n-Kernel Orthogonal Polynomials on the Dirichlet, Dirichlet-Multinomial, Poisson-Dirichlet and Ewens' sampling distributions, and positive-definite sequences.

Robert C. Griffiths

Department of Statistics
1 South Parks Road Oxford OX1 3TG, UK
e-mail: griff@stats.ox.ac.uk

Dario Spanò

Department of Statistics
University of Warwick, Coventry CV4 7AL, UK
e-mail: d.spano@warwick.ac.uk

Abstract: We consider a multivariate version of the so-called Lancaster problem of characterizing canonical correlation coefficients of symmetric bivariate distributions with identical marginals and orthogonal polynomial expansions. The marginal distributions examined in this paper are the Dirichlet and the Dirichlet-Multinomial distribution, respectively on the continuous and the N-discrete d-dimensional simplex. Their infinite-dimensional limit distributions, respectively the Poisson-Dirichlet distribution and the Ewens' sampling formula, are considered as well. We study in particular the possibility of mapping canonical correlations on the d-dimensional continuous simplex (i) to canonical correlation sequences on the $d+1$-dimensional simplex and/or (ii) to canonical correlations on the discrete simplex, and viceversa. Driven by this motivation, the first half of the paper is devoted to providing a full characterization and probabilistic interpretation of $|n|$-orthogonal polynomial kernels (i.e. sums of products of orthogonal polynomials of the same degree $|n|$) with respect to the mentioned marginal distributions. Orthogonal polynomial kernels are important to overcome some non-uniqueness difficulties arising when dealing with multivariate orthogonal (or bi-orthogonal) polynomials. We establish several identities and some integral representations which are multivariate extensions of important results known for the case $d=2$ since the 1970's. These results, along with a common interpretation of the mentioned kernels in terms of dependent Polya urns, are shown to be key features leading to several non-trivial solutions to Lancaster's problem, many of which can be extended naturally to the limit as $d \to \infty$.

AMS 2000 subject classifications: 33C50, 60E05, 60G07.
Keywords and phrases: Multivariate orthogonal polynomials, Orthogonal polynomial kernels, Jacobi, Hahn, Dirichlet distribution, Dirichlet-Multinomial, Poisson-Dirichlet, Ewens' sampling formula, Canonical correlations, Positive-definite sequences, Lancaster's problem

.

1. Introduction.

Let π be a probability measure on some Borel space $(E, \mathcal{E})$ with $E \subseteq \mathbb{R}$. Consider an exchangeable pair (X, Y) of random variables with given marginal law π. Modeling tractable joint distributions for (X, Y) with π as given marginals is a classical problem in Mathematical Statistics. One possible approach, introduced by Oliver Lancaster [29] is in terms of so-called *canonical correlations*. Let $\{P_n\}_{n=0}^{\infty}$ be a family of orthogonal polynomials with weight measure π i.e. such that

$$\mathbb{E}_\pi \left(P_n(X)P_m(X)\right) = \frac{1}{c_m}\delta_{nm}, \qquad n, m \in \mathbb{Z}_+$$

for a sequence of positive constants $\{c_m\}$. Here $\delta_{mn} = 1$ if $n = m$ and 0 otherwise, and $\mathbb{E}_\pi$ denotes the expectation taken with respect to π.
A sequence $\rho = \{\rho_n\}$ is the sequence of canonical correlation coefficients for the pair (X, Y) if it is possible

to write the joint law of (X, Y) as

$$g_\rho(dx, dy) = \pi(dx)\pi(dy) \left\{ \sum_{n \in \mathbb{Z}_+}^{\infty} \rho_n c_n P_n(x) P_n(y) \right\}. \tag{1.1}$$

Suppose that the system $\{P_n\}$ is *complete* with respect to $L_2(\pi)$, that is, every function f with finite π-variance admits a representation

$$f(x) = \sum_{n=0}^{\infty} \widehat{f}(n) c_n P_n(x) \tag{1.2}$$

where

$$\widehat{f}(n) = \mathbb{E}_\pi \left[f(X) P_n(X) \right], \qquad n = 0, 1, 2, \ldots \tag{1.3}$$

Define the regression operator by

$$Tf(x) := \mathbb{E}(f(Y)|X = x).$$

If (X, Y) have canonical correlations $\{\rho_n\}$ then, for every f with finite variance,

$$T_\rho f(x) = \sum_{n=0}^{\infty} \rho_n \widehat{f}(n) c_n P_n(x).$$

In other words,

$$\widehat{T_\rho f}(n) = \rho_n \widehat{f}(n), \quad n = 0, 1, 2, \ldots.$$

In particular,

$$T_\rho P_n = \rho_n P_n, \qquad n = 0, 1, \ldots$$

and

$$\widehat{T_\rho P_m}(n) = \delta_{mn} \rho_m P_m, \qquad m, n = 0, 1, \ldots$$

which means that the polynomials $\{P_n\}$ are the eigenfunctions and ρ is the sequence of eigenvalues of T. Lancaster's problem is therefore a spectral problem whereby regression operators with given eigenfunctions are uniquely characterized by their eigenvalues. Because T_ρ maps positive functions to positive functions, the problem of identifying canonical correlation sequences ρ is strictly related to the problem of characterizing so-called *positive-definite sequences*.

In this paper we consider a multivariate version of Lancaster's problem, when π is taken to be the either the Dirichlet or the Dirichlet-Multinomial distribution (notation: D_α and $DM_{\alpha,N}$, with $\alpha \in \mathbb{R}_+^d$ and $N \in \mathbb{Z}_+$) on the $(d-1)$-dimensional continuous and N-discrete simplex, respectively:

$$\Delta_{(d-1)} := \{x \in [0,1]^d : |x| = 1\}$$

and

$$N\Delta_{(d-1)} := \{m \in \mathbb{Z}_+^d : |m| = N\}.$$

The eigenfunctions will be therefore represented by multivariate Jacobi or Hahn polynomials, respectively. One difficulty arising when $d > 2$ is that the orthogonal polynomials $P_n = P_{n_1 n_2 \ldots n_d}$ are multi-indexed. The degree of every polynomial P_n is $|n| := n_1 + \cdot + n_d$. For every integer $|n|$ there are

$$\binom{|n| + d - 1}{d - 1}$$

polynomials with degree $|n|$, so when $d > 2$ there is not an unique way to introduce a total order in the space of all polynomials. One way to overcome such a difficulty is by working with *orthogonal polynomial kernels*. By $|n|$-orthogonal polynomial kernels with respect to π we mean functions of the form

$$P_{|n|}(x, y) = \sum_{m \in \mathbb{Z}_+^d : |m| = |n|} c_m P_m(x) P_m(y), \qquad |n| = 0, 1, 2, \ldots \tag{1.4}$$

where $|m| = m_1 + \cdots + m_d$ for every $m \in \mathbb{Z}_+^d$. Polynomial kernels are uniquely defined and totally ordered. A representation equivalent to (1.2) in term of polynomial kernels is:

$$f(x) = \sum_{|n|=0}^{\infty} \mathbb{E}_\pi(f(Y)P_{|n|}(x,Y)). \tag{1.5}$$

If f is a polynomial of order $|m|$ the series terminates at $|m|$. Consequently, for general $d \geq 2$, the individual orthogonal polynomials $P_n(x)$ are uniquely determined by their leading coefficients of degree $|n|$ and $P_{|n|}(x,y)$. If a leading term is

$$\sum_{\{k:|k|=|n|\}} b_{nk} \prod_1^d x_i^{k_i}$$

then

$$P_n(x) = \sum_{\{k:|k|=|n|\}} b_{nk} \mathbb{E}\Big[\prod_1^d Y_i^{k_i} P_{|n|}(x,Y)\Big], \tag{1.6}$$

where Y has distribution π.
It is easy to check that

$$\mathbb{E}_\pi\left[P_{|n|}(x,Y)P_{|m|}(z,Y)\right] = P_{|n|}(x,z)\delta_{|m||n|}.$$

$P_{|n|}(x,y)$ also has an expansion in terms of any complete sets of biorthogonal polynomials of degree $|n|$. That is, if $\{P_n^\diamond(x)\}$ and $\{P_n^\circ(x)\}$ are polynomials orthogonal to polynomials of degree less that $|n|$ and

$$\mathbb{E}\Big[P_n^\diamond(X)P_{n'}^\circ(X)\Big] = \delta_{nn'},$$

then

$$P_{|n|}(x,y) = \sum_{\{n:|n| \text{ fixed}\}} P_n^\diamond(x)P_n^\circ(y). \tag{1.7}$$

Biorthogonal polynomials always have expansions

$$\begin{aligned} P_n^\diamond(x) &= \sum_{\{m:|m|=|n|\}} c_{nm}^\diamond P_m(x) \\ P_n^\circ(x) &= \sum_{\{m:|m|=|n|\}} c_{nm}^\circ P_m(x), \end{aligned} \tag{1.8}$$

where the matrices of coefficients satisfy

$$C^{\diamond T} C^\circ = I, \text{ equivalent to } C^\circ = C^{\diamond -1^T}.$$

Similar expressions to (1.6) hold for $P_n^\diamond(x)$ and $P_n^\circ(x)$ using their respective leading coefficients. This can be shown by using their expansions in an orthonormal polynomial set and applying (1.6).
The polynomial kernels with respect to D_α and $DM_{\alpha,N}$ will be denoted by $Q_{|n|}^\alpha(x,y)$ and $H_{|n|}^\alpha(r,s)$, and called Jacobi and Hahn kernels, respectively.

This paper is divided in two parts. The goal of the first part is to describe Jacobi and Hahn kernels under a unified view: we will first provide a probabilistic description of their structure and mutual relationship, then we will investigate their symmetrized and infinite-dimensional versions.
In the second part of the paper we will turn our attention to the problem of identifying canonical correlation sequences with respect to D_α and DM_α. We will restrict our focus on sequences ρ such that, for every $n \in \mathbb{Z}_+^d$,

$$\rho_n = \rho_{|n|}.$$

For these sequences, Jacobi or Hahn polynomial kernels will be used to find out conditions for a sequence $\{\rho_{|n|}\}$ to satisfy the inequality

$$\sum_{|n|=0}^{\infty} \rho_{|m|} P_{|m|}(u, v) \geq 0. \tag{1.9}$$

Since T_ρ is required to map constant functions to constant functions, a straightforward necessary condition is always that

$$\rho_0 = 1.$$

For every $d = 2, 3, \ldots$ and every $\alpha \in \mathbb{R}_+^d$ we will call any solution to (1.9) an *α-Jacobi positive definite sequence (α-JPDS)* if $\pi = D_\alpha$ and an *α-Hahn postive-definite sequence (α-HPDS)* if $\pi = DM_\alpha$.
We are interested, in particular, in studying if and when one or both the following statements are true.

(P1) For every d and $\alpha \in \mathbb{R}_+^d$ ρ is α-JPDS $\Leftrightarrow$ ρ is $\widetilde{\alpha}$-JPDS for every $\widetilde{\alpha} \in \mathbb{R}_+^{d+1} : |\widetilde{\alpha}| = |\alpha|$;
(P2) For every d and $\alpha \in \mathbb{R}_+^d$ ρ is α-JPDS $\Leftrightarrow$ ρ is α-HPDS.

Regarding (P1), it will be clear in Section 7 that the sufficiency part ($\Leftarrow$) always holds. To find conditions for the necessity part ($\Rightarrow$) of (P1), we will use two alternative approaches. The first one is based on a multivariate extension of a powerful product formula for the Jacobi polynomials, due to Koornwinder and finalized by Gasper in the early 1970's: for α, β in a "certain region"(see Theorem 3.51 further on), the integral representation

$$\frac{P_n^{\alpha,\beta}(x)}{P_n^{\alpha,\beta}(1)} \frac{P_n^{\alpha,\beta}(y)}{P_n^{\alpha,\beta}(1)} = \int_0^1 \frac{P_n^{\alpha,\beta}(z)}{P_n^{\alpha,\beta}(1)} m_{x,y}(dz), \qquad x, y \in (0,1), n \in \mathbb{N}$$

holds for a probability measure $m_{x,y}$ on [0,1]. Our extension for multivariate polynomial kernels, of non-easy derivation, is found in Proposition 5.4 to be

$$Q_{|n|}^{\alpha}(x, y) = \mathbb{E}\left[Q_{|n|}^{\alpha_d, |\alpha| - \alpha_d}(Z_d, 1)\right], \quad |n| = 0, 1, \ldots \tag{1.10}$$

for every d and $\alpha \in \mathbb{R}_+^d$ in a "certain region,"and for a particular $[0,1]$-valued random variable Z_d. Here, for every $j = 1, \ldots, d$, $e_j = (0, 0, \ldots, 1, 0, \ldots, 0) \in \mathbb{R}^d$ is the vector with all zero components except for the j-th coordinate which is equal to 1. Integral representations such as (1.10) are useful in that they map immediately univariate positive functions to the type of bivariate distribution we are looking for:

$$f(x) \geq 0 \Longrightarrow \mathbb{E}\left[f(Z_d)\right] = \sum_{|n|} \widehat{f}(n) Q_{|n|}(x, y) \geq 0.$$

In fact, whenever (1.10) holds true, we will be able to conclude that (P1) is true.
The identity (1.10) holds only with particular choices of the parameter α. At the best one needs one of the α_j's to be greater than 2. This makes it hard to use (P1) to build canonical correlations with respect to Poisson-Dirichlet limit marginals on the infinite simplex. The latter would be a desirable aspect for modeling dependent measures on the infinite symmetric group or for applications e.g. in Nonparametric Bayesian Statistics.
On the other hand, there are several examples in the literature of positive-definite sequences satisfying (P1) for every choice of α, even in the limit case of $|\alpha| = 0$. Two notable and well-known instances are

(i)
$$\rho_{|n|}(t) = e^{-\frac{1}{2}|n|(|n|+|\alpha|-1)t}, \qquad |n| = 0, 1, \ldots$$
arising as the eigenvalues of the transition semigroup of the so-called *d-type, neutral Wright-Fisher diffusion process in Population Genetics.*

(ii)
$$\rho_{|n|}(z) = z^{|n|}, \qquad |n| = 0, 1, \ldots$$
i.e. the eigenvalues of the so-called Poisson kernel, whose positivity is a well-known result in Special Functions Theory (see e.g. [18],[7]).

A probabilistic account of the relationship existing between examples (i) and (ii) is given in [12].
It is therefore natural to ask when (P1) holds with no constraints on the parameter α.
Our second approach to Lancaster's problem will answer in part to this question. This approach is heavily based on the probabilistic interpretation of the Jacobi and Hahn polynomial kernels shown in the first part of the paper. We will prove in Proposition 8.1 that, if $\{d_{|m|} : m = 0, 1, 2, \ldots\}$ is a probability mass function (pmf) on $\mathbb{Z}_+$, then every positive-definite sequence $\{\rho_{|n|}\}_{|n|=0}^{\infty}$ of the form

$$\rho_{|n|} = \sum_{|m|=|n|}^{\infty} \frac{|m|!\Gamma(|\alpha| + |m|)}{(|m| - |n|)!\Gamma(|\alpha| + |m| + |n|)} d_{|m|}, \qquad |m| = 0, 1, \ldots \tag{1.11}$$

satisfies (P1) for every choice of α, therefore (P1) can be used to model canonical correlations with respect to the Poisson-Dirichlet distribution.
In Section 9 we investigate the possibility of a converse result, i.e. will find a set of conditions on a JPD sequence ρ to be of the form (1.11) for a pmf $\{d_{|m|}\}$.

As for Hahn positive-definite sequences and (P2), our results will be mostly consequence of Proposition 3.1, where we establish the following representation of Hahn kernels as mixtures of Jacobi kernels:

$$H^{\alpha}_{|n|}(r, s) = \frac{(|N| - |n|)!\Gamma(|\alpha| + |N| + |n|)}{|N|!\Gamma(|\alpha| + |N|)} \mathbb{E}\left[Q^{\alpha}_{|n|}(X, Y) \mid r, s\right] \quad |n| = 0, 1, \ldots$$

for every $N \in \mathbb{Z}_+$ and $r, s \in N\Delta_{(d-1)}$, where the expectation on the right-hand side is taken with respect to $D_{\alpha+r} \otimes D_{\alpha+s}$ i.e. a product of *posterior* Dirichlet probability measures. A similar result was proven by [17] to hold for individual Hahn polynomials as well.
We will also show (Proposition 6.1) that a discrete version of (1.10) (but with the appearance of an extra coefficient) holds for Hahn polynomial kernels.
Based on these findings, we will be able to prove in Section 7.2 some results "close to"(P2): we will show that JPDSs can be viewed as a map from HPDSs and also the other way around, but such mappings are not in general the identity (i.e. (P2)).
On the other way, we will show (Proposition 7.8) that every JPDS is indeed the limit of a sequence of (P2)-positive-definite sequences.
Our final result on HPDSs is in Proposition 8.6, where we prove that if, for every N, a JPDS ρ is of the form 1.11, for a probability distribution $d^{(N)} = \{d^{(N)}_{|m|}\}_{|m| \in \mathbb{Z}_+}$ such that $d^{(N)}_l = 0$ for $l > N$, then (P2) holds properly. Such sequences also satisfy (P1) and admit infinite-dimensional Poisson-Dirichlet (and Ewens' sampling distribution) limits.
The key for the proof of Proposition 8.6 is provided by Proposition 3.5, where we show the connection between our representation of Hahn kernels and a kernel generalization of a product formula for Hahn polynomials, proved by Gasper [11] in 1973. Proposition 3.5 is, in our opinion, of some interest even independently of its application.

1.1. Outline of the paper.

The paper is organized as follows. Section 1.2 will conclude this Introduction by recalling some basic properties and definitions of the probability distribution we are going to deal with. In Section 2 an explicit description of $Q^{\alpha}_{|n|}$ is given in terms of mixtures of products of Multinomial probability distributions. We will next obtain (Section 3) an explicit representation for $H^{\alpha}_{|n|}$ as *posterior mixtures* of $Q^{\alpha}_{|n|}$. This is done by applying an analogous relationship which was proved in [17] to hold for individual orthogonal polynomials. In the same section we will generalize Gasper's product formula to an alternative representation of $H^{\alpha}_{|n|}$ and will describe the connection coefficients in the two representations. In Sections 4-4.2 we will then show that similar structure and probabilistic descriptions also hold for kernels with respect to the ranked versions of D_α and DM_α, and to their infinite-dimensional limits, known as the Poisson-Dirichlet and the Ewens'

sampling distribution, respectively. As an immediate application, symmetrized Jacobi kernels will be used in section 4.1 to characterize *individual* (i.e. not kernels) orthogonal polynomials with respect to the *ranked* Dirichlet distribution. This will conclude the first part.
Sections 5-6 will be the bridge between the first and the second part of the paper. We will prove the identity (1.10) for the Jacobi product formula and its Hahn equivalent. We will point out the connection between (1.10) and another multivariate Jacobi product formula due to Koornwinder and Schwartz [27]. In section 7 we will use the results of Section 5 to characterize sequences obeying to (P1), with constraints on α, and will investigate the existence of sequences satisfying (P2).
In section 8 we will find sufficient conditions for (P1) to hold with no constraints on the parameters, when a JPDS can be expressed as a linear functional of a probability distribution on $\mathbb{Z}_+$. We will discuss the possibility of a converse mapping from JPDSs to probability mass functions in Section 9.
Finally, in Section 8.2 we will make a similar use of probability mass functions to find sufficient conditions for a proper version of (P2).

Acknowledgements.

Part of the material included in this paper (especially the first part) has been informally circulating for quite a while, in form of notes, among other Authors. Some of them have also used it for several interesting applications in Statistics and Probability (see [33], [21]). Here we wish to thank them for their helpful comments.

1.2. Elements from Distribution Theory.

We briefly list the main definitions and properties of the probability distributions that will be used in the paper. We also refer to [17] for further properties and related distributions. For $\alpha, n \in \mathbb{R}^d$ denote

$$x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}, \qquad \Gamma(\alpha) = \prod_{i=1}^d \Gamma(\alpha_i)$$

and

$$\binom{|n|}{n} = \frac{|n|!}{\prod_{i=1}^d n_i!}.$$

Also, we will use

$$(a)_{(x)} = \tfrac{\Gamma(a+x)}{\Gamma(a)}$$

$$(a)_{[x]} = \tfrac{\Gamma(a+\underline{1})}{\Gamma(a+\underline{1}-x)},$$

whenever the ratios are well defined. Here $\underline{1} := (1, 1, \ldots, 1)$.

If $x \in \mathbb{Z}_+$ then $(a)_{(x)} = a(a+1)\cdots(a+x-1)$ and $(a)_{[x]} = a(a-1)\cdots(a-x+1)$. $\mathbb{E}_\mu$ will denote the expectation under the probability distribution μ. The subscript will be omitted when there is no risk of confusion.

Definition 1.1.

(i) Dirichlet (α) *distribution,* $\alpha \in \mathbb{R}_+^d$*:*

$$D_\alpha(dx) := \frac{\Gamma(|\alpha|)x^{\alpha-\underline{1}}}{\Gamma(\alpha)}\mathbb{I}(x \in \Delta_{(d-1)})dx.$$

(ii) Dirichlet-Multinomial (α, N) *distribution,* $\alpha \in \mathbb{R}_+^d, N \in \mathbb{Z}_+$:

$$\begin{aligned} DM_\alpha(r;N) &= \mathbb{E}_{D_\alpha}\left[\binom{|r|}{r} X^r\right] \\ &= \binom{|r|}{r}\frac{(\alpha)_{(r)}}{(|\alpha|)_{(N)}}, \qquad r \in N\Delta_{(d-1)}. \end{aligned} \tag{1.12}$$

Define the *ranking function* $\psi : \mathbb{R}^d \to \mathbb{R}^d$ as the function reordering the elements of any vector $y \in \mathbb{R}^d$ in decreasing order. Denote its image by

$$\psi(y) = y^\downarrow = (y_1^\downarrow, \dots, y_d^\downarrow).$$

The ranked continuous and discrete simplex will be denoted by $\Delta_{d-1}^\downarrow = \psi(\Delta_{d-1})$ and $N\Delta_{d-1}^\downarrow = \psi(N\Delta_{d-1})$, respectively.

Definition 1.2. *The* Ranked Dirichlet *distribution with parameter* $\alpha \in \mathbb{R}_+^d$, *is*

$$D_\alpha^\downarrow(x) := D_\alpha \circ \psi^{-1}(x^\downarrow) = \frac{1}{d!}\sum_{\sigma \in S_d} D_\alpha(\sigma x^\downarrow),$$

where S_d *is the group of all permutations on* $\{1, \dots, d\}$ *and* $\sigma x = (x_{\sigma(1)}, \dots, x_{\sigma(d)})$.
Similarly

$$DM_\alpha^\downarrow(\ ;N) := DM_\alpha(\ ;N) \circ \psi^{-1}$$

defines the Ranked Dirichlet Multinomial *distribution.*

Ranked symmetric Dirichlet and Dirichlet-Multinomial measures can be interpreted as distributions on random partitions.
For every $r \in \mathbb{Z}_+^d$ let $\beta_j = \beta_j(r)$ be the number of elements in $r \in \mathbb{Z}_+^d$ equal to j, so $\sum \beta_j(r) = k(r)$ is the number of strictly positive components of r and $\sum_{i=1}^{|r|} i\beta_i(r) = |r|$.
For each $x \in \Delta_{(d-1)}$ denote

$$[x, r]_d := \sum_{(i_1,\dots,i_k)\subseteq\{1,\dots,d\}} \prod_{j=1}^{k} x_{i_j}^{r_j}$$

where the sum is over all $d_{[k]}$ subsequences of k distinct integers, and let $[x, r]$ be its extension to $x \in \Delta_\infty$. Take a collection $(\xi_1, \dots \xi_{|r|})$ of independent, identically distributed random variables, with values in a space of d "colors" ($d \le \infty$), and assume that x_j is the common probability of any ξ_i of being of color j. The function $[x, r]_d$ can be interpreted as the probability distribution of any such sample realization giving rise to $k(r)$ distinct values whose *unordered* frequencies count $\beta_1(r)$ singletons, $\beta_2(r)$ doubletons and so on.
There is a bijection between $r^\downarrow = \psi(r)$ and $\beta(r) = (\beta_1(r), \dots, \beta_{|r|}(r))$, both maximal invariant functions with respect to S_d, both representing partitions of $|r|$ in $k(r)$ parts. Note that $[x, r]_d$ is invariant too, for every $d \le \infty$. It is well-known that, for every $x \in \Delta_d^\downarrow$,

$$\sum_{r^\downarrow \in |r|\Delta_{(d-1)}^\downarrow} \binom{|r|}{r^\downarrow}\frac{1}{\prod_{i\ge1}\beta_i(r^\downarrow)!}[x, r^\downarrow]_d = 1, \tag{1.13}$$

that is, for every x,

$$\binom{|r|}{r^\downarrow}\frac{1}{\prod_{i\ge1}\beta_i(r^\downarrow)!}[x, r^\downarrow]_d$$

represents a probability distribution on the space of random partitions of $|r|$.
For $|\alpha| > 0$, let $D_{|\alpha|,d}$, $DM_{|\alpha|,d}$ denote the Dirichlet and Dirichlet-Multinomial distributions with symmetric

parameter $(|\alpha|/d, \ldots, |\alpha|/d)$. Then

$$DM^{\downarrow}_{|\alpha|,d}(r^{\downarrow}; N) = \mathbb{E}_{D^{\downarrow}_{|\alpha|,d}}\left\{\binom{N}{r^{\downarrow}} \frac{1}{\prod_{i\geq 1}\beta_i(r^{\downarrow})!}[X^{\downarrow}, r^{\downarrow}]_d\right\} \tag{1.14}$$

$$= d_{[k]}\frac{r!}{\prod_1^N j!^{\beta_j}\beta_j!} \cdot \frac{\prod_1^r(|\alpha|/d)^{\beta_j}_{(j)}}{|\alpha|_{(|r|)}}$$

$$\underset{d\to\infty}{\longrightarrow} \frac{r!}{\prod_1^r j^{\beta_j}\beta_j!} \cdot \frac{|\alpha|^k}{|\alpha|_{(|r|)}} := ESF_{|\alpha|}(r) \tag{1.15}$$

Definition 1.3. *The limit distribution $ESF_{|\alpha|}(r)$ in (1.15) is called the* Ewens Sampling Formula *with parameter $|\alpha|$.*

Poisson-Dirichlet point process ([23]).

Let $Y^{\infty} = (Y_1, Y_2, \ldots)$ be the sequence of points of a non-homogeneous point process with intensity measure

$$N_{|\alpha|}(y) = |\alpha| y^{-1} e^{-y}.$$

The probability generating functional is

$$\mathcal{F}_{|\alpha|}(\xi) = \mathbb{E}_{|\alpha|}\left(\exp\left\{\int \log\xi(y) N_{|\alpha|}(dy)\right\}\right) = \exp\left\{|\alpha|\int_0^{\infty}(\xi(y)-1)y^{-1}e^{-y}dy\right\}, \tag{1.16}$$

for suitable functions $\xi : \mathbb{R} \to [0,1]$. Then $|Y^{\infty}|$ is a Gamma($|\alpha|$) random variable and is independent of the sequence of ranked, normalized points

$$X^{\downarrow\infty} = \frac{\psi(Y^{\infty})}{|Y^{\infty}|}.$$

Definition 1.4. *The distribution of $X^{\downarrow\infty}$, is called the* Poisson-Dirichlet *distribution with parameter $|\alpha|$.*

Proposition 1.5. *(i) The* Poisson-Dirichlet ($|\alpha|$) *distribution on Δ_{∞} is the limit*

$$PD_{|\alpha|} = \lim_{d\to\infty} D^{\downarrow}_{|\alpha|,d}.$$

(ii) The relationship between D_{α} and DM_{α} is replicated by ESF, which arises as the (symmetric) moment formula for the PD distribution:

$$ESF_{|\alpha|}(r; N) = \mathbb{E}_{PD_{|\alpha|}}\left\{\binom{|r|}{r^{\downarrow}}\frac{1}{\prod_{i\geq 1}\beta_i(r^{\downarrow})!}[x, r^{\downarrow}]\right\}, \quad r \in N\Delta^{\downarrow}. \tag{1.17}$$

Proof. If $Y = (Y_1, \ldots, Y_d)$ is a collection of d independent random variables with identical distribution Gamma($|\alpha|/d, 1$), then their sum $|y|$ is a Gamma($|\alpha|$) random variable independent of $Y/|Y|$, which has distribution $D_{\alpha|,d}$. The probability generating functional of Y is ([13])

$$\mathcal{F}_{|\alpha|,d}(\xi) = \left(1 + \int_0^{\infty}(\xi(y)-1)\frac{|\alpha|}{d}\frac{y^{\frac{|\alpha|}{d}-1}e^{-y}}{\Gamma(\frac{|\alpha|}{d}+1)}dy\right)^d$$

$$\underset{d\to\infty}{\rightarrow} \mathcal{F}_{|\alpha|}(\xi) \tag{1.18}$$

which, by continuity of the ordering function ψ, implies that *if $X^{\downarrow d}$ has distribution $D^{\downarrow}_{|\theta|,d}$, then*

$$X^{\downarrow d} \overset{\mathcal{D}}{\to} X^{\downarrow\infty}.$$

This proves (i). For the proof of (ii) we refer to [13]. ☐

2. Polynomial kernels in the Dirichlet distribution.

2.1. Polynomial kernels for $d \geq 2$.

The aim of this section is to prove the following

Proposition 2.1. *For every $\alpha \in \mathbb{R}_+^d$ and every integer $|n|$, the $|n|$-th orthogonal polynomial kernel with respect to D_α is given by*

$$Q_{|n|}^{\alpha}(x,y) = \sum_{|m|=0}^{|n|} a_{|n||m|}^{|\alpha|} \xi_{|m|}^{\alpha}(x,y), \tag{2.19}$$

where

$$a_{|n||m|}^{|\alpha|} = (|\alpha| + 2|n| - 1)(-1)^{|n|-|m|} \frac{(|\alpha| + |m|)_{(|n|-1)}}{|m|!(|n| - |m|)!} \tag{2.20}$$

form a lower-triangular, invertible system, and

$$\xi_{|m|}^{\alpha}(x,y) = \sum_{|l|=|m|} \binom{|m|}{l} \frac{|\alpha|_{(|m|)}}{\prod_1^d \alpha_{i(l_i)}} \prod_1^d (x_i y_i)^{l_i} \tag{2.21}$$

$$= \sum_{|l|=|m|} \frac{\binom{|m|}{l} x^l \binom{|m|}{l} y^l}{DM_\alpha(l;|m|)}. \tag{2.22}$$

An inverse relationship is

$$\xi_{|m|}(x,y) = 1 + \sum_{|n|=1}^{|m|} \frac{|m|_{[|n|]}}{(|\alpha| + |m|)_{(|n|)}} Q_{|n|}^{\alpha}(x,y). \tag{2.23}$$

A first construction of the Kernel polynomials was given by [13]. We provide here a revised proof.

Proof. Let $\{Q_n^\circ\}$ be a system of *orthonormal* polynomials with respect to D_α (i.e. such $\mathbb{E}({Q_n^\circ}^2) = 1$). We need to show that, for independent Dirichlet distributed vectors X, Y, if $|n|, |k| \leq |m|$, then

$$E\Big(\xi_{|m|}^{\alpha}(X,Y) Q_n^\circ(X) Q_k^\circ(Y)\Big) = \delta_{nk} \frac{|m|_{[|n|]}}{(|\alpha| + |m|)_{(|n|)}}. \tag{2.24}$$

If this is true, an expansion is therefore

$$\begin{aligned} \xi_{|m|}^{\alpha}(x,y) &= 1 + \sum_{|n|=1}^{|m|} \frac{|m|_{[|n|]}}{(|\alpha| + |m|)_{(|n|)}} \sum_{\{n:|n| \text{ fixed }\}} Q_n^\circ(x) Q_n^\circ(y) \\ &= 1 + \sum_{|n|=1}^{|m|} \frac{|m|_{[|n|]}}{(|\alpha| + |m|)_{(|n|)}} Q_{|n|}^{\alpha}(x,y). \end{aligned} \tag{2.25}$$

Inverting the triangular matrix with (m,n)th element

$$\frac{|m|_{[|n|]}}{(|\alpha| + |m|)_{(|n|)}}$$

gives (2.19) from (2.23). The inverse matrix is triangular with $(|m|, |n|)$th element

$$(|\alpha| + 2|n| - 1)(-1)^{|n|-|m|} \frac{(|\alpha| + |m|)_{(|n|-1)}}{|m|!(|n| - |m|)!}, \qquad |n| \geq |m|$$

and the proof will be complete.

Proof of (2.24). Write

$$\mathbb{E}\Big(\prod_1^{d-1} {X_i}^{n_i}\xi^\alpha_{|m|}(X,Y)\mid Y\Big) = \sum_{\{l:|l|=|m|\}}\binom{|m|}{l}\prod_1^d Y_i^{l_i}\frac{\prod_1^{d-1}(l_i+\alpha_i)_{(n_i)}}{(|\alpha|+|m|)_{(|n|)}}. \tag{2.26}$$

Expressing the last product in (2.26) as

$$\prod_1^{d-1}(l_i+\alpha_i)_{(n_i)} = \prod_1^{d-1} l_{i[n_i]} + \sum_{\{k:|k|<|n|\}} b_{nk}\prod_1^{d-1} l_{i[k_i]}$$

for constants b_{nk}, from the identity

$$(l^*+\alpha^*)_{(n^*)} = \sum_{k^*=0}^{n^*}\binom{n^*}{k^*}(k^*+\alpha^*)_{(n^*-k^*)}l^*_{[k^*]},$$

shows that

$$\mathbb{E}\Big(\prod_1^{d-1} {X_i}^{n_i}\xi^\alpha_{|m|}(X,Y)\mid Y\Big) = \frac{|m|_{[n]}}{(|\alpha|+|m|)_{(n)}}\prod_1^{d-1} Y_i^{n_i} + R_0(Y). \tag{2.27}$$

Thus if $|n|\le|k|\le|m|$,

$$\begin{aligned}\mathbb{E}\Big(\xi^\alpha_{|m|}(X,Y)Q^\circ_n(X)\mid Y\Big) &= \frac{|m|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}}\sum_{\{k:|k|=|n|\}} a_{nk}\prod_1^{d-1}{Y_i}^{k_i} + R_1(Y)\\ &= \frac{|m|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}}Q^\circ_n(Y) + R_2(Y),\end{aligned} \tag{2.28}$$

where

$$\sum_{\{k:|k|=|n|\}} a_{nk}\prod_1^{d-1}{X_i}^{k_i}$$

are terms of leading degree $|n|$ in $Q^\circ_n(X)$ and $R_j(Y)$, $j=0,1,2$ are polynomials of degree less than $|n|$ in Y. Thus if $|n|\le|k|\le m$,

$$\begin{aligned}E\Big(\xi^\alpha_{|m|}(X,Y)Q^\circ_n(X)Q^\circ_k(Y)\Big) &= E\Big(Q^\circ_k(Y)\Big\{\frac{|m|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}}Q^\circ_{|n|}(Y)+R_2(Y)\Big\}\Big)\\ &= \frac{|m|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}}\delta_{nk}\end{aligned} \tag{2.29}$$

By symmetry (2.29) holds for all n,k such that $|n|,|k|\le m$. $\square$

2.2. *Some properties of the kernel polynomials.*

2.2.1. *Particular cases.*

$$\begin{aligned}Q^\alpha_0 &= 1,\\ ,Q^\alpha_1 &= (|\alpha|+1)(\xi_1-1)\\ &= (|\alpha|+1)\Big(|\alpha|\sum_1^d x_iy_i/\alpha_i - 1\Big)\\ Q^\alpha_2 &= \frac{1}{2}(|\alpha|+3)\Big((|\alpha|+2)\xi_2 - 2(|\alpha|+1)\xi_1 + |\alpha|\Big),\end{aligned}$$

where

$$\xi_2 = |\alpha|(|\alpha|+1)\Big(\sum_1^d (x_i y_i)^2/\alpha_i(\alpha_i+1) + 2\sum_{i<j} x_i x_j y_i y_j/\alpha_i\alpha_j\Big).$$

2.2.2. The j-th coordinate kernel.

A well-known property of Dirichlet measures is that, if Y is a Dirichlet(α) vector in $\Delta_{(d-1)}$ then its j-th coordinate Y_j has distribution $D_{\alpha_j,|\alpha|-\alpha_j}$. Such a property is reflected in the Jacobi polynomial kernels. For every d let e_j be the vector in $\mathbb{R}^d$ with every i-th coordinate equal δ_{ij}, $i,j=1,\dots,d$. Then

$$\xi^\alpha_{|m|}(y,e_j) = \frac{(|\alpha|)_{(|m|)}}{(\alpha_j)_{(|m|)}} y_j^{|m|}, \qquad |m|\in\mathbb{Z}_+,\ y\in\Delta_{(d-1)}. \tag{2.30}$$

In particular,

$$\xi^\alpha_{|m|}(e_j,e_k) = \frac{(|\alpha|)_{(|m|)}}{(\alpha_j)_{(|m|)}}\delta_{jk}. \tag{2.31}$$

Therefore, for every d and $\alpha\in\mathbb{R}^d_+$, (2.30) implies

$$\begin{aligned} Q^\alpha_{|n|}(y,e_j) &= \sum_{|m|=0}^{|n|} a^{|\alpha|}_{|n||m|}\xi^\alpha_{|m|}(e_j,y) \\ &= Q^{\alpha_j,|\alpha|-\alpha_j}_{|n|}(y_j,1) \\ &= \zeta^{\alpha_j,|\alpha|-\alpha_j}_{|n|} R^{\alpha_j,|\alpha|-\alpha_j}_{|n|}(y_j), \qquad j=1,\dots,d,\ y\in\Delta_{(d-1)}.. \end{aligned} \tag{2.32}$$

where

$$R^{\alpha,\beta}_n(x) = \frac{Q^{\alpha.\beta}_n(x,1)}{Q^{\alpha,\beta}_n(1,1)} = {}_2F_1\left(\begin{matrix}-n,\ n+\theta-1\\ \beta\end{matrix}\ \middle|\ 1-x\right) \qquad n=0,1,2,\dots \tag{2.33}$$

are univariate Jacobi Polynomials ($\alpha>0,\beta>0$) normalized by their value at 1 and

$$\frac{1}{\zeta^{\alpha,\beta}_{|n|}} := E\left[R^{\alpha,\beta}_{|n|}(x)\right]^2.$$

In (2.33), ${}_pF_q$, $p,q\in\mathbb{N}$, denotes the Hypergeometric function (see [1] for basic properties).

Remark 2.2. *For $\alpha,\beta\in\mathbb{R}_+$, let $\theta=\alpha+\beta$. It is known (e.g. [17], (3.25)) that*

$$\frac{1}{\zeta^{\alpha,\beta}_{|n|}} = n!\frac{1}{(\theta+2n-1)(\theta)_{(n-1)}}\frac{(\alpha)_{(n)}}{(\beta)_{(n)}}. \tag{2.34}$$

On the other hand, for every $\alpha=(\alpha_1,\dots,\alpha_d)$,

$$\zeta^{\alpha_j,|\alpha|-\alpha_j}_{|n|} = Q^\alpha_{|n|}(e_j,e_j) = \sum_{|m|=0}^{|n|} a^{|\alpha|}_{|n||m|}\frac{(|\alpha|)_{(|m|)}}{(\alpha_j)_{(|m|)}}. \tag{2.35}$$

Thus the identity

$$\frac{(|\alpha|-\alpha_j)_{(|n|)}}{(\alpha_j)_{(|n|)}} = \sum_{|m|=0}^{|n|}\binom{|n|}{|m|}(-1)^{|n|-|m|}\frac{(|\alpha|+|n|-1)_{(|m|)}}{(\alpha_j)_{(|m|)}} \tag{2.36}$$

holds for every $|\alpha|$ positive and $0<\alpha_j<|\alpha|$. In the limit as $\alpha_j\to|\alpha|$, this implies

$$\sum_{|m|=0}^{|n|}\binom{|n|}{|m|}(-1)^{|n|-|m|}(|\alpha|+|m|)_{(|n|-1)} = 0. \tag{2.37}$$

2.2.3. Addition of variables in x.

Let A be a $d' \times d$ ($d' < d$) 0-1 matrix whose rows are orthogonal. A known property of the Dirichlet distribution is that, if X has distribution D_α, then AX has a $D_{A\alpha}$ distribution. Similarly, with some easy computation

$$\mathbb{E}(\xi^{\alpha}_{|m|}(X,y) \mid AX = ax) = \xi^{A\alpha}_{|m|}(AX, Ay).$$

One has therefore the following

Proposition 2.3. *A representation for Polynomial kernels in $D_{A\alpha}$ is:*

$$Q^{A\alpha}_{|n|}(Ax, Ay) = \mathbb{E}\Big[Q^{\alpha}_{|n|}(X,y) \mid AX = Ax\Big]. \tag{2.38}$$

Example 2.4. *For any $\alpha \in \mathbb{R}^d$ and $k \leq d$, suppose $AX = (X_1 + \cdots + X_k, X_{k+1} + \cdots + X_d) = X'$. Then, denoting $\alpha' = \alpha_1 + \cdots + \alpha_k$ and $\beta' = \alpha_{k+1} + \cdots + \alpha_d$, one has*

$$Q^{A\alpha}_{|n|}(x', y') = \zeta^{\alpha',\beta'}_{|n|} R^{\alpha',\beta'}_{|n|}(x') R^{\alpha',\beta'}_{|n|}(y') = \mathbb{E}\Big[Q^{\alpha}_{|n|}(X,y) \mid X' = x'\Big].$$

3. Kernel Polynomials on the Dirichlet-Multinomial distribution.

For the Dirichlet-Multinomial distribution, it is possible to derive an explicit formula for the kernel polynomials by considering that Hahn polynomials can be expressed as *posterior* mixtures of Jacobi polynomials (cf. [17], 5.2). Let $\{Q^\circ_n(x)\}$ be a orthonormal polynomial set on the Dirichlet, considered as functions of $(x_1, \ldots, x_{d-1})$. Define

$$h^\circ_n(r\,;|r|) = \int Q^\circ_n(x) D_{\alpha+r}(dx), \tag{3.39}$$

then $\{h^\circ_n\}$ is a system of multivariate orthogonal polynomials with respect to DM_α with constant of orthogonality

$$\mathbb{E}_{\alpha,|r|}\left[h^\circ_n(R\,;|r|)^2\right] = \delta_{nn'} \frac{|r|_{[|n|]}}{(|\alpha| + |r|)_{(|n|)}}. \tag{3.40}$$

Note also that if $|r| \to \infty$ with $r_i/|r| \to x_i$, $i = 1, \ldots, d$, then

$$\lim_{|r|\to\infty} h^\circ_n(r\,;|r|) = Q^\circ_n(x).$$

.

Proposition 3.1. *The Hahn Kernel polynomials with respect to $DM_\alpha(\cdot \,||r|)$ are*

$$H^{\alpha}_{|n|}(r,s) = \frac{(\alpha + |r|)_{(|n|)}}{|r|_{[|n|]}} \int\int Q_{|n|}(x,y) D_{\alpha+r}(dx) D_{\alpha+s}(dy) \tag{3.41}$$

for $r = (r_1, \ldots, r_d)$, $s = (s_1, \ldots, s_d)$, $|r| = |s|$ fixed, and $|n| = 0, 1, \ldots, |r|$.
An explicit expression is

$$H^{\alpha}_{|n|}(r,s) = \frac{(|\alpha| + |r|)_{(|n|)}}{|r|_{[|n|]}} \cdot \sum_{m=0}^{n} a^{|\alpha|}_{|n||m|} \xi^{H,\alpha}_{|m|}(r,s), \tag{3.42}$$

where $(a^{|\alpha|}_{|n||m|})$ *is as in* (2.20) *and*

$$\xi^{H,\alpha}_{|m|}(r,s) = \sum_{|l|=|m|}\binom{|m|}{l}\frac{|\alpha|_{(|m|)}}{\prod_1^d \alpha_{i(l_i)}}\frac{\prod_1^d(\alpha_i+r_i)_{(l_i)}(\alpha_i+s_i)_{(l_i)}}{(|\alpha|+|r|)_{(|m|)}(|\alpha|+|s|)_{(|m|)}} \tag{3.43}$$

$$= \sum_{|l|=|m|}\frac{DM_{\alpha+r}(l;|m|)DM_{\alpha+s}(l;|m|)}{DM_\alpha(l;|m|)}. \tag{3.44}$$

Proof. The Kernel sum is by definition

$$H^\alpha_{|n|}(r,s) = \frac{(|\alpha|+|r|)_{(|n|)}}{|r|_{[|n|]}}\sum_{\{n:|n|\ \text{fixed}\}} h^\circ_n(r\ ;\ |r|)h^\circ_n(s\ ;\ |r|) \tag{3.45}$$

and from (3.41), (3.42) follows. The form of $\xi^{H,\alpha}_{|m|}$ is obtained by taking the expectation $\xi^\alpha_{|m|}(X,Y)$, appearing in the representation (2.19) of $Q^\alpha_{|n|}$, with respect to the product measure $D_{\alpha+r}D_{\alpha+s}$. ☐

The first polynomial kernel is

$$H^\alpha_1(r,s) = \frac{(|\alpha|+1)(|\alpha|+r)}{|\alpha|}\Big(\frac{|\alpha|}{(|\alpha|+|r|)^2}\sum_1^d\frac{(\alpha_i+r_i)(\alpha_i+s_i)}{\alpha_i} - 1\Big).$$

Projections on one coordinate.

As in the Jacobi case, the connection with Hahn polynomials on $\{0,\dots,N\}$ is given by marginalization on one coordinate.

Proposition 3.2. *For* $|r|\in\mathbb{N}$ *and* $d\in\mathbb{N}$, *denote* $\hat{r}_{j,1} = e_j|r|\in\mathbb{N}^d$, *where* $e_j=(0,\dots,0,1,0,\dots,0)$ *with* 1 *only at the* j*-th coordinate.*
For every $\alpha\in\mathbb{N}^d$,

$$H^\alpha_{|n|}(s,e_j|r|) = \frac{1}{c^{|\alpha|}_{|r|,|n|}} h^{\circ(\alpha_j,|\alpha|-\alpha_j)}_{|n|}(s_j;|r|)h^{\circ(\alpha_j,|\alpha|-\alpha_j)}_{|n|}(|r|;|r|), \qquad |s|=|r|. \tag{3.46}$$

where

$$c^{|\alpha|}_{|r|,|n|} := \frac{|r|_{[|n|]}}{(|\alpha|+|r|)_{(|n|)}} = \mathbb{E}\left[h^{\circ(\alpha,\beta)}_{|n|}(R;|r|)^2\right]$$

and $\{h^{\circ,j}_{|n|}\}$ *are orthogonal polynomials with respect to* $DM_{\alpha_j,|\alpha|-\alpha_j}(\cdot;|r|)$.

Proof. Because for every d and $\alpha\in\mathbb{R}^d_+$

$$H^\alpha_{|n|}(s,r) = \frac{1}{c^{|\alpha|}_{|r|,|n|}}\sum_{|m|=0}^{|n|} a^{|\alpha|}_{|n||m|}\xi^{H,\alpha}_{|m|}(r,s)$$

for $d=2$ and $\alpha,\beta>0$ with $\alpha+\beta=|\alpha|$,

$$h^{\circ(\alpha,\beta)}_{|n|}(k;|s|)h^{\circ(\alpha,\beta)}_{|n|}(j;|s|) = \sum_{|m|=0}^{|n|} a^{|\alpha|}_{|n||m|}\xi^{H,\alpha,\beta}_{|m|}(k,j), \qquad k,j=0,\dots,|s|.$$

Now rewrite $\xi^{H,\alpha}_{|m|}$ as

$$\begin{aligned}\xi^{H,\alpha}_{|m|}(s,r) &= \sum_{|l|=|m|}\frac{DM_{\alpha+s}(l;|m|)DM_{\alpha+r}(l;|m|)}{DM_\alpha(l;|m|)}\\ &= \sum_{|l|=|m|}DM_{\alpha+s}(l;|m|)\frac{DM_{\alpha+l}(r;|s|)}{DM_\alpha(r;|s|)}.\end{aligned} \tag{3.47}$$

Consider without loss of generality the case $j=1$. Since for every α

$$DM_\alpha(l;|m|) = DM_{\alpha_1,|\alpha|-\alpha_1}(l_1;|m|)DM_{\alpha_2,\dots,\alpha_d}(l_2,\dots,l_d;|m|-l_1)$$

then, for $|r|=|s|$

$$\begin{aligned}\xi^{H,\alpha}_{|m|}(s,e_1|r|) &= \sum_{l_1=0}^{|m|}DM_{\alpha_1+s_1,|\alpha|-\alpha_1+|m|-s_1}(l_1;|m|)\frac{DM_{\alpha_1+l_1,|\alpha|-\alpha_1+|m|-l_1}(|r|;|r|)}{DM_{\alpha_1,|\alpha|-\alpha_1}(|r|;|r|)}\\ &\quad\times\sum_{|u|=|m|-l_1}DM_{\alpha'+s'}(u;|m|-l_1)\frac{DM_{\alpha+l}(0:0)}{DM_\alpha(0;0)}\\ &= \sum_{l_1=0}^{|m|}DM_{\alpha+s}(l_1;|m|)\frac{DM_{\alpha_1+l_1}(|r|;|r|)}{DM_\alpha(|r|;|r|)}\sum_{|u|=|m|-l_1}DM_{\alpha'+s'}(u;|m|-l_1)\\ &= \sum_{l_1=0}^{|m|}DM_{\alpha+s}(l_1;|m|)\frac{DM_{\alpha_1+l_1}(|r|;|r|)}{DM_\alpha(|r|;|r|)} &(3.48)\\ &= \xi^{H,\alpha_1,|\alpha|-\alpha_1}_{|m|}(s_1,|r|) &(3.49)\end{aligned}$$

Then (3.46) follows immediately. □

3.1. Generalization of Gasper's product formula for Hahn polynomials.

For $d=2$ and $\alpha,\beta>0$ the Hahn polynomials

$$h^{\alpha,\beta}_{|n|}(r;N) = {}_3F_2\left(\begin{matrix}-|n|,|n|+\theta-1,-r\\ \alpha,-N\end{matrix}\;\middle|\;1\right),\qquad |n|=0,1,\dots,N. \tag{3.50}$$

with $\theta=\alpha+\beta$, have constant of orthogonality

$$\frac{1}{u^{\alpha,\beta}_{N,n}} := \sum_{r=0}^{N}\left[h^{\alpha,\beta}_n(r;N)\right]^2 DM_{\alpha,\beta}(n;N) = \frac{1}{\binom{N}{n}}\frac{(\theta+N)_{(n)}}{(\theta)_{(n-1)}}\frac{1}{\theta+2n-1}\frac{(\beta)_{(n)}}{(\alpha)_{(n)}}.$$

The following product formula was found by Gasper [10]:

$$h^{\alpha,\beta}_{|n|}(r;N)h^{\alpha,\beta}_{|n|}(s;N) = \frac{(-1)^{|n|}(\beta)_{(|n|)}}{(\alpha)_{(|n|)}}\sum_{l=0}^{|n|}\sum_{k=0}^{|n|-l}\frac{(-1)^{l+k}|n|_{[l+k]}(\theta+|n|-1)_{(l+k)}r_{[l]}s_{[l]}(N-r)_{[k]}(N-s)_{[k]}}{l!k!N_{[l+k]}N_{[l+k]}(\alpha)_{(l)}(\beta)_{(k)}}. \tag{3.51}$$

Thus

$$\begin{aligned}u^{\alpha,\beta}_{N,n}h^{\alpha,\beta}_{|n|}(r;N)h^{\alpha,\beta}_{|n|}(s;N) &= \frac{N_{[|n|]}}{(\theta+N)_{(|n|)}}\sum_{|m|=0}^{|n|}\frac{(-1)^{|n|-|m|}(\theta)_{(|n|-1)}(\theta+|n|-1)_{(|m|)}(\theta+2|n|-1)}{|m|!(|n|-|m|)!(\theta)_{(|m|)}}\chi^{H,\alpha,\beta}_{|m|}(r,s)\\ &= \frac{N_{[|n|]}}{(\theta+N)_{(|n|)}}\sum_{|m|=0}^{|n|}a^\theta_{|n||m|}\chi^{H,\alpha,\beta}_{|m|}(r,s),\end{aligned} \tag{3.52}$$

where

$$\chi^{H,\alpha,\beta}_{|m|}(r,s) := \sum_{j=0}^{|m|} \frac{1}{DM_{\alpha,\beta}(j;|m|)} \left[\frac{\binom{|m|}{j} r_{[j]}(N-r)_{[|m|-j]}}{N_{[|m|]}} \right] \left[\frac{\binom{|m|}{j} s_{[j]}(N-s)_{[|m|-j]}}{N_{[|m|]}} \right]. \tag{3.53}$$

By uniqueness of polynomial kernels, we can identify the connection coefficients between the functions ξ and χ:

Proposition 3.3. *For every $|m|, |n| \in \mathbb{Z}_+$, and every $r, s \in \{0, \dots, N\}$,*

$$\xi^{H,\alpha,\beta}_{|m|}(r,s) = \sum_{|l|=0}^{|m|} b_{|m||l|} \chi^{H,\alpha,\beta}_{|l|}(r,s), \tag{3.54}$$

where

$$b_{|m||l|} = \sum_{|n|=|l|}^{|m|} \left(\frac{N_{[|n|]}}{(\theta+N)_{(|n|)}} \right)^2 \frac{m_{[|n|]}}{(\theta+m)_{(|n|)}} a^{\theta}_{|n||l|}. \tag{3.55}$$

Proof. From 3.42,

$$u^{\alpha,\beta}_{N,n} h^{\alpha,\beta}_{|n|}(r;N) h^{\alpha,\beta}_{|n|}(s;N) = H^{\alpha,\beta}_{|n|}(r,s) = \frac{(\theta+N)_{(|n|)}}{N_{[|n|]}} \sum_{|m|=0}^{|n|} a^{\theta}_{|n||m|} \xi^{H,\alpha,\beta}_{|m|}(r,s). \tag{3.56}$$

Since the array $A = (a^{\theta}_{|n||m|})$ has inverse $C = A^{-1}$ with entries

$$c^{\theta}_{|m||n|} = \left(\frac{m_{[|n|]}}{(\theta+m)_{(|n|)}} \right), \tag{3.57}$$

then equating (3.56) and (3.52) leads to

$$\begin{aligned} \xi^{H,\alpha,\beta}_{|m|} &= \sum_{|n|=0}^{|m|} c^{\theta}_{|m||n|} \frac{N_{[|n|]}}{(\theta+N)_{(|n|)}} H^{\alpha,\beta}_{|n|} \\ &= \sum_{|n|=0}^{|m|} c^{\theta}_{|m||n|} \left(\frac{N_{[|n|]}}{(\theta+N)_{(|n|)}} \right)^2 \sum_{|l|=0}^{|n|} a^{\theta}_{|n||l|} \chi^{H,\alpha,\beta}_{|l|} \\ &= \sum_{|l|=0}^{|m|} b_{|m||l|} \chi^{H,\alpha,\beta}_{|l|}. \end{aligned}$$

□

The following Corollary is then straightforward.

Corollary 3.4.

$$\mathbb{E}\left[\xi^{H,\alpha,\beta}_{|m|} \chi^{H,\alpha,\beta}_{|l|} \right] = \mathbb{E}\left[\xi^{H,\alpha,\beta}_{|l|} \chi^{H,\alpha,\beta}_{|m|} \right] = \sum_{|n|=0}^{|m|\wedge|l|} \frac{|m|_{[|l|]} |l|_{[|n|]}}{(\theta+|m|)_{(|n|)} (\theta+|l|)_{(|n|)}}.$$

For every $r \in N\Delta_{(d-1)}$ and $m \in \mathbb{Z}^d_+$ define

$$p_m(r) = \prod_{i=1}^{d} (r_i)_{[m_i]}.$$

Gasper's product formula (3.51), or rather the representation (3.52) has a multivariate extension in the following.

Proposition 3.5. *For every d, $\alpha \in \mathbb{R}_+^d$ and $N \in \mathbb{Z}_+$, the Hahn polynomial kernels admit the following representation:*

$$H_{|n|}^{\alpha}(r,s) = \frac{N_{[|n|]}}{(|\alpha|+N)_{(|n|)}} \sum_{|m|=0}^{|n|} a_{|n||m|}^{|\alpha|} \chi_{|m|}^{H,\alpha}(r,s), \qquad r,s \in N\Delta_{(d-1)},\ |n| = 0,1,\dots \tag{3.58}$$

where

$$\chi_{|m|}^{H,\alpha}(r,s) := \sum_{l:|l|=|m|} \frac{1}{DM_\alpha(l;|m|)} \left(\frac{\binom{|m|}{l} p_l(r)}{N_{[|m|]}} \right) \left(\frac{\binom{|m|}{l} p_l(s)}{N_{[|m|]}} \right). \tag{3.59}$$

Proof. If we prove that, for every $|m|$ and $|n|$,

$$\chi_{|m|}^{H,\alpha}(r,s) = \sum_{|n|=0}^{|m|} \frac{c_{|m||n|}^{|\alpha|}}{c_{|N||n|}^{|\alpha|}} H_{|n|}^{\alpha}(r,s)$$

where $c_{|i||j|}^{|\alpha|}$ are given by (3.57) (independent of d!), then the proof follows by inversion.
Consider the orthonormal multivariate Jacobi polynomials $Q_n^{\circ}(x)$. The functions

$$h_n^{\circ}(r;\ N) := \int_{\Delta_{(d-1)}} Q_n^{\circ}(x) D_{\alpha+r}(dx)$$

satisfy the identity

$$\mathbb{E}\left[h_n^{\circ}(R\,;N) \binom{|m|}{l} p_l(R) \right] = N_{[|m|]} h_n^{\circ}(l\ ;|m|) DM_\alpha(l;|m|), \qquad l \in |m|\Delta_{(d-1)}, n \in \mathbb{Z}_+^d. \tag{3.60}$$

([12], (5.71)).
Then for every fixed s,

$$\mathbb{E}\left[\chi_{|m|}^{H,\alpha}(R,s) h_n^{\circ}(R;N) \right] = \sum_{|l|=|m|} \binom{|m|}{l} \frac{p_l(s)}{N_{[|m|]}} h_{|n|}^{\circ}(l;\ |m|), \tag{3.61}$$

so iterating the argument we can write

$$\mathbb{E}\left[\chi_{|m|}^{H,\alpha}(R,S) h_n^{\circ}(R\,;N) h_n^{\circ}(S\,;N) \right] = c_{|m||n|}. \tag{3.62}$$

Now by uniqueness of the polynomial kernel,

$$H_{|n|}^{\alpha}(r,s) = \sum_{|n|=0}^{\infty} \frac{1}{c_{N,n}^{|\alpha|}} h_n^{\circ}(r\,;N) h_n^{\circ}(s\,;N),$$

therefore

$$\chi_{|m|}^{H,\alpha}(r,s) = \sum_{|n|=0}^{|m|} \frac{c_{|m||n|}^{|\alpha|}}{c_{|N||n|}^{|\alpha|}} H_{|n|}^{\alpha}(r,s)$$

and the proof is complete. ☐

The connection coefficients between $\xi_{|m|}^{H,\alpha}$ and $\xi_{|m|}^{\alpha}$ are, for every d, the same as for the two-dimensional case:

Corollary 3.6. *For every d and $\alpha \in \mathbb{R}_+^d$,*

(i)

$$\xi_{|m|}^{H,\alpha,\beta}(r,s) = \sum_{|l|=0}^{|m|} b_{|m||l|} \chi_{|l|}^{H,\alpha,\beta}(r,s), \tag{3.63}$$

where $(b_{|m||l|})$ are given by (3.55).

(ii)

$$\mathbb{E}\left[\xi^{H,\alpha}_{|m|}\chi^{H,\alpha}_{|l|}\right] = \mathbb{E}\left[\xi^{H,\alpha}_{|l|}\chi^{H,\alpha}_{|m|}\right] = \sum_{|n|=0}^{|m|\wedge|l|} \frac{|m|_{[|l|]}|l|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}(|\alpha|+|l|)_{(|n|)}} \qquad |m|,|l| = 0,1,2,\ldots$$

3.2. Polynomial kernels on the Hypergeometric distribution.

Note that there is a direct alternative proof of orthogonality of $H^{\alpha}_{|n|}(r,s)$ similar to that for $Q_{|n|}(x)$. In the Hahn analogous proof orthogonality does not depend on the fact that $|\alpha|>0$. In particular we obtain *Kernels on the hypergeometric distribution:*

$$\frac{\binom{c_1}{r_1}\cdots\binom{c_d}{r_d}}{\binom{|c|}{|r|}} \tag{3.64}$$

by replacing α by $-c$ in (3.42) and (3.43). Again a direct proof similar to that for $Q_{|n|}(x)$ would be possible.

4. Symmetric kernels on ranked Dirichlet and Poisson-Dirichlet measures.

Let $D_{|\alpha|,d}$ be the Dirichlet distribution on d points with symmetric parameters $(|\alpha|/d,\ldots,|\alpha|/d)$, and $D^{\downarrow}_{|\alpha|,d}$ its ranked version. Denote with $Q^{(|\alpha|,d)}_{|n|}$ and $Q^{(|\alpha|,d)\downarrow}_{|n|}$ the corresponding $|n|$-kernels.

Proposition 4.1.

$$Q^{(|\alpha|,d)\downarrow}_{|n|} = (d!)^{-1}\sum_{\sigma} Q^{(|\alpha|,d)}_{|n|}(\sigma(x),y),$$

where summation is over all permutations σ of $1,\ldots,d$. The Kernel polynomials have a similar form to $Q^{(|\alpha|,d)}_{|n|}$, but with $\xi^{(|\alpha|,d)}_{m}$ replaced by

$$\xi^{(|\alpha|,d)\downarrow}_{|m|} = \sum_{l\in|m|\Delta^{\downarrow}_{(d-1)}} \frac{|m|!|\theta|_{(m)}(d-k)!\left(\prod_1^m \beta_i(l)!\right)[x;l][y;l]}{d!\prod_1^m\left[j!(|\theta|/d)_{(j)}\right]^{\beta_j(l)}} \tag{4.65}$$

$$= \sum_{l\in|m|\Delta^{\downarrow}_{(d-1)}} \frac{\sharp(l)[x;l]\sharp(l)[y;l]}{DM^{\downarrow}_{|\alpha|,d}(l;|m|)}. \tag{4.66}$$

where

$$\sharp(l) := \binom{|l|}{l}\frac{1}{\prod_{i\geq 1}\beta_i(l)!}.$$

Proof. Note that

$$
\begin{aligned}
Q^{(|\alpha|,d)\downarrow}_{|n|}(x,y) &= \frac{1}{d!}\sum_{\sigma\in\mathcal{G}_d} Q^{(|\alpha|,d)}_{|n|}(\sigma x,y) \\
&= \frac{1}{d!}\sum_{\sigma\in\mathcal{G}_d}\sum_{|m|\le|n|} a^{|\alpha|}_{|n||m|}\xi^{(|\alpha|,d)}_{|m|}(\sigma x,y) \\
&= d!\sum_{|m|\le|n|} a^{|\alpha|}_{|n||m|}\frac{1}{(d!)^2}\sum_{\sigma\in\mathcal{G}_d}\sum_{|l|=|m|}\frac{\binom{|m|}{l}^2(\sigma x)^l y^l}{DM_{|\alpha|,d}(l;|m|)} \\
&= \sum_{|m|\le|n|} a^{|\alpha|}_{|n||m|}\frac{1}{(d!)^2}\sum_{\sigma,\tau\in\mathcal{G}_d}\sum_{|l|=|m|}\frac{\binom{|m|}{l}^2(\sigma\tau x)^l (y)^l}{DM_{|\alpha|}(l;|m|)} \\
&= \sum_{|m|\le|n|} a^{|\alpha|}_{|n||m|}\frac{1}{(d!)^2}\sum_{\sigma,\tau\in\mathcal{G}_d}\sum_{|l|=|m|}\frac{\binom{|m|}{l}^2(\sigma x)^l (\tau y)^l}{DM_{|\alpha|,d}(l;|m|)} \qquad (4.67) \\
&= \frac{1}{(d!)^2}\sum_{\sigma,\tau\in\mathcal{G}_d} Q^{(|\alpha|,d)}_{|n|}(\sigma x,\tau y) \qquad (4.68)
\end{aligned}
$$

Now,

$$
\begin{aligned}
\mathbb{E}_{D^{\downarrow}_{(|\alpha|,d)}}\left[Q^{(|\alpha|,d)\downarrow}_{|n|}(x,Y)Q^{(|\alpha|,d)\downarrow}_{|m|}(z,Y)\right] &= \frac{1}{d!}\sum_{\sigma\in\mathcal{G}_d} Q^{(|\alpha|,d)}_{|n|}(\sigma x,z)\delta_{|n||m|} \\
&= Q^{(|\alpha|,d)\downarrow}_{|n|}(x,z)\delta_{|n||m|}, \qquad (4.69)
\end{aligned}
$$

hence $Q^{(|\alpha|,d)\downarrow}_{|n|}$ is the $|n|$ polynomial kernel with respect to $D^{\downarrow}_{(|\alpha|,d)}$. The second part of the theorem, involving the identity (4.66), is just another way of rewriting (4.67). ☐

Remark 4.2. *The first polynomial is* $Q^{(|\alpha|,d)\downarrow}_{|1|}\equiv 0$.

4.0.1. Infinite-dimensional limit.

As $d\to\infty$, $\xi^{(|\alpha|,d)\downarrow}_{|m|}\to\xi^{(|\alpha|,\infty)\downarrow}_{|m|}$, with

$$
\begin{aligned}
\xi^{(|\alpha|,\infty)\downarrow}_{|m|} &= |\alpha|_{(|m|)}\sum\frac{|m|!\,(\prod_1^m b_i!)\,[x;l][y;l]}{|\alpha|^k[0!1!]^{b_1}\cdots[(k-1)!k!]^{b_k}} \qquad (4.70) \\
&= \sum\frac{\sharp(l)[x;l]\binom{|m|}{l}\sharp(l)[y;l]}{ESF_{|\alpha|}(l)}. \qquad (4.71)
\end{aligned}
$$

Proposition 4.3. *The* $|n|$*-polynomial kernel with respect to the Poisson-Dirichlet point process is given by*

$$
Q^{(|\alpha|,\infty)\downarrow}_{|n|}=\sum_{|m|=0}^{|n|} a^{|\alpha|}_{|n||m|}\xi^{(|\alpha|,\infty)\downarrow}_{|m|}. \qquad (4.72)
$$

The first polynomial is zero and the second polynomial is

$$
Q^{\infty}_2=(F_1-\mu)(F_2-\mu)/\sigma^2\,,
$$

where

$$
F_1=\sum_1^\infty x^2_{(i)},\ F_2=\sum_1^\infty y^2_{(i)},
$$

and

$$\mu = \frac{1}{1+|\alpha|}, \;\; \sigma^2 = \frac{2|\alpha|}{(|\alpha|+3)(|\alpha|+2)(|\alpha|+1)^2}.$$

4.1. Symmetric polynomials.

We can use the symmetric kernel derived in Section 4 to obtain a system of orthogonal polynomials with weight measure given by the distribution $D^{\downarrow}_{|\alpha|,d}$. We first need to observe a uniqueness property of general polynomial kernels.

Lemma 4.4. *Let π be any measure on $\mathbb{R}^d$, and assume that $\mathbb{P} = \{P_n : n \in \mathbb{N}^d\}$ is a system of polynomials such that*

$$P_{|n|}(x,y) = \sum_{\{|m|=|n|\}} P_m(x)P_m(y), \quad |n| = 0,1,2,\dots$$

forms a complete orthogonal kernel system with respect to π.
Then $\mathbb{P}$ is a system of orthonormal polynomials with respect to π.

Proof. Because $P_{|n|}(x,y)$ is an orthogonal kernel, we have

$$\mathbb{E}\left[P_{|k|}(x,Y)P_{|l|}(z,Y)\right] = \delta_{|k||l|}P_{|l|}(x,z), \qquad |k|,|l| \in \mathbb{N}.$$

This can be written as

$$\sum_{\{|m|=|k|\}} P_m(x)\mathbb{E}\left[P_m(Y)P_{|l|}(z,Y)\right] = \delta_{|k||l|}P_{|l|}(x,z), \tag{4.73}$$

which shows that for every x

$$\mathbb{E}\left[P_m(Y)P_{|l|}(x,Y)\right] = P_m(x)\delta_{|m||l|}.$$

That is,

$$P_k(x) = \sum_{|m|=|k|} c_{km}P_m(x) \tag{4.74}$$

with

$$c_{km} = \mathbb{E}\left[P_k(Y)P_m(Y)\right].$$

Identity (4.74) implies that $P_k(x)$ is "biorthogonal to itself", i.e.

$$c_{km} = \delta_{km},$$

and the proof is complete. □

Now it is easy to derive orthogonal polynomials with respect to the symmetric measure $D^{\downarrow}_{|\alpha|,d}$.

Proposition 4.5. *Let $\{P_n^{(|\alpha|,d)}\}$ be an orthonormal system of polynomials with respect to $D_{|\alpha|,d}$. Then the system $\{P_n^{(|\alpha|,d)\downarrow}\}$ defined by*

$$P_n^{(|\alpha|,d)\downarrow}(x) = \frac{1}{d!}\sum_{\sigma\in\mathcal{G}_d} P_n^{(|\alpha|,d)}(\sigma x), \tag{4.75}$$

where $\mathcal{G}_d$ is the group of all permutations of $\{1,\dots,d\}$, is orthonormal with respect to $D^{\downarrow}_{|\alpha|,d}$.

Proof. In Section 4 we have shown that the polynomial kernel relative to $D^{\downarrow}_{|\alpha|,d}$ is of the form

$$Q_{|n|}^{(|\alpha|,d)\downarrow}(x,y) = \frac{1}{(d!)^2}\sum_{\sigma,\tau\in\mathcal{G}_d} Q_{|n|}^{(|\alpha|,d)}(\sigma x,\tau y) = \sum_{|m|=|n|}\left(\sum_\sigma \frac{P_m^{(|\alpha|,d)}(\sigma x)}{d!}\right)\left(\sum_\tau \frac{P_m^{(|\alpha|,d)}(\tau y)}{d!}\right). \tag{4.76}$$

The proof is then completed by Lemma 4.4. □

4.2. Kernel Polynomials on the Ewens' sampling distribution.

The Ewens' sampling distribution can be obtained as a limit distribution from the unordered Dirichlet-Multinomial distribution $DM^{\downarrow}_{|\alpha|,d}$ as $d \to \infty$. The proof of the following proposition can be obtained by the same arguments used to prove Proposition 4.1.

Proposition 4.6. *(i) The polynomial kernels with respect to $DM^{\downarrow}_{|\alpha|,d}$ are of the same form as (3.42) but with $\xi^{H,(|\alpha|,d)}_{|m|}$ replaced by*

$$\xi^{H,(|\alpha|,d)\downarrow}_{|m|} := (d!)^{-1} \sum_{\pi} \xi^{H,(|\alpha|,d)}_{|m|}(\pi(r), s). \tag{4.77}$$

(ii) The Kernel polynomials with respect to $ESF_{|\alpha|}$ are derived by considering the limit form $\xi^{H,|\alpha|\downarrow}_{|m|}$ of $\xi^{H,(|\alpha|,d)\downarrow}_{|m|}$ This has the same form as $\xi^{|\alpha|\downarrow}_{|m|}$, (4.71), with $[x;b][y;b]$ replaced by $[r;b]'[s;b]'$, where

$$[r;b]' = (|\alpha| + |r|)^{-1}_{(|m|)} \sum r_{i_1(l_1)} \cdots r_{i_k(l_k)}$$

and summation is over $\sum_1^{|m|} jb_j = |m|$, $\sum_1^{|m|} b_j = k$, $k = 1, \ldots |m|$. The Kernel polynomials have the same form as (3.42) with $\xi^{H,(|\alpha|,d)}_{|m|}$ replaced by $\xi^{H,|\alpha|\downarrow}_{|m|}$. The first polynomial is identically zero under this symmetrization.

5. Integral representation for Jacobi polynomial kernels.

This section is a bridge between the first and the second part of the paper. We provide an integral representation for Jacobi and Hahn polynomial kernels, extending to $d \geq 2$ the well-known Jacobi and Hahn product formulae found by Koornwinder and Gasper's for $d = 2$ ([24], GAS72). It will be a key tool to identify, under certain conditions on the parameters, positive-definite sequences on the discrete and continuous multi-dimensional simplex.

5.1. Product formula for Jacobi polynomials when $d = 2$.

For $d = 2$, consider the shifted Jacobi polynomials normalized by their value at 1:

$$R_n^{\alpha,\beta}(x) = \frac{Q_n^{\alpha.\beta}(x,1)}{Q_n^{\alpha,\beta}(1,1)}. \tag{5.78}$$

They can also be obtained from the ordinary Jacobi polynomials $P_n^{a,b}$ $(a, b > -1)$ with Beta weight measure

$$w_{a,b} = (1-x)^a (1+x)^b dx, \qquad x \in [-1,1]$$

via the transformation:

$$R_n^{\alpha,\beta}(x) = \frac{P_n^{\beta-1,\alpha-1}(2x-1)}{P_n^{\beta-1,\alpha-1}(1)}. \tag{5.79}$$

The constant of orthogonality $\zeta_n^{(\alpha,\beta)}$ is given by (2.34).
A crucial property of Jacobi polynomials is that, under certain conditions on the parameters, products of Jacobi polynomials have an integral representation with respect to a positive (probability) measure. The following theorem is part of a more general result of Gasper [10].

Theorem 5.1. *(Gasper (1972)) A necessary and sufficient condition for the equality*

$$\frac{P_n^{a,b}(x)}{P_n^{a,b}(1)}\frac{P_n^{a,b}(y)}{P_n^{a,b}(1)} = \int_{-1}^{1} \frac{P_n^{a,b}(z)}{P_n^{a,b}(1)} \widetilde{m}_{x,y;a,b}(dz), \tag{5.80}$$

to hold for a positive measure $d\widetilde{m}_{x,y}$, is that $a \geq b > -1$ and either $b \geq 1/2$ or $a+b \geq 0$. If $a+b > -1$ or if $a > -1/2$ and $a+b = -1$ with $x \neq -y$, then $\widetilde{m}_{x,y;a,b}$ is absolutely continuous with respect to $w_{a,b}$ with density of the form

$$\frac{d\widetilde{m}_{x,y;a,b}}{dw_{a,b}}(z) = \sum_{n=0}^{\infty} \phi_n \frac{P_n^{a,b}(x)}{P_n^{a,b}(1)}\frac{P_n^{a,b}(y)}{P_n^{a,b}(1)}\frac{P_n^{a,b}(z)}{P_n^{a,b}(1)}, \tag{5.81}$$

with $\phi_n = P_n^{a,b}(1)^2/\mathbb{E}\left[P_n^{a,b}(X)\right]$.

An explicit formula for the density (5.81) is possible when $a \geq b > -1/2$:

$$\frac{P_n^{a,b}(x)}{P_n^{a,b}(1)}\frac{P_n^{a,b}(y)}{P_n^{a,b}(1)} = \int_0^1 \int_0^{\pi} \frac{P_n^{a,b}(\psi)}{P_n^{a,b}(1)} \widetilde{m}_{a,b}(du, d\omega), \tag{5.82}$$

where

$$\psi(x,y;u,\omega) = \{(1+x)(1+y) + (1-x)(1-y)\}/2 + u\cos\omega\sqrt{(1-x^2)(1-y^2)} - 1$$

and

$$\widetilde{m}_{a,b}(du, d\omega) = \frac{2\Gamma(a+1)}{\sqrt{\pi}\Gamma(a-b)\Gamma(b+\frac{1}{2})}(1-u^2)^{a-b-1}u^{2b+1}(\sin\omega)^{2b}dud\omega. \tag{5.83}$$

See [24] for an analytic proof of this formula. Note that $\phi(1,1;u,\omega) = 1$, so $d\widetilde{m}_{a,b}(u,\omega)$ is a probability measure.
Gasper's theorem can be rewritten in an obvious way in terms of the shifted Jacobi polynomials $R_n^{\alpha,\beta}(x)$ on $[0,1]$:

Corollary 5.2. *For $\alpha, \beta > 0$ the product formula*

$$R_n^{\alpha,\beta}(x)R_n^{\alpha,\beta}(y) = \int_0^1 R_n^{\alpha,\beta}(z) m_{x,y;\ \alpha,\beta}(dz) \tag{5.84}$$

holds for a positive measure $m_{x,y;\ \alpha,\beta}$ if and only if $\beta \geq \alpha$ and either $\alpha \geq 1/2$ or $\alpha + \beta \geq 2$. In this case $m_{x,y}^{(\alpha,\beta)} = \widetilde{m}_{2x-1,2y-1;\beta-1,\alpha-1}$ where $d\widetilde{m}$ is defined by (5.81). The measure is absolutely continuous if $\alpha+\beta \geq 2$ or if $\beta > 1/2$ and $\alpha + \beta > 1$ with $x \neq y$. In this case

$$m_{x,y}^{(\alpha,\beta)}(dz) = K(x,y,z)D_{\alpha,\beta}(dz)$$

where

$$K(x,y,z) = \sum_{n=0}^{\infty} \zeta_n^{\alpha,\beta} R_n^{\alpha,\beta}(x)R_n^{\alpha,\beta}(y)R_n^{\alpha,\beta}(z) \geq 0. \tag{5.85}$$

Remark 5.3. *When α, β satisfy the constraints of Corollary 5.2, we will say that α, β* satisfy Gasper's conditions.

When $\alpha \geq 1/2$, an explicit integral identity follows from (5.82)-(5.83). Let $m_{\alpha\beta}(du, d\omega) = \widetilde{m}_{\beta-1,\alpha-1}(du, d\omega)$. Then

$$R_n^{\alpha,\beta}(x)R_n^{\alpha,\beta}(y) = \int_0^1 \int_0^{\pi} R_n^{\alpha,\beta}(\varphi) m_{\alpha\beta}(du, d\omega), \tag{5.86}$$

where for $x, y \in [0,1]$

$$\varphi(x,y;u,\omega) = xy + (1-x)(1-y) + 2u\cos\omega\sqrt{x(1-x)y(1-y)}. \tag{5.87}$$

In ϕ set $x \leftarrow 2x-1, y \leftarrow 2y-1$ to obtain (5.87).

5.2. Integral representation for d > 2.

An extension of the product formula (5.84) is possible for the kernel Q_n^α for the bivariate Dirichlet of any dimension d.

Proposition 5.4. *Let $\alpha \in \mathbb{R}_+^d$ such that, for every $j = 1, \dots, d$, $\alpha_j \le \sum_{i=1}^{j-1} \alpha_i$ and $1/2 \le \alpha_j$, or $\sum_{i=1}^{j} \alpha_i \ge 2$. Then, for every $x, y \in \Delta_{(d-1)}$ and every integer $|n|$,*

$$Q_{|n|}^{\alpha}(x,y) = \mathbb{E}\left[Q_{|n|}^{\alpha_d,|\alpha|-\alpha_d}(Z_d,1) \mid x,y\right] \tag{5.88}$$

where, for every $x, y \in \Delta_{(d-1)}$, Z_d is the $[0,1]$ random variable defined by the recursion:

$$Z_1 \equiv 1; \qquad Z_j = \Phi_j D_j Z_{j-1} \qquad j = 2, \dots, d \tag{5.89}$$

with

$$D_j := \frac{(1-x_j)(1-y_j)}{(1-X_j^*)(1-Y_j^*)}; \qquad X_j^* := \frac{x_j}{1-x_j(1-\sqrt{Z}_{j-1})}; \qquad Y_j^* := \frac{y_j}{1-y_j(1-\sqrt{Z}_{j-1})} \tag{5.90}$$

where Φ_j is a random variable in $[0,1]$, with distribution

$$dm_{x_j^*, y_j^*; \alpha_j, \sum_{i=1}^{j-1} \alpha_i}$$

where $dm_{x,y;\ \alpha,\beta}$ is defined as in Corollary 5.2.

The Proposition makes it natural to order the parameters of the Dirichlet in a decreasing way, so that it is sufficient to assume that $\alpha_{(1)} + \alpha_{(2)} \ge 2$ to obtain the representation (5.88).
Since the matrix $A = \{a_{nm}\}$ is invertible, the proof of Proposition 5.4 only depends on the properties of the function ξ. The following lemma is in fact all we need.

Lemma 5.5. *For every $|m| \in \mathbb{N}, d = 2, 3, \dots$ and $\alpha \in \mathbb{R}^d$ satysfying the assumptions of Proposition 5.4,*

$$\xi_{|m|}^{\alpha}(x,y) = \frac{|\alpha|_{(|m|)}}{(\alpha_d)_{(|m|)}} \mathbb{E}\left[\ Z_d^{|m|} \mid x,y\right], \tag{5.91}$$

where Z_d is defined as in Proposition 5.4.

Let $\theta = \alpha + \beta$. Assume the Lemma is true. From (5.86) and (5.93) we know that, for every $n = 0, 1, \dots$ and every $s \in [0,1]$,

$$Q_{|n|}^{\alpha,\beta}(s,1) = \sum_{|m| \le |n|} a_{|n||m|}^{\theta} \frac{(\theta)_{(|m|)}}{\alpha_{(|m|)}} s^{|m|}.$$

Thus from (5.91)

$$\begin{aligned} Q_{|n|}^{\alpha}(x,y) &= \mathbb{E}\left[\sum_{|m| \le |n|} a_{|n||m|}^{|\alpha|} \frac{(|\alpha|)_{(|m|)}}{\alpha_{d(|m|)}} Z_d^{|m|} \,\middle|\, x,y\right] \\ &= \left[Q_{|n|}^{\alpha_d,|\alpha|-\alpha_d}(Z_d,1) \mid x,y\right] \end{aligned}$$

which is what is claimed in Proposition 5.4.
Now we proceed with the proof of the Lemma.

Proof. The proof is by induction.

If $d = 2$, $x, y \in [0, 1]$,

$$\xi_{|m|}^{(\alpha,\beta)}(x, y) = \sum_{j=0}^{|m|} \binom{|m|}{j} \frac{(\alpha+\beta)_{(|m|)}}{(\alpha)_{(j)}(\beta)_{(|m|-j)}} (xy)^j [(1-x)(1-y)]^{|m|-j}. \tag{5.92}$$

Setting $y = 1$, the only positive addend in (5.92) is the one with $j = |m|$ so

$$\xi_{|m|}^{(\alpha,\beta)}(x, 1) = \frac{(\alpha+\beta)_{(|m|)}}{(\alpha)_{(|m|)}} \ z^{|m|}. \tag{5.93}$$

Therefore, if $\theta = \alpha + \beta$, from (5.86) and (5.93) we conclude

$$\begin{aligned} \xi_{|m|}^{\alpha,\beta}(x, y) &= \sum_{j=0}^{|m|} \binom{|m|}{j} \frac{(\theta)_{(|m|)}}{(\alpha)_{(j)}(\beta)_{(|m|-j)}} (xy)^j [(1-x)(1-y)]^{|m|-j} \\ &= \frac{(\theta)_{(|m|)}}{(\alpha)_{(|m|)}} \int_{[0,1]} \ z^{|m|} m_{x,y;\ \alpha,\beta}(dz). \end{aligned} \tag{5.94}$$

Thus the proposition is true for $d = 2$.
To prove the result for any general $d > 2$, consider

$$\begin{aligned} \xi_{|m|}^{\alpha}(x, y) &= \sum_{m_d=0}^{|m|} \binom{|m|}{m_d} (x_d y_d)^{m_d} [(1-x_d)(1-y_d)]^{|m|-m_d} \frac{(|\alpha|)_{|m|}}{(\alpha_d)_{(m_d)}(|\alpha|-\alpha_d)_{(|m|-m_d)}} \\ &\times \sum_{\tilde{m}\in\mathbb{N}^{d-1}:|\tilde{m}|=|m|-m_d} \binom{|m|-m_d}{\tilde{m}} \frac{(|\alpha|-\alpha_d)_{(|m|-m_d)}}{\prod_{i=1}^{d-1}(\alpha_i)_{(\tilde{m}_i)}} \prod_{i=1}^{d-1} (\tilde{x}_i \tilde{y}_i)^{\tilde{m}_i}, \end{aligned} \tag{5.95}$$

where $\tilde{x}_i = \frac{x_i}{1-x_d}$, $\tilde{y}_i = \frac{y_i}{1-y_d}$, $(i = 1, \ldots, d-1)$.

Now assume the proposition is true for $d - 1$. Then the inner sum of (5.95) has a representation like (5.91) and we can write

$$\begin{aligned} \xi_{|m|}^{\alpha}(x, y) &= \sum_{m_d=0}^{|m|} \binom{|m|}{m_d} (x_d y_d)^{m_d} [(1-x_d)(1-y_d)]^{|m|-m_d} \frac{(|\alpha|)_{|m|}}{(\alpha_d)_{(m_d)}(|\alpha|-\alpha_d)_{(|m|-m_d)}} \\ &\times \frac{(|\alpha|-\alpha_d)_{(|m|-m_d)}}{(\alpha_{d-1})_{(|m|-m_d)}} \mathbb{E}\left[Z_{d-1}^{|m|-m_d} \ |\widetilde{x}, \widetilde{y}\right], \end{aligned} \tag{5.96}$$

where the distribution of Z_{d-1}, given $\widetilde{x}, \widetilde{y}$, depends only on $\tilde{\alpha} = (\alpha_1, \ldots, \alpha_{d-1})$. Now, set

$$\frac{X_d^*}{1-X_d^*} = \frac{x_d}{(1-x_d)\sqrt{Z_{d-1}}};$$
$$\frac{Y_d^*}{1-Y_d^*} = \frac{y_d}{(1-y_d)\sqrt{Z_{d-1}}},$$

and define the random variable

$$D_d := \frac{(1-x_d)(1-y_d)}{(1-X_d^*)(1-Y_d^*)}. \tag{5.97}$$

Then simple algebra leads to rewriting equation (5.96) as:

$$\xi^{\alpha}_{|m|}(x,y) = \mathbb{E}\left[\frac{|\alpha|_{(|m|)}(D_d Z_{d-1})^{|m|}}{(\alpha_{d-1}+\alpha_d)_{(|m|)}}\left(\sum_{m_d=0}^{|m|}\binom{|m|}{m_d}\frac{(\alpha_{d-1}+\alpha_d)_{(|m|)}}{(\alpha_d)_{(m_d)}(\alpha_{d-1})_{(|m|-m_d)}}(X_d^* X_d^*)^{m_d}[(1-X_d^*)(1-Y_d^*)]^{|m|-m_d}\right)\,\middle|\, x,y\right]. \tag{5.98}$$

Now the sum in (5.98) is of the form (5.92), with $\alpha = \alpha_{d-1}$, $\beta = \alpha_d$, with m replaced by $m - m_d$ and the pair (x, y) replaced by (x_d^*, y_d^*). Therefore we can use the equality (5.94) to obtain

$$\begin{aligned}\xi^{\alpha}_{|m|}(x,y) &= \mathbb{E}\left[\frac{(|\alpha|)_{(|m|)}}{(\alpha_d)_{(|m|)}}(D_d Z_{d-1})^{|m|}\mathbb{E}\left(\Phi_d^{|m|}\ |X_d^*, Y_d^*\right)\,\middle|\, x,y\right]\\ &= \frac{(|\alpha|)_{(|m|)}}{(\alpha_d)_{(|m|)}}\mathbb{E}\left[Z_d^{|m|}\mid x,y\right].\end{aligned} \tag{5.99}$$

(the inner conditional expectation being a function of Z_{d-1}) so the proof is complete. ☐

5.3. Connection with a multivariate product formula by Koornwinder and Schwartz.

For the individual, multivariate Jacobi polynomials orthogonal with respect to $D_\alpha : \alpha \in \mathbb{R}^d$, a product formula is proved in [27]. For every $x \in \Delta_{(d-1)}$, $\alpha \in \mathbb{R}^d_+$ and $n = (n_1, \ldots, n_{d-1}) : |n| = n$, these polynomials can be written as

$$R^{\alpha}_n(x) = \prod_{j=1}^{d-1}\left[R^{\alpha_j, E_j+2N_j}_{n_j}\left(\frac{x_j}{1-\sum_{i=1}^{j-1}x_i}\right)\right]\left(1-\frac{x_j}{1-\sum_{i=1}^{j-1}x_i}\right)^{N_j} \tag{5.100}$$

where $E_j = |\alpha| - \sum_{i=1}^{j}\alpha_i$ and $N_j = n - \sum_{i=1}^{j} n_i$. The normalization is such that $R^{\alpha}_n(e_d) = 1$, where $e_d := (0, 0, \ldots, 1) \in \mathbb{R}^d$. For an account of such polynomials see also [17].

Theorem 5.6. (Koornwinder and Schwartz) *Let $\alpha \in \mathbb{R}^d$ satisfy $\alpha_d > 1/2$ and, for every $j = 1, \ldots, d$, $\alpha_j \geq \sum_{i=j+1}^{d}\alpha_i$. Then, for every $x, y \in \Delta_{(d-1)}$ there exists a positive probability measure $dm^*_{x,y;\alpha}$ such that, for every $n \in \mathbb{N}^d_+$,*

$$R^{\alpha}_n(x)R^{\alpha}_n(y) = \int_{\Delta_{(d-1)}} R^{\alpha}_n(z)m^*_{x,y;\alpha}(dz). \tag{5.101}$$

Note that Theorem 5.6 holds for conditions on α which are stronger that our Proposition 5.4. This is the price to pay for the measure $m^*_{x,y;\alpha}$ of Koornwinder and Schwartz to have an explicit description (we omit it here), extending (5.83). It is possible to establish a relation between the measure $m^*_{x,y;\alpha}(z)$ of Theorem 5.6 and the distribution of Z_d of Proposition 5.4.

Proposition 5.7. *Let α obey the conditions of Theorem 5.6. Denote with $m_{x,y;\alpha}$ the probability distribution of Z_d of Proposition 5.4 and $m^*_{x,y;\alpha}$ the mixing measure in Theorem 5.6. Then*

$$m^*_{x,y;\alpha} = m_{x,y;\alpha}.$$

Proof. Notice that both $m^*_{x,y;\alpha}$ and $m_{x,y;\alpha}$ are absolutely continuous with respect to $D_{\alpha_d,|\alpha|-\alpha_d}$. Denote their densities with $\mu^*_{x,y;\alpha}(z) := \frac{dm^*_{x,y;\alpha}}{dD_{\alpha_d,|\alpha|-\alpha_d}}(z)$ and $\mu_{x,y;\alpha}(z) = \frac{dm_{x,y;\alpha}}{dD_{\alpha_d,|\alpha|-\alpha_d}}(z)$. From Proposition 5.4,

$$Q^{\alpha}_{|n|}(x,y) = \zeta^{\alpha_d,|\alpha|-\alpha_d}_{|n|}\mathbb{E}\left(R^{\alpha_d,|\alpha|-\alpha_d}_{|n|}(Z_d)\mu_{x,y;\alpha}(Z_d)\right).$$

Now, by uniqueness,

$$\begin{aligned}Q^{\alpha}_{|n|}(x,y) &= \sum_{|m|=|n|} Q^{\alpha}_m(x)Q^{\alpha}_m(y)\\ &= \sum_{|m|=|n|}\zeta^{\alpha}_m R^{\alpha}_m(x)R^{\alpha}_m(y)\end{aligned} \tag{5.102}$$

where $\zeta_n^\alpha := \mathbb{E}(R_n^\alpha)^{-2}$.
So, by Theorem 5.6 and because $R_n(e_d) = 1$,

$$\begin{aligned} Q^\alpha_{|n|}(x,y) &= \int \left(\sum_{|m|=|n|} \zeta^\alpha_m R^\alpha_m(z) \right) dm^*_{x,y;\alpha}(z) \\ &= \int Q^\alpha_{|n|}(z, e_d)\, dm^*_{x,y;\alpha}(z), \end{aligned} \tag{5.103}$$

where Q_n^α are orthonormal polynomials. But we know that

$$Q^\alpha_{|n|}(z, e_d) = \zeta^{\alpha_d, |\alpha|-\alpha_d}_{|n|} R^{\alpha_d,|\alpha|-\alpha_d}_{|n|}(z_d)$$

so

$$Q^\alpha_{|n|}(x,y) = \zeta^{\alpha_d,|\alpha|-\alpha_d}_{|n|} \mathbb{E}\left(R^{\alpha_d,|\alpha|-\alpha_d}_{|n|}(Z_d) \mu_{x,y;\alpha}(Z_d) \right) = \zeta^{\alpha_d,|\alpha|-\alpha_d}_{|n|} \mathbb{E}\left(R^{\alpha_d,|\alpha|-\alpha_d}_{|n|}(Z_d) \mu^*_{x,y;\alpha}(Z_d) \right). \tag{5.104}$$

Thus both $\mu_{x,y;\alpha}(z)$ and $\mu^*_{x,y;\alpha}(z)$ have the same Riesz-Fourier expansion

$$\sum_{|n|=0}^{\infty} Q^\alpha_{|n|}(x,y) R^{\alpha_d,|\alpha|-\alpha_d}_{|n|}(z)$$

and this completes the proof. □

6. Integral representations for Hahn polynomial kernels.

Intuitively it is easy now to guess that a discrete integral representation for Hahn polynomial kernels, similar to that shown by Proposition 5.4 for Jacobi kernels, should hold for any $d \geq 2$. We can indeed use Proposition 5.4 to derive such a representation. We need to reconsider formula (3.41) for Hahn polynomial in the following version:

$$\widetilde{h}^\alpha_n(m;|r|) := \int R^\alpha_n(x) D_{\alpha+r}(dx) = \frac{h^0_n(m;|r|)}{\sqrt{\zeta^\alpha_n}}, \tag{6.105}$$

with the new coefficient of orthogonality

$$\frac{1}{\omega^\alpha_{n,|r|}} := \mathbb{E}\left[\widetilde{h}^\alpha_n(M;|r|)\right]^2 = \frac{|r|_{[|n|]}}{(|\alpha|+|r|)_{(|n|)}} \frac{1}{\zeta^\alpha_n}. \tag{6.106}$$

Formula (6.105) is equivalent to

$$R^\alpha_n(x) = \frac{(|\alpha|+|r|)_{(|n|)}}{|r|_{[|n|]}} \sum_{|m|=|r|} \widetilde{h}^\alpha_{|n|}(m;|r|) \binom{|r|}{m} x^m, \qquad \alpha \in \mathbb{R}^d, x \in \Delta_{(d-1)} \tag{6.107}$$

(see [17], 5.2.1 for a proof).

Proposition 6.1. *For $\alpha \in \mathbb{R}^d$ satisfying the same conditions as in Proposition 5.4, a representation for the Hahn polynomial kernels is:*

$$H^\alpha_{|n|}(r,s) = \omega^\alpha_{n,|r|} \frac{(|\alpha|+|r|)_{(|n|)}}{|r|_{[|n|]}} \mathbb{E}_{r,s}\left[\widetilde{h}^{\alpha_d,|\alpha|-\alpha_d}_{|n|}(K;|r|)\right], \qquad |n| \leq |r| = |s|, \alpha \in \mathbb{R}^d, \tag{6.108}$$

where the expectation is taken with respect to the measure:

$$u_{r,s;\alpha}(k) := \int_{\Delta_{(d-1)}} \int_{\Delta_{(d-1)}} \mathbb{E}\left[\binom{|r|}{k} Z_d^k (1-Z_d)^{|r|-k} \mid x, y\right] D_{\alpha+r}(dx) D_{\alpha+s}(dy), \tag{6.109}$$

where Z_d, for every x, y, is the random variable defined recursively as in Proposition 5.4.

Proof. From (3.41),

$$H^{\alpha}_{|n|}(r,s) = \frac{(|\alpha|+|r|)_{(|n|)}}{|r|_{[|n|]}} \int_{\Delta_{(d-1)}} \int_{\Delta_{(d-1)}} Q^{\alpha}_{|n|}(x,y) D_{\alpha+r}(dx) D_{\alpha+s}(dy).$$

Then (5.88) implies

$$H^{\alpha}_{|n|}(r,s) = \frac{\zeta^{\alpha_d,|\alpha|-\alpha_d}_{|n|}(|\alpha|+|r|)_{(|n|)}}{|r|_{[|n|]}} \int_{\Delta_{(d-1)}} \int_{\Delta_{(d-1)}} \int_0^1 R^{\alpha_d,|\alpha|-\alpha_d}_{|n|}(z_d) m_{x,y;\alpha}(dz_d) D_{\alpha+r}(dx) D_{\alpha+s}(dy)$$

so by (6.107)

$$\begin{aligned} H^{\alpha}_{|n|}(r,s) &= \zeta^{\alpha_d,|\alpha|-\alpha_d}_{|n|} \left(\frac{(|\alpha|+|r|)_{(|n|)}}{|r|_{[|n|]}} \right)^2 \\ &\times \sum_{|k|\le|r|} \widetilde{h}^{\alpha_d,|\alpha|-\alpha_d}_{|n|}(k;|r|) \int_{\Delta_{(d-1)}} \int_{\Delta_{(d-1)}} \int_0^1 \binom{|r|}{k} z_d^k (1-z_d)^{|r|-k} m_{x,y;\alpha}(dz_d) D_{\alpha+r}(dx) D_{\alpha+s}(dy), \end{aligned}$$

and the proof is complete. □

7. Positive-definite sequences and polynomial kernels.

We can now turn our attention to the problem of identifying and possibly characterizing positive-definite sequences with respect to the Dirichlet or Dirichlet-Multinomial probability distribution. We will agree with the following definition which restricts the attention to multivariate positive-definite sequences $\{\rho_n : n \in \mathbb{Z}^d_+\}$ which depend on n only via $|n|$.

Definition 7.1. *For every $d \ge 2$ and $\alpha \in \mathbb{R}^d_+$, call a sequence $\{\rho_{|n|}\}_{|n|=0}^{\infty}$ an α-Jacobi positive-definite sequence (α-JPDS) if $\rho_0 = 1$ and, for every $x, y \in \Delta_{(d-1)}$,*

$$p(x,y) = \sum_{|n|=0}^{\infty} \rho_{|n|} Q^{\alpha}_{|n|}(x,y) \ge 0. \tag{7.110}$$

For every $d \ge 2$, $\alpha \in \mathbb{R}^d_+$ and $|r| \in \mathbb{Z}_+$, call a sequence $\{\rho_{|n|}\}_{|n|=0}^{\infty}$ an α-Hahn positive-definite sequence (α-HPDS) if $\rho_0 = 1$ and, for every $r, s \in |r|\Delta_{(d-1)}$,

$$p^H(r,s) = \sum_{|n|=0}^{\infty} \rho_{|n|} H_{|n|}(r,s) \ge 0. \tag{7.111}$$

7.1. Jacobi Positivity from the integral representation.

A consequence of the product formulae (5.84) and (5.86) is a characterization of positive-definite sequences for the Beta distribution.
The following is a $[0,1]$-version of a theorem proved by Gasper with respect to Beta measures on $[-1,1]$.

Theorem 7.2. (Bochner [4]), Gasper [10]). *Let $D_{\alpha,\beta}$ be the Beta distribution on $[0,1]$ with $\alpha \le \beta$. If either $1/2 \le \alpha$ or $\alpha+\beta \ge 2$, then a sequence ρ_n is positive-definite for $D_{\alpha,\beta}$ if and only if*

$$\rho_n = \int R^{\alpha,\beta}_n(z) \nu_{\alpha,\beta}(z) \tag{7.112}$$

for a positive measure ν with support on $[0,1]$. Moreover, if

$$u(x) = \sum_{n=0}^{\infty} \zeta_n^{\alpha,\beta} \rho_n R_n(x) \geq 0$$

with

$$\sum_{n=0}^{\infty} \zeta_n^{\alpha,\beta} |\rho_n| < \infty,$$

then

$$\nu(A) = \int_A u(x) D_{\alpha,\beta}(dx) \tag{7.113}$$

for every Borel set $A \subseteq [0,1]$.

We refer to [4], [10] for the technicalities of the proof. To emphasize the key role played by (5.84), just observe that the positivity of ν and (7.112) entails the representation

$$p(x,y) := \sum_{n=0}^{\infty} \zeta_n \rho_n R_n^{\alpha,\beta}(x) R_n^{\alpha,\beta}(y) = \int_0^1 u(z) m_{x,y;\alpha,\beta}(dz) \geq 0,$$

and $u(z) = p(z,1)$, whenever $u(1)$ is absolutely convergent.

To see the full extent of the characterization, we recall, in a Lemma, an important property of Jacobi polynomials, namely: two different systems of Jacobi polynomials are connected by an integral formula if their parameters share the same total sum.

Lemma 7.3. *For $\mu > 0$,*

$$\int_0^1 R_n^{\alpha,\beta}(1-(1-x)z)\ D_{\beta,\mu}(dz) = R_n^{\alpha-\mu,\beta+\mu}(x) \tag{7.114}$$

and

$$\int_0^1 R_n^{\alpha,\beta}(xz)\ D_{\alpha,\mu}(dz) = \frac{\zeta_n^{\alpha+\mu,\beta-\mu}}{\zeta_n^{\alpha,\beta}} R_n^{a+\mu,b-\mu}(x). \tag{7.115}$$

Proof. We provide here a probabilistic proof in terms of polynomial kernels $Q_{|n|}^{\alpha,\beta}(x,y)$, even though the two integrals can also be view as a reformulation, in terms of the shifted polynomials $R_n^{\alpha,\beta}$, of known integral representations for the Jacobi polynomials $\{P_n^{a,b}\}$ on $[-1,1]$ $(a,b>-1)$ (see e.g. AA 7.392.3 and 7.392.4).

Let us start with (7.115). The moments of a Beta (α,β) distribution on $[0,1]$ are, for every integer $m \leq n = 0,1,\ldots$

$$\mathbb{E}[X^m(1-X)^{n-m}] = \frac{\alpha_{(m)}\beta_{(n-m)}}{(\alpha+\beta)_{(n)}}.$$

Now, for every $n \in \mathbb{N}$,

$$\begin{aligned}
\int_0^1 \zeta_n^{\alpha,\beta} R_n^{\alpha,\beta}(xz) D_{\alpha,\mu}(dz) &= \int_0^1 Q_n^{\alpha,\beta}(xz,1) D_{\alpha,\mu}(dz) \\
&= \sum_{m\leq n} a_{nm} \frac{(\alpha+\beta)_{(m)}}{(\alpha)_{(m)}} \int_0^1 (xz)^m D_{\alpha,\mu}(dz) \\
&= \sum_{m\leq n} a_{nm} \frac{(\alpha+\beta)_{(m)}}{(\alpha)_{(m)}} \frac{(\alpha)_{(m)}}{(\alpha+\mu)_{(m)}} x^m = \zeta_n^{\alpha+\mu,\beta-\mu} R_n^{a+\mu,b-\mu}(x),
\end{aligned} \tag{7.116}$$

and this proves (7.115).

To prove (7.114), simply remember (see e.g. [17], 3.1) that

$$R_n^{\alpha,\beta}(0) = (-1)^n \frac{\alpha_{(n)}}{\beta_{(n)}}$$

and that

$$R_n^{\alpha,\beta}(x) = \frac{R_n^{\beta,\alpha}(1-x)}{R_n^{\beta,\alpha}(0)}.$$

So we can use (7.115) to see that

$$\int_0^1 \frac{R_n^{\beta,\alpha}((1-x)z)}{R_n^{\beta,\alpha}(0)} D_{\beta,\mu}(dz) \quad = \quad (-1)^n \frac{\alpha_{(n)}}{\beta_{(n)}} \frac{\zeta_n^{\beta+\mu,\alpha-\mu}}{\zeta_n^{\beta,\alpha}} R_n^{\beta+\mu,\alpha-\mu}(1-x) = \zeta_n^{\alpha-\mu,\beta+\mu}(x), \tag{7.117}$$

and the proof is complete. □

Lemma 7.3 completes Theorem 7.2:

Corollary 7.4. *Let $\alpha \leq \beta$ with $\alpha + \beta \geq 2$. If a sequence ρ_n is positive-definite for $D_{\alpha,\beta}$, then it is positive-definite for $D_{\alpha+\mu,\beta-\mu}$, for any $0 \leq \mu \leq \beta$.*

Proof. By Theorem 7.2 ρ_n is positive-definite for $D_{\alpha,\beta}$ if and only if

$$\sum_n \zeta_n^{\alpha,\beta} \rho_n R_n^{\alpha,\beta}(x) \geq 0.$$

So (7.115) implies also

$$\sum_n \zeta_n^{\alpha,\beta} \rho_n \frac{\zeta_n^{\alpha+\mu,\beta-\mu}}{\zeta_n^{\alpha,\beta}} R_n^{\alpha+\mu,\beta-\mu}(x) \geq 0.$$

The case for $D_{\alpha-\mu,\beta+\mu}$ is proved similarly, but using (7.114) instead of (7.115).

□

For $d > 2$ Proposition 5.4 leads to a similar characterization of all positive-definite sequences, for the Dirichlet distribution, which are indexed only by their total degree, i.e. all sequences $\rho_n = \rho_{|n|}$.

Proposition 7.5. *Let $\alpha \in R^d$ satisfy the same conditions as in Proposition 5.4. A sequence $\{\rho_n = \rho_{|n|} : n \in \mathbb{N}\}$ is positive-definite for the Dirichlet (α) distribution if and only if it is positive-definite for $D_{c|\alpha|,(1-c)|\alpha|}$, for every $c \in (0,1)$.*

Proof. Sufficiency. First notice that, since

$$Q_n^{\alpha,\beta}(x,y) = Q_n^{\beta,\alpha}(1-x, 1-y), \tag{7.118}$$

then a sequence is positive-definite for $D_{\alpha,\beta}$ if and only if it is positive definite for $D_{\beta,\alpha}$, so that we can assume, without loss of generality, that $c|\alpha| \leq (1-c)|\alpha|$. Let $\alpha = (\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_d)$ satisfy the conditions of Proposition 5.4 (again, the decreasing order is assumed for simplicity) and let

$$\sum_{|n|=0}^{\infty} \rho_{|n|} Q_{|n|}^{c|\alpha|,(1-c)|\alpha|}(u,v) \geq 0 \qquad\qquad u, v \in [0,1].$$

If $\alpha_d > c_{|\alpha|}$ then Corollary 7.112, applied with $\mu = \alpha_d - c_{|\alpha|}$ implies that

$$\sum_{|n|=0}^{\infty} \rho_{|n|} Q_{|n|}^{\alpha_d,|\alpha|-\alpha_d}(u,v) \geq 0$$

so by Proposition 5.4

$$0 \leq \int \left[\sum_{|n|=0}^{\infty} \rho_{|n|} Q_{|n|}^{\alpha_d,|\alpha|-\alpha_d}(z_d, 1) \right] m_{x,y;\alpha}(dz_d) = \sum_{|n|=0}^{\infty} \rho_{|n|} Q_{|n|}^{\alpha}(x,y) \qquad\qquad x, y \in \Delta_{(d-1)}. \tag{7.119}$$

If $\alpha_d < c_{|\alpha|}$, then apply Corollary 7.112 with $\mu = |\alpha|(1-c) - \alpha_d$ to obtain

$$\sum_{|n|=0}^{\infty} \rho_{|n|} Q_{|n|}^{|\alpha|-\alpha_d,\alpha_d}(u,v) = \sum_{|n|=0}^{\infty} \rho_{|n|} Q_{|n|}^{\alpha_d,|\alpha|-\alpha_d}(1-u,1-v) \geq 0.$$

which implies again (7.119), thus $\{\rho_{|n|}\}$ is positive-definite for D_α.

Necessity. For $I \subseteq \{1,\ldots,d\}$, the random variables

$$X_I = \sum_{j\in I} X_j; \qquad Y_I = \sum_{j\in I} Y_j$$

have a Beta$(\alpha_I, |\alpha| - \alpha_I)$ distribution, where $\alpha_I = \sum_{j\in I}\alpha_j$. Since

$$E(Q_n^\alpha(X,Y)|Y_I = z) = Q_n^{\alpha_I}(z),$$

then for arbitrary $x, y \in \Delta_{(n-1)}$,

$$\sum_{|n|=0}^{\infty} \rho_{|n|} Q_{|n|}^{\alpha}(x,y) \geq 0$$

implies

$$Q_{|n|}^{\alpha_I,|\alpha|-\alpha_I}(x,y) = Q_{|n|}^{|\alpha|-\alpha_I,\alpha_I}(1-x,1-y) \geq 0.$$

Now we can apply once again Corollary 7.112 with $\mu = \pm(c|\alpha| - \alpha_I)$ (whichever is positive) to obtain, with the possible help of (7.118),

$$\sum_{|n|=0}^{\infty} \rho_{|n|} Q_{|n|}^{c|\alpha|,(1-c)|\alpha|}(u,v) \geq 0 \qquad u,v \in [0,1].$$

□

7.2. Positive definite sequences in the Dirichlet-Multinomial distribution.

In this Section we aim to investigate the relationship existing between JPDS and HPDS . In particular, we wish to understand when (P2) is true, i.e. when a sequence is both HPDS and JPDS for a given α. It turns out that, by using the results in Sections 3 and 6, it is possible to define several (sometimes striking) mappings from JPDS and HPDS and viceversa, but we could prove (P2) only for particular subclasses of positive-definite sequences. In Proposition 7.8 we prove that every JPDS is a limit of (P2)-sequences. Later in Proposition 8.6 we will identify another (P2)-family of positive-definite sequences, as a proper subfamily of the JPDSs, derived in Section 8 as the image, under a specific bijection, of a probability on $\mathbb{Z}_+$.

The first proposition holds with no constraints on α or d.

Proposition 7.6. *For every d and $\alpha \in \mathbb{R}_+^d$, let $\rho = \{\rho_{|n|}\}$ be a α-JPDS. Then*

$$\rho_{|n|}\frac{N_{[|n|]}}{(|\alpha|+N)_{(|n|)}}, \qquad n = 0,1,2,\ldots \tag{7.120}$$

is a positive-definite sequence for $DM_\alpha(\cdot;N)$ for every $N = 1,2,\ldots$.

Proof. From Proposition 3.1, if

$$\sum_{|n|=0}^{\infty} \rho_{|n|} Q^{\alpha}_{|n|}(x,y) \geq 0$$

then for every $r, s \in \mathbb{N}^d : |r| = |s|$,

$$\sum_{|n|=0}^{\infty} \rho_{|n|} \int\int Q^{\alpha}_{|n|}(x,y) D_{\alpha+r}(dx) D_{\alpha+s}(dy) = \sum_{|n|=0}^{\infty} \rho_{|n|} \frac{|r|_{[|n|]}}{(|\alpha|+|r|)_{(|n|)}} H^{\alpha}_{|n|}(r,s) \geq 0.$$

□

Two important HPDSs are given in the following Lemma.

Lemma 7.7. *For every d, every $|m| \leq N$ and every $\alpha \in \mathbb{R}^d_+$, both sequences*

$$\left\{ \frac{|m|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}} \frac{(|\alpha|+N)_{(|n|)}}{N_{[|n|]}} \right\}_{|n| \in \mathbb{Z}_+} \tag{7.121}$$

and

$$\left\{ \frac{|m|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}} \right\}_{|n| \in \mathbb{Z}_+} \tag{7.122}$$

are α-HPDSs for $DM_\alpha(\cdot; N)$.

Proof. From Proposition 3.5, by inverting (3.58) we know that, for $|m| = 0, \ldots, N$

$$0 \leq \chi^{H,\alpha}_{|m|} = \sum_{|n|=0}^{|m|} \frac{|m|_{[|n|]}}{(|\alpha|+m)_{(|n|)}} \frac{(|\alpha|+N)_{(|n|)}}{|N|_{[|n|]}} H^{\alpha}_{|n|}$$

so

$$\left\{ \frac{|m|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}} \frac{(|\alpha|+N)_{(|n|)}}{|N|_{[|n|]}} \right\}$$

is a HPDS.
Now let $\widetilde{\rho}_{|n|}$ be a JPDS. By proposition 7.6, the sequence

$$\left\{ \widetilde{\rho}_{|n|} \frac{|N|_{[|n|]}}{(|\alpha|+N)_{(|n|)}} \right\}$$

is α-HPDS. By multiplication,

$$\left\{ \widetilde{\rho}_{|n|} \frac{N_{[|n|]}}{(|\alpha|+N)_{(|n|)}} \frac{|m|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}} \frac{(|\alpha|+N)_{(|n|)}}{N_{[|n|]}} \right\} = \left\{ \widetilde{\rho}_{|n|} \frac{|m|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}} \right\}$$

is HPDS as well. This also implies that

$$\left\{ \frac{|m|_{[|n|]}}{(|\alpha|+|m|)_{(|n|)}} \right\}$$

is HPDS (to convince oneself, take $\widetilde{\rho}$ as in Example 8.3 or in Example 8.4 and take the limit as $t \to 0$ or $z \to 1$ respectively). □

We are now ready for our first result on (P2)-sequences.

Proposition 7.8. *For every d and $\alpha \in \mathbb{R}^d_+$, let $\rho = \{\rho_{|n|}\}$ be a α-JPDS. Then there exist a sequence $\{\rho^N_{|n|} : |n| \in \mathbb{Z}_+\}_{N=0}^{\infty}$ such that,* *(i) for every $|n|$,*

$$\rho_{|n|} = \lim_{N\to\infty} \rho^N_{|n|}$$

(ii) for every N, the sequence $\{\rho^N_{|n|}\}$ is both HPDS and JPDS.

Proof. We show the proof for $d = 2$. For $d > 2$ the proof is essentially the same, with all distributions obviously replaced by their multivariate versions. Take I, J two independent $DM_{\alpha,\beta}(\cdot; N)$ and $DM_{\alpha,\beta}(\cdot; M)$ random variables. As a result of de Finetti's Representation Theorem, conditionally on the event $\{\lim_{N\to\infty}(\frac{I}{N}\frac{J}{M}) = (x,y)\}$, the (I,J) are independent Binomial r.v.'s with parameter (N,x) and (M,y), respectively.
Let $f : [0,1]^2 \to \mathbb{R}$ be a positive continuous function. The function

$$B_{N,M}f(x,y) := \mathbb{E}\left[f\left(\frac{I}{N}, \frac{J}{M}\right) \mid x, y\right], \qquad N, M = 0, 1, \ldots$$

is positive as well and as $N, M \to \infty$,

$$B_{N,M}f(x,y) \longrightarrow f(x,y).$$

Now take

$$p_\rho(x,y) = \sum_{|n|} \rho_{|n|} Q^{\alpha,\beta}_{|n|}(x,y) \geq 0$$

for every $x, y \in [0,1]$. Then, for X, Y independent $D_{\alpha,\beta}$,

$$\begin{aligned}
\rho_{|n|} &= \mathbb{E}\left[Q^{\alpha,\beta}_{|n|}(X,Y)p_\rho(X,Y)\right] \\
&= \mathbb{E}\left[Q^{\alpha,\beta}_{|n|}(X,Y) \lim_{N\to\infty} B_{N,N}p_\rho(X,Y)\right] \\
&= \lim_{N\to\infty} \mathbb{E}\left[Q^{\alpha,\beta}_{|n|}(X,Y)B_{N,N}p_\rho(X,Y)\right] \\
&= \lim_{N\to\infty} \rho^N_{|n|},
\end{aligned}$$

where

$$\rho^N_{|n|} := \mathbb{E}\left[Q^{\alpha,\beta}_{|n|}(X,Y)B_{N,N}p_\rho(X,Y)\right].$$

But $B_{N,N}p_\rho$ is positive, so (i) is proved.
Now rewrite

$$\begin{aligned}
\rho^N_{|n|} &= \int_0^1\int_0^1 \sum_{i=1}^N\sum_{j=1}^N p_\rho\left(\frac{i}{N}, \frac{j}{N}\right) Q^{\alpha,\beta}_{|n|}(x,y)x^i(1-x)^{N-i}\binom{N}{i}y^j(1-y)^{N-j}D_\alpha(dx)D_\alpha(dy) \\
&= \sum_{i=1}^N\sum_{j=1}^N DM_\alpha(i;N)DM_\alpha(j;N)p_\rho\left(\frac{i}{N}, \frac{j}{N}\right)\mathbb{E}\left[Q^{\alpha,\beta}_{|n|}(X,Y) \mid i,j\right] \\
&= \frac{N_{[n]}}{(\alpha+\beta+N)_{(|n|)}}\mathbb{E}\left[p_\rho\left(\frac{I}{N}, \frac{J}{N}\right)H^{\alpha,\beta}_{|n|}(I,J)\right] \qquad (7.123)
\end{aligned}$$

for I, J are independent $DM_{\alpha,\beta}(\cdot; N)$ random variables. The last equality follows from (3.41). Since p_ρ is positive, from (7.123) it follows that

$$\left\{\rho^N_{|n|}\frac{(\alpha+\beta+N)_{(|n|)}}{N_{[n]}}\right\}$$

is, for every N, α-HPDS. But by Lemma 7.7, we can multiply every term of the sequence by the HPDS (7.122) where we set $|m| = N$, to obtain (ii).

□

The next Proposition shows some mappings from Hahn to Jacobi PDSs. It is in some sense a converse of Proposition 7.6 under the usual (extended) Gasper constraints on α.

Proposition 7.9. *If α satisfies the conditions of Proposition 5.4, let $\{\rho_{|n|}\}$ be α-HPDS for some integer N. Then both $\{\rho_{|n|}\}$ and (7.120) are positive-definite for D_α.*

Proof. If

$$\sum_{|n|=0}^{\infty} \rho_{|n|} H_{|n|}^{\alpha}(r,s) \geq 0$$

for every $r, s \in N\Delta_{(d-1)}$, then Proposition 3.2 implies that

$$\sum_{|n|=0}^{\infty} \rho_{|n|} H_{|n|}^{\alpha_1,|\alpha|-\alpha_1}(r_1, N) \geq 0.$$

Now consider the Hahn polynomials re-normalized so that

$$\widetilde{h}_{|n|}^{\alpha_1,|\alpha|-\alpha_1}(r;N) = \int_0^1 \frac{Q_{|n|}^{\alpha_1,|\alpha|-\alpha_1}(x,1)}{Q_{|n|}^{\alpha_1,|\alpha|-\alpha_1}(1,1)} D_{\alpha+r}(dx)$$

Then it is easy to prove that

$$\widetilde{h}_{|n|}^{\alpha_1,|\alpha|-\alpha_1}(N;N) = 1$$

and

$$\mathbb{E}\left[\widetilde{h}_{|n|}^{\alpha_1,|\alpha|-\alpha_1}(R;N)\right]^2 = \frac{N_{[|n|]}}{(|\alpha|+N)_{(|n|)}} \frac{1}{\zeta_{|n|}^{\alpha_1,|\alpha|-\alpha_1}}, \qquad |n| = 0, 1, \ldots$$

(see also [17], (5.65).) Hence

$$\begin{aligned} 0 &\leq \sum_{|n|=0}^{\infty} \rho_{|n|} H_{|n|}^{\alpha_1,|\alpha|-\alpha_1}(r_1, N) \\ &= \sum_{|n|=0}^{\infty} \rho_{|n|} \frac{(|\alpha|+N)_{(|n|)}}{N_{[|n|]}} \zeta_{|n|}^{\alpha_1,|\alpha|-\alpha_1} \widetilde{h}_{|n|}^{\alpha_1,|\alpha|-\alpha_1}(r_1;N) =: f_N(r). \end{aligned}$$

So for every $|n|$,

$$\begin{aligned} \rho_{|n|} &= \mathbb{E}\left[f_N(R)\widetilde{h}_{|n|}^{\alpha_1,|\alpha|-\alpha_1}(R;N)\right] \\ &= \int_0^1 \phi_N(x) R_{|n|}(x) D_\alpha(dx) \end{aligned} \tag{7.124}$$

where

$$\phi_N(x) = \sum_{r=0}^{N} \binom{N}{r} x^r (1-x)^{N-r} f_N(r) \geq 0$$

hence, by Gasper's theorem (Theorem 7.2) $\rho_{|n|}$ is $(\alpha_1, |\alpha|-\alpha_1)$-JPDS. Therefore, by Proposition 7.5, it is also α-JPDS. Finally, from the form of $\xi_{|m|}^{\alpha}$ we know that

$$|r|_{[|n|]}/(|\alpha|+|r|)_{(|n|)} = \widehat{\xi_{|N|}^{\alpha}}(n)$$

is α-JPDS, thus (7.120) is JPDS.

□

Remark 7.10. *Notice that*

$$\frac{|r|_{[|n|]}}{(|\alpha|+|r|)_{(|n|)}}$$

is itself a positive-definite sequence for D_α. This is easy to see directly from the representation (2.23) of $\xi^\alpha_{|m|}$ (we will consider more of it in Section 8).
Since products of positive-definite sequences are positive definite sequences, then we have, as a completion to all previous results,

If $\{\rho_{|n|}\}$ is positive-definite for D_α, then (7.120) positive-definite for both D_α and DM_α.

8. A probabilistic derivation of Jacobi positive-definite sequences.

In the previous sections we have found characterizations of Dirichlet positive-definite sequences holding only if the parameters satisfied a particular set of constraints. Here we show some sufficient conditions for a sequence to be α-JPDS, not requiring any constraints on α. This is done by exploiting the probabilistic interpretation of the Orthogonal Polynomial kernels. Let us reconsider the function $\xi^\alpha_{|m|}$. From (2.21) we can write, for every $|m|\in\mathbb{N}$, and $x,y\in\Delta_{d-1}$,

$$\xi_{|m|}(x,y)D_\alpha(dy)=\sum_{|l|=|m|}\binom{|m|}{l}x^l D_{\alpha+l}(dy) \tag{8.125}$$

This is, for every $|m|$, a transition kernel, expressed as a mixture of posterior Dirichlet distributions, with multinomial mixing measure. Similarly, because of the symmetry of $\xi^\alpha_{|m|}$, for any fixed $|m|$, the bivariate measure

$$\mathcal{BD}_{\alpha,|m|}(dx,dy):=\xi_{|m|}(x,y)D_\alpha(dx)D_\alpha(dy) \tag{8.126}$$

so $\xi^\alpha_{|m|}(x,y)$ has the interpretation as a (exchangeable) copula for the joint law of two vectors (X,Y), with identical Dirichlet marginal distribution. Such a joint law can be simulated via the following Gibbs sampling scheme:

(i) Generate a vector X of Dirichlet(α) random frequencies on d points.
(ii) Conditional on the observed $X=x$, sample $|m|$ iid observations with common law x.
(iii) Given the vector l of counts obtained at step (ii), take Y as stochastically independent of X and with distribution $D_{\alpha+l}(dy)$.

The bivariate measure $\mathcal{BD}_{\alpha,|m|}$ and its infinite-dimensional extension has found several applications in Bayesian Statistics e.g. by [31], but no connections were made with orthogonal kernel and canonical correlation sequences. A recent important development of this direction is in [33].

Now, let us allow the number $|m|$ in the above procedure to be random, i.e. for a probability distribution $\{d_{|m|}:|m|=0,1,2,\ldots\}$ on $\mathbb{N}$ we modify step (ii) to

(ii)' Conditional on the observed $X=x$, with probability $d_{|m|}$ sample $|m|$ iid observations with common law x.
Then the above Gibbs-sampling procedure leads us to a new exchangeable joint distribution, with identical Dirichlet marginals and copula given by

$$\mathcal{BD}_{\alpha,d}(dx,dy)=\mathbb{E}\left[\mathcal{BD}_{\alpha,|M|}(dx,dy)\right]=\sum_{|m|=0}^{\infty}d_{|m|}\xi^\alpha_{|m|}(x,y)D_\alpha(dx)D_\alpha(dy). \tag{8.127}$$

The probabilistic construction has just led us to prove the following

Proposition 8.1. *Let $\{d_{|m|} : m = 0, 1, \ldots\}$ be a probability measure on $\{0, 1, 2, \ldots\}$. Suppose that, for every $|\theta| \geq 0$, the series*

$$\rho_{|n|} = \sum_{|m| \geq |n|} \frac{|m|_{[|n|]}}{(|\theta| + |m|)_{(|n|)}} d_{|m|}, \qquad |n| = 0, 1, 2, \ldots \tag{8.128}$$

converges. Then $\{\rho_{|n|}\}$ is α-JPDS for every d and every $\alpha \in \mathbb{R}^d$ such that $|\alpha| = |\theta|$.

Proof. Note that

$$\rho_0 = \sum_{|m|=0}^{\infty} d_{|m|} = 1$$

is always true for every probability measure $\{d_{|m|}\}$.
Now reconsider the form (2.23) for the (positive) function $\xi^{\alpha}_{|m|}$: we can rewrite (8.127) as

$$\begin{aligned} 0 &\leq \sum_{|m|=0}^{\infty} d_{|m|} \xi^{\alpha}_{|m|}(x, y) \\ &= \sum_{|m|=0}^{\infty} d_{|m|} \sum_{|n| \leq |m|} \frac{|m|_{[|n|]}}{(|\theta| + |m|)_{(|n|)}} Q^{\alpha}_{|n|}(x, y) \\ &= \sum_{|n|=0}^{\infty} \left[\sum_{|m| \geq |n|} \frac{|m|_{[|n|]}}{(|\theta| + |m|)_{(|n|)}} d_{|m|} \right] Q^{\alpha}_{|m|}(x, y) \\ &= \sum_{|n|=0}^{\infty} \rho_{|n|} Q^{\alpha}_{|m|}(x, y), \end{aligned} \tag{8.129}$$

and since (8.129) does not depend on the dimension of α, then the proposition is proved for D_{α}. □

Example 8.2. *Take $d_{|m|} = \delta_{|m||l|}$, the probability assigning full mass to $|l|$. Then*

$$\rho_{|n|} = \sum_{|m| \geq |n|} \frac{|m|_{[|n|]}}{(|\theta| + |m|)_{(|n|)}} \delta_{|m||l|} = \frac{|l|_{[|n|]}}{(|\theta| + |l|)_{(|n|)}} \mathbb{I}(|l| \geq |n|). \tag{8.130}$$

and by Proposition 2.1,

$$\sum_{|n|=0}^{\infty} \rho_{|n|} Q^{\alpha}_{|n|}(x, y) = \sum_{|n|=0}^{|l|} \frac{|l|_{[|n|]}}{(|\theta| + |l|)_{(|n|)}} Q^{\alpha}_{|n|}(x, y) = \xi^{\alpha}_{|l|}(x, y). \tag{8.131}$$

Example 8.3. *Consider, for every $t \geq 0$, the probability distribution*

$$d_{|m|}(t) = \sum_{|n| \geq |n|} a^{|\alpha|}_{|m||n|} e^{-\frac{1}{2}|n|(|n|+|\alpha|-1)t}, \quad |m| = 0, 1, 2 \ldots \tag{8.132}$$

where $(a^{|\alpha|}_{|m||n|})$ is the invertible triangular system (2.20) defining the polynomial kernels $Q^{\alpha}_{|n|}$ in Proposition 2.1. Since the coefficients of the inverse system are exactly of the form

$$\frac{|m|_{[|n|]}}{(|\theta| + |m|)_{(|n|)}}, \quad |m|, |n| = 0, 1, 2, \ldots$$

then

$$\rho_{|n|}(t) = e^{-\frac{1}{2}|n|(|n|+|\alpha|-1)t}$$

is for every t a positive-definite sequence. In particular, it is the one characterizing the neutral Wright-Fisher diffusion in Population Genetics, mentioned in section 1.1., whose generator has eigenvalues $-\frac{1}{2}|n|(|n|+|\alpha|-1)$ and orthogonal polynomial eigenfunctions.
The distribution (8.132) is the so-called coalescent lineage distribution *(see [15],[16]), i.e. the probability distribution of the number of lineages surviving up to time t back in the past, when the total mutation rate is $|\alpha|$ and the allele frequencies of d phenotypes in the whole population are governed by $A_{|\alpha|,d}$.*

Example 8.4 (Perfect independence and dependence)**.** *Extreme cases of perfect dependence or perfect independence can be obtained from Example 8.3 when we take the limit as $t \to 0$ or $t \to \infty$, respectively. In the former case, $d_{|m|}(0) = \delta_{|m|\infty}$ so that $\rho_{|n|}(0) = 1$ for every $|n|$. The corresponding bivariate distribution is such that*

$$\mathbb{E}_0(Q_n(Y)|X = x) = Q_n(x)$$

so that, for every square-integrable function

$$f = \sum_n c_n Q_n$$

we have

$$\mathbb{E}_0(f(Y)|X = x) = \sum_n c_n Q_n(x) = f(x)$$

that is, $\mathcal{BD}_{\alpha,\{0\}}$ is in fact the Dirac measure $\delta(y - x)$.
In the latter case, $d_{|m|}(\infty) = \delta_{|m|0}$ so that $\rho_{|n|}(\infty) = 0$ for every $|n| > 1$ and $\mathbb{E}_0(Q_n(Y)|X = x) = \mathbb{E}[Q_n(Y)]$ implying that

$$\mathbb{E}_\infty(f(Y)|X = x) = \mathbb{E}[f(Y)]$$

i.e. X, Y are stochastically independent.

8.1. The infinite-dimensional case.

Proposition 8.1 also extends to Poisson-Dirichlet measures. The argument and construction are the same, once one replaces $\xi^{\alpha}_{|m|}$ with $\xi^{\downarrow|\theta|,\infty}_{|m|}$. We only need to observe that because the functions

$$\binom{|m|}{l} \sharp(l)[x, l],$$

forming the terms in $\xi^{\downarrow|\theta|,\infty}_{|m|}$ (see (4.66)), are probability measures on $|m|\Delta^{\downarrow}_{\infty}$, then the kernel

$$\xi^{\downarrow|\theta|,\infty}_{|m|}(x, y) D^{\downarrow}_{|\theta|,\infty}(dy)$$

defines, for every x, a proper transition probability function on $\Delta^{\downarrow}_{\infty}$, allowing for the Gibbs-sampling interpretation as in Section 8, but modified as follow:

(i) Generate a point X in $\Delta^{\downarrow}_{\infty}$ with distribution $PD(|\theta|)$.
(ii) Conditional on the observed $X = x$, sample a partition of $|m|$ with distribution function $\binom{|m|}{l}\sharp(l)[x, l]$.
(iii) Conditionally on the parition l obtained at step (ii), take Y as stochastically independent of X and with distribution

$$\frac{\binom{|m|}{l}\sharp(l)[x, l] PD_\theta(dy)}{ESF_{|\theta|}(l)}$$

.

Thus the proof of the following statement is now obvious.

Proposition 8.5. *Let $\{d_{|m|} : m = 0, 1, \dots\}$ be a probability measure on $\{0, 1, 2, \dots\}$. Suppose that, for every $|\theta| \geq 0$, the series*

$$\rho_{|n|} = \sum_{|m| \geq |n|} \frac{|m|_{[|n|]}}{(|\theta| + |m|)_{(|n|)}} d_{|m|}, \qquad |n| = 0, 1, 2, \dots \tag{8.133}$$

converges. Then $\{\rho_{|n|}\}$ is a positive-definite sequence for the Poisson-Dirichlet point process with parameter $|\theta|$

8.2. From Jacobi to Hahn positive-definite sequences via discrete distributions.

We have seen in proposition 7.9 that Jacobi positive-definite sequences $\{\rho_{|n|}\}$ can always be mapped to Hahn positive-definite sequences of the form $\{\rho_{|n|} \frac{N_{[|n|]}}{(|\alpha|+N)_{(|n|)}}\}$. We now show that a JPDS $\{\rho_{|n|}\}$ is also HPDS when it is the image, via (8.133), of a particular class of discrete probability measures.

Proposition 8.6. *For every N and $|\theta| > 0$, let $\rho^{(N)} = \{\rho^{(N)}_{|n|} : |n| \in \mathbb{Z}_+\}$ be of the same form (8.133) for a probability mass function $d^{(N)} = \{d_{|m|} : |m| \in \mathbb{Z}_+\}$ such that $d_|l| = 0$ for every $|l| > N$. Then $\rho^{(N)}$ is $\widetilde{\alpha}$-JPDS if and only if it is $\widetilde{\alpha}$-HPDS for every d and $\alpha \in \mathbb{R}^d_+$ such that $|\alpha| = |\theta|$.*

Proof. By Lemma 7.7, the sequence

$$\left\{ \frac{|m|_{[|n|]}}{(|\alpha| + m)_{(|n|)}} \right\}$$

is HPDS (to convince oneself, take $\widetilde{\rho}$ as in Example 8.3 or in Example 8.4 and take the limit as $t \to 0$ or $z \to 1$ respectively).
Now replace $|m|$ with a random M with distribution given by $d^{(N)}$. Then

$$\begin{aligned} 0 \quad &\leq \quad \mathbb{E}\left[\sum_{|n|=0}^{|m|} \frac{|M|_{[|n|]}}{(|\alpha| + M)_{(|n|)}} H^{\alpha}_{|n|} \right] \\ &= \quad \sum_{|n|=0}^{N} \left(\sum_{|m|=|n|}^{N} d^{(N)}_{|m|} \frac{|M|_{[|n|]}}{(|\alpha| + M)_{(|n|)}} \right) H^{\alpha}_{|n|}. \end{aligned}$$

which proves the "Hahn"part of the claim. The "Jacobi"part is obviously proved by Proposition 8.133. ☐

9. From positive-definite sequences to probability measures.

In the previous Section we have seen that it is possible to map probability distributions on $\mathbb{Z}_+$ to Jacobi positive-definite sequences. It is natural to ask if, on the other way around, JPDS's $\{\rho_{|n|}\}$ can be mapped to probability distributions $\{d_{|m|}\}$ on $\mathbb{Z}_+$, for every $m = 0, 1, \dots,$ via the inversion

$$d_{|m|}(\rho) = \sum_{n=m}^{\infty} a^{|\alpha|}_{|n||m|} \rho_{|n|} \tag{9.134}$$

In this Section we give some sufficient conditions on ρ for $d_{|m|}(\rho)$ to be nonnegative for every $|m| = 0, 1, \dots,$ and an important counterexample showing that not all JPDS can be associated to probabilities. We restrict our attention to the Beta case ($d = 2$) as we now know that, if associated to a probability on $\mathbb{Z}_+$, any JPDS for $d = 2$ is also JPDS for $d > 2$.
Suppose $\rho = \{\rho_{|n|}\}_{|n|=0}^{\infty}$ satisfies

$$p_\rho(x, y) := \sum_{|n|=0}^{\infty} \rho_{|n|} Q^{\alpha,\beta}_{|n|}(x, y) \geq 0 \tag{9.135}$$

and in particular

$$p_\rho(x) := p_\rho(x,1) \geq 0. \tag{9.136}$$

Proposition 9.1. *If all the derivatives of $p_\rho(x)$ exist, then $d_{|m|}(\rho) \geq 0$ for every $|m| \in \mathbb{Z}_+$ if and only if all derivatives of $p_\rho(x)$ are non-negative.*

Proof. Rewrite $d_{|m|}(\rho)$ as

$$\begin{aligned} d_{|m|}(\rho) &= \sum_{|v|=0}^{\infty} a^{|\theta|}_{|v|+|m|,|m|}\rho_{|v|+|m|} \\ &= \frac{(|\theta|+|m|)_{(|m|)}}{|m|!}\sum_{|v|=0}^{\infty} a^{|\theta|+2|m|}_{|v|0}\rho_{|v|+|m|}, \qquad |m|=0,1,\ldots. \end{aligned} \tag{9.137}$$

This follows from the general identity:

$$a^{|\theta|}_{|v|+|j|,|u|+|j|} = a^{|\theta|+2|j|}_{|v|,|u|}\frac{|u|!}{(|u|+|j|)!}(|\theta|+|u|+|j|)_{(|j|)}. \tag{9.138}$$

Now consider the expansion of Jacobi polynomials. We know that

$$\begin{aligned} \zeta^{\alpha,\beta}_{|n|}R^{\alpha,\beta}_{|n|}(x)R^{\alpha,\beta}_{|n|}(y) &= Q^{\alpha,\beta}_{|n|}(x,y) \\ &= \sum_{|m|=0}^{n} a^{|\theta|}_{|n||m|}\xi^{\alpha,\beta}_{|m|}(x,y), \end{aligned} \tag{9.139}$$

where

$$\xi^{\alpha,\beta}_{|m|}(x,y) = \sum_{i=0}^{m}\left[\binom{|m|}{i}(xy)^i[(1-x)(1-y)]^{|m|-i}\right] \Big/ \left[\frac{\alpha_{(i)}\beta_{(|m|-i)}}{|\theta|_{(|m|)}}\right].$$

Since $R^{\alpha,\beta}_{|n|}(1)=1$ and $\xi^{\alpha,\beta}_{|m|}(0,1)=\delta_{|m|0}$ then

$$\zeta^{\alpha,\beta}_{|n|}R^{\alpha,\beta}_{|n|}(0) = Q^{\alpha,\beta}_{|n|}(0,1) = a^{|\theta|}_{|n|0} \tag{9.140}$$

Therefore (9.137) becomes

$$d_{|m|}(\rho) = \frac{(|\theta|+|m|)_{(|m|)}}{|m|!}\sum_{|v|=0}^{\infty}\zeta_v^{\alpha+|m|,\beta+|m|}R^{\alpha+|m|,\beta+|m|}_{|v|}(0)\rho_{|v|+|m|}, \qquad m=0,1,\ldots \tag{9.141}$$

Now apply e.g. [[18], (4.3.2)] to deduce

$$\frac{d^{|m|}}{dy^{|m|}}\left[D_{\alpha+|m|,\beta+|m|}(y)R^{\alpha+|m|,\beta+|m|}_{|v|}(y)\right] = (-1)^{|m|}\frac{\theta_{(2|m|)}}{\alpha_{(|m|)}}R^{\alpha,\beta}_{|v|+|m|}(y)D_{\alpha,\beta}(y). \tag{9.142}$$

For $|m|=1$,

$$\begin{aligned} \rho_{|v|+1} &= \int_0^1 p_\rho(x)R^{\alpha,\beta}_{|v|+1}(x)D_{\alpha,\beta}(x)dx \\ &= -\frac{\alpha}{|\theta|_{(2)}}\int_0^1 p_\rho(x)\left[\frac{d}{dx}R^{\alpha+1,\beta+1}_{|v|}(x)D_{\alpha+1,\beta+1}(x)\right]dx \\ &= \frac{\alpha}{|\theta|_{(2)}}\int_0^1\left(\frac{d}{dx}p_\rho(x)\right)R^{\alpha+1,\beta+1}_{|v|}(x)D_{\alpha+1,\beta+1}(x)dx. \end{aligned}$$

The last equality is obtained after integrating by parts. Similarly, denoted

$$p^{(|m|)}_\rho(x) := \frac{d^{|m|}}{dx^{|m|}}p_\rho(x), \qquad |m|=0,1,\ldots$$

it is easy to prove that

$$\rho_{|v|+|m|} = \frac{|m|!\alpha_{(|m|)}}{|\theta|_{(2|m|)}} \int_0^1 p_\rho^{(|m|)}(x) R_{|v|}^{\alpha+|m|,\beta+|m|}(x) D_{\alpha+|m|,\beta+|m|}(x)dx, \tag{9.143}$$

So we can write

$$d_{|m|}(\rho) = \frac{\alpha_{(|m|)}}{|\theta|_{(|m|)}} p_\rho^{(|m|)}(0).$$

Thus if $p_\rho^{(|m|)} \geq 0$, then $d_{|m|}(\rho)$ is, for every $|m|$, non-negative and this proves the sufficiency.

For the necessity, assume, without loss of generality, that $\{d_{|m|}(\rho) : m \in \mathbb{Z}_+\}$ is a probability mass function on $\mathbb{Z}_+$. Then its probability generating function (*pgf*) must have all derivatives nonnegative. For every $0 < \gamma < |\theta|$, the pgf has the representation:

$$\begin{aligned}
\varphi(s) &= \sum_{|m|=0}^{\infty} d_{|m|}(\rho) s^{|m|} \\
&= \mathbb{E}_{\gamma,|\theta|-\gamma}\left[\sum_{|m|=0}^{\infty} d_{|m|}(\rho)\xi_{|m|}^{\gamma,|\theta|-\gamma}(sZ,1)\right] \\
&= \mathbb{E}_{\gamma,|\theta|-\gamma}\left[\sum_{|m|=0}^{\infty} \rho_{|n|}\zeta_{|n|}^{\gamma,|\theta|-\gamma} R_{|n|}^{\gamma,|\theta|-\gamma}(sZ)\right] \\
&= \mathbb{E}_{\gamma,|\theta|-\gamma}\left[p_\rho(sZ)\right]
\end{aligned} \tag{9.144}$$

where Z is a Beta$(\gamma, |\theta| - \gamma)$ random variable. Here the second equality follows from the identity:

$$\frac{|\theta|_{(|m|)}}{\alpha_{(|m|)}} x^{|m|} = \xi_{|m|}^{\alpha,\beta}(x,1), \quad \alpha,\beta > 0 \tag{9.145}$$

and the third equality comes from (9.139).
So, for every $k = 0, 1, \ldots,$

$$0 \leq \frac{d^{|k|}}{ds^{|k|}}\varphi(s) = \mathbb{E}_{\gamma,|\theta|-\gamma}\left[Z^{|k|} p_\rho^{(|k|)}(sZ)\right] \tag{9.146}$$

for every $\gamma \in (0, |\theta|)$. Now if we take the limit as $\gamma \to |\theta|$, $Z \xrightarrow{d} 1$ so, by continuity,

$$\mathbb{E}_{\gamma,|\theta|-\gamma}\left[Z^{|k|} p_\rho^{(|k|)}(sZ)\right] \underset{\gamma\to|\theta|}{\to} p^{(|k|)}(s)$$

preserving the positivity, which completes the proof. □

9.1. A counterexample.

In Gasper's representation (Theorem 7.2), every positive-definite sequence is a mixture of Jacobi polynomials, normalized with respect to their value at 1. It is natural to ask whether these extreme points lead themselves to probability measures on $\mathbb{Z}_+$. A positive answer would imply that all positive-definite sequences, under Gasper's conditions, are coupled with probabilities on the integers. Rather surprisingly, the answer is negative.

Proposition 9.2. *Let $\alpha, \beta > 0$ satisfy Gasper's conditions. The function*

$$d_{|m|} = \sum_{|n|\geq|m|} a_{|n||m|}^{|\theta|} R_{|n|}^{\alpha,\beta}(x), \qquad |m| = 0, 1, 2, \ldots$$

is not a probability measure.

Proof. Rewrite

$$\begin{aligned}\phi_x(s) &= \sum_{n=0}^{\infty} R_n^{\alpha,\beta}(x) \sum_{|m|=0}^{|n|} a_{|n||m|}^{|\theta|} s^{|m|} \\ &= \mathbb{E} \sum_{|n|=0}^{\infty} \zeta_{|n|}^{\alpha,\beta} R_{|n|}^{\alpha,\beta}(x) R_{|n|}^{\alpha,\beta}(Ws),\end{aligned} \tag{9.147}$$

where W is a Beta (α, β) random variable. This also shows that, for every x,

$$\frac{dD_{\alpha,\beta}(y)}{dy} \sum_{n=0}^{\infty} \zeta_{|n|}^{\alpha,\beta} R_{|n|}^{\alpha,\beta}(x) R_{|n|}^{\alpha,\beta}(y) = \delta_x(y),$$

i.e. the Dirac measure putting all its unit mass on x (see also Example 8.4).
Now, if $\phi_x(s)$ is a probability generating function, then, for every positive L_2 function g, any mixture of the form

$$\begin{aligned}q(s) &= \int_0^1 g(x)\phi_x(s) \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} dx \\ &= \int_0^1 g(ws) \frac{w^{\alpha-1}(1-w)^{\beta-1}}{B(\alpha,\beta)} dw\end{aligned} \tag{9.148}$$

must be a probability generating function, i.e. it must have all derivatives positive. However, if we choose $g(x) = e^{-\lambda x}$, then we know that, g being completely monotone, the derivatives of q will have alternating sign, which proves the claim. □

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[2] G. E. Andrews, R. Askey, and R. Roy. *Special functions*, volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1999.

[3] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic combinatorial structures: a probabilistic approach*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2003.

[4] S. Bochner. Positive zonal functions on spheres. *Proc. Nat. Acad. Sci. U. S. A.*, 40:1141–1147, 1954.

[5] P. L. Butzer. On two-dimensional Bernstein polynomials. *Canadian J. Math.*, 5:107–113, 1953.

[6] K. Doksum. Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probability*, 2:183–201, 1974.

[7] C. F. Dunkl and Y. Xu. *Orthogonal polynomials of several variables*, volume 81 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2001.

[8] A. Erdélyi. On some expansions in Laguerre polynomials. *J. London Math. Soc.*, s1-13(2):154–156, 1938.

[9] T. S. Ferguson. A Bayesian analysis of some nonparametric problems. *Ann. Statist.*, 1:209–230, 1973.

[10] G. Gasper. Banach algebras for Jacobi series and positivity of a kernel. *Ann. of Math. (2)*, 95:261–280, 1972.

[11] G. Gasper. Nonnegativity of a discrete Poisson kernel for the Hahn polynomials. *J. Math. Anal. Appl.*, 42:438–451, 1973. Collection of articles dedicated to Salomon Bochner.

[12] R. Griffiths and D. Spanò. *Diffusion Processes and Coalescent Trees*. Chapter 15 of *Probability and Mathematical Genetics: Papers in Honour of Sir John Kingman*. London Mathematical Society Lecture Notes Series. Cambridge University Press, 2009/10,. *to appear*.

[13] R. C. Griffiths. On the distribution of allele frequencies in a diffusion model. *Theoret. Population Biol.*, 15(1):140–158, 1979.

[14] R. C. Griffiths. A transition density expansion for a multi-allele diffusion model. *Adv. in Appl. Probab.*, 11(2):310–325, 1979.

[15] R. C. Griffiths. Lines of descent in the diffusion approximation of neutral Wright-Fisher models. *Theoret. Population Biol.*, 17(1):37–50, 1980.

[16] R. C. Griffiths. Coalescent lineage distributions. *Adv. in Appl. Probab.*, 38(2):405–429, 2006.

[17] R. C. Griffiths and D. Spanò. Multivariate Jacobi and Laguerre polynomials, infinite-dimensional extensions, and their probabilistic connections with multivariate Hahn and Meixner polynomials. *http://arxiv.org/abs/0809.1431*, 2008.

[18] M. E. H. Ismail. *Classical and quantum orthogonal polynomials in one variable*, volume 98 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005. With two chapters by Walter Van Assche, With a foreword by Richard A. Askey.

[19] S. Karlin and J. McGregor. Linear growth models with many types and multidimensional Hahn polynomials. In *Theory and application of special functions (Proc. Advanced Sem., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1975)*, pages 261–288. Math. Res. Center, Univ. Wisconsin, Publ. No. 35. Academic Press, New York, 1975.

[20] S. Karlin and J. L. McGregor. The Hahn polynomials, formulas and an application. *Scripta Math.*, 26:33–46, 1961.

[21] K. Khare and H. Zhou. Rates of convergence of some multivariate Markov chains with polynomial eigenfunctions. *Ann. Appl. Probab.*, 19(2):737–777, 2009.

[22] J. F. C. Kingman. The population structure associated with the Ewens sampling formula. *Theoret. Population Biology*, 11(2):274–283, 1977.

[23] J. F. C. Kingman, S. J. Taylor, A. G. Hawkes, A. M. Walker, D. R. Cox, A. F. M. Smith, B. M. Hill, P. J. Burville, and T. Leonard. Random discrete distribution. *J. Roy. Statist. Soc. Ser. B*, 37:1–22, 1975. With a discussion by S. J. Taylor, A. G. Hawkes, A. M. Walker, D. R. Cox, A. F. M. Smith, B. M. Hill, P. J. Burville, T. Leonard and a reply by the author.

[24] T. Koornwinder. Jacobi polynomials. II. An analytic proof of the product formula. *SIAM J. Math. Anal.*, 5:125–137, 1974.

[25] T. Koornwinder. Jacobi polynomials. II. An analytic proof of the product formula. *SIAM J. Math. Anal.*, 5:125–137, 1974.

[26] T. Koornwinder. Two-variable analogues of the classical orthogonal polynomials. In *Theory and application of special functions (Proc. Advanced Sem., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1975)*, pages 435–495. Math. Res. Center, Univ. Wisconsin, Publ. No. 35. Academic Press, New York, 1975.

[27] T. H. Koornwinder and A. L. Schwartz. Product formulas and associated hypergroups for orthogonal polynomials on the simplex and on a parabolic biangle. *Constr. Approx.*, 13(4):537–567, 1997.

[28] A. E. Koudou. Probabilités de Lancaster. *Exposition. Math.*, 14(3):247–275, 1996.

[29] H. O. Lancaster. The structure of bivariate distributions. *Ann. Math. Statist.*, 29:719–736, 1958.

[30] G. Lauricella. Sulle funzioni ipergeometriche a piu‘ variabili. *Rend. Circ. Mat. Palermo*, 7:111–158, 1893.

[31] P. Muliere, P. Secchi, and S. Walker. Partially exchangeable processes indexed by the vertices of a k-tree constructed via reinforcement. *Stochastic Process. Appl.*, 115(4):661–677, 2005.

[32] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.

[33] L. Saloff-Coste, P. Diaconis, and K. Khare. Gibbs sampling, exponential families and orthogonal polynomials. to appear in *Statistical Science*, 2009.

[34] T. Sauer. Jacobi polynomials in Bernstein form. *J. Comput. Appl. Math.*, 199(1):149–158, 2007.

[35] S. Waldron. On the Bernstein-Bézier form of Jacobi polynomials on a simplex. *J. Approx. Theory*, 140(1):86–99, 2006.

[36] G. A. Watterson. Correction to: “The stationary distribution of the infinitely-many neutral alleles diffusion model”. *J. Appl. Probability*, 14(4):897, 1977.

[37] G. A. Watterson. Lines of descent and the coalescent. *Theoret. Population Biol.*, 26(1):77–92, 1984.