Limit Theorems on Hypergroups

DISSERTATION

Zur Erlangung des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt
der Fakultät für Mathematik
der Technischen Universität Dortmund

von
Dipl.-Math. Waldemar Grundmann

Dortmund, November 2012

To my wife Natalie and my son Daniel

Acknowledgments

For the successful completion of this work, I wish to thank my thesis advisor Prof. Dr. Michael Voit. He suggested this interesting topic to me and gave me constant encouragement and inspiration throughout the development of this thesis. I am much indebted to Prof. em. Dr. Wilfried Hazod for agreeing to review this work. I also wish to thank all of my colleagues from the chair “Analysis und Stochastik” for their assistance, friendship and many happy and enjoyable times.

Last but not least, I would like to thank my wife Natalie for her love and for tirelessly supporting and enduring me at all times.

Contents

1 Introduction **1**

2 Random walks on hypergroups **4**
2.1 Preliminaries 4
2.2 Concretization of hypergroups 14
2.3 Moments on hypergroups 18

3 Moment functions and limit theorems for Jacobi hypergroups **28**
3.1 Hyperbolic spaces 28
3.2 Jacobi functions and Jacobi hypergroups 30
3.3 Moment functions on Jacobi hypergroups 33
3.4 Limit theorems for growing parameters 38

4 Radial limit theorems on $\mathbb{R}^p$ for growing dimensions p **45**

5 Radial random walks on $p \times q$ matrices for $p \to \infty$ **68**
5.1 Matrix algebra 70
5.2 Permutations on a multiset 72
5.3 Bessel functions of matrix argument 74
5.4 Polynomials of matrix argument 75
5.5 Random Matrices 76
5.6 Method of moments for random matrices 79
5.7 Matrix variate normal distribution 82
5.8 Radial measures on $\mathbb{M}_{p,q}$ and their moments 87
5.9 Radial limit theorems on $\mathbb{M}_{p,q}$ for $p \to \infty$ 93

Bibliography **108**

Chapter 1

Introduction

The study of limit theorems on hypergroups began in the 1960s with Haldane's and Kingman's articles [15] and [23]. They studied methods which allowed the investigation of rotation-invariant random vectors and generalised them to non-integral "dimensions". For this purpose, a continuous series of commutative hypergroups on $[0, \infty[$ was introduced. These hypergroups structures are closely related with a product formula for Bessel functions and are therefore called Bessel-Kingman hypergroup; c.f. Section 3.4 of [6]. In 1976, the mathematical structure of a hypergroup was reintroduced, under the name of a "convo," and systematically studied by R.I. Jewett [19] in his article "Spaces with an Abstract Convolution of Measures". It is unsurprising that probability theory on this class of hypergroups is so well developed.

Let $(\mathrm{K}, *)$ be a hypergroup in the sense of Jewett [19]. The convolution $*$ allows the notion of random walks on $(\mathrm{K}, *)$ by saying that a (time-homogeneous) Markov chain $(S_n)_{n\geq 0}$ is a random walk on $(\mathrm{K}, *)$ with law $\nu \in \mathcal{M}^1(\mathrm{K})$ if

$$\mathbb{P}(S_{n+1} \in A | S_n = x) = (\delta_x * \nu)(A) \tag{1.0.1}$$

for all $n \geq 0$, $x \in \mathrm{K}$ and Borel sets $A \subset \mathrm{K}$. A lot of research was carried out in this setting, such as, for example, on recurrence, laws of large numbers (LLNs), large deviation principles (LDPs), and central limit theorems (CLTs); see e.g. [5], [23], and [38]-[43]. However, there are also many issues where the representation of S_n as a sum of independent and identically distributed (i.i.d.) random variables is useful. Typical examples are laws of large numbers where truncation methods are used, see [46] and Section 7.3 of [6].

On a hypergroup, there is in general no deterministic operation corresponding to the convolution of measures as it is in the classical case of locally compact Hausdorff groups. Consequently, sums of hypergroup-valued random variables cannot be defined directly. This obstacle was overcome in a particular case by Kingman [23]. He applied the concept of randomized addition, in such a way that, as in the classical case, the distribution of the sum of two independent K-valued random variables equals the convolution of the distributions of the summands. Later, this construction was generalised by Zeuner [46] under the name of concretisation.

In the study of limit theorems, the modified moments of a random variable, which are adapted to the hypergroup operation, were introduced to formulate the conditions under which a particular limit theorem holds, and to calculate the actual value of the limit. The notions of moments of the first and second order and of dispersion were introduced for special cases by Tutubalin [37]. Later, the idea of dispersion appeared in the work of Faraut [9] and Trimèche [36]. The modified moment functions on Sturm-Liouville hypergroups and polynomial hypergroups were studied by Zeuner [46] and Voit [38], respectively. A systematic study of this subject on an arbitrary hypergroup has been carried out by Zeuner [47].

Today, there is a wide range of limit theorems for random walks on hypergroups in general as well as for special cases available. For an overview on these results, we refer to the monograph [6] and references cited there.

Outline of this Thesis and main Results

In this thesis, we shall investigate a special type of limit theorem for some classes of hypergroups. To describe the common roots of these limit theorems, let us consider the following example: Let $\nu \in \mathcal{M}^1([0,\infty[)$ be a fixed probability measure. Then for each dimension $d \in \mathbb{N}$ there is a unique rotation-invariant probability measure $\nu_d \in \mathcal{M}^1(\mathbb{R}^d)$ with $\varphi_d(\nu_d) = \nu$, where $\varphi_d(x) = \|x\|_2$ is the norm mapping. For each $d \in \mathbb{N}$ consider i.i.d. $\mathbb{R}^d$-valued random variables X_k^d, $k \in \mathbb{N}$, with law ν_d as well as the associated radial random walks $(S_n^d := \sum_{k=1}^n X_k^d)_{n\geq 0}$ on $\mathbb{R}^d$. The process $(\|S_n^d\|_2)_{n\geq 1}$ is a random walk on the so-called Bessel-Kingman hypergroup $([0,\infty[, *_\alpha)$ of parameter $\alpha := d/2 - 1$. The aim is to find limit theorems for the random walks $(S_n^\alpha)_{n\geq 0}$ on the Bessel-Kingman hypergroup $([0,\infty[, *_\alpha)$ for n, $\alpha \to \infty$ in a suitably coupled way.

More generally, let $(\mathrm{K}, *_\alpha)$ be a sequence of hypergroups, where the convolution $*_\alpha$ depends on a parameter $\alpha \in I$ with $I \subset \mathbb{R}$. For each “dimension” parameter α consider a time-homogeneous Markov-chain $(S_n^\alpha)_{n\geq 1}$ on the hypergroup $(\mathrm{K}, *_\alpha)$.

We now ask for limit theorems for $(S_n^\alpha)_{n\geq 1}$ when n, $\alpha \to \infty$ in a suitably coupled way.

This thesis is set out in three main sections. In Chapter 2, after recalling necessary preliminaries on hypergroups in general, we derive in Proposition 2.3.9 useful algebraic identities for the modified moments of randomized sums. Moreover, we establish some inequalities between moment functions of the first and second order (see 2.3.10). As an application of these inequalities, we prove a weak LLN and the associated strong LLN for randomized sums on a general hypergroup in an efficient manner.

In Chapter 3, we present sharp estimates and asymptotic results for moment functions on so-called hypergroups of the Jacobi type on $[0, \infty[$ of index (α, β). Further, we use these estimates to prove a central limit theorem for random walks $(S_n^{(\alpha,\beta)})_{n\geq 1}$ on Jacobi hypergroups $([0, \infty[, *_{\alpha,\beta})$ with growing parameter $\alpha \to \infty$. As a special case, we obtain a CLT for radial, time-homogeneous random walks on hyperbolic spaces $\mathrm{H}_d(\mathbb{F})$ of growing dimension d over the fields $\mathbb{F} = \mathbb{R}$, $\mathbb{C}$ or the quaternions $\mathbb{H}$. In addition to these results, we also derive associated strong laws of large numbers.

Chapter 4 is devoted to the study of rotation invariant random walks $(S_n^d)_{n\geq 1}$ on $\mathbb{R}^d$, which are defined as in the representative example from above. Here, for an arbitrary sequence of dimensions $d_n \to \infty$, we investigate the asymptotic behaviour of random variables $\|S_n^{d_n}\|_2$ as $n \to \infty$. In this chapter we derive two complementary CLTs for the functional $\|S_n^{d_n}\|_2$ with normal limits under disjoint growth conditions for d_n, namely for $n/d_n \to \infty$ and $n/d_n \to 0$. Moreover, we present a CLT for the case $n/d_n \to c \in (0, \infty)$. Besides these results, we shall also prove associated strong laws of large numbers. An essential ingredient for the proofs of all limit results is an explicit formula for moments of radial distributed random variables on $\mathbb{R}^d$ with $d \to \infty$.

In the last chapter we deal with orthogonal invariant random walks on the space $\mathbb{M}_{p,q}(\mathbb{R})$ of $p \times q$ matrices instead of $\mathbb{R}^p$ for $p \to \infty$ and some fixed dimension $q > 1$. The main result in this chapter generalises CLTs from Chapter 4 and is presented in Theorem 5.9.1. The proof of this result relies on asymptotic formulas for moment functions of orthogonal invariant probability measures on $\mathbb{M}_{p,q}$ as well as on some identities for matrix-variate normal distributions.

Part of this thesis has been accepted by the Journal of Theoretical Probability and is due to appear under the title "Moment Functions and Central Limit Theorem for Jacobi Hypergroups on $[0, \infty[$".

Chapter 2

Random walks on hypergroups

In this chapter we first give a brief introduction to the concepts of concretisation and randomized sums, which are fundamental for the construction of random walks on hypergroups. Although on a hypergroup is in general no deterministic pointwise operation, these concepts enable us to construct "sums" of hypergroup valued variables which are consistent with the underlying convolution structure. Furthermore, we recall the concept of moment functions on hypergroups, which we then use to study limit behaviour of partial randomized sums S_n of hypergroup-valued random variables. In particular, we derive weak LLNs and the associated strong LLNs for S_n as $n \to \infty$.

2.1 Preliminaries

In this section we collect the definitions and properties of hypergroups which we will need throughout this work. The material is mainly taken from [6]. For a general background on hypergroups, the reader is referred to the fundamental article [19].

For a nonvoid locally compact Hausdorff space K let $\mathcal{M}^1(\mathrm{K})$ and $\mathcal{M}_b(\mathrm{K})$ denote the set of probability measures on the Borel algebra $\mathcal{B}(\mathrm{K})$ and the set of bounded Radon measures, respectively. The space of continuous functions on K is denoted by $\mathcal{C}(\mathrm{K})$. Superscripts are used to declare differentiability while subscripts are used to declare other properties: $\mathcal{C}^k(\mathrm{K})$ are the functions which are k-times continuously differentiable. $\mathcal{C}_b(\mathrm{K})$ resp. $\mathcal{C}_c(\mathrm{K})$ are continuous functions which are bounded resp. with compact support. The combination of upper and lower indices is to be read as follows: $\mathcal{C}_b^2(\mathrm{K})$ are the bounded twice continuously differentiable functions.

Definition 2.1.1. Let K be a a nonvoid locally compact Hausdorff space and $*$ a bilinear, binary operation on $\mathcal{M}_b(\mathrm{K})$, such that $(\mathcal{M}_b(\mathrm{K}), +, *)$ is an algebra. Then $(\mathrm{K}, *)$ will be called a hypergoup if the following conditions are satisfied:

(H1) Given $x, y \in \mathrm{K}$, $\delta_x * \delta_y \in \mathcal{M}^1(\mathrm{K})$ and $\mathrm{supp}(\delta_x * \delta_y)$ is compact.

(H2) The mapping $(x, y) \mapsto \delta_x * \delta_y$ of $\mathrm{K} \times \mathrm{K}$ into $\mathcal{M}^1(\mathrm{K})$ is continuous with respect to the weak topology on $\mathcal{M}^1(\mathrm{K})$.

(H3) The mapping $(x, y) \mapsto \mathrm{supp}(\delta_x * \delta_y)$ is continuous from $\mathrm{K} \times \mathrm{K}$ into the space of compact subsets of K provided with the *Michael topology*, cf. [6, Section 1.1] and [27].

(H4) There exists a neutral element e of K such that $\delta_x * \delta_e = \delta_e * \delta_x = \delta_x$ for all $x \in \mathrm{K}$.

(H5) There exists an involution (a homeomorphism $x \mapsto \check{x}$ of K onto itself with the property $(\check{x})^\vee = x$ for all $x \in \mathrm{K}$) such that $(\delta_x * \delta_y)^\vee = \delta_{\check{y}} * \delta_{\check{x}}$ for all $x, y \in \mathrm{K}$ where $\mu^\vee$ denotes the image of μ under this involution.

(H6) For $x, y \in \mathrm{K}$, $e \in \mathrm{supp}(\delta_x * \delta_y)$ if and only if $x = \check{y}$.

The operation $*$ will be called *convolution*. A hypergroup K is called *commutative* if $\delta_x * \delta_y = \delta_y * \delta_x$ for $x, y \in \mathrm{K}$; i.e. if $(\mathcal{M}_b(\mathrm{K}), +, *)$ is a commutative algebra. A hypergroup for which the involution is the identity mapping is called *hermitian*. It is evident that every hermitian hypergroup is commutative, since

$$\delta_x * \delta_y = (\delta_x * \delta_y)^\vee = \delta_{\check{y}} * \delta_{\check{x}} = \delta_y * \delta_x.$$

We will often write for a hypergroup $(\mathrm{K}, *)$ simply K when no confusion can arise. Moreover, we will denote the n-fold convolution power of a measure $\nu \in \mathcal{M}^1(\mathrm{K})$ with respect to the convolution $*$ by ν^n. It is clear that the neutral element e and the involution $\vee$ are necessarily unique. Furthermore the convolution product of a hypergroup is uniquely determined if $\delta_x * \delta_y$ is known for all $x, y \in \mathrm{K}$. Obviously, conditions (H4) and (H5) carry over to arbitrary bounded measures instead of the point masses δ_x and δ_y. Namely, for μ, $\nu \in \mathcal{M}_b(\mathrm{K})$ we have

$$\mu * \nu = \int_{\mathrm{K}} \int_{\mathrm{K}} \delta_x * \delta_y \mathrm{d}\mu(x) \mathrm{d}\nu(y),$$

see [19] Sections 2.3-2.5 for more details.

Remark 2.1.2. **a)** Every locally compact group G is a hypergroup with the usual convolution structure on $\mathcal{M}_b(G)$.

b) In general, hypergroups have no algebraic structure of their own.

Definition 2.1.3. A nonzero measure $\omega_{\mathrm{K}} \in \mathrm{K}$ is called *left* (resp. *right*) *Haar measure* if $\delta_x * \omega_{\mathrm{K}} = \omega_{\mathrm{K}}$ ($\omega_{\mathrm{K}} * \delta_x = \omega_{\mathrm{K}}$) holds for all $x \in \mathrm{K}$. A left and right Haar measure is called Haar measure.

The following theorem ensures the existence of Haar measure for two large classes of hypergroups.

Theorem 2.1.4. *Let* K *be commutative or compact hypergroup. Then* K *admits a Haar measure.*

It is well known that a Haar measure on an arbitrary hypergroup is unique up to a positive multiplicative constant. We remark that if K is a compact hypergroup and $\omega_{\mathrm{K}}(\mathrm{K}) = 1$ then $\omega_{\mathrm{K}} * \omega_{\mathrm{K}} = \omega_{\mathrm{K}}$, that is the *normalized Haar measure* of a compact hypergroup is an idempotent.

Definition 2.1.5. A closed nonvoid subset H of a hypergroup K will be called a *subhypergroup* if $\check{\mathrm{H}} = \mathrm{H}$ and $\mathrm{H} * \mathrm{H} \subset \mathrm{H}$.

It is clear that $e \in \mathrm{H}$ for every subhypergroup H. If μ, $\nu \in \mathcal{M}^b(\mathrm{H})$ then μ, ν can be regarded as members of $\mathcal{M}^b(\mathrm{K})$, and since H is closed under convolution it follows that $\mu * \nu$ may be regarded as a member of $\mathcal{M}^b(\mathrm{H})$. With this operation H is also a hypergroup having the same identity and involution mapping as K.

For commutative hypergroups the powerful tool of the Fourier transform is available. As in the classical case, its enables to describe the weak convergence of probability measures.

Definition 2.1.6. Let K be a commutative hypergroup. A continuous function $\chi : \mathrm{K} \to \mathbb{C}$ is called *multiplicative* if

a) $\chi(e) = 1$ and

b) $\chi(x * y) = \chi(x)\chi(y)$ for all x, $y \in \mathrm{K}$.

If in addition χ is bounded and $\chi(\check{x}) = \overline{\chi(x)}$ for all $x \in \mathrm{K}$ then χ is called a *character*. The dual $\hat{\mathrm{K}}$ of K is the set of characters on K and endowed with the topology of the uniform convergence on compact sets a locally compact Hausdorff space.

Definition 2.1.7. Let K be a commutative hypergroup with Haar measure ω_{K}.

a) The *Fourier transform* $\hat{\mu}$ of a measure $\mu \in \mathcal{M}_b(\mathrm{K})$ is defined as the complex valued continuous function on $\hat{\mathrm{K}}$ by

$$\hat{\mu}(\chi) := \int_{\mathrm{K}} \overline{\chi} \mathrm{d}\mu, \quad \chi \in \hat{\mathrm{K}}.$$

The map $\mu \mapsto \hat{\mu}$, $\mathcal{M}_b(\mathrm{K}) \to \mathcal{C}_b(\hat{\mathrm{K}})$ is known also as the Fourier transform (on K).

b) The *Fourier transform* of a function $f \in L^1(\mathrm{K})$ is defined on $\hat{\mathrm{K}}$ by

$$\hat{f}(\chi) := (f\omega_{\mathrm{K}})^\wedge(\chi) = \int_{\mathrm{K}} f\overline{\chi} \mathrm{d}\omega_{\mathrm{K}}, \quad \chi \in \hat{\mathrm{K}}.$$

Theorem 2.1.8 (Levitan-Plancherel)**.** *Let* K *be a commutative hypergroup with Haar measure* ω_{K}*. There exists only one possible measure* $\pi_{\mathrm{K}} \in \mathcal{M}_+(\hat{\mathrm{K}})$*, such that*

$$\int_{\mathrm{K}} |f|^2 \, \mathrm{d}\omega_{\mathrm{K}} = \int_{\hat{\mathrm{K}}} \left|\hat{f}\right|^2 \mathrm{d}\pi_{\mathrm{K}}$$

for all $f \in L^1(\mathrm{K}) \cap L^2(\mathrm{K})$*. The measure* π_{K} *is referred to as the Plancherel measure of the hypergroup* K*.*

Definition 2.1.9. Let K be a commutative hypergroup with Haar measure ω_{K} and Plancherel measure π_{K}.

a) The *inverse Fourier transform* of $\sigma \in \mathcal{M}_b(\hat{\mathrm{K}})$ is defined on K by

$$\check{\sigma}(x) := \int_{\hat{\mathrm{K}}} \chi(x) \mathrm{d}\sigma(\chi), \quad x \in \mathrm{K}.$$

b) The *inverse Fourier transform* of a function $f \in L^1(\hat{\mathrm{K}})$ is defined on K by

$$\check{f}(x) := \int_{\hat{\mathrm{K}}} f(\chi)\chi(x) \mathrm{d}\pi_{\mathrm{K}}(\chi), \quad x \in \mathrm{K}.$$

For the rest of this section we give some concrete examples of hypergroups and their properties, which shall be required later.

Double coset hypergroups

Let K be a hypergroup and $\mathrm{H} \subset \mathrm{K}$ a compact subhypergroup with Haar measure $\omega_{\mathrm{H}} \in \mathcal{M}^1(\mathrm{H})$. The *double cosets* of H are just the sets $\mathrm{H} * \{x\} * \mathrm{H}$ where $x \in \mathrm{K}$.

For simplicity of notation, we denote them by HxH. The collection of double cosets of H will be denoted by K//H, i.e.

$$\mathrm{K}//\mathrm{H} := \{\mathrm{H}x\mathrm{H} : \quad x \in \mathrm{K}\}.$$

It is clear that K//H is a decomposition of K into compact subsets. The collection K//H will be given the quotient topology with respect to the natural projection defined by

$$p_{\mathrm{H}} : \mathrm{K} \to \mathrm{K}//\mathrm{H}, \quad x \mapsto \mathrm{H}x\mathrm{H}. \tag{2.1.1}$$

With this topology K//H is closed in $\mathcal{C}(\mathrm{K})$ and hence locally compact. The projection p_{H} is open, proper and closed. Moreover, p_{H} can be extended to a mapping

$$p_{\mathrm{H}} : \mathcal{M}_b(\mathrm{K}) \to \mathcal{M}_b(\mathrm{K}//\mathrm{H}), \quad \mu \mapsto p_{\mathrm{H}}(\mu)$$

where $p_{\mathrm{H}}(\mu)$ is the image measure of μ under p_{H}. The mapping p_{H} restricted to the space of H-invariant bounded measures on K

$$\mathcal{M}^b_{\mathrm{H}}(\mathrm{K}) := \{\mu \in \mathcal{M}_b(\mathrm{K}) : \quad \omega_H * \mu * \omega_{\mathrm{H}} = \mu\}$$

is bijective. The inverse of this restricted mapping is the unique continuous positive linear mapping $p^*_{\mathrm{H}} : \mathcal{M}_b(\mathrm{K}//\mathrm{H}) \to \mathcal{M}^b_{\mathrm{H}}(\mathrm{K})$ that satisfies $p^*_{\mathrm{H}}(\delta_{\mathrm{H}x\mathrm{H}}) = \omega_{\mathrm{H}} * \delta_x * \omega_{\mathrm{H}}$ for each $x \in \mathrm{K}$. A convolution structure can be defined on $\mathcal{M}_b(\mathrm{K}//\mathrm{H})$ via

$$\mu * \nu := p_{\mathrm{H}}(p^*_{\mathrm{H}}(\mu) * p^*_{\mathrm{H}}(\nu)), \quad \mu,\ \nu \in \mathcal{M}_b(\mathrm{K}//\mathrm{H}).$$

With this convolution K//H is a hypergroup with identity H and involution given by $(\mathrm{H}x\mathrm{H})^\vee = \mathrm{H}\check{x}\mathrm{H}$. Moreover, if there exists a Haar measure ω_{K} on K then $\omega'_{\mathrm{K}} := \int_{\mathrm{K}} \delta_{\mathrm{H}x\mathrm{H}} d\omega_{\mathrm{K}}(x)$ is a Haar measure on K//H.

Sturm-Liouville hypergroups

Sturm-Liouville hypergroups represent important class of hypergroups, which arise from Sturm-Liouville boundary value problems on the nonnegative real line. In order to build up the Sturm-Liouville operator basic to the construction on hypergroups one introduces the Sturm-Lioville functions.

Definition 2.1.10. a) A function $A : \mathbb{R}_+ \to \mathbb{R}$ a called *Sturm-Liouville function* if A is continuous on $[0, \infty[$, differentiable and strictly positive on $]0, \infty[$.

b) A *Sturm-Liouville operator* $\mathcal{L}_A$ associated with a given Sturm-Lioville function

A is defined for all $f \in \mathcal{C}^2(\mathbb{R}_+^*)$ as

$$\mathcal{L}_A f := -f'' - \frac{A'}{A} f', \quad f \in C^2(\mathbb{R}_+^*).$$

The differential operator l on $\mathcal{C}^2(]0,\infty[^2)$ is defined by

$$l[u](x,y) := (\mathcal{L}_A)_x u(x,y) - (\mathcal{L}_A)_y u(x,y)$$

for all $x, y > 0$ and $u \in \mathcal{C}^2([0,\infty[^2)$, where the subscripts indicate variables with respect to which $\mathcal{L}_A$ is taken.

Let A be a Sturm-Liouville function with

$$\frac{A'(x)}{A(x)} = \frac{\alpha_0}{x} + \alpha_1(x) \tag{2.1.2}$$

for all x in a neighborhood of 0, with $\alpha_0 \geq 0$ and such that

(SL1) One of the following additional conditions holds:

- **(SL1a)** $\alpha_0 > 0$ and $\alpha_1 \in \mathcal{C}^\infty(\mathbb{R})$ is an odd function (which implies that $A(0) = 0$).
- **(SL1b)** $\alpha_0 = 0$ and $\alpha_1 \in C^1(\mathbb{R}_+)$ (which implies that $A(0) > 0$).

(SL2) There exists a function $\beta \in \mathcal{C}^1(\mathbb{R}_+)$ such that $\beta(0) \geq 0$, and $\frac{A'}{A} - \beta$ is nonnegative and decreasing on $]0,\infty[$, and $q := \frac{1}{2}\beta' - \frac{1}{4}\beta^2 + \frac{A'}{2A}\beta$ is decreasing on $]0,\infty[$.

(SL3) $\frac{A'}{A} \geq 0$ is decreasing and A is increasing with $\lim_{x\to\infty} A(x) = \infty$.

Definition 2.1.11. A hypergroup $(\mathbb{R}_+, *)$ is said to be a *Sturm-Liouville hypergroup* if there exists a Sturm-Liouville function A such that for every real-valued function f on $\mathbb{R}_+$, which is the restriction of an even nonnegative $\mathcal{C}^\infty$-function on $\mathbb{R}$, the function u_f defined by

$$u_f(x,y) := \int_{\mathbb{R}_+} f \mathrm{d}(\delta_x * \delta_y) \quad x,\ y \in \mathbb{R}_+$$

is twice differentiable and satisfies the partial differential equation

$$l[u_f] = 0, \quad (u_f)_y(x,0) = 0 \text{ for all } x > 0.$$

In the following theorem we collect some known facts, which can be found in [6] and [47].

Theorem 2.1.12. *Let A be a Sturm-Liouville function satisfying conditions (2.1.2) and (SL2). Then there exists a unique hypergroup operation $*$ on $\mathbb{R}_+$ such that $(\mathbb{R}_+, *)$ is a Sturm-Liouville hypergroup. Moreover, $(\mathbb{R}_+, *)$ has following properties*

a) *The neutral element is 0.*

b) *The hypergroup is hermitian.*

c) $\omega_{\mathbb{R}_+} := A\lambda_{\mathbb{R}_+}$ *is a Haar measure of $(\mathbb{R}_+, *)$.*

d) $\rho := \lim\limits_{x\to\infty} \frac{A'(x)}{2A(x)}$ *exists with $\rho \geq 0$; it is called the index of $(\mathbb{R}_+, *)$. Moreover, the growth of $(\mathbb{R}_+, *)$ is determined by ρ. If $\rho > 0$ then the hypergroup is of exponential growth; if $\rho = 0$ then it is exponentially bounded.*

e) *A continuous function $\varphi : \mathbb{R}_+ \to \mathbb{C}$ is multiplicative if and only if φ is a solution of the initial value problem*

$$\mathcal{L}_A\varphi = s\varphi, \ \varphi(0) = 1, \ \varphi'(0) = 0 \quad \text{with } s \in \mathbb{C}.$$

A multiplicative function is a character on $(\mathbb{R}_+, *)$ if and only if the eigenvalue s of $\mathcal{L}_A$ lies in $\mathbb{R}_+$. The dual of $(\mathbb{R}_+, *)$ consists of the functions φ_λ, where $\lambda \in \hat{\mathbb{R}}_+ := \mathbb{R}_+ \cup i[0, \rho]$ and φ_λ is the unique solution of the initial value problem

$$\mathcal{L}_A\varphi_\lambda = (\lambda^2 + \rho^2)\varphi_\lambda, \ \varphi_\lambda(0) = 1 \text{ and } \varphi'_\lambda(0) = 0.$$

f) *The hypergroup admits a Laplace representation, i.e. every character φ_λ has the following representation: For every $x \in \mathbb{R}_+$ there exists a probability measure ν_x on $[-x, x]$ such that*

$$\varphi_\lambda(x) = \int e^{-t(\rho+i\lambda)}\mathrm{d}\nu_x(t). \tag{2.1.3}$$

Furthermore, the measure $\tau_x(t) := e^{-\rho t}d\nu_x(t)$ is a symmetric subprobability measure on $\mathbb{R}$ which depends continuously on x in the weak topology on $\mathcal{M}^1(\mathbb{R})$.

Definition 2.1.13. A Sturm-Liouville hypergroup with defining Sturm-Liouville function A which satisfies the conditions (2.1.2), (SL1a) and (SL3) will be called *Chébli-Trimèche hypergroup.*

Let $(\mathbb{R}_+, *)$ be a Chébli-Trimèche hypergroup with defining function A. It is proved in [7] that $\lambda \mapsto \varphi_\lambda(x)$ is an analytic function for every $x \in \mathbb{R}_+$. The derivatives of $\varphi_\lambda(x)$ with respect to λ were established as the most important tool for defining modified moments for each probability measure on $\mathbb{R}_+$ in a way, which is consistent with the convolution structure (cf. Section 2.3 and [6, Section 7.2]).

Definition 2.1.14. For every $\lambda \in \mathbb{C}$, $x \in \mathbb{R}_+$ and $k \in \mathbb{N}_0$ let

$$\varphi_{k,\lambda}(x) := \frac{\partial^k}{\partial \mu^k} \varphi_{\lambda+i\mu}(x)_{|\mu=0} \quad \text{and} \quad \mathrm{m}_k(x) := \varphi_{k,i\rho}(x). \tag{2.1.4}$$

We recapitulate from [6, Section 7.2] some facts about $\varphi_{k,\lambda}$ and m_k. By differentiating the equation $\mathcal{L}_A \varphi_\lambda = (\rho^2 + \lambda^2)\varphi_\lambda$ with respect to λ we obtain

$$\begin{cases} \mathcal{L}_A \varphi_{k,\lambda} = (\rho^2 + \lambda^2)\varphi_{k,\lambda} + 2ik\lambda\varphi_{k-1,\lambda} - k(k-1)\varphi_{k-2,\lambda}, \\ \varphi_{k,\lambda}(0) = \varphi'_{k,\lambda}(0) = 0 \quad \text{for} \quad k \geq 1. \end{cases}$$

and especially

$$\begin{cases} \mathcal{L}_A \mathrm{m}_k = -2k\rho \mathrm{m}_{k-1} - k(k-1)\mathrm{m}_{k-2}, \\ \mathrm{m}_k(0) = \mathrm{m}'_k(0) = 0 \quad \text{for} \quad k \geq 1. \end{cases} \tag{2.1.5}$$

For $k = 0$ we have $\varphi_{i\rho} \equiv 1$ and thus $\mathrm{m}_0 \equiv 1$. It follows from the Laplace representation (2.1.3) that

$$\mathrm{m}_k(x) = \int_{-x}^{x} t^k \mathrm{d}\nu_x(t) = \int_0^x t^k \left(e^{t\rho} + (-1)^k e^{-t\rho}\right) \mathrm{d}\tau_x(t) \tag{2.1.6}$$

for $x \in \mathbb{R}_+$, $\lambda \in \mathbb{C}$ and $k \geq 1$. In particular, m_k is non-negative and in the case $\rho = 0$ it is clear that $\mathrm{m}_k = 0$ if k is odd.

In the following proposition we collect some properties of functions m_k.

Proposition 2.1.15. *Let $(\mathbb{R}_+, *)$ be a Chébli-Trimèche hypergroup with defining function A. The functions m_k, $(k \in \mathbb{N}_0)$ defined by (2.1.4) are real valued and analytic on $]0, \infty[$. Moreover, $(\mathrm{m}_k)_{k \in \mathbb{N}_0}$ forms a sequence of functions with following properties*

a) $\int_K \mathrm{m}_k(z)\, \mathrm{d}\delta_x * \delta_y(z) = \sum_{j=0}^{k} \binom{k}{j} \mathrm{m}_j(x)\mathrm{m}_{k-j}(y)$ *for all $k \in \mathbb{N}$ and $m_0 \equiv 1$.*

b) $\mathrm{m}_1^2(x) \leq \mathrm{m}_2(x)$ *for all $x \in \mathbb{R}_+$.*

and in the case of subexponential growth (i.e. $\rho = 0$)

c) $m_{2k-1} \equiv 0$ *for all* $k \geq 1$ *and* $\lim_{x\to\infty} m_2(x)/x = \infty$

or in the case of exponential growth (i.e. $\rho > 0$*)*

d) $\lim_{x\to\infty} m_1'(x) = 1$

e) m_1^{-1} *exists and is differentiable.*

f) $\lim_{x\to\infty} (m_1^{-1})'(x) = 1$

Proof: We only proof the properties c) and d); all the rest can be found in [6, Section 7.2]. At first we prove that $m_1'(x) > 0$ for all $x \in]0, \infty[$. Suppose that m_1' takes values in $] - \infty, 0]$. Since $m_1''(0) = 2\rho/(\alpha + 1) > 0$ there exists $\xi > 0$ such that $m_1'(\xi) > 0$, $m_1''(\xi) < 0$ and $m_1'''(\xi) < 0$. This implies that $m_1' \frac{A'}{A}$ and m_1'' are strictly decreasing in a neighborhood of ξ. But this is impossible since $m_1'' + m_1' \frac{A'}{A} = 2\rho$ by (2.1.5). From this contradiction we conclude that $m_1'(x) > 0$ for all $x \in]0, \infty[$. In particular, m_1 is strictly increasing and hence m_1^{-1} exists on $\mathbb{R}_+$. From c) we get $\lim_{x\to\infty} m_1(x) = \infty$ and therefore $\lim_{y\to\infty} m_1^{-1}(y) = \infty$. By the inverse function theorem and c) we obtain

$$(m_1^{-1})'(y) = \frac{1}{m_1'(m_1^{-1}(y))} \searrow 1 \quad \text{as } y \to \infty \tag{2.1.7}$$

□

Orbital morphisms

Let $(K, *_K)$ and $(J, *_J)$ be hypergroups and $\Phi : J \to K$ a mapping.

Definition 2.1.16. a) $\Phi : J \to K$ is called *orbital mapping* if it is a proper and open continuous surjection. In this case, the compact sets $\Phi^{-1}(y)$ will be called the Φ-*orbits*.

b) A *recomposition* of an orbital mapping $\Phi : J \to K$ is a weakly continuous mapping $x \mapsto q_x$, $K \to \mathcal{M}^1(J)$ such that $supp(q_x) = \Phi^{-1}(x)$ for all $x \in K$. If there exists a measure $l \in \mathcal{M}_+(J)$ such that

$$l = \int_J q_{\Phi(y)} dl(y)$$

then this recomposition is said to be *consistent with* l.

c) A orbital mapping $\Phi : J \to \mathrm{K}$ is called a *generalised orbital morphism associated with the recomposition* $(q_x)_{x\in\mathrm{K}}$ if $q_{x^-} = q_x^-$ and $\Phi(q_x *_J q_y) = \delta_x *_K \delta_y$ whenever $x, y \in \mathrm{K}$. If $(q_x)_{x\in\mathrm{K}}$ is consistent with a Haar measure of J (provided this exists) then the associated generalised orbital morphism Φ is said to be an *orbital morphism*.

Obviously, every injective generalised orbital morphism is a hypergroup isomorphism.

Radial random walks

Let G be a locally compact Hausdorff group, H a compact subgroup of G and π the canonical projection from the left cosets G/H onto double cosets $G//\mathrm{H}$

$$\pi : G/\mathrm{H} \to G//\mathrm{H}, \quad gH \mapsto HgH \quad (g \in G).$$

Definition 2.1.17. a) A measure $\mu \in \mathcal{M}^1(G/\mathrm{H})$ is called H-*radial* or H-*invariant* if

$$\mu(hB) = \mu(B) \quad \text{for all } h \in \mathrm{H} \text{ and } B \in \mathcal{B}(G/\mathrm{H}).$$

The set of all H-radial measures will be denoted by $\mathcal{M}^1_{rad}(G/\mathrm{H})$.

b) A measure $\nu \in \mathcal{M}^1(G//\mathrm{H})$ on $G//\mathrm{H}$ is called *radial part* of $\mu \in \mathcal{M}^1_{rad}(G/\mathrm{H})$ if

$$\pi(\mu)(B) = \nu(B) \quad \text{for all } B \in \mathcal{B}(G//\mathrm{H}).$$

c) Let $\mu \in \mathcal{M}^1_{rad}(G/\mathrm{H})$ with radial part $\nu \in \mathcal{M}^1(G//\mathrm{H})$. A stochastic process $(S_n)_{n\in\mathbb{N}}$ on G/H with start in H (i.e. $S_0 = \mathrm{H}$) is called *radial random walk associated with* μ if

$$\mathbb{P}(S_{n+1} \in HB|\ S_n = Hx) = \mu(Hx^{-1}B) = \nu(Hx^{-1}BH).$$

Theorem 2.1.18. *Let* $\nu \in \mathcal{M}^1(G//\mathrm{H})$. *Then there exists a unique* H-*radial measure* $\mu \in \mathcal{M}^1_{rad}(G/\mathrm{H})$ *such that* ν *is radial part of* μ.

Proof: The proof is given in [42]. □

2.2 Concretization of hypergroups

The forming of sums of K-valued random variables is not directly possible, as there is no deterministic operation on K in general. It is clear, that the convolution $\mathbb{P}_X * \mathbb{P}_Y$ of two independent random variables X and Y should be the law of an analogy for their *sum*. In this section we recapitulate the construction of the randomized sum of K-valued random variables using the concept of the concretization of hypergroups (see [6, Chapter 7]). We start with the definition of an increment process on the hypergroup K.

Definition 2.2.1. Let $I \neq \emptyset$ be a totally ordered parameter set and $(X_t)_{t\in I}$ a K-valued stochastic process on $(\Omega, \mathcal{A}, \mathbb{P})$ with canonical filtration $(\mathcal{F}_t)_{t\in I}$. The process $(X_t)_{t\in I}$ is said to be an *increment process* on $(\mathrm{K}, *)$ if for all $s,\ t \in I$ with $s < t$ there exists $\eta_{s,t} \in \mathcal{M}^1(\mathrm{K})$ such that

$$\mathbb{P}(X_t \in A|\mathcal{F}_s) = (\eta_{s,t} * \delta_{X_s})(A) \quad \mathbb{P} - a.s.$$

whenever $A \in \mathcal{B}(\mathrm{K})$. In the case $I \subset \mathbb{R}_+$ an increment process $(X_t)_{t\in I}$ is called *stationary* if $\eta_{s,t} = \eta_{0,t-s}$ for all $s < t$ in I.

It can be easily checked that increment processes on K have the elementary Markov property, i.e.

$$\mathbb{P}(X_t \in A|\mathcal{F}_s) = \mathbb{P}(X_t \in A|X_s) \quad \mathbb{P} - a.s.$$

whenever $A \in \mathcal{B}(X)$ and $s < t$ from I. The probability measures $\eta_{s,t} \in \mathcal{M}^1(\mathrm{K})$ form a so-called hemigroup, i.e. for $s \leq t \leq u \in I$, we have $\eta_{s,t} * \eta_{t,u} = \eta_{s,u}$. Moreover, for $t \in I$, the law of X_t is given by $\mathbb{P}_{X_0} * \eta_{0,t}$ as in the group case. On the other hand, standard arguments on the construction of Markov processes ensure that for a given initial law and a given hemigroup $(\eta_{s,t})_{s\leq t\in I} \subset \mathcal{M}^1(\mathrm{K})$ there always exists an associated increment process on a suitable probability space.

Definition 2.2.2. A time-homogeneous Markov process $(S_n)_{n\in\mathbb{N}_0}$ is called a *random walk* on the hypergroup $(\mathrm{K}, *)$ with law $\nu \in \mathcal{M}^1(\mathrm{K})$ if

$$\mathbb{P}(S_{n+1} \in A|S_n = x) = (\delta_x * \nu)(A)$$

for all $n \in \mathbb{N}_0$, $x \in \mathrm{K}$, and Borel sets $A \subset \mathrm{K}$.

Clearly, a random walk $(S_n)_{n\geq 1}$ on $(\mathrm{K}, *)$ is an increment process with $I = \mathbb{N}$

and $\eta_{m,n}$ ($m \leq n \in \mathbb{N}$) is the $n-m$ fold convolution product of some probability measure $\eta \in \mathcal{M}^1(\mathrm{K})$, i.e. $\eta_{m,n} = \eta^{n-m}$.

We next turn to the concept of concretisation on hypergroups, which makes possible to define an analogy of a sum for K-valued random variables.

Definition 2.2.3. Let K be a hypergroup, μ a probability measure on a compact set M, and let $\Phi : \mathrm{K} \times \mathrm{K} \times \mathrm{M} \to \mathrm{K}$ be Borel-measurable. The triple (M, μ, Φ) is called a *concretization* of K if

$$\mu\{\Phi(x, y, \cdot) \in A\} = (\delta_x * \delta_y)(A)$$

for all x, $y \in \mathrm{K}$ and $A \in \mathcal{B}(\mathrm{K})$.
Since $\delta_x * \delta_e = \delta_e * \delta_x = \delta_x$ for all $x \in \mathrm{K}$, we obviously have

$$\Phi(e, x, \cdot) = \Phi(x, e, \cdot) = x \quad \mu - a.s. \tag{2.2.1}$$

For a list of examples of concretisation we refer to [6, Section 7.1] and [46]. The following result guarantees the existence of concretization for a large class of hypergroups.

Theorem 2.2.4. *Let* K *be a second countable hypergroup. Then there exists a measurable mapping* Φ *from* $\mathrm{K} \times \mathrm{K} \times [0, 1]$ *into* K *such that* $([0, 1], \lambda_{[0,1]}, \Phi)$ *is a concretization of* K.

The proof of this theorem is given in [6, Theorem 7.1.3] and in [46, 47]. In the sequel, let (M, μ, Φ) be a concretization of the hypergroup $(\mathrm{K}, *)$.

Definition 2.2.5. For any two K-valued random variables X and Y on $(\Omega, \mathcal{A}, \mathbb{P})$ and an auxiliary M-valued random variable Λ on $(\Omega, \mathcal{A}, \mathbb{P})$ such that Λ is independent of the random variable (X, Y) and has distribution $\mathcal{L}(\Lambda) =: \mu$ we define the *randomized sum* of X with Y by

$$X \overset{\Lambda}{+} Y := \Phi(X, Y, \Lambda).$$

Obviously, $X \overset{\Lambda}{+} Y$ is a K-valued random variable on $(\Omega, \mathcal{A}, \mathbb{P})$. As in the classical case, one has the following relation.

Proposition 2.2.6. *Let* X, Y *and* Λ *be as in the Definition 2.2.5. If* X, Y *and* Λ *are independent, then*

$$\mathcal{L}(X \overset{\Lambda}{+} Y) = \mathcal{L}(X) * \mathcal{L}(Y). \tag{2.2.2}$$

For a proof see [6, Proposition 7.1.6] and [46].

Remark 2.2.7. **a)** A concretisation is not uniquely determined by the hypergroup and hence the randomized sum $X \overset{\Lambda}{+} Y$ depends on the particular choice of the underlying concretisation of K.

b) The joint distribution of the random variables X, Y and $X \overset{\Lambda}{+} Y$ is independent from the particular choice of concretisation. (see [6, Proposition 7.1.8] and [46].)

c) In contrast to the group case the randomized sum of deterministic random variables need not to be deterministic.

In order to define partial sums S_n with values in K we will need the following assumption.

Assumption 2.2.8. Let $(X_n)_{n\geq 1}$ be a sequence of K-valued random variables and $(\Lambda_n)_{n\geq 1}$ of M-valued random variables on $(\Omega, \mathcal{A}, \mathbb{P})$ such that X_1, Λ_1, X_2, Λ_2, $\dots$ are independent, $\mathbb{P}_{X_j} = \nu_j$ and $\mathbb{P}_{\Lambda_j} = \mu$ for all $j \geq 1$.

The definition of randomized sum in 2.2.5 can now be extended to sequences $(X_n)_{n\geq 1}$ of K valued random variables and $(\Lambda_n)_{n\geq 1}$ of M-valued auxiliary random variables as in the assumption above.

Definition 2.2.9. The randomized sum S_n is recursively defined by $S_0 = e$ and

$$S_n := \begin{cases} X_1 & \text{for } n = 1, \\ \Phi(S_{n-1}, X_n, \Lambda_{n-1}) & \text{for } n > 1. \end{cases} \tag{2.2.3}$$

Clearly, S_n is a K-valued random variable on $(\Omega, \mathcal{A}, \mathbb{P})$ with

$$\begin{aligned} S_n &= S_{n-1} \overset{\Lambda_{n-1}}{+} X_n = (S_{n-2} \overset{\Lambda_{n-2}}{+} X_{n-1}) \overset{\Lambda_{n-1}}{+} X_n \\ &= (\dots(((X_1 \overset{\Lambda_1}{+} X_2) \overset{\Lambda_2}{+} X_3) \overset{\Lambda_3}{+} X_4) \overset{\Lambda_4}{+} \dots) \overset{\Lambda_{n-1}}{+} X_n. \end{aligned}$$

We see that $(S_n)_{n\geq 1}$ is a (in general non homogeneous) Markov chain with transition kernels

$$N_n(x, A) := \delta_x * \nu_n(A) \quad \text{for } \mathbb{P}_{S_{n-1}}\text{-almost all } x$$

on $(\mathrm{K}, \mathcal{B}(\mathrm{K}))$, which is clear from

$$\begin{aligned}(\delta_x * \nu_n)(A) &= \int_{\mathrm{K}} \delta_x * \delta_y(A) \mathrm{d}\nu_n(y) = \int_{\mathrm{K}} \mu(\Phi(x, y, \cdot) \in A) \mathrm{d}\nu_n(y) \\ &= \mathbb{P}\left(\Phi(x, X_n, \Lambda_{n-1}) \in A\right) = \mathbb{P}\left(S_n \in B \middle| S_{n-1} = x\right).\end{aligned}$$

Remark 2.2.10. a) If the random variables $X_1, X_2, \ldots$ are identically distributed, then $(S_n)_{n\in\mathbb{N}}$ is a random walk on K with law $\nu := \mathbb{P}_{X_1}$ in the sense of the Definition 2.2.2.

b) A random walk $(S_n)_{n\in\mathbb{N}}$ on K is a stationary increment process with $\eta_{m,n} := \nu^{n-m}$ for $m < n$ where $\nu := \mathbb{P}_{X_1}$

c) Forming randomized sums is in general not an associative operation although convolution of distributions obviously is.

Definition 2.2.11. Let $j \leq n$ from $\mathbb{N}$. The randomized sum $S_n(j, R)$ where the j-th term X_j is replaced by a K-valued random variable R on $(\Omega, \mathcal{A}, \mathbb{P})$ is recursively defined by

$$S_n(j, R) := \begin{cases} \Phi(S_{n-1}(j, R), X_n, \Lambda_{n-1}) & \text{for } j < n, \\ \Phi(S_{j-1}, R, \Lambda_{j-1}) & \text{for } j = n, \\ S_n & \text{for } j > n. \end{cases} \tag{2.2.4}$$

For instance, if $n = 5$ and $j = 3$ then

$$S_5(3, R) = S_4(3, R) \overset{\Lambda_4}{+} X_5 = \ldots = \left(\left((X_1 \overset{\Lambda_1}{+} X_2) \overset{\Lambda_2}{+} R\right) \overset{\Lambda_3}{+} X_4\right) \overset{\Lambda_4}{+} X_5.$$

Clearly, for $R \equiv e$ the randomized sum $S_n(j, e)$ coincides $\mathbb{P}$-a.s. with the randomized sum S_n, where the j-th term is omitted. It is easy to check that the distribution of $\mathbb{P}_{S_n(j,R)}$ is given by

$$\mathbb{P}_{S_n(j,R)} = \nu_1 * \ldots * \nu_{j-1} * \mathbb{P}_R * \nu_{j+1} * \ldots * \nu_n.$$

For $\nu \in M^1(\mathrm{K})$ and $n \in \mathbb{N}$ we denote the n-fold convolution power of ν with respect to the convolution $*$ by ν^n. If $X_1, X_2, \ldots$ are identically distributed, say $\mathbb{P}_{X_n} = \nu$ for $n \geq 1$ then

$$\mathbb{P}_{S_n} = \nu^n \quad \text{and} \quad \mathbb{P}_{S_n(j,R)} = \nu^{j-1} * \mathbb{P}_R * \nu^{n-j} \quad (j, n \in \mathbb{N},\ j \leq n).$$

2.3 Moments on hypergroups

We recapitulate the concept of moment function introduced by Zeuner; see [46, 47] and [6, Section 7.2].

Definition 2.3.1. Define the function $\mathrm{m}_0(x) = 1$ for $x \in \mathrm{K}$.

a) A finite sequence $(\mathrm{m}_i)_{i=1,\ldots,n}$ of measurable and locally-bounded functions $\mathrm{m}_j : \mathrm{K} \to \mathbb{C}$ $(j = 1, \ldots, n)$ is called a sequence of moment functions of length $n \in \mathbb{N}$ if

$$\int_{\mathrm{K}} \mathrm{m}_i(z)\, \mathrm{d}\delta_x * \delta_y(z) = \sum_{j=0}^{i} \binom{i}{j} \mathrm{m}_j(x)\mathrm{m}_{i-j}(y) \tag{2.3.1}$$

for all $i = 1, \ldots, n$ and all $x, y \in \mathrm{K}$. Moreover, the function m_j $(j = 1, \ldots, n)$ is called a moment function of j-th order (associated with the sequence $(\mathrm{m}_i)_{i=1,\ldots,n}$).

b) A sequence of moment functions $(\mathrm{m}_i)_{i=1,\ldots,n}$ has the property $(MF1)$ if m_i are real valued for all $i \in \{1, \ldots, n\}$ and $\mathrm{m}_1^2(x) \leq \mathrm{m}_2(x)$ for all $x \in \mathrm{K}$.

c) For a fixed sequence $(\mathrm{m}_i)_{i=1,\ldots,n}$ of moment functions, introduce the space

$$\mathcal{M}^1_n(\mathrm{K}) := \left\{\mu \in \mathcal{M}^1(\mathrm{K}) :\ \mathrm{m}_i \in L^1(\mathrm{K}, \mu) \text{ for all } 0 \leq i \leq n\right\}$$

of probability measures for which all moments up to the n-th exist.

Proposition 2.3.2. *Let* $(\mathrm{m}_i)_{i=1,\ldots,n}$ *be a sequence of moment functions on* K *of length* $n \in \mathbb{N}$ *and* μ, $\nu \in \mathcal{M}^1(\mathrm{K})$. *Then*

a) $\mathrm{m}_i(e) = 0$ *for* $i = 1, \ldots, n$.

b) *All* m_i *are continuous*

c) *If* m_k *is bounded for an* $k \in \{1, \ldots, n\}$ *then* $\mathrm{m}_i \equiv 0$ *for all* $i \leq k$.

d) *Let* $k \in \{1, \ldots, n\}$. *Then* $\mu * \nu \in \mathcal{M}^1_k(\mathrm{K})$ *if and only if* μ, $\nu \in \mathcal{M}^1_k(\mathrm{K})$.

Proof: Properties a),b) and c) are known from [28, Section 4].
Property *d)* will be shown by induction on $k \in \{1, \ldots, n\}$. The case $k = 1$ is proven in [47, Lemma 5.9]. Assume that assertion holds when k is replaced by $k - 1 \in \{1, \ldots, n-1\}$. For $y \in \mathrm{K}$ we set

$$f_k(y) := \int_{\mathrm{K}} |\sum_{j=0}^{k} \binom{k}{j} \mathrm{m}_j(x)\mathrm{m}_{k-j}(y)|\mathrm{d}\mu(x) \in [0, \infty].$$

When $\mu * \nu \in M_k^1(\mathrm{K})$, we obtain the inequality

$$\begin{aligned}\int_{\mathrm{K}} f_k(y)\mathrm{d}\nu(y) &= \int_{\mathrm{K}}\int_{\mathrm{K}}\left|\int_{\mathrm{K}} \mathrm{m}_k \mathrm{d}\delta_x * \delta_y\right|\mathrm{d}\mu(x)\mathrm{d}\nu(y) \\ &\le \int_{\mathrm{K}}\int_{\mathrm{K}} |\mathrm{m}_k|\mathrm{d}\delta_x * \delta_y \mathrm{d}\mu(x)\mathrm{d}\nu(y) = \int_{\mathrm{K}} |\mathrm{m}_k|\mathrm{d}\mu * \nu < \infty\end{aligned}$$

Hence, by Fubini's theorem there exists $y_0 \in \mathrm{K}$ with $f_k(y_0) < \infty$. This implies

$$\begin{aligned}\int_{\mathrm{K}} |\mathrm{m}_k(x)|\mathrm{d}\mu(x) &\le \int_{\mathrm{K}}\left|\sum_{j=0}^{k}\binom{k}{j}\mathrm{m}_j(x)\mathrm{m}_{k-j}(y_0)\right| + \left|\sum_{j=0}^{k-1}\binom{k}{j}\mathrm{m}_j(x)\mathrm{m}_{k-j}(y_0)\right|\mathrm{d}\mu(x) \\ &\le f_k(y_0) + \sum_{j=0}^{k-1}\binom{k}{j}|\mathrm{m}_{k-j}(y_0)|\int_{\mathrm{K}}|\mathrm{m}_j(x)|\mathrm{d}\mu(x) < \infty\end{aligned}$$

and by symmetrie, $\int_{\mathrm{K}} |\mathrm{m}_k|\mathrm{d}\nu < \infty$. The reverse implication is clear. □

Here and subsequently, let $(\mathrm{m}_i)_{i=1,\ldots,n}$ be a sequence of moment functions of length $n \ge 2$ with the Property $(M1)$; cf. Definition 2.3.1, b). We now turn to the introduction of modified moments of K-valued random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ following H. Zeuner's notion from [47].

Definition 2.3.3. Let $k \in \{1, \ldots, n\}$ and X, Y be a K-valued random variables on $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\mathrm{m}_1(X)$ and $\mathrm{m}_1(Y)$ are integrable.

a) If $\mathrm{m}_k(X) \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ then

$$\mathbb{E}_k^*(X) := \mathbb{E}(\mathrm{m}_k(X)) = \int \mathrm{m}_k(X)\mathrm{d}\mathbb{P}$$

will be called the *k-th modified moment* of X (with respect to the moment function m_k). We write $\mathbb{E}_*(X)$ for a modified moment of first order and refer to a *modified expectation*.

b) X and Y are called $*_k$-*uncorrelated* if and only if $\mathbb{P}_X, \mathbb{P}_Y \in \mathcal{M}_{k+1}^1(\mathrm{K})$ and

$$\mathbb{E}(\mathrm{m}_i(X)\mathrm{m}_{j-i}(Y)) = \mathbb{E}_i^*(X)\mathbb{E}_{j-i}^*(Y)$$

for all i, $j \in \mathbb{N}_0$ with $i \le j \le k$. If $k = 0$, then $*_0$-uncorrelatedness just means that $\mathrm{m}_1(X)$ and $\mathrm{m}_1(Y)$ are integrable.

Obviously, if two random variables X and Y are independent and $\mathbb{P}_X$, $\mathbb{P}_Y \in \mathcal{M}_k^1(\mathrm{K})$ for some $k \in \mathbb{N}$, then they are $*_k$-uncorrelated.

Remark 2.3.4. In a certain sense, the concept of modified moments on hypergroups generalised the classical concept of moment functions. In fact, the functions $\mathrm{m}_k(x) := x^k$, $k \in \mathbb{N}$ form a sequence of moment functions on $\mathrm{K} := \mathbb{R}$, where $\mathcal{M}^b(\mathbb{R})$ is equipped with the usual convolution.

As in the classical case, there is a binomial identity for a modified moments of a randomized sum $X \overset{\Lambda}{+} Y$ of independent random variables X, Y and Λ, which has been proven in [46]. The same identity but under the weakened assumptions of the $*_k$-uncorrelatedness instead of the independence, is established by our next proposition.

Proposition 2.3.5. *Let X, Y and Λ be random variables as in the Definition 2.2.5 such that X and Y are $*_{n-1}$-uncorrelated for an $n \in \mathbb{N}$. Then $\mathrm{m}_n(X \overset{\Lambda}{+} Y)$ is integrable and*

$$\mathbb{E}_n^*(X \overset{\Lambda}{+} Y) = \sum_{k=0}^{n} \binom{n}{k} \mathbb{E}_k^*(X)\mathbb{E}_{n-k}^*(Y). \tag{2.3.2}$$

In particular, $\mathbb{E}_*(X \overset{\Lambda}{+} Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

Proof: By assumption and the Proposition 2.3.2, $\mathrm{m}_n(X \overset{\Lambda}{+} Y)$ is integrable. Moreover, the independence of (X, Y) and Λ together with the Equation (2.3.1) and $*_{n-1}$-uncorrelatedness of X and Y implies that

$$\begin{aligned} \mathbb{E}_n^*(X \overset{\Lambda}{+} Y) &= \mathbb{E}\Big(\mathbb{E}(\mathrm{m}_n(X \overset{\Lambda}{+} Y)|X, Y)\Big) = \mathbb{E}\Big(\int_{\mathrm{K}} \mathrm{m}_n \mathrm{d}(\delta_X * \delta_Y)\Big) \\ &= \mathbb{E}\Big(\sum_{k=0}^{n} \binom{n}{k} \mathrm{m}_j(X)\mathrm{m}_{n-j}(Y)\Big) = \sum_{k=0}^{n} \binom{n}{k} \mathbb{E}_k^*(X)\mathbb{E}_{n-k}^*(Y) \end{aligned}$$

□

Remark 2.3.6. Let $\mu_i \in \mathcal{M}_k^1(\mathrm{K})$ $(i = 1, 2, 3)$. By a straightforward calculation we obtain following commutativity property

$$\int_{\mathrm{K}} \mathrm{m}_k(x)\mathrm{d}(\mu_1 * \mu_2 * \mu_3)(x) = \int_{\mathrm{K}} \mathrm{m}_k(x)\mathrm{d}(\mu_{\sigma(1)} * \mu_{\sigma(2)} * \mu_{\sigma(3)})(x) \tag{2.3.3}$$

for any permutation σ of the set $\{1, 2, 3\}$.

In order to define the modified variance of a random variable X on a hypergroup K, a function v has been introduced by Zeuner in [47, Section 6] by

$$v : \mathrm{K} \times \mathbb{R} \to \mathbb{C}, v(x, \xi) := \mathrm{m}_2(x) - 2\xi\mathrm{m}_1(x) + \xi^2.$$

Definition 2.3.7. For any K-valued random variable X such that $\mathbb{P}_X \in \mathcal{M}_2^1(\mathrm{K})$ we call

$$V_*(X) := \mathbb{E}\Big(v(X, \mathbb{E}_*(X))\Big) = \mathbb{E}_2^*(X) - \mathbb{E}_*(X)^2$$

the *modified variance* of X (associated with $(\mathrm{m}_1, \mathrm{m}_2)$).

Is $(\mathrm{m}_1, \mathrm{m}_2)$ a sequence of moments with Property $(MF1)$ (cf. the Definition 2.3.1), then the function v is non negative. Moreover, for every random variable X such that $\mathbb{P}_X \in \mathcal{M}_2^1(\mathrm{K})$ the function $\xi \mapsto \mathbb{E}(v(X,\xi))$ on $\mathbb{R}$ takes its minimum at $\xi = \mathbb{E}_*(X)$, this value being $V_*(X) \geq 0$. Clearly, the modified variance of a deterministic random variable is in general different from 0. This is consistent with the fact that randomized sums of deterministic random variables need not to be deterministic.

Remark 2.3.8. Let X, Y and Λ be random variables as in the Definition 2.2.5 such that X and Y are $*_1$-uncorrelated (w.r.t. the moment functions m_1, m_2). Then, by the same arguments as in the Proposition 2.3.5 we obtain

$$\mathbb{V}_*(X \overset{\Lambda}{+} Y) = \mathbb{V}_*(X) + \mathbb{V}_*(Y). \tag{2.3.4}$$

If $\mathrm{m}_1^2 \leq \mathrm{m}_2$ on K, then the usual variance of $\mathrm{m}_1(X)$ is upper bounded by the modified variance of X, i.e.

$$\mathbb{V}ar(\mathrm{m}_1(X)) = \mathbb{E}(\mathrm{m}_1(X)^2) - \mathbb{E}(\mathrm{m}_1(X))^2 \leq \mathbb{V}_*(X). \tag{2.3.5}$$

In the remainder of this chapter let $(X_n)_{n\geq1}$ and $(\Lambda_n)_{n\geq1}$ be sequences as in the Assumption 2.2.8. By $(S_n)_{n\geq1}$ we abbreviate the associated Markov chain introduced in Definition 2.2.9. The following observation will be the key for the moment estimate in Theorem 2.3.10 below.

Proposition 2.3.9. *Let $(X_n)_{n\geq1}$ and $(\Lambda_n)_{n\geq1}$ be sequences as in Assumption 2.2.8 such that $\mathbb{P}_{X_i} \in \mathcal{M}_k^1(\mathrm{K})$ for all $i \in \mathbb{N}$ and some $k \in \mathbb{N}$. Moreover, let S_n be the corresponding randomized sum. Then*

$$\mathbb{E}(\mathrm{m}_k(S_n)|X_j) = \sum_{l=0}^{k} \binom{k}{l} \mathrm{m}_l(X_j)\mathbb{E}_{k-l}^*(S_n(j,e)) \quad \mathbb{P}-a.s.$$

for all $n \in \mathbb{N}$, $j \in \{1,\ldots,n\}$. In particular, for $k=1$ we have

$$\mathbb{E}(\mathrm{m}_1(S_n)|X_j) = \mathrm{m}_1(X_j) + \mathbb{E}_*(S_n(j,e)) \quad \mathbb{P}-a.s.$$

Proof: For a set $B \in \mathcal{B}(\mathrm{K})$ consider the truncation map

$$\chi_B : \mathrm{K} \to \mathrm{K}, \qquad \chi_B(x) = \begin{cases} x & \text{for } x \in B, \\ e & \text{for } x \notin B \end{cases}$$

as well as the random variable $S_n(j, \chi_B(X_j))$. Obviously, by the definition of χ_B we have

$$S_n(j, \chi_B(X_j))(\omega) = \begin{cases} S_n(\omega) & \text{for } \omega \in \{X_j \in B\} \\ S_n(j,e)(\omega) & \text{for } \omega \notin \{X_j \in B\}\,. \end{cases}$$

This and the independence of X_j and $S_n(j,e)$ clearly forces

$$\begin{aligned} \mathbb{E}(1_{\{X_j \in B\}} \mathrm{m}_k(S_n)) &= \mathbb{E}^*_k(S_n(j, \chi_B(X_j))) - \mathbb{E}(1_{\{X_j \notin B\}} \cdot \mathrm{m}_k(S_n(j,e))) \\ &= \mathbb{E}^*_k(S_n(j, \chi_B(X_j))) - \mathbb{P}(X_j \notin B)\mathbb{E}^*_k(S_n(j,e)). \end{aligned} \tag{2.3.6}$$

On the other side, by Remark 2.2.10, Eq. (2.3.2) and (2.3.3) it follows that

$$\begin{aligned} \mathbb{E}^*_k(S_n(j, \chi_B(X_j))) &= \int_\Omega \mathrm{m}_k(x) \mathrm{d}(\mathbb{P}_{\chi_B(X_j)} * \nu^{n-1})(x) \\ &= \sum_{\alpha=0}^{k} \binom{k}{\alpha} \mathbb{E}^*_\alpha(\chi_B(X_j)) \mathbb{E}^*_{k-\alpha}(S_n(j,e)). \end{aligned} \tag{2.3.7}$$

Taking (2.3.6) and (2.3.7) into account we obtain

$$\mathbb{E}(1_{\{X_j \in B\}} \mathrm{m}_k(S_n)) = \sum_{\alpha=0}^{k} \binom{k}{\alpha} \mathbb{E}^*_\alpha(\chi_B(X_j)) \mathbb{E}^*_{k-\alpha}(S_n(j,e)) - \mathbb{P}(X_j \notin B)\mathbb{E}^*_k(S_n(j,e)).$$

Since $\mathrm{m}_0 \equiv 1$, we see at once that $\mathbb{E}^*_0(\chi_B(X_j)) - \mathbb{P}(X_j \notin B) = \mathbb{P}(X_j \in B)$. Moreover, we have $1_{\{X_j \in B\}} \mathrm{m}_l(X_j) = \mathrm{m}_l(\chi_B(X_j))$ for all $l \in \mathbb{N}$, which is due to the fact that $\mathrm{m}_l(e) = 0$. This gives

$$\begin{aligned} \mathbb{E}(1_{\{X_j \in B\}} \mathrm{m}_k(S_n)) &= \sum_{\alpha=1}^{k} \binom{k}{\alpha} \mathbb{E}^*_\alpha(\chi_B(X_j)) \mathbb{E}^*_{k-\alpha}(S_n(j,e)) + \mathbb{P}(X_j \in B)\mathbb{E}^*_k(S_n(j,e)) \\ &= \sum_{\alpha=0}^{k} \binom{k}{\alpha} \mathbb{E}(1_{\{X_j \in B\}} \mathrm{m}_\alpha(X_j)) \mathbb{E}^*_{k-\alpha}(S_n(j,e)). \end{aligned}$$

□

Clearly, a moment function m_1 of the first order is in general not linear. An upper bound of the deviation of $\mathrm{m}_1(S_n)$ from the linearisation $\sum \mathrm{m}_1(X_j)$ in L^2-sense is established by our next theorem.

Theorem 2.3.10. *Let* $(\mathrm{m}_1, \mathrm{m}_2)$ *be a sequence of moment functions such that* $0 \le \mathrm{m}_1^2(\mathrm{x}) \le \mathrm{m}_2(\mathrm{x})$ *for all* $x \in \mathrm{K}$. *Let* $(X_n)_{n\ge1}$ *and* $(\Lambda_n)_{n\ge1}$ *be sequences as in the Assumption 2.2.8 such that* $\mathbb{P}_{X_i} \in \mathcal{M}_2^1(\mathrm{K})$ *for all* $i \in \mathbb{N}$. *Moreover, let* $(S_n)_{n\ge1}$ *be the associated process of randomized sums. Then*

$$\mathbb{V}ar\Big(\mathrm{m}_1(S_n) - \sum_{j=1}^n \mathrm{m}_1(X_j)\Big) \le \sum_{j=1}^n \mathbb{E}(\mathrm{m}_2(X_j) - \mathrm{m}_1(X_j)^2) \tag{2.3.8}$$

Proof: We define $Z_n := \mathrm{m}_1(S_n) - \sum_{j=1}^n \mathrm{m}_1(X_j)$ and calculate

$$Z_n^2 = \mathrm{m}_1\left(S_n\right)^2 - 2\mathrm{m}_1\left(S_n\right) \cdot \sum_{j=1}^n \mathrm{m}_1\left(X_j\right) + \left\{\sum_{j=1}^n \mathrm{m}_1\left(X_j\right)\right\}^2.$$

Since the random variables $X_1, \dots, X_n$ are independent, the property $(MF1)$ $(\mathrm{m}_1^2 \le \mathrm{m}_2)$ yields

$$\mathbb{E}(Z_n^2) \le \mathbb{E}_{*,2}(S_n) - 2\sum_{j=1}^n \mathbb{E}(\mathrm{m}_1(S_n)\mathrm{m}_1(X_j)) + \tag{2.3.9}$$
$$+ \sum_{i\ne j}^n \mathbb{E}_*(X_i)\mathbb{E}_*(X_j) + \sum_{j=1}^n \mathbb{E}(\mathrm{m}_1(X_j)^2).$$

Using Proposition 2.3.9 and Equation (2.3.2) we obtain for $j \in \{1, \dots, n\}$,

$$\begin{aligned}\mathbb{E}(\mathrm{m}_1(S_n)\mathrm{m}_1(X_j)) &= \mathbb{E}(\mathrm{m}_1(X_j)\mathbb{E}(\mathrm{m}_1(S_n)|X_j)) \\ &= \mathbb{E}(\mathrm{m}_1(X_j)\{\mathrm{m}_1(X_j) + \mathbb{E}_*(S_n(j,e))\}) \\ &= \mathbb{E}(\mathrm{m}_1(X_j)^2) + \sum_{i\in\{1,\dots,n\}\setminus\{j\}} \mathbb{E}_*(X_j)\mathbb{E}_*(X_i).\end{aligned}$$

Iterative application of (2.3.2) to $\mathbb{E}\mathrm{m}_2(S_j \overset{\Lambda_j}{+} X_{j+1})$ $(j = 1, \dots, n-1)$ leads to

$$\begin{aligned}\mathbb{E}_2^*(S_n) &= \mathbb{E}_2^*(S_{n-1}) + 2\mathbb{E}_*(X_n)\mathbb{E}_*(S_{n-1}) + \mathbb{E}_2^*(X_n) \\ &= \mathbb{E}_2^*(S_{n-2}) + 2\mathbb{E}_*(X_{n-1})\mathbb{E}_*(S_{n-2}) + \mathbb{E}_2^*(X_{n-1}) + \\ &+ \mathbb{E}_2^*(S_{n-1}) + 2\mathbb{E}_*(X_n)\mathbb{E}_*(S_{n-1}) + \mathbb{E}_2^*(X_n) = \dots = \\ &= \sum_{i\ne j}^n \mathbb{E}_*(X_i)\mathbb{E}_*(X_j) + \sum_{j=1}^n \mathbb{E}_{*,2}(X_j)\end{aligned}$$

Therefore, we obtain from (2.3.9)

$$\mathbb{E}(Z_n^2) \le \sum_{j=1}^{n} \mathbb{E}(\mathrm{m}_2(X_j) - \mathrm{m}_1(X_j)^2).$$

On the other hand, by Proposition 2.3.5 we have $\mathbb{E}_*(S_n) = \sum_{k=1}^n \mathbb{E}_*(X_k)$ and hence $\mathbb{V}ar(Z_n) = \mathbb{E}(Z_n^2)$. □

Remark 2.3.11. a) In the situation of the theorem above, we have for identically distributed random variables X_j, $j \in \mathbb{N}$

$$\mathbb{V}ar(\mathrm{m}_1(S_n) - \sum_{j=1}^{n} \mathrm{m}_1(X_j)) \le n\mathbb{E}(\mathrm{m}_2(X_1) - \mathrm{m}_1(X_1)^2) \le n\mathbb{V}_*(X_1)$$

b) While the randomized sum S_n clearly depends on the particular choice of the underlying concretization on K, this is not the case for the conditional expectation in Proposition 2.3.9 and thus in the estimation (2.3.8).

Corollary 2.3.12. *Suppose that all assumptions of Theorem 2.3.10 hold. Let $(b_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers with limit ∞ such that $\sum_{n=1}^{\infty} \frac{1}{b_n}\mathbb{V}_*(X_n) < \infty$. Then*

$$\frac{1}{\sqrt{b_n}}\Big(\mathrm{m}_1(S_n) - \sum_{k=1}^{n} \mathrm{m}_1(X_k)\Big) \to 0 \quad \textit{in } L^2.$$

Proof: Let $Z_n := |\mathrm{m}_1(S_n) - \sum_{j=1}^n \mathrm{m}_1(X_j)|/\sqrt{b_n}$. By Theorem 2.3.10 we obtain

$$\mathbb{E}(Z_n^2) \le \frac{1}{b_n}\sum_{j=1}^{n} \mathbb{V}_*(X_j).$$

Thus, by Kronecker's Lemma, Z_n converges in L^2 to 0 as $n \to \infty$.

□

Corollary 2.3.13. *Suppose that all assumptions of Theorem 2.3.10 hold and $\sum_{n=1}^{\infty} \frac{1}{n^2}\mathbb{V}_*(X_n) < \infty$. Then*

$$\frac{1}{n}\Big(\mathrm{m}_1(S_n) - \mathbb{E}_*(S_n)\Big) \to 0 \quad \textit{in } L^2 \textit{ and a.s.}$$

Proof: By Inequality (2.3.5) we get $\sum_{n=1}^{\infty} \frac{1}{n^2}\mathbb{V}ar(\mathrm{m}_1(X_n)) < \infty$. Hence

$$\frac{1}{n}\sum_{k=1}^{n}(\mathrm{m}_1(X_k) - \mathbb{E}\mathrm{m}_1(X_k)) \to 0 \quad \mathbb{P} - a.s. \tag{2.3.10}$$

by the classical law of large numbers for independent random variables. On the other hand, according to the above corollary, we have

$$\frac{1}{n}\Big(\mathrm{m}_1(S_n) - \sum_{k=1}^{n} \mathrm{m}_1(X_k)\Big) \to 0 \quad \text{in } L^2. \tag{2.3.11}$$

Combining this with (2.3.10), the corollary follows. □

If $\mathrm{m}_1 \equiv 0$ then the statement (2.3.8) being empty. For this case we prove the following theorem.

Theorem 2.3.14. *Let* $(\mathrm{m}_k)_k$ *be a sequence of moment functions such that* $\mathrm{m}_1 \equiv 0$ *and* $0 \le \mathrm{m}_2^2 \le \mathrm{m}_4$. *Suppose that assumptions of Proposition 2.3.9 hold. Then*

$$\begin{aligned}\mathbb{E}\Big(\{\mathrm{m}_2(S_n) - \sum_{j=1}^{n} \mathrm{m}_2(X_j)\}^2\Big) \le \sum_{j=1}^{n} &\mathbb{E}(\mathrm{m}_4(X_j) - \mathrm{m}_2(X_j)^2) \\ &+ 2\sum_{j\neq i} \mathbb{E}_2^*(X_j)\mathbb{E}_2^*(X_i).\end{aligned} \tag{2.3.12}$$

Proof: Let $Z_n := \mathrm{m}_2(S_n) - \sum_{j=1}^{n} \mathrm{m}_2(X_j)$. Since $\mathrm{m}_2^2 \le \mathrm{m}_4$ and $\mathrm{m}_1 \equiv 0$ we calculate

$$\begin{aligned}\mathbb{E}(Z_n^2) \le \mathbb{E}_4^*(S_n) - 2\sum_{j=1}^{n} &\mathbb{E}(\mathrm{m}_2(S_n)\mathrm{m}_2(X_j)) + \\ + \sum_{i\neq j}^{n} \mathbb{E}_2^*(X_i)\mathbb{E}_2^*(X_j) &+ \sum_{j=1}^{n} \mathbb{E}(\mathrm{m}_2(X_j)^2).\end{aligned} \tag{2.3.13}$$

In the same way as in the proof of Theorem 2.3.10 we obtain

$$\mathbb{E}(\mathrm{m}_2(S_n)\mathrm{m}_2(X_j)) = \mathbb{E}(\mathrm{m}_2(X_j)^2) + \sum_{i\in\{1,\ldots,n\}\setminus\{j\}} \mathbb{E}_2^*(X_j)\mathbb{E}_2^*(X_i).$$

Iterative application of (2.3.2) to $\mathbb{E}(\mathrm{m}_4(S_j) \overset{\Lambda_j}{+} X_{j+1})$ $(j = 1, \ldots, n-1)$ leads to

$$\begin{aligned}\mathbb{E}_4^*(S_n) &= \mathbb{E}_4^*(S_{n-1}) + 6\mathbb{E}_2^*(X_n)\mathbb{E}_2^*(S_{n-1}) + \mathbb{E}_4^*(X_n) \\ &= \mathbb{E}_4^*(S_{n-2}) + 6\mathbb{E}_2^*(X_{n-1})\mathbb{E}_2^*(S_{n-2}) + \mathbb{E}_4^*(X_{n-1}) + \\ &+ \mathbb{E}_4^*(S_{n-1}) + 6\mathbb{E}_2^*(X_n)\mathbb{E}_2^*(S_{n-1}) + \mathbb{E}_4^*(X_n) = \ldots = \\ &= 3\sum_{i\neq j}^{n} \mathbb{E}_2^*(X_i)\mathbb{E}_2^*(X_j) + \sum_{j=1}^{n} \mathbb{E}_4^*(X_j)\end{aligned}$$

Therefore, we obtain from (2.3.13) the asserted inequality (2.3.12). □

Let $(S_n)_{n\geq 1}$ be a random walk on a hypergroup K. As in the most classical case where $\mathrm{K} = \mathbb{R}$, one can construct martingales in a canonical way from the random walk $(S_n)_{n\geq 1}$ involving moment functions on K. Here is a result for the first and second moment function due to [46], which can be easily checked.

Theorem 2.3.15. *For any increment process $(S_t)_{t\in I}$ on a hypergroup* K *we have the following statements.*

If $\mathbb{E}_*(S_t)$ exists for all $t \in I$ then $(\mathrm{m_1(S_t)} - \mathbb{E}_*(\mathrm{S_t}))_{\mathrm{t\in I}}$ is a martingale. In addition if $I := \mathbb{R}_+$ or $\mathbb{Z}_+$ and $\mathbb{E}_*(S_{t_0}) \geq 0$ for some $t_0 \in I$ then $(\mathrm{m_1(S_t)})_{\mathrm{t\in I}}$ is a submartingale.

If $\mathbb{V}_*(S_t)$ exists for all $t \in I$ then $(v(S_t, \mathbb{E}_*(S_t)) - \mathbb{V}_*(S_t))_{t\in I}$ is a martingale. In addition $(v(S_t, \mathbb{E}_*(S_t)))_{t\in I}$ is a submartingale.

In the following proposition we derive a sufficient condition for the convergence of submartingales.

Proposition 2.3.16. *Let $(X_k)_{k\in\mathbb{N}}$ be a nonnegative submartingale. If there exists an $\alpha < 2$ such that $\mathbb{E}(X_k^2) = O(k^\alpha)$ for all k, then X_k/k converges a.s. to 0. Moreover, for any $\varepsilon > 0$ one has*

$$\mathbb{P}(\sup_{m\leq k<\infty} X_k/k > \varepsilon) = O(m^{\alpha-2}).$$

Proof: For an $\varepsilon > 0$ and $m \in \mathbb{N}$ let $\tau_m := \inf\{k \in \mathbb{N} : \ X_k/k > \varepsilon \text{ and } k \geq m\}$. It is clear that

$$\begin{aligned}\varepsilon^2\mathbb{P}(m \leq \tau_m \leq n) &\leq \sum_{k=m}^{n} \frac{1}{k^2}\int_{\tau_m=k} X_k^2 d\mathbb{P} = \sum_{\mathrm{k=m}}^{\mathrm{n}} \frac{1}{\mathrm{k}^2}\Big(\int_{\tau_\mathrm{m}>\mathrm{k}-1} - \int_{\tau_\mathrm{m}>\mathrm{k}}\Big)\mathrm{X_k^2 d}\mathbb{P} \\ &\leq \int_{\tau_m>m-1} \frac{X_m^2}{m^2} d\mathbb{P} + \sum_{\mathrm{k=m}}^{\mathrm{n-1}} \Big[\frac{1}{(\mathrm{k}+1)^2}\int_{\tau_\mathrm{m}>\mathrm{k}} \mathrm{X_{k+1}^2 d}\mathbb{P} - \frac{1}{\mathrm{k}^2}\int_{\tau_\mathrm{m}>\mathrm{k}} \mathrm{X_k^2 d}\mathbb{P}\Big]\end{aligned}$$

for m, $n \in \mathbb{N}$ with $m < n$. Therefore, noting that

$$\frac{1}{(k+1)^2}\int_{\tau_m>k} X_{k+1}^2 d\mathbb{P} - \frac{1}{\mathrm{k}^2}\int_{\tau_\mathrm{m}>\mathrm{k}} \mathrm{X_k^2 d}\mathbb{P} \leq \frac{1}{(k+1)^2}\int_{\tau_m>k} (X_{k+1}^2 - X_k^2) d\mathbb{P}$$

for $m \le k \le n-1$, we get

$$\varepsilon^2 \mathbb{P}(m \le \tau_m \le n) \le \frac{1}{m^2}\mathbb{E}(X_m^2) + \sum_{k=m}^{n-1} \frac{1}{(k+1)^2} \int_{\tau_m > k} (X_{k+1}^2 - X_k^2) \mathrm{d}\mathbb{P}.$$

By submartingale property and assumption $\mathbb{E}(X_k^2) = O(k^\alpha)$ it follows that

$$\begin{aligned}
\varepsilon^2 \mathbb{P}(m \le \tau_m \le n) &\le \frac{1}{m^2}\mathbb{E}(X_m^2) + \sum_{k=m}^{n-1} \frac{1}{(k+1)^2}\mathbb{E}(X_{k+1}^2 - X_k^2) \\
&\le \sum_{k=m}^{n-1} \left[\frac{1}{k^2} - \frac{1}{(k+1)^2}\right] \cdot \mathbb{E}(X_k^2) + \frac{1}{n^2}\mathbb{E}(X_n^2) \\
&\le C \cdot \sum_{k=m}^{n-1} \frac{k^\alpha}{k^3} + \frac{1}{n^2}\mathbb{E}(X_n^2),
\end{aligned}$$

where $C > 0$ is a positive constant independent from m, n and k. Now, as $n \to \infty$, we obtain

$$\mathbb{P}(\sup_{m \le k < \infty} X_k/k > \varepsilon) \le \mathbb{P}(m \le \tau_m < \infty) \le \frac{C}{\varepsilon^2} m^{\alpha-2+\delta} \cdot \sum_{k=m}^{\infty} \frac{1}{k^{1+\delta}}$$

for $0 < \delta < 2-\alpha$. Thus, X_k/k converges a.s. to zero with the claimed convergence rate. □

Corollary 2.3.17. *Suppose that assumptions of Theorem 2.3.10 hold and X_1, $X_2, \ldots$ are identically distributed. Then $|\mathrm{m}_1(S_n) - \mathbb{E}_*(S_n)|/n$ converges a.s. to 0. Moreover, for any $\varepsilon > 0$ and $\delta > 0$ one has*

$$\mathbb{P}(\sup_{m \le k < \infty} \frac{1}{n}|\mathrm{m}_1(S_n) - \mathbb{E}_*(S_n)| > \varepsilon) = O(m^{\delta-1}).$$

Proof: Let $Z_n := \mathrm{m}_1(S_n) - \sum_{j=1}^n \mathrm{m}_1(X_j)$ and $Y_n := \sum_{j=1}^n \mathrm{m}_1(X_j) - \mathbb{E}_*(S_n)$. Since the random variables $\mathrm{m}_1(X_j)$, $j = 1, 2, \ldots$ are i.i.d., we have

$$\mathbb{E}(Y_n^2) = n \cdot \mathbb{E}((\mathrm{m}_1(X_1) - \mathbb{E}_*(X_1))^2) = O(n).$$

Thus, by Cauchy-Schwarz inequality and Theorem 2.3.10 we obtain

$$\mathbb{E}((\mathrm{m}_1(S_n) - \mathbb{E}_*(S_n))^2) \le \mathbb{E}(Z_n^2) + 2\sqrt{\mathbb{E}(Z_n^2)\mathbb{E}(Y_n^2)} + \mathbb{E}(Y_n^2) = O(n).$$

Now, the assertion follows by Proposition 2.3.16. □

Chapter 3

Moment functions and limit theorems for Jacobi hypergroups

In this chapter, we derive sharp estimates and asymptotic results for moment functions on so-called Jacobi type hypergroups on $[0,\infty[$. Moreover, we use these estimates to prove limit theorems for random walks on Jacobi hypergroups of index (α,β) where α tends to infinity. As a special case, we obtain limit results for radial, time-homogeneous random walks on hyperbolic spaces of growing dimensions.

We first collect the necessary background information on hyperbolic spaces and Jacobi hypergroups and indicate the connection between them. The material is mainly taken from [25]. We also refer to [9] and [17].

3.1 Hyperbolic spaces

Let $k \geq 2$, and $\mathbb{F} = \mathbb{R}$, $\mathbb{C}$, or the skew field of the quaternions $\mathbb{H}$ with real dimension $d = 1$, 2 or 4. We denote with $\mathrm{U}(k,\mathbb{F})$ the orthogonal, unitary or symplectic group, respectively. Moreover, we consider

$$\mathrm{U}(1,k,\mathbb{F}) := \Big\{A \in \mathrm{GL}(k+1,\mathbb{F}) : A^* I_{1,k} A = I_{1,k}\Big\},$$

with the diagonal matrix $\mathrm{I}_{1,k} = \mathrm{diag}(-1,1,\ldots,1)$. It is easy to see that $\mathrm{U}(1,k,\mathbb{F})$ is a group of right linear operators on $\mathbb{F}^{k+1}$ which leave invariant the Hermitian form

$$q(x,y) := -x_0\overline{y_0} + x_1\overline{y_1} + \ldots + x_k\overline{y_k}, \quad x,\ y \in \mathbb{F}^{k+1}.$$

The map $\theta : \mathrm{U}(1,k,\mathbb{F}) \to \mathrm{U}(1,k,\mathbb{F})$, $\theta(g) = (g^*)^{-1}$ (g^* denoting $\mathbb{F}$-hermitian adjoint of g) is an involutive automorphism of $\mathrm{U}(1,k,\mathbb{F})$. Let K_k be the subgroup of elements fixed under θ, i.e.

$$\mathrm{K}_k := \mathrm{U}(1,\mathbb{F}) \times \mathrm{U}(k,\mathbb{F}) := \left\{ \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} | x \in \mathrm{U}(1,\mathbb{F}),\ y \in \mathrm{U}(k,\mathbb{F}) \right\}.$$

The *hyperbolic space* of dimension k over $\mathbb{F}$ may be regarded as the symmetric space

$$\mathrm{H}_k(\mathbb{F}) := \mathrm{G}_k/\mathrm{K}_k,$$

where $\mathrm{G}_k := \mathrm{U}(1,k,\mathbb{F})$. In all cases, the double coset space $\mathrm{G}_k//\mathrm{K}_k$ can be regarded as the interval $[0,\infty[$ by identifying $t \geq 0$ with the double coset

$$\mathrm{K}_k a_t \mathrm{K}_k \quad \text{with} \quad a_t = \begin{pmatrix} \mathrm{ch}\,(t) & 0 & \dots & 0 & \mathrm{sh}\,(t) \\ 0 & & & & 0 \\ \vdots & & I_{k-1} & & \vdots \\ 0 & & & & 0 \\ \mathrm{sh}\,(t) & 0 & \dots & 0 & \mathrm{ch}\,(t) \end{pmatrix};$$

(see [9] and [25, Ch. 3]) We define the *hyperbolic distance* on $\mathrm{H}_k(\mathbb{F})$ by

$$\mathrm{dist}(x\mathrm{K}_k, y\mathrm{K}_k) := \varphi_k(\mathrm{K}_k y^{-1} x \mathrm{K}_k), \quad \text{for } x,\ y\ \in\ \mathrm{G}_k,$$

where $\varphi_k : \mathrm{K}_k a_t \mathrm{K}_k \mapsto t$ is the homeomorphism between $\mathrm{G}_k//\mathrm{K}_k$ and $[0,\infty[$. Moreover, let π_k be the canonical projection from $\mathrm{G}_k/\mathrm{K}_k = \mathrm{H}_k(\mathbb{F})$ onto the double coset space $\mathrm{G}_k//\mathrm{K}_k$. It is clear that $\varphi_k(\pi_k(y)) = \mathrm{dist}(y, \mathrm{K}_k)$ for all $y \in \mathrm{G}_k/\mathrm{K}_k$.

Let us now fix a probability measure $\nu \in \mathcal{M}^1([0,\infty[)$. Then there exists a unique radial (i.e. K_k-invariant) measure $\nu_k \in \mathcal{M}^1(\mathrm{H}_k(\mathbb{F}))$ with $\varphi_k(\pi_k(\nu_k)) = \nu$ (see in a more general context [42] and references cited there). In this way, we introduce the time-homogeneous radial random walks $(S_n^k)_{n\geq 0}$ associated with the ν_k by $S_0^k := K_k \in \mathrm{H}_k(\mathbb{F})$ and

$$\mathbb{P}(S_{n+1}^k \in A \cdot \mathrm{K}_k \mid S_n^k = x \cdot \mathrm{K}_k) = \nu_k(x^{-1} A \mathrm{K}_k)$$

for $n \geq 0$, $x \in \mathrm{G}_k$, and $A \subset \mathrm{G}_k$ a Borel set. It is well known (see e.g. Lemma 4.4 of [31] or [42]) that the image process $(\mathrm{dist}(S_n^k, \mathrm{K}_k))_{n\geq 1}$ is a time-homogeneous Markov chain on $[0,\infty[$ starting at time 0 in $\mathrm{dist}(S_0^k, \mathrm{K}_k) = 0$ with transition probabilities

$$\mathbb{P}(\mathrm{dist}(S_{n+1}^k, \mathrm{K}_k) \in A \mid \mathrm{dist}(S_n^k, \mathrm{K}_k) = x) = \delta_x *_{k,d} \nu(A) \tag{3.1.1}$$

for $n \geq 0$, $x \geq 0$ and $A \subset [0, \infty[$ a Borel set, where $*_{k,d}$ denotes the double coset convolution on $\mathrm{G}_k//\mathrm{K}_k \simeq [0, \infty[$.

Among other results, we shall derive the following central limit theorem for the random walk $(S_n^k)_{n\geq 0}$ on $\mathrm{H}_k(\mathbb{F})$ where for a fixed field $\mathbb{F}$, the dimension k and the number of steps n tends to infinity.

Theorem 3.1.1. *Let $(k_n)_{n\geq 1} \subset \mathbb{N}$ be an increasing sequence of dimensions with $\lim_{n\to\infty} k_n = \infty$ and fix $\mathbb{F}$ as above. Let $\nu \in \mathcal{M}^1([0, \infty[)$ with a finite second moment. For each dimension $k \geq 2$, consider the K_k-invariant time-homogeneous random walk $(S_n^k)_{n\geq 0}$ on $\mathrm{H}_k(\mathbb{F})$ such that for all n, k, the variables $\mathrm{dist}(S_{n+1}^k, S_n^k)$ have distribution ν. Then, $r_j := \int_0^\infty (\ln(\mathrm{ch}\, x))^j d\nu(x) < \infty$ exist for $j = 1, 2$, and*

$$\frac{1}{\sqrt{n}}\Big(\mathrm{dist}(S_n^{k_n}, S_0^{k_n}) - nr_1\Big)$$

tends in distribution for $n \to \infty$ to the normal distribution $\mathcal{N}(0, r_2 - r_1^2)$.

The above theorem will be proven by considering the moments of the distributions of $\Big(\mathrm{dist}(S_n^k, S_0^k)\Big)_{n\geq 0}$ on the spaces $\mathrm{G}_k//\mathrm{K}_k \simeq [0, \infty[$ equipped with the associated double coset convolutions $*_{k,\mathbb{F}}$. These convolutions may be regarded as special cases of the so-called *Jacobi convolution* $*_{\alpha,\beta}$ on $[0, \infty[$ depending on indices $\alpha \geq \beta \geq -1/2$ which were investigated by Koornwinder in [25]. The same result, but with some restrictions on the growth of $k = k(n)$ in dependence of n was derived by M. Voit in [42, 43] by using different methods.

3.2 Jacobi functions and Jacobi hypergroups

For fixed parameters $\alpha \geq \beta \geq -\frac{1}{2}$ we define the function $A_{\alpha,\beta} : \mathbb{R}_+ \to \mathbb{R}$ by

$$A_{\alpha,\beta}(x) := \mathrm{sh}\,(x)^{2\alpha+1} \mathrm{ch}\,(x)^{2\beta+1}, \quad (x \in \mathbb{R}_+).$$

It is easily seen that $A_{\alpha,\beta}$ is a Sturm-Liouville function which satisfies the conditions (2.1.2), $(SL1)$, $(SL2)$ and $(SL3)$ of the Section 2.1. The associated Sturm-Lioville operator $\mathcal{L}_{\alpha,\beta} := \mathcal{L}_{A_{\alpha,\beta}}$ is given by

$$\mathcal{L}_{\alpha,\beta} f = -f'' - \frac{A'_{\alpha,\beta}}{A_{\alpha,\beta}} f' = -f'' - \Big((2\alpha + 1)\mathrm{coth} + (2\beta + 1)\mathrm{tanh}\Big) f' \qquad (3.2.1)$$

for $f \in \mathcal{C}^2([0, \infty[)$ with $f'(0) = 0$.

According to the Theorem 2.1.12 there exists a unique hypergroup operation $*_{\alpha,\beta}$ on $\mathrm{K} := [0,\infty[$ such that $(\mathrm{K}, *_{\alpha,\beta})$ is a hypergroup; c.f. [25] as well as [7], [36], [46] in a general context of Chébli-Trimèche hypergroups. We denote $(\mathrm{K}, *_{\alpha,\beta})$ and $*_{\alpha,\beta}$ as the *Jacobi hypergroup* and the *Jacobi convolution* on $[0,\infty[$ with parameter (α,β), respectively.

For certain parameters $\alpha := \alpha(k,d)$ and $\beta := \beta(d)$ with

$$\alpha = d \cdot k/2 - 1, \qquad \beta = d/2 - 1, \tag{3.2.2}$$

the operator $-\mathcal{L}_{\alpha,\beta}$ is the radial part of the Laplace-Beltrami operator on $\mathrm{H}_k(\mathbb{F}) = \mathrm{G}_k/\mathrm{K}_k$, and the Jacobi functions $\varphi_\lambda^{(\alpha,\beta)}$ are spherical functions of the Gelfand pair $(\mathrm{G}_k, \mathrm{K}_k)$. Moreover, the double coset convolutions $*_{k,d}$ on $\mathrm{G}_k//\mathrm{K}_k \simeq [0,\infty[$ for the hyperbolic space $\mathrm{H}_k(\mathbb{F})$ are given by the Jacobi convolutions $*_{\alpha,\beta}$ with the parameters α and β as in (3.2.2).

In the following proposition, we collect some known facts about Jacobi hypergroups which can be found in [6] and [25].

Proposition 3.2.1. *Let $\alpha \geq \beta \geq -\frac{1}{2}$ and $(\mathrm{K}, *_{\alpha,\beta})$ be the Jacobi hypergroup of index (α,β). Then the neutral element of this hypergroup is 0, and the inversion is the identity mapping. Moreover, $(\mathrm{K}, *_{\alpha,\beta})$ has the following properties:*

a) The hypergroup $(\mathrm{K}, *_{\alpha,\beta})$ admits a Lebesgue absolutely continuous convolution which for Dirac measures δ_x and δ_y with $x,\ y \in [0,\infty[$ can be represented as

$$\delta_x *_{\alpha,\beta} \delta_y(f) = \int_0^1 \int_0^\pi f(\mathrm{arch}\,|\mathrm{ch}\, x \cdot \mathrm{ch}\, y + re^{i\varphi}\mathrm{sh}\, x \cdot \mathrm{sh}\, y|)\mathrm{d}m_{\alpha,\beta}(r,\varphi)$$

for $f \in \mathcal{C}_b([0,\infty[)$ and for the probability measure $m_{\alpha,\beta}$ with

$$\mathrm{d}m_{\alpha,\beta}(r,\varphi) = \frac{2\Gamma(\alpha+1)(1-r^2)^{\alpha-\beta-1}(r\sin\varphi)^{2\beta}\cdot r\mathrm{d}r\mathrm{d}\varphi}{\Gamma(1/2)\Gamma(\alpha-\beta)\Gamma(\beta+1/2)} \tag{3.2.3}$$

for $\alpha > \beta > -1/2$. For $\alpha > \beta = -1/2$, the measure degenerates into

$$\mathrm{d}m_{\alpha,-1/2}(r,\varphi) = \frac{\Gamma(\alpha+1)(1-r^2)^{\alpha-1/2}\mathrm{d}r\cdot \mathrm{d}(\delta_0+\delta_\pi)(\varphi)}{\Gamma(1/2)\Gamma(\alpha+1/2)} \tag{3.2.4}$$

and for $\alpha = \beta > -1/2$ into

$$\mathrm{d}m_{\alpha,\alpha}(r,\varphi) = \frac{2\Gamma(\alpha+1)\sin^{2\alpha}\varphi\mathrm{d}\varphi\cdot \mathrm{d}\delta_0(r)}{\Gamma(1/2)\Gamma(\alpha+1/2)}. \tag{3.2.5}$$

b) *The index of this hypergroup is given by* $\rho_{\alpha,\beta} := \alpha + \beta + 1$.

c) *The multiplicative functions are precisely the Jacobi functions*

$$\varphi_\lambda^{(\alpha,\beta)}(t) := {}_2F_1((\rho_{\alpha,\beta} - i\lambda)/2, (\rho_{\alpha,\beta} + i\lambda)/2; \alpha + 1; -\mathrm{sh}^2(t)), \quad (\lambda \in \mathbb{C})$$

which are the unique solutions to the Sturm Lioville problem

$$\mathcal{L}_{\alpha,\beta}\varphi_\lambda^{(\alpha,\beta)}(x) = (\rho_{\alpha,\beta}^2 + \lambda^2)\varphi_\lambda^{(\alpha,\beta)}, \quad \varphi_\lambda^{(\alpha,\beta)}(0) = 1, \quad \left(\varphi_\lambda^{(\alpha,\beta)}\right)'(0) = 0. \quad (3.2.6)$$

Moreover the dual of $(\mathbb{R}_+, *_{\alpha,\beta})$ *is given by*

$$\hat{\mathbb{R}}_+ = \left\{\varphi_\lambda^{(\alpha,\beta)} | \ \lambda \in \mathbb{R}_+ \cup i[0, \rho_{\alpha,\beta}]\right\}.$$

d) *The Jacobi function* $\varphi_\lambda^{(\alpha,\beta)}$ *admits the following Laplace representation*

$$\varphi_\lambda^{(\alpha,\beta)}(t) = \int_0^1 \int_0^\pi |\mathrm{ch}\, t + re^{i\varphi}\mathrm{sh}\, t|^{i\lambda - \rho_{\alpha,\beta}} \mathrm{d}m_{\alpha,\beta}(r, \varphi) \qquad (3.2.7)$$

with the probability measure $m_{\alpha,\beta}$ *introduced in (3.2.3), (3.2.4) and (3.2.5), respectively.*

e) $\omega_{\mathrm{K}} := A_{\alpha,\beta}\lambda^1_{|[0,\infty[}$ *is a Haar measure of* $(\mathrm{K}, *_{\alpha,\beta})$.

f) *Plancherel measure* π_{K} *associated with Haar measure* ω_K *is given by*

$$\mathrm{d}\pi_K(t) := \frac{1}{|c_{\alpha,\beta}(t)|^2}\mathrm{d}\lambda_{\mathrm{K}}(t)$$

with Harish-Chandra's c-function

$$c_{\alpha,\beta}(t) := \frac{\sqrt{2\pi}2^{-it}\Gamma(it)\Gamma(\alpha+1)}{\Gamma\left(\frac{\rho_{\alpha,\beta}+it}{2}\right)\Gamma\left(\frac{\rho_{\alpha,\beta}+it}{2} - \beta\right)}, \quad (t \in [0, \infty[).$$

For a fixed measure $\nu \in \mathcal{M}^1([0,\infty[)$ consider the associated random walk $(S_n^{(\alpha,\beta)})_{n\geq 0}$ with law ν on the Jacobi hypergroup of index (α,β) (on the construction of random walks on arbitrary hypergroups cf. Section 2.2). Theorem 3.1.1 then can be regarded as a special case of the following central limit theorem:

Theorem 3.2.2. *Let* $\beta \geq -\frac{1}{2}$ *and let* $(\alpha_n)_{n\in\mathbb{N}} \subset [\beta, \infty[$ *be an arbitrary increasing sequence with* $\lim_{n\to\infty} \alpha_n = \infty$. *Let* $\nu \in \mathcal{M}^1([0,\infty[)$ *with a finite second moment*

$\int_0^\infty x^2 \mathrm{d}\nu(x) < \infty$. *Consider the associated Jacobi random walk* $(S_n^{(\alpha_n,\beta)})_{n\geq 0}$ *on* $[0,\infty[$ *with law* ν. *Then*

$$\frac{1}{\sqrt{n}}\Big(S_n^{(\alpha_n,\beta)} - \mathbb{E}(S_n^{(\alpha_n,\beta)})\Big)$$

tends in distribution for $n \to \infty$ *to* $\mathcal{N}(0, r_2 - r_1^2)$, *where* r_1 *and* r_2 *are defined as in Theorem 3.1.1.*

The proof is essentially based on the general Theorem 2.3.10 as well as on some sharp estimates and asymptotic results (for large indices α) for moment functions on Jacobi hypergroups.

3.3 Moment functions on Jacobi hypergroups

Let $(\mathrm{K}, *_{\alpha,\beta})$ be a Jacobi hypergroup on $[0,\infty[$ with defining function $A_{\alpha,\beta}$ for parameters $\alpha \geq \beta \geq -1/2$. Since $(\mathrm{K}, *_{\alpha,\beta})$ belongs to the general class of Chébli-Trimèche hypergroups, the Jacobi function $\varphi_\lambda^{(\alpha,\beta)}$ is an analytic function of λ (see 2.1). Let $(\mathrm{m}_k)_{k=0,\ldots,n}$, $(n \in \mathbb{N})$ be the sequence of functions defined by

$$\mathrm{m}_k^{(\alpha,\beta)}(x) := \frac{\partial^k}{\partial \mu^k}\varphi_{i(\rho_{\alpha,\beta}+\mu)}(x)|_{\mu=0},$$

as in Definition 2.1.14. If no confusion is possible, we will suppress the parameters α, β in expressions concerning $(\mathrm{K}, *_{\alpha,\beta})$, i.e. we will simply write $*$, $\mathcal{L}$, φ_λ, ρ, $\mathrm{m}_k, \cdots$

From Section 2.1 we obtain the following facts about m_k. For $k=0$ we have $\varphi_{i\rho} \equiv 1$ and thus $\mathrm{m}_0 \equiv 1$. It is easily verified that for any $n \in \mathbb{N}$ the tuple $(\mathrm{m}_k)_{k=1,\ldots,n}$ is a sequence of moment functions of the length n in the sense of Definition 2.3.1. The cases $n=1$ and $n=2$ are proven in [47, Section 5 and 6]. By differentiating the equation (3.2.6) with respect to λ, we obtain

$$\mathcal{L}\mathrm{m}_k = -2k\rho\mathrm{m}_{k-1} - k(k-1)\mathrm{m}_{k-2}, \quad \mathrm{m}_k(0) = \mathrm{m}_k'(0) = 0 \tag{3.3.1}$$

for $k \geq 1$. It follows from the Laplace integral representation for Jacobi functions (3.2.7) that

$$\mathrm{m}_k(x) = \int_0^1 \int_0^\pi \Big(\ln|\mathrm{ch}\, x + r \cdot e^{i\varphi}\mathrm{sh}\, x|\Big)^k \mathrm{d}m_{\alpha,\beta}(r,\varphi) \tag{3.3.2}$$

for $x \geq 0$ and $k \geq 1$. In particular, m_k is non-negative.

Next, we prove a series of statements about the moments m_k, which are needed in the following section.

Lemma 3.3.1. *For all $k \in \mathbb{N}$ the functions m_k are recursively given by*

$$\mathrm{m}_k(x) = \int_0^x \int_0^y \frac{A_{\alpha,\beta}(z)}{A_{\alpha,\beta}(y)} \left(2k\rho \mathrm{m}_{k-1}(z) + k(k-1)\mathrm{m}_{k-2}(z)\right) \mathrm{d}z\mathrm{d}y. \tag{3.3.3}$$

Proof: We first notice that the integral in (3.3.3) exists as for $0 < z < y$ we have $0 \leq a_{\alpha,\beta}(z)/a_{\alpha,\beta}(y) \leq 1$. Now let w denote the function on $[0, \infty[$ defined by

$$w : [0, \infty[\to \mathbb{R}, \quad w(x) := \int_0^x b(t)F(t, x)\mathrm{d}t,$$

where $b(t) := 2k\rho \mathrm{m}_{k-1}(t) + k(k-1)\mathrm{m}_{k-2}(t)$ and $F(t, x) = A_{\alpha,\beta}(t)/A_{\alpha,\beta}(x)$. By using Leibniz integral rule, it is easy to check that the function w satisfies the differential equation

$$y' = -\frac{A'_{\alpha,\beta}}{A_{\alpha,\beta}} y + b, \quad y(0) = 0. \tag{3.3.4}$$

On the other hand, we conclude from (3.3.1) that $t \mapsto \mathrm{m}'_k(t)$ also satisfies the initial value problem (3.3.4). Hence, Picard-Lindelöf theorem leads to $w \equiv \mathrm{m}'_k$. Now, by integrating the equation above, we obtain the asserted recursion formula for m_k. □

Remark 3.3.2. Let $\alpha \geq \beta \geq -1/2$ and $x \geq 0$. By using the recursion formula (3.3.3), we obtain for the moment function of the first order

$$\mathrm{m}_1^{(\alpha,\beta)}(x) = 2\rho_{\alpha,\beta} \int_0^x \int_0^y \frac{A_{\alpha,\beta}(z)}{A_{\alpha,\beta}(y)} \mathrm{d}z\mathrm{d}y. \tag{3.3.5}$$

Let $B_{\alpha,\beta}$ denote the function $y \mapsto \int_0^y A_{\alpha,\beta}(z)\mathrm{d}z$ on $[0, x]$. For $\beta = 0$ we check at once that $B_{\alpha,0}(y) = \mathrm{sh}\,(y)^{2(\alpha+1)}/(2\alpha + 2)$ and hence

$$\mathrm{m}_1^{(\alpha,0)}(x) = 2\rho_{\alpha,0} \int_0^x \frac{B_{\alpha,0}(y)}{A_{\alpha,0}(y)} \mathrm{d}y = \ln(\mathrm{ch}\, x). \tag{3.3.6}$$

For $\alpha = \beta = -1/2$ the identity (3.3.5) makes it obvious that $\mathrm{m}_1^{(-1/2,-1/2)} \equiv 0$ and hence $\mathrm{m}_2^{(-1/2,-1/2)}(x) = x^2$ by (3.3.3).

Lemma 3.3.3. *For all $k,\ l \in \mathbb{N}$ with $l > 1$ we have*

$$\mathrm{m}_k(x)^l \le \mathrm{m}_{kl}(x) \le x^{kl} \quad \textit{for every} \quad x \ge 0.$$

In particular, m_1 *and* m_2 *satisfy the growth condition* $(MF1)$ *of Definition 2.3.1.*

Proof: According to the Laplace representation of m_k in Theorem 2.1.12, for every $x \ge 0$ there exists a probability measure ν_x on $[-x, x]$ such that $\mathrm{m}_k(x) = \int_{-x}^{x} t^k \mathrm{d}\nu_x(t)$ for all $k \in \mathbb{N}$. Furthermore, the measure τ_x with $\mathrm{d}\tau_x(t) = e^{-\rho t} \cdot \mathrm{d}\nu_x(t)$ is a symmetric subprobability measure on $\mathbb{R}$. Thus,

$$\mathrm{m}_k(x) = \int_0^x t^k (e^{\rho t} + (-1)^k e^{-\rho t}) \mathrm{d}\tau_x(t).$$

Since the integrand $t \mapsto t^k(e^{\rho t} + (-1)^k e^{-\rho t})$ is convex on $[0, \infty[$, the first inequality follows from Jensen's inequality. The second inequality is a consequence of the fact that the measure ν_x in (2.1.6) is supported by $[-x, x]$. □

Lemma 3.3.4. *For $\beta \ge -1/2$ there exists a positive constant c_β (dependent only on β) such that for all $x \ge 0$ and all $\alpha \ge \beta + 1$,*

$$\Big(1 - \frac{c_\beta}{\alpha}\Big) \ln(\operatorname{ch} x) \le \mathrm{m}_1^{(\alpha,\beta)}(x) \le \Big(1 + \frac{c_\beta}{\alpha}\Big) \ln(\operatorname{ch} x). \tag{3.3.7}$$

Proof: Let $x \in [0, \infty[$. Firstly, we consider the case $0 \le \beta \le \alpha$. By the monotonicity of ch , we see at once that $A_{\alpha,\beta}(z)/A_{\alpha,\beta}(y) \le A_{\alpha,0}(z)/A_{\alpha,0}(y)$ for all $0 \le z \le y$. Hence, from (3.3.5) and (3.3.6) we deduce

$$\mathrm{m}_1^{(\alpha,\beta)}(x) \le 2\rho_{\alpha,\beta} \int_0^x \int_0^y \frac{A_{\alpha,0}(z)}{A_{\alpha,0}(y)} \mathrm{d}z\mathrm{d}y = \frac{2\rho_{\alpha,\beta}}{2\rho_{\alpha,0}} \mathrm{m}_1^{(\alpha,0)}(x) = \Big(1 + \frac{\beta}{\alpha + 1}\Big) \ln(\operatorname{ch} x).$$

On the other hand, using $\operatorname{ch} z / \operatorname{ch} y \ge \operatorname{sh} z / \operatorname{sh} y$ for $0 \le z \le y$ we have

$$\mathrm{m}_1^{(\alpha,\beta)}(x) \ge 2\rho_{\alpha,\beta} \int_0^x \int_0^y \frac{A_{\alpha+\beta,0}(z)}{A_{\alpha+\beta,0}(y)} \mathrm{d}z\mathrm{d}y = \frac{\rho_{\alpha,\beta}}{\rho_{\alpha+\beta,0}} \mathrm{m}_1^{(\alpha+\beta,0)}(x) = \ln(\operatorname{ch} x). \tag{3.3.8}$$

We now turn to the case $\beta \in [-1/2, 0[$. In a similar manner as above, by using the inequality $A_{\alpha,\beta}(z)/A_{\alpha,\beta}(y) \le A_{\alpha+\beta,0}(z)/A_{\alpha+\beta,0}(y)$ for $0 \le z \le y$, we conclude from (3.3.5) that

$$\mathrm{m}_1^{(\alpha,\beta)}(x) \le 2\rho_{\alpha,\beta} \int_0^x \int_0^y \frac{A_{\alpha+\beta,0}(z)}{A_{\alpha+\beta,0}(y)} \mathrm{d}z\mathrm{d}y = \frac{\rho_{\alpha,\beta}}{\rho_{\alpha+\beta,0}} \mathrm{m}_1^{(\alpha+\beta,0)}(x) = \ln(\operatorname{ch} x).$$

For $0 \le z \le y$ we get $(\operatorname{ch} z/\operatorname{ch} y)^{2\beta} \ge 1$. Hence,

$$\mathrm{m}_1^{(\alpha,\beta)}(x) \ge \frac{\rho_{\alpha,\beta}}{\rho_{\alpha,0}} \mathrm{m}_1^{(\alpha,0)}(x) = \Big(1 + \frac{\beta}{\alpha+1}\Big) \ln(\operatorname{ch} x),$$

which completes the proof. □

Lemma 3.3.5. *There is a constant $C > 0$ such that for all $x \ge 0$ and all α, $\beta \in \mathbb{R}$ with $\alpha \ge \beta + 3/2 \ge 1$,*

$$-C \le \mathrm{m}_1^{(\alpha,\beta)}(x) - \ln(\operatorname{ch} x) \le C.$$

Proof: The inequality $0 \le x - \ln(\operatorname{ch} x) \le \ln(2)$ and Lemma 3.3.3 imply $\mathrm{m}_1^{\alpha,\beta}(x) - \ln(\operatorname{ch} x) \le \ln(2)$.

We now turn to the second inequality. In the case $\beta \ge 0$, we deduce from (3.3.8) that $\mathrm{m}_1^{(\alpha,\beta)}(x) - \ln(\operatorname{ch} x) \ge 0$.

Finally, we consider the case $\beta \in [-1/2, 0]$. Let $x \ge 0$ and $r \in [0, 1[$. Then

$$\begin{aligned} &\ln(1 + r\tanh(x)) \ge \ln(1+r) \ge 0 \quad \text{and} \\ &\ln(1 - r\tanh(x)) \ge \ln(1-r) \ge -(e \cdot (1-r))^{-1}. \end{aligned}$$

Therefore, using the integral representation in 3.3.2 of $\mathrm{m}_1^{(\alpha,-1/2)}$, we obtain

$$\begin{aligned} \mathrm{m}_1^{(\alpha,-1/2)}(x) - \ln(\operatorname{ch} x) &= \int_0^1 \int_0^\pi \ln(|1 + re^{i\varphi}\tanh(x)|) \mathrm{d}m_{\alpha,-1/2}(r,\varphi) \\ &\ge -\frac{\Gamma(\alpha+1)}{\Gamma(1/2)\Gamma(\alpha+1/2)} \frac{1}{e} \int_0^1 (1-r)^{-1}(1-r^2)^{\alpha-1/2} \mathrm{d}r \\ &\ge -\frac{\Gamma(\alpha+1)}{\Gamma(1/2)\Gamma(\alpha+1/2)} \frac{2}{e} \int_0^1 (1-r^2)^{\alpha-3/2} \mathrm{d}r. \end{aligned}$$

After substitution $r^2 = u$, we see that the last integral is given (up to a constant) by the beta function $\mathrm{B}(\alpha - 1/2, 1/2) = \Gamma(\alpha-1/2)\Gamma(1/2)/\Gamma(\alpha)$. Hence, using the functional equation for gamma function, we get

$$\mathrm{m}_1^{(\alpha,-1/2)}(x) - \ln(\operatorname{ch} x) \ge -\alpha/(e(\alpha - 1/2)) \ge -2/e.$$

From (3.3.5) we see at once that

$$\mathrm{m}_1^{(\alpha,\beta)}(x) \ge \rho_{\alpha,\beta} \int_0^x \int_0^y \frac{A_{\alpha+\beta+1/2,-1/2}(z)}{A_{\alpha+\beta+1/2,-1/2}(y)} \mathrm{d}z \mathrm{d}y = \mathrm{m}_1^{(\alpha+\beta+1/2,-1/2)}(x).$$

It follows that $\mathrm{m}_1^{(\alpha,\beta)}(x) - \ln(\operatorname{ch} x) \geq -1$, which completes the proof. □

Lemma 3.3.6. *There is a constant $C > 0$ such that for all $x \geq 0$ and all α, $\beta \in \mathbb{R}$ with $\alpha \geq \beta + 3/2 \geq 1$,*

$$-C \leq \left(\mathrm{m}^{(\alpha,\beta)}\right)^{-1}(x) - x \leq C.$$

Proof: By using Lemmas 3.3.3 and 3.3.5 as well as the inequality $0 \leq x - \ln(\operatorname{ch} x) \leq \ln(2)$, we obtain

$$0 \leq x - \mathrm{m}_1^{(\alpha,\beta)}(x) \leq (x - \ln(\operatorname{ch} x)) + (\ln(\operatorname{ch} x) - \mathrm{m}_1^{(\alpha,\beta)}(x)) \leq C \tag{3.3.9}$$

with a suitable constant $C > 0$. Since the graph of $(\mathrm{m}_1^{(\alpha,\beta)})^{-1}$ is obtained by reflecting the graph of $\mathrm{m}_1^{(\alpha,\beta)}$ across the line $y = x$, the inequality follows immediately from (3.3.9). □

Lemma 3.3.7. *Let $\alpha \geq \beta \geq -1/2$ with $(\alpha, \beta) \neq (-1/2, -1/2)$ and $x \geq 0$. Then*

$$\mathrm{m}_1(x)^2 \leq \mathrm{m}_2(x) \leq \mathrm{m}_1(x)^2 + \frac{1}{\rho}\mathrm{m}_1(x), \quad (x \geq 0). \tag{3.3.10}$$

Proof: Because of Lemma 3.3.3 we have only to verify the second inequality. From Remark 3.3.2 we obtain $\mathrm{m}_1'(x) = 2\rho A_{\alpha,\beta}(x)/a_{\alpha,\beta}(x)$. Hence, by Lemma 3.3.1 and integration by parts, we observe

$$\begin{aligned}
\mathrm{m}_2(x) &= 4\rho \int_0^x \int_0^y \frac{a_{\alpha,\beta}(z)}{a_{\alpha,\beta}(y)} \mathrm{m}_1(z) \mathrm{d}z \mathrm{d}y + \frac{1}{\rho}\mathrm{m}_1(x) \\
&= 4\rho \int_0^x \frac{A_{\alpha,\beta}(y)}{a_{\alpha,\beta}(y)} \mathrm{m}_1(y) \mathrm{d}y - 4\rho \int_0^x \int_0^y \frac{A_{\alpha,\beta}(z)}{a_{\alpha,\beta}(y)} \mathrm{m}_1'(z) \mathrm{d}z \mathrm{d}y + \frac{1}{\rho}\mathrm{m}_1(x) \\
&\leq 2 \int_0^x \mathrm{m}_1'(y) \mathrm{m}_1(y) \mathrm{d}y + \frac{1}{\rho}\mathrm{m}_1(x) = \mathrm{m}_1(x)^2 + \frac{1}{\rho}\mathrm{m}_1(x).
\end{aligned}$$

□

In order to formulate and to prove limit theorems in Section 3.4 the following notation is useful. For $j \in \mathbb{N}_0$, $-1/2 \leq \beta \leq \alpha$ and $\nu \in \mathcal{M}^1([0, \infty[)$ we define

$$\begin{aligned}
r_j &:= \int_0^\infty \ln(\operatorname{ch} x)^j \mathrm{d}\nu(x), & \hat{r}_j(\alpha) &:= \int_0^\infty m_j^{(\alpha,\beta)}(x) \mathrm{d}\nu(x), \\
\tilde{r}_j &:= \int_0^\infty x^j \mathrm{d}\nu(x), & \check{r}_j(\alpha) &:= \int_0^\infty m_1^{(\alpha,\beta)}(x)^j \mathrm{d}\nu(x).
\end{aligned}$$

Remark 3.3.8. a) If $\nu \in \mathcal{M}_1([0,\infty[)$ admits j-th moment, i.e. $\int_0^\infty x^j \mathrm{d}\nu(x) < \infty$, then r_j, $\hat{r}_j(\alpha)$ and $\check{r}_j(\alpha)$ exist in $[0,\infty[$, which is clear from Lemma 3.3.3.

b) From Lemmas 3.3.4 and 3.3.7 we obtain for $k = 1, 2$

$$\lim_{\alpha\to\infty} \hat{r}_k(\alpha) = r_k = \lim_{\alpha\to\infty} \check{r}_k(\alpha). \tag{3.3.11}$$

Lemma 3.3.9. *Let $k \in \mathbb{N}_0$ and $\alpha \geq -1/2$. Then*

$$\mathrm{m}_k^{(\alpha,\alpha)}(x) = 2^{-k}\mathrm{m}_k^{(\alpha,-\frac{1}{2})}(2x) \quad (x \geq 0). \tag{3.3.12}$$

Proof: The idea of the following proof goes back to Koornwinder (see Section 5.3 of [25]). For $\alpha \geq \beta \geq -1/2$ let $\mathcal{L}_{(\alpha,\beta)}$ be the differential operator as in (3.2.1). For a function $g \in \mathcal{C}^2(\mathbb{R}_+)$ with $g'(0) = 0$ we define a function $\tilde{g}$ by $\tilde{g}(t) := g(2t)$, $(t \geq 0)$. By a straightforward calculation one obtains

$$\Big(\mathcal{L}_{(\alpha,\alpha)}\tilde{g}\Big)(t) = 4\Big(\mathcal{L}_{(\alpha,-\frac{1}{2})}g\Big)(2t). \tag{3.3.13}$$

For $k = 0$ the Formula (3.3.12) is obviously true. Let $k > 0$; we set $f(t) := \mathrm{m}_k^{(\alpha,\alpha)}(t)$ and $h(t) := 2^{-k}\mathrm{m}_k^{(\alpha,-1/2)}(2t)$. Since (2.1.5) we have

$$\Big(\mathcal{L}_{(\alpha,\alpha)}f\Big)(t) = -2k(2\alpha+1)\mathrm{m}_{k-1}^{(\alpha,\alpha)}(t) - k(k-1)\mathrm{m}_{k-2}^{(\alpha,\alpha)}(t).$$

On the other hand, we calculate

$$\begin{aligned}
\Big(\mathcal{L}_{(\alpha,\alpha)}h\Big)(t) &= 4\Big(\mathcal{L}_{(\alpha,-\frac{1}{2})}\tilde{h}\Big)(2t) = 4\cdot 2^{-k}\Big(\mathcal{L}_{(\alpha,-\frac{1}{2})}\mathrm{m}_k^{(\alpha,-\frac{1}{2})}\Big)(2t) \\
&= 4\cdot 2^{-k}\Big(-2k(\alpha+\frac{1}{2})\mathrm{m}_{k-1}^{(\alpha,-\frac{1}{2})}(2t) - k(k-1)\mathrm{m}_{k-2}^{(\alpha,-\frac{1}{2})}(2t)\Big) \\
&= 4\cdot 2^{-k}\Big(-2k(\alpha+\frac{1}{2})2^{k-1}\mathrm{m}_{k-1}^{(\alpha,\alpha)}(t) - k(k-1)2^{k-2}\mathrm{m}_{k-2}^{(\alpha,\alpha)}(t)\Big) \\
&= -2k(2\alpha+1)\mathrm{m}_{k-1}^{(\alpha,\alpha)}(t) - k(k-1)\mathrm{m}_{k-2}^{(\alpha,\alpha)}(t).
\end{aligned}$$

By the uniqueness of the solution of the underlying initial value problem, we finally conclude that $f \equiv h$. □

3.4 Limit theorems for growing parameters

Let $(S_n^{(\alpha,\beta)})_{n\geq 0}$ be the time-homogeneous random walk on Jacobi hypergroup $([0,\infty[,*_{\alpha,\beta})$ with law ν. In particular, $(S_n^{(\alpha,\beta)})_{n\geq 0}$ is a time-homogeneous Markov process on $[0,\infty[$ starting in 0 with transition probability

$$\mathbb{P}(S_{n+1}^{(\alpha,\beta)} \in A \mid S_n^{(\alpha,\beta)} = x) = \delta_x *_{\alpha,\beta} \nu(A) \tag{3.4.1}$$

for $n \geq 0$, $x \geq 0$ and $A \subset [0,\infty[$ a Borel set.

In this section, we study the asymptotic behaviour of $(S_n^{(\alpha,\beta)})_{n\geq 0}$ for increasing "dimension" parameter α. From this moment onwards, we will suppose that the variables $X_1,\ X_2,\ \ldots$ are i.i.d. with finite second moment $\int_0^\infty x^2 \mathrm{d}\nu(x) < \infty$. It is known that for a fixed parameter $\beta \geq -1/2$, in the case of the finite second moment $\int_0^\infty x^2 \mathrm{d}\nu(x) < \infty$ under strong requirements on the growth of the sequence $(\alpha_n)_{n\in\mathbb{N}} \subset [\beta,\infty[$, namely $n/\sqrt{\alpha_n} \to 0$, the random variable

$$\frac{1}{\sqrt{n}}\left\{S_n^{(\alpha_n,\beta)} - n \cdot r_1\right\}$$

tends in distribution for $n \to \infty$ to the normal distribution $\mathcal{N}(0, r_2 - r_1^2)$ (see [42, Theorem 4.2]). In analogy with the radial limit theorems on $\mathbb{R}^{\alpha_n}$ for $\alpha_n \to \infty$ (see [41, 42]) one might suspect that, also in our situation, the case $n >> \alpha_n$ would establish another limit distribution as in the case $n << \alpha_n$. However, we shall prove:

Theorem 3.4.1. *Fix any $\beta \geq -1/2$ and let $(\alpha_n)_{n\in\mathbb{N}} \subset [\beta,\infty[$ be an arbitrary increasing sequence with $\lim_{n\to\infty} \alpha_n = \infty$. Let $\nu \in \mathcal{M}^1([0,\infty[)$ be a probability measure with a finite second moment $\int_0^\infty x^2 \mathrm{d}\nu(x) < \infty$ and consider the associated Jacobi random walks $(S_n^{(\alpha_n,\beta)})_{n\geq 0}$ on $[0,\infty[$ with law ν. Then*

$$\frac{1}{\sqrt{n}}\left(\mathrm{m}_1^{(\alpha_n,\beta)}(S_n^{(\alpha_n,\beta)}) - n\hat{r}_1(\alpha_n)\right) \tag{3.4.2}$$

tends in distribution for $n \to \infty$ to the normal distribution $\mathcal{N}(0, r_2 - r_1^2)$.

Proof: In the first step, we show that the random variables

$$Z_n := \frac{1}{\sqrt{n}}\left(\mathrm{m}_1^{(\alpha_n,\beta)}(S_n^{(\alpha_n,\beta)}) - \sum_{j=1}^{n} \mathrm{m}_1^{(\alpha_n,\beta)}(X_j)\right)$$

converge to zero in the L^2-sense. For this purpose, we conclude from Theorem 2.3.10 that

$$\mathbb{E}\left(Z_n^2\right) \leq \mathbb{E}\left(\mathrm{m}_2(X_1)\right) - \mathbb{E}\left(\mathrm{m}_1(X_1)^2\right) = \hat{r}_2(\alpha_n) - \check{r}_2(\alpha_n).$$

Now, the claimed convergence follows by using Remark 3.3.8.

Let Y_n denote the random variable in (3.4.2). For an $\alpha > \beta$ we define $\Xi_{n,\alpha}$ by

$$\Xi_{n,\alpha} := \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\xi_{j,\alpha} \quad \text{with} \quad \xi_{j,\alpha} := \mathrm{m}_1^{(\alpha,\beta)}(X_j) - \hat{r}_1(\alpha)$$

and denote the distribution of $\Xi_{n,\alpha}$ by $\mu_{n,\alpha}$. In the following, we write τ, τ_α instead of the normal distribution $\mathcal{N}(0, \check{r}_2(\alpha) - \hat{r}_1(\alpha)^2)$ and $\mathcal{N}(0, r_2 - r_1^2)$, respectively. Since $Y_n = Z_n - \Xi_{n,\alpha_n}$ it remains to prove that for all bounded and uniformly continuous functions $f \in \mathcal{C}_b^u(\mathbb{R})$ on $\mathbb{R}$ the integral $\int_0^\infty f\mathrm{d}\mu_{n,\alpha_n}$ tends for $n \to \infty$ to the integral $\int_0^\infty f\mathrm{d}\tau$.

Let $\varepsilon > 0$ and $f \in \mathcal{C}_b^u(\mathbb{R})$. For an $\alpha \geq -1/2$ and an $n \in \mathbb{N}$ we have

$$\begin{aligned}\Big|\int f d\mu_{n,\alpha_n} - \int f\mathrm{d}\tau\Big| \leq \Big|\int f\mathrm{d}\mu_{n,\alpha_n} - \int f\mathrm{d}\mu_{n,\alpha}\Big| + \\ + \Big|\int f\mathrm{d}\mu_{n,\alpha} - \int f\mathrm{d}\tau_\alpha\Big| + \Big|\int f\mathrm{d}\tau_\alpha - \int f\mathrm{d}\tau\Big|.\end{aligned} \tag{3.4.3}$$

Now we show that each of the three terms on the right hand side of (3.4.3) become arbitrarily small provided the involved parameters α and n are large enough. We start with the first term. Let α and γ be fixed parameters greater or equal to β. The random variables $\xi_{i,\alpha}$ and $\xi_{j,\gamma}$ are independent for any $i \neq j$ and in the case $\alpha = \gamma$, they are identically distributed. Moreover, $\xi_{j,\alpha}$ are centered with variance $\check{r}_2(\alpha) - \hat{r}_1(\alpha)^2$. From this it follows that

$$\begin{aligned}0 \leq \mathbb{E}\Big((\Xi_{n,\alpha_n} - \Xi_{n,\alpha})^2\Big) &= \mathbb{E}\Big(\Xi_{n,\alpha_n}^2 - 2\Xi_{n,\alpha_n}\Xi_{n,\alpha} + \Xi_{n,\alpha}^2\Big) \\ &= \frac{1}{n}\Big\{n\mathbb{E}(\xi_{1,\alpha_n}^2) - 2n\mathbb{E}(\xi_{1,\alpha_n}\xi_{1,\alpha}) + n\mathbb{E}(\xi_{1,\alpha}^2)\Big\} \\ &= \check{r}_2(\alpha_n) - \hat{r}_1(\alpha_n)^2 - 2\mathbb{E}(\xi_{1,\alpha_n}\xi_{1,\alpha}) + \check{r}_2(\alpha) - \hat{r}_1(\alpha)^2.\end{aligned} \tag{3.4.4}$$

By using the estimate of m_1 in Lemma 3.3.4 we get

$$\begin{aligned}\mathbb{E}(\xi_{1,\alpha_n}\xi_{1,\alpha}) &= \mathbb{E}(\mathrm{m}_1^{(\alpha_n,\beta)}(X_1)\mathrm{m}_1^{(\alpha,\beta)}(X_1)) - \hat{r}_1(\alpha_n)\hat{r}_1(\alpha) \\ &\geq (1 - \frac{c_\beta}{\alpha_n})(1 - \frac{c_\beta}{\alpha})\mathbb{E}(\ln(\operatorname{ch} X_1)^2) - \hat{r}_1(\alpha_n)\hat{r}_1(\alpha) \\ &\geq r_2 - \hat{r}_1(\alpha_n)\hat{r}_1(\alpha) - C_\beta r_2/\min(\alpha_n, \alpha),\end{aligned} \tag{3.4.5}$$

where c_β and C_β are positive constants dependent only on β. Therefore, taking

(3.4.4) and (3.4.5) into account, we obtain

$$\begin{aligned} 0 &\le \mathbb{E}\Big((\Xi_{n,\alpha_n} - \Xi_{n,\alpha})^2\Big) \qquad (3.4.6)\\ &\le \check{r}_2(\alpha_n) - \hat{r}_1(\alpha_n)^2 + \check{r}_2(\alpha) - \hat{r}_1(\alpha)^2 - 2r_2 + 2\hat{r}_1(\alpha_n)\hat{r}_1(\alpha) + \tilde{c}_\beta/\min(\alpha_n, \alpha), \end{aligned}$$

where $\tilde{c}_\beta$ is a positive constant depending only on β.

For $\delta > 0$ we set $A_\delta := \{|\Xi_{n,\alpha_n} - \Xi_{n,\alpha}| \le \delta\}$. Clearly, for an $\varepsilon > 0$ and an $f \in \mathcal{C}_b^u(\mathbb{R})$ there exists a $\delta > 0$ such that $\int_{A_\delta} |f \circ \Xi_{n,\alpha_n} - f \circ \Xi_{n,\alpha}| \mathrm{d}\mathbb{P} \le \varepsilon$. On the other hand, by Chebyshev's inequality we get

$$\int_{\Omega\setminus A_\delta} |f \circ \Xi_{n,\alpha_n} - f \circ \Xi_{n,\alpha}|\, \mathrm{d}\mathbb{P} \le 2\,\|f\|_\infty \frac{1}{\delta^2}\mathbb{E}((\Xi_{n,\alpha_n} - \Xi_{n,\alpha})^2).$$

Therefore, from (3.4.6) and (3.3.11) we conclude that there exist $n_0 := n_0(\varepsilon, f)$ > 0 and $\alpha_0 := \alpha_0(\varepsilon, f) > 0$ such that

$$\int_{\Omega\setminus A_\delta} |f \circ \Xi_{n,\alpha_n} - f \circ \Xi_{n,\alpha}| \mathrm{d}\mathbb{P} \le 2\varepsilon\,\|f\|_\infty \quad \forall\ \alpha \ge \alpha_0,\ n \ge n_0.$$

Hence, we obtain the following estimation for the first term in (3.4.3)

$$\left|\int f\mathrm{d}\mu_{n,\alpha_n} - \int f\mathrm{d}\mu_{n,\alpha}\right| \le \varepsilon(1 + 2\,\|f\|_\infty) \quad \forall\ \alpha \ge \alpha_0,\ n \ge n_0.$$

From the classical CLT, we deduce the following estimation for the second term in (3.4.3):

$$\forall\ \alpha\ \exists\ n_1 := n_1(\alpha, \varepsilon, f) : \quad \left|\int f\mathrm{d}\mu_{n,\alpha} - \int f\mathrm{d}\tau_\alpha\right| \le \varepsilon \quad \forall\ n \ge n_1.$$

Since $\check{r}_2(\alpha) - \hat{r}_1^2(\alpha)$ approaches to $r_2 - r_1^2$ as $\alpha \to \infty$, the sequence of measures $(\tau_\alpha)_\alpha$ converges weakly to τ, and thus we obtain for the last term in (3.4.3)

$$\exists\ \alpha_1 := \alpha_1(\varepsilon, f) : \quad \left|\int f\mathrm{d}\tau_\alpha - \int f\mathrm{d}\tau\right| \le \varepsilon \quad \forall\ \alpha \ge \alpha_1.$$

In summary, for $\varepsilon > 0$ and $f \in \mathcal{C}_b^u(\mathbb{R})$ there exists $n_0 := n_0(\varepsilon, f) > 0$ such that

$$\left|\int f\mathrm{d}\mu_{n,\alpha_n} - \int f\mathrm{d}\tau\right| \le \varepsilon(3 + 2\,\|f\|_\infty) \quad \forall n \ge n_0.$$

Hence, Ξ_{n,α_n} and therefore, finally Y_n, converges to $\mathcal{N}(0, r_2 - r_1^2)$. □

Corollary 3.4.2. *In the situation of Theorem 3.4.1,*

$$\frac{1}{\sqrt{n}}\left\{S_n^{(\alpha_n,\beta)} - \left(\mathrm{m}_1^{(\alpha_n,\beta)}\right)^{-1}(n\hat{r}_1(\alpha_n))\right\} \tag{3.4.7}$$

tends in distribution for $n \to \infty$ to the normal distribution $\mathcal{N}(0, r_2 - r_1^2)$.

Proof: Let $x_n := \mathrm{m}_1^{(\alpha_n,\beta)}\left(S_n^{(\alpha_n,\beta)}\right)$ and $y_n := n\hat{r}_1(\alpha_n)$. Adapted from the mean value theorem there is a ξ between x_n and y_n such that

$$\left|(x_n - y_n) - (\mathrm{m}_1^{-1}(x_n) - \mathrm{m}_1^{-1}(y_n))\right| = \left|x_n - y_n\right| \cdot \left|1 - (\mathrm{m}_1^{-1})'(\xi)\right|.$$

Since $\left(\mathrm{m}_1^{-1}\right)'(x) \searrow 1$ as $x \to \infty$ (see [46, proof of Lemma 5.7]) we obtain

$$\left|(x_n - y_n) - (\mathrm{m}_1^{-1}(x_n) - \mathrm{m}_1^{-1}(y_n))\right| \leq \left(\left(\mathrm{m}_1^{-1}\right)'\left(\min\{x_n, y_n\}\right) - 1\right) \cdot \left|x_n - y_n\right|.$$

Therefore, by the preceding theorem, the variable in (3.4.7) tends in distribution for $n \to \infty$ to the normal distribution $\mathcal{N}(0, r_2 - r_1^2)$. □

Corollary 3.4.3. *In the situation of Theorem 3.4.1,*

$$\frac{1}{\sqrt{n}}\left(S_n^{(\alpha_n,\beta)} - n\hat{r}_1(\alpha_n)\right) \tag{3.4.8}$$

tends in distribution for $n \to \infty$ to the normal distribution $\mathcal{N}(0, r_2 - r_1^2)$.

Proof: Let Y_n and Z_n denote the random variable in (3.4.7) and (3.4.8), respectively. Then, using the Lemma 3.3.6, we see that

$$Y_n - Z_n = \frac{1}{\sqrt{n}}\left(n\hat{r}_1(\alpha_n) - (\mathrm{m}_1^{(\alpha_n,\beta)})^{-1}(n\hat{r}_1(\alpha_n))\right)$$

converges to zero for $n \to \infty$. According to the Corollary 3.3.6, Y_n tends in distribution to $\mathcal{N}(0, r_2 - r_1^2)$, and hence so does Z_n. □

Remark 3.4.4. a) Theorem 3.2.2 follows by combining estimate (3.3.9) with Corollary 3.4.3.

b) Let $\mathbb{F}$ be the field of $\mathbb{R}$, $\mathbb{C}$ or the quaternions $\mathbb{H}$ with real dimension $d = 1$, 2 or 4. For a dimension $k \in \mathbb{N}$ consider the Jacobi random walk $(S_n^{(\alpha,\beta)})_{n\geq 1}$ with parameters $\alpha := \alpha(k, d)$ and $\beta := \beta(d)$ as in (3.2.2). Moreover, let $(S_n^k)_{n\geq 1}$ be the homogeneous radial random walk on hyperbolic space $\mathrm{H}_k(\mathbb{F})$

of the dimension k as it is described in Section 3.2. Since the double coset convolution $*_{k,d}$ on $\mathrm{G}_d//\mathrm{K}_d \simeq [0,\infty[$ coincides with the Jacobi convolution $*_{\alpha,\beta}$, we conclude from (3.1.1) and (3.4.1) that the process $(\mathrm{dist}(S_n^d, S_0^d))_{n\geq 1}$ is also a Jacobi random walk of parameter (α,β). Therefore, Theorem 3.1.1 is a direct consequence of the Corollary 3.4.3.

In the CLT above, β is fixed and α_n tends to infinity. It is natural to think of variants of theorem 3.4.1 for α_n and $\beta_n \to \infty$ in certain coupled ways (see [42]). Usually, such kinds of CLT no longer have a geometric interpretation. Nevertheless, we present here a CLT for the case $\beta_n \to \infty$ and $\alpha_n = \beta_n + c$ for some constant $c \geq 0$:

Theorem 3.4.5. *Let $c \geq 0$ be a constant, and let $(\beta_n)_{n\in\mathbb{N}} \subset [-1/2,\infty[$ be an arbitrary, increasing sequence of indices. Let ν be a probability measure on $[0,\infty[$ with second moment $\int_0^\infty x^2 \mathrm{d}\nu(x) < \infty$. Then, $\eta := \int_0^\infty \ln \mathrm{ch}\,(2x)\mathrm{d}\nu(x) < \infty$ and $\sigma^2 := \int_0^\infty \Big(\ln \mathrm{ch}\,(2x)\Big)^2 \mathrm{d}\nu(x) < \infty$ exist, and*

$$\frac{1}{\sqrt{n}}\Big(\mathrm{m}_1^{(\beta_n+c,\beta_n)}\Big(S_n^{(\beta_n+c,\beta_n)}\Big) - n\eta\Big)$$

tends in distribution for $n \to \infty$ to $\mathcal{N}(0, \sigma^2 - \eta^2)$.

Proof: By (3.3.5), monotonicity of sh and Lemma 3.3.9 we obtain

$$\mathrm{m}_1^{(\beta_n+c,\beta_n)}(x) \leq \frac{\rho_{\beta_n+c,\beta_n}}{\rho_{\beta_n,\beta_n}} \mathrm{m}_1^{(\beta_n,\beta_n)}(x) = \frac{\rho_{\beta_n+c,\beta}}{\rho_{\beta_n,\beta_n}} \frac{1}{2} \mathrm{m}_1^{(\beta_n,-\frac{1}{2})}(2x).$$

On the other hand, by monotonicity of ch we get

$$\mathrm{m}_1^{(\beta_n+c,\beta_n)}(x) \geq \frac{\rho_{\beta_n+c,\beta_n}}{\rho_{\beta_n+c,\beta_n+c}} \mathrm{m}_1^{(\beta_n+c,\beta_n+c)}(x) = \frac{\rho_{\beta_n+c,\beta_n}}{\rho_{\beta_n+c,\beta_n+c}} \frac{1}{2} \mathrm{m}_1^{(\beta_n+c,-\frac{1}{2})}(2x).$$

From Lemmas 3.3.4 and 3.3.7 it follows that

$$\lim_{n\to\infty} \mathrm{m}_j^{(\beta_n+c,\beta_n)}(x) = \Big(\frac{\ln \mathrm{ch}\,(2x)}{2}\Big)^j \quad \text{for } j = 1,2 \text{ and } x \geq 0.$$

The proof of Theorem 3.4.1 can now be transferred word by word to the setting above, which then leads to the proof of the assertion. □

As a by-product of our previous results, we obtain the following LLN for random walks on Jacobi type hypergroups $(\mathrm{K}, *_{\alpha,\beta})$ with growing dimension α.

Theorem 3.4.6. *Fix any* $\beta \geq -1/2$ *and let* $(\alpha_n)_{n\in\mathbb{N}} \subset [\beta, \infty[$ *be an increasing sequence for which there exists an* $\varepsilon > 0$ *such that* $n^{-\varepsilon}/\alpha_n \to 0$ *as* $n \to \infty$*. Let* $\nu \in \mathcal{M}^1([0,\infty[)$ *be a probability measure with a finite second moment* $\int_0^\infty x^2 \mathrm{d}\nu(x) < \infty$*. Then*

$$\frac{1}{n}\left\{m_1^{(\alpha_n,\beta)}(S_n^{(\alpha_n,\beta)}) - \mathbb{E}(m_1^{(\alpha_n,\beta)}(S_n^{(\alpha_n,\beta)}))\right\} \to 0 \quad \mathbb{P}-a.s.$$

Proof: Consider the random variables $Z_n := m_1^{(\alpha_n,\beta)}(S_n^{(\alpha_n,\beta)}) - \sum_{j=1}^n m_1^{(\alpha_n,\beta)}(X_j)$. By Remark 2.3.11 and Lemma 3.3.7 we have

$$0 \leq \mathbb{V}ar\Big(\frac{1}{n}Z_n\Big) \leq \frac{1}{n}\Big(\mathbb{E}\Big(\mathrm{m}_2^{(\alpha_n,\beta)}(X_1) - \mathrm{m}_1^{(\alpha_n,\beta)}(X_1)^2\Big) \leq \frac{1}{n\cdot\rho_{\alpha_n,\beta}}\mathbb{E}\Big(\mathrm{m}_1^{(\alpha_n,\beta)}(X_1)\Big).$$

Therefore, by using the assumption $n^{\varepsilon}/\alpha_n \to 0$, we conclude that

$$0 \leq \sum_{n=1}^{\infty} \mathbb{V}ar(\frac{1}{n}Z_n) \leq C\sum_{n=1}^{\infty} n^{-1-\varepsilon} < \infty$$

for some constant $C > 0$. Hence $Z_n \to 0$ $\mathbb{P}$-a.s. by the Borel-Cantelli lemma. On the other hand,

$$\frac{1}{n}\sum_{j=1}^{n}\Big(m_1^{(\alpha_n,\beta)}(X_j) - \mathbb{E}(m_1^{(\alpha_n,\beta)}(X_j))\Big) \to 0 \quad \mathbb{P}-a.s.$$

which follows from the classical law of large numbers for i.i.d. random variables with finite expected value. Combining this with the convergence of Z_n, the theorem follows. □

Corollary 3.4.7. *In the situation of the preceding theorem one has*

$$\frac{1}{n}(S_n^{(\alpha_n,\beta)} - n\hat{r}_1(\alpha_n)) \to 0 \quad \mathbb{P}-a.s.$$

Proof: The proof is carried out using the same arguments as in the proofs of Corollary 3.4.2 and 3.4.3 combined with the preceding theorem. □

Chapter 4

Radial limit theorems on $\mathbb{R}^p$ for growing dimensions p

The results in this chapter are motivated by the following problem: Let $\nu \in \mathcal{M}^1([0,\infty[)$ be a fixed probability measure. Then for each dimension $p \in \mathbb{N}$ there is a unique rotation invariant probability measure $\nu_p \in \mathcal{M}^1(\mathbb{R}^p)$ with $\varphi_p(\nu_p) = \nu$, where $\varphi_p(x) := \|x\|_2$ is the norm mapping; i.e. ν is radial part of ν_p. For each dimension $p \in \mathbb{N}$ consider i.i.d. $\mathbb{R}^p$-valued random variables $X_1^p, X_2^p, \ldots$ with law ν_p as well as the associated radial random walks

$$\left(S_n^p := \sum_{k=1}^{n} X_k^p\right)_{n\geq 0}$$

on $\mathbb{R}^p$. We are interested in finding limit theorems for the $[0,\infty[$-valued random variables $\|S_n^p\|_2$ for $n, p \to \infty$ coupled in a suitable way.

In the first part of this chapter, we derive the following two associated central limit theorems (CLTs) under disjoint growth conditions for $p = p_n$.

Theorem 4.0.8. *Assume that $\nu \in \mathcal{M}^1([0,\infty[)$ with $\nu \neq \delta_0$ admits a finite fourth moment. Let $m_k(\nu) := \int_0^\infty x^k \mathrm{d}\nu(x)$, $k \leq 4$ and $(p_n)_n$ be a sequence of dimensions with $\lim_{n\to\infty} p_n = \infty$.*
CLT I: *If $\lim_{n\to\infty} n/p_n = \infty$, then*

$$\frac{\sqrt{p_n}}{n}\left(\|S_n^{p_n}\|_2^2 - n m_2(\nu)\right)$$

tends in distribution for $n \to \infty$ to the normal distribution $\mathcal{N}(0, 2m_2(\nu))$.

CLT II: *If* $\lim_{n\to\infty} n/p_n = c \in [0,\infty[$*, then*

$$\frac{1}{\sqrt{n}}\Big(\|S_n^{p_n}\|_2^2 - nm_2(\nu)\Big)$$

tends in distribution for $n \to \infty$ *to the normal distribution* $\mathcal{N}(0, m_4(\nu) - 2cm_2(\nu)^2)$.

Remark 4.0.9. Parts of this theorem were derived in [41] by using completely different methods. More precisely, CLTs above were proven for sequences $(p_n)_n$ with some strong restrictions. The first CLT with the restriction $n/p_n^3 \to \infty$, i.e. $n >> p_n$ was identified by M. Voit as an obvious consequence of Berry-Esseen estimates on $\mathbb{R}^p$ with explicit constants depending on the dimension p, which are due to Bentkus and Götze [2, 3]. The proof of the second CLT with the restrictions $n^2/p_n \to 0$, i.e. $n << p_n$ was derived in [41] as a consequence of asymptotic properties of so called Bessel convolutions (for a survey about the Bessel convolutions on $\mathbb{R}_+$ we recommend [10]).

With the approach used in [41] one is not able to get rid of the strong conditions on the growth of $p = p_n$. In particular, the mixed case $p_n = c \cdot n$ for some constant c, which builds a bridge between the CLTs with $n << p_n$ and $n >> p_n$, was stated there as an open problem.

In the second part of this chapter, we shall show that for all sequences $(p_n)_{n\in\mathbb{N}}$ with $p_n \to \infty$,

$$\frac{1}{n}\,\|S_n^{p_n}\|_2^2 \to m_2(\nu)$$

in probability, provided the second moment of ν exists. Moreover, we derive associated strong laws of large numbers (LLNs) for $n >> p$, $n << p$ and $n \sim p$ (cf. Theorem 4.0.22). In the case $p_n >> n$, with an additional very strong restriction that the dimensions p_n grow faster than any polynomial, a strong LLN has been proved by M. Voit and M. Rösler in [30].

Remark 4.0.10. The preceding problem (c.f. the beginning of this chapter) can be generalised as follows: For a fix dimension $p \in \mathbb{N}$ the usual convolution on $\mathbb{R}^p$ induces a probability-preserving Banach-$*$-algebra isomorphism between the space $\mathcal{M}_{b,\mathrm{rad}}(\mathbb{R}^p)$ of all bounded, rotation invariant Borel measures on $\mathbb{R}^p$ and the space $\mathcal{M}_b([0,\infty[)$ of bounded Borel measures on $[0,\infty[$ via the norm map $\varphi : x \mapsto \|x\|_2$. The space $[0,\infty[$ together with this new convolution becomes a commutative orbit hypergroup; see [6] and [19]. Moreover, one can easily verify that the convolution of point measures on $[0,\infty[$, induced from $\mathbb{R}^p$, is given by

$$\delta_r *_\alpha \delta_s(f) := \frac{\Gamma(\alpha+1)}{\Gamma(1/2)\Gamma(\alpha+1/2)} \int_{-1}^{1} f(\sqrt{r^2+s^2-2rst})(1-t^2)^{\alpha-1/2}\mathrm{d}t, \quad (4.0.1)$$

with $\alpha := p/2-1$. The convolution on $\mathcal{M}^1([0,\infty[)$ is just given by bilinear, weakly continuous extension.

It was observed in [10] that equation (4.0.1) defines a commutative hypergroup $([0,\infty[, *_\alpha)$ for all indices $\alpha \geq -1/2$, where 0 is the neutral element and the involution is the identity mapping. The characters φ_λ of the hypergroup $([0,\infty[, *_\alpha)$ are defined by $\varphi_\lambda(x) := \Lambda_\alpha(\lambda x)$ for all $x \in [0,\infty[$ where Λ_α denotes the modified Bessel function of order α. This type of hypergroups and associated random walks were systematically studied by Kingman [23] and later by many others; cf. [6] and references there. Therefore, $([0,\infty[, *_\alpha)$ is called a Bessel-Kingman hypergroup of index α.

To get a first feeling of possible results, we begin with the well known case where the dimension is fixed and $n \to \infty$. Let us assume that ν admits a finite second moment $m_2(\nu) \in (0,\infty)$. The classical CLT on $\mathbb{R}^p$ implies that for $n \to \infty$, the random variables $S_n^p/\sqrt{n}$ tend in distribution to some normal distribution $\mathcal{N}(M,\Sigma)$, with $M = \mathbb{E}(X_1^p) \in \mathbb{R}^p$ and $\Sigma = \mathbb{C}ov(X_1^p) \in \mathbb{R}^{p\times p}$. Since $\nu_p = \mathbb{P}_{X_1^p}$ is rotation invariant, the covariance matrix Σ is invariant under all conjugations with respect to orthogonal transformations. Therefore, $\Sigma = c_p \mathrm{I}_p$ with identity matrix I_p and some constant c_p. In particular, $M = \mathbf{0} \in \mathbb{R}^p$. As ν is the radial part of ν_p, we obtain

$$m_2(\nu) = \int_0^\infty x^2 \mathrm{d}\nu(x) = \int_{\mathbb{R}^p} \|y\|_2^2 \, d\nu_p(y) = p \cdot c_p,$$

and hence $\Sigma = \frac{m_2(\nu)}{p}\mathrm{I}_p$. Now, the relation between standard normal distribution on $\mathbb{R}^p$ and the χ^2-distribution χ_p^2 with p degrees of freedom, clearly forces that the random variables $\frac{p}{nm_2(\nu)}\|S_n^p\|$ converge in distribution to χ_p^2. Moreover, for distribution function $F_{n,p}$ of $\frac{p}{nm_2(\nu)}\|S_n^p\|$ and F_p of the χ_p^2-distribution, one has the following Berry-Esseen-type estimation on $\mathbb{R}^p$ with explicit constants depending on the dimension p

$$\|F_{n,p} - F_p\|_\infty \leq C \cdot \frac{p^{3/2}}{\sqrt{n}}$$

for n, $p \in \mathbb{N}$ with a global constant C; cf. Theorem 2 of Bentkus [2].

On the other hand, it is well known that for χ_p^2-distributed random variables

Z_p (with $\mathbb{E}(Z_p)$ and $\mathbb{V}ar(Z_p) = 2p$), the random variables $(Z_p - p)/(2p)$ tend to the standard normal distribution $\mathcal{N}(0,1)$ for $p \to \infty$. A combination of the above results implies the first CLT under an additional condition: $n/p_n^3 \to \infty$.

Before beginning our systematic study, we need some notations. Let $f : \mathbb{C} \to \mathbb{C}$ be a function and $z = (z_1, \ldots, z_p) \in \mathbb{C}^p$. We write $f(z)$ for the tuple $(f(z_1), \ldots, f(z_p)) \in \mathbb{C}^p$.

Definition 4.0.11. Let $X = (x_1, \ldots, x_p)$ be a $\mathbb{R}^p$ valued random variable with distribution $\mu := \mathbb{P}_X \in \mathcal{M}^1(\mathbb{R}^p)$.

a) We say that μ (or X) admits a k-th moment ($k \in \mathbb{N}$) if

$$|\mu|_k := \int_{\mathbb{R}^p} \|x\|_2^k \, \mathrm{d}\mu(x) < \infty. \tag{4.0.2}$$

b) Under the condition (4.0.2), for an $\kappa \in \mathbb{N}_0^{p\times q}$ with $|\kappa| = \kappa_1 + \ldots + \kappa_p = k$ we define the *κ-th moment* of μ (and also of X) by

$$m_\kappa(\mu) := \int_{\mathbb{R}^p} z^\kappa \mathrm{d}\mu(z) \in \mathbb{R}.$$

In particular, for $\kappa \in \mathbb{N}_0$ and $\mu \in \mathcal{M}^1(\mathbb{R})$ one has $m_\kappa(\mu) = \int_{\mathbb{R}} x^\kappa \mathrm{d}\mu(x)$, as a usual κ-th moment of μ. If necessary, we may use the notation $m_\kappa(X)$ without risk of confusion.

c) For a fixed $k \in \mathbb{N}_0$, introduce the space

$$\mathcal{M}_k^1(\mathbb{R}^p) := \Big\{\mu \in \mathcal{M}^1(\mathbb{R}^p) : \quad |\mu|_k < \infty\Big\}$$

of probability measures for which all κ-moments with $|\kappa| \le k$ exist. Moreover, let $\mathcal{M}_{rad}^1(\mathbb{R}^p)$ denote the space of all radial (i.e. rotation invariant) measures on $\mathbb{R}^p$ and $\mathcal{M}_{rad,k}^1(\mathbb{R}^p) := \mathcal{M}_{rad}^1(\mathbb{R}^p) \cap \mathcal{M}_k^1(\mathbb{R}^p)$.

Here and subsequently, we consider the following geometric situation: Let $\nu \in \mathcal{M}^1([0,\infty[)$ be a fixed probability measure with finite second moment $m_2(\nu) < \infty$. Moreover, let $(p_n)_{n\ge1} \subset \mathbb{N}$ be a sequence with $\lim_{n\to\infty} p_n = \infty$. For each $n \in \mathbb{N}$ let ν_{p_n} be the unique radial probability measure on $\mathbb{R}^{p_n}$ with radial part ν, i.e. $\varphi_{p_n}(\nu_{p_n}) = \nu$, where $\varphi_{p_n} : \mathbb{R}^{p_n} \to [0,\infty[,\ x \mapsto \|x\|_2$ is the norm mapping. Now, we consider triangular arrays of independent random variables

$$\mathbb{X}_j^{p_n} = (X_j^{(1)}, \dots, X_j^{(p_n)}) : \Omega \longrightarrow \mathbb{R}^{p_n}, \quad j \le n, \ n \in \mathbb{N},$$

with radial distributions $\mathbb{X}_j^{p_n} \sim \nu_{p_n}$, $n \in \mathbb{N}$ as well as the associated partial sums $S_n^{p_n} = \sum_{j=1}^n \mathbb{X}_j^{p_n}$, $n \in \mathbb{N}$. If no confusion can arise we will omit the index p_n. The main object of study is the stochastic process

$$\Big(\Xi_n(\nu) := \|S_n^{p_n}\|_2^2 - n m_2(\nu)\Big)_{n\in\mathbb{N}}.$$

Notice that the distribution of $\Xi_n(\nu)$ depends only on the measure $\nu \in \mathcal{M}^1([0,\infty[)$.

The proof of Theorem 4.0.8 will be divided into two main steps: In the first step we prove a reduced form of Theorem 4.0.8 assuming that ν has a compact support. Along with this, we shall use the method of moments to establish the desired convergence of the distributions $\sqrt{p_n}/n \cdot \Xi_n(\nu)$ and $\Xi_n(\nu)/\sqrt{n}$, respectively. In the second step, we will show how to get rid of the support condition for ν. Both steps are based on the decomposition of $\Xi_n(\nu)$ via

$$\Xi_n(\nu) = \mathfrak{A}_n(\nu) + \mathfrak{B}_n(\nu) \tag{4.0.3}$$

where

$$\mathfrak{A}_n(\nu) := \sum_{i=1}^{n} A_i, \quad \text{with } A_i := \|\mathbb{X}_i\|^2 - m_2(\nu), \tag{4.0.4}$$

$$\text{and} \quad \mathfrak{B}_n(\nu) := \sum_{i=1}^{p_n} B_i, \quad \text{with } B_i := \sum_{\alpha,\beta=1,\dots,n;\ \alpha\neq\beta} X_\alpha^{(i)} X_\beta^{(i)}. \tag{4.0.5}$$

Next, we shall prove that the random variables $\mathfrak{A}_n(\nu)$ and $\mathfrak{B}_n(\nu)$, up to some suitable normalization factors, converge in distribution to some normal distributions (see Propositions 4.0.14 and 4.0.16). For a measure $\nu \in \mathcal{M}^1([0,\infty[)$ with compact support, we establish these convergences by the method of moments.

Theorem 4.0.12 (Method of moments)**.** *Let $Y, Y_1, Y_2, \dots$ be real valued random variables. Suppose that the distribution of Y is determined by its moments $m_k(Y)$ ($k \in \mathbb{N}$), that the Y_n have moments $m_k(Y_n)$ of all orders, and that*

$$\lim_{n\to\infty} m_k(Y_n) = m_k(Y)$$

for $k = 1, 2, \dots$. Then the sequence $(Y_n)_n$ converges to Y in distribution.

Proof: See [4][Theorem 30.2.].

By calculating the moments of $\mathfrak{A}_n(\nu)$ and $\mathfrak{B}_n(\nu)$, we shall often deal with expressions of the form $(x_1+\ldots+x_n)^k$, $k \in \mathbb{N}$. By using the multinomial formula it is a simple matter to see that

$$(x_1+\ldots+x_n)^k = \sum_{u=1}^{k} \sum_{\lambda \in C(k,u)} \frac{k!}{\lambda_1! \cdots \lambda_n!} \sum_{\mu \in W(n,u)} x_{\mu_1}^{\lambda_1} \cdots x_{\mu_u}^{\lambda_u}, \tag{4.0.6}$$

where

$$C(k,u) = \Big\{\lambda = (\lambda_1,\ldots,\lambda_u) \in \mathbb{N}^u : \quad |\lambda| := \sum_{i=1}^{u} \lambda_i = k\Big\},$$
$$W(n,u) = \{\mu = (\mu_1,\ldots,\mu_u) \in \{1,\ldots,n\}^u : \mu_1 < \mu_2 < \cdots < \mu_u\}.$$

A generalisation of Formula (4.0.6) shall be proven in Theorem 5.2.1.

In the following theorem, we compute the moments of a radial measure ν_p. The asymptotic behaviour of these moments for $p \to \infty$ shall be exploited in the proofs of Theorems 4.0.8 and 4.0.22.

Theorem 4.0.13. *Let $\kappa = (\kappa_1,\ldots,\kappa_p) \in \mathbb{N}_0^p$, $l := |\kappa|/2$, $\nu \in \mathcal{M}^1([0,\infty[)$ and $\nu_p \in \mathcal{M}^1(\mathbb{R}^p)$ be the corresponding radial probability measure on $\mathbb{R}^p$ which admits the κ-th order moment. Then*

$$m_\kappa(\nu_p) = \begin{cases} 0, & \text{if } \exists\, j: \ \kappa_j \text{ is odd}, \\ \frac{m_{2l}(\nu)}{4^l (\frac{p}{2})_l} \prod_{j=1}^{p} \frac{(2l_j)!}{l_j!}, & \text{if } \kappa = (2l_1,\ldots,2l_p), \end{cases} \tag{4.0.7}$$

where $(p/2)_l = (p/2)(p/2+1)\cdots(p/2+l-1)$ is the usual Pochhammer symbol. In particular,

$$m_\kappa(\nu_p) = O(p^{-l}) \text{ as } p \to \infty. \tag{4.0.8}$$

Proof: Let $\kappa = (\kappa_1,\ldots,\kappa_p) \in \mathbb{N}_0^p$ and $l := |\kappa|/2$. If there is a $j \in \{1,\ldots,p\}$ such that κ_j is odd, then the κ-th moment of ν_p is zero, which is due to the fact that ν_p is a radial probability measure.

Suppose that $\kappa = (2l_1,\ldots,2l_p)$ with some $l_i \in \mathbb{N}_0$. Moreover, let U_r^p denote the uniform distribution on the Euclidean sphere $S_r^{p-1} \subset \mathbb{R}^p$ with radius $r > 0$. One can easily show that ν_p enables the decomposition

$$\nu_p(\cdot) = \int_{\mathbb{R}^p} \mathrm{U}_{\varphi_p(x)}^p(\cdot)\mathrm{d}\nu_p(x) = \int_0^\infty \mathrm{U}_r^p(\cdot)\mathrm{d}\nu(r) \in \mathcal{M}^1(\mathbb{R}^p).$$

Hence, invoking Fubini's theorem, we see that

$$m_\kappa(\nu_p) = \int_{\mathbb{R}^p} x_1^{\kappa_1} \cdots x_p^{\kappa_p} \mathrm{d}\nu_p(x) = \int_0^\infty m_\kappa(\mathrm{U}_r^p)\mathrm{d}\nu(r). \tag{4.0.9}$$

In order to calculate the κ-th moment of ν_p, we first compute $\widehat{\mathrm{U}_r^p}(y)$ for $y \in \mathbb{R}^p$. Since U_r^p is a radial measure on $\mathbb{R}^p$ (i.e. invariant under rotations), there is no loss of generality in assuming $y = (0, \dots, 0, t) \in \mathbb{R}^p$, with $t > 0$. By straightforward calculation we obtain

$$\widehat{\mathrm{U}_r^p}(y) = \int_{\mathbb{R}^p} e^{itx_p} \mathrm{d}\mathrm{U}_r^p(x_1, \dots, x_p) = \frac{\Gamma(\frac{p}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{p-1}{2})} \int_{-1}^1 e^{itru(1-u^2)^{p/2-3/2}} du.$$

Let $\mathcal{J}_\nu$ and Λ_ν denote the usual and modified Bessel functions of the first kind of index ν, respectively i.e., $\Lambda_\nu(z) = \Gamma(\nu+1)\left(\frac{2}{z}\right)^\nu \mathcal{J}_\nu(z)$ with $\Re(\nu) > -1/2$ and $|\arg z| < \pi$. From integral representation [26, Formula 5.10.2] of $\mathcal{J}_\nu$ (for $\nu = p/2-1$), we see that

$$\widehat{\mathrm{U}_r^p}(y) = \Gamma(\tfrac{p}{2})\left(\tfrac{2}{tr}\right)^{p/2+1} \mathcal{J}_{\frac{p}{2}-1}(tr) = \Lambda_{\frac{p}{2}-1}(tr).$$

Now let D_κ denote the differential operator

$$\mathrm{D}_\kappa := \frac{\partial^{\kappa_1}}{\partial z_1^{\kappa_1}} \cdot \ldots \cdot \frac{\partial^{\kappa_p}}{\partial z_p^{\kappa_p}}.$$

The well known relation between moments of a probability measure and partial derivatives of its characteristic function as well as the theorem about the interchanging of differentiation and summation yield

$$\begin{aligned} m_\kappa(\mathrm{U}_r^p) &= \int_{\mathbb{R}^p} x_1^{2l_1} \cdots x_p^{2l_p} \mathrm{d}\mathrm{U}_r^p(x) = (-1)^l \mathrm{D}_\kappa \Lambda_{\frac{p}{2}-1}(r\,\|z\|)_{|z=0} \\ &= (-1)^l \mathrm{D}_\kappa \Big(\sum_{k=0}^\infty \frac{(-1)^k r^{2k}}{(p/2)_k k! 4^k} \|z\|^{2k} \Big)_{|z=0} \\ &= (-1)^l \sum_{k=0}^\infty \frac{(-1)^k r^{2k}}{(p/2)_k k! 4^k} \mathrm{D}_\kappa \left(\|z\|^{2k} \right)_{|z=0}. \end{aligned}$$

By multinomial theorem we have

$$\begin{aligned}\mathrm{D}_\kappa\Big(\|z\|^{2k}\Big)_{|z=0} &= \sum_{i_1+\ldots+i_p=k}\binom{k}{i_1,\ldots,i_p}\mathrm{D}_\kappa\Big(z_1^{2i_1}\cdots z_p^{2i_p}\Big)_{|z=0}\\ &= \delta_{l,k}\cdot\binom{k}{l_1,\ldots,l_p}(2l_1)!\cdots(2l_p)!\end{aligned}$$

where $\delta_{l,k}=1$ if $l=k$ and 0 otherwise. Thus,

$$m_\kappa(\mathrm{U}_r^p) = \frac{r^{2l}}{4^l(\frac{p}{2})_l}\prod_{j=1}^p\frac{(2l_j)!}{l_j!} \tag{4.0.10}$$

Finally, we conclude from (4.0.9) that

$$m_\kappa(\nu_p) = \frac{1}{4^l(\frac{p}{2})_l}\prod_{j=1}^p\frac{(2l_j)!}{l_j!}\int_0^\infty r^{2l}\mathrm{d}\nu(r) = \frac{m_{2l}(\nu)}{4^l(\frac{p}{2})_l}\prod_{j=1}^p\frac{(2l_j)!}{l_j!}.$$

□

The formula (4.0.7) yields the following covariance structure of a ν_{p_n} distributed random vector $\mathbb{X}=(X_1,\ldots,X_{p_n})$:

$$\mathbb{E}(X_\alpha)=0,\quad \mathbb{C}ov(X_{\alpha,},X_\beta)=\delta_{\alpha,\beta}\cdot m_2(\nu)/p_n. \tag{4.0.11}$$

Now we compute the covariance structure of $\mathfrak{A}_n(\nu)$ and $\mathfrak{B}_n(\nu)$: Since the random variables A_i $(i=1,2,\ldots)$ are independent and identically distributed, we have

$$\mathbb{E}(A_k)=0,\quad \mathbb{C}ov(A_i,A_j)=\delta_{i,j}(m_4(\nu)-m_2(\nu)^2). \tag{4.0.12}$$

This gives $\mathbb{V}ar(\mathfrak{A}_n(\nu))/n = m_4(\nu)-m_2(\nu)^2$. By the independence of random variables $\mathbb{X}_k$, $k\in\mathbb{N}$ and (4.0.11), we obtain

$$\mathbb{E}(B_k)=0,\quad \mathbb{C}ov(B_i,B_j)=\delta_{i,j}\frac{n(n-1)}{p_n^2}2m_2(\nu)^2. \tag{4.0.13}$$

Thus, we get $\lim_{n\to\infty}p_n/n^2\cdot\mathbb{C}ov(\mathfrak{B}_n(\nu))=2m_2(\nu)^2$.

We are now ready to establish CLTs for the random variables $\mathfrak{A}_n(\nu)$ and $\mathfrak{B}_n(\nu)$, respectively. We start with $\mathfrak{A}_n(\nu)$.

Proposition 4.0.14. *Assume that $\nu\in\mathcal{M}^1([0,\infty[)$ has compact support. Then the asymptotic behaviour of $\mathfrak{A}_n:=\mathfrak{A}_n(\nu)$ is given as follows:*

a) If $n/p_n\to c\in[0,\infty[$ as $n\to\infty$, then $\mathfrak{A}_n/\sqrt{n}$ tends in distribution to

$\mathcal{N}(0, m_4(\nu) - m_2(\nu)^2)$.

b) If $n/p_n \to \infty$ as $n \to \infty$, then $\sqrt{p_n}/n \cdot \mathfrak{A}_n$ tends in distribution to δ_0.

Proof: By the method of moments 4.0.12, it suffices to show that the k-th moments of $\mathfrak{A}_n/\sqrt{n}$ and $\sqrt{p_n}/n \cdot \mathfrak{A}_n$ tend to the k-th moments of the corresponding limiting distributions for every integer $k \in \mathbb{N}$. By the multinomial formula (4.0.6), we have

$$\mathfrak{A}_n^k = \sum_{u=1}^{k} \sum_{\lambda \in C(k,u)} \frac{k!}{\lambda_1! \cdots \lambda_u!} \sum_{\mu \in W(n,u)} A_{\mu_1}^{\lambda_1} \cdots A_{\mu_u}^{\lambda_u}.$$

Since the random variables A_j, $j = 1, 2, \ldots$ are identically distributed, it follows that $\mathbb{E}(A_{\mu_1}^{\lambda_1} \cdots A_{\mu_u}^{\lambda_u}) = \mathbb{E}(A_1^{\lambda_1} \cdots A_u^{\lambda_u})$ for all tuples $\mu = (\mu_1, \ldots, \mu_u) \in W(n, u)$. Thus

$$\mathbb{E}\Big(\mathfrak{A}_n^k\Big) = \sum_{u=1}^{k} \sum_{\lambda \in C(k,u)} \frac{k!}{\lambda_1! \cdots \lambda_u!} \binom{n}{u} \mathbb{E}\Big(A_1^{\lambda_1} \cdots A_u^{\lambda_u}\Big).$$

For $u \in \{1, \ldots, k\}$ and $\lambda = (\lambda_1, \ldots, \lambda_u) \in C(k, u)$ we consider

$$T(\lambda) := \binom{n}{u} \mathbb{E}\Big(A_1^{\lambda_1} \cdots A_u^{\lambda_u}\Big). \tag{4.0.14}$$

If $\lambda_\alpha = 1$ for some α, i.e. A_α appears with multiplicity one in $A_1^{\lambda_1} \cdots A_u^{\lambda_u}$, then $T(\lambda) = 0$ holds because $\mathbb{E}(A_\alpha) = 0$ and $A_1, A_2, \ldots$ are independent random variables.

Suppose that $\lambda_\alpha \geq 2$ for each α and $\lambda_\alpha > 2$ for some α. Then $k > 2u$, and since ν has a compact support, we get $T(\lambda) = O(n^u)$. This clearly forces $T(\lambda)/n^{k/2}$ and $p_n^{k/2}/n^k \cdot T(\lambda)$ in the cases $n/p_n \to c \in [0, \infty[$ and $n/p_n \to \infty$, respectively tend to zero as $n \to \infty$.

Now we turn to the case $\lambda = (2, \ldots, 2)$, in particular $k = 2u$. It follows that

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n^u} \mathbb{E}(\mathfrak{A}_n^{2u}) = \lim_{n\to\infty} \frac{(2u)!}{2! \cdots 2!} T((2, \ldots, 2)) &= \frac{(2u)!}{2^u u!} \mathbb{E}(A_1^2) \cdots \mathbb{E}(A_u^2) \\ &= 1 \cdot 3 \cdots (2u-3) \cdot (2u-1) \cdot (m_4(\nu) - m_2(\nu)^2)^u,\end{aligned}$$

i.e. the $2u$-th moment of $\mathfrak{A}_n/\sqrt{n}$ converges to the $2u$-th moment of $\mathcal{N}(0, m_4(\nu) - m_2(\nu)^2)$. Now, by the method of moments we obtain a).

If $n/p_n \to \infty$, then $p_n^{k/2}/n^k \cdot T(\lambda) = p_n^u/n^k \cdot T(\lambda)$ converges to zero as $n \to \infty$ and therefore $\sqrt{p_n^k/n^{2k}} \cdot \mathbb{E}(\mathfrak{A}_n^k)$ does so. This completes the proof. □

Now we turn to the treatment of the random variable $\mathfrak{B}_n(\nu)$. First, we need the following definition.

Definition 4.0.15. Let $p \in \mathbb{N}$. We say that a monomial

$$M : \mathbb{R}^p \to \mathbb{R}, \quad x \mapsto x^\kappa = c \cdot x^{\kappa_1} \cdots x^{\kappa_p}, \quad (c \in \mathbb{R},\ \kappa = (\kappa_1, \ldots, \kappa_p) \in \mathbb{N}_0^p)$$

is even if all κ_i are even. Moreover, we say that a polynomial $P : \mathbb{R}^p \to \mathbb{R}$ is even, if P is a linear combination of even monomials.

Proposition 4.0.16. *Assume that $\nu \in \mathcal{M}^1(\Pi_q)$ has compact support. Then the asymptotic behaviour of $\mathfrak{B}_n := \mathfrak{B}_n(\nu)$ is given as follows:*

(a) If $n/p_n \to 0$ as $n \to \infty$, then $\mathfrak{B}_n/\sqrt{n}$ tends in distribution to $\delta_{\mathbf{0}}$.

(b) If $n/p_n \to c \in]0, \infty]$ as $n \to \infty$, then $\sqrt{p_n}/n \cdot \mathfrak{B}_n$ tends in distribution to the normal distribution $\mathcal{N}(0, T^2(\nu))$.

Proof: As in the proof of Theorem 4.0.14, we show the convergence of the corresponding moments, i.e. we prove that the k-th moments $p_n^{k/2}/n^k \cdot \mathbb{E}(\mathfrak{B}_n^k)$ and $\mathbb{E}(\mathfrak{B}_n^k)/n^{k/2}$ converge to the corresponding moments of the limiting distributions as $n \to \infty$ for any $k \in \mathbb{N}$. By the multinomial formula (4.0.6), we have

$$\mathfrak{B}_n^k = \sum_{v=1}^{k} \sum_{\mu \in C(k,v)} \frac{k!}{\mu_1! \cdots \mu_v!} \sum_{\eta \in W(p_n,v)} B_{\eta_1}^{\mu_1} \cdots B_{\eta_v}^{\mu_v}.$$

Since the random vectors $\mathbb{X}_1, \mathbb{X}_2, \ldots$ are independent and identically distributed, it follows immediately from the definition of B_i $(i = 1, \ldots, p_n)$ that $B_{\eta_1}^{\mu_1} \cdots B_{\eta_v}^{\mu_v} \stackrel{d}{=} B_1^{\mu_1} \cdots B_v^{\mu_v}$ for all tuples $\eta = (\eta_1, \ldots, \eta_v) \in W(n, v)$. Thus

$$\mathbb{E}\Big(\mathfrak{B}_n^k\Big) = \sum_{v=1}^{k} \sum_{\mu \in C(k,v)} \frac{k!}{\mu_1! \cdots \mu_v!} \binom{p_n}{v} \mathbb{E}\Big(B_1^{\mu_1} \cdots B_v^{\mu_v}\Big).$$

Let $v \in \{1, \ldots, k\}$ and $\mu = (\mu_1, \ldots, \mu_v) \in C(k, v)$. To investigate the product $B_1^{\mu_1} \cdots B_v^{\mu_v}$, we introduce the set

$$\mathcal{I}_{k,n} = \Big\{(i_1, j_1, i_2, j_2, \ldots, i_k, j_k) \in \mathbb{N}^{2k} : \quad 1 \le i_\alpha, j_\alpha \le n \text{ and } i_\alpha \ne j_\alpha \ \forall\ \alpha\Big\}.$$

We see at once that

$$B_1^{\mu_1} \cdots B_v^{\mu_v} = \prod_{j=1}^{v} \Big(\sum_{\alpha \ne \beta} X_\alpha^{(1)} X_\beta^{(1)}\Big)^{\mu_j} = \sum_{I \in \mathcal{I}_{k,n}} \mathrm{S}(I, \mu) \qquad (4.0.15)$$

where for an $I = (i_1, j_1, i_2, j_2, \dots, i_k, j_k) \in \mathcal{I}_{k,n}$,

$$\mathrm{S}(I,\mu) = \prod_{\alpha=1}^{\mu_1} X_{i_\alpha}^{(1)} X_{j_\alpha}^{(1)} \cdot \prod_{\alpha=\mu_1+1}^{\mu_1+\mu_2} X_{i_\alpha}^{(2)} X_{j_\alpha}^{(2)} \cdot \ldots \cdot \prod_{\alpha=\mu_1+\ldots+\mu_{v-1}+1}^{\mu_1+\ldots+\mu_v} X_{i_\alpha}^{(v)} X_{j_\alpha}^{(v)}. \quad (4.0.16)$$

For an $a \in \{1, \dots, n\}$ and $b \in \{1, \dots, v\}$ let $\mathrm{mult}_{I,\lambda}(a,b)$ be the number of occurrences of $X_a^{(b)}$ in the product $\mathrm{S}(I,\lambda)$, i.e.

$$\mathrm{mult}_{I,\mu}(a,b) = \left|\left\{ l \in \{1, \dots, k\} : \quad X_{i_l}^{(b)} = X_a^{(b)} \text{ or } X_{j_l}^{(b)} = X_a^{(b)} \right\}\right| \in \{0, \dots, \mu_b\}.$$

The product $\mathrm{S}(I,\lambda)$ may be regarded as a monomial in the variables $X_a^{(1)}, \dots, X_a^{(v)}$, while the random variables coming from other indices are considered as constant, i.e. with the notation of $\mathrm{mult}_{I,\lambda}$ we have

$$\mathrm{S}(I,\lambda) = \prod_{b=1}^{v} (X_a^{(b)})^{\mathrm{mult}_{I,\lambda}(a,b)} \cdot R = \mathbb{X}_a^{\kappa} \cdot R,$$

where $\kappa = (\mathrm{mult}_{I,\lambda}(a,1), \dots, \mathrm{mult}_{I,\lambda}(a,v)) \in \mathbb{N}_0^v$ and R is the product of some $X_\alpha^{(\beta)}$ with $\alpha \in \{1, \dots, n\} \setminus \{a\}$, $\beta \in \{1, \dots, p_n\}$. By the independence of $\mathbb{X}_1, \mathbb{X}_2, \dots$ it is clear that $\mathbb{E}\mathrm{S}(I,\mu) = \mathbb{E}(\mathbb{X}_a^{\kappa}) \cdot \mathbb{E}(R)$. Now, if $\mathrm{S}(I,\lambda) = \mathbb{X}_a^{\kappa} \cdot R$ is not an even monomial (i.e. if there exists $b \in \mathbb{N}$ such that $\mathrm{mult}_{I,\lambda}(a,b)$ is odd), we obtain from Theorem 4.0.13 that $\mathbb{E}\mathrm{S}(I,\lambda) = \mathbb{E}(\mathbb{X}_a^{\kappa})\mathbb{E}(R) = 0$. Hence, defining

$$\mathrm{J}^o := \{ I \in \mathcal{I}_{k,n} : \quad \exists\ a,\ b \in \mathbb{N} : \quad \mathrm{mult}_{I,\lambda}(a,b) \text{ is odd } \},$$

we observe that

$$\mathbb{E}\mathrm{S}(I,\lambda) = 0 \quad \forall\ I \in \mathrm{J}^o. \quad (4.0.17)$$

Let $\mathrm{d}(I)$ be the number of distinct elements in $\{I\} := \{i_1, j_1, \dots, i_k, j_k\}$ and

$$\mathrm{J}_m := \{ I \in \mathcal{I}_{k,n} : \ \mathrm{d}(I) = m \}, \quad (m \in \mathbb{N}).$$

It is easy to check that $\mathrm{d}(I) \in \{2, \dots, 2k\}$ for all $I \in \mathcal{I}_{k,n}$ and $\mathrm{J}_m \subset \mathrm{J}^o$ for all $m > k$. Therefore,

$$\mathbb{E}\left(B_1^{\mu_1} \cdots B_v^{\mu_v}\right) = \sum_{m=2}^{k} \sum_{I \in \mathrm{J}_m} \mathbb{E}\mathrm{S}(I,\lambda). \quad (4.0.18)$$

Since $|\mathrm{J}_m| \le C \cdot n^m$ for a positive constant C, the number of terms in the last sum is at most of the order $O(n^m)$. Moreover, according to (4.0.8), each term $\mathbb{E}\mathrm{S}(I,\mu)$

in (4.0.18) is of the order $O(p_n^{-k})$. This gives

$$\sum_{I\in \mathrm{J}_m} \mathbb{E}\mathrm{S}(I,\lambda) = O\Big(\frac{n^m}{p_n^k}\Big), \quad (m\in\mathbb{N}). \tag{4.0.19}$$

For $v\in\{1,\dots,k\}$ and $\mu=(\mu_1,\dots,\mu_v)\in C(k,v)$ we consider

$$T(\mu) := \binom{p_n}{v}\mathbb{E}\Big(B_1^{\mu_1}\cdots B_v^{\mu_v}\Big).$$

If $\mu_\alpha = 1$ for some α, i.e. B_α appears with multiplicity one in $B_1^{\mu_1}\cdots B_v^{\mu_v}$, and therefore each $I\in\mathcal{I}_{k,n}$ from the Representation (4.0.15) is necessarily from J^o, and hence (4.0.17) gives $T(\mu)=0$.

Suppose that $\mu_\alpha \ge 2$ for each α and $\mu_\alpha > 2$ for some α, i.e. in particular $k>2v$. From (4.0.18) and (4.0.19) we conclude that $T(\mu) = O(n^k/p_n^{k-v})$. Thus, $n^{-k/2}T(\mu)$ and $p_n^{k/2}n^{-k}T(\mu)$ in the cases $n/p_n\to c\in[0,\infty[$ and $n/p_n\to\infty$, respectively tend to zero as $n\to\infty$.

Now we turn to the case $\mu=(2,\dots,2)$, in particular $k=2v$. From (4.0.18) and (4.0.19) we get

$$\begin{aligned}\lim_{n\to\infty}\frac{p_n^{k/2}}{n^k}T(\mu) &= \lim_{n\to\infty}\frac{p_n^{k/2}}{n^k}\binom{p_n}{v}\mathbb{E}\Big(B_1^2\cdots B_v^2\Big)\\ &= \lim_{n\to\infty}\frac{p_n^{k/2}}{n^k}\binom{p_n}{v}\sum_{I\in\mathrm{J}_k}\mathbb{E}\mathrm{S}(I,\mu)\end{aligned}$$

Since $\mathbb{X}_1,\mathbb{X}_2,\dots$ are i.i.d., we have

$$\sum_{I\in\mathrm{J}_k}\mathbb{E}\mathrm{S}(I,\mu) = \binom{n}{k}\sum_{I\in\mathrm{J}_k\cap\{1,\dots,k\}^{2k}}\mathbb{E}\mathrm{S}(I,\mu). \tag{4.0.20}$$

Let $I=(i_1,j_1,\dots,i_k,j_k)\in\mathrm{J}_k\cap\{1,\dots,k\}^{2k}$. If there is an $\alpha\in\{1,\dots,v\}$ with $\{i_{2\alpha-1},j_{2\alpha-1}\}\neq\{i_{2\alpha},j_{2\alpha}\}$, then either $\mathrm{mult}_{I,\lambda}(i_{2\alpha},\alpha)=1$ or $\mathrm{mult}_{I,\lambda}(j_{2\alpha},\alpha)=1$. Thus, by (4.0.17) we obtain $\mathbb{E}\mathrm{S}(I,\mu)=0$. Hence,

$$\sum_{I\in\mathrm{J}_k}\mathbb{E}\mathrm{S}(I,\mu) = \binom{n}{k}\sum{}' \,\mathbb{E}\mathrm{S}(I,\mu),$$

where $\sum'$ extends over the tuples $I=(i_1,j_1,\dots,i_k,j_k)\in\mathrm{J}_k\cap\{1,\dots,k\}^{2k}$ with $\{i_{2\alpha-1},j_{2\alpha-1}\}=\{i_{2\alpha},j_{2\alpha}\}$ for all $\alpha\in\{1,\dots,v\}$. An easy combinatorial computation shows that the number of terms in the sum $\sum'$ is $2^v k!$. Moreover, by independence of $\mathbb{X}_1, \mathbb{X}_2, \ldots$ and Formula (4.0.11), it is clear that

$$\mathbb{E}\mathrm{S}(I,\mu) = \prod_{j=1}^{v} \mathbb{E}\Big(\big(X_1^{(1)}\big)^2\Big)\mathbb{E}\Big(\big(X_2^{(1)}\big)^2\Big) = m_2(\nu)^k/p_n^k.$$

Summarizing what has been shown before, we get

$$\begin{aligned}
\lim_{n\to\infty} \frac{p_n^{k/2}}{n^k}\mathbb{E}\Big(\mathfrak{B}_n^k\Big) &= \lim_{n\to\infty} \frac{p_n^{k/2}}{n^k}\frac{k!}{\mu_1!\cdots\mu_v!}T(\mu) \\
&= \lim_{n\to\infty} \frac{p_n^{k/2}}{n^k}\frac{k!}{\mu_1!\cdots\mu_v!}\binom{p_n}{v}\binom{n}{k}2^v k! p_n^{-k} m_2(\nu)^k \\
&= \frac{k!}{2^v v!}(2m_2(\nu)^2)^v = 1\cdot 3\cdots(2k-1)\cdot(2m_2(\nu)^2)^v, \qquad (4.0.21)
\end{aligned}$$

i.e., the k-th moment of $\sqrt{p_n}/n \cdot \mathfrak{B}_n$ converges to that of a centered normal distribution with variance $2m_2(\nu)^2$.

The proof is completed by showing that in the case $n/p_n \to 0$ the k-th moment of $\mathfrak{B}_n/\sqrt{n}$ tends to zero as $n\to\infty$. The case where $k\in\mathbb{N}$ is odd, has already been discussed. If k is even, (4.0.21) yields

$$\lim_{n\to\infty} \frac{1}{n^{k/2}}\mathbb{E}\Big(\mathfrak{B}_n^k\Big) = \lim_{n\to\infty} \frac{n^{k/2}}{p_n^{k/2}}\frac{p_n^{k/2}}{n^k}\mathbb{E}\Big(\mathfrak{B}_n^k\Big) = 0.$$

□

The following asymptotic uncorrelation of random variables $\mathfrak{A}_n$ and $\mathfrak{B}_n$ shall play a crucial role in the subsequent proof of Theorem 4.0.8.

Proposition 4.0.17. *Assume that $\nu \in \mathcal{M}^1([0,\infty[)$ has compact support and that* $\lim_{n\to\infty} n/p_n = c \in]0,\infty[$. *Then $\mathfrak{A}_n$ and $\mathfrak{B}_n$ are asymptotically uncorrelated, i.e. for all $0\le l\le k$*

$$\frac{1}{n^{k/2}}\Big(\mathbb{E}(\mathfrak{A}_n^l)\cdot\mathbb{E}(\mathfrak{B}_n^{k-l}) - \mathbb{E}(\mathfrak{A}_n^l\cdot\mathfrak{B}_n^{k-l})\Big)$$

tends to zero as $n\to\infty$.

Proof: From multinomial Formula (4.0.6), by using symmetry argument, we conclude

$$F_n(l,k) := \frac{1}{n^{k/2}}\Big(\mathbb{E}(\mathfrak{A}^l)\cdot\mathbb{E}(\mathfrak{B}^{k-l}) - \mathbb{E}(\mathfrak{A}^l\cdot\mathfrak{B}^{k-l})\Big)$$
$$= \frac{1}{n^{k/2}}\sum_{u=1}^{l}\sum_{\lambda\in C(l,u)}\frac{l!}{\lambda_1!\cdots\lambda_u!}\binom{n}{u}\cdot\sum_{v=1}^{k-l}\sum_{\mu\in C(k-l,u)}\frac{(k-l)!}{\mu_1!\cdots\mu_v!}\binom{p_n}{v}H_{\lambda,\mu},$$

with

$$H_{\lambda,\mu} = \mathbb{E}\Big(A_1^{\lambda_1}\cdots A_u^{\lambda_u}\Big)\cdot\mathbb{E}\Big(B_1^{\mu_1}\cdots B_v^{\mu_v}\Big) - \mathbb{E}\Big(A_1^{\lambda_1}\cdots A_u^{\lambda_u}\cdot B_1^{\mu_1}\cdots B_v^{\mu_v}\Big).$$

Let us keep the notations of Proposition 4.0.16. We have $B_1^{\mu_1}\cdots B_v^{\mu_v} = \sum' \mathrm{S}(I,\mu)$, where $\sum'$ extends over the tuples $I = (i_1, j_1, \ldots, i_{k-l}, j_{k-l}) \in \mathcal{I}_{k-l,n}$ and $\mathrm{S}(I,\mu)$ is defined as in (4.0.16).

If $\mu_\alpha = 1$ for some $\alpha \in \{1,\ldots,v\}$, then each term $\mathrm{S}(I,\mu)$ in the sum above is not an even monomial ($\mathrm{mult}_{I,\mu}(a,\mu_\alpha) = 1$ for some $a \in \mathbb{N}$) and thus, neither is $A_1^{\lambda_1}\cdots A_u^{\lambda_u}\cdot\mathrm{S}(I,\mu)$. Therefore, $H_{\lambda,\mu} = 0$ by Theorem 4.0.13.

Suppose that $\mu_\alpha \geq 2$ for each α. By Eq. (4.0.18) we have

$$H_{\lambda,\mu} = \sum_{I\in \mathrm{J}_2\cup\ldots\cup\mathrm{J}_{k-l}}\Big\{\mathbb{E}\Big(A_1^{\lambda_1}\cdots A_u^{\lambda_u}\Big)\cdot\mathbb{E}\mathrm{S}(I,\mu) + \\ - \mathbb{E}\Big(A_1^{\lambda_1}\cdots A_u^{\lambda_u}\cdot\mathrm{S}(I,\mu)\Big)\Big\}. \tag{4.0.22}$$

For $M := \{1,\ldots,u\}$ and $G := \{\alpha \in M:\ \lambda_\alpha = 1\}$ we define the following subsets of $\mathcal{I}_{k-l,n}$:

$$\mathrm{J}_M := \{I \in \mathrm{J}_2\cup\ldots\cup\mathrm{J}_{k-l}:\quad \{I\}\cap M \neq \emptyset\},$$
$$\mathcal{K}_G := \{I \in \mathrm{J}_2\cup\ldots\cup\mathrm{J}_{k-l}:\quad G\subset\{I\}\}.$$

It is easily checked that for the cardinalities of J_M and $\mathcal{K}_G$, we have

$$|\mathrm{J}_M| \leq Cn^{k-l-1} \quad\text{and}\quad |\mathcal{K}_G| \leq Cn^{k-l-|G|} \tag{4.0.23}$$

with some constant $C = C(k,l)$.

We consider the I-th term in the sum (4.0.22), which will be denoted by $T(I)$. Is $I \notin \mathrm{J}_M$, i.e. $\{I\}\cap M = \emptyset$, and thus $A_1,\ldots,A_u$ are independent from $\mathrm{S}(I,\mu)$. This clearly forces $T(I) = 0$. Is $I \notin \mathcal{K}_G$, i.e. there exists $\tau \in G$ with $\tau \notin \{I\}$, and therefore A_τ is independent from A_i ($i \in M\setminus\{\tau\}$) and $\mathrm{S}(I,\mu)$. Thus, we get $T(I) = 0$ from (4.0.12).

Taking (4.0.23) into account, we see that the number of nonzero summands in (4.0.22) is bounded above $\min(n^{k-l-1}, n^{k-l-|G|})$. On the other hand, Theorem 4.0.13 yields that each of them is bounded above C/p_n^{k-l} where $C > 0$ is a suitable global constant. In summary, we get

$$|H_{\lambda,\mu}| \leq C \cdot \min(n^{-1}, n^{-|G|}). \tag{4.0.24}$$

Since $\mu \in C(k-l, v)$ with $\mu_\alpha \geq 2$ for all $\alpha \in \{1, \cdots, v\}$ we have $k - l \geq 2v$. Moreover, since $\lambda \in C(l, u)$ we get $l \geq 2u - |G|$. Hence, by straightforward calculation using $n/p_n \to c \in]0, \infty[$, we conclude from (4.0.24) that for suitable constants C_i,

$$\begin{aligned}|F_n(l,k)| &\leq \frac{C_1}{n^{k/2}} \sum_{u=1}^{l} \sum_{v=1}^{k-l} \sum_{\lambda \in C(l,u)} \sum_{\mu \in C(k-l,v)} \binom{n}{u} \binom{p_n}{v} \min(n^{-1}, n^{-|G|}) \\ &\leq \frac{C_2}{n^{k/2}} \sum_{u=1}^{l} \sum_{v=1}^{k-l} \sum_{\lambda \in C(l,u)} n^{u+v} \min(n^{-1}, n^{-|G|}) \leq \frac{C_3}{\sqrt{n}}.\end{aligned}$$

This completes the proof. □

Proof of Theorem 4.0.8 for $\nu \in \mathcal{M}^1([0,\infty[)$ with compact support: If $n/p_n \to \infty$ then $\sqrt{p_n}/n \cdot \mathfrak{A}_n \xrightarrow{d} \delta_0$ and $\sqrt{p_n}/n \cdot \mathfrak{B}_n \xrightarrow{d} \mathcal{N}(0, 2m_2(\nu)^2)$ according to Propositions 4.0.14 and 4.0.16. This clearly forces $\sqrt{p_n}/n \cdot \Xi_n(\nu) \xrightarrow{d} \mathcal{N}(0, 2m_2(\nu)^2)$ by Slutsky's Theorem. Suppose that $n/p_n \to 0$. Then we get as above $\Xi_n(\nu)/\sqrt{n} \xrightarrow{d} \mathcal{N}(0, m_4(\nu) - m_2(\nu)^2)$. It only remains to check the convergence in the case $n/p_n \to c \in]0, \infty[$. Let $k \in \mathbb{N}$. By (4.0.3) and Proposition 4.0.17

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}\Big(\big(\tfrac{1}{\sqrt{n}}\Xi_n(\nu)\big)^k\Big) &= \lim_{n\to\infty} \frac{1}{n^{k/2}} \sum_{l=0}^{k} \binom{k}{l} \mathbb{E}\left(\mathfrak{A}_n^l \mathfrak{B}_n^{k-l}\right) \\ &= \lim_{n\to\infty} \frac{1}{n^{k/2}} \sum_{l=0}^{k} \binom{k}{l} \mathbb{E}\left(\mathfrak{A}_n^l\right) \cdot \mathbb{E}\left(\mathfrak{B}_n^{k-l}\right).\end{aligned}$$

Consider independent random variables Z_i with distributions $P_{Z_i} = \mathcal{N}(0, \sigma_i^2)$, $i = 1, 2$. It is clear that $Z_1 + Z_2$ is $\mathcal{N}(0, \sigma_1^2 + \sigma_2^2)$ distributed. Furthermore, for the k-th moment of $Z_1 + Z_2$ we have

$$\begin{aligned} m_k(\mathcal{N}(0, \sigma_1^2 + \sigma_2^2)) &= \sum_{l=0}^{k} \binom{k}{l} \mathbb{E}(Z_1^l)\mathbb{E}(Z_2^{k-l}) \\ &= \sum_{l=0}^{k} \binom{k}{l} m_l(\mathcal{N}(0, \sigma_1^2)) m_{k-l}(\mathcal{N}(0, \sigma_2^2)). \end{aligned}$$

Propositions 4.0.14 and 4.0.16 now lead to

$$\lim_{n\to\infty} \mathbb{E}\Big(\big(\tfrac{1}{\sqrt{n}}\Xi_n(\nu)\big)^k\Big) = m_k(\mathcal{N}(0, m_4(\nu) - m_2(\nu)^2 + 2cm_2(\nu)^2)).$$

□

In order to get rid of the assumption that $\mathrm{supp}(\nu)$ is compact, we introduce for an $a > 0$ the truncated $\mathbb{R}^{p_n}$-valued random variables

$$\mathbb{X}_{k,a}^{p_n} := \begin{cases} \mathbb{X}_k^{p_n}, & \text{if } \|\mathbb{X}_k^{p_n}\|_2 \leq a, \\ \mathbf{0}, & \text{otherwise} \end{cases} \qquad k = 1, 2, \ldots$$

Let us denote by ν_a the distribution of $\varphi_{p_n}(\mathbb{X}_{1,a}^{p_n})$ (which is not dependent on p_n). Obviously, the sequence $\mathbb{X}_{k,a}^{p_n}$, $k \in \mathbb{N}$, are i.i.d. with the radial law $\nu_{p_n,a} \in \mathcal{M}^1(\mathbb{R}^{p_n})$ which corresponds to $\nu_a \in \mathcal{M}^1([0, \infty[)$. We define $\Xi_n(\nu_a)$, $\mathfrak{A}_n(\nu_a)$, $A_{j,a}$ ($j = 1, \ldots, n$), $\mathfrak{B}_n(\nu_a)$ and $B_{j,a}$ ($j = 1, \ldots, p_n$) according to (4.0.3), (4.0.4) and (4.0.5), respectively by taking $\mathbb{X}_{k,a}^{p_n}$ instead of $\mathbb{X}_k^{p_n}$, $k \in \mathbb{N}$. As it used to be, we will omit the index p_n when no confusion can arise. Clearly, we have $\Xi_n(\nu_a) = \mathfrak{A}_n(\nu_a) + \mathfrak{B}_n(\nu_a)$.

In the following we show that $\Xi_n(\nu_a)$ is a "good" approximation of $\Xi_n(\nu)$. To formulate this exactly, we first fix some $\delta > 0$ and a sequence $(p_n)_n$; we then introduce the sequence $(\delta_n)_n$ by

$$\delta_n := \begin{cases} \delta \cdot \sqrt{n}, & \text{if } \frac{n}{p_n} \to c \in [0, \infty[, \\ \delta \cdot \frac{n}{\sqrt{p_n}}, & \text{if } \frac{n}{p_n} \to \infty. \end{cases} \tag{4.0.25}$$

In the next lemmas we show that the events

$$\{|\mathfrak{A}_n(\nu_a) - \mathfrak{A}_n(\nu)| > \delta_n\} \text{ and } \{|\mathfrak{B}_n(\nu_a) - \mathfrak{B}_n(\nu)| > \delta_n\}$$

have arbitrary small probabilities for a and n large enough.

Lemma 4.0.18. *For all $\varepsilon > 0$, $\delta > 0$ there exist a_0, $n_0 \in \mathbb{N}$ such that for all*

$n,\ a \in \mathbb{N}$ *with* $a \geq a_0$ *and* $n \geq n_0$

$$\mathbb{P}\left(|\mathfrak{A}_n(\nu) - \mathfrak{A}_n(\nu_a)| > \delta_n\right) \leq \varepsilon.$$

Proof: Let $\delta > 0$ and $(\delta_n)_n$ be a sequence as in (4.0.25). Since $(A_i - A_{i,a})$, $(i = 1, 2, \ldots)$ are i.i.d., it follows by Chebychev inequality that

$$\mathbb{P}\left(|\mathfrak{A}_n(\nu) - \mathfrak{A}_n(\nu_a)| \geq \delta_n\right) \leq \frac{n}{\delta_n^2} \mathbb{V}ar\left(A_1 - A_{1,a}\right). \tag{4.0.26}$$

Using triangle inequality, we obtain

$$\sup\left\{A_{1,a}^2 :\ a \in \mathbb{N}\right\} \leq \left(\|\mathbb{X}_1\|_2^2 + m_2(\nu)\right)^2 \in L^1(\Omega),$$

Therefore, the set $\{A_{1,a}^2 :\ a \in \mathbb{N}\}$ is uniformly integrable. On the other hand, since the random variable $|A_1|$ is almost surely finite, $A_{1,a}$ converges almost surely to A_1 as $a \to \infty$. Thus, we get

$$A_{1,a}^2 \longrightarrow A_1^2 \quad \text{in } L^1. \tag{4.0.27}$$

By taking (4.0.26) and (4.0.27) into account, the lemma follows. □

Lemma 4.0.19. *For all* $\varepsilon > 0$, $\delta > 0$ *there exist* $a_0,\ n_0 \in \mathbb{N}$ *such that for all* $n,\ a \in \mathbb{N}$ *with* $a \geq a_0$ *and* $n \geq n_0$

$$\mathbb{P}\left(|\mathfrak{B}_n(\nu) - \mathfrak{B}_n(\nu_a)| > \delta_n\right) \leq \varepsilon. \tag{4.0.28}$$

Proof: Let $\delta > 0$ and $(\delta_n)_n$ be a sequence as in (4.0.25). By Chebychev inequality it follows that

$$\mathbb{P}\left(|\mathfrak{B}_n(\nu) - \mathfrak{B}_n(\nu_a)| \geq \delta_n\right) \leq \frac{1}{\delta_n^2} \sum_{j,i=1}^{p_n} \mathbb{C}ov\left(B_i - B_{i,a}, B_j - B_{j,a}\right). \tag{4.0.29}$$

Since $\mathbb{X}_1, \mathbb{X}_2, \ldots$ and $\mathbb{X}_{1,a}, \mathbb{X}_{2,a}, \ldots$ are both sequences of $\mathbb{R}^{p_n}$-valued i.i.d. random variables with radial distributions ν and ν_a, respectively, one can easily see that

$$\mathbb{C}ov\left(X_l^{(i)}, X_{k,a}^{(j)}\right) = \delta_{l,k} \cdot \delta_{i,j} \cdot \mathbb{C}ov\left(X_1^{(1)}, X_{1,a}^{(1)}\right) \tag{4.0.30}$$

holds for all indices l, k, i and j. Moreover, by (4.0.11) we obtain

$$\mathbb{C}ov\left(X_1^{(1)}, X_{1,a}^{(1)}\right) = \frac{m_2(\nu)}{p_n} - \int_{\{\|\mathbb{X}_1\| > a\}} X_1^{(1)} X_1^{(1)} \mathrm{d}\mathbb{P}.$$

Cauchy-Schwarz inequality and Theorem 4.0.13 lead to

$$\int_{\{\|\mathbb{X}_1\|>a\}} X_1^{(l)} X_1^{(1)} \mathrm{d}\mathbb{P} \leq \sqrt{\mathbb{P}\Big(\|\mathbb{X}_1\| > a\Big)} \cdot \sqrt{m_{(4,0,\ldots,0)}(\nu_{p_n})} \leq \frac{C}{a \cdot p_n}$$

with a suitable constant $C := C(\nu) > 0$. Thus, we get

$$\begin{aligned}\mathbb{C}ov\Big(B_i, B_{j,a}\Big) &= \sum_{l_1 \neq l_2} \sum_{k_1 \neq k_2} \mathbb{C}ov\Big(X_{l_1}^{(i)} \cdot X_{l_2}^{(i)}, X_{k_1,a}^{(j)} \cdot X_{k_2,a}^{(j)}\Big) \\ &= \delta_{i,j} \cdot 2n(n-1) \cdot \mathbb{C}ov\Big(X_1^{(1)}, X_{1,a}^{(1)}\Big)^2 \\ &\geq \delta_{i,j} \cdot 2n(n-1) \cdot \Big(\frac{m_2(\nu)}{p_n} - \frac{C}{a \cdot p_n}\Big)^2.\end{aligned}$$

Hence, Equation (4.0.13) yields

$$\begin{aligned}0 &\leq \mathbb{C}ov\left(B_i - B_{i,a}, B_j - B_{j,a}\right) = \delta_{i,j} \cdot \mathbb{C}ov\left(B_1 - B_{1,a}, B_1 - B_{1,a}\right) \\ &= \delta_{i,j} \cdot \Big(\mathbb{V}ar(B_1) - 2\mathbb{C}ov(B_1, B_{1,a}) + \mathbb{V}ar(B_{1,a})\Big) \\ &\leq \delta_{i,j} \cdot 2n(n-1)\frac{1}{p_n^2}\Big(m_2(\nu)^2 - 2 \cdot \Big(m_2(\nu) - C/a\Big)^2 + m_2(\nu_a)^2\Big) \\ &\leq \delta_{i,j} \cdot 2n(n-1)\frac{1}{p_n^2}\Big(m_2^2(\nu_a) - m_2(\nu) + 4Cm_2^2(\nu)/a - 2C^2/a^2\Big). \qquad (4.0.31)\end{aligned}$$

Taking (4.0.29) and (4.0.31) into account, we obtain (4.0.28). □

Corollary 4.0.20. *For all $\varepsilon > 0$, $\delta > 0$ there exist $a_0,\ n_0 \in \mathbb{N}$ such that for all $n,\ a \in \mathbb{N}$ with $a \geq a_0$ and $n \geq n_0$*

$$\mathbb{P}\left(|\Xi_n(\nu) - \Xi_n(\nu_a)| > \delta_n\right) \leq \varepsilon.$$

Proof: For an $\delta > 0$ we observe

$$\mathbb{P}\left(|\Xi_n(\nu) - \Xi_n(\nu_a)| > \delta_n\right) \leq \mathbb{P}\Big(|\mathfrak{A}_n - \mathfrak{A}_{n,a}| > \frac{\delta_n}{2}\Big) + \mathbb{P}\Big(|\mathfrak{B}_n - \mathfrak{B}_{n,a}| > \frac{\delta_n}{2}\Big).$$

Combining this with Lemmas 4.0.18 and 4.0.19, the corollary follows. □

Proof of Theorem 4.0.8: Let us first prove the CLT I. In this case the normalization is given by $\sqrt{p_n}/n$ and for the growth of p_n we have the condition $n/p_n \to \infty$ as $n \to \infty$. We set $\xi_n := \sqrt{p_n}/n \cdot \Xi_n(\nu)$, and $\xi_{n,a} = \sqrt{p_n}/n \cdot \Xi_n(\nu_a)$ and denote their distributions by μ_n and $\mu_{n,a}$, respectively. Moreover, we write τ_ν instead of

$\mathcal{N}(0, 2m_2(\nu))$. Using triangle inequality, we deduce that

$$\left|\int f \mathrm{d}\mu_n - \int f \mathrm{d}\tau_\nu\right| \le \left|\int f \mathrm{d}\mu_n - \int f \mathrm{d}\mu_{n,a}\right| + \\ + \left|\int f \mathrm{d}\mu_{n,a} - \int f \mathrm{d}\tau_{\nu_a}\right| + \left|\int f \mathrm{d}\tau_{\nu_a} - \int f \mathrm{d}\tau_\nu\right|. \tag{4.0.32}$$

Let $\varepsilon > 0$, $f \in \mathcal{C}_b^u([0, \infty[)$ be a bounded uniformly continuous function on $[0, \infty[$ and $A_\delta := \{|\xi_n - \xi_{n,a}| \le \delta\}$ $(\delta > 0)$. It follows that

$$\exists\ \delta > 0 : \quad \int_{A_\delta} |f \circ \xi_n - f \circ \xi_{n,a}| \,\mathrm{d}\mathbb{P} \le \varepsilon.$$

On the other hand, by Corollary 4.0.20,

$$\exists\ a_0,\ n_0 > 0 : \quad \int_{\Omega \setminus A_\delta} |f \circ \xi_n - f \circ \xi_{n,a}| \,\mathrm{d}\mathbb{P} \le 2\varepsilon \|f\|_\infty \quad \forall\ a \ge a_0,\ n \ge n_0.$$

This gives us the following estimation for the first summand in (4.0.32):

$$\exists\ a_0,\ n_0 > 0 : \left|\int f \mathrm{d}\mu_n - \int f \mathrm{d}\mu_{n,a}\right| \le \varepsilon(1 + 2\|f\|_\infty) \quad \forall\ a \ge a_0,\ n \ge n_0. \tag{4.0.33}$$

Since ν_a has a compact support, we conclude from 4.0.8 that $\mu_{n,a}$ converges weakly to τ_{ν_a} $(a > 0)$. Hence,

$$\forall\ a > 0\ \exists\ n_0 > 0 : \left|\int f \mathrm{d}\mu_{n,a} - \int f \mathrm{d}\tau_{\nu_a}\right| \le \varepsilon \quad \forall\ n \ge n_0. \tag{4.0.34}$$

Finally, it is evident that

$$\exists\ a_0 > 0 : \left|\int f \mathrm{d}\tau_{\nu_a} - \int f \mathrm{d}\tau_\nu\right| \le \varepsilon \quad \forall\ a \ge a_0. \tag{4.0.35}$$

Taking (4.0.33), (4.0.34) and (4.0.35) into account, we obtain

$$\exists\ n_0 > 0 : \left|\int f \mathrm{d}\mu_n - \int f \mathrm{d}\tau_\nu\right| \le \varepsilon(3 + 2\|f\|_\infty) \quad \forall\ n \ge n_0,$$

which completes the proof of CLT I in Theorem 4.0.8. The same proof works for CLT II. □

For the rest of this chapter, we devote ourselves to the laws of large numbers for the functionals $\|S_n^{p_n}\|^2$. Our first result in this direction is the following weak law of large numbers:

Theorem 4.0.21. *Let* $\nu \in \mathcal{M}^1([0,\infty[)$ *with finite fourth moment* $m_4(\nu) < \infty$ *and* $(p_n)_n$ *be a sequence of dimensions with* $\lim_{n\to\infty} p_n = \infty$*. Then*

$$\frac{1}{n}\,\|S_n^{p_n}\|_2^2 \longrightarrow m_2(\nu)$$

in probability as $n \to \infty$*.*

Proof: By Cauchy-Schwarz inequality, we have for any $\varepsilon > 0$,

$$\mathbb{P}\left(|\Xi_n(\nu)| \geq \varepsilon \cdot n\right) \leq \frac{1}{\varepsilon^2 \cdot n^2}\mathbb{E}\Big(\Xi_n(\nu)^2\Big).$$

From (4.0.12) and (4.0.13) we get

$$\begin{aligned}\mathbb{E}\Big(\Xi_n(\nu)^2\Big) &= \mathbb{E}\Big(\mathfrak{A}_n(\nu)^2 + 2\mathfrak{A}_n(\nu)\mathfrak{B}_n(\nu) + \mathfrak{B}_n(\nu)^2\Big)\\ &= n(m_4(\nu) - m_2(\nu)^2) + 0 + 2n(n-1)p_n^{-1}m_2(\nu)^2,\end{aligned}$$

and so the theorem follows. □

We next turn to an associated strong law of large numbers.

Theorem 4.0.22. *Let* $\nu \in \mathcal{M}^1([0,\infty[)$ *with finite eighth moment* $m_8(\nu) < \infty$ *and* $(p_n)_{n\geq 1}$ *admissible sequence of dimensions. Moreover, let* $(a_n)_{n\geq 1}$ *be a sequence in* $]0,\infty[$ *with* $a_n = n$ *if* $n/p_n \to c \in [0,\infty[$ *and* $a_n = n^2/p_n$ *if* $n/p_n \to \infty$*. Then*

$$\lim_{n\to\infty} \Xi_n(\nu)/a_n = 0 \quad \mathbb{P} - a.s.$$

Proof: Let $\varepsilon > 0$. By Markov inequality,

$$\begin{aligned}\mathbb{P}\left(|\Xi_n(\nu)| \geq \varepsilon \cdot a_n\right) &\leq \mathbb{P}\Big(|\mathfrak{A}_n(\nu)| \geq \frac{\varepsilon \cdot a_n}{2}\Big) + \mathbb{P}\Big(|\mathfrak{B}_n(\nu)| \geq \frac{\varepsilon \cdot a_n}{2}\Big)\\ &\leq \Big(\frac{2}{\varepsilon a_n}\Big)^4\Big(\int_\Omega \mathfrak{A}_n(\nu)^4 \mathrm{d}\mathbb{P} + \int_\Omega \mathfrak{B}_n(\nu)^4 \mathrm{d}\mathbb{P}\Big).\end{aligned} \tag{4.0.36}$$

Let us consider the first integral in (4.0.36). The integrand is

$$\mathfrak{A}_n(\nu)^4 = \sum_{\alpha,\beta,\gamma,\delta} A_\alpha A_\beta A_\gamma A_\delta \tag{4.0.37}$$

where the four indices range independently from 1 to n. Depending on how the indices match up, each term in this sum reduces to one of the following five forms, where in each case the indices are now distinct

$$A_i^4; \qquad A_i^2 A_j^2; \qquad A_i^3 A_j; \qquad A_i^2 A_j A_k; \qquad A_i A_j A_k A_l.$$

Since $\int A_i \mathrm{d}\mathbb{P} = 0$, it follows by the independence of $A_1, A_2, \ldots$ that the expected value of the summand vanishes if there is one index different from the three others. Therefore, we have only to consider the terms of the form A_i^4 and $A_i^2 A_j^2$. The number of occurrences in the sum (4.0.37) of the first form is n. The number of occurrences of the second form is $3n(n-1)$.

By Hölder's inequality we get $m_i(\nu) \leq m_8(\nu)^{i/8}$, $(i \leq 8)$. Hence, from inequality $A_i \leq \|\mathbb{X}_i\|^2 - m_2(\nu)$,

$$\mathbb{E}(A_i^4) \leq \sum_{i=0}^{4} \binom{4}{i} m_{2i}(\nu) \cdot m_2(\nu)^{4-i} \leq 2^4 m_8(\nu).$$

Thus, by Cauchy-Schwarz inequality we obtain

$$\mathbb{E}(A_i^2 A_j^2) \leq \mathbb{E}(A_1^4) \leq 2^4 m_8(\nu).$$

A term by term integration of (4.0.37) therefore gives

$$\int \mathfrak{A}_n(\nu)^4 \mathrm{d}\mathbb{P} \leq 2^4 m_8(\nu)(n + 3n(n-1)) \leq Cn^2 \tag{4.0.38}$$

with a suitable constant C>0.

We now turn to the second integral in (4.0.36). In analogy with (4.0.37),

$$\int \mathfrak{B}_n(\nu)^4 \mathrm{d}\mathbb{P} = \sum_{\alpha,\beta,\gamma,\delta} \int B_\alpha B_\beta B_\gamma B_\delta \mathrm{d}\mathbb{P}$$

where the four indices ranging independently from 1 to p_n. By Theorem 4.0.13, the summand $\int B_\alpha B_\beta B_\gamma B_\delta \mathrm{d}\mathbb{P}$ vanishes if there is one index different from the three others. This leaves terms of the form $\int B_i^4 \mathrm{d}\mathbb{P}$, of which there are p_n, and terms of the form $\int B_i^2 B_j^2 \mathrm{d}\mathbb{P}$ for $i \neq j$, of which there are $3p_n(p_n - 1)$. Hence, by symmetry,

$$\int \mathfrak{B}_n(\nu)^4 \mathrm{d}\mathbb{P} = p_n \int B_1^4 \mathrm{d}\mathbb{P} + 3p_n(p_n - 1) \int B_1^2 B_2^2 \mathrm{d}\mathbb{P}. \tag{4.0.39}$$

We proceed now with calculation of the integrals in (4.0.39). By the definition of B_i in (4.0.5),

$$\int B_1^4 \mathrm{d}\mathbb{P} = \sum_{\alpha_1 \neq \beta_1} \sum_{\alpha_2 \neq \beta_2} \sum_{\alpha_3 \neq \beta_3} \sum_{\alpha_4 \neq \beta_4} \int \prod_{i=1}^{4} X_{\alpha_i}^{(1)} X_{\beta_i}^{(1)} \mathrm{d}\mathbb{P}, \tag{4.0.40}$$

where the indices α_i, β_i with $\alpha_i \neq \beta_i$, in each sum $i = 1, 2, 3, 4$, range from 1 to n. Since $\mathbb{E}((X_\alpha^{(1)})^l) = 0$ for all α and for all odd $l \in \mathbb{N}$, it is clear that the expected value of the product $X_{\alpha_1}^{(1)} X_{\beta_1}^{(1)} X_{\alpha_2}^{(1)} X_{\beta_2}^{(1)} X_{\alpha_3}^{(1)} X_{\beta_3}^{(1)} X_{\alpha_4}^{(1)} X_{\beta_4}^{(1)}$ vanishes if there is one index $\gamma \in \{\alpha_1, \beta_1, \dots, \alpha_4, \beta_4\}$, such that the number

$$|\{v \in \{\alpha_1, \beta_1, \dots, \alpha_4, \beta_4\} : v = \gamma\}| \in \{1, \dots, 4\}$$

is odd (i.e. $X_\gamma^{(1)}$ appears in the product with odd multiplicity). Therefore, the sum in (4.0.40) reduces to one of the following three forms, where in each case the indices are now distinct

$$\int (X_i^{(1)} X_j^{(1)})^4 d\mathbb{P}; \qquad \int (X_i^{(1)})^4 (X_j^{(1)} X_k^{(1)})^2 d\mathbb{P}; \qquad \int (X_i^{(1)} X_j^{(1)} X_k^{(1)} X_l^{(1)})^2 \mathrm{d}\mathbb{P}.$$

The number of occurrences in the sum (4.0.40) of the first, the second and the third form is bounded by $C \cdot n^2$, $C \cdot n^3$ and $C \cdot n^4$, respectively where $C > 0$ is a suitable constant.

By using the Theorem 4.0.13, we obtain that there exists a universal constant $C > 0$ such that each term in the sum (4.0.40) is bounded from above by $C \cdot p_n^{-4}$. Consequently,

$$0 \leq \int B_1^4 \mathrm{d}\mathbb{P} \leq C_1 \frac{n^4}{p_n^4} \tag{4.0.41}$$

for some constant $C_1 > 0$. In the same manner we can see that

$$0 \leq \int B_1^2 B_2^2 \mathrm{d}\mathbb{P} \leq C_2 \frac{n^4}{p_n^4} \tag{4.0.42}$$

for some constant $C_2 > 0$. Taking (4.0.41), (4.0.42) and (4.0.39) into account, we obtain that

$$0 \leq \int \mathfrak{B}_n(\nu)^4 \mathrm{d}\mathbb{P} \leq C \frac{n^4}{p_n^2} \tag{4.0.43}$$

for some constant $C > 0$.

Combining (4.0.36) with (4.0.38) and with the preceding inequality we get

$$P\left(|\Xi_n(\nu)| \geq \varepsilon \cdot a_n\right) \leq C \varepsilon^{-4} a_n^{-4} \Big(n^2 + n^4/p_n^2\Big).$$

Now, we obtain that $P(|\Xi_n(\nu)| \geq \varepsilon \cdot a_n) \leq Cn^{-2}\varepsilon^{-4}$. And so by the Borel-Cantelli lemma, $P(|\Xi_n(\nu)| \geq \varepsilon \cdot a_n \text{ i.o.}) = 0$ for each positive ε. This completes the proof. $\square$

Chapter 5

Radial random walks on $p \times q$ matrices for $p \to \infty$

The main goal of this chapter is to derive a generalisation of the radial limit theorems presented in Chapter 4. The extension concerns a matrix-valued version. We consider the following geometric situation: For $p,\ q \in \mathbb{N}$ we will denote by $\mathbb{M}_{p,q}$ the space of $p \times q$-matrices over the field of real numbers $\mathbb{R}$. Furthermore, let $\mathbb{S}_q$ be the space of symmetric $q \times q$-matrices. We will denote by Π_q the cone of positive semidefinite $q \times q$ matrices in $\mathbb{S}_q$. We regard $\mathbb{M}_{p,q}$ as a real vector space of dimension pq, equipped with the Euclidean scalar product $\langle x, y\rangle := \mathrm{tr}(x'y)$ and norm $\|x\| = \sqrt{\mathrm{tr}(x'x)}$ where x' is transpose of x and tr is the trace in $\mathbb{M}_q := \mathbb{M}_{q,q}$. In the square case $p = q$, $\|\cdot\|$ is just the Frobenius norm. The orthogonal group $\mathbb{O}_p$ acts on $\mathbb{M}_{p,q}$ by left multiplication,

$$\mathbb{O}_p \times \mathbb{M}_{p,q} \to \mathbb{M}_{p,q}, \quad (O, x) \mapsto Ox. \tag{5.0.1}$$

By uniqueness of the polar decomposition, two matrices $x,\ y \in \mathbb{M}_{p,q}$ belong to the same $\mathbb{O}_p$-orbit if and only if $x'x = y'y$. Thus the space $\mathbb{M}_{p,q}^{\mathbb{O}_p}$ of $\mathbb{O}_p$-orbits in $\mathbb{M}_{p,q}$ is naturally parameterized by the cone Π_q via the map

$$O_p x \mapsto \sqrt{x'x} =: |x|\,, \quad \mathbb{M}_{p,q}^{\mathbb{O}_p} \to \Pi_q,$$

where for $r \in \Pi_q$, the matrix $\sqrt{r} \in \Pi_q$ denotes the unique positive semidefinite square root of r. According to this, the map

$$\varphi_p : \mathbb{M}_{p,q} \to \Pi_q, \quad x \mapsto \sqrt{x'x}$$

will be regarded as the canonical projection $\mathbb{M}_{p,q} \to \mathbb{M}_{p,q}^{\mathbb{O}_p}$.

In the case $q = 1$ we have $\mathbb{M}_{p,1} \cong \mathbb{R}^p$, $\mathbb{S}_1 = \mathbb{R}$, $\Pi_1 = [0, \infty[$ and φ_p is the usual norm mapping $\| \cdot \|_2 : \mathbb{R}^p \to [0, \infty[$.

Let us now fix a parameter $q \in \mathbb{N}$. By taking images of measures, φ_p induces a Banach space isomorphism between the space $\mathcal{M}_b^{\mathbb{O}_p}(\mathbb{M}_{p,q})$ of all bounded radial (i.e. $\mathbb{O}_p$ invariant) Borel measures on $\mathbb{M}_{p,q}$ and the space $\mathcal{M}_b(\Pi_q)$ of bounded Borel measures on the cone Π_q. In particular, for each measure $\nu \in \mathcal{M}^1(\Pi_q)$ and parameter p there is a unique radial probability measure $\nu_p := \nu_{p,q} \in \mathcal{M}^1(\mathbb{M}_{p,q})$ with $\varphi_p(\nu_p) = \nu$.

Let $\nu \in \mathcal{M}^1(\Pi_q)$ be a fixed probability measure and $q \in \mathbb{N}$. As in the case $q = 1$, we now consider for each "dimension" $p \in \mathbb{N}$ the associated radial measures ν_p on $\mathbb{M}_{p,q}$ and the radial random walks $(S_n^p := \sum_{k=1}^n X_k^p)_{n \geq 0}$, i.e. X_k^p, $k \in \mathbb{N}$ are independent ν_p-distributed random variables.

In this chapter, we are going to derive limit theorems for the Π_q-valued random variable (up to a normalization)

$$\varphi_p^2(S_n^p) = (S_n^p)' S_n^p : \Omega \longrightarrow \Pi_q$$

for $n, p \to \infty$ coupled in a suitable way. The proofs will rely on asymptotic results for moment functions of so called radial distributed random variables on $\mathbb{M}_{p,q}$ for $p \to \infty$ as well as on some identities for matrix variate normal distributions. Parts of these limit results were derived in [38] and [41] by using very different methods (cf. Remark 4.0.9).

The organization of this chapter is as follows: In the first two sections, some preliminaries for the proof of the main results 5.9.1 and 5.9.12 are presented. More precisely, in Section 5.1, after recalling some basic facts about relevant matrix algebra, we derive a generalisation of so-called permutation equivalence property for Kronecker products. In 5.2, we introduce the notation of permutation on a multiset and derive a multinomial theorem for non commutative operation. In Sections 5.3 and 5.4, the background on Bessel functions on the cone Π_q and on polynomials of matrix argument is provided. In the following two Sections, we recall the notion of random matrix and give a short overview over the concept of moments for random matrices. Then we generalise the method of moments to matrix-variate distributions. Sections 5.7 and 5.8 are devoted to the study of matrix variate normal distributions and moments of radial measures, respectively. In Section 5.9, our main result is formulated and proven.

5.1 Matrix algebra

The aim of this section is to provide relevant facts about matrix algebra. The material is mainly taken from [18]. We start with the notion of the Kronecker product.

Kronecker product: Let $\otimes$ denote the *Kronecker product* over the field of real numbers $\mathbb{R}$, i.e. $\otimes$ is an operation on two matrices of arbitrary size over $\mathbb{R}$ resulting in a block matrix. It is the matrix of the tensor product with respect to a standard choice of basis. With that, the Kronecker product of $A = [a_{ij}] \in \mathbb{M}_{m,n}$ and $B = [b_{ij}] \in \mathbb{M}_{p,q}$ is the block matrix

$$A \otimes B := [a_{ij}B] \in \mathbb{M}_{mp,nq}.$$

The Kronecker product is bilinear and associative but not commutative. However, $A \otimes B$ and $B \otimes A$ are *permutation equivalent*, meaning that there exist permutation matrices P and Q such that

$$A \otimes B = P \cdot (B \otimes A) \cdot Q. \tag{5.1.1}$$

If A and B are square matrices, then $A \otimes B$ and $B \otimes A$ are even *permutation similar*, meaning that we can take $P = Q'$. If A, B, C and D are matrices of such size that one can form the matrix products $A \cdot C$ and $B \cdot D$, then

$$(A \otimes B) \cdot (C \otimes D) = A \cdot C \otimes B \cdot D. \tag{5.1.2}$$

This is called the *mixed-product property*, because it mixes the ordinary matrix product and the Kronecker product. If two matrices P and Q are permutation, orthogonal or positive definite matrices then it is also the Kronecker product $P \otimes Q$.

The k-th Kronecker power $A^{\otimes k}$ is defined inductively for all positive integers k by $A^{\otimes 1} = A$ and $A^{\otimes k} = A \otimes A^{\otimes (k-1)}$ for $k = 2, 3, \ldots$. This definition implies that for $A \in \mathbb{M}_{m,n}$, we have $A^{\otimes k} \in \mathbb{M}_{m^k,n^k}$.

We now derive a generalisation of the permutation equivalence property, which will be required for the proof of the Theorem 5.9.1.

Lemma 5.1.1. *Let $A_i \in \mathbb{M}_{p_i,q_i}$ $(i = 1, \ldots, k)$, $p := p_1 \cdots p_k$ and $q := q_1 \cdots q_k$. Then, for each permutation $\sigma \in Sym(\{1, \ldots, k\})$ there exist permutation matrices*

$P \in \mathbb{M}_{p,p}$ and $Q \in \mathbb{M}_{q,q}$ such that

$$A_{\sigma(1)} \otimes \ldots \otimes A_{\sigma(k)} = P \cdot (A_1 \otimes \ldots \otimes A_k) \cdot Q.$$

Proof: Without loss of generality we can assume that $k = 4$, for the Kronecker product is associative. Since $(1) \otimes M = M = M \otimes (1)$ for any matrices M, it suffices to show that $A_1 \otimes A_3 \otimes A_2 \otimes A_4$ is permutation equivalent to $A_1 \otimes A_2 \otimes A_3 \otimes A_4$. For a matrix M let I_M and I^M denote the identity matrices of such size that one can form the matrix products $\mathrm{I}_M \cdot M$ and $M \cdot \mathrm{I}^M$. By the property (5.1.1) there exist permutation matrices P and Q with $A_3 \otimes A_2 = P(A_2 \otimes A_3)Q$. Therefore, using (5.1.2) we obtain by an easy computation

$$A_1 \otimes A_3 \otimes A_2 \otimes A_4 = (\mathrm{I}_{A_1} \otimes P \otimes \mathrm{I}_{A_4}) \cdot (A_1 \otimes A_2 \otimes A_3 \otimes A_4) \cdot \left(\mathrm{I}^{A_1} \otimes Q \otimes \mathrm{I}^{A_4}\right).$$

Clearly, both $\mathrm{I}_{A_1} \otimes P \otimes \mathrm{I}_{A_4}$ and $\mathrm{I}^{A_1} \otimes Q \otimes \mathrm{I}^{A_4}$ are permutation matrices. This completes the proof. □

Hadamard product Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two matrices of the same dimensions, say in $\mathbb{M}_{p,q}$. The *Hadamard product*, also known as the entrywise product of A and B is denoted by $A \circ B$ and is defined to be the matrix

$$A \circ B := [a_{ij} b_{ij}] \in \mathbb{M}_{p,q}.$$

The Hadamard product is commutative, associative and distributive over addition, and is a principal submatrix of the Kronecker product. Moreover, the identity matrix under the Hadamard multiplication of two matrices is a matrix (of the same dimension) where all elements are equal to 1.

For a matrix M , let us denote by $\mathbf{1}_M$ the 1-matrix of the same dimension as M, i.e $\mathbf{1}_M = (c_{ij})_{ij}$ with $c_{ij} = 1$ for all i, j. We will write it simply $\mathbf{1}$ when no confusion will arise. It is clear that

$$A \otimes B = (A \otimes \mathbf{1}) \circ (\mathbf{1} \otimes B), \tag{5.1.3}$$

$$B \otimes A = (\mathbf{1} \otimes A) \circ (B \otimes \mathbf{1}). \tag{5.1.4}$$

Let P and Q be permutation matrices of such size that one can form the matrix products $P \cdot A$ and $A \cdot Q$. It is easy to check that

$$P(A \circ B)Q = (PAQ) \circ (PBQ). \tag{5.1.5}$$

Vectorization of a matrix: The *vectorization of a matrix* is a linear transformation which converts the matrix into a column vector. In particular, the vectorization of a matrix $X \in \mathbb{M}_{p,q}$, denoted by $\mathrm{vec}(X)$, is the $p \cdot q \times 1$ vector defined as

$$\mathrm{vec}(X) = (x_1', \ldots, x_q')' \in \mathbb{R}^{pq},$$

where x_i, $i = 1, \ldots, q$ is the i-th column of X. The operation vec is compatible with the inner product $\langle , \rangle$ on $\mathbb{M}_p$ within the meaning

$$\langle A, B \rangle = \mathrm{vec}(A)'\mathrm{vec}(B) = \mathrm{tr}(\mathrm{vec}(A')\mathrm{vec}(B')') \tag{5.1.6}$$

for all A, $B \in \mathbb{M}_{p,q}$. We shall also use the vectorization together with the Kronecker product: for matrices $A \in \mathbb{M}_{p,q}$, $B \in \mathbb{M}_{q,n}$ and $C \in \mathbb{M}_{n,m}$ one has

$$\mathrm{vec}(ABC) = (C' \otimes A)\mathrm{vec}(B) = (\mathrm{I}_m \otimes AB)\mathrm{vec}(C) = (C'B' \otimes \mathrm{I}_p)\mathrm{vec}(A).$$

Furthermore, vectorization is an algebra homomorphism from the space $\mathbb{M}_{p,q}$ to $\mathbb{R}^{pq}$ both equipped with the Hadamard product $\circ$, i.e.

$$\mathrm{vec}(A \circ B) = \mathrm{vec}(A) \circ \mathrm{vec}(B), \quad A,\ B \in \mathbb{M}_{p,q}.$$

5.2 Permutations on a multiset

In this section, we generalise the multinomial formula (4.0.6) in terms of the Kronecker product instead of the usual multiplication. In order to do this, we first recall the notion of the permutation on a multiset from [32, Chapter 1].

Let $u \in \mathbb{N}$ and $k \in \mathbb{N}_0$. We denote by $C_0(k, u)$ the set of all u-compositions of k, i.e.

$$C_0(k,u) = \Big\{ \lambda \in \mathbb{N}_0^u : \quad |\lambda| := \sum_{i=1}^{u} \lambda_i = k \Big\},$$

and write $C(k, u)$ instead of $C_0(k, u) \cap \mathbb{N}^u$. Moreover, we set $M_u := \{1, 2, \ldots, u\}$. For a $\lambda \in C(k, u)$ a *finite multiset* $Mult(\lambda)$ on the ordered set M_u is a set, where i is contained with the multiplicity λ_i for all $i \in M_u$. One regards λ_i as the number of repetitions of i.

A *permutation* $\pi = (\pi_1\pi_2 \ldots \pi_k)$ on $Mult(\lambda)$ can be defined as a linear ordering of the elements of $Mult(\lambda)$, i.e. an element $i \in M$ appears exactly λ_i times in the permutation π. The set of all permutation on $Mult(\lambda)$ will be denoted by $\mathfrak{S}(\lambda)$. A permutation $\pi = (\pi_1\pi_2 \cdots \pi_k)$ on $Mult(\lambda)$ can be regarded as a way to place k-distinguishable balls in u distinguishable boxes such that the i-th box contains λ_i balls. Indeed, if i $(i = 1, \ldots, u)$ appears in position $j \in \{1, \ldots, k\}$ of the permutation π, then we put the "ball" π_j into the box i. For instance let $u = 3$, $\lambda := (1, 3, 2) \in C(k, u)$ be a 3-composition of $k = 6$ and $\pi = (2\ 1\ 2\ 3\ 3\ 2) =: (\pi_1\ \pi_2 \ldots \pi_6)$ be a permutation on $Mult(\lambda)$ then we put π_2 in the first box, π_1, π_3, π_6 in the second box and π_4, π_5 in the third box. It is clear that

$$|\mathfrak{S}(\lambda)| = \binom{k}{\lambda_1, \ldots, \lambda_u} = \frac{k!}{\lambda_1! \ldots \lambda_u!}$$

Let $m_i \in \mathbb{M}_{p_i,q_i}$ $(i = 1, \ldots, u)$, $\lambda \in C(k, u)$ and $\pi = (\pi_1, \ldots, \pi_k) \in \mathfrak{S}(\lambda)$. We will write $\pi(m_1, \ldots, m_u)$ instead of $m_{\pi_1} \otimes m_{\pi_2} \otimes \cdots \otimes m_{\pi_k}$. In the following theorem, which will be used in Section 5.9 several times, we expand a Kronecker power of a matrix sum in terms of powers of the terms in that sum.

Theorem 5.2.1. *Let $k \in \mathbb{N}$ and $x_1, \ldots, x_n \in \mathbb{M}_{p,q}$. Then*

$$\left(\sum_{i=1}^{n} x_i\right)^{\otimes,k} = \sum_{u=1}^{k} \sum_{\lambda \in C(k,u)} \sum_{\mu \in W(n,u)} \sum_{\pi \in \mathfrak{S}(\lambda)} \pi(x_{\mu_1}, \ldots, x_{\mu_u}), \tag{5.2.1}$$

where $W(n, u) = \{\mu = (\mu_1, \ldots, \mu_u) \in \{1, \ldots, n\}^u : \mu_1 < \mu_2 < \cdots < \mu_u\}$.

For $p = q = 1$ the Kronecker product coincides with the usual multiplication on $\mathbb{R}$ and therefore (5.2.1) generalises multinomial formula (4.0.6). For indices $u \in \{1, \ldots, k\}$, $\mu = (\mu_1, \ldots, \mu_n) \in W(n, u)$, $\lambda \in C(k, u)$ and $\pi \in \mathfrak{S}(\lambda)$ let us consider the associated summand

$$\pi(x_{\mu_1}, \ldots, x_{\mu_u}) = x_{\mu_{\pi_1}} \otimes \ldots \otimes x_{\mu_{\pi_k}} \tag{5.2.2}$$

from (5.2.1). It is clear that the different matrices $x_{\mu_1}, \ldots, x_{\mu_u}$, the numbers of their repetitions and their exact positions in the Kronecker product (5.2.2) are described by $\mu = (\mu_1, \ldots, \mu_u) \in W(n, u)$, $\lambda = (\lambda_1, \ldots, \lambda_u) \in C(k, u)$ and $\pi = (\pi_1, \ldots, \pi_k) \in \mathfrak{S}(\lambda)$, respectively. *Proof:* We proceed by induction on k. For $k = 1$ there is nothing to prove. Next, suppose as induction hypothesis that (5.2.1) holds with $k - 1$ instead of k. It gives

$$\begin{aligned} \left(\sum_{i=1}^{n} x_i\right)^{\otimes,k} &= \sum_{u=1}^{k-1} \sum_{\lambda \in C(k-1,u)} \sum_{\mu \in W(n,u)} \sum_{\pi \in \mathfrak{S}(\lambda)} \pi(x_{\mu_1}, \ldots, x_{\mu_u}) \otimes \sum_{j=1}^{n} x_j \\ &= \sum_{j=1}^{n} \sum_{u=1}^{k-1} \sum_{\lambda \in C(k-1,u)} \sum_{\mu \in W(n,u)} \sum_{\pi \in \mathfrak{S}(\lambda)} \pi(x_{\mu_1}, \ldots, x_{\mu_u}) \otimes x_j. \end{aligned} \tag{5.2.3}$$

Consider a summand $\pi(x_{\mu_1}, \ldots, x_{\mu_u}) \otimes x_j$ of the sum above, i.e. $j \in \{1, \ldots, n\}$, $u \in \{1, \ldots, k-1\}$, $\lambda \in C(k-1, u)$, $\mu \in W(n, u)$ and $\pi \in \mathfrak{S}(\lambda)$. If there is $\beta \in \{1, \ldots, u\}$ with $j = \mu_\beta$ then $\pi(x_{\mu_1}, \ldots, x_{\mu_u}) \otimes x_j$ corresponds to a summand in (5.2.1) associated with indices $\tilde{u} = u$, $\tilde{\lambda} = (\lambda_1, \ldots, \lambda_{\beta-1}, \lambda_\beta + 1, \lambda_{\beta+1}, \ldots, \lambda_u)$, $\tilde{\mu} = \mu$ and $\tilde{\pi} = (\pi_1, \ldots, \pi_{k-1}, \beta)$. In the other case, i.e. if $j \in (\mu_{\beta-1}, \mu_\beta)$ for a $\beta \in \{1, \ldots, u+1\}$ with the convention $\mu_0 := 0$ and $\mu_{u+1} = \infty$ the term $\pi(x_{\mu_1}, \ldots, x_{\mu_u}) \otimes x_j$ corresponds to a summand in (5.2.1) associated with indices $\tilde{u} = u+1$, $\tilde{\lambda} = (\lambda_1, \ldots, \lambda_{\beta-1}, 1, \lambda_\beta, \ldots, \lambda_u)$, $\tilde{\mu} = (\mu_1, \ldots, \mu_{\beta-1}, j, \mu_\beta, \ldots, \mu_u)$ and $\tilde{\pi} = (\pi_1, \ldots, \pi_{k-1}, \beta)$. As the number of summands in both (5.2.1) and (5.2.3) is equal to n^k, the induction step follows. $\square$

5.3 Bessel functions of matrix argument

In the following, we introduce $\mathbb{S}_q$-valued Bessel functions which generalise the usual Bessel functions in one variable. The material is mainly taken from [29]. For a general background on matrix Bessel functions, the reader is referred to the fundamental article [16] and the monograph [11].

Definition 5.3.1 (Spherical polynomials on $\mathbb{S}_q$)**.**

(a) A *partition* $\lambda = (\lambda_1, \ldots, \lambda_q)$ is an q-tuple ($q \in \mathbb{N}$) of non-negative integers such that $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_q$. We write $\lambda \geq 0$ for short.

(b) For a partition λ the *power function* on $\mathbb{S}_q$ is defined by

$$\Delta_\lambda(x) := \Delta_1(x)^{\lambda_1-\lambda_2}\Delta_2(x)^{\lambda_2-\lambda_3} \cdot \ldots \cdot \Delta_{q-1}(x)^{\lambda_{q-1}-\lambda_q}\Delta_q(x)^{\lambda_q}, \qquad (5.3.1)$$

where $\Delta_i(x)$ are the principal minors of the determinant $\Delta(x) = \det(x)$.

(c) For a partition λ the *spherical polynomial* on $\mathbb{S}_q$ is defined by

$$\Phi_\lambda(x) = \int_{\mathbb{O}_q} \Delta_\lambda(uxu^{-1})\mathrm{d}u, \qquad (5.3.2)$$

where $\mathrm{d}u$ is the normalized Haar measure on $\mathbb{O}_q$.

Remark 5.3.2. From the definition (5.3.1), we immediately see that power functions Δ_λ are homogeneous of degree $|\lambda| = \lambda_1 + \ldots + \lambda_q$, (i.e. $\Delta_\lambda(tx) = t^{|\lambda|}\Delta_\lambda(x)$ for all $t \in \mathbb{R}$ and $x \in \mathbb{S}_q$) and hence so are Φ_λ. Moreover, (5.3.2) makes it obvious that Φ_λ are invariant under conjugation by $\mathbb{O}_q$. For the identity matrix I_q, we have $\Phi_\lambda(\mathrm{I}_q) = 1$.

There is a renormalization $Z_\lambda = c_{\lambda,q}\Phi_\lambda$ with constants $c_{\lambda,q}$ depending only on λ and q such that

$$\mathrm{tr}(x)^k = \sum_{|\lambda|=k} Z_\lambda(x) \quad (x \in \mathbb{S}_q,\ k \in \mathbb{N}_0);$$

see Section XI.5. of [11]. The functions Z_λ are called *zonal polynomials*. By construction, the Z_λ are homogeneous polynomials, which are invariant under conjugation by $\mathbb{O}_q$ and thus depend only on the eigenvalues of their argument. More precisely, for $x \in \mathbb{S}_q$ with eigenvalues $\xi = (\xi_1, \dots, \xi_q) \in \mathbb{R}^q$, one has

$$Z_\lambda(x) = C_\lambda^\alpha(\xi) \quad \text{with} \quad \alpha = 2$$

where the C_λ^α are the *Jack polynomials* of index α in a suitable normalization (c.f. [11],[29]). The Jack polynomials C_λ^α are homogeneous of degree $|\lambda|$ and symmetric in their arguments.

Definition 5.3.3. [Bessel functions on $\mathbb{S}_q$]

(a) For a partition λ and $\alpha > 0$ the *generalised Pochhammer symbol* is defined by

$$(\mu)_\lambda^\alpha = \prod_{j=1}^{q} \left(\mu - \frac{1}{\alpha}(j-1)\right)_{\lambda_j} \qquad (\mu \in \mathbb{C}),$$

where $(\cdot)_j$ denotes the usual Pochhammer symbol.

(b) For an index $\mu \in \mathbb{C}$ satisfying $(\mu)_\lambda^\alpha \neq 0$ for all $\lambda \geq 0$ the *matrix Bessel function* associated with the cone Π_q of index μ is defined as ${}_0F_1$-hypergeometric series in terms of the Z_λ by

$$\mathcal{J}_\mu(x) = \sum_{\lambda \geq 0} \frac{(-1)^{|\lambda|}}{(\mu)_\lambda^{1/2} |\lambda|!} Z_\lambda(x). \tag{5.3.3}$$

Remark 5.3.4. If $q = 1$, then $\Pi_q = [0, \infty[$ and we have $\mathcal{J}_\mu(x^2/4) = \Lambda_{\mu-1}(x)$, where $\Lambda_\kappa = {}_0F_1(\kappa+1; -z^2/4)$ is the usual modified Bessel function in one variable (cf. Theorem 4.0.13).

5.4 Polynomials of matrix argument

Let $p,\ q \in \mathbb{N}$. For $\kappa = (\kappa_{ij})_{i,j} \in \mathbb{N}_0^{p\times q}$ (a composition) we set $|\kappa| := \sum_{i,j} \kappa_{ij}$ and $R_i(\kappa) := \sum_{j=1}^q \kappa_{ij}$, $i = 1, \dots, p$. Moreover, we write $z^\kappa := \prod_{i,j} z_{ij}^{\kappa_{ij}}$; z^κ is a monomial of degree $|\kappa|$. The spaces of *polynomials* and *even-row polynomials* are defined by

$$\mathcal{P} := span\Big\{x^{\kappa} : \ \kappa \in \mathbb{N}_0^{p\times q}\Big\},$$
$$\mathcal{P}_e := span\Big\{x^{\kappa} : \ \kappa \in \mathbb{N}_0^{p\times q}, \forall\ i\ R(i) \text{ is even } \Big\},$$

respectively.

For the proof of our main result 5.9.1 we need the following elementary lemma.

Lemma 5.4.1. *Let $r \in \Pi_q$, and $\kappa \in \mathbb{N}_0^{p\times q}$. Then*

$$\Psi_{r,\kappa} : \mathbb{M}_{p,q} \to \mathbb{R}, \quad \Psi_{r,\kappa}(z) := ((zr)'(zr))^{\kappa}$$

is an even polynomial of degree $2|\kappa|$.

Proof: Since the product of two even-row polynomials is also a even-row polynomial, the proof follows easily by induction on $n = |\kappa|$. □

5.5 Random Matrices

In this section, we introduce the notion of random matrices and the concept of moments in this surrounding. Part of this section corresponds to standard treatment of the univariate case and its generalisation to a multivariate one. A good references for matrix variate distributions is the book [14], for multivariate statistical analysis see also [1].

We begin with the notion of a random matrix.

Definition 5.5.1. A $p\times q$ matrix X consisting of $q\cdot p$ real valued random variables $x_{11}, x_{12}, \ldots, x_{pq}$ on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$, is called a (real) $p \times q$ random matrix.

Let $X = (x_{ij})$ be $p \times q$ random matrix and $h = (h_{ij}) : \mathbb{M}_{p,q} \to \mathbb{M}_{r,s}$, i.e. $h_{ij} : \mathbb{M}_{p,q} \to \mathbb{R}$, $i = 1, \ldots, r$, $j = 1, \ldots, s$. Then the expected value of the function $h(X)$ is a $r \times s$ matrix defined by

$$\mathbb{E}(h(X)) = \Big(\mathbb{E}(h_{ij}(X))\Big)_{1\leq i\leq r,\ 1\leq j\leq s}$$

when $\mathbb{E}(h_{ij}(X))$ exists. From above it is an easy consequence that

(i) $\mathbb{E}(A) = A$, A constant matrix,

(ii) for $A \in \mathbb{M}_{p,r}$ and $B \in \mathbb{M}_{s,q}$

$$\mathbb{E}(Ah(X)B) = A\mathbb{E}(h(X))B,$$

(iii) for $A,\ B \in \mathbb{M}_{r,s}$

$$\mathbb{E}(A \circ h(X) \circ B) = A \circ \mathbb{E}(h(X)) \circ B,$$

(iv) for constant matrices A and B

$$\mathbb{E}(A \otimes h(X) \otimes B) = A \otimes \mathbb{E}(h(X)) \otimes B.$$

(v) for $g = (g_{ij})$ of the same order as h, i.e. $g : \mathbb{M}_{p,q} \to \mathbb{M}_{r,s}$,

$$\mathbb{E}\Big(h(X) + g(X)\Big) = \mathbb{E}\Big(h(X)\Big) + \mathbb{E}\Big(g(X)\Big).$$

Thus, for the $p \times q$ random matrix X, the mean matrix is given by

$$\mathbb{E}(X) := \Big(\mathbb{E}(x_{ij})\Big)_{1 \leq i,j \leq q} \in \mathbb{M}_{p,q}.$$

The $pq \times rs$ covariance matrix of the random matrices $X = (x_{ij})$ in $\mathbb{M}_{p,q}$ and $Y = (y_{ij})$ in $\mathbb{M}_{r,s}$ is defined by

$$\mathbb{C}ov(X, Y) := \mathbb{C}ov(\text{vec}(X'), \text{vec}(Y')) \in \mathbb{M}_{pq,rs}.$$

Let $\mathbf{x}_i'$ and $\mathbf{y}_j'$ denote the i-th and j-th rows of the matrices X and Y, respectively, $i = 1, \ldots, p$ and $j = 1, \ldots, r$. By standard calculation we obtain

$$\mathbb{C}ov(X, Y) = \begin{pmatrix} \mathbb{C}ov(\mathbf{x}_1, \mathbf{y}_1) & \ldots & \mathbb{C}ov(\mathbf{x}_1, \mathbf{y}_r) \\ \vdots & & \vdots \\ \mathbb{C}ov(\mathbf{x}_p, \mathbf{y}_1) & \ldots & \mathbb{C}ov(\mathbf{x}_p, \mathbf{y}_r) \end{pmatrix}$$

with $\mathbb{C}ov(\mathbf{x}_i, \mathbf{y}_j) \in \mathbb{M}_{q,s}$. Therefore, we consider $\mathbb{C}ov(X, Y)$ as a block matrix with p row and r column partitions. Moreover, we identify the entries of $\mathbb{C}ov(X, Y)$ by

$$\mathbb{C}ov(X, Y)_{(\alpha,\gamma),(\beta,\delta)} := \Big(\mathbb{C}ov(\mathbf{x}_\alpha, \mathbf{y}_\beta)\Big)_{\gamma,\delta} = \mathbb{C}ov(x_{\alpha\gamma}, y_{\beta\delta}),$$

where $1 \leq \alpha \leq p$, $1 \leq \beta \leq r$, $1 \leq \gamma \leq q$ and $1 \leq \delta \leq s$. As a special case, we get the covariance matrix of X as

$$\mathbb{C}ov(X) = \mathbb{C}ov(X, X) = \begin{pmatrix} \mathbb{C}ov(\mathbf{x}_1, \mathbf{x}_1) & \dots & \mathbb{C}ov(\mathbf{x}_1, \mathbf{x}_p) \\ \vdots & \ddots & \vdots \\ \mathbb{C}ov(\mathbf{x}_p, \mathbf{x}_1) & \dots & \mathbb{C}ov(\mathbf{x}_p, \mathbf{x}_p) \end{pmatrix} \in \mathbb{M}_{pq,pq}. \quad (5.5.1)$$

Since $\mathbb{C}ov(X)$ is symmetric, we have $\mathbb{C}ov(X)_{(\alpha,\gamma),(\beta,\delta)} = \mathbb{C}ov(X)_{(\beta,\delta),(\alpha,\gamma)}$ for all admissible indices α, β, γ and δ. Moreover, it is clear that $\mathbb{C}ov(X)$ is positive-semidefinite. The familiar calculating rules for covariance matrices in multidimensional case can easily be transferred to matrix variate one. For $p \times q$ random matrices X_1, X_2 and matrices $A \in \mathbb{M}_{t,p}$, $B \in \mathbb{M}_{w,r}$, $C \in \mathbb{M}_{t,q}$ we have

$$\begin{aligned} \mathbb{C}ov(X, Y) &= \mathbb{C}ov(Y, X)' \\ \mathbb{C}ov(AX + C, BY) &= A \otimes \mathrm{I}_q \cdot \mathbb{C}ov(X, Y) \cdot B' \otimes \mathrm{I}_s \\ \mathbb{C}ov(X_1 + X_2, Y) &= \mathbb{C}ov(X_1, Y) + \mathbb{C}ov(X_2, Y). \end{aligned}$$

By using Equation (5.1.6) we easily calculate the following useful relation between covariance matrices and inner product on $\mathbb{M}_{p,q}$:

$$\mathrm{tr}(\mathbb{C}ov(X_1, X_2)) = \mathbb{E} \langle X_1, X_2 \rangle - \langle \mathbb{E}(X_1), \mathbb{E}(X_2) \rangle . \quad (5.5.2)$$

Now, we introduce the notion of the higher order moments for matrix valued random variables.

Definition 5.5.2. Let $X = (x_{ij})$ be $\mathbb{M}_{p,q}$ valued random matrix with distribution $\mu := \mathbb{P}_X \in \mathcal{M}^1(\mathbb{M}_{p,q})$. We say that μ (or X) admits a k-th moment ($k \in \mathbb{N}$) if

$$|\mu|_k := \int_{\mathbb{M}_{p,q}} \|x\|^k \, \mathrm{d}\mu(x) < \infty. \quad (5.5.3)$$

Under the condition (5.5.3), we define the *k-th moment* of μ (and also of X) by

$$\mathrm{M}_k(\mu) := \int_{\mathbb{M}_{p,q}} x^{\otimes,k} \mathrm{d}\mu(x) \in \mathbb{M}_{p^k,q^k}.$$

Moreover, for a $\kappa \in \mathbb{N}_0^{p \times q}$ with $|\kappa| = k$ we set

$$m_\kappa(\mu) := \int_{\mathbb{M}_{p,q}} z^\kappa \mathrm{d}\mu(z) \in \mathbb{R},$$

and also call $m_\kappa(\mu)$ the *κ-th moment* of μ (or of X). If necessary, we may use the notation $\mathrm{M}_k(X)$ and $m_\kappa(X)$ without risk of confusion.

Obviously, $m_\kappa(\mu)$ represents an entry of the matrix $\mathrm{M}_k(\mu)$. Let us now consider a tuple $I = ((i_1, j_1), \ldots, (i_k, j_k))$ with $i_\alpha \in \{1, \ldots, p\}$ and $j_\alpha \in \{1, \ldots, q\}$ for $\alpha \in \{1, \ldots, k\}$. Then the I-th component $\mathrm{M}_k(X)_I$ of $\mathrm{M}_k(X)$ is given by

$$\mathrm{M}_k(X)_I = \mathbb{E}\Big(x_{i_1 j_1} \cdots x_{i_k j_k}\Big) \in \mathbb{R}.$$

In the following, $\widehat{\mu}$ denotes the characteristic function of a probability measure μ on $\mathbb{M}_{p,q}$, i.e.

$$\widehat{\mu}(x) = \int_{\mathbb{M}_{p,q}} \exp(i \langle x, y \rangle) \mathrm{d}\mu(y).$$

Let $k \in \mathbb{N}_0$ and $\kappa \in \mathbb{N}_0^{p \times q}$ with $|\kappa| = k$. If μ admits a k-th moment then we have a well known relation between the characteristic function of μ and moment $m_\kappa(\mu)$, namely

$$m_\kappa(\mu) = (-i)^{|\kappa|} \mathrm{D}_\kappa \widehat{\mu}(x)|_{x=0}, \tag{5.5.4}$$

where D_κ is the differential operator $\frac{\partial^{\kappa_{11}}}{\partial x_{11}^{\kappa_{11}}} \cdot \frac{\partial^{\kappa_{12}}}{\partial x_{12}^{\kappa_{12}}} \cdots \frac{\partial^{\kappa_{pq}}}{\partial x_{pq}^{\kappa_{pq}}}$.

5.6 Method of moments for random matrices

For some distributions, the characteristic function is unmanageable but moments can nonetheless be calculated. In these cases, it is sometimes possible to prove weak convergence of distributions by establishing that the moments converge. For example, in the univariate case it is well known that a sequence of probability measures $(\mu_n)_n$ on $\mathbb{R}$ converges weakly to a measure $\mu \in \mathcal{M}^1(\mathbb{R})$ if the k-th moment of μ_n converges to those of μ, provided all moments of μ_n and μ exist, and μ is determined by its moments. In the literature, this approach is called the method of moments; see [4].

The goal of this section is to generalise the method of moments to the matrix variate case. For the convenience of the reader, we first repeat some basic definitions from the classical probability theory in terms of our matrix variate context.

Definition 5.6.1. Let $\mathcal{S}$ be a subset of positive bounded Borel measures on $p \times q$ matrices over the field $\mathbb{R}$.

(a) A real valued function $g \in \mathcal{C}(\mathbb{M}_{p,q})$ is said to be *uniformly integrable* with respect to (w.r.t.) $\mathcal{S}$ if

$$\lim_{r\to\infty} \sup_{\mu\in\mathcal{S}} \int_{\{\|x\|\geq r\}} |g(x)| \,\mathrm{d}\mu(x) = 0.$$

(a) $\mathcal{S}$ is said to be *tight* if the function $g \equiv 1$ is uniformly integrable.

The following two lemmas, which are a slight modification of [8], Section 8.1, Corollary 6 and Corollary 7, respectively, will be needed for the proof of the generalised method of moments 5.6.4.

Lemma 5.6.2. *Let μ, $\mu_n \in \mathcal{M}_b^+(\mathbb{M}_{p,q})$ $(n \in \mathbb{N})$ such that $\mu_n \to \mu$ vague. Let $g \in \mathcal{C}(\mathbb{M}_{p,q})$ be uniformly integrable w.r.t. $\{\mu_n :\ n \in \mathbb{N}\}$. Then $g \in L^1(\mathbb{M}_{p,q}, \mu_k) \cap L^1(\mathbb{M}_{p,q}, \mu)$ for all $n \in \mathbb{N}$ and*

$$\lim_{n\to\infty} \int_{\mathbb{M}_{p,q}} g\mathrm{d}\mu_n = \int_{\mathbb{M}_{p,q}} g\mathrm{d}\mu.$$

Proof: Without loss of generality we can assume that g is positive-valued. Let $(f_n)_{n\in\mathbb{N}} \subset \mathcal{C}_c(\mathbb{M}_{p,q})$ with $0 \leq f_n \leq 1$ for all $n \in \mathbb{N}$ and $f_n \nearrow 1$. It is clear that for all $k \in \mathbb{N}$ the sequence $(f_k \cdot \mu_n)_n$ converges vaguely to $f_k \cdot \mu$. By dominated convergence we obtain

$$\begin{aligned} 0 \leq \int g\mathrm{d}\mu &= \lim_{k\to\infty} \int f_k g\mathrm{d}\mu = \lim_{k\to\infty}\lim_{n\to\infty} \int f_k g\mathrm{d}\mu_n \\ &\leq \lim_{n\to\infty} \int_{\{\|x\|<r\}} g\mathrm{d}\mu_n + \lim_{n\to\infty} \int_{\{\|x\|\geq r\}} g\mathrm{d}\mu_n < \infty. \end{aligned}$$

For an $\varepsilon > 0$ there exists a $u_\varepsilon \in \mathcal{C}_c(\mathbb{M}_{p,q})$ with $0 \leq u_\varepsilon \leq 1$ and $\int g(1-u_\varepsilon)\mathrm{d}\mu_n \leq \varepsilon$ for all $n \in \mathbb{N}$. Proceeding as in the inequality chain above with the function $g(1-u_\varepsilon)$ instead of g, we get $\int g(1-u_\varepsilon)\mathrm{d}\mu \leq \varepsilon$. Finally, by triangle inequality we deduce that

$$\begin{aligned} \left|\int g\mathrm{d}\mu_n - \int g\mathrm{d}\mu\right| &\leq \left|\int g(1-u_\varepsilon)\mathrm{d}\mu_n\right| \\ &\quad + \left|\int g(1-u_\varepsilon)\mathrm{d}\mu\right| + \left|\int gu_\varepsilon\mathrm{d}\mu_n - \int gu_\varepsilon\mathrm{d}\mu\right| \leq 3\varepsilon \end{aligned}$$

for n large enough. □

Lemma 5.6.3. *Let μ, $\mu_n \in \mathcal{M}_b^+(\mathbb{M}_{p,q})$ $(n \in \mathbb{N})$ such that $\mu_n \to \mu$ vague. Let $a > 0$ with $\sup_{n\in\mathbb{N}} \int \|x\|^a \,\mathrm{d}\mu_n(x) < \infty$. Then for all $\kappa \in \mathbb{N}_0^{p\times q}$ with $|\kappa| < a$,*

$$\lim_{n\to\infty} m_\kappa(\mu_n) = m_\kappa(\mu).$$

Proof: Let $\kappa \in \mathbb{N}_0^{p\times q}$ with $b := |\kappa| < a$. By assumption, for an $r > 0$ we obtain

$$\int_{\{\|x\|\geq r\}} \|x\|^b \,\mathrm{d}\mu_n(x) \leq \int_{\{\|x\|\geq r\}} \frac{\|x\|^a}{r^{a-b}}\mathrm{d}\mu_n(x) \leq \int_{\mathbb{M}_{p,q}} \frac{\|x\|^a}{r^{a-b}}\mathrm{d}\mu_n(x) \leq \frac{C}{r^{a-b}}$$

with a constant $C > 0$ independent from n. We thus get $g : x \to x^\kappa$ is uniformly integrable w.r.t. $\{\mu_n : n \in \mathbb{N}\}$. The lemma is now a consequence of Lemma 5.6.2. □

Now we turn to the generalisation of the (univariate) method of moments [4, Theorem 30.2].

Theorem 5.6.4 (Method of moments)**.** *Let $Y, Y_1, Y_2, \ldots$ be $\mathbb{M}_{p,q}$ valued random variables. Suppose that the distribution of Y is determined by its moments $\mathrm{M}_k(Y)$ ($k \in \mathbb{N}$), that the Y_n have moments $\mathrm{M}_k(Y_n)$ of all orders, and that*

$$\lim_{n\to\infty} \mathrm{M}_k(Y_n) = \mathrm{M}_k(Y) \tag{5.6.1}$$

for $k = 1, 2, \ldots$. Then the sequence $(Y_n)_n$ converges to Y in distribution.

Proof: Let μ_n and μ be the distributions of Y_n and Y. Since all moments of Y_n exist and converge, there exists a constant $C > 0$ such that $\sup_{n\in\mathbb{N}} \mathbb{E}(\|Y_n\|) < C$. Hence, by Markov's inequality follows that the sequence $\{\mu_n : n \in \mathbb{N}\}$ is tight.

Now let us assume that μ_n does not converge to μ. In particularly, there is a sub-sequence $N' \subset \mathbb{N}$ such that for every sub-sequences $N'' \subset N'$ we have $\mu_n \nrightarrow \mu$ along N''. On the other hand, every tight sequence of probability measures contains a subsequence that converges weakly to a probability measure. Thus, there exist a subsequence $N_0'' \subset N'$ and a Borel measure $\tilde{\mu}$ such that $\mu_n \to \tilde{\mu}$ along N_0''.

For all $\kappa \in \mathbb{N}_0^{p\times q}$ we have $\sup_{n\in\mathbb{N}} \int \|x\|^{2|\kappa|} \,\mathrm{d}\mu_n(x) < \infty$. Therefore, $m_\kappa(\mu_n) \to m_\kappa(\tilde{\mu})$ along N'' by Corollary 5.6.3. Hence, $\mathrm{M}_k(\mu_n) \to \mathrm{M}_k(\tilde{\mu})$ along N'' for all $k \in \mathbb{N}$. However, by the Assumption (5.6.1) we also have $\mathrm{M}_k(\mu_n) \to \mathrm{M}_k(\mu)$. Since μ is determined by its moments we have $\tilde{\mu} = \mu$, a contradiction. □

By the application of the method of moments one has a problem to decide whether a distribution is uniquely determined by its moments or not. This problem, known as the moment problem, is reasonably well developed and understood. A good general work here is [24], see also the references given there.

Let us now outline one usefull condition under which a distribution is uniquely determined by its moments.

Theorem 5.6.5 (Cramér condition). *Let X be $\mathbb{M}_{p,q}$-valued random matrix. If the moment generating function $G_X(t) = \mathbb{E}(e^{\langle t,X\rangle})$ is well-defined in a proper neighborhood of zero in $\mathbb{M}_{p,q}$, then the set of moments $\mathfrak{M} := \left\{m_\kappa(X) : \ \kappa \in \mathbb{N}_0^{p\times q}\right\}$ is well-defined. Moreover, the distribution of X is determined by its moments $\mathfrak{M}$.*

Remark 5.6.6. The distributions with compact support are determined by their moments.

5.7 Matrix variate normal distribution

In this section, we derive some results for moments of matrix variate normal distributions. As a preparation, a short account on the relevant notions and facts is included in the present section. For a general background on matrix variate distributions, the reader is referred to the monograph [14].

Let $X = (x_1, \ldots, x_n)'$ have the multivariate normal distribution, denoted by $X \sim \mathcal{N}_n(\mu, C)$, with mean vector $\mu \in \mathbb{R}^n$ and covariance matrix $C = (c_{ij}) \in \mathbb{M}_{n,n}$. The moment formulas $m_\kappa(X)$, $(\kappa \in \mathbb{N}_0^n)$ are well studied in the literature (see [35] and references cited there). The Isserli's theorem [35, Theorem 1] allows one to compute higher-order moments of the multivariate normal distribution in terms of its covariance matrix.

Theorem 5.7.1 (L. Isserlis). *Let $X = (x_1, \cdots, x_n)$ be a zero mean multivariate normal random vector with covariance matrix $C = (c_{ij}) \in \mathbb{M}_{n,n}$. The k-th order moment $\mathbb{E}(x_{i_1} \cdots x_{i_k})$, with $i_1, \ldots, i_k \in \{1, \ldots, n\}$ not necessarily distinct, of the variable X is given as follows:*

(a) If k is odd, $\mathbb{E}(x_{i_1} \cdots x_{i_k}) = 0$.

(b) If k is even with $k = 2 \cdot u$ $(u \geq 1)$, then it is

$$\mathbb{E}(x_{i_1} \cdots x_{i_{2u}}) = \sum \prod \mathbb{E}(x_{i_\alpha} x_{i_\beta}) = \sum \prod c_{i_\alpha, i_\beta},$$

where the notation $\sum \prod$ means summing over all distinct ways of partitioning $\{x_{i_1}, \ldots, x_{i_{2u}}\}$ into pairs. This yields $\frac{1}{u!}\binom{2u}{2,\ldots,2} = \frac{(2u-1)!}{2^{u-1}(u-1)!}$ terms in the sum.

For instance, we compute the 4-order moments $\mathbb{E}(x_{i_1} x_{i_2} x_{i_3} x_{i_4})$ of X, i.e. $k = 4$,

$u = 2$ and $\frac{1}{2!}\binom{4}{2,2} = 3$. One has the following five cases

$$
\begin{aligned}
\mathbb{E}(x_\alpha^4) &= 3c_{\alpha\alpha}^2 && (\alpha = i_1 = i_2 = i_3 = i_4), \\
\mathbb{E}(x_\alpha^3 x_\beta) &= 3c_{\alpha\alpha}c_{\alpha\beta} && (\alpha = i_1 = i_2 = i_3,\ \beta = i_4), \\
\mathbb{E}(x_\alpha^2 x_\beta^2) &= c_{\alpha\alpha}c_{\beta\beta} + 2c_{\alpha\beta}^2 && (\alpha = i_1 = i_2,\ \beta = i_3 = i_4), \\
\mathbb{E}(x_\alpha^2 x_\beta x_\gamma) &= c_{\alpha\alpha}c_{\beta\gamma} + 2c_{\alpha\beta}c_{\alpha\gamma} && (\alpha = i_1 = i_2,\ \beta = i_3,\ \gamma = i_4), \\
\mathbb{E}(x_\alpha x_\beta x_\gamma x_\delta) &= c_{\alpha\beta}c_{\gamma\delta} + c_{\alpha\gamma}c_{\beta\delta} + c_{\alpha\delta}c_{\beta\gamma} && (\alpha = i_1,\ \beta = i_2,\ \gamma = i_3,\ \delta = i_4).
\end{aligned}
$$

These equations clearly show the mechanism of the above theorem. No matter how high the order of central moments is required, it is remarkably easy to calculate them, as long as the covariances c_{ij} are given. In the most classical case $n = 1$, i.e. $\mathbb{P}_X$ is a centered Gaussian distribution on $\mathbb{R}$ with covariance $\sigma^2 > 0$, the Isserlis' identity in 5.7.1 reduces to the well known formula

$$
\mathbb{E}(X^k) = \begin{cases} 0, & \text{if } k \text{ is odd}, \\ \frac{(2u)!}{2^u u!}\sigma^{2u}, & \text{if } k = 2u. \end{cases}
$$

Now we turn to the generalisation of the multivariate case to a matrix variate one. We begin with the definition of matrix variate distribution.

Definition 5.7.2. The random matrix Z on $\mathbb{M}_{p,q}$ is said to have matrix variate normal distribution with mean matrix $M \in \mathbb{M}_{p,q}$ and covariance matrix $\Sigma \in \Pi_{p\cdot q}$ if $\mathrm{vec}(Z') \sim \mathcal{N}_{p\cdot q}(\mathrm{vec}(M'), \Sigma)$. We shall use the notation $Z \sim \mathcal{N}_{p,q}(M, \Sigma)$.

We now compute the moment generation function of the random matrix Z.

Lemma 5.7.3. *If Z is $\mathcal{N}_{p,q}(M, \Sigma)$-distributed random matrix, then the moment generation function $G_Z(t) := \mathbb{E}(e^{\langle t, Z\rangle})$, $t \in \mathbb{M}_{p,q}$ of Z is given by*

$$
G_Z(t) = \exp\Big(\langle t, M\rangle + \frac{1}{2}\mathrm{vec}(t')'\Sigma\mathrm{vec}(t')\Big), \quad t \in \mathbb{M}_{p,n}.
$$

Proof: We know that $\mathrm{vec}(Z') \sim \mathcal{N}_{p\cdot q}(\mathrm{vec}(M'), \Sigma)$. Hence, from the moment generation function of a multivariate normal distribution, we get

$$\begin{aligned}
G_Z(t) &= \mathbb{E}\Big(e^{\langle t,Z\rangle}\Big) = \mathbb{E}\Big(e^{\langle \mathrm{vec}(t'),\mathrm{vec}(Z')\rangle}\Big) \\
&= \exp\Big(\langle \mathrm{vec}(M'),\mathrm{vec}(t')\rangle + \frac{1}{2}\mathrm{vec}(t')'\Sigma\mathrm{vec}(t')\Big) \\
&= \exp\Big(\langle t, M\rangle + \frac{1}{2}\mathrm{vec}(t')'\Sigma\mathrm{vec}(t')\Big).
\end{aligned}$$

□

Remark 5.7.4. The moment generating function $G_Z : \mathbb{M}_{p,q} \to \mathbb{R}$, $t \mapsto G_Z(t)$ of $Z \sim \mathcal{N}_{p,q}(M,\Sigma)$ is well defined, and so the Cramér condition 5.6.5 implies that all moments $m_\kappa(Z)$, $\kappa \in \mathbb{N}_0^{p\times q}$ exist and the distribution $\mathcal{N}_{p,q}(M,\Sigma)$ is determined by its moments.

In the following we consider a $p \cdot q \times p \cdot q$ matrix as a block matrix with p row and q column partitions, i.e.

$$\Sigma = \begin{pmatrix} \Sigma^{(1,1)} & \dots & \Sigma^{(1,q)} \\ \vdots & \ddots & \vdots \\ \Sigma^{(q,1)} & \dots & \Sigma^{(q,q)} \end{pmatrix} \qquad \text{with } \Sigma^{(i,k)} \in \mathbb{M}_{p,p},\ i,k \in \{1,\dots,q\},$$

and identify the entries of Σ by

$$\Sigma_{(i,j),(k,l)} := \Big(\Sigma^{(i,k)}\Big)_{j,l}, \quad 1 \le i,k \le q,\ 1 \le j,l \le p.$$

If Σ is symmetric, then we clearly have $\Sigma_{(i,j),(k,l)} = \Sigma_{(k,l),(i,j)}$ for all admissible indices i,j,l and k.

In order to generalise the Theorem 5.7.1 we need some unusual notations: Let $u \in \mathbb{N}$, $k = 2u$, $I = ((i_1,j_1),\dots,(i_k,j_k)) \in (\{1,\dots,n\} \times \{1,\dots,p\})^k$, $\lambda = (2,\dots,2) \in C(k,u)$ and $\pi = (\pi_1,\dots,\pi_k) \in \mathfrak{S}(\lambda)$. For a tuple $v = (v_1,\dots,v_j)$, we will write $\{v\}$ instead of the set $\{v_1,\dots,v_j\}$. We consider the sets

$$\pi(I)_i := \{(i_\mu,j_\mu) \in \{I\} :\ \pi_\mu = i\} \quad (i = 1,\dots,u).$$

Obviously, $\pi(I)_i$ $(i = 1,\dots,u)$ forms a partition of $\{I\}$ with $|\pi(I)_i| = 2$. We define for π, I and a $p\cdot n \times p\cdot n$ positive definite and symmetric covariance matrix Σ,

$$\pi(\Sigma)_I := \prod_{i=1}^{u} \Sigma_{(\alpha_i,\beta_i),(\gamma_i,\delta_i)} \text{ where } \{(\alpha_i,\beta_i),(\gamma_i,\delta_i)\} = \pi(I)_i.$$

Example 5.7.5. Let $u = 2$, $I = \{(2,1),(2,2),(3,2),(2,1)\}$, $\lambda := (2,2) \in C(4,2)$ and $\pi \in \mathfrak{S}(\lambda)$. Then

$$\begin{aligned}
\pi &= (1\ 1\ 2\ 2), & \pi(I)_1 &= \{(2,1),(2,2)\}, & \pi(\Sigma)_I &= \Sigma_{(2,1),(2,2)} \cdot \Sigma_{(3,2),(2,1)} \\
\pi &= (1\ 2\ 1\ 2), & \pi(I)_1 &= \{(2,1),(3,2)\}, & \pi(\Sigma)_I &= \Sigma_{(2,1),(3,2)} \cdot \Sigma_{(2,2),(2,1)} \\
\pi &= (1\ 2\ 2\ 1), & \pi(I)_1 &= \{(2,1),(2,1)\}, & \pi(\Sigma)_I &= \Sigma_{(2,1),(2,1)} \cdot \Sigma_{(2,2),(3,2)} \\
\pi &= (2\ 1\ 1\ 2), & \pi(I)_1 &= \{(2,2),(3,2)\}, & \pi(\Sigma)_I &= \Sigma_{(2,2),(3,2)} \cdot \Sigma_{(2,1),(2,1)} \\
\pi &= (2\ 1\ 2\ 1), & \pi(I)_1 &= \{(2,2),(2,1)\}, & \pi(\Sigma)_I &= \Sigma_{(2,2),(2,1)} \cdot \Sigma_{(2,1),(3,2)} \\
\pi &= (2\ 2\ 1\ 1), & \pi(I)_1 &= \{(3,2),(2,1)\}, & \pi(\Sigma)_I &= \Sigma_{(3,2),(2,1)} \cdot \Sigma_{(2,1),(2,2)}
\end{aligned}$$

We can now formulate the generalisation of Isserlis' Theorem 5.7.1.

Theorem 5.7.6. *Let $Z = (z_{ij})$ be $p \times q$ matrix variate normal distributed random variable with mean matrix zero and covariance matrix $\Sigma \in \Pi_{p\cdot q}$. The k-th order moments $\mathrm{M}_k(Z)_I = \mathbb{E}(z_{i_1,j_1} \cdots z_{i_k,j_k})$ with $I = ((i_1,j_1),\dots,(i_k,j_k)) \in (\{1,\dots,n\} \times \{1,\dots,p\})^k$ are given as follows:*

$$\mathrm{M}_k(Z)_I = \begin{cases} 0, & \textit{if } k \textit{ is odd}, \\ \frac{1}{u!} \sum\limits_{\pi \in \mathfrak{S}(\lambda)} \pi(\Sigma)_I, & \textit{if } k = 2u,\ \lambda = (2,\dots,2) \in C(k,u). \end{cases} \tag{5.7.1}$$

Proof: Let $\{(\alpha_1,\beta_1),(\gamma_1,\delta_1)\},\dots,\{(\alpha_u,\beta_u),(\gamma_u,\delta_u)\}$ be a partition of $\{I\}$ into two pairs. It is evident, that there are $u!$ permutations $\pi \in \mathfrak{S}(\lambda)$ with

$$\Sigma_{(\alpha_1,\beta_1),(\gamma_1,\delta_1)} \cdots \Sigma_{(\alpha_u,\beta_u),(\gamma_u,\delta_u)} = \pi(\Sigma)_I.$$

Therefore, by 5.7.1 and definition of matrix variate normal distribution, the theorem follows. □

For instance, we compute the I-th entry of the 4th order moments $\mathrm{M}_4(Z)$ with $I = \{(2,1),(2,2),(3,2),(2,1)\}$ (cf. Example 5.7.5).

$$\begin{aligned}
\mathrm{M}_4(Z)_I &= \frac{1}{2!} \sum\nolimits_{\pi \in \mathfrak{S}(\lambda)} \pi(\Sigma)_I \\
&= \Sigma_{(2,1),(2,2)}\Sigma_{(3,2),(2,1)} + \Sigma_{(2,1),(3,2)}\Sigma_{(2,2),(2,1)} + \Sigma_{(2,1),(2,1)}\Sigma_{(2,2),(3,2)}.
\end{aligned}$$

The following two simple observations concerning the k-th moment of normal distributed random matrix and a sum of two independent, normal distributed random matrices, respectively will be needed for the proof of Theorem 5.9.1.

Lemma 5.7.7. *Let $Z = (z_{ij})$ be $\mathcal{N}_{p,q}(\mathbf{0}, \Sigma)$-distributed random variable and $Z_1, Z_2, \ldots$ independent copies of Z. The k-order moment of Z is given by*

$$\mathrm{M}_k(Z) = \begin{cases} \mathbf{0}, & \text{if } k \text{ is odd,} \\ \frac{1}{u!} \sum\limits_{\pi \in \mathfrak{S}(\lambda)} \mathbb{E}\left(\pi(Z_1, \ldots, Z_u)\right), & \text{if } k = 2u \end{cases}$$

where $\lambda = (2, \ldots, 2) \in C(2u, u)$.

Proof: Let $k \in \mathbb{N}$ and $I = ((i_1, j_1), \ldots, (i_k, j_k)) \in (\{1, \ldots, n\} \times \{1, \ldots, p\})^k$. If k is odd, then it follows by (5.7.1) that $\mathrm{M}_k(Z)_I = 0$. Suppose that $k = 2u$, $(u \in \mathbb{N})$. For $\pi \in \mathfrak{S}(\lambda)$, $\lambda = (2, \ldots, 2) \in C(k, u)$ and I as above, we have $\pi(Z_1, \ldots, Z_u)_I = (Z_{\pi_1} \otimes \ldots \otimes Z_{\pi_k})_I$. Let $\{(\alpha_i, \beta_i), (\gamma_i, \delta_i)\} = \pi(I)_i$, $i = 1, \ldots, u$. By independence, it follows

$$\begin{aligned} \mathbb{E}\left(\pi(Z_1, \ldots, Z_u)_I\right) &= \mathbb{E}\left((Z_{\pi_1} \otimes \ldots \otimes Z_{\pi_k})_I\right) \\ &= \mathbb{E}\Big(Z_1 \otimes Z_1\Big)_{(\alpha_1,\beta_1),(\gamma_1,\delta_1)} \cdots \mathbb{E}\Big(Z_u \otimes Z_u\Big)_{(\alpha_u,\beta_u),(\gamma_u,\delta_u)} \\ &= \Sigma_{(\alpha_1,\beta_1),(\gamma_1,\delta_1)} \cdots \Sigma_{(\alpha_u,\beta_u),(\gamma_u,\delta_u)} = \pi(\Sigma)_I. \end{aligned}$$

The lemma is now a consequence of Eq. (5.7.1). □

Lemma 5.7.8. *Let Z_i $(i = 1,\ 2)$ be independent random variables with distributions $\mathcal{N}_{p,q}(\mathbf{0}, \Sigma_i)$. Then*

$$\mathbb{E}\left(\left(Z_1 + Z_2\right)^{\otimes,k}\right) = \sum_{l=0}^{k} \sum_{\pi \in \mathfrak{S}((l,k-l))} \mathbb{E}\left(\pi(Z_1, \mathbf{1})\right) \circ \mathbb{E}\left(\pi(\mathbf{1}, Z_2)\right). \tag{5.7.2}$$

Proof: By the definition of $\circ$-product and independence of Z_1 and Z_2 we have

$$\begin{aligned} \mathbb{E}\left((Z_1 + Z_2)^{\otimes,k}\right) &= \sum_{l=0}^{k} \sum_{\pi \in \mathfrak{S}((l,k-l))} \mathbb{E}\left(\pi(Z_1, Z_2)\right) \\ &= \sum_{l=0}^{k} \sum_{\pi \in \mathfrak{S}((l,k-l))} \mathbb{E}\left(\pi(Z_1, \mathbf{1}) \circ \pi(\mathbf{1}, Z_2)\right) \\ &= \sum_{l=0}^{k} \sum_{\pi \in \mathfrak{S}((l,k-l))} \mathbb{E}\left(\pi(Z_1, \mathbf{1})\right) \circ \mathbb{E}\left(\pi(\mathbf{1}, Z_2)\right). \end{aligned}$$

□

In the situation of the above lemma, for $q = p = 1$ it is obvious that

$$m_k(\mathcal{N}_1(0, \sigma_1^2 + \sigma_2^2) = \sum_{l=0}^{k} \binom{k}{l} m_l(\mathcal{N}_1(0, \sigma_1^2) m_{k-l}(\mathcal{N}_1(0, \sigma_2^2).$$

5.8 Radial measures on $\mathbb{M}_{p,q}$ and their moments

In this section, we study radial measures on the space $\mathbb{M}_{p,q}$. In particular, we derive asymptotic results for their moments as $p \to \infty$. These results will play a key role in the proof of Theorem 5.9.1. We start with the definition of a radial measure on $\mathbb{M}_{p,q}$.

Definition 5.8.1 (radial functions and measures on $\mathbb{M}_{p,q}$)**.**

(a) A function $f : \mathbb{M}_{p,q} \to \mathbb{C}$ is called *radial* if

$$f(Ox) = f(x) \quad \forall\, x \in \mathbb{M}_{p,q},\ O \in \mathbb{O}_p.$$

(b) A measure ν_p on $\mathbb{M}_{p,q}$ is called *radial* if it is invariant under orthogonal transformation, i.e.

$$O(\nu_p) = \nu_p \quad \forall\, O \in \mathbb{O}_p.$$

Remark 5.8.2. For the classical case $q = 1$ the space $\mathbb{M}_{p,q}$ corresponds to $\mathbb{R}^p$. Here a function $f : \mathbb{R}^p \to \mathbb{C}$ is radial if and only if it is constant on each sphere $S_r^{p-1} \subset \mathbb{R}^p$ with radius $r > 0$. In particular, for each $F : \mathbb{R}_+ \to \mathbb{C}$ there exist a unique radial function $f : \mathbb{R}^p \to \mathbb{C}$ with $f(x) = F(\|x\|)$. Thus, there is one to one correspondence between radial functions on $\mathbb{R}^p$ and functions on $\mathbb{R}_+$.

In order to study radial measures on $\mathbb{M}_{p,q}$ for $q > 1$ we need an analogue of a sphere in our higher rank setting. For an $r \in \Pi_q$ we define a sphere of radius r as the set

$$\Sigma_{p,q}^r = \left\{ x \in \mathbb{M}_{p,q} :\ \sqrt{x'x} = r \right\}.$$

Clearly, $\Sigma_{p,q}^r$ is the orbit of the block matrix $\sigma_r := (r\ 0)' \in \mathbb{M}_{p,q}$ according to the operation of the orthogonal group $\mathbb{O}_p$ on $\mathbb{M}_{p,q}$ via left multiplication $(O, x) \mapsto Ox$. By the definition 5.8.1, a function $f : \mathbb{M}_{p,q} \to \mathbb{C}$ is radial if and only if it is constant on each sphere $\Sigma_{p,q}^r \subset \mathbb{M}_{p,q}$ with radius $r \in \Pi_q$. Let us consider the map

$$\varphi_p : \mathbb{M}_{p,q} \to \Pi_q, \quad \varphi_p(x) = \sqrt{x'x}.$$

By analogy to the case $q = 1$, there is a one to one correspondence between functions $F : \Pi_q \to \mathbb{C}$ and radial functions $f : \mathbb{M}_{p,q} \to \mathbb{C}$ via

$$f(x) = F(\varphi_p(x)), \quad x \in \mathbb{M}_{p,q}.$$

Accordingly, there is one to one relationship on the level of measures. Namely, each measure ν on Π_q corresponds to a (unique) radial measure ν_p on $\mathbb{M}_{p,q}$ via $\varphi_p(\nu_p) = \nu$; the measure ν is called *radial part* of ν_p.

In the following, for simplicity of notation, we write $\Sigma_{p,q}$ instead of $\Sigma_{p,q}^{\mathrm{I}_q}$, where $\mathrm{I}_q \in \mathbb{R}^{q\times q}$ denotes the identity matrix. In the case $q = 1$ we identify $\Sigma_{p,1}^r$ with the Euclidean sphere of radius $r \in [0, \infty[$. Moreover, let us denote by U_p^r the uniform distribution on a sphere $\Sigma_{p,q}^r$.

One can easily show that a radial probability measure ν_p with its radial part $\nu \in \mathcal{M}^1(\Pi_q)$ enables the decomposition

$$\nu_p(\cdot) = \int_{\mathbb{M}_{p,q}} \mathrm{U}_p^{\varphi_p(x)}(\cdot)\mathrm{d}\nu_p(x) = \int_{\Pi_q} \mathrm{U}_p^r(\cdot)\mathrm{d}\nu(r) \in \mathcal{M}^1(\mathbb{M}_{p,q}). \tag{5.8.1}$$

In the sense of Jewett [19], the formula above is an example of a decomposition of a measure (here ν_p) according to so-called orbital morphism (here φ_p). More precisely, φ_p is an orbital mapping which is a proper and open continuous surjection from $\mathbb{M}_{p,q}$ onto Π_q. The mapping $r \mapsto \mathrm{U}_p^r$ from Π_q to $\mathcal{M}^1(\mathbb{M}_{p,q})$ is a recomposition of φ_p which means that each U_p^r is a probability measure on $\mathbb{M}_{p,q}$ with support equal to $\varphi_p^{-1}(r)$ (here $= \Sigma_{p,q}^r$), and such that $\nu_p = \int_{\mathbb{M}_{p,q}} \mathrm{U}_p^{\varphi_p(x)}\mathrm{d}\nu_p(x)$.

Here and subsequently, ν_p denotes a radial probability measure on $\mathbb{M}_{p,q}$ with the corresponding radial part $\nu \in \mathcal{M}^1(\Pi_q)$ and X is $\mathbb{M}_{p,q}$-valued random variable with radial distribution ν_p. If ν admits a k-th moment ($k \in \mathbb{N}$), then

$$r_k(\nu) := \int_{\Pi_q} x^k d\nu(x) = \int_{\mathbb{M}_{p,q}} \varphi_p(x)^k \mathrm{d}\nu_p(x)$$

exists in Π_q. In this case, we call $r_k(\nu)$ the k-th modified moment of ν.

In the next lemmas, we explore the covariance structure of X and we compute the asymptotic behaviour of the moments of ν_p for large dimensions p.

Lemma 5.8.3. *Let $\nu \in \mathcal{M}^1(\Pi_q)$ and $\nu_p \in \mathcal{M}^1(\mathbb{M}_{p,q})$ be the corresponding radial probability measure on $\mathbb{M}_{p,q}$ which admits second moment. Moreover, let $X =$*

$(x_{ij})_{i,j}$ *be* $\mathbb{M}_{p,q}$*-valued random variable with distribution* ν_p*. Then*

$$\mathbb{E}\left(X\right) = \mathbf{0} \quad \textit{and} \quad \mathbb{C}ov(X) = \frac{1}{p}\mathrm{I}_p \otimes r_2(\nu).$$

Proof: For $r \in \mathbb{R} \setminus \{0\}$ let $M_{j,r}$ and $S_{i,j}$ be $p \times p$ matrices produced by multiplying all elements of row j of the identity matrix by r and by exchanging row i and row j of the identity matrix, respectively. As $S_{i,j}$ is a symmetric involution on $\mathbb{M}_p$, we have $S_{i,j} \in \mathbb{O}_p$. For $r = \pm 1$, the matrix $M_{j,r}$ is also orthogonal. By assumption, X and OX are identically distributed for any $O \in \mathbb{O}_p$. Therefore, we have

$$\mathbb{E}\left(X\right)_{ij} = \mathbb{E}\left(M_{j,-1}X\right)_{ij} = -\mathbb{E}\left(X\right)_{ij}.$$

Thus, $\mathbb{E}(x_{ij}) = 0$ for all indices i and j and hence $\mathbb{E}(X) = \mathbf{0} \in \mathbb{M}_{p,q}$.

Choose $i, k \in \{1, \ldots, q\}$ and $j, l \in \{1, \ldots, p\}$. If $j \neq l$, then we conclude from

$$\mathbb{E}\left(x_{ji}x_{lk}\right) = \mathbb{E}\left((M_{j,-1}X)_{ji}(M_{j,-1}X)_{lk}\right) = -E\left(x_{ji}x_{lk}\right)$$

that $\mathbb{E}(x_{ji}x_{lk}) = 0$. Now, suppose that $j = l$. The transformation $\mathbb{M}_{p,q} \to \mathbb{M}_{p,q}$, $A \mapsto S_{i,j}A$, switches all matrix elements on row i with their counterparts on row j. Therefore, from radiality of $P_X = \nu_p$ it follows that

$$\mathbb{E}(x_{ji}x_{jk}) = \mathbb{E}((S_{j,1}X)_{ji}(S_{j,1}X)_{jk}) = \mathbb{E}(x_{1i}x_{1k}) \quad \text{for } i, k \in \{1, \ldots, q\}.$$

Let $\mathbf{x}_i'$ denote the i-th row of the matrix X, $i = 1, \ldots, p$. From what has already been proven, we conclude that

$$r_2(\nu) = \int_{\mathbb{M}_{p,q}} X'X \mathrm{d}\nu_p(x) = \sum_{i,j=1}^{p} \mathbb{E}(\mathbf{x}_i\mathbf{x}_i') = p \cdot \mathbb{E}(\mathbf{x}_1\mathbf{x}_1') \in \mathbb{M}_q.$$

Hence,

$$\mathbb{C}ov(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{p} \cdot \delta_{ij} r_2(\nu) \in \mathbb{M}_q$$

for all $i,\ j \in \{1, \ldots, p\}$. Now, the desired formula for the covariance of X is a consequence of the Equation (5.5.1).

□

To derive formulas for the moments of radial distributed random variable, as in the case $q = 1$, we first compute the characteristic function of the uniform distribution on $\Sigma_{p,q}^r$ for $r \in \Pi_q$.

Lemma 5.8.4. *The characteristic function for the uniform distribution U_p^r on the sphere $\Sigma_{p,q}^r$ of radius $r \in \Pi_q$ is given by*

$$\widehat{\mathrm{U}_p^r}(z) = \mathcal{J}_\mu\left(\frac{1}{4}(zr)^*(zr)\right), \quad z \in \mathbb{M}_{p,q} \tag{5.8.2}$$

where $\mu = p/2$ and $\mathcal{J}_\mu$ is the Bessel function of index μ as in (5.3.3).

Proof: Let $r \in \Pi_q$. Consider the map

$$T_r : \Sigma_{p,q} \to \Sigma_{p,q}^r, \quad y \mapsto yr.$$

Since $T_r(\mathrm{U}_p^{\mathrm{I}_q}) = \mathrm{U}_p^r$, we get by substitution formula

$$\widehat{\mathrm{U}_p^r}(z) = \int_{\mathbb{M}_{p,q}} e^{i\langle z,y\rangle} \mathrm{dU}_p^r(y) = \int_{\Sigma_{p,q}} e^{i\langle z,yr\rangle} \mathrm{dU}_p^{\mathrm{I}_q}(y).$$

On the other hand, according to Proposition XVI.2.3. of [11], we have

$$\int_{\Sigma_{p,q}} e^{i\langle y,x\rangle} \mathrm{dU}_p^{\mathrm{I}_q}(y) = \mathcal{J}_\mu\left(\frac{1}{4}x^*x\right), \quad \mu = \frac{p}{2},$$

for $x \in \mathbb{M}_{p,q}$. By taking these two identities above into account, (5.8.2) follows as claimed. □

Remark 5.8.5. In the case $q = 1$, we obtain from Lemma 5.8.4 and Remark 5.3.4 the following formula for the characteristic function of the uniform distribution $\mathrm{U}_p^r \in \mathcal{M}^1(\mathbb{R}^p)$ on the sphere $S_r^{p-1} \subset \mathbb{R}^p$

$$\mathrm{U}_p^r(z) = \Lambda_{\frac{p}{2}-1}(r \cdot \|z\|), \quad z \in \mathbb{R}^p,$$

where Λ_λ is the modified Bessel function of index λ.

Lemma 5.8.6. *Let $\kappa \in \mathbb{N}_0^{p\times q}$, $l = |\kappa|/2$ and $\mu = p/2$. The κ-th moment $m_\kappa(\mathrm{U}_p^r)$ of the uniform distribution on $\Sigma_{p,q}^r$ is given as follows:*

(a) If $\mathrm{R}_i(\kappa) = \sum_{j=1}^q \kappa_{ij}$ is even for all $i = 1, \dots, p$, then $l \in \mathbb{N}_0$ and

$$m_\kappa(\mathrm{U}_p^r) = \frac{1}{4^l\,|\kappa|!} \sum_{\lambda \in C_0(l,q)} \frac{1}{(\mu)_\lambda^{1/2}} \mathrm{D}_\kappa\Big(\mathrm{Z}_\lambda((zr)'(zr))\Big)_{|z=0}. \tag{5.8.3}$$

(b) If $\mathrm{R}_i(\kappa)$ *is not even for some* $i = 1, \ldots, p$*, then* $m_\kappa(\mathrm{U}_p^r) = 0$.

Proof: By the Identity (5.5.4), the preceding lemma and (5.3.3) we have

$$m_\kappa(\mathrm{U}_p^r) = (-i)^{|\kappa|} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \sum_{\lambda \in C_0(j,q)} \frac{1}{(\mu)_\lambda^{1/2}} \mathrm{D}_\kappa\Big(\mathrm{Z}_\lambda\Big(\frac{1}{4}(zr)^*(zr)\Big)\Big)_{|z=0}. \tag{5.8.4}$$

Let $\lambda \in \mathbb{N}_0^q$ and $p_r : z \mapsto \mathrm{Z}_\lambda\left((zr)^*(zr)\right)$. Since Z_λ is a homogeneous polynomial of degree $|\lambda|$, Lemma 5.4.1 shows that p_r is a homogeneous, even-row polynomial of degree $2\,|\lambda|$. Therefore, each term on the right-hand side of (5.8.4) vanishes if $\kappa \in \mathbb{N}_0^{p\times q}$ with $\mathrm{R}_i(\kappa)$ is odd for some $i \in \{1, \ldots, p\}$ or if $|\kappa| \neq 2\,|\lambda|$. This proves the assertion. □

Unfortunately, the formula (5.8.3) for $q > 1$ can not be simplified as in the case $q = 1$ (cf. 4.0.10). Nevertheless, we can easily indicate the asymptotic behavior of $m_\kappa(\mathrm{U}_p^r)$ for $p \to \infty$: since each generalised Pochhammer symbol in the finite sum (5.8.3) is of the order $O(p^{-l})$, so is $m_\kappa(\mathrm{U}_p^r)$. We now compute an upper bound for the κ-th moment of a radial distribution ν_p as $p \to \infty$.

Theorem 5.8.7. *Let* $k,\ q,\ p \in \mathbb{N}$ *with* $p > 2q$ *and* ν_p *be a radial probability measure on* $\mathbb{M}_{p,q}$ *which admits* k*-th moment. Then there is a constant* $C > 0$ *such that*

$$\|\mathrm{M}_k(\nu_p)\|_\infty \leq C \cdot p^{-k/2}.$$

Proof: Let $\kappa \in \mathbb{N}_0^{p\times q}$ with $|\kappa| = k$. Since ν_p admits k-th moment, it is clear that $m_\kappa(\nu_p)$ exists. By the decomposition (5.8.1), we obtain

$$m_\kappa(\nu_p) = \int_{\Pi_q} m_\kappa(\mathrm{U}_p^r)\mathrm{d}\nu(r) \tag{5.8.5}$$

where U_p^r is the uniform distribution on $\Sigma_{p,q}^r$ and ν the radial part of ν_p. Thus, according to the above lemma, we have

$$0 \leq |m_\kappa(\nu_p)| \leq \frac{1}{2^k k!} \sum_{\lambda \in C_0(l,q)} \frac{1}{(p/2)_\lambda^{1/2}} \int_{\Pi_q} |\mathrm{P}_\lambda(r)|\,\mathrm{d}\nu(r), \tag{5.8.6}$$

where $l = k/2$ and P_λ are functions on Π_q defined by

$$\mathrm{P}_\lambda(r) = \mathrm{D}_\kappa\Big(\mathrm{Z}_\lambda((zr)^*(zr))\Big)_{|z=0}, \quad \lambda \in C_0(l,q).$$

For a polynomial P on $\mathbb{M}_q$ let $C_{\max}(P)$ be the highest absolute value of the coefficients of P in its monomial expansion. It is clear that P_λ is homogeneous polynomial in the variable r on Π_q of degree k and

$$C_{\max}(\mathrm{P}_\lambda) \leq k! C_{\max}(\mathrm{Z}_\lambda).$$

Since the number of terms in the monomial expansion of P_λ, $\lambda \in C_0(l, q)$ depends only on q and k, there exists a constant $C_1 = C_1(q, k)$ such that

$$\int_{\Pi_q} |\mathrm{P}_\lambda(r)| \, \mathrm{d}\nu(r) \leq C_1 \cdot |\nu|_k \, .$$

Hence, the definition of the generalised Pochhammer symbol in 5.3.3 and Inequality (5.8.6) yield that for sufficiently large p,

$$|m_\kappa(\nu_p)| \leq C_2 \cdot p^{-k/2}$$

with a suitable constant $C_2 = C_2(\nu, q, k)$ which is independent of κ and p. From this, we finally conclude that

$$\|\mathrm{M}_k(\nu_p)\|_\infty = \sup_{\kappa \in \mathbb{N}_0^{p \times q} : |\kappa| = k} |m_\kappa(\nu_p)| \leq C_2 \cdot p^{-k/2}.$$

□

Corollary 5.8.8. *Let $k \in \mathbb{N}$, $\kappa \in \mathbb{N}_0^{p \times q}$ with $k = |\kappa|$ and $\nu_p \in \mathcal{M}^1(\mathbb{M}_{p,q})$ be a radial probability measure on $\mathbb{M}_{p,q}$ which admits k-th moment. Then $m_\kappa(\nu_p)$ exists in $\mathbb{R}$ and has the following asymptotic as $p \to \infty$:*

(a) If $\mathrm{R}_i(\kappa)$ is even for all $i = 1, \ldots, p$, then $m_\kappa(\nu_p) = O\left(p^{-k/2}\right)$.

(b) If $\mathrm{R}_i(\kappa)$ is not even for some $i = 1, \ldots, p$, then $m_\kappa(\nu_p) = 0$.

Proof: (a) is clear by the Theorem 5.8.7. The assertion (b) follows immediately from (5.8.5) and Lemma 5.8.6. □

Definition 5.8.9. Let $(D_n)_{n \in \mathbb{N}}$, $(d_n)_{n \in \mathbb{N}}$ be sequences of matrices from $\mathbb{M}_q$ and positive real numbers, respectively. We write $D_n = O(d_n)$ as $n \to \infty$, if and only if $\|D_n\|_\infty = O(d_n)$ as $n \to \infty$.

Remark 5.8.10. Let $(p_n)_{n \in \mathbb{N}} \subset \mathbb{N}$ with $p_n \to \infty$. For an $n \in \mathbb{N}$ we consider radial measure $\nu_{p_n} \in \mathcal{M}^1(\mathbb{M}_{p_n,q})$ associated with $\nu \in \mathcal{M}^1(\Pi_q)$ and p_n. Then $(\mathrm{M}_k(\nu_{p_n}))_{n \geq 1} \subset \mathbb{M}_{q^k}$ and by Theorem 5.8.7, we have $\mathrm{M}_k(\nu_{p_n}) = O(p_n^{-k})$ for $n \to \infty$.

5.9 Radial limit theorems on $\mathbb{M}_{p,q}$ for $p \to \infty$

Let $\nu \in \mathcal{M}^1(\Pi_q)$ be a fixed probability measure which admits the second order moment. In particular, the modified moment $r_2(\nu) = \int_{\Pi_q} x^2 d\nu(x)$ exists in Π_q. As we have mentioned in Section 5.8, for each dimension $p \in \mathbb{N}$ there is a unique radial probability measure $\nu_p \in \mathcal{M}^1(\mathbb{M}_{p,q})$ with ν as its radial part, i.e., $\nu = \varphi_p(\nu_p)$. For each $p \in \mathbb{N}$ consider i.i.d. $\mathbb{M}_{p,q}$-valued random variables

$$\mathbb{X}_k := \left(X_k^{(i,j)}\right)_{1\leq i\leq p,\ 1\leq j\leq q}, \quad k \in \mathbb{N}$$

with law ν_p as well as the random variables

$$\Xi_n^p(\nu) := \varphi_p(S_n^p)^2 - nr_2(\nu), \tag{5.9.1}$$

where $S_n^p := \sum_{k=1}^n \mathbb{X}_k$. Let $(p_n)_{n\in\mathbb{N}} \subset \mathbb{N}$ be a sequence with $\lim_{n\to\infty} p_n = \infty$. In the first part of this section, we derive the following two complementary CLTs for $\mathbb{M}_q$-valued random variables $\Xi_n(\nu) := \Xi_n^{p_n}(\nu)$ under disjoint growth conditions for the dimensions p_n.

Theorem 5.9.1. *Assume that $\nu \in \mathcal{M}^1(\Pi_q)$ admits finite fourth moment.*
CLT I: *If $\lim_{n\to\infty} n/p_n = \infty$, then $\sqrt{p_n}/n \cdot \Xi_n(\nu)$ tends in distribution to the centered matrix variate normal distribution $\mathcal{N}_{q,q}(\mathbf{0}, T(\nu))$ on $\mathbb{M}_q$ with covariance matrix $T(\nu) := T_1(\nu) + T_2(\nu) \in \Pi_{q^2}$ where*

$$T_1(\nu)_{(i,j),(k,l)} = r_2(\nu)_{i,k} r_2(\nu)_{j,l} \quad \text{and} \quad T_2(\nu)_{(i,j),(k,l)} = r_2(\nu)_{i,l} r_2(\nu)_{j,k}.$$

CLT II: *If $\lim_{n\to\infty} n/p_n = c \in [0,\infty[$, then $\Xi_n(\nu)/\sqrt{n}$ tends to the centered matrix variate normal distribution $\mathcal{N}_{q,q}(\mathbf{0}, \Sigma(\nu) + cT(\nu))$ on $\mathbb{M}_q$ where $\Sigma(\nu) \in \Pi_{q^2}$ is the covariance matrix of $\varphi_{p_n}^2(\nu_{p_n})$ which is independent of p_n.*

The proof will be divided into two main steps: In the first step, we prove a reduced form of Theorem 5.9.1 assuming that ν has a compact support. In the second step, we will show how to dispense with the assumption on the support of ν. Both steps are based on the decomposition of $\Xi_n(\nu)$ via

$$\Xi_n(\nu) = \mathfrak{A}_n(\nu) + \mathfrak{B}_n(\nu)$$

where

$$\mathfrak{A}_n(\nu) := \sum_{i=1}^{n} A_i, \quad \text{with } A_i := \varphi_{p_n}(\mathbb{X}_i)^2 - r_2(\nu), \tag{5.9.2}$$

$$\mathfrak{B}_n(\nu) := \sum_{i=1}^{p_n} B_i, \quad \text{with } B_i := \sum_{\alpha\neq\beta}^{n} \left[X_\alpha^{(i,j)} X_\beta^{(i,l)}\right]_{1\leq j,l\leq q}. \tag{5.9.3}$$

We compute the covariance structure of $\mathfrak{A}_n(\nu)$ and $\mathfrak{B}_n(\nu)$. Since the random variables A_i, $i \in \mathbb{N}$ are independent and identically distributed, it is easily seen that

$$\mathbb{E}(A_k) = \mathbf{0} \quad \text{and} \quad \mathbb{C}ov(A_i, A_j) = \delta_{i,j}\Sigma(\nu). \tag{5.9.4}$$

This gives $\mathbb{C}ov(\mathfrak{A}_n(\nu)) = n \cdot \Sigma(\nu)$. By the independence of random variables $\mathbb{X}_k$, $k \in \mathbb{N}$, using the Lemma 5.8.3, we obtain

$$\mathbb{E}(B_k) = \mathbf{0} \quad \text{and} \quad \mathbb{C}ov(B_i, B_j) = \delta_{i,j}\frac{n(n-1)}{p_n^2}T(\nu). \tag{5.9.5}$$

Hence, $\mathbb{C}ov(\mathfrak{B}_n(\nu)) = n(n-1)/p_n \cdot T(\nu)$. Corollary 5.8.8, b) yields $\mathbb{C}ov(A_i, B_j) = \mathbf{0}$ for all i and j. This clearly forces

$$\mathbb{C}ov(\Xi_n(\nu)) = \mathbb{C}ov(\mathfrak{A}_n(\nu)) + \mathbb{C}ov(\mathfrak{B}_n(\nu)) = n\Sigma(\nu) + \frac{n(n-1)}{p_n}T(\nu). \tag{5.9.6}$$

In the following we will establish convergence in distribution of $\mathfrak{A}_n(\nu)$ and $\mathfrak{B}_n(\nu)$ (after appropriate scaling) by the method of moments 5.6.4.

Proposition 5.9.2. *Assume that $\nu \in \mathcal{M}^1(\Pi_q)$ has compact support. Then the asymptotic behaviour of $\mathfrak{A}_n := \mathfrak{A}_n(\nu)$ is given as follows:*

a) If $n/p_n \to c \in [0, \infty[$ as $n \to \infty$, then $\mathfrak{A}_n/\sqrt{n}$ tends in distribution to matrix variate normal distribution $\mathcal{N}_{q,q}(\mathbf{0}, \Sigma(\nu))$.

b) If $n/p_n \to \infty$ as $n \to \infty$, then $\sqrt{p_n}/n \cdot \mathfrak{A}_n$ tends in distribution to $\delta_{\mathbf{0}}$.

Proof: If we prove that for all $k \in \mathbb{N}_0$, the k-th order moments

$$\mathrm{M}_k(\tfrac{1}{\sqrt{n}}\mathfrak{A}_n) = \frac{1}{n^{k/2}}E(\mathfrak{A}_n^{\otimes,k}) \quad \text{and} \quad \mathrm{M}_k(\tfrac{\sqrt{p_n}}{n}\mathfrak{A}_n) = \frac{p_n^{k/2}}{n^k}E(\mathfrak{A}_n^{\otimes,k}) \tag{5.9.7}$$

tend to the k-th order moment of the corresponding limit distribution, then the assertion follows by the method of moments 5.6.4. Therefore, we calculate (5.9.7) as $n \to \infty$. Since the random variables A_j are identically distributed, Theorem 5.2.1 shows that

$$\mathbb{E}(\mathfrak{A}_n^{\otimes,k}) = \sum_{u=1}^{k} \sum_{\lambda \in C(k,u)} \binom{n}{u} \sum_{\pi \in \mathfrak{S}(\lambda)} \mathbb{E}\pi(A_1, \dots, A_u).$$

For $u \in \{1, \dots, k\}$ and $\lambda \in C(k,u)$ we consider

$$T(\lambda) := \binom{n}{u} \sum_{\pi \in \mathfrak{S}(\lambda)} \mathbb{E}\pi(A_1, \dots, A_u). \tag{5.9.8}$$

If $\lambda_\alpha = 1$ for some α, i.e. A_α appears exactly once in $\pi(A_1, \dots, A_u)$, then each summand in (5.9.8) vanishes, which is due to the facts that $\mathbb{E}(A_\alpha) = \mathbf{0} \in \mathbb{M}_q$ and A_i are independent.

Suppose that $\lambda_\alpha \geq 2$ for each α and $\lambda_\alpha > 2$ for some α. Then $k > 2u$, and since ν has compact support, we get $T(\lambda) = O(n^u)$. This clearly forces $n^{-k/2}T(\lambda)$ and $p_n^{k/2} n^{-k} T(\lambda)$ in the cases $n/p_n \to c \in [0, \infty[$ and $n/p_n \to \infty$, respectively tend to zero as $n \to \infty$.

Now we turn to the case $\lambda = (2, \dots, 2)$, in particular $k = 2u$. Let $Z_1, \dots, Z_u$ be independent and $\mathcal{N}_{q,q}(\mathbf{0}, \Sigma(\nu))$ distributed random variables. By Lemma 5.1.1, for any $\pi \in \mathfrak{S}(\lambda)$ there exist permutation matrices P_π and Q_π with

$$\begin{aligned} P_\pi \cdot \mathbb{E}\pi(A_1, \dots, A_u) \cdot Q_\pi &= \mathbb{E}(A_1 \otimes A_1 \otimes \ldots \otimes A_u \otimes A_u) \\ &= \mathbb{E}(Z_1 \otimes Z_1 \otimes \ldots \otimes Z_u \otimes Z_u) = P_\pi \cdot \mathbb{E}\pi(Z_1, \dots, Z_u) \cdot Q_\pi. \end{aligned}$$

Hence $\mathbb{E}\pi(A_1, \dots, A_u) = \mathbb{E}\pi(Z_1, \dots, Z_u)$ for all $\pi \in \mathfrak{S}(\lambda)$. Therefore, according to the Lemma 5.7.7, we have

$$T(\lambda) = \binom{n}{u} \sum_{\pi \in \mathfrak{S}(\lambda)} \mathbb{E}\pi(Z_1, \dots, Z_u) = \frac{n!}{(n-u)!} \mathrm{M}_k(Z_1).$$

This proves that the moments in (5.9.7) converge to those of $\mathcal{N}_{q,q}(\mathbf{0}, \Sigma(\nu))$ and $\delta_{\mathbf{0}}$ distributions on $\mathbb{M}_q$, respectively. □

We proceed to state the corresponding convergence result for the random variable $\mathfrak{B}_n(\nu)$ as $n \to \infty$.

Proposition 5.9.3. *Assume that $\nu \in \mathcal{M}^1(\Pi_q)$ has compact support. Then the asymptotic behaviour of $\mathfrak{B}_n := \mathfrak{B}_n(\nu)$ is given as follows:*

a) If $n/p_n \to 0$ as $n \to \infty$, then $\mathfrak{B}_n/\sqrt{n}$ tends in distribution to $\delta_{\mathbf{0}}$.

b) If $n/p_n \to c \in]0,\infty]$ as $n \to \infty$, then $\sqrt{p_n}/n \cdot \mathfrak{B}_n$ tends in distribution to the matrix variate normal distribution $\mathcal{N}_{q,q}(\mathbf{0}, T(\nu))$.

For the proof we need a couple of simple lemmas. We first introduce some notation. Let $k,\ n \in \mathbb{N}$ and $\mathcal{I}_{k,n}$ the set of all $2k$-tuples $(i_1, j_1, \ldots, i_k, j_k)$ of positive integers less or equal n such that $i_\alpha \neq j_\alpha$ for all $\alpha = 1, \ldots, k$. For an $I \in \mathcal{I}_{k,n}$ and $\pi = (\pi_1, \ldots, \pi_k) \in \mathbb{N}^k$ we consider the random matrix

$$\mathrm{S}(I,\pi) := \bigotimes_{l=1}^{k} \left[X_{i_l}^{(\pi_l,\alpha_l)} X_{j_l}^{(\pi_l,\beta_l)}\right]_{1\le\alpha_l,\beta_l\le q}. \tag{5.9.9}$$

Each entry of $\mathrm{S}(I,\pi) \in \mathbb{M}_{q^k}$ is a product with k factors and corresponds to the tuple

$$\Big((i_1,\pi_1,\alpha_1),(j_1,\pi_1,\beta_1),\ldots,(i_k,\pi_k,\alpha_k),(j_k,\pi_k,\beta_k)\Big). \tag{5.9.10}$$

For (5.9.10) and two integers a, b we define

$$\mathrm{mult}_{I,\pi}(a,b) = \left|\{\tau \in \{1,\ldots,k\} : (i_\tau,\pi_\tau) = (a,b) \text{ or } (j_\tau,\pi_\tau) = (a,b)\}\right|.$$

It is clear that $\mathrm{mult}_{I,\pi}(a,b)$ does not depend on the indices α_τ and β_τ. Therefore, $\mathrm{mult}_{I,\pi}(a,b)$ is the number of factors in an arbitrary entry of the matrix $\mathrm{S}(I,\pi)$ which are coming from the b-th row of $\mathbb{X}_a$. Moreover, we write $\mathrm{d}(I)$ for the number of distinct elements in $\{I\}$. We consider the following subsets of $\mathcal{I}_{k,n}$

$$\begin{aligned}
\mathrm{J}_m &:= \{I \in \mathcal{I}_{k,n} : \quad \mathrm{d}(I) = m\}, \\
\tilde{\mathrm{J}}_m &:= \{I \in \mathrm{J}_m : \quad \{I\} = \{1,\ldots,m\}\}, \\
\mathrm{J}^o(\pi) &:= \{I \in \mathcal{I}_{k,n} : \quad \exists\ a,\ b \in \mathbb{N} : \quad \mathrm{mult}_{I,\pi}(a,b) \text{ is odd}\}.
\end{aligned} \tag{5.9.11}$$

It is easy to check that $\mathrm{d}(I) \in \{2,\ldots,2k\}$ for all $I \in \mathcal{I}_{k,n}$ and that the cardinality of J_m is at most of the order $O(n^m)$ for $n \to \infty$.

Lemma 5.9.4. *Let $\lambda = (\lambda_1,\ldots,\lambda_v) \in C(k,v)$, $\mu = (\mu_1,\ldots,\mu_v) \in W(n,v)$ and $\pi \in \mathfrak{S}(\lambda)$. Then*

$$\pi(B_{\mu_1},\ldots,B_{\mu_v}) \stackrel{\mathcal{L}}{=} \pi(B_1,\ldots,B_v).$$

Proof: By Theorem 5.1.1 it is sufficient to show that

$$B_{\mu_1}^{\otimes,\lambda_1} \otimes \ldots B_{\mu_v}^{\otimes,\lambda_v} \stackrel{\mathcal{L}}{=} B_1^{\otimes,\lambda_1} \otimes \ldots B_v^{\otimes,\lambda_v}. \tag{5.9.12}$$

Since $\mathbb{X}_k,\ k \in \mathbb{N}$ are i.i.d. with radial law ν_{p_n} we have $B_\mu \overset{\mathcal{L}}{=} B_1$ for any $\mu \in \{1, \ldots, p_n\}$. Thus, we get $(B_{\mu_1}, \ldots, B_{\mu_v}) \overset{\mathcal{L}}{=} (B_1, \ldots, B_v)$. This clearly forces 5.9.12. □

Lemma 5.9.5. *Let $\lambda = (\lambda_1, \ldots, \lambda_v) \in C(k, v)$ and $\pi \in \mathfrak{S}(\lambda)$. Then there is a constant C such that*

$$\sup_{I \in \mathcal{I}_{k,n}} \|\mathbb{E}S(I, \pi)\|_\infty \le C \cdot p_n^{-k}.$$

Proof: Let $I \in \mathcal{I}_{k,n}$ and π be as above. By the definition of $S(I, \pi)$ and Theorem 5.8.7, we obtain

$$\|\mathbb{E}S(I, \pi)\|_\infty \le \prod_{j=1}^{v} \left\|\mathrm{M}_{2\lambda_j}(\nu_{p_n})\right\|_\infty \le C \cdot p_n^{-k}$$

for some suitable constant $C > 0$ which is independent of I and p. □

Lemma 5.9.6. *Let $v \in \mathbb{N}$, $k = 2v$, $\mu = (2, \ldots, 2) \in C(k, v)$, $\pi \in \mathfrak{S}(\mu)$ and $Z_1, \ldots, Z_v$ be independent $\mathcal{N}(\mathbf{0}, T(\nu))$ distributed random variables. Then*

$$\mathbb{E}\pi(Z_1, \ldots, Z_v) = \frac{p_n^k}{k!} \sum_{I \in \tilde{\mathrm{J}}_k} \mathbb{E}\mathrm{S}(I, \pi). \tag{5.9.13}$$

Proof: According to the Theorem 5.1.1 there is no loss of generality in assuming that $\pi = (1, 1, \ldots, v, v)$. We set

$$\mathrm{J}_{k,\pi} := \left\{(i_1, j_1, \ldots, i_k, j_k) \in \tilde{\mathrm{J}}_k : \quad \{i_\alpha, j_\alpha\} = \{i_\beta, j_\beta\} \text{ if } \pi_\alpha = \pi_\beta\right\}.$$

It is easy to check that $\tilde{\mathrm{J}}_k \setminus \mathrm{J}_{k,\pi} \subset J^o(\pi)$. Therefore, by Eq. (5.9.16),

$$\sum_{I \in \tilde{\mathrm{J}}_k} \mathbb{E}\mathrm{S}(I, \pi) = \sum_{I \in \mathrm{J}_{k,\pi}} \mathbb{E}\mathrm{S}(I, \pi).$$

For a permutation $\sigma \in \mathrm{Sym}(\{1, \ldots, k\}) =: \mathrm{S}_k$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_u) \in \mathbb{Z}_2^v$ we consider the following functions:

$$\begin{aligned} \varphi_\sigma :& \mathrm{J}_{k,\pi} \longrightarrow \mathrm{J}_{k,\pi}, \qquad (i_1, j_1, \ldots, i_k, j_k) \mapsto (\sigma(i_1), \sigma(j_1), \ldots, \sigma(i_k), \sigma(j_k)) \\ \theta_\varepsilon :& \mathrm{J}_{k,\pi} \longrightarrow \mathrm{J}_{k,\pi}, \qquad (i_1, j_1, \ldots, i_k, j_k) \mapsto (r_1, t_1, \ldots, r_k, t_k), \end{aligned}$$

where $(r_1, t_1, \ldots, r_k, t_k)$ is defined as follows: for any $\alpha, \beta \in M_k$ with $\alpha < \beta$ and

$\pi_\alpha = \pi_\beta \in M_v$ we have

$$(r_\alpha, t_\alpha, r_\beta, t_\beta) = \begin{cases} (i_\alpha, j_\alpha, i_\beta, j_\beta), & \text{if } \varepsilon_{\pi_\alpha} = 0, \\ (i_\alpha, j_\alpha, j_\beta, i_\beta), & \text{if } \varepsilon_{\pi_\alpha} = 1. \end{cases}$$

It is easily seen that φ_σ and θ_ε are well defined. Let $I_0 := (1, 2, 1, 2, 3, 4, 3, 4, \ldots, k-1, k, k-1, k) \in \mathrm{J}_{k,\pi}$. By standard verification we obtain a one to one correspondence between $\mathrm{S}_k \times \mathbb{Z}_2^v$ and $\mathrm{J}_{k,\pi}$ via the map $\Psi : (\sigma, \varepsilon) \mapsto \varphi_\sigma(\theta_\varepsilon(I_0))$. Since $\mathbb{X}_1, \mathbb{X}_2, \ldots$ are i.i.d., we have for all $\sigma \in \mathrm{S}_k$

$$\mathbb{E}S(\varphi_\sigma(I), \pi) = ES(I, \pi) \quad \forall\ I \in \mathrm{J}_{k,\pi}. \tag{5.9.14}$$

For an $\varepsilon \in \mathbb{Z}_2$ we consider algebraic operation

$$\varepsilon(a, b) = \begin{cases} a, & \text{if } \varepsilon = 0, \\ b, & \text{if } \varepsilon = 1. \end{cases}$$

Let $\mathrm{R}(\pi)$ denote the right-hand side of Equation (5.9.13). Then, by Equation (5.9.14), it follows that

$$\begin{aligned}
\mathrm{R}(\pi) &= \frac{p_n^k}{k!} \sum_{(\sigma,\varepsilon)\in \mathrm{S}_k \times \mathbb{Z}_2^v} \mathbb{E}S(\Psi(\sigma, \varepsilon), \pi) \\
&= p_n^k \sum_{\varepsilon \in \mathbb{Z}_2^u} \mathbb{E}S(\Psi(id, \varepsilon), \pi) = p_n^k \sum_{\varepsilon \in \mathbb{Z}_2^u} \mathbb{E}S(\theta_\varepsilon(I_0), \pi) \\
&= \sum_{\varepsilon \in \mathbb{Z}_2^u} \bigotimes_{j=1}^{v} p_n^2 \mathbb{E}\left(\left[X_{2j-1}^{(j,\alpha)} X_{2j}^{(j,\beta)} \right]_{1\le\alpha,\beta\le q} \otimes \left[X_{\varepsilon_j(2j-1,2j)}^{(j,\alpha)} X_{\varepsilon_j(2j,2j-1)}^{(j,\beta)} \right]_{1\le\alpha,\beta\le q} \right) \\
&= \sum_{\varepsilon \in \mathbb{Z}_2^v} \bigotimes_{j=1}^{v} \varepsilon_j(T_1(\nu), T_2(\nu)).
\end{aligned}$$

On the other hand, by the independence of $Z_1, \ldots, Z_v$, we see that

$$\begin{aligned}
\mathbb{E}\pi(Z_1, \ldots, Z_v) &= \bigotimes_{j=1}^{v} \mathbb{E}(Z_j \otimes Z_j) = \bigotimes_{j=1}^{v} T(\nu) \\
&= \bigotimes_{j=1}^{v} (T_1(\nu) + T_2(\nu)) = \sum_{\varepsilon \in \mathbb{Z}_2^v} \bigotimes_{j=1}^{v} \varepsilon_j(T_1(\nu), T_2(\nu)).
\end{aligned}$$

This completes the proof. □

Proof of Proposition 5.9.3: According to the Theorem 5.6.4, it suffices to show that the k-th moments of $\mathfrak{B}_n/\sqrt{n}$ and of $\sqrt{p_n}/n \cdot \mathfrak{B}_n$ tend to the corresponding ones of the limiting distributions as $n \to \infty$. By using very similar arguments as in the proof of the Theorem 5.8.3, it is easily seen that B_i $(i = 1, 2, \ldots)$ are identically distributed. From this and Theorem 5.2.1, we conclude that

$$\mathbb{E}\Big(\mathfrak{B}_n^{\otimes,k}\Big) = \sum_{v=1}^{k} \sum_{\mu \in C(k,v)} \binom{p_n}{v} \sum_{\pi \in \mathfrak{S}(\mu)} \mathbb{E}\pi(B_1, \ldots, B_v).$$

For a $v \in \{1, \ldots, k\}$, $\lambda \in C(k, v)$ and $\pi \in \mathfrak{S}(\mu)$ we consider $\pi(B_1, \ldots, B_v)$. The definition of B_a $(a \in \{1, \ldots, v\})$ in (5.9.3) enables us to write

$$\pi(B_1, \ldots, B_v) = B_{\pi_1} \otimes \ldots \otimes B_{\pi_k} = \sum_{I \in \mathcal{I}_{k,n}} \mathrm{S}(I, \pi), \tag{5.9.15}$$

where each term $\mathrm{S}(I, \pi)$ with $I = (i_1, j_1, \ldots, i_k, j_k)$ is given by (5.9.9). For a selected index $c \in M_n$, each entry of the k-fold Kronecker product $\mathrm{S}(I, \pi)$ may be regarded as a monomial in the variables $\mathbb{X}_c$ (i.e. in $X_c^{(\alpha,\beta)}$ with $\alpha \in M_{p_n}$, $\beta \in M_q$), while the random variables coming from other indices are considered as constant. In this view, for any $I \in \mathrm{J}^o(\pi)$, each entry of $\mathrm{S}(I, \pi)$ is for some $c \in M_n$ and $a \in M_v$ a monomial in the variable $\mathbb{X}_c$, which is not even in row a. Hence, Corollary 5.8.8 clearly forces

$$\mathbb{E}\mathrm{S}(I, \pi) = \mathbf{0} \quad \forall\, I \in \mathrm{J}^o(\pi). \tag{5.9.16}$$

Therefore, since $\mathrm{J}_m \subset \mathrm{J}^o(\pi)$ for $m > k$, we conclude from (5.9.15) that

$$\mathbb{E}\pi(B_1, \ldots, B_v) = \sum_{m=2}^{k} \sum_{I \in \mathrm{J}_m} \mathbb{E}\mathrm{S}(I, \pi). \tag{5.9.17}$$

By the definition of J_m in (5.9.11), it is obvious that the number of terms in the last sum is at most of the order $O(n^m)$. Therefore, by Lemma 5.9.5 and Definition (5.8.9),

$$E\pi(B_1, \ldots, B_v) = \sum_{I \in \mathrm{J}_k} \mathbb{E}\mathrm{S}(I, \pi) + O\left(\frac{n^{k-1}}{p_n^k}\right) = O\left(\frac{n^k}{p_n^k}\right). \tag{5.9.18}$$

For $v \in \{1, \ldots, k\}$ and $\mu \in C(k, v)$ let us consider

$$T(\mu) := \binom{p_n}{v} \sum_{\pi \in \mathfrak{S}(\mu)} \mathbb{E}\pi(B_1, \ldots, B_v). \tag{5.9.19}$$

If $\mu_\alpha = 1$ for some α, i.e for any $\pi \in \mathfrak{S}(\mu)$ the factor B_α appears exactly once in the product $\pi\Big(B_1, \dots, B_v\Big)$, and therefore, each $I \in \mathcal{I}_{k,n}$ from the Representation (5.9.15) of $\pi\Big(B_1, \dots, B_v\Big)$ is necessarily from $\mathrm{J}^o(\pi)$. Hence, (5.9.16) gives $T(\mu) = 0$.

Suppose that $\mu_\alpha \geq 2$ for each α and $\mu_\alpha > 2$ for some α, i.e. in particular $k > 2v$. From (5.9.18), we conclude that $n^{-k/2}T(\mu) = O\left(n^{k/2}p_n^{v-k}\right)$ and $p_n^{k/2}n^{-k}T(\mu) = O\left(p_n^{v-k/2}\right)$ tend to $\mathbf{0}$ in the case (a) $n/p_n \to 0$ and case (b) $p_n/n \to 0$, respectively as $n \to \infty$.

We now turn to the case $\mu = (2, \dots, 2)$, in particular $k = 2v$. By Equation (5.9.18), it follows in the case (a) that $n^{-v}T(\mu) = O((n/p_n)^{k-v})$. Hence, $n^{-v}T(\mu)$ converges to zero as $n \to \infty$. Since $\mathbb{X}_1, \mathbb{X}_2, \dots$ are i.i.d., we have

$$\sum_{I \in \mathrm{J}_k} \mathbb{E}\mathrm{S}(I, \pi) = \binom{n}{k} \sum_{I \in \tilde{\mathrm{J}}_k} \mathbb{E}\mathrm{S}(I, \pi).$$

Therefore, by using Equation (5.9.18),

$$T(\mu) = \frac{p_n!}{(p_n - v)!} \frac{n!}{(n-k)!} \frac{1}{v!} \sum_{\pi \in \mathfrak{S}(\mu)} \frac{1}{p_n^k} \frac{p_n^k}{k!} \sum_{I \in \tilde{\mathrm{J}}_k} \mathbb{E}\mathrm{S}(I, \pi) + O\left(\frac{n^{k-1}}{p_n^{k-v}}\right).$$

Let $Z_1, \dots, Z_v$ be independent and $\mathcal{N}(\mathbf{0}, T(\nu))$ distributed random variables. Lemma 5.9.6 and Theorem 5.7.7 now lead to

$$\lim_{n \to \infty} \frac{p_n^v}{n^k} T(\mu) = \frac{1}{v!} \sum_{\pi \in \mathfrak{S}(\mu)} \mathbb{E}\pi(Z_1, \dots, Z_v) = \mathrm{M}_k(Z_1).$$

The required result then follows from the method of moments 5.6.4. □

Proposition 5.9.7. *Assume that $\nu \in \mathcal{M}^1(\Pi_q)$ has compact support and that n/p_n tends to some positive constant c as $n \to \infty$. Then $\mathfrak{A}_n$ and $\mathfrak{B}_n$ are asymptotically uncorrelated, that is, for all $0 \leq l \leq k$ and all $\sigma \in \mathfrak{S}((l, k-l))$ the random variable*

$$\frac{1}{n^{k/2}} \left[\mathbb{E}\sigma(\mathfrak{A}_n, \mathbf{1}) \circ \mathbb{E}\sigma(\mathbf{1}, \mathfrak{B}_n) - \mathbb{E}\sigma(\mathfrak{A}_n, \mathfrak{B}_n)\right] \tag{5.9.20}$$

tends to zero as $n \to \infty$.

Proof: Let $F(n, \sigma)$ denote the expression (5.9.20). According to the Theorem 5.1.1, there is no loss of generality in assuming that $\sigma = (1, \dots 1, 2, \dots, 2)$. From

Theorem 5.2.1, by using symmetry argument, we conclude that

$$\begin{aligned} F(n;\sigma) &= \frac{1}{n^{k/2}}\Big(\mathbb{E}(\mathfrak{A}^{\otimes,l}) \otimes \mathbb{E}(\mathfrak{B}^{\otimes,k-l}) - \mathbb{E}(\mathfrak{A}^{\otimes,l} \otimes \mathfrak{B}^{\otimes,k-l})\Big) \\ &= \frac{1}{n^{k/2}} \sum_{u=1}^{l} \sum_{v=1}^{k-l} \sum_{\lambda \in C(l,u)} \sum_{\mu \in C(k-l,v)} \binom{n}{u}\binom{p_n}{v} \sum_{\pi \in \mathfrak{S}(\lambda)} \sum_{\pi' \in \mathfrak{S}(\mu)} H(\pi,\pi'), \end{aligned}$$

with

$$H(\pi,\pi') = \mathbb{E}\pi(A_1,\ldots,A_u) \otimes \mathbb{E}\pi'(B_1,\ldots,B_v) - \mathbb{E}\pi(A_1,\ldots,A_u) \otimes \pi'(B_1,\ldots,B_v).$$

If $\mu_\alpha = 1$ for some $\alpha \in \{1,\ldots,v\}$, then each entry of $\pi'(B_1,\ldots,B_v)$ is not an even polynomial and thus neither is $\pi(A_1,\ldots,A_u) \otimes \pi'(B_1,\ldots,B_v)$. Therefore, $H(\pi,\pi') = \mathbf{0}$ by Corollary 5.8.8, (b).

Suppose that $\mu_\alpha \geq 2$ for each α. By Equation (5.9.17), we have

$$\begin{aligned} H(\pi,\pi') = \sum_{I \in \mathrm{J}_2 \cup \ldots \cup \mathrm{J}_k} \{\, & \mathbb{E}\pi(A_1,\ldots,A_u) \otimes \mathbb{E}S(I,\pi') + \\ & - \mathbb{E}(\pi(A_1,\ldots,A_u) \otimes S(I,\pi'))\}\,. \end{aligned} \tag{5.9.21}$$

For $M := \{1,\ldots,u\}$ and $G := \{\alpha \in M :\ \lambda_\alpha = 1\}$ we define

$$\begin{aligned} \mathrm{J}^{\exists}(M) &:= \{I \in \mathrm{J}_2 \cup \ldots \cup \mathrm{J}_{k-l} : \quad \{I\} \cap M \neq \emptyset\} \subset \mathcal{I}_{k-l,n}, \\ \mathrm{J}^{\forall}(G) &:= \{I \in \mathrm{J}_2 \cup \ldots \cup \mathrm{J}_{k-l} : \quad G \subset \{I\}\} \subset \mathcal{I}_{k-l,n}. \end{aligned}$$

It is easily checked that for the cardinalities of $\mathrm{J}^{\exists}(M)$ and $\mathrm{J}^{\forall}(G)$ we have

$$\left|\mathrm{J}^{\exists}(M)\right| \leq Cn^{k-l-1} \quad \text{and} \quad \left|\mathrm{J}^{\forall}(G)\right| \leq Cn^{k-l-|G|}, \tag{5.9.22}$$

with some constant $C = C(k,l)$. We consider the I-th term in the sum above, which will be denoted by $T(I)$. Is $I \notin \mathrm{J}^{\exists}(M)$, that is $\{I\} \cap M = \emptyset$, and thus $A_1,\ldots,A_u$ are independent from $\mathrm{S}(I,\pi')$. This clearly forces $T(I) = 0$.

Is $I \notin \mathrm{J}^{\forall}(G)$, i.e. there exists $\tau \in G$ with $\tau \notin \{I\}$, and therefore A_τ is independent from A_i $(i \in M \setminus \{\tau\})$ and $S(I,\pi')$. We thus get $T(I) = 0$ from (5.9.4).

Taking (5.9.22), the number of nonzero summands in (5.9.21) is bounded above $\min(n^{k-l-1}, n^{k-l-|G|})$. On the other hand, Lemma 5.9.5 yields that each of them is bounded above C/p_n^{k-l} where $C > 0$ is a suitable global constant. Summarized we get

$$\|H(\pi,\pi')\|_\infty \le C\cdot \min(n^{-1}, n^{-|G|}). \tag{5.9.23}$$

Since $\mu \in C(k-l, v)$ with $\mu_\alpha \ge 2$ for all $\alpha \in \{1,\cdots,v\}$ we have $k-l \ge 2v$. Moreover, since $\lambda \in C(l,u)$, we get $l \ge 2u - |G|$. And hence, by straightforward calculation and by using $n/p_n \to c \in]0,\infty[$, we conclude from (4.0.24) that for suitable constants C_i,

$$\begin{aligned}\|F_n(l,k)\|_\infty &\le \frac{C_1}{n^{k/2}} \sum_{u=1}^{l}\sum_{v=1}^{k-l}\sum_{\lambda\in C(l,u)}\sum_{\mu\in C(k-l,v)} \binom{n}{u}\binom{p_n}{v}\min(n^{-1}, n^{-|G|}) \\ &\le \frac{C_2}{n^{k/2}} \sum_{u=1}^{l}\sum_{v=1}^{k-l}\sum_{\lambda\in C(l,u)} n^{u+v}\min(n^{-1}, n^{-|G|}) \le \frac{C_3}{\sqrt{n}}.\end{aligned}$$

This completes the proof. □

proof of Theorem 5.9.1 for $\nu \in \mathcal{M}^1(\Pi_q)$ with compact support.
If $n/p_n \to \infty$ then $\sqrt{p_n}/n \cdot \mathfrak{A}_n \xrightarrow{d} 0$ and $\sqrt{p_n}/n \cdot \mathfrak{B}_n \xrightarrow{d} \mathcal{N}(\mathbf{0}, T(\nu))$ according to Propositions 5.9.2 and 5.9.3. This clearly forces $\sqrt{p_n}/n \cdot \Xi_n(\nu) \xrightarrow{d} \mathcal{N}(\mathbf{0}, T(\nu))$ by Slutsky's Theorem. Suppose that $n/p_n \to \infty$. Then we get as above $\Xi_n(\nu)/\sqrt{n} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma(\nu))$. It remains only to check the convergence in the case $n/p_n \to c \in]0,\infty[$. Let $k \in \mathbb{N}$. By Theorem 5.2.1,

$$\mathrm{M}_k(\Xi_n(\nu)) = \mathbb{E}\left((\mathfrak{A}_n + \mathfrak{B}_n)^{\otimes k}\right) = \sum_{l=0}^{k}\sum_{\pi\in\mathfrak{S}((l,k-l))} \mathbb{E}\pi(\mathfrak{A}_n, \mathfrak{B}_n).$$

Therefore, by Proposition 5.9.7,

$$\lim_{n\to\infty} \mathrm{M}_k\left(\frac{1}{\sqrt{n}}\Xi_n(\nu)\right) = \lim_{n\to\infty} \frac{1}{n^{k/2}} \sum_{l=0}^{k}\sum_{\pi\in\mathfrak{S}((l,k-l))} \mathbb{E}\pi(\mathfrak{A}_n, \mathbf{1}) \circ \mathbb{E}\pi(\mathbf{1}, \mathfrak{B}_n).$$

Consider independent random variables Z_1, Z_2 and Z with distributions $\mathcal{N}(\mathbf{0}, \Sigma(\nu))$, $\mathcal{N}(\mathbf{0}, cT(\nu))$ and $\mathcal{N}(\mathbf{0}, \Sigma(\nu) + cT(\nu))$, respectively. Propositions 5.9.2, 5.9.3 and Lemma 5.7.8 now lead to

$$\lim_{n\to\infty} \mathrm{M}_k(\Xi_n(\nu)) = \sum_{l=0}^{k}\sum_{\pi\in\mathfrak{S}((l,k-l))} \mathbb{E}\left(\pi(Z_1, \mathbf{1})\right) \circ \mathbb{E}\left(\pi(\mathbf{1}, Z_2)\right) = \mathrm{M}_k(Z).$$

□

To dispense with the assumption: $\operatorname{supp}(\nu)$ is compact, we introduce for an $a>0$ the truncated $\mathbb{M}_{p_n,q}$-valued random variables

$$\mathbb{X}_{k,a} := \begin{cases} \mathbb{X}_k, & \text{if } \|\varphi_{p_n}(\mathbb{X}_k)\| \le a, \\ \mathbf{0}, & \text{else} \end{cases} \qquad k = 1, 2, \ldots$$

Let us denote by ν_a the distribution of $\varphi_{p_n}(\mathbb{X}_{1,a})$ (which is not dependent on p_n). Obviously, the sequence $\mathbb{X}_{k,a}$, $k \in \mathbb{N}$, are i.i.d. with the radial law $\nu_{p_n,a} \in \mathcal{M}(\mathbb{M}_{p_n,q})$ which corresponds to ν_a. We define $\Xi_n(\nu_a)$, $\mathfrak{A}_n(\nu_a)$, $A_{j,a}$ $(j = 1, \ldots, n)$, $\mathfrak{B}_n(\nu_a)$ and $B_{j,a}$ $(j = 1, \ldots, p_n)$ according to (5.9.1), (5.9.2) and (5.9.3), respectively, by taking $\mathbb{X}_{k,a}$ instead of $\mathbb{X}_k$, $k \in \mathbb{N}$. Clearly, we have $\Xi_n(\nu_a) = \mathfrak{A}_n(\nu_a) + \mathfrak{B}_n(\nu_a)$.

Lemma 5.9.8. *For all $\varepsilon > 0$, $\delta > 0$ there exist a_0, $n_0 \in \mathbb{N}$ such that for all n, $a \in \mathbb{N}$ with $a \ge a_0$ and $n \ge n_0$*

$$\mathbb{P}\left(\|\mathfrak{A}_n(\nu) - \mathfrak{A}_n(\nu_a)\| > \delta_n\right) \le \varepsilon,$$

where $\delta_n = \delta\sqrt{n}$ if $n/p_n \to c \in [0, \infty[$ and $\delta_n = \delta n/\sqrt{p_n}$ if $n/p_n \to \infty$.

Proof: Let $\delta > 0$ and $(\delta_n)_n$ be a sequence with $\delta_n > 0$ for all $n \in \mathbb{N}$. Since $(A_i - A_{i,a})$, $(i = 1, 2, \ldots)$ are i.i.d., it follows by Chebychev inequality

$$\mathbb{P}\left(\|\mathfrak{A}_n(\nu) - \mathfrak{A}_n(\nu_a)\| \ge \delta_n\right) \le \frac{n}{\delta_n^2}\mathbb{E}\left(\|A_1 - A_{1,a}\|^2\right). \tag{5.9.24}$$

By using triangle inequality, we obtain

$$\sup_{a\in\mathbb{N}} \|A_{1,a}\|^2 \le \left(\left\|\varphi_{p_n}^2(X_1)\right\| + \|r_2(\nu)\|\right)^2 \in L^1(\Omega),$$

Therefore, the set $\{\|A_{1,a}\|^2 : \; a \in \mathbb{N}\}$ is uniformly integrable. On the other hand, since the random variable $\|A_1\|$ is almost surely finite, $\|A_{1,a}\|^2$ converges a.s. to $\|A_1\|^2$ as $a \to \infty$. Thus, we get

$$\|A_{1,a}\|^2 \longrightarrow \|A_1\|^2 \quad \text{in } L^1. \tag{5.9.25}$$

Taking (5.9.24) with δ_n, as in the assumption (i.e. $\delta_n = \delta\sqrt{n}$ or $\delta_n = \delta n/\sqrt{p_n}$) and (5.9.25) into account, completes the proof. $\square$

Lemma 5.9.9. *For all $\varepsilon > 0$, $\delta > 0$ there exist a_0, $n_0 \in \mathbb{N}$ such that for all*

$n, \ a \in \mathbb{N}$ *with* $a \geq a_0$ *and* $n \geq n_0$

$$\mathbb{P}\left(\|\mathfrak{B}_n(\nu) - \mathfrak{B}_n(\nu_a)\| > \delta_n\right) \leq \varepsilon, \tag{5.9.26}$$

where $\delta_n = \delta\sqrt{n}$ *if* $n/p_n \to c \in [0, \infty[$ *and* $\delta_n = \delta n/\sqrt{p_n}$ *if* $n/p_n \to \infty$.

Proof: Let $\delta > 0$ and $(\delta_n)_n$ be a sequence with $\delta_n > 0$ for all $n \in \mathbb{N}$. By Chebychev inequality follows

$$\mathbb{P}\left(\|\mathfrak{B}_n(\nu) - \mathfrak{B}_n(\nu_a)\| \geq \delta_n\right) \leq \frac{1}{\delta_n^2} \sum_{j,i=1}^{p_n} \mathbb{E}\left(\langle B_i - B_{i,a}, B_j - B_{j,a}\rangle\right). \tag{5.9.27}$$

Using Lemma 5.8.3, one can easily compute that

$$\mathbb{E}\left(\langle B_i, B_j\rangle\right) = \delta_{ij} \cdot \frac{n(n-1)}{p_n^2} \sum_{l,k=1}^{q} r_2(\nu)_{l,l} r_2(\nu)_{k,k} + r_2(\nu)_{l,k} r_2(\nu)_{l,k} \tag{5.9.28}$$

$$\mathbb{E}\left(\langle B_{i,a}, B_{j,a}\rangle\right) = \delta_{ij} \cdot \frac{n(n-1)}{p_n^2} \sum_{l,k=1}^{q} r_2(\nu_a)_{l,l} r_2(\nu_a)_{k,k} + r_2(\nu_a)_{l,k} r_2(\nu_a)_{l,k} \tag{5.9.29}$$

With the notation

$$\tilde{r}_2(a;n) := \left(\mathbb{E}\left(X_{1,a}^{(1,l)} X_1^{(1,k)}\right)\right)_{1 \leq l,k \leq q},$$

we see at once that

$$E\left(\langle B_i, B_{j,a}\rangle\right) = \delta_{ij} n(n-1) \sum_{l,k=1}^{q} \tilde{r}_2(a;n)_{l,l} \tilde{r}_2(a;n)_{k,k} + \tilde{r}_2(a;n)_{l,k} \tilde{r}_2(n;a)_{l,k}.$$

For $l, \ k \in \{1, \ldots, q\}$, we obtain

$$\tilde{r}_2(a;n)_{l,k} = \frac{1}{p_n} r_2(\nu)_{l,k} - \int_{\{\|X_1\| > a\}} X_1^{(1,l)} X_1^{(1,k)} \mathrm{d}\mathbb{P}. \tag{5.9.30}$$

By Cauchy-Schwarz inequality and straightforward calculation we get

$$0 \leq \left| \int_{\{\|X_1\| > a\}} X_1^{(1,l)} X_1^{(1,k)} \mathrm{d}\mathbb{P} \right| \leq \frac{c}{a p_n},$$

with some constant $c > 0$. From this and (5.9.30), we deduce

$$p_n \tilde{r}_2(a;n) = r_2(\nu) + O\left(\frac{1}{a}\mathbf{1}\right)$$

and hence

$$\forall\, \varepsilon > 0\ \exists\, M > 0\ \forall\, n \geq M,\ \forall\, a \geq M: \quad 0 \leq \frac{p_n^2}{n^2}\mathbb{E}\left(\|B_i - B_{i,a}\|\right) \leq \varepsilon.$$

Finally, this and (5.9.24) lead to the claim. □

Corollary 5.9.10. *For all $\varepsilon > 0$, $\delta > 0$ there exist $a_0,\ n_0 \in \mathbb{N}$ such that for all $n,\ a \in \mathbb{N}$ with $a \geq a_0$ and $n \geq n_0$*

$$\mathbb{P}\left(\|\Xi_n(\nu) - \Xi_n(\nu_a)\| > \delta_n\right) \leq \varepsilon, \tag{5.9.31}$$

where $\delta_n = \delta\sqrt{n}$ if $n/p_n \to c \in [0, \infty[$ and $\delta_n = \delta n/\sqrt{p_n}$ if $n/p_n \to \infty$.

Proof: For an $\delta > 0$ we observe that

$$\mathbb{P}\left(\|\Xi_n(\nu) - \Xi_n(\nu_a)\| > \delta_n\right) \leq \mathbb{P}\Big(\|\mathfrak{A}_n - \mathfrak{A}_{n,a}\| > \frac{\delta_n}{2}\Big) + \mathbb{P}\Big(\|\mathfrak{B}_n - \mathfrak{B}_{n,a}\| > \frac{\delta_n}{2}\Big).$$

Combining this with Lemmas 5.9.8 and 5.9.9, the corollary follows. □

proof of Theorem 5.9.1: Let us first prove the CLT I. In this case the normalisation is given by $\sqrt{p_n}/n$ and for the growth of p_n we have the condition $n/p_n \to \infty$ as $n \to \infty$. We set $\xi_n := \sqrt{p_n}/n \cdot \Xi_n(\nu)$ and $\xi_{n,a} = \sqrt{p_n}/n \cdot \Xi_n(\nu_a)$ and denote their distributions by μ_n and $\mu_{n,a}$, respectively. Moreover, we write τ_ν instead of $\mathcal{N}(\mathbf{0}, T(\nu))$. Using triangle inequality, we deduce that

$$\begin{aligned}\left|\int f \mathrm{d}\mu_n - \int f \mathrm{d}\tau_\nu\right| \leq &\left|\int f \mathrm{d}\mu_n - \int f \mathrm{d}\mu_{n,a}\right| + \\ &+ \left|\int f \mathrm{d}\mu_{n,a} - \int f \mathrm{d}\tau_{\nu_a}\right| + \left|\int f \mathrm{d}\tau_{\nu_a} - \int f \mathrm{d}\tau_\nu\right|.\end{aligned} \tag{5.9.32}$$

Let $\varepsilon > 0$, $f \in \mathcal{C}_b^u(\Pi_q)$ be a bounded uniformly continuous function on Π_q and $A_\delta := \{\|\xi_n - \xi_{n,a}\| \leq \delta\}$ $(\delta > 0)$. It follows that

$$\exists\, \delta > 0: \quad \int_{A_\delta} |f \circ \xi_n - f \circ \xi_{n,a}|\, \mathrm{d}\mathbb{P} \leq \varepsilon.$$

On the other hand, by Corollary 5.9.10,

$$\exists\, a_0,\ n_0 > 0: \quad \int_{\Omega \setminus A_\delta} |f \circ \xi_n - f \circ \xi_{n,a}|\, \mathrm{d}\mathbb{P} \leq 2\varepsilon\, \|f\|_\infty \quad \forall\, a \geq a_0,\ n \geq n_0.$$

This gives us the following estimation for the first summand in (5.9.32):

$$\exists\, a_0,\ n_0 > 0 : \left| \int f \mathrm{d}\mu_n - \int f \mathrm{d}\mu_{n,a} \right| \leq \varepsilon(1+2\,\|f\|_\infty) \quad \forall\, a \geq a_0,\ n \geq n_0. \tag{5.9.33}$$

Since ν_a has a compact support, we conclude from 5.9.1 that $\mu_{n,a}$ converges weakly to τ_{ν_a} $(a > 0)$, hence

$$\forall\, a > 0\ \exists\, n_0 > 0 : \left| \int f \mathrm{d}\mu_{n,a} - \int f \mathrm{d}\tau_{\nu_a} \right| \leq \varepsilon \quad \forall\, n \geq n_0. \tag{5.9.34}$$

Finally, it is evident that

$$\exists\, a_0 > 0 : \left| \int f \mathrm{d}\tau_{\nu_a} - \int f \mathrm{d}\tau_\nu \right| \leq \varepsilon \quad \forall\, a \geq a_0. \tag{5.9.35}$$

Taking (5.9.33), (5.9.34) and (5.9.35) into account, we obtain

$$\exists\, n_0 > 0 : \left| \int f \mathrm{d}\mu_n - \int f \mathrm{d}\tau_\nu \right| \leq \varepsilon(3 + 2\,\|f\|_\infty) \quad \forall\, n \geq n_0,$$

which completes the proof of CLT I in Theorem 5.9.1. The same proof works for CLT II. □

For the rest of this chapter, we devote ourselves to the laws of large numbers for the functionals $\Xi_n(\nu)$. Our first result in this direction is the following weak law of large numbers:

Theorem 5.9.11. *Assume that $\nu \in \mathcal{M}^1(\Pi_q)$ admits forth moment. Then the random variables $n^{-1}\Xi_n(\nu)$ converge in probability to $\mathbf{0} \in \mathbb{M}_{q,q}$.*

Proof: For any $\varepsilon > 0$, Cauchy-Schwarz inequality and Equation (5.5.2) yield

$$\begin{aligned}\mathbb{P}\left(\|\Xi_n(\nu)\| \geq \varepsilon \cdot n\right) &\leq \frac{1}{\varepsilon^2 \cdot n^2} \mathbb{E}\left\langle \Xi_n(\nu), \Xi_n(\nu) \right\rangle = \frac{1}{\varepsilon^2 \cdot n^2} \mathrm{tr}\Big(\mathbb{C}ov(\Xi_n(\nu)\Big) \\ &\leq \frac{1}{\varepsilon^2}\Big(\frac{1}{n}\mathrm{tr}(\Sigma(\nu)) + \frac{1}{p_n}\mathrm{tr}(T(\nu))\Big).\end{aligned}$$

and so the theorem follows. □

Theorem 5.9.12. *Assume that $\nu \in \mathcal{M}^1(\Pi_q)$ admits eights moment and let $(a_n)_{n\geq 1}$ be a sequence in $]0, \infty[$ with $a_n = n$ if $n/p_n \to c \in [0, \infty[$ and $a_n = n^2/p_n$ if $n/p_n \to \infty$. Then $a_n^{-1}\Xi_n(\nu) \to \mathbf{0}$ almost surely as $n \to \infty$.*

Proof: For any $\varepsilon > 0$, by Markov inequality,

$$\mathbb{P}\left(\|\Xi_n(\nu)\| \geq \varepsilon \cdot a_n\right) \leq \mathbb{P}\Big(\|\mathfrak{A}_n\| \geq \frac{\varepsilon \cdot a_n}{2}\Big) + \mathbb{P}\Big(\|\mathfrak{B}_n\| \geq \frac{\varepsilon \cdot a_n}{2}\Big)$$
$$\leq \Big(\frac{2}{\varepsilon a_n}\Big)^4 \mathbb{E}\Big(\|\mathfrak{A}_n(\nu)\|^4 + \|\mathfrak{B}_n(\nu)\|^4\Big). \tag{5.9.36}$$

In a similar way as in the proof of Theorem 4.0.22, it is easy to check that

$$\mathbb{E}\Big(\|\mathfrak{A}_n(\nu)\|^4 + \|\mathfrak{B}_n(\nu)\|^4\Big) \leq C \cdot |\nu|_8 \cdot (n^2 + n^4/p_n^2) \tag{5.9.37}$$

for some constant $C > 0$. Combining (5.9.36) with (5.9.37) we get

$$P\left(\|\Xi_n(\nu)\| \geq \varepsilon \cdot a_n\right) \leq C\varepsilon^{-4} a_n^{-4}\Big(n^2 + n^4/p_n^2\Big) = O(n^{-2}).$$

By using the Borel-Cantelli lemma, $P(|\Xi_n(\nu)| \geq \varepsilon \cdot a_n \text{ i.o.}) = 0$ for each positive ε. This completes the proof. □

Bibliography

[1] T. W. Anderson (1984), *An Introduction to Multivariate Statistical Analysis*, 2nd edition John Wiley & Sons, New York.

[2] V. Bentkus (1986), Dependence of the Berry-Esseen estimate on the dimension. *Lithuanian Math. J.* **26**, 110-113.

[3] V. Bentkus, F. Götze (1997), Uniform rates of convergence in the CLT for quadratic forms. *Probab. Theory Related Fields* **109**, 367-416.

[4] P. Billingsley (1986), *Probability and Measure*, 2nd edition John Wiley & Sons, New York.

[5] N. H. Bingham (1972), Random walks on spheres. *Z. Wahrsch. verw. Geb.* **22**, 169-192 .

[6] W.R. Bloom, H. Heyer (1995), *Harmonic analysis of probability measures on hypergroups*, De Gruyter Studies in Mathematics 20, de Gruyter-Verlag Berlin, New York.

[7] H. Chébli (1974), *Positivité des opérateurs de "translation généralisées" associées à un opérateur de Sturm-Liouville et quelques applications à l'analyse harmonique.* Thèse de Doctorat d'Etat, Université Louis Pasteur, Strasbourg.

[8] Y.S. Chow, H. Teicher (1997), *Probability theory; Independence, Interchangeability, Martingales*, 3rd edition Springer Texts in Statistics, New York.

[9] J. Faraut (1982), *Analyse harmonique sur les pairs de Gelfand et les espaces hyperboliques.* Analyse Harmonique, Chapitre IV, C.I.M.P.A., Nice.

[10] U. Finckh (1986), *Beiträge zur Wahrscheinlichkeitstheorie auf einer Kingman-Struktur.* Dissertation, Tübingen.

[11] J. Faraut, A. Korányi (1994), *Analysis on symmetric cones.* Oxford Science Publications, Clarendon press, Oxford.

[12] W. Grundmann (2012), Method of moments and limit theorems on $p \times q$ matrices for $p \to \infty$. Preprint 2012, arXiv:1210.7090v1.

[13] W. Grundmann (2012), Moment Functions and Central Limit Theorem for Jacobi Hypergroups on $[0, \infty[$. *J. Theor. Probab.,* (forthcoming).

[14] A.K. Gupta, D.K.Nagar (1994), *Matrix variate distributions.* Oxford Science Publications, Clarendon press, Oxford.

[15] J.B.S Haldane (1960), The addition of random vectors. *Indian J. Stat.* **22**, 213-220

[16] C.S. Herz (1955), Bessel functions of matrix argument. *Ann. Math.* **61**, 474-523.

[17] S. Helgason (2000), Groups and Geometric Analysis. *Mathematical Surveys and Monographs* **83**, AMS 2000.

[18] R.A. Horn, C.R. Johnson (1991), *Topics in matrix analysis.* Cambridge University Press, Cambridge.

[19] R.I. Jewett (1975), Spaces with an abstract convolution of measures. *Adv. Math.* **18**, 1-101.

[20] Huang Jizheng, LIU Heping, Liu Jianming (2009), An Analogue of Beurling's theorem for the Jacobi Transform, *Acta Mathematica Sinica* **25**, 85-94.

[21] O. Kallenberg (2001), *Foundations of modern probability*, Probability and its Applications, Springer, New York.

[22] F.I. Karpelevich, V.N. Tutubalin, M.G. Shur (1959), Limit theorems for the compositions of distributions in the Lobachevski plane and space. *Theory Probab. Appl.* **4**, 399-402.

[23] J.F.C. Kingman (1963), Random walks with spherical symmetry. *Acta Math.* **109**, 11-53.

[24] C. Kleiber, J. Stoyanov (2011), Multivariate distributions and the moment problem. *Journal of Multivariate Analysis*, In Press.

[25] T. Koornwinder (1984), *Jacobi functions and analysis on noncompact semisimple Lie groups.* Special Functions: group theoretical aspects and applications, Eds. Richard Askey et al., D. Reidel, Dordrecht-Boston-Lancaster, 1-85.

[26] N.N. Lebedev (1973), *Spezielle Funktionen und ihre Anwendung.* Bibliographisches Institut AG, Zürich.

[27] E.A. Michael (1976), Topologies on spaces of subsets, *Trans. Amer. Math. Soc.* **71**, 152-182.

[28] C. Rentzsch, M. Voit (2000), Lévy Processes on Commutative Hypergroups, *Contemporary Mathematics* **261**, 83-105.

[29] M. Rösler (2007), Bessel convolutions on matrix cones, *Compos. Math.* **143**, 749-779.

[30] M. Rösler, M. Voit (2011), Limit theorems for radial random walks on $p \times q$ matrices as p tends to infinity. *Math. Nachr.* **284**, 87–104.

[31] M. Rösler, M. Voit (2008), Dunkl theory, convolution algebras, and related Markov processes. In: P. Graczyk et al., Harmonic and Stochastic Analysis of Dunkl Processes, Hermann, Paris, 1–112.

[32] R.P. Stanley (2005), *Enumerative Combinatorics*, Cambridge studies in advanced mathematics 49, Cambridge University Press, Cambridge.

[33] R.P. Stanley, (1989) Some combinatorial properties of Jack symmetric functions. *Adv. Math.* **77**, 76-115.

[34] A. Terras, (1985) *Harmonic Analysis on Symmetric Spaces and Applications*, Springer-Verlag.

[35] K. Triantafyllopoulos (2003), On the central moments of the multidimensional Gaussian distribution, *The Mahtematical Scientist* **28**, 125-128.

[36] K. Trimèche (1978), Probabilités indéfiniment divisibles et théorème de la limite centrale pour une convolution généralisée sur la demi-droite, (English summary). C.R. Acad. Sci. Paris Sér. A-B **286**, no. 1, 63-66.

[37] V.N. Tutubalin (1962), On the limit behaviour of compositions of measures in the plane and space of Lobachevski. *Theory Probab. Appl.* **7**, 189-196.

[38] M. Voit (1990), Laws of large numbers for polynomial hypergroups and some applications. *J. Theor. Probab.* **3**, no. 2, 245-266.

[39] M. Voit (1995), Central Limit Theorems for Jacobi Hypergroups. *Contemporary Mathematics* **183**, 395-411.

[40] M. Voit (1995), A limit theorem for isotropic random walks on $\mathbb{R}^d$ for $d \to \infty$. *Russian J. Math. Phys.* **3**, 535-539.

[41] M. Voit (2012), Central limit theorem for radial random walks on $p \times q$ matrices for $p \to \infty$. Adv. Pure Appl. Math. **3**, 231-246.

[42] M. Voit (2009), Limit theorems for radial random walks on homogeneous spaces with growing dimensions. J. Hilgert, Joachim (ed.) et al., Proc. symp. on infinite dimensional harmonic analysis IV. On the interplay between representation theory, random matrices, special functions, and probability, Tokyo, World Scientific. 308-326.

[43] M. Voit (2012), Central limit theorems for hyperbolic spaces and Jacobi processes on $[0, \infty[$. Preprint 2012, arXiv:1201.3490.

[44] H. Zeuner (1986), *On hyperbolic hypergroups.* Probability measures on group VIII, H. Heyer (ed.). Lecture Notes Math. **1210**, Springer, Berlin Heidelberg New York. 216-224.

[45] H. Zeuner (1989), The central limit theorem for Chébli-Trimèche hypergroups. *J. Theor. Probab.* **2**, no.1, 51-63.

[46] H. Zeuner (1989), Laws of Large Numbers of Hypergroups on $\mathbb{R}_+$. *Math. Ann.* **283**, 657-678.

[47] H. Zeuner (1992), Moment functions and laws of large numbers on hypergroups. *Math. Z.* **211**, 369-407.