathbf{K}} \right\}, \tag{1.4}$$

and

$$\mathcal{P}(\mathbf{K}, A) := \left\{ \mu \in \mathcal{P}(\mathbf{K}) : \int_{\mathbf{K}} \Phi_k^{\mathbf{K}} \, \mathrm{d}\mu \, \in A \right\}.$$

Throughout the paper we follow the convention that the infimum over an empty set is $+\infty$.

We complete these large deviation results by establishing a LDP in infinite dimension settings. In the multi-dimensional moment frame we state the LDP on the space of infinite real multi-sequence (under a certain bound constrain) equipped with a norm topology. In the trigonometric moment frame we state the LDP for random measures representing X_n ($n \in \mathbb{N}$). This last result, as in [GLC04], leads to obtain the useful expression for $I(.)$ as in (1.4).

The paper is organized as follows. To be self contained in next section we recall some definitions and basic results on large deviation and Kullback information. In Section 3 and 4 we state our results for the multi-dimensional and trigonometric moment setting respectively. Section 5 is devoted to comparing our result with the previous one in [GLC04]. The proofs of all results are deferred to the last section.

2. Large Deviations and Kullback Information

2.1. **Large Deviation Principle.** Let us first recall what a LDP is (see for example [DZ98]). Let (u_n) be a positive sequence of real numbers decreasing to 0.

Definition 2.1. We say that a sequence (R_n) of probability measures on a measurable Hausdorff space $(U, \mathcal{B}(U))$ satisfies a LDP with rate function I and speed (u_n) if:

i) I is lower semicontinuous, with values in $\mathbb{R}^+ \cup \{+\infty\}$.

ii) For any measurable set A of U:

$$-I(\operatorname{int} A) \leq \liminf_{n\to\infty} u_n \log R_n(A) \leq \limsup_{n\to\infty} u_n \log R_n(A) \leq -I(\operatorname{clo} A),$$

where $I(A) = \inf_{\xi\in A} I(\xi)$ and $\operatorname{int} A$ (resp. $\operatorname{clo} A$) is the interior (resp. the closure) of A.

We say that the rate function I is good if its level sets $\{x \in U : \; I(x) \leq a\}$ are compact for any $a \geq 0$. More generally, a sequence of U-valued random variables is said to satisfy a LDP if their distributions satisfy a LDP.

To be self-contained let us recall some facts and tools on large deviations which will be useful in the paper (we refer to [DZ98] for more on large deviations).

- **Contraction principle**. Assume that (R_n) satisfies a LDP on $(U, \mathcal{B}(U))$ with good rate function I and speed (u_n). Let T be a continuous mapping from U to another space V. Then, $(R_n \circ T^{-1})$ satisfies a LDP on $(V, \mathcal{B}(V))$ with good rate function
$$I'(y) = \inf_{x:T(x)=y} I(x), \quad y \in V,$$
and speed (u_n).
- **Exponential approximation**. Assume that U is a metric space and let d denotes the distance on U. Let (X_n) be a U-valued random sequence satisfying a LDP with good rate function I and speed (u_n). Let (Y_n) be another U-valued random sequence. If for any $\xi > 0$
$$\limsup_{n\to\infty} u_n \log \mathbb{P}(\, \mathrm{d}(X_n, Y_n) > \xi) = -\infty,$$
then (Y_n) satisfies the same LDP as (X_n).

Let (X_n) be a sequence of random variables on $(U, \mathcal{B}(U))$, and X a fixed point of U. If, for all sequence (v_n) decreasing to 0 such that $u_n/v_n \to 0$, we have a LDP for the sequence of random variables $(\sqrt{v_n/u_n}(X_n - X))$ with speed (v_n) we say that (X_n) satisfies a moderate deviation principle (MDP) [DZ98].

2.2. **Kullback and Reversed Kullback Information.** Let μ and ν be probabilities on certain measurable space U. The *Kullback information* or cross entropy of μ with respect to ν is defined by

$$\mathcal{K}(\mu|\nu) = \begin{cases} \displaystyle\int_U \ln \frac{\mathrm{d}\mu}{\mathrm{d}\nu}\, \mathrm{d}\mu, & \text{if } \mu \ll \nu \text{ and } \ln \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \in L^1(\mu) \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.1}$$

Properties of $\mathcal{K}$ as a function of μ may be found in [Bre79]. $\mathcal{K}(\cdot\,|\nu)$ is the rate function for Sanov large deviations theorem (see Section 6.2 of [DZ98]) where ν is a probability generating the random variables. In this paper, the rate function involved is the Kullback information with μ as a fixed measure and ν as variable measure (See (2.1)). That is why, we refers to it as the *reversed* Kullback information. The role of μ will be played by $\nu^{\mathcal{S}}$ (resp. $\nu^{\mathbb{T}}$) for the multi-dimensional (resp. trigonometric) moment setting.

In order to identify the rate function that controls the LDP in the multi-dimensional moment framework (Lemma 6.4) we exploit an optimization result on measure space [BL93, Theorem 3.4]. Namely, this result establish the duality relation between an optimization problem on measure space and the corresponding dual optimization problem on continuous function space. For the Kullback information (defined for general Borel measures), this relation can be paraphrased as follows. For $x \in \mathbb{R}^{|\mathcal{L}_k|}$,

$$\inf\left\{\mathcal{K}(\nu|\mu) + \mu(U) \,:\, \mu \text{ regular Borel finite measure}, \int_U \Phi_k \,\mathrm{d}\mu = x\right\}$$
$$= \sup\left\{\langle \Lambda, x\rangle + \int_U \ln(1 - \langle \Phi_k, \Lambda\rangle)\,\mathrm{d}\nu \,:\, \Lambda \in \mathbb{R}^{|\mathcal{L}_k|}\right\}$$

where $\langle\cdot,\cdot\rangle$ denotes the usual scalar product in $\mathbb{R}^{|\mathcal{L}_k|}$. We refer the reader to [BL93] for general statement and related results.

3. Multidimensional moment problem

We recall the *multi-dimensional moment space* of order n

$$M_n^{Simplexm} := \left\{\left(c_{\boldsymbol{\beta}} = \int_{\mathcal{S}} x^{\boldsymbol{\beta}}\,\mathrm{d}\mu(x) : \boldsymbol{\beta} \in \mathcal{L}_n\right) : \mu \in \mathcal{P}(\mathcal{S})\right\}$$

where $x^{\boldsymbol{\beta}} := x_1^{\beta_1} x_2^{\beta_2} ... x_m^{\beta_m}$ (with $0^0 = 1$) for $x = (x_1, x_2, ..., x_m) \in \mathbb{R}^m$ and $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_m) \in \mathbb{N}^m$. For $\boldsymbol{\beta} \in \mathbb{N}^m$, let $|\boldsymbol{\beta}| := \sum_{i=1}^m \beta_i$.

Definition 3.1. A multi-sequence $(c_{\boldsymbol{\beta}} : |\boldsymbol{\beta}| \geq 1)$ is called a *completely monotone multi-sequence* if for all $\beta_0 \in \mathbb{N}$ and $\boldsymbol{\beta} \in \mathbb{N}^m$

$$(-1)^{\beta_0}\Delta^{\beta_0} c_{\boldsymbol{\beta}} \geq 0 \tag{3.1}$$

where, with $c_{(0,0,...,0)} = 1$,

$$\begin{aligned}\Delta c_{\boldsymbol{\beta}} = {} & c_{(\beta_1+1,\beta_2,...,\beta_m)} + c_{(\beta_1,\beta_2+1,...,\beta_m)} + ... \\ & + c_{(\beta_1,\beta_2,...,\beta_m+1)} - c_{\boldsymbol{\beta}}\end{aligned}$$

(Δ^{β_0} stands for β_0 iterates of the operator Δ). A finite multi-sequence $(c_{\boldsymbol{\beta}} : \boldsymbol{\beta} \in \mathcal{L}_n)$ is called a *partially monotone multi-sequence of order* n if for all $\beta_0 \in \mathbb{N}$ and $\boldsymbol{\beta} \in \mathbb{N}^m$ such that $\beta_0 \leq n - |\boldsymbol{\beta}|$ (3.1) holds. Following [Gup00] we will call the completely (resp. partially) monotone multi-sequences as G-sequences (resp. partial G-sequences).

The G-sequences are the multi-dimensional version of the completely monotone sequences on $\mathbb{R}$. These sequences are the solution of the Hausdorff moment problem, i.e. the power moment problem on $[0, 1]$ [ST43]. The following theorem, due to [Gup99a], solves the multi-dimensional moment problem on $\mathcal{S}$.

Let $\mathcal{R}_\infty$ be the set of all infinite real multi-sequences. We recall that

$$\mathcal{L} = \bigcup_n \mathcal{L}_n = \mathbb{N}^m \setminus \{(0,0,...,0)\}.$$

Theorem 3.2 (Theorem 2.2 [Gup99a]). *Let $\mathbf{c} = (c_{\boldsymbol{\beta}} : \boldsymbol{\beta} \in \mathcal{L} \in \mathcal{R}_\infty$. There exists $\mu \in \mathcal{P}(\mathcal{S})$ such that*

$$c_{\boldsymbol{\beta}} = \int_{\mathcal{S}} x^{\boldsymbol{\beta}} \,\mathrm{d}\mu(x), \quad \forall \boldsymbol{\beta} \in \mathcal{L}.$$

if and only if $\mathbf{c}$ is a G-sequence.

We denote by $\mathcal{R}_n$ the set of all multi-sequence of order n, i.e. $\mathcal{R}_n = \mathbb{R}^{|\mathcal{L}_n|}$. For $n,k \in \mathbb{N}^*$ with $k < n$, let $\Pi^{\mathcal{S}}_{n,k} : \mathcal{R}_n \to \mathcal{R}_k$ $(n \geq k)$ be the projection map defined as

$$(c_{\boldsymbol{\beta}} : \boldsymbol{\beta} \in \mathcal{L}_n) \mapsto (c_{\boldsymbol{\beta}} : \boldsymbol{\beta} \in \mathcal{L}_k).$$

Let $\mathcal{G}_n$ denote the set of all partial G-sequences of order n. Thus, by the previous theorem $M_k^{Simplexm} \subset \Pi^{\mathcal{S}}_{n,k}(\mathcal{G}_n) \subset \Pi^{\mathcal{S}}_{n+1,k}(\mathcal{G}_{n+1})$ for all $n > k$. Moreover, $M_k^{Simplexm} = \bigcap_{n \geq k} \Pi^{\mathcal{S}}_{n,k}(\mathcal{G}_n)$.

The sets $\mathcal{G}_n$ are completely described in [Gup00]. In particular, they are bounded with nonzero Lebesgue measure . So, we can endow them with the corresponding uniform probability measure. Let $X_n^{\mathcal{S}}$ be a random vector of $\mathcal{G}_n$ having uniform distribution. We define the random multi-sequence $(X_n^{\mathcal{S},k} : n \in \mathbb{N}^*)$ of $\mathcal{G}_k$ by

$$X_n^{\mathcal{S},k} := \Pi^{\mathcal{S}}_{n,k} X_n^{\mathcal{S}}.$$

As we have said in the introduction $\nu^{\mathcal{S}}$, the uniform distribution on $\mathcal{S}$, plays a major role in the asymptotic behavior of $\left(X_n^{\mathcal{S},k} : n \in \mathbb{N}\right)$. Let $\mathbf{c}^{\mathcal{S},k} = (\mathbf{c}^{\mathcal{S}}_{\boldsymbol{\beta}} : \boldsymbol{\beta} \in \mathcal{L}_k)$ denotes the moment multi-sequence of order k of $\nu^{\mathcal{S}}$, i.e.

$$\mathbf{c}^{\mathcal{S}}_{\boldsymbol{\beta}} := \int_{\mathcal{S}} x^{\boldsymbol{\beta}} \,\mathrm{d}\nu^{\mathcal{S}}(x) = \frac{m! \prod_{i=1}^m \beta_i}{(m + |\boldsymbol{\beta}|)!}, \qquad \boldsymbol{\beta} \in \mathcal{L}_k.$$

Gupta [Gup00] stated a normal central limit for $(X_n^{\mathcal{S},k})$ that is not completely certain. He made a little mistake that lead him to obtain and incorrect variance matrix for the limit normal distribution. Here we present the correct statement of his theorem and we prove it. Particularly, we will use combinatorial results established in [Gup00].

Theorem 3.3.

$$\sqrt{\binom{n+m}{n}} \left(X_n^{\mathcal{S},k} - \mathbf{c}^{\mathcal{S},k}\right) \xrightarrow[n\to\infty]{Law} \mathcal{N}_{N_k}(0, \Sigma_k^{\mathcal{S}}),$$

where

$$\Sigma_k^{\mathcal{S}} = \left(\mathbf{c}^{\mathcal{S}}_{\boldsymbol{\gamma}+\boldsymbol{\beta}} - \mathbf{c}^{\mathcal{S}}_{\boldsymbol{\gamma}} \mathbf{c}^{\mathcal{S}}_{\boldsymbol{\beta}}\right)_{\boldsymbol{\gamma},\boldsymbol{\beta} \in \mathcal{L}_k}.$$

Our main result in this context concerns the large deviation behavior of the sequences of random multi-sequences $\left(X_n^{\mathcal{S},k}\right)$.

Theorem 3.4. *The random vector sequence $\left(X_n^{\mathcal{S},k} : n \in \mathbb{N}\right)$ satisfies a LDP with speed $\left(\binom{n+m}{n}^{-1} : n \in \mathbb{N}\right)$ and good rate function*

$$I^k_{\mathcal{S}}(x) = \inf_{\mu \in \mathcal{P}(\mathcal{S},\{x\})} \mathcal{K}(\nu^{\mathcal{S}} | \mu), \quad x \in \mathcal{R}_k.$$

Remark 1. Using the dual relation in Section 2.2 the rate function $I_{\mathcal{S}}^k$ could be expressed as, for $x \in M_k^{Simplexm}$,

$$I_{\mathcal{S}}^k(x) = \sup_{(\Lambda_0,\Lambda)\in\mathbb{R}\times\mathcal{R}_k} \left\{ \Lambda_0 - 1 + \langle x, \Lambda\rangle + \int_{\mathcal{S}} \ln\left(1 - \Lambda_0 - \sum_{\boldsymbol{\beta}\in\mathcal{L}_k} \Lambda_{\boldsymbol{\beta}} x^{\boldsymbol{\beta}}\right) \mathrm{d}\nu^{\mathcal{S}}(x) \right\}$$

The following result gives an estimation of convergence rate between those of the Theorem 3.3 and Theorem 3.4.

Theorem 3.5. *Let* $\big(M(n) : n \in \mathbb{N}\big)$ *be a sequence of positive real numbers growing to* $+\infty$ *such that* $M(n) = o\big(\binom{n+m}{n}\big)$. *The sequence of random vectors*

$$\widetilde{X}_n^{\mathcal{S},k} := \sqrt{\frac{\binom{n+m}{n}}{M(n)}} \left(X_n^{\mathcal{S},k} - \mathbf{c}^{\mathcal{S},k} \right)$$

satisfies the LDP with speed $(M(n)^{-1})$ *and good rate function*

$$H_{\mathcal{S}}(\mathbf{x}) := \frac{1}{2} \mathbf{x}^T \left(\Sigma_k^{\mathcal{S}}\right)^{-1} \mathbf{x}, \quad \mathbf{x} \in \mathcal{R}_k.$$

Finally, we present LDP in an infinite dimensional setup. Let $M_\infty^{\mathcal{S}}$ be the set of all moment multi-sequences, i.e. in view of Theorem 3.2, the set of all G-sequences. Take on $\mathcal{R}_\infty$ the norm defined by

$$\|\mathbf{c}\| := \sum_{\boldsymbol{\beta}\in\mathcal{L}} \frac{1}{\binom{|\boldsymbol{\beta}|+m}{m} |\boldsymbol{\beta}|^2} |c_{\boldsymbol{\beta}}|, \quad \big(\mathbf{c} = (c_{\boldsymbol{\beta}} : \boldsymbol{\beta} \in \mathcal{L})\big). \tag{3.2}$$

Equip $\mathcal{R}_\infty$ with the corresponding borelian σ-field. Endow $M_\infty^{\mathcal{S}}$ with the induced topology and σ-field. Further, let $\mathcal{C}_\infty \subset \mathcal{R}_\infty$ defined as

$$\mathcal{C}_\infty := \{\mathbf{c} \in \mathcal{R}_\infty : \forall \boldsymbol{\beta} \;\, 0 \le c_{\boldsymbol{\beta}} \le 1\}.$$

Obviously, $M_\infty^{\mathcal{S}} \subset \mathcal{C}_\infty$. It is well known that the moment problem on $\mathcal{S}$ is determined [ST43, Corollary 1.1]. In others words, given $\mathbf{c} \in M_\infty^{\mathcal{S}}$ there exists an unique probability measure $\mu_{\mathbf{c}}$ verifying

$$c_{\boldsymbol{\beta}} = \int_{\mathcal{S}} x^{\boldsymbol{\beta}} \, \mathrm{d}\mu_{\mathbf{c}}(x), \quad \forall\, \boldsymbol{\beta} \in \mathcal{L}.$$

For $n \in \mathbb{N}^*$, let $\Pi_{\infty,n}^{\mathcal{S}} : \mathcal{R}_\infty \to \mathcal{R}_n$ be the projection map.

Theorem 3.6. *Let* (μ_n) *be a sequence of probability measures defined on* $\mathcal{C}_\infty$ *such that*

$$\mu_n \circ (\Pi_{\infty,n}^{\mathcal{S}})^{-1}$$

is the uniform probability measure on $\mathcal{G}_n$. *Then,* (μ_n) *verifies the LDP on* $\mathcal{R}_\infty$ *endowed with the topology induced by the norm* (3.2). *The good rate function is*

$$I_{\mathcal{G}}(\mathbf{c}) = \begin{cases} \mathcal{K}(\nu^{\mathcal{S}} | \mu_{\mathbf{c}}), & \text{if } \mathbf{c} \in M_\infty^{\mathcal{S}}, \\ +\infty, & \text{otherwise} \end{cases} \tag{3.3}$$

and the speed is $\left(\binom{n+m}{n}^{-1} : n \in \mathbb{N}\right)$.

Remark 2. As can be seen in the proof of this theorem, it remains valid if we consider as $\mathcal{C}_\infty$ any bounded set (in the infinity norm) containing $M_\infty^{\mathcal{S}}$. Furthermore, the theorem holds for the family of norms

$$\|\mathbf{c}\|_l := \sum_{\boldsymbol{\beta}\in\mathcal{L}} \frac{1}{\binom{|\boldsymbol{\beta}|+m}{m}|\boldsymbol{\beta}|^l}|c_{\boldsymbol{\beta}}|, \quad \big(\mathbf{c} = (c_{\boldsymbol{\beta}} : \boldsymbol{\beta} \in \mathcal{L})\big)$$

with $l > 1$. (3.2) corresponds to the case $l = 2$.

4. Trigonometric moment problem

For $\mu \in \mathcal{P}(\mathbb{T})$ the trigonometric moments on $\mathbb{T}$ are given by

$$t_k(\mu) = \int_{\mathbb{T}} z^k \,\mathrm{d}\mu(z), \quad k \in \mathbb{Z}.$$

Since $t_k(\cdot) = \bar{t}_{-k}(\cdot)$ it suffices to consider $k \geq 0$ in order to study the corresponding moment spaces. We recall that the (n-th) *moment space* is

$$M_n^{\mathbb{T}} = \Big\{\mathbf{t}_n := \big(t_1(\mu), t_2(\mu), ..., t_n(\mu)\big) : \mu \in \mathcal{P}(\mathbb{T})\Big\}.$$

For $\mathbf{z}_n = (z_1, z_2, ..., z_n) \in \mathbb{C}^{\,n}$, let

$$\mathcal{T}_n(\mathbf{z}_n) := \begin{pmatrix} 1 & z_1 & z_2 & \dots & z_n \\ \bar{z}_1 & 1 & z_1 & \dots & z_{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \bar{z}_n & \bar{z}_{n-1} & \bar{z}_{n-2} & \dots & 1 \end{pmatrix}.$$

We denote by $\Delta_n(\mathbf{z}_k)$ the determinant of $\mathcal{T}_n(\mathbf{z}_k)$. For $\mu \in \mathcal{P}(\mathbb{T})$, the matrices $\mathcal{T}_n(\mu) := \mathcal{T}_n(\mathbf{t}_n(\mu))$, $n \in \mathbb{N}^*$ are called *Toeplitz matrices*. They play a major role in the theory of moments on $\mathbb{T}$.

Let $X_n^{\mathbb{T}}$ be a random vector of $M_n^{\mathbb{T}}$ uniformly distributed. Let $\Pi_{n,k}^{\mathbb{T}}$ denotes the projection map from $\mathbb{C}^{\,n}$ to $\mathbb{C}^{\,k}$ ($k \leq n$). Let $X_n^{\mathbb{T},k}$ be random vector of $M_k^{\mathbb{T}}$ defined as

$$X_n^{\mathbb{T},k} := \Pi_{n,k}^{\mathbb{T}} X_n^{\mathbb{T}}.$$

Our first result give the limit distribution of $\big(\sqrt{n}X_n^{\mathbb{T},k} : n \in \mathbb{N}^*\big)$. Consequently, we obtain a weak law of large numbers for $\big(X_n^{\mathbb{T},k}\big)$. Namely,

$$X_n^{\mathbb{T},k} \xrightarrow[n]{\mathbb{P}} (0, 0, ..., 0) \in \mathbb{C}^{\,k}.$$

Note that the zero vector $(0, 0, ..., 0) \in \mathbb{C}^{\,[k]}$ appears as a limit because for all $j \in \mathbb{N}^*$, $t_j(\nu^{\mathbb{T}}) = 0$.

For $n \in \mathbb{N}^*$, we denote by ρ_n the measure the Lebesgue on $\mathbb{C}^{\,n}$.

Theorem 4.1. *For any measurable set A of $\mathbb{C}^{\,k}$,*

$$\lim_{n\to\infty} \mathbb{P}(\sqrt{n}\, X_n^{\mathbb{T},k} \in A) = \frac{1}{\pi^k} \int_A \exp\big(-\|z\|^2\big)\,\mathrm{d}\rho_1(z)$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{C}^{\,k}$.

Remark 3. In others words, the limiting distribution of $\left(\sqrt{n}\, X_n^{\mathbb{T},k} : n \in \mathbb{N}^*\right)$ is a k-dimensional complex normal distribution. Each component are pairwise independent with independent real and imaginary part following normal distribution of expectation zero and variance $1/2$. As we see in previous section (and in [CKS93]) the variance of the limiting distribution are very related with the moments of the limit probability. In fact, in this case, as the variance is the identity matrix it is hidden its relation with $\nu^{\mathbb{T}}$. A more descriptive formulation of the previous theorem could be

$$\sqrt{n} X_n^{\mathbb{T},k} \xrightarrow[n\to\infty]{Law} \mathcal{N}_k^{\mathbb{C}}(0, \mathcal{T}_k(\nu^{\mathbb{T}}))$$

where $\mathcal{N}_k^{\mathbb{C}}(\cdot,\cdot)$ denotes the k-dimensional complex normal distribution.

An interesting byproduct of the proof of this theorem is the "asymptotic" volume of n-th moment space. Similar results are given in [CKS93] and [Gup00]. Here, we will obtain that

$$\mathrm{Volume}_{\mathbb{R}^{2n}}\left(M_n^{\mathbb{T}}\right) = \frac{\pi^n}{n!}. \tag{4.1}$$

Since the Stirling's formula for large n, $n! = \sqrt{2\pi n}\exp(-n) n^{n+1/2}(1+o(1))$, we have

$$\mathrm{Volume}_{\mathbb{R}^{2n}}\left(M_n^{\mathbb{T}}\right) = \exp(-n \ln n[1+o(1)]).$$

Our next results concern the large and moderate deviation behavior of the sequence $\left(X_n^{\mathbb{T},k} : n \in \mathbb{N}\right)$.

Theorem 4.2. *The sequence $\left(X_n^{\mathbb{T},k} : n \in \mathbb{N}\right)$ satisfies a LDP with speed $\left(n^{-1}\right)$ and good rate function*

$$I_{\mathbb{T}}^{k(\mathbf{z}_k)} = \begin{cases} -\ln \frac{\Delta_k(\mathbf{z}_k)}{\Delta_{k-1}(\mathbf{z}_k)} & \text{if } \mathbf{z}_k \in \operatorname{int} M_k^{\mathbb{T}}, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.2}$$

Theorem 4.3. *Let (u_n) be a given decreasing sequence to 0 such that $n^{-1} = o(u_n)$. The sequence of random vectors $\left(\sqrt{nu_n}\, X_n^{\mathbb{T},k} : n \in \mathbb{N}\right)$ satisfies a LDP on $\mathbb{C}^{\,k}$ with speed $\left(u_n\right)$. The good rate function is, for all $\mathbf{z} \in \mathbb{C}^{\,k}$,*

$$H_{\mathbb{T}}(\mathbf{z}_k) := \|\mathbf{z}\|^2.$$

Finally we give a LDP for random probability measures. This result follows from Theorem 4.2 and make possible to understand the rate function in (4.2). For every n, let $\mathbb{Q}_{\,n}$ be a probability measure defined on $\mathcal{P}(\mathbb{T})$ such that if μ_n has distribution $\mathbb{Q}_{\,n}$ then the random vector $\mathbf{t}_n(\mu_n)$ has distribution $\mathbb{P}_n$.

Theorem 4.4. *The sequence of random probability measures (μ_n) satisfies a LDP on $\mathcal{P}(\mathbb{T})$ with good rate function $\mathcal{K}(\nu^{\mathbb{T}}|\cdot)$.*

Remark 4. There is a natural way to obtain such $\mathbb{Q}_{\,n}$. For $n \in \mathbb{N}^*$ and $\mathbf{z}_n \in M_n^{\mathbb{T}}$, let $\mathbb{W}_n^{\mathbf{z}_n}$ be a probability distribution on $\mathcal{P}(\mathbb{T}, \{\mathbf{z}\})$. The mixture probability measure defined as

$$\mathbb{Q}_{\,n}(\cdot) = \int_{M_n^{\mathbb{T}}} \mathbb{W}_n^{\mathbf{z}_n}(\cdot)\, \mathrm{d}\mathbb{P}_n(\mathbf{z}_n)$$

verifies our constraint. In [GLC04], in the frame of the power moment problem on $[0,1]$, the authors considered for $\mathbb{W}_n^{\mathbf{z}_n}$ the probability measure concentrated on the upper canonical representation of the moment vector $\mathbf{z}_n$. See [KN77] for definition. Applying the contraction principle, we obtain

Corollary 4.5.

$$I_{\mathbb{T}}^{k(\mathbf{z}_k)} = \inf_{\mu \in \mathcal{P}(\mathbb{T}, \{\mathbf{z}_k\})} \mathcal{K}(\nu^{\mathbb{T}} | \mu).$$

5. Related LDPs.

The results on LDP obtained in previous sections can be seen in a common frame: the interval $[0,1]$. We will compare each of these LDP with the previous one obtained in [GLC04]. This analysis will give a first approach to a slight generalization of our formulation of the asymptotics analysis on moment spaces. In particular, endowing $\widetilde{M}_n$ (or M_n) with general probability measures instead of the uniform one.

5.1. **The general $[a,b]$-moment spaces.** We recall briefly certain aspects developed in [CKS93] and [GLC04]. First, we want to remark that the all results obtained there are valid if one consider any real bounded interval of the real line. Moreover, the results are valid if one consider the moments associated to a given sequence of polynomial $\mathbf{P} = (P_n)$ with $\deg P_n = n$, $n \in \mathbb{N}^*$ instead of the ordinary power moments. For any interval $[a,b]$ $(a < b)$ of the real line and $n \in \mathbb{N}^*$, the n-th moment space is

$$M_k^{[a,b]}(\mathbf{P}) = \Big\{ (\widehat{c}_1(\mu), \widehat{c}_2(\mu), ..., \widehat{c}_k(\mu)) \in \mathbb{R}^k : \mu \in \mathcal{P}([a,b]) \Big\}.$$

where, for $\mu \in \mathcal{P}([a,b])$ and $j \in \mathbb{N}^*$,

$$\widehat{c}_j(\mu) = \int_a^b P_j(x)\, \mathrm{d}\mu(x).$$

The reason of the extensibility of the result on $[0,1]$ to any interval $[a,b]$ and any family $\{P_n\}$ is explained by the canonical moments which are defined as follows. Given $\widehat{\mathbf{c}}^j = (\widehat{c}_1, \widehat{c}_2, ..., \widehat{c}_j) \in \operatorname{int} M_j^{[a,b]}(\mathbf{P})$ we define, for $j \in \mathbb{N}$,

$$\begin{aligned} c_{j+1}^{+}(\widehat{\mathbf{c}}^j) &= \max \Big\{ c \in \mathbb{R} : ((\widehat{c}_1, \widehat{c}_2, ..., \widehat{c}_j, c) \in M_{j+1}^{[a,b]}(\mathbf{P}) \Big\} \\ c_{j+1}^{-}(\widehat{\mathbf{c}}^j) &= \min \Big\{ c \in \mathbb{R} : ((\widehat{c}_1, \widehat{c}_2, ..., \widehat{c}_j, c) \in M_{j+1}^{[a,b]}(\mathbf{P}) \Big\}. \end{aligned}$$

For $i \in \mathbb{N}^*$, the i-th canonical moment is defined as

$$\begin{aligned} \widehat{p}_i(\widehat{\mathbf{c}}^k) &= \widehat{p}_i(\widehat{\mathbf{c}}^i) \\ &= \frac{\widehat{c}_i - c_i^{-}(\widehat{\mathbf{c}}^{i-1})}{c_i^{+}(\widehat{\mathbf{c}}^{i-1}) - c_i^{-}(\widehat{\mathbf{c}}^{i-1})}. \end{aligned}$$

The canonical moments are independent of the particular choice of the sequence $\mathbf{P}$ [CKS93]. Furthermore, they are invariant by linear transformation (with positive slope) of the support interval. Namely, if $\mu \in \mathcal{P}([0,1])$ is linearly transformed measure with positive slope to $\mu' \in \mathcal{P}([a,b])$, then their corresponding canonical moments are the same [Ski69, Theorem 5]. These properties make very attractive the study of the canonical moments. We refers the reader to [DS97], a very complete monograph about the subject.

The results of [CKS93] and [GLC04] were based on the knowledge of the exact distribution of the canonical moments. Theorem 1.3 of [CKS93] establish that the uniform probability measure on $M_n^{[a,b]}(\mathbf{P})$ ($n \in \mathbb{N}^*$) is equivalent to the first n canonical moment are independent and $\widehat{p}_j$ has $\beta(n-j, n-j)$-distribution, $j = 1, 2, ..., n$.

5.2. **Deriving LDPs on $\mathbb{R}$ from LDP of Section 4.** We denoted by $F : \mathbb{T} \to [-1, 1]$ the map defined as $z \mapsto \Re z$. For $\mu \in \mathcal{P}(\mathbb{T})$, let

$$\mu_F := \mu \circ F^{-1} \in \mathcal{P}([-1, 1]). \tag{5.1}$$

For $k \in \mathbb{N}^*$, let $\widehat{F}_k$ be the map from $M_k^{\mathbb{T}}$ to $M_k([-1, 1])$ defined by (5.1). More precisely, for $\mathbf{z}_k \in M_k^{\mathbb{T}}$,

$$\widehat{F}(\mathbf{z}_k) = \widehat{\mathbf{c}}^k(\mu_F) := (\widehat{c}_1(\mu_F), \widehat{c}_2(\mu_F), ..., \widehat{c}_k(\mu_F)) \in M_k([0, 1]),$$

where μ is any measure on $\mathcal{P}(\mathbb{T})$ representing $\mathbf{z}_k$. We will see in the following that this map is independent of the selection of μ. Hence, it is well-defined. For $j = 0, 1, ..., k$

$$\begin{aligned} \Re t_j(\mu) = \Re\left(\int_{\mathbb{T}} z^j \,\mathrm{d}\mu(z)\right) &= \int_{\mathbb{T}} \cos(j \arccos F(z)) \,\mathrm{d}\mu(z) \\ &= \int_{-1}^{1} T_j(x) \,\mathrm{d}\mu_F(x) \end{aligned}$$

where $T_j = \cos(j \arccos x)$ is the j-th Chebyshev polynomial of the first kind. These polynomials can be expressed as $T_j(x) = \sum_{j=0}^{m} a_{mj} x^j$ where

$$a_{mj} = \begin{cases} \dfrac{m}{2} 2^j (-1)^{\frac{m-j}{2}} \dbinom{m+j}{j+1} & \text{if } m-j \text{ is even,} \\ 0, & \text{otherwise .} \end{cases}$$

Therefore, for $\mu \in \mathcal{P}([-1, 1])$ and for $\mathbf{s} = (s_0, s_1, ..., s_k)$ with

$$s_j = s_j(\mu) := \int_{-1}^{1} T_j \,\mathrm{d}\mu, \quad j = 0, 1, ..., k,$$

and $\mathbf{c} = (1, \widehat{c}_1(\mu), \widehat{c}_2(\mu), ..., \widehat{c}_k(\mu))$ we have $\mathbf{c} = A^{-1}\mathbf{s}$ with $A = (a_{mj} : m, j = 1, 2, ..., k)$. Thus, the relation $(1, \widehat{F}(\mathbf{z}_k)) = A^{-1}(1, \Re \mathbf{z}_k)$ define $\widehat{F}(\cdot)$ independently of the measure representing $\mathbf{z}_k$.

By the contraction principle and the δ-method can be derived from the results of Section 4 the asymptotic behavior of sequences of random vectors

$$\widehat{Z}_n^{(k)} := \widehat{F}(X_n^{\mathbb{T},k}) = \Pi_{n,k}^{\mathcal{S}} \widehat{F}(\widehat{Z}_n^{(k)}).$$

Surprisingly, we obtain the same asymptotic behavior (after obvious translation to $[0, 1]$) of the random moment vector described in [CKS93] and [GLC04]. The explanation is given by the behavior of canonical moments. It is no hard to prove that the uniform probability measure on $M_n^{\mathbb{T}}$ yields that the first n (random) canonical moments $\widehat{p}_j$ (related to $\widehat{F}(X_n^{\mathbb{T},n})$) are independent with $\beta(n-j+1/2, n-j+1/2)$-distribution, $j = 1, 2, ..., n$. Compare this result with the corresponding one in previous section.

We can use the asymptotics results to study a moment space on an interval of the real line in different way. Actually, this is an equivalent form of the same study. Consider in the interval $[-\pi,\pi)$ the system of continuous functions

$$\Phi_n^{[-\pi,\pi)} := \big(\cos\theta,\, \sin\theta,\, \cos 2\theta,\, \sin 2\theta, ...,\, \cos n\theta,\, \sin n\theta\big), \quad n\in\mathbb{N}^*.$$

For $\mu\in\mathcal{P}([-\pi,\pi))$ the corresponding *real trigonometric moments* are defined as $\alpha_0=1$,

$$\alpha_k(\mu)=\int_{-\pi}^{\pi}\cos k\theta\,\mathrm{d}\mu(\theta), \quad \beta_k(\mu)=\int_{-\pi}^{\pi}\sin k\theta\,\mathrm{d}\mu(\theta), \quad k\in\mathbb{N}.$$

The *moment space* $M_n^{[-\pi,\pi)}$ is defined as

$$M_n^{[-\pi,\pi)}=\Big\{(\alpha_1(\mu),\beta_1(\mu),...,\alpha_n(\mu),\beta_n(\mu))^T\in\mathbb{R}^{2n} : \mu\in\mathcal{P}([-\pi,\pi))\Big\}.$$

We may identify the set $\mathcal{P}(\mathbb{T})$ with $\mathcal{P}([-\pi,\pi))$ by the relation $z=\exp(i\theta)$, for $z\in\mathbb{T}$ and $\theta=[-\pi,\pi)$. Using the convention $\mathrm{d}\mu(\theta):=\mathrm{d}\mu(e^{i\theta})$, we have

$$t_k(\mu)=\int_{-\pi}^{\pi}e^{ik\theta}\,\mathrm{d}\mu(\theta), \quad k\in\mathbb{N}.$$

The power moments t_k are related with the trigonometric moments in the following way:

$$t_k(\mu)=\alpha_k(\mu)+i\beta_k(\mu).$$

for $\mu\in\mathcal{P}(\mathbb{T})$ and $k\in\mathbb{N}$. Taking $\mathbb{P}_n$ on $M_n^{\mathbb{T}}$ is equivalent to endow the set $M_n^{[-\pi,\pi)}$ with the corresponding uniform probability measure. Therefore, this problem is into the general setting described in the introduction of papers. Moreover, all the results for the random power moment problem on $\mathbb{T}$ can easily translate for the random real trigonometric moments problem on $[-\pi,\pi)$. The normal limit can be translate using a δ-method [VdV98] and the large deviations by the contraction principle. In particular, we can formulate the following result.

Let $\widehat{T}_n$ be a random vector of $M_n^{[-\pi,\pi)}$ uniformly distributed. Let $\widehat{T}_n^k$ be the random vector of $M_k^{[-\pi,\pi)}$ formed as projection of $\widehat{T}_n$ on $M_k^{[-\pi,\pi)}$.

Corollary 5.1. *The sequence of random vectors $\big(\widehat{T}_n^k\big)$ satisfies a LDP with speed $\big((2n)^{-1}\big)$ and good rate function*

$$I_{[-\pi,\pi)}^k(x)=\inf_{\mu\in\mathcal{P}([-\pi,\pi),\{x\})}\mathcal{K}\left(\nu^{[-\pi,\pi)}|\mu\right),$$

where $\nu^{[-\pi,\pi)}$ denotes the uniform probability measure on $[-\pi,\pi)$.

5.3. **LDP for one-dimensional G-sequences.** Considering $m=1$, in the Section 3 we have an LDP involving ordinary power moments on $[0,1]$.

In [Gup00] the description of the sets $\mathcal{G}_n$, $(n\in\mathbb{N}^*)$ is given. In particular, for $m=1$, the set $\mathcal{G}_n$ is the convex hull of the finite set of $n+1$ points

$$c_n^i=(c_{n,1}^i,c_{n,2}^i,...,c_{n,n}^i), \quad i=0,1,...,n,$$

with

$$c_{n,j}^i=\begin{cases}\dfrac{\binom{i}{j}}{\binom{n}{j}} & \text{if } i\geq j\\ 0 & \text{otherwise,}\end{cases} \qquad j=1,...,n.$$

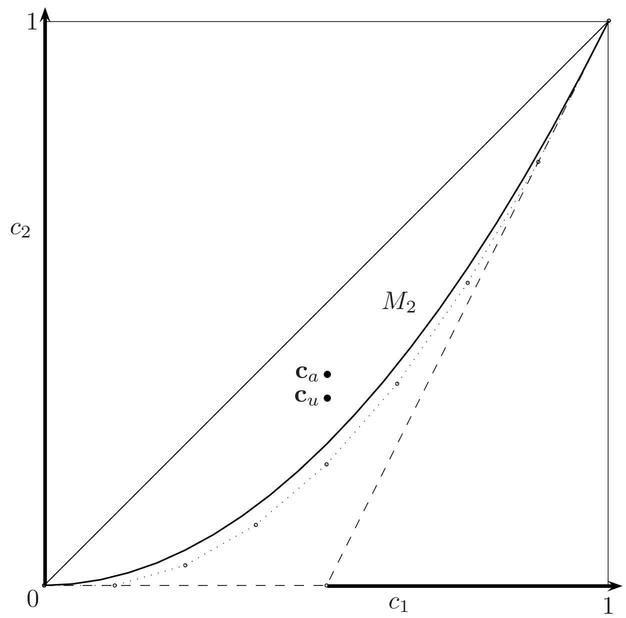

FIGURE 1. The moment space M_2 (region limited by the solid line) and the approximating spaces $\mathcal{G}_2$, $\Pi^{\mathcal{S}}_{8,2}\mathcal{G}_8$ limited by the segment with extreme points $(0,0)$ and $(1,1)$ (represented in solid line) and the dashed polygonal and dotted polygonal respectively. c_a and c_u the 2-dimensional moments of arcsine and uniform distribution respectively.

In Figure 1 we sketch the moment space M_2 and the sets $\Pi^{\mathcal{S}}_{n,2}\mathcal{G}_n$, $n = 2, 8$. We highlight the points $\Pi^{\mathcal{S}}_{8,2}c^i_8$, $i = 0, 1, ..., 8$. As can be seen in the figure the "enlargement" of the extended moment spaces under the curve $\{x, x^2 : x \in [0,1]\}$ yields that the point limit c_u be "underneath" to the point c_a.

There is a different frame where the sets $\mathcal{G}_n$ appear: Bayesian binomial estimation. We refer the reader to Section 10.2 of [DS97] and the paper of Skibinsky [Ski68] to know about this subject.

5.4. **LDP on canonical moment spaces.** The arguments used to proofs the results in [CKS93], [GLC04] and the robust properties of the canonical moment suggest to study the asymptotics related to random moment vector endowing with probability measures the canonical moment spaces instead of the ordinary moments space. We present in the following a very simple construction of a probability measure on canonical moment spaces. Let $(\mu_n) \subset \mathcal{P}([0,1])$ satisfying the LDP with rate function I_C. Equip $[0,1]^n$ (the n-th canonical moment space) with

$$\overline{\mu}_n := \mu_n \otimes \mu_{n-1} \otimes \cdot \overset{\cdot}{_1} \cdot \otimes \mu_1 \in \mathcal{P}([0,1]^n).$$

Let $\widetilde{\Pi}_{n,k}$ denote the projection map from $[0,1]^n$ to $[0,1]^k$, taking the first n components. Following similar arguments to those of [GLC04] (and Section 7 in the context of complex moments) we have that $\big(\mu^k_n := \overline{\mu}_n \circ (\widetilde{\Pi}_{n,k})^{-1} : n \in \mathbb{N}\big)$ verify the LDP with good rate function $I^k_C(p_1, p_2, ..., p_k) = \sum_{i=1}^k I_C(p_i)$.

Let $\mathbf{p}_n$ denote the map that transforms the n-dimensional ordinary moment vector on the corresponding n-dimensional canonical moment vector. Let $\widehat{U}_n := \overline{\mu}_n \circ \mathbf{p}_n^{-1}$. The sequence $(\hat{\mu}_n^{(k)} := \widehat{U}_n \circ (\widetilde{\Pi}_{n,k})^{-1}) \subset \mathcal{P}(M_n^{[a,b]})$ satisfy the LDP with $\overline{I}^k(c) = I_C^k(\mathbf{p}_k(c))$, where $c \in M_k^{[a,b]}$.

Similar constructions could be make in the complex canonical moment spaces. In a forthcoming paper we will develop these ideas.

6. Proofs of Results of Section 3

6.1. **Notation and previous results.** We will use the notation of [Gup00]. For $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_m) \in \mathbb{N}^m$ and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_m) \in \mathbb{N}^m$,

$$\boldsymbol{\beta}! := \prod_{i=1}^m \beta_i!,$$

$$\binom{n}{\boldsymbol{\beta}} := \frac{n!}{\boldsymbol{\beta}!(n-|\boldsymbol{\beta}|)!},$$

$$\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} := \prod_{i=1}^m \binom{\alpha_i}{\beta_i},$$

$$N(n) := |\mathcal{L}_n| = \binom{n+m}{m} - 1.$$

For $n \in \mathbb{N}$,

$$\overline{\mathcal{L}}_n := \mathcal{L}_n \cup \{(0,0,...,0)\}$$

and $\overline{\mathcal{R}}_n = \mathbb{R}^{|\overline{\mathcal{L}}_n|}$, i.e, $\overline{\mathcal{R}}_n$ is the set of the multi-sequences of order n with an additional component indexed by $(0,0,...,0)$. In this section, for sake of simplify of notation X_n^k stands for $X_n^{\mathcal{S},k}$.

Proposition 3.3 of [Gup00] gives a useful expression for the distribution of X_n^k. This result is the core of all the proofs. Let $(E_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \overline{\mathcal{L}}_n)$ be a multi-sequence of i.i.d. standard exponential random variables Exp(1) on a certain probability space.

Lemma 6.1. *The law of X_n^k is same as law of*

$$Y_{n,\boldsymbol{\beta}} := \frac{\displaystyle\sum_{\boldsymbol{\delta} \in \overline{\mathcal{L}}_{n-|\boldsymbol{\beta}|}} \binom{\boldsymbol{\beta}+\boldsymbol{\delta}}{\boldsymbol{\delta}} E_{\boldsymbol{\beta}+\boldsymbol{\delta}}}{\displaystyle\binom{n}{\boldsymbol{\beta}} \sum_{\boldsymbol{\delta} \in \overline{\mathcal{L}}_n} E_{\boldsymbol{\alpha}}}, \quad \boldsymbol{\beta} \in \mathcal{L}_k. \tag{6.1}$$

6.2. **Proof of Theorem 3.3.** We have the following relations due to Gupta. The proof can be found in [Gup00, pag. 427–428].

Lemma 6.2. *For $\boldsymbol{\beta} \in \overline{\mathcal{L}}_k$, $k < n$,*

$$\frac{1}{N_n} \sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n} \frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}}{\binom{n}{\boldsymbol{\beta}}} \mathbb{I}_{\{\boldsymbol{\beta} \leq \boldsymbol{\alpha}\}} = \mathbf{c}_{\boldsymbol{\beta}}^{\mathcal{S}}. \tag{6.2}$$

For $\boldsymbol{\beta}, \boldsymbol{\gamma} \in \overline{\mathcal{L}}_k$, $k < n$,

$$\frac{1}{N_n} \sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n} \frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}}{\binom{n}{\boldsymbol{\beta}}} \mathbb{I}_{\{\boldsymbol{\beta} \le \boldsymbol{\alpha}\}} \frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}}}{\binom{n}{\boldsymbol{\gamma}}} \mathbb{I}_{\{\boldsymbol{\gamma} \le \boldsymbol{\alpha}\}} \underset{n}{\rightarrow} \mathbf{c}^{\mathcal{S}}_{\boldsymbol{\beta}+\boldsymbol{\gamma}}. \tag{6.3}$$

Define the random finite multisequence $(\widetilde{Y}_n)$ by

$$\widetilde{Y}_{n,\boldsymbol{\beta}} := \frac{\displaystyle\sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n : \boldsymbol{\alpha} \ge \boldsymbol{\beta}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} E_{\boldsymbol{\alpha}}}{\displaystyle\binom{n}{\boldsymbol{\beta}}\binom{n+k}{n}}, \quad \boldsymbol{\beta} \in \overline{\mathcal{L}}_k \tag{6.4}$$

where $(E_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \overline{\mathcal{L}}_n)$ is the same as in (6.1).

For make a shorter notation we write N instead of $N(n)$. Define the sequence of random multisequences $\big(Z_n : n \in \mathbb{N}\big)$ as

$$Z_{n,\boldsymbol{\beta}} = \sqrt{N}\left(\widetilde{Y}_{n,\boldsymbol{\beta}} - \widetilde{Y}_{n,(0,\ldots,0)}\mathbf{c}^{\mathcal{S}}_{\boldsymbol{\beta}}\right), \quad \boldsymbol{\beta} \in \mathcal{L}_k.$$

As $\widetilde{Y}_{n,(0,\ldots,0)} \overset{P}{\to} 1$, it follows that (Z_n) and $\left(\sqrt{N}\left(Y_n - \mathbf{c}^{\mathcal{S},k}\right)\right)$ have the same limit law.

We will to computed the characteristic function associated to the variables Z_n. For all $\Theta \in \mathcal{R}_k$,

$$\begin{aligned}
&\mathbb{E}\left[\exp i\langle \Theta, Z_n\rangle)\right] \\
&= \int_{\mathbb{R}^N_+} \exp\left(i \sum_{\boldsymbol{\beta} \in \mathcal{L}_k} \Theta_{\boldsymbol{\beta}} \sqrt{N}\left(\sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n} \frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} x_{\boldsymbol{\alpha}}}{\binom{n}{\boldsymbol{\beta}} N} \mathbb{I}_{\{\boldsymbol{\beta} \le \boldsymbol{\alpha}\}} - \frac{\sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n} x_{\boldsymbol{\alpha}}}{N} \mathbf{c}^{\mathcal{S}}_{\boldsymbol{\beta}}\right)\right. \\
&\qquad \left. - \sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n} x_{\boldsymbol{\alpha}}\right) \prod_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n} \mathrm{d}x_{\boldsymbol{\alpha}} \\
&= \int_{\mathbb{R}^N_+} \exp\left(\sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n} x_{\boldsymbol{\alpha}}\left(\frac{i}{\sqrt{N}} \sum_{\boldsymbol{\beta} \in \mathcal{L}_k} \Theta_{\boldsymbol{\beta}}\left(\frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}}{\binom{n}{\boldsymbol{\beta}}} \mathbb{I}_{\{\boldsymbol{\beta} \le \boldsymbol{\alpha}\}} - \mathbf{c}^{\mathcal{S}}_{\boldsymbol{\beta}}\right) - 1\right)\right) \prod_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n} \mathrm{d}x_{\boldsymbol{\alpha}} \\
&= \prod_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n}\left(1 - \frac{i}{\sqrt{N}} B(\Theta, n, \boldsymbol{\alpha})\right)^{-1},
\end{aligned}$$

where

$$B(\Theta, n, \boldsymbol{\alpha}) = \sum_{\boldsymbol{\beta} \in \mathcal{L}_k} \Theta_{\boldsymbol{\beta}}\left(\frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}}{\binom{n}{\boldsymbol{\beta}}} \mathbb{I}_{\{\boldsymbol{\beta} \le \boldsymbol{\alpha}\}} - \mathbf{c}^{\mathcal{S}}_{\boldsymbol{\beta}}\right). \tag{6.5}$$

So, using the Taylor expansion $\ln(1+r) = r - 1/2r^2 + o(r^2)$ valid in a neighborhood of 0,

$$\begin{aligned}
\ln\left(\mathbb{E}\left[\exp i\langle \Theta, Z_n\rangle\right]\right) &= -\sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n} \ln\left(1 - \frac{i}{\sqrt{N}} B(\Theta, n, \boldsymbol{\alpha})\right) \\
&= -\sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n}\left(-\frac{i}{\sqrt{N}} B(\Theta, n, \boldsymbol{\alpha}) + \frac{1}{2N}(B(\Theta, n, \boldsymbol{\alpha}))^2 + o\left(\frac{1}{N}\right)\right).
\end{aligned}$$

By Lemma 6.2,

$$\sum_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} B(\Theta, n, \boldsymbol{\alpha}) = 0.$$

and

$$\lim_n \frac{1}{2N} \sum_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} (B(\Theta, n, \boldsymbol{\alpha}))^2 = \sum_{\boldsymbol{\beta}\in\mathcal{L}_k} \sum_{\boldsymbol{\gamma}\in\mathcal{L}_k} \Theta_{\boldsymbol{\beta}}\Theta_{\boldsymbol{\gamma}} \left(\mathbf{c}^{\mathcal{S}}_{\boldsymbol{\beta}+\boldsymbol{\gamma}} - \mathbf{c}^{\mathcal{S}}_{\boldsymbol{\beta}}\mathbf{c}^{\mathcal{S}}_{\boldsymbol{\gamma}}\right).$$

Therefore, by Lévy's continuity theorem [VdV98, Theorem 2.13], the limit

$$\lim_n \mathbb{E}\left[\exp i\langle\Theta, Z_n\rangle\right] = \exp\left(-\Theta^T \Sigma^{\mathcal{S}}_k \Theta\right)$$

implies that

$$Z_n \xrightarrow[n]{Law} \mathcal{N}_{N_k}\left(\mathbf{0}, \Sigma^{\mathcal{S}}_k\right).$$

6.3. **Proof of Theorem 3.4.** Let $\left(\widehat{Y}_n : n \in \mathbb{N}\right)$ be the random finite multi-sequence of $\overline{\mathcal{R}}_k$ defined by

$$\widehat{Y}_{n,\boldsymbol{\beta}} := \frac{1}{N(n)} \sum_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} \left(\frac{1}{n}\boldsymbol{\alpha}\right)^{\boldsymbol{\beta}} E_{\boldsymbol{\alpha}}, \quad \boldsymbol{\beta} \in \mathcal{L}_k.$$

The proof of theorem is based on the LDP for the sum of weighted random variables. In this case, it is not possible to derive it directly from the LDP for weighted empirical means obtained in [GG97], but we may use the result of [Naj02].

From Theorem 2.2 of [Naj02] we have the LDP for the sequence $(\widehat{Y}_n)$ with good rate function

$$I^k_1(x) = \sup_{\Lambda\in\overline{\mathcal{R}}_k} \left\{\langle\Lambda, x\rangle + \int_{\mathcal{S}} \ln(1 - P_\Lambda)\,\mathrm{d}\nu^{\mathcal{S}}\right\}, \quad x \in \overline{\mathcal{R}}_k,$$

where, for $\Lambda \in \overline{\mathcal{R}}_k$, P_Λ is the polynomial on $\mathcal{S}$

$$P_\Lambda(x) = \sum_{\boldsymbol{\beta}\in\overline{\mathcal{L}}_k} \Lambda_{\boldsymbol{\beta}} x^{\boldsymbol{\beta}}.$$

Consider on $\overline{\mathcal{R}}_k$ the following metric

$$\mathrm{d}(\mathbf{x}, \mathbf{y}) := \sum_{\boldsymbol{\beta}\in\overline{\mathcal{L}}_k} |\mathbf{x}_{\boldsymbol{\beta}} - \mathbf{y}_{\boldsymbol{\beta}}|, \quad \mathbf{x}, \mathbf{y} \in \overline{\mathcal{R}}_k$$

compatible with the topology on it.

Recall the definition of the sequence $(\widetilde{Y}_n)$ in (6.4). We have the following lemma.

Lemma 6.3. *The sequences $(\widehat{Y}_n)$ and $(\widetilde{Y}_n)$ are exponentially equivalent on $(\overline{\mathcal{R}}_k, \mathrm{d})$.*

Proof. Let $\epsilon > 0$. Set

$$\overline{\mathcal{L}}_{\epsilon,n} = \{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n : \forall\, i\; \alpha_i \geq \lfloor \epsilon n \rfloor\},$$

where $\lfloor w \rfloor$ denotes the integer part of w. For n big enough such that $k < \epsilon n$ and for $\boldsymbol{\alpha} \in \overline{\mathcal{L}}_{\epsilon,n}$

$$\frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}}{\binom{n}{\boldsymbol{\beta}}} \geq \prod_{i=1}^{m} \left(\frac{\alpha_i - \beta_i}{n} \right)^{\beta_i} \geq \left(\frac{1}{n}\boldsymbol{\alpha} \right)^{\boldsymbol{\beta}} \left(1 - \frac{k}{\epsilon n} \right)^{k}$$

and

$$\frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}}{\binom{n}{\boldsymbol{\beta}}} \leq \prod_{i=1}^{m} \left(\frac{\alpha_i}{n - |\boldsymbol{\beta}|} \right)^{\beta_i} \leq \left(\frac{1}{n}\boldsymbol{\alpha} \right)^{\boldsymbol{\beta}} \left(1 + \frac{k}{n-k} \right)^{k}. \tag{6.6}$$

These two inequalities leads to

$$\left(1 - \frac{k}{\epsilon n} \right)^{k} - 1 \leq \frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}}{\binom{n}{\boldsymbol{\beta}}} - \left(\frac{1}{n}\boldsymbol{\alpha} \right)^{\boldsymbol{\beta}} \leq \left(1 + \frac{k}{n-k} \right)^{k} - 1. \tag{6.7}$$

If we set

$$Q_0(\epsilon, n) := \max \left(1 - \left(1 - \frac{k}{\epsilon n} \right)^{k}, \left(1 + \frac{k}{n-k} \right)^{k} - 1 \right) \tag{6.8}$$

then

$$\left| \frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}}{\binom{n}{\boldsymbol{\beta}}} - \left(\frac{1}{n}\boldsymbol{\alpha} \right)^{\boldsymbol{\beta}} \right| \leq Q_0(\epsilon, n) \to 0 \quad \text{when } n \to \infty.$$

On the other hand, we have

$$\begin{aligned} |\overline{\mathcal{L}}_{\epsilon,n}| &= \left| \left\{ (\lfloor \epsilon n \rfloor, \lfloor \epsilon n \rfloor, ..., \lfloor \epsilon n \rfloor) + \boldsymbol{\delta} : \boldsymbol{\delta} \in \overline{\mathcal{L}}_{n-[\epsilon n]} \right\} \right| \\ &= N(n - \lfloor \epsilon n \rfloor). \end{aligned}$$

By straightforward calculations

$$\left(\frac{1-\epsilon}{1 + m/n} \right)^{m} \leq \frac{|\overline{\mathcal{L}}_{\epsilon,n}|}{N(n)} \leq \left(1 - \epsilon + \frac{m}{n} \right)^{m}. \tag{6.9}$$

For n big enough, by (6.6), (6.7) and (6.8) we have the bound

$$\mathrm{d}\big(\widehat{Y}_n, \widetilde{Y}_n\big) \leq \frac{1}{N(n)} \left(2 \sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n \setminus \overline{\mathcal{L}}_{\epsilon,n}} E_{\boldsymbol{\alpha}} + Q_0(\epsilon, n) \sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_{\epsilon,n}} E_{\boldsymbol{\alpha}} \right)$$

Using (6.9) and the fact that $|\overline{\mathcal{L}}_n \setminus \overline{\mathcal{L}}_{\epsilon,n}| = N(n) - |\overline{\mathcal{L}}_{\epsilon,n}|$,

$$\begin{aligned} \mathrm{d}\big(\widehat{Y}_n, \widetilde{Y}_n\big) \leq{} & 2 \left(1 - \left(\frac{1-\epsilon}{1+m/n} \right)^{m} \right) \frac{1}{|\overline{\mathcal{L}}_n \setminus \overline{\mathcal{L}}_{\epsilon,n}|} \sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n \setminus \overline{\mathcal{L}}_{\epsilon,n}} E_{\boldsymbol{\alpha}} \\ & + Q_0(\epsilon, n) \, (1 - \epsilon + m/n)^{m} \, \frac{1}{|\overline{\mathcal{L}}_{\epsilon,n}|} \sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_{\epsilon,n}} E_{\boldsymbol{\alpha}}. \end{aligned}$$

For all $Q_1 > 0$, there exists $\epsilon > 0$ such that for all n big enough

$$2 \left(1 - \left(\frac{1-\epsilon}{1+m/n} \right)^{m} \right) < \frac{t}{2Q_1}$$

and

$$Q_0(\epsilon, n) \left(1 - \epsilon + \frac{m}{n} \right)^{m} < \frac{t}{2Q_1}.$$

Therefore

$$\mathbb{P}\left(\mathrm{d}\big(\widehat{Y}_n, \widetilde{Y}_n\big) > t\right) \leq \mathbb{P}\left(\frac{1}{|\overline{\mathcal{L}}_n \setminus \overline{\mathcal{L}}_{\epsilon,n}|} \sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_n \setminus \overline{\mathcal{L}}_{\epsilon,n}} E_{\boldsymbol{\alpha}} > Q_1\right) + \mathbb{P}\left(\frac{1}{|\overline{\mathcal{L}}_{\epsilon,n}|} \sum_{\boldsymbol{\alpha} \in \overline{\mathcal{L}}_{\epsilon,n}} E_{\boldsymbol{\alpha}} > Q_1\right). \tag{6.10}$$

By the LDP for empirical mean of independent standard exponentially distributed random variables (particular application of Cramér's Theorem [DZ98, Theorem 2.2.3]), follows

$$\lim_n \frac{1}{n} \ln \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i \geq Q_1\right) \leq -Q_1 + 1 + \ln Q_1.$$

Hence, by Lemma 1.2.15 of [DZ98], (6.9) and (6.10)

$$\begin{aligned}
\limsup \frac{1}{N(n)} \ln \mathbb{P}&\left(\mathrm{d}\big(\widehat{Y}_n, \widetilde{Y}_n\big) > t\right) \\
&\leq \max\Big(\epsilon(1 + \ln Q_1 - Q_1), (1-\epsilon)(1 + \ln Q_1 - Q_1)\Big) \\
&\leq 1 + \ln Q_1 - Q_1
\end{aligned}$$

The exponentially equivalence is then consequence of the arbitrariness of t and Q_1. □

By the previous lemma $\big(\widetilde{Y}_n\big)$ verifies the same LDP as $\big(\widehat{Y}_n\big)$. The contraction principle with the continuous map from $\overline{\mathcal{R}}_k$ to $\mathcal{R}_k$ defined by $(x_{\boldsymbol{\beta}} : \boldsymbol{\beta} \in \overline{\mathcal{L}}_k) \mapsto \big(x_{\boldsymbol{\beta}}/x_{(0,0,\ldots,0)} : \boldsymbol{\beta} \in \mathcal{L}_k\big)$ yields the LDP for

$$Y_{n,\boldsymbol{\beta}} = \frac{\widetilde{Y}_{n,\boldsymbol{\beta}}}{\widetilde{Y}_{n,(0,0,\ldots,0)}}, \quad \boldsymbol{\beta} \in \mathcal{L}_k.$$

Therefore, in view of Lemma 6.1, $\big(X_n^k\big)$ verifies a LDP with good rate function

$$I_2^k(x) := \inf_{y \in \overline{\mathcal{R}}_k} \left\{ I_1^k(y) : \frac{y_{\boldsymbol{\beta}}}{y_{(0,0,\ldots,0)}} = x_{\boldsymbol{\beta}},\, \boldsymbol{\beta} \in \mathcal{L}_k \right\}.$$

It remains to proof that $I_{\mathcal{S}}^k \equiv I_2^k$.

Lemma 6.4. $I_{\mathcal{S}}^k \equiv I_2^k$.

Proof. The core of the proof is a particular application of a duality theorem concerning optimization on measure spaces. Let $\mathcal{M}(\mathcal{S})$ (resp. $\mathcal{M}_+(\mathcal{S})$) be the space of all Borel finite (resp. finite positive) measures on $\mathcal{S}$. By Theorem 7.1.3 of [Dud02] the set $\mathcal{M}(\mathcal{S})$ coincide with the set of all regular Borel finite measures on $\mathcal{S}$. For $\mu \in \mathcal{M}(\mathcal{S})$, let $\mu = \mu_a + \mu_\sigma$ be the Lebesgue decomposition of μ respect to the uniform probability measure on $\mathcal{S}$. For sake of simplicity we drop the index of $\nu^{\mathcal{S}}$ and denote it by ν. Let

$$\Upsilon_1(\mu) := -\int_{\mathcal{S}} \ln\left(\frac{\mathrm{d}\mu_a}{\mathrm{d}\nu}\right) \mathrm{d}\nu^{\mathcal{S}} + \infty \cdot \mu_\sigma^-(\mathcal{S}), \quad \mu \in \mathcal{M}(\mathcal{S}),$$

$$\Upsilon_2(\Lambda) := \int_{\mathcal{S}} \ln\left(1 - P_\Lambda\right) \mathrm{d}\nu^{\mathcal{S}} - 1, \quad \Lambda \in \overline{\mathcal{R}}_k.$$

From Theorem 3.4 of [BL93], for $x \in \overline{\mathcal{R}}_k$,

$$\inf_{\mu \in \mathcal{M}(\mathcal{S},x)} \{\Upsilon_1(\mu) + \mu(\mathcal{S})\} = \sup_{\Lambda \in \overline{\mathcal{R}}_k} \{\langle x, \Lambda\rangle + \Upsilon_2(\Lambda)\}, \tag{6.11}$$

where, for $x \in \overline{\mathcal{R}}_k$,

$$\mathcal{M}(\mathcal{S}, x) := \left\{\mu \in \mathcal{M}(\mathcal{S}) \,:\, \int_{\mathcal{S}} y^{\boldsymbol{\beta}} \, \mathrm{d}\mu(y) = x_{\boldsymbol{\beta}},\, \boldsymbol{\beta} \in \overline{\mathcal{L}}_k\right\}.$$

If $\mu \notin \mathcal{M}_+(\mathcal{S})$, then either $\mu_\sigma^-(\mathcal{S}) > 0$ or $\nu^{\mathcal{S}}\left(\left\{\frac{\mathrm{d}\mu_a}{\mathrm{d}\nu^{\mathcal{S}}} > 0\right\}\right) > 0$. Hence, $\Upsilon_1(\mu) = +\infty$ for all $\mu \notin \mathcal{M}_+(\mathcal{S})$. For $\mu \in \mathcal{M}_+(\mathcal{S})$, straightforward calculations yield

$$-\int_{\mathcal{S}} \ln\left(\frac{\mathrm{d}\mu_a}{\mathrm{d}\nu^{\mathcal{S}}}(s)\right) \mathrm{d}\nu^{\mathcal{S}} = \mathcal{K}(\nu^{\mathcal{S}}|\mu) \tag{6.12}$$

where $\mathcal{K}$ is the corresponding Kullback information for positive measures.

Then, by (6.11) and (6.3)

$$I_1^k(x) = \inf_{\mu \in \mathcal{M}_+(\mathcal{S},x)} \mathcal{K}(\nu^{\mathcal{S}}|\mu) + x_{(0,0,\ldots,0)} - 1, \tag{6.13}$$

where $\mathcal{M}_+(\mathcal{S}, x) = \mathcal{M}(\mathcal{S}, x) \cap \mathcal{M}_+(\mathcal{S})$. Let $z \in \overline{\mathcal{R}}_k$ define $\hat{z} \in \overline{\mathcal{R}}_k$ as

$$\hat{z}_{(0,0,\ldots,0)} = 1,$$
$$\hat{z}_{\boldsymbol{\beta}} = z_{\boldsymbol{\beta}}, \quad \boldsymbol{\beta} \in \mathcal{L}_k.$$

Then, by (6.13)

$$\begin{aligned} I_2^k(z) &= \inf_{r \in \mathbb{R}_+} \left\{I_1^k(r\hat{z})\right\} \\ &= \inf_{r \in \mathbb{R}_+} \left\{\inf_{\mu \in \mathcal{P}(\mathcal{S},\{z\})} \mathcal{K}(\nu^{\mathcal{S}}|\mu) - \ln r + r - 1\right\}. \end{aligned}$$

The fact that the minimum value of the real function $r \to r - \ln r$ is 1 yields

$$I_2^k(z) = \inf_{\mu \in \mathcal{P}(\mathcal{S},\{z\})} \mathcal{K}(\nu^{\mathcal{S}}|\mu).$$

□

6.4. **Proof of Theorem 3.5.** Throughout this section $(M(n) : n \in \mathbb{N})$ is a sequence of positive real numbers increasing to $+\infty$ such that $M(n) = o(N(n))$. For make a shorter notation we write M (resp. N) instead of $M(n)$ (resp. $N(n)$).

Define the sequence of random multi-sequences $\left(\widehat{Z}_n : n \in \mathbb{N}\right)$ as

$$\widehat{Z}_{n,(0,\ldots,0)} = \widetilde{Y}_{n,(0,\ldots,0)},$$
$$\widehat{Z}_{n,\boldsymbol{\beta}} = \sqrt{\frac{N(n)}{M(n)}}\left(\widetilde{Y}_{n,\boldsymbol{\beta}} - \mathbf{c}_{\boldsymbol{\beta}}^{\mathcal{S}}\right), \quad \boldsymbol{\beta} \in \mathcal{L}_k$$

where $\widetilde{Y}_n$ is as (6.4). The proof is based on the Gärtner-Ellis theorem [DZ98, Theorem 2.3.6] which involves the limit of the logarithmic moment generating function. Hence, for all $\widetilde{\Lambda} \in \overline{\mathcal{R}}_k$, we have to calculate

$$\widetilde{\Theta}(\widetilde{\Lambda}) := \lim_n \frac{1}{M} \ln \mathbb{E}\left[\exp(M\langle\widetilde{\Lambda}, \widehat{Z}_n\rangle)\right].$$

Given $\widetilde{\Lambda} \in \overline{\mathcal{R}}_k$, we denote by Λ the multi-sequence $\Lambda = \big(\widetilde{\Lambda}_{\boldsymbol{\beta}} : \boldsymbol{\beta} \in \mathcal{L}_k\big)$. For n big enough,

$$\begin{aligned}
&\mathbb{E}\left[\exp(M\langle\widetilde{\Lambda}, \widehat{Z}_n\rangle)\right] \\
&\quad = \int_{\mathbb{R}_+^N} \exp\left(M\left(\sum_{\boldsymbol{\beta}\in\mathcal{L}_k} \widetilde{\Lambda}_{\boldsymbol{\beta}} \sqrt{\frac{M}{N}} \left(\sum_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} \frac{\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}}{N\binom{n}{\boldsymbol{\beta}}} \mathbb{I}_{\{\boldsymbol{\beta}\le\boldsymbol{\alpha}\}} x_{\boldsymbol{\alpha}} - \sum_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} \frac{x_{\boldsymbol{\alpha}}}{N}\mathbf{c}_{\boldsymbol{\beta}}^{\mathcal{S}}\right)\right.\right. \\
&\qquad \left.\left. + \widetilde{\Lambda}_{(0,0,\ldots,0)} \sum_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} \frac{x_{\boldsymbol{\alpha}}}{N}\right) - \sum_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} x_{\boldsymbol{\alpha}}\right) \prod_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} \mathrm{d}x_{\boldsymbol{\alpha}} \\
&\quad = \left(1 - \sqrt{\frac{M}{N}}\,(B(\Lambda, n, \boldsymbol{\alpha})) - \frac{M}{N}\widetilde{\Lambda}_{(0,0,\ldots,0)}\right)^{-1},
\end{aligned}$$

where $B(\Lambda, n, \boldsymbol{\alpha})$ are as in (6.5). Using the Taylor expansion of $\ln(1+x)$,

$$\begin{aligned}
&\frac{1}{M}\ln\mathbb{E}\left[\exp(M\langle\widetilde{\Lambda}, \widehat{Z}_n\rangle)\right] \\
&\qquad = -\frac{1}{M}\sum_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} \ln\left(1 - \sqrt{\frac{M}{N}}\,(B(\Lambda, n, \boldsymbol{\alpha})) - \frac{M}{N}\widetilde{\Lambda}_{(0,0,\ldots,0)}\right) \\
&\qquad = \widetilde{\Lambda}_{(0,0,\ldots,0)} + \frac{1}{\sqrt{NM}}\sum_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} B(\Lambda, n, \boldsymbol{\alpha}) + \frac{1}{2N}\sum_{\boldsymbol{\alpha}\in\overline{\mathcal{L}}_n} (B(\Lambda, n, \boldsymbol{\alpha}))^2 \\
&\qquad\quad + o\left(\left(\frac{M}{N}\right)^{\frac{1}{2}}\right).
\end{aligned}$$

Lemma 6.2 implies that

$$\widetilde{\Theta}(\widetilde{\Lambda}) = \widetilde{\Lambda}_{(0,0,\ldots,0)} + \frac{1}{2}\Lambda^T D\Lambda.$$

From Gärtner-Ellis theorem we have the LDP for $(\widehat{Z}_n)$ with good rate function

$$\widetilde{H}(\widetilde{x}) = \sup_{\widetilde{\Lambda}\in\overline{\mathcal{R}}_k} \left\{\langle\widetilde{\Lambda}, \widetilde{x}\rangle - \widetilde{\Theta}(\widetilde{\Lambda})\right\}, \quad \widetilde{x} \in \overline{\mathcal{R}}_k.$$

By straightforward calculations

$$\widetilde{H}(\widetilde{x}) = \begin{cases} +\infty & \text{if } \widetilde{x}_{(0,0,\ldots,0)} \neq 1, \\ \frac{1}{2}x^T D^{-1} x & \text{if } \widetilde{x}_{(0,0,\ldots,0)} = 1, \end{cases} \quad \widetilde{x} \in \overline{\mathcal{R}}_k$$

where $x = (\widetilde{x}_{\boldsymbol{\beta}},\, \boldsymbol{\beta} \in \mathcal{L}_k) \in \mathcal{R}_k$. Now, by the contraction principle it follows a LDP for $\big(\widetilde{Z}_n\big)$ with good rate function defined as, for $y \in \mathcal{R}_k$,

$$H_{\mathcal{S}}(y) = \inf\left\{\widetilde{H}(\widetilde{x}) : y_{\boldsymbol{\beta}} = \frac{\widetilde{x}_{\boldsymbol{\beta}}}{\widetilde{x}_{(0,0,\ldots,0)}}, \quad \boldsymbol{\beta} \in \mathcal{L}_k\right\} = \frac{1}{2}y^T D^{-1} y$$

and speed $\big(M^{-1}\big)$.

6.5. **Proof of Theorem 3.6.**

Lemma 6.5. *(μ_n) verifies a LDP on $\mathcal{R}_\infty$ endowed with the topology induced by pointwise convergence with the good rate function $I_\mathcal{G}$ and the speed $\left(\binom{n+m}{n}^{-1}\right)$.*

Proof. For $k < n$ define $\mu_n^{(k)} := \mu_n \circ (\Pi^\mathcal{S}_{\infty,k})^{-1} \in \mathcal{P}(\mathcal{R}_k)$. The fact that $\Pi^\mathcal{S}_{\infty,k} = \Pi^\mathcal{S}_{n,k} \circ \Pi^\mathcal{S}_{\infty,n}$ implies that $\mu_n^{(k)} = U_n \circ (\Pi^\mathcal{S}_{n,k})^{-1}$, the law of X_n^k. Consequently $\left(\mu_n^{(k)}\right)$ verifies the LDP with good rate function $I^k_\mathcal{S}$. By Dawson-Gärtner theorem we have the LDP for $\left(\mu_n\right)$ on the projective limit of $\mathcal{R}_k$, that is $\mathcal{R}_\infty$ (equip with the pointwise convergence topology) with good rate function

$$\mathbf{c} \mapsto \sup_{k\in\mathbb{N}} I^k_\mathcal{S}(\Pi^\mathcal{S}_{\infty,k}\mathbf{c}), \quad \mathbf{c} \in \mathcal{R}_\infty. \tag{6.14}$$

By definition of $I_\mathcal{G}$, for all $k \in \mathbb{N}^*$,

$$I^k_\mathcal{S}(x) = \inf\left\{I_\mathcal{G}(\mathbf{c}) : \Pi^\mathcal{S}_{\infty,k}\mathbf{c} = x\right\}, \quad \forall\, x \in M^\mathcal{S}_k.$$

Therefore, in view of Lemma 4.6.5 of [DZ98] we have that the function in (6.14) is exactly $I_\mathcal{G}$. □

In view of Corollary 4.2.6 of [DZ98] it is sufficient to prove that $\mathcal{C}_\infty$ is compact in the topology induced by the norm.

Lemma 6.6. *The set $\mathcal{C}_\infty$ is compact.*

Proof. The proof follows elemental arguments. Since $\mathcal{C}_\infty$ is a closed set it is sufficient to proof that it is totally bounded, i.e. for all $\xi > 0$ there is a finite set $\{\mathbf{y}_i : i \in \mathcal{I}\} \subset \mathcal{C}_\infty$ such that

$$\forall \mathbf{c} \in \mathcal{C}_\infty \exists i \in \mathcal{I} : \|\mathbf{c} - \mathbf{y}_i\| \leq \xi.$$

Let $\xi > 0$ and N_ξ such that

$$\sum_{n > N_\xi} \frac{1}{n^2} < \frac{\xi}{2}. \tag{6.15}$$

For $n \in \mathbb{N}$, let

$$\mathcal{C}_n := \left\{\left(c_{\boldsymbol\beta} : 0 \leq c_{\boldsymbol\beta} \leq 1,\, \boldsymbol\beta \in \mathcal{L}_n\right)\right\} \subset \mathcal{R}_n.$$

The set $\mathcal{C}_{N_\xi}$ is a compact set in $\mathcal{R}_{N_\xi}$ equipped with the norm

$$\|\mathbf{a}\|_{N_\xi} := \sum_{\boldsymbol\beta \in \mathcal{L}_{N_\xi}} \frac{1}{\binom{|\boldsymbol\beta|+m}{m}|\boldsymbol\beta|^2} |c_{\boldsymbol\beta}| \quad \text{for } \mathbf{a} \in \mathcal{R}_{N_\xi}.$$

Consequently, $\mathcal{C}_{N_\xi}$ is totally bounded. Thus, there is a finite subset $\{\mathbf{x}_i : i \in \mathcal{I}\}$ of $\mathcal{C}_{N_\xi}$ such that for all $\mathbf{a} \in \mathcal{C}_{N_\xi}$, there is $\mathbf{x}_i$ with

$$\|\mathbf{a} - \mathbf{x}_i\|_{N_\xi} \leq \frac{\xi}{2}. \tag{6.16}$$

For all $i \in \mathcal{I}$, define $\mathbf{y}_i \in \mathcal{C}_\infty$ by

$$\mathbf{y}_{i,\boldsymbol\beta} := \begin{cases} \mathbf{x}_{i,\boldsymbol\beta} & \text{if } \boldsymbol\beta \in \mathcal{L}_{N_\xi}, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathbf{c} \in \mathcal{C}_\infty$. There exists $i \in \mathcal{I}$ such that (6.16) holds with $\mathbf{a} := \Pi_{\infty, N_\xi^{\mathcal{S}}} \mathbf{c}$. Hence,

$$\begin{aligned} \|\mathbf{c} - \mathbf{y}_i\| &= \sum_{\boldsymbol{\beta} \in \mathcal{L}_{N_\xi}} \frac{|c_{\boldsymbol{\beta}} - \mathbf{y}_{i,\boldsymbol{\beta}}|}{\binom{|\boldsymbol{\beta}|+m}{m} |\boldsymbol{\beta}|^2} + \sum_{|\boldsymbol{\beta}| > N_\xi} \frac{|c_{\boldsymbol{\beta}} - \mathbf{y}_{i,\boldsymbol{\beta}}|}{\binom{|\boldsymbol{\beta}|+m}{m} |\boldsymbol{\beta}|^2} \\ &\leq \left\| \Pi_{\infty, N_\xi^{\mathcal{S}}} \mathbf{c} - \mathbf{x}_i \right\|_{N_\xi} + \sum_{|\boldsymbol{\beta}| > N_\xi} \frac{1}{\binom{|\boldsymbol{\beta}|+m}{m} |\boldsymbol{\beta}|^2}. \end{aligned} \tag{6.17}$$

Since $|\{\boldsymbol{\beta} : |\boldsymbol{\beta}| = n\}| \leq |\overline{\mathcal{L}}_n| = \binom{n+m}{m}$, (6.15) yields the bound

$$\sum_{|\boldsymbol{\beta}| > N_\xi} \frac{1}{\binom{|\boldsymbol{\beta}|+m}{m} |\boldsymbol{\beta}|^2} = \sum_{n > N_\xi} \sum_{|\boldsymbol{\beta}| = n} \frac{1}{\binom{n+m}{m} n^2} < \frac{\xi}{2}.$$

This bound, (6.17) and (6.16) imply

$$\|\mathbf{c} - \mathbf{y}_i\| \leq \xi.$$

□

7. Proofs of results of Section 4

Throughout this section X_n^k stands for $X_n^{\mathbb{T},k}$.

7.1. **Trigonometric and canonical moments.** We consecrate this subsection to the canonical moments. These, like in [CKS93] and [GLC04], are the principal tool to proof the results. This time, in the trigonometric moment context. The notation and preliminary results follows Chapter 9 of [DS97]. Throughout this subsection consider $\mu \in \mathcal{P}(\mathbb{T})$ fixed. Further, for $n \in \mathbb{N}^*$, $\mathbf{t}_n := \mathbf{t}_n(\mu) = (t_1, t_2, ..., t_n)$ and $\Delta_n := \Delta_n(\mathbf{t}_n)$.

For $w \in \mathbb{C}$, let

$$R_n(w) := \begin{vmatrix} t_1 & t_2 & \dots & t_n & w \\ t_0 & t_1 & \dots & t_{n-1} & t_n \\ \vdots & \vdots & & \vdots & \vdots \\ t_{-n+1} & t_{-n+2} & \dots & t_0 & t_1 \end{vmatrix}$$

For $\mathbf{t}_n \in M_n^{\mathbb{T}}$, the range of the $(n+1)$-th moment is the circle

$$\{|z - s_{n+1}| \leq r_{n+1}\}$$

where $s_{n+1} = (-1)^{n+1} R_n(0)/\Delta_{n-1}$ and $r_{n+1} = \Delta_n / \Delta_{n+1}$.

Let $\mathbf{t}_n \in \operatorname{int} M_n^{\mathbb{T}}$ the first n *canonical moments* of the measure μ are defined by

$$p_k = \frac{t_k - s_k}{r_k} \in \mathbb{D}, \quad k = 1, 2, ..., n, \tag{7.1}$$

where $\mathbb{D} = \{z \in \mathbb{C} \ : |z| \leq 1\}$. They can be alternatively expressed as

$$p_k = \frac{(-1)^{k-1}}{\Delta_{k-1}} R_{k-1}(t_k), \quad k = 1, 2, ..., n.$$

We will denote by $\mathbf{p}_n$ the map from $\operatorname{int} M_n^{\mathbb{T}}$ to $\mathbb{D}^n$ representing (7.1).

Consequently

$$\frac{\partial p_i}{\partial t_j} = \begin{cases} 0, & \text{if } i < j, \\ (-1)^{j-1} \frac{\Delta_{j-2}}{\Delta_{j-1}}, & \text{if } i = j; \end{cases}$$

and

$$\left|\frac{\partial p_i}{\partial t_j}\right|_{i,j=1\ldots n} = \prod_{k=1}^{n-1}(-1)^j \frac{\Delta_{j-1}}{\Delta_j} = \frac{(-1)^{\frac{n(n-1)}{2}}}{\Delta_{n-1}}. \tag{7.2}$$

We have then the following lemma:

Lemma 7.1 (Lemma 10.7.1 of [DS97])**.** *The first-order Taylor expansion of the inverse of* (7.1) *around* $(0, 0, ..., 0)$ *is given by*

$$t_n(p_1, p_2, ..., p_n) = p_n + O\left(\sum_{i=1}^{n} |p_i|^2\right).$$

Lemma 7.2 ([DS97, Lemma 9.3.4, Corollary 9.3.5])**.**

$$\frac{\Delta_{n-1}}{\Delta_n} = \prod_{j=1}^{n}(1 - |p_j|^2), \tag{7.3}$$

$$\Delta_n = \prod_{j=1}^{n}(1 - |p_j|^2)^{n-j+1}. \tag{7.4}$$

7.2. **Proof of Theorem 4.1.** Let η_n denote the probability measure on $\mathbb{D}$ defined by

$$\mathrm{d}\eta_n(z) = \frac{n+1}{\pi}\left(1 - |z|^2\right)^n \mathrm{d}\rho_1(z).$$

Recall that ρ_n denotes the Lebesgue measure on $\mathbb{C}^n$.

Lemma 7.3. *Endowing* $M_n^{\mathbb{T}}$ *with the uniform probability measure is equivalent to the* n *first canonical coordinates* p_j, $j = 1, 2, ..., n$ *being independent random variables in such a way that* p_j *has distribution law* η_{n-j}.

Proof. Let f an arbitrary function on $M_n^{\mathbb{T}}$, then

$$\mathbb{E}_n f = C_n \int_{M_n^{\mathbb{T}}} f(\mathbf{t}_n)\, \mathrm{d}\rho_n(\mathbf{t}_n)$$

with $C_n := 1/\rho_n(M_n^{\mathbb{T}})$. Writing the integral in the variables p_i (using (7.2) and (7.4)) we have that

$$\mathbb{E}_n f = C_n \int_{\mathbb{D}^n} f(p_1, ..., p_n) \prod_{j=1}^{n-1}(1 - |p_j|^2)^{n-j}\, \mathrm{d}\rho_1(p_1) ... \mathrm{d}\rho_1(p_n).$$

The proof is completed calculating

$$\int_{\mathbb{D}} (1 - |z|^2)^n\, \mathrm{d}\rho_1(z) = \int_0^{2\pi}\int_0^1 (1 - r^2)^n r\, \mathrm{d}r\, \mathrm{d}\phi = \pi/(n+1).$$

□

The relation (4.1) follows from

$$\mathrm{Volume}_{\mathbb{R}^{2n}}(M_n^{\mathbb{T}}) = \rho_n(M_n^{\mathbb{T}}) = \left[\prod_{j=1}^{n-1}\int_{\mathbb{D}}(1 - |z|^2)^j\, \mathrm{d}\rho_1(z)\right]^{-1}.$$

Lemma 7.4. *Let Y_n be a complex random variable following law η_n. The sequence $(\sqrt{n}Y_n)$ converges in distribution to a normal complex distribution in such a way that the real and imaginary parts are independent with expectation* 0 *and variance* 1/2.

Proof. The density function of $(\sqrt{n}Y_n)$ is

$$\frac{n+1}{n\pi}\left(1-\frac{|z|^2}{n}\right)^n.$$

By the bound $e^{-x} > (1-(x/n))^n$, for all positive integer n and real x it follows

$$\frac{n+1}{n\pi}\left(1-|z|^2\right)^n \leq \frac{2}{\pi}\exp(-|z|^2).$$

In view of the dominated convergence Lebesgue's theorem, the previous bound and the pointwise limit $(n+1)\left(1-|z|^2\right)^n/(n\pi) \underset{n}{\longrightarrow} e^{-|z|^2}/\pi$ (for all $z \in \mathbb{C}$) imply that, for all borelian set of $\mathbb{C}$ A,

$$\lim_n \mathbb{P}(\sqrt{n}Y_n \in A) = \frac{1}{\pi}\int_A e^{-\|\mathbf{z}\|^2}\,\mathrm{d}\rho_1(\mathbf{z}).$$

The last implies the Lemma. □

Proof of Theorem 4.1. Let $Y_n = (Y_{1,n}, Y_{2,n}, ..., Y_{k,n})$ where $Y_{i,n}$, $i = 1, 2, ..., k$ are independent and $Y_{i,n}$ has distribution η_{n-i}, $i = 1, 2, ..., k$. By Lemma 7.4, it follows that $(\sqrt{n}Y_n)$ has a central complex normal limit distribution of variance matrix $1/2I_k$.

By Lemma 7.1 and Lemma 7.3, we may write

$$\sqrt{n}\,X_n^{\mathbb{T},k} \overset{Law}{=} \sqrt{n}Y_n + O(\sqrt{n}\|Y_n\|^2).$$

Therefore, the limit law of $\sqrt{n}\,X_n^{\mathbb{T},k}$ is the same of $\sqrt{n}Y_n$. □

7.3. **Proof of Theorem 4.2.**

Lemma 7.5. *The sequence of distribution laws (η_n) satisfies a LDP with good rate function*

$$J_1(z) = -\ln(1-|z|^2) \quad (z \in \mathbb{C}).$$

Proof. Define, for $z_0 \in \mathbb{D}\setminus\{0\}$, $r_0 > 0$ and $0 < \theta_0 \leq \pi$,

$$V(z_0, r_0, \theta_0) = \{z \in \mathbb{D} : \big||z| - |z_0|\big| < r_0, \quad |\mathrm{Arg}(z) - \mathrm{Arg}(z_0)| < \theta_0\}.$$

Let $z_0 \in \mathbb{D}\setminus\{0\}$, then for $0 < r_0 < |z_0|$ and $0 < \theta_0 \leq \pi$, we have

$$\begin{aligned}\eta_n(V(z_0, r_0, \theta_0)) &:= \int_{V(z_0,r_0,\theta_0)} \frac{n+1}{\pi}\left(1-|z|^2\right)^n \mathrm{d}\nu(z)\\ &\leq \frac{2\theta_0(n+1)}{\pi}\left(1-(|z_0|-r)^2\right)^n.\end{aligned}$$

This leads to the bound

$$\inf_{U\in\mathcal{V}(z_0)}\left\{\limsup_n \frac{1}{n}\ln\eta_n(U)\right\} \leq 1-|z_0|^2. \tag{7.5}$$

This last inequality is obviously also true for $z_0 = 0$. Now, for $z_0 \in \mathbb{D} \setminus \{0\}$ and $U \in \mathcal{V}(z_0)$ with $U \subset \mathbb{D}$, there exists $r_0 > 0$ and $0 < \theta_0 \leq \pi$ such that $V(z_0, r_0, \theta_0) \subset U$. Then

$$\begin{aligned}\eta_n(U) &\geq \eta_n(V(z_0, r_0, \theta_0)) \\ &\geq \frac{\theta_0(n+1)}{\pi} \left(1 - (|z_0| + r)^2\right)^n .\end{aligned}$$

Consequently

$$\inf_{U \in \mathcal{V}(z_0)} \left\{ \liminf_n \frac{1}{n} \ln \eta_n(U) \right\} \geq 1 - |z_0|^2 \tag{7.6}$$

for $z_0 \neq 0$. For $U \in \mathcal{V}(0)$ there exists $r_0 > 0$ such that $\{|z| < r_0\} \subset U$ then, for $r < r_0$,

$$\eta_n(U) \geq \eta_n(\{|z| < r\}) \geq (n+1) r^2 (1 - r^2)^n.$$

So the inequality (7.6) is also true for $z_0 = 0$. In view of Theorem 4.1.11 of [DZ98], (7.5) and (7.6) we have the weak LDP. The full LDP follows from the fact that the distributions η_n are supported in the compact set $\mathbb{D}$. $\square$

Exercise 4.2.7 of [DZ98], Lemma 7.3 and Lemma 7.5 imply that the sequence of random variables $\left(\mathbf{p}_k(X_n^k) : n \in \mathbb{N}\right)$ satisfies the LDP on $\mathbb{D}^k$ with the good rate function

$$J_2^k(\mathbf{z}_k) = \begin{cases} -\sum_{i=1}^k \ln(1 - |z_i|^2) & \text{if } \mathbf{z}_k \in \operatorname{int} \mathbb{D}^k, \\ +\infty & \text{otherwise;} \end{cases}$$

where $\mathbf{z}_k = (z_1, z_2, ..., z_k) \in \mathbb{C}^{\,k}$.

Now, in view of the contraction principle (write X_n^k as $\mathbf{p}_k^{-1}(\mathbf{p}_k(X_n^k))$ and use that $\mathbf{p}_k^{-1}(\cdot)$ is a continuous bijection), we have the LDP for (X_n^k) with the good rate function

$$I_{\mathbb{T}}^{k(\mathbf{t}k)} = \begin{cases} -\ln \prod_{i=1}^k (1 - |z_i|^2) & \text{if } \mathbf{t}k \in \operatorname{int} M_k^{\mathbb{T}}, \\ +\infty & \text{otherwise,} \end{cases}$$

for $\mathbf{t}k = \mathbf{p}_k^{-1}(\mathbf{z}_k)$. The expression (4.2) for $I_{\mathbb{T}}^k$ follows from the relation (7.3).

7.4. **Proof of Theorem 4.3.**

Lemma 7.6. *Let Y_n be a random variable having distribution η_{n-k} ($k \leq n$). Let*

$$\widetilde{Y}_n := \sqrt{n u_n}\, Y_n.$$

Then $(\widetilde{Y}_n)$ satisfies a LDP on $\mathbb{C}$ with good rate function

$$J_1(z) = |z|^2$$

and speed (u_n).

Proof. Let $z_0 \in \mathbb{C} \setminus \{0\}$, $0 < r_0 < |z_0|$ and $0 < \theta_0 \leq \pi$ then we have

$$\begin{aligned}\mathbb{P}(\widetilde{Y}_n \in V(z_0, r_0, \theta_0)) =& \mathbb{P}\left(Y_n \in V\left(\frac{z_0}{\sqrt{n u_n}}, \frac{r_0}{\sqrt{n u_n}}, \theta_0\right)\right) \\ =& \frac{\theta_0(n-k+1)}{\pi} \int_{(n u_n)^{-1}(|z_0| - r_0)^2}^{(n u_n)^{-1}(|z_0| + r_0)^2} (1 - \rho)^{n-k} \, \mathrm{d}\rho.\end{aligned}$$

Consequently

$$\mathbb{P}(\widetilde{Y}_n \in V(z_0, r_0, \theta_0)) \leq \frac{\theta_0(n-k+1)}{\pi n u_n} r_0 |z_0| \left(1 - \frac{(|z_0| - r_0)^2}{n u_n}\right)^{n-k}.$$

Then

$$\limsup_n u_n \ln \mathbb{P}(\widetilde{Y}_n \in V(z_0, r_0, \theta_0)) \leq -(|z_0| - r_0)^2.$$

This leads to

$$\inf_{U \in \mathcal{V}(z_0)} \left\{ \limsup_n u_n \ln \mathbb{P}(\widetilde{Y}_n \in U) \right\} \leq -|z_0|^2, \quad z_0 \in \mathbb{C}. \tag{7.7}$$

In the other hand, using similar arguments, we obtain

$$\inf_{U \in \mathcal{V}(z_0)} \left\{ \liminf_n u_n \ln \mathbb{P}(\widetilde{Y}_n \in U) \right\} \geq -|z_0|^2, \quad z_0 \in \mathbb{C}. \tag{7.8}$$

In view of Theorem 4.1.11 of [DZ98], (7.7) and (7.8) we have the weak LDP. In order to have the full LDP it suffices to show that the distributions of $\widetilde{Y}_n$, $(n \in \mathbb{N})$ are exponentially tight. Let $0 < r < \sqrt{n u_n}$, then

$$\mathbb{P}(|\widetilde{Y}_n| \leq r) = 1 - \left(1 - \frac{r^2}{n u_n}\right)^{n-k+1}$$

and consequently

$$\limsup_n u_n \ln \mathbb{P}(\widetilde{Y}_n > r) \leq -r^2.$$

The arbitrariness of r implies the exponential tightness. $\square$

For all $n > k$, let

$$\widehat{Z}_n^k k := \sqrt{n u_n} \mathbf{p}_k \left(X_n^k\right),$$
$$Z_n^k k := \sqrt{n u_n}\, X_n^k.$$

By Exercise 4.2.7 of [DZ98] and Lemma 7.6 we have a LDP for the sequence of random vectors $\left(\widehat{Z}_n^k k : n \in \mathbb{N}\right)$ with rate function

$$H_{\mathbb{T}}^k(\mathbf{p}k) = \sum_{i=1}^k J_1(p_i) = \sum_{i=1}^k p_i^2, \quad (\mathbf{p}k = (p_1, p_2, ..., p_k) \in \mathbb{C}^{\ k}).$$

A LDP for $\left(Z_n^k k : n \in \mathbb{N}^*\right)$ follows from the following Lemma.

Lemma 7.7. *The random vector sequence $\left(\widehat{Z}_n^k k\right)$ and $\left(Z_n^k k\right)$ are exponentially equivalent.*

Proof. By Lemma 7.1 there are constants $\epsilon, Q > 0$ such that $\|X_n^k - \mathbf{p}_k\left(X_n^k\right)\| \leq Q \|\mathbf{p}_k\left(X_n^k\right)\|^2$ whenever $\|\mathbf{p}_k\left(X_n^k\right)\| \leq \epsilon$. Therefore, for any $\xi > 0$,

$$\begin{aligned}
\mathbb{P}\left(\|\widehat{Z}_n^k k - Z_n^k k\| > \xi\right) =& \mathbb{P}\left(\|\widehat{Z}_n^k k - Z_n^k k\| > \xi, \|\mathbf{p}_k(X_n^k)\| \leq \epsilon\right) \\
&+ \mathbb{P}\left(\|\widehat{Z}_n^k k - Z_n^k k\| > \xi, \|\mathbf{p}_k(X_n^k)\| > \epsilon\right) \\
\leq& \mathbb{P}\left(\|\widehat{Z}_n^k k\|^2 > \sqrt{n u_n} \xi Q^{-1}\right) + \mathbb{P}\left(\|\widehat{Z}_n^k k\| > \sqrt{n u_n} \epsilon\right).
\end{aligned}$$

Let $Q_0 > 0$. For n big enough we have $\sqrt{nu_n} > Q_0$ and consequently

$$\mathbb{P}\Big(\|\widehat{Z}_n^k k\| > \sqrt{nu_n}\xi Q^{-1}\Big) \le \mathbb{P}\Big(\|\widehat{Z}_n^k k\|^2 > \Big(\frac{\xi Q_0}{Q}\Big)^{1/2}\Big)$$
$$\mathbb{P}\big(\|\widehat{Z}_n^k k\| > \sqrt{nu_n}\epsilon\big) \le \mathbb{P}\big(\|\widehat{Z}_n^k k\| > Q_0\epsilon\big).$$

Let $q(Q_0) := \min\{Q_0\epsilon, \sqrt{\xi Q_0 Q^{-1}}\}$. By the LDP for $\big(\widehat{Z}_n^k k\big)$ with function J_2^k we have

$$\begin{aligned}\limsup_n u_n \ln \mathbb{P}\big(\|\widehat{Z}_n^k k - Z_n^k k\| > \xi\big) &\le \limsup_n u_n \ln(\|\widehat{Z}_n^k k\| > q(Q_0))\\ &\le -q(Q_0).\end{aligned}$$

From the fact that $Q_0 \to \infty$ yields $q(Q_0) \to \infty$ follows the exponential tightness. □

7.5. **Proof of Theorem 4.4.** Let $M_\infty^{\mathbb{T}}$ denote the *space of infinite moment sequences* defined as

$$M_\infty^{\mathbb{T}} = \big\{\mathbf{t} = (t_1, t_2, ...) : (t_1, t_2, ..., t_n) \in M_n^{\mathbb{T}}, \quad \text{for all } n \in \mathbb{N}^*\big\}.$$

It is well-know the bijection that exists between this space and $\mathcal{P}(\mathbb{T})$. More exactly, given $\mathbf{t} = (t_1, t_2, ...) \in M_\infty^{\mathbb{T}}$ there exists a unique measure $\mu_{\mathbf{t}}$ in $\mathcal{P}(\mathbb{T})$ such that $t_j(\mu_{\mathbf{t}}) = t_j$ for all j [ST43]. Moreover, it can be easily proved that the map $\mathbf{t} \mapsto \mu_{\mathbf{t}}$ is continuous.

Equip $M_\infty^{\mathbb{T}}$ with the product algebra. For $n \in \mathbb{N}^*$, let $\widetilde{\mathbb{Q}}_n \in \mathcal{P}(M_\infty^{\mathbb{T}})$ be the measure image of $\mathbb{Q}_n$ by the bijection $\mathbf{t} \mapsto \mu_{\mathbf{t}}$. Let $\Pi_{\infty,k}^{\mathbb{T}}$ denote the projection map from $M_\infty^{\mathbb{T}}$ to $M_k^{\mathbb{T}}$. We have that $\widetilde{\mathbb{Q}}_n \circ \Big(\Pi_{\infty,k}^{\mathbb{T}}\Big)^{-1}$ is exactly the law of X_n^k. Then $\left(\widetilde{\mathbb{Q}}_n \circ \Big(\Pi_{\infty,k}^{\mathbb{T}}\Big)^{-1} : n \in \mathbb{N}^*\right)$ satisfies the LDP (Theorem 4.2) with good rate function $I_{\mathbb{T}}^k$. By Dawson-Gärtner's Theorem we have a LDP for $\big(\widetilde{\mathbb{Q}}_n\big)$ on $M_\infty^{\mathbb{T}}$ with good rate function

$$\begin{aligned} I_{\mathbb{T}}^\infty(\mathbf{t}) &= \sup_{k\in\mathbb{N}} I_{\mathbb{T}}^k(\Pi_{\infty,k}^{\mathbb{T}}\mathbf{t}_\infty)\\ &= \begin{cases} \sup\limits_{k\in\mathbb{N}} -\ln\left(\dfrac{\Delta_k(\Pi_{\infty,k}^{\mathbb{T}}\mathbf{t})}{\Delta_{k-1}(\Pi_{\infty,k-1}^{\mathbb{T}}\mathbf{t})}\right) & \text{if } \mathbf{t} \in \operatorname{int} M_\infty^{\mathbb{T}},\\ +\infty & \text{otherwise.}\end{cases}\end{aligned}$$

From (7.3) we have that

$$\sup_k -\ln \frac{\Delta_k(\Pi_{\infty,k}^{\mathbb{T}}\mathbf{t})}{\Delta_{k-1}(\Pi_{\infty,k-1}^{\mathbb{T}}\mathbf{t})} = \lim_k -\ln \frac{\Delta_k(\Pi_{\infty,k}^{\mathbb{T}}\mathbf{t})}{\Delta_{k-1}(\Pi_{\infty,k-1}^{\mathbb{T}}\mathbf{t})}.$$

Now, by the Grenander-Szegö's Theorem [GS58, §5.2] we have

$$\lim_k -\ln \frac{\Delta_k(\Pi_{\infty,k}^{\mathbb{T}}\mathbf{t})}{\Delta_{k-1}(\Pi_{\infty,k-1}^{\mathbb{T}}\mathbf{t})} = \int_{\mathbb{T}} \ln \frac{d\nu^{\mathbb{T}}}{d\mu_{\mathbf{t}}}\, d\nu^{\mathbb{T}}.$$

In view of the contraction principle with the continuous map $\mathbf{t} \mapsto \mu_{\mathbf{t}}$ we have a LDP for $\big(\mathbb{Q}_n\big)$.

The proof of the Corollary 4.5 follows from the LDP for $(\mathbb{Q}_n)$ using the last expression for the rate function and the contraction principle on other sense.

Acknowledgment. I thank Professor Fabrice Gamboa for his advice. These paper is based on research I did toward my Ph.D. title under Gamboa's supervision. I also thank Professor Alain Rouault for refering to me the papers of J. C. Gupta. Finally I thank Professor Mauro Piccioni for useful comments about this subject.

References

[Bil99] P. Billingsley. *Convergence of probability measures.* John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[BL93] J. M. Borwein and A. S. Lewis. Partially-finite programming in L_1 and the existence of maximum entropy estimates. *SIAM J. Optim.*, 3(2):248–267, 1993.

[Bre79] J. Bretagnolle. Formule de Chernoff pour les lois empiriques de variables a valeurs dans des espaces generaux. *Asterisque*, 68:32–52, 1979.

[CKS93] F. C. Chang, J. H. B. Kemperman, and W. J. Studden. A normal limit theorem for moment sequences. *Ann. Probab.*, 21(3):1295–1309, 1993.

[DS97] H. Dette and W. J. Studden. *The theory of canonical moments with applications in statistics, probability, and analysis.* John Wiley & Sons Inc., New York, 1997. A Wiley-Interscience Publication.

[Dud02] R. M. Dudley. *Real analysis and probability.* Cambridge Studies in Advanced Mathematics.Cambridge University Press, 2002.

[DZ98] A. Dembo and O. Zeitouni. *Large deviations techniques and applications.* Springer-Verlag, New York, second edition, 1998.

[GG97] F. Gamboa and E. Gassiat. Bayesian methods and maximum entropy for ill-posed inverse problems. *Ann. Statist.*, 25(1):328–350, 1997.

[GLC04] F. Gamboa and L. Lozada-Chang. Large deviation for random power moment problem. *Ann. Probab.*, 34(3B), 2004.

[GS58] U. Grenander and G. Szegö. *Toeplitz forms and their applications.* University of California Press, Berkeley, 1958.

[Gup99a] J. C. Gupta. The moment problem for the standard k-dimensional simplex. *Sankhyā Ser. A*, 61(2):286–291, 1999.

[Gup99b] J. C. Gupta. Partial Hausdorff sequences and symmetric probabilities on finite products of $\{0, 1\}$. *Sankhyā Ser. A*, 61(3):347–357, 1999.

[Gup00] J. C. Gupta. Completely monotone multisequences, symmetric probabilities and a normal limit theorem. *Proc. Indian Acad. Sci. Math. Sci.*, 110(4):415–430, 2000.

[KN77] M. G. Kreĭn and A. A. Nudelman. *The Markov moment problem and extremal problems.* American Mathematical Society, Providence, R.I., 1977. Ideas and problems of P. L. Chebyshev and A. A. Markov and their further development, Translated from the Russian by D. Louvish, Translations of Mathematical Monographs, Vol. 50.

[Naj02] J. Najim. A Cramér type theorem for weighted random variables. *Electron. J. Probab.*, 7(4):32 pp (electronic), 2002.

[Ski68] M. Skibinsky. Minimax estimation of a random probability whose first n moments are known. *Ann. Math. Statist.*, 39(2):492–501, 1968.

[Ski69] M. Skibinsky. Some striking properties of binomial and beta moments. *Ann. Math. Statist.*, 40:1753–1764, 1969.

[ST43] J. A. Shohat and D. J. Tamarkin. *The problem of moments*, volume 1 of *Mathematical Surveys.* American Mathematical Society (AMS), Providence, R.I., 1943.

[VdV98] A.W. Van der Vaart. *Asymptotic statistics.* Cambridge Series on Statistical and Probabilistic Mathematics,3., Cambridge Univ. Press., 1998.