Bernoulli **19**(2), 2013, 548–598
DOI: 10.3150/11-BEJ403

Orthogonal polynomial kernels and canonical correlations for Dirichlet measures

ROBERT C. GRIFFITHS[1] and DARIO SPANÒ[2]

[1]*Department of Statistics, University of Oxford, 1 South Parks Road Oxford OX1 3TG, UK. E-mail: griff@stats.ox.ac.uk*
[2]*Department of Statistics, University of Warwick, Coventry CV4 7AL, UK. E-mail: d.spano@warwick.ac.uk*

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 Ewens's 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 vice versa. 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. We establish several identities and some integral representations which are multivariate extensions of important results known for the case $d=2$ since the 1970s. These results, along with a common interpretation of the mentioned kernels in terms of dependent Pólya 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 $d \to \infty$.

Keywords: canonical correlations; Dirichlet distribution; Dirichlet-multinomial distribution; Ewens's sampling formula; Hahn; Jacobi; Lancaster's problem; multivariate orthogonal polynomials; orthogonal polynomial kernels; Poisson–Dirichlet distribution; Pólya urns; positive-definite sequences

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 Henry Oliver Lancaster [22] is in terms of so-called *canonical correlations*. Let $\{P_n\}_{n=0}^{\infty}$ be a family of orthogonal polynomials with weight measure π, that is, such that

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

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.

1350-7265

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

$$g_\rho(\mathrm{d}x, \mathrm{d}y) = \pi(\mathrm{d}x)\pi(\mathrm{d}y)\left\{\sum_{n=0}^{\infty} \rho_n c_n P_n(x) P_n(y)\right\}, \tag{1.1}$$

where $\rho_0 = 1$. 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

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

Define the conditional expectation operator by

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

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 particular,

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

that is, the polynomials $\{P_n\}$ are the eigenfunctions, and ρ is the sequence of eigenvalues of T_ρ. Lancaster's problem is therefore a spectral problem, whereby conditional expectation 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 either the Dirichlet or the Dirichlet multinomial distribution (notation: D_α and $\mathrm{DM}_{\alpha,N}$, with $\alpha \in \mathbb{R}_+^d$ and $N \in \mathbb{Z}_+$) on the $(d-1)$-dimensional continuous and N-discrete simplex, respectively. The eigenfunctions will be the multivariate Jacobi or Hahn polynomials, respectively. One difficulty arising when $d > 2$ is that the orthogonal polynomials $P_{\mathbf{n}} = P_{n_1 n_2 \cdots n_d}$ are multi-indexed. The degree of every polynomial $P_{\mathbf{n}}$ is $|\mathbf{n}| := n_1 + \cdots + n_d$ (throughout the paper, for every vector $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$, we will denote its length by $|\mathbf{x}|$). There are

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

polynomials with degree n, so, when $d > 2$, there is no unique way to introduce a total order in the space of all polynomials. *Orthogonal polynomial kernels* are instead uniquely defined and totally ordered.

By orthogonal polynomial kernels of degree n, with respect to π, we mean functions of the form

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

where $\{P_{\mathbf{n}}\colon \mathbf{n}\in\mathbb{Z}_+^d\}$ is a system of orthogonal polynomials with weight measure π.

It is easy to check that

$$\mathbb{E}_\pi[P_n(\mathbf{x},\mathbf{Y})P_m(\mathbf{z},\mathbf{Y})] = P_n(\mathbf{x},\mathbf{z})\delta_{mn}.$$

A representation equivalent to (1.2) in term of polynomial kernels is

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

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

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

then

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

where Y has distribution π.

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

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

then

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

Similar expressions to (1.6) hold for $P_{\mathbf{n}}^\diamond(\mathbf{x})$ and $P_{\mathbf{n}}^\circ(\mathbf{x})$, 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 $\mathrm{DM}_{\alpha,N}$ will be denoted by $Q_n^\alpha(\mathbf{x},\mathbf{y})$ and $H_n^\alpha(\mathbf{r},\mathbf{s})$, 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.

We will show that all the kernels under study can be constructed via systems of bivariate Pólya urns with random samples in common. This remarkable property assimilates the Dirichlet "world" to other distributions, within the so-called Meixner class, whose orthogonal polynomial kernels admit a representation in terms of bivariate sums with random elements in common, a fact known since the 1960s (see [6,7]. See also [4] for a modern Bayesian approach).

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 $\mathrm{DM}_{\alpha,N}$. We will restrict our focus on sequences ρ such that, for every $\mathbf{n} \in \mathbb{Z}_+^d$, $\rho_{\mathbf{n}}$ depends on $\mathbf{n}$ only through its total length $|\mathbf{n}| = \sum_{i=1}^d n_i$:

$$\rho_{\mathbf{n}} = \rho_n \qquad \forall \mathbf{n} \in \mathbb{Z}_+^d\colon\ |\mathbf{n}| = n.$$

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

$$\sum_{n=0}^{\infty} \rho_n P_n(\mathbf{u}, \mathbf{v}) \geq 0. \tag{1.8}$$

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 (1.8) an *α-Jacobi positive-definite sequence* (*α-JPDS*), if $\pi = D_\alpha$, and an *(α, N)-Hahn positive-definite sequence* (*(α, N)-HPDS*), if $\pi = \mathrm{DM}_{\alpha,N}$.

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

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 1970s: for α, β in a "certain region" (see Theorem 5.1 further on), the integral representation

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

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

$$Q_n^\alpha(\mathbf{x}, \mathbf{y}) = \mathbb{E}\big[Q_n^{\alpha_d, |\alpha| - \alpha_d}(Z_d, 1) | \mathbf{x}, \mathbf{y}\big], \qquad |n| = 0, 1, \ldots \tag{1.9}$$

for every $d \in \mathbb{N}$, every α in a "certain region" of $\mathbb{R}_+^d$, and for a particular [0, 1]-valued random variable Z_d. Here, for every $j = 1, \ldots, d$, $\mathbf{e}_j = (0, 0, \ldots, 1, 0, \ldots, 0) \in \mathbb{R}^d$ is the vector with all zero components, except for the jth coordinate which is equal to 1. Integral representations such as (1.9) are useful in that they map immediately univariate positive functions to the type of bivariate distribution we are looking for,

$$f(\mathbf{x}) \geq 0 \quad \Longrightarrow \quad \sum_n \widehat{f}(n) Q_n(\mathbf{x}, \mathbf{y}) = \mathbb{E}[f(Z_d)|\mathbf{x}, \mathbf{y}] \geq 0.$$

In fact, whenever (1.9) holds true, we will be able to conclude that (P1) is true.

Identity (1.9), however, holds only with particular choices of the parameter α. At best, one needs one of the α_js to be greater than 2. This makes it hard to use (P1) to build, in the limit as $d \to \infty$, canonical correlations with respect to Poisson–Dirichlet marginals on the infinite simplex. The latter would be a desirable aspect for modeling dependent measures on the infinite symmetric group or for applications, for example, 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)

$$\rho_n(t) = \mathrm{e}^{-(1/2)n(n+|\alpha|-1)t}, \qquad n = 0, 1, \ldots, \tag{1.10}$$

arising as the eigenvalues of the transition semigroup of the so-called *d-type, neutral Wright–Fisher diffusion process in population genetics; see, for example*, [11,14,27]. The generator of the diffusion process $\{\mathbf{X}(t), t \geq 0\}$ describing the relative frequencies of genes with type space $\{1, \ldots, d\}$ is

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{\mathrm{d}} x_i(\delta_{ij} - x_j) \frac{\partial^2}{\partial x_i \, \partial x_j} + \frac{1}{2} \sum_{i=1}^{d} (\alpha_i - |\alpha| x_i) \frac{\partial}{\partial x_i}.$$

In this model, mutation is parent-independent from type i to j at rate $\alpha_j/2$, $j \in \{1, \ldots, d\}$. Assuming that $\alpha > 0$, the stationary distribution of the process is D_α, and the transition density has an expansion

$$f(\mathbf{x}, \mathbf{y}; t) = D_\alpha(y) \left\{ 1 + \sum_{n=1}^{\infty} \rho_n(t) Q_n^\alpha(\mathbf{x}, \mathbf{y}) \right\}.$$

The limit model as $d \to \infty$ with $\alpha = |\alpha|/d$ is the infinitely-many-alleles-model, where mutation is always to a novel type. The stationary distribution is Poisson–Dirichlet(α).

The same sequence (1.10) is also a HPDS playing a role in population genetics [17]: it is the eigenvalue sequence of the so-called Moran model with type space $\{1, \ldots, d\}$. In a population of N individuals, $\{\mathbf{Z}(t), t \geq 0\}$ denotes the number of individuals of each type at t, $|\mathbf{Z}(t)| = N$. In reproduction events, an individual is chosen at random to reproduce with one child, and another is chosen at random to die. The offspring of a parent of type i does not mutate with probability $1-\mu$, or mutates in a parent independent way to type j with probability μp_j, $j \in \{1,\dots,d\}$, where $|\mathbf{p}|=1$. The generator of the process is described by

$$\mathcal{L}f(\mathbf{z}) = \sum_{i=1}^{d}\sum_{j=1}^{d} z_i\left(\frac{\lambda}{N}z_j + \mu p_j\right)\left[f(\mathbf{z}-\mathbf{e}_i+\mathbf{e}_j)-f(\mathbf{z})\right].$$

Setting $\alpha = M\mu\mathbf{p}/\lambda$, $\lambda = N/2$, the stationary distribution of the process is $\mathrm{DM}_{\alpha,N}$, the eigenvalues are (1.10), and the transition density is

$$P\big(Z(t)=s|Z(0)=r\big) = \mathrm{DM}_{\alpha,N}(s)\left\{1+\sum_{n=1}^{N}\rho_n(t)H_n(r,s)\right\}.$$

Thus (1.10) is an example of positive-definite sequence satisfying both (P1) and (P2).

(ii)

$$\rho_n(z) = z^n, \qquad n = 0,1,\dots;$$

that is, the eigenvalues of the so-called Poisson kernel, whose positivity is a well-known result in special functions theory (see, e.g., [5,16]).

An interpretation of Poisson kernels as Markov transition semigroups is in [14], where it is shown that (ii) can be obtained via an appropriate subordination of the genetic model (i).

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, this question. This approach is heavily based on the probabilistic interpretation (Pólya urns with random draws in common) of the Jacobi and Hahn polynomial kernels shown in the first part of the paper. We will prove in Proposition 8.1 that, if $\{d_m\colon m=0,1,2,\dots\}$ is a probability mass function (p.m.f.) on $\mathbb{Z}_+$, then every positive-definite sequence $\{\rho_n\}_{n=0}^{\infty}$ of the form

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

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, that is, will find a set of conditions on a JPD sequence ρ to be of the form (1.11) for a p.m.f. $\{d_m\}$.

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

$$H_n^{\alpha}(\mathbf{r},\mathbf{s}) = \frac{(N-n)!\Gamma(|\alpha|+N+n)}{N!\Gamma(|\alpha|+N)}\mathbb{E}\left[Q_n^{\alpha}(\mathbf{X},\mathbf{Y})|\mathbf{r},\mathbf{s}\right], \qquad n=0,1,\dots$$

for every $N\in\mathbb{Z}_+$ and $\mathbf{r},\mathbf{s}\in N\Delta_{(d-1)}$, where the expectation on the right-hand side is taken with respect to $D_{\alpha+\mathbf{r}}\otimes D_{\alpha+\mathbf{s}}$, that is, a product of *posterior* Dirichlet probability measures. A similar result was proven by [15] to hold for individual Hahn polynomials as well. The interpretation is again in terms of dependent Pólya sequences with random elements in common.

We will also show (Proposition 6.1) that a discrete version of (1.9) (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 10 some results "close to" (P2): that JPDSs can indeed be viewed as a map from HPDSs, and vice versa, but such mappings, in general, are not the inverse of each other.

On the other hand, we will show (Proposition 10.4) that every JPDS is in fact the limit of a sequence of (P2)-positive-definite sequences.

Our final result on HPDSs is in Proposition 10.8, where we prove that if, for fixed N, $d^{(N)} = \{d_m^{(N)}\}_{m\in\mathbb{Z}_+}$ is a probability distribution such that $d_l^{(N)} = 0$ for $l > N$, then (P2) holds properly for the JPDS ρ of the form (1.11). Such sequences also satisfy (P1) and admit infinite-dimensional Poisson–Dirichlet (and Ewens's sampling distribution) limits.

The key for the proof of Proposition 10.8 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 [9] 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_n^α is given in terms of mixtures of products of multinomial probability distributions arising from dependent Pólya urns with random elements in common. We will next obtain (Section 3) an explicit representation for H_n^α as *posterior mixtures* of Q_n^α. In the same section we will generalize Gasper's product formula to an alternative representation of H_n^α 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 $\mathrm{DM}_{\alpha,N}$, and to their infinite-dimensional limits, known as the Poisson–Dirichlet and Ewens's sampling distribution, respectively. 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 identity (1.9) for the Jacobi product formula and its Hahn equivalent. We will point out the connection between (1.9) and another multivariate Jacobi product formula due to Koornwinder and Schwartz [21].

In Section 7 we will focus more closely on positive-definite sequences (canonical correlations). We will use results of Section 5 (first approach) to characterize sequences obeying to (P1), with constraints on α.

In Section 8 we will use a second probabilistic approach to 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 $\mathbb{Z}_+$. Every such sequence will be determined by a probability mass function on the integers. We will discuss the possibility of a converse mapping from JPDSs to probability mass functions in Section 9.

In the remaining sections we will investigate the existence of sequences satisfying (P2). In particular, in Section 10.1 we will make a similar use of probability mass functions to find sufficient conditions for a proper version of (P2).

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 [15] for further properties and related distributions. For $\alpha, \mathbf{x} \in \mathbb{R}^d$ and $\mathbf{n} \in \mathbb{Z}_+^d$, denote

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

and

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

Also, we will use

$$(\mathbf{a})_{(\mathbf{x})} = \frac{\Gamma(\mathbf{a}+\mathbf{x})}{\Gamma(\mathbf{a})},$$

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

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

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) *The* Dirichlet(α) *distribution*, $\alpha \in \mathbb{R}_+^d$, *on the* d*-dimensional simplex*

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

is given by

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

(ii) *The Dirichlet multinomial* (α, N) *distribution*, $\alpha \in \mathbb{R}_+^d, N \in \mathbb{Z}_+$ *on the* $(d-1)$*-dimensional discrete simplex*

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

is given by the probability mass function

$$\mathrm{DM}_{\alpha,N}(\mathbf{r}) = \binom{N}{\mathbf{r}} \frac{(\alpha)_{(\mathbf{r})}}{(|\alpha|)_{(N)}}, \qquad \mathbf{r} \in N\Delta_{(d-1)}. \tag{1.12}$$

1.2.1. *Pólya sampling distribution*

$\mathrm{DM}_{\alpha,N}$ can be thought as the moment formula (sampling distribution) of D_α,

$$\mathbb{E}_{D_\alpha}\left[\binom{N}{\mathbf{r}} \mathbf{X}^{\mathbf{r}}\right],$$

so $\mathrm{DM}_{\alpha,N}$ can be interpreted as the probability distribution of a sample of N random variables in $\{1,\dots,d\}$, which are conditionally independent and identically distributed with common law $\mathbf{X}$, the latter being a random distribution with distribution D_α. The probability distribution of $\mathbf{X}$, conditional on a sample of N such individuals, is, by Bayes's theorem, again Dirichlet with different parameters

$$D_{\alpha+\mathbf{r}}(d\mathbf{x}) = \frac{\binom{N}{\mathbf{r}}\mathbf{x}^{\mathbf{r}}}{\mathrm{DM}_{\alpha,N}(\mathbf{r})} D_\alpha(d\mathbf{x}). \tag{1.13}$$

As $N \to \infty$, the measure $\mathrm{DM}_{\alpha,N}$ tends to D_α. The Dirichlet multinomial distribution can also be thought as the distribution of color frequencies arising in a sample of size N from a d-color Pólya urn. This sampling scheme can be described as follows: in an urn there are $|\alpha|$ balls of which α_i are of color i, $i = 1,\dots,d$ (for this interpretation one may assume, without loss of generality, that $\alpha \in \mathbb{Z}_+^d$). Pick a ball uniformly at random, note its color, then return the ball in the urn and add another ball of the same color. The probability of the first sample to be of color i is $\alpha_i/|\alpha|$. After simple combinatorics one sees that the distribution of the color frequencies after M draws is $\mathrm{DM}_{\alpha,M}$. Conditional on having observed $\mathbf{r}$ as frequencies in the first M draws, the probability distribution of observing $\mathbf{s}$ in the next $N-M$ draws is

$$\mathrm{DM}_{\alpha+\mathbf{r},N-M}(\mathbf{s}) = D_{\alpha,N-M}(\mathbf{s}) \frac{D_{\alpha+\mathbf{s},M}(\mathbf{r})}{D_{\alpha,M}(\mathbf{r})}. \tag{1.14}$$

1.2.2. *Ranked frequencies and limit distributions*

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

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

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 one with density, with respect to the* d*-dimensional Lebesgue measure*

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

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

$$\mathrm{DM}_{\alpha,N}^{\downarrow} := \mathrm{DM}_{\alpha,N} \circ \psi^{-1}$$

defines the ranked Dirichlet multinomial *distribution.*

With a slight abuse of notation, we will use $D^{\downarrow}$ to indicate both the ranked Dirichlet measure and its density. Ranked symmetric Dirichlet and Dirichlet multinomial measures can be interpreted as distributions on random partitions. For every $\mathbf{r} \in N\Delta_{(d-1)}$ let $\beta_j = \beta_j(\mathbf{r})$ be the number of elements in $\mathbf{r}$ equal to j and $k(\mathbf{r}) = \sum \beta_j(\mathbf{r})$ the number of strictly positive components of $\mathbf{r}$. Thus $\sum_{i=1}^r i\beta_i(\mathbf{r}) = N$.

For each $\mathbf{x} \in \Delta_{(d-1)}$ denote the monomial symmetric polynomials by

$$[\mathbf{x}, \mathbf{r}]_d := \sum_{i_1 \neq \cdots \neq i_k \in \{1,\dots,d\}^k} \prod_{j=1}^k x_{i_j}^{r_j},$$

where the sum is over all $d_{[k]}$ subsequences of k distinct integers, and let $[\mathbf{x}, \mathbf{r}]$ be its extension to $\mathbf{x} \in \Delta_\infty$. Take a collection $(\xi_1, \dots, \xi_N)$ 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 $[\mathbf{x}, \mathbf{r}]_d$ can be interpreted as the probability distribution of any such sample realization giving rise to $k(\mathbf{r})$ distinct values whose *unordered* frequencies count $\beta_1(\mathbf{r})$ singletons, $\beta_2(\mathbf{r})$ doubletons and so on.

There is a bijection between $\mathbf{r}^{\downarrow} = \psi(r)$ and $\beta(\mathbf{r}) = (\beta_1(\mathbf{r}), \dots, \beta_N(\mathbf{r}))$, both maximal invariant functions with respect to permutations of coordinates, both representing partitions of N in $k(\mathbf{r})$ parts. Note that $[\mathbf{x}, \mathbf{r}]_d$ is invariant too, for every $d \le \infty$. It is well known that, for every $\mathbf{x} \in \Delta_d^{\downarrow}$,

$$\sum_{\mathbf{r}^{\downarrow} \in N\Delta_{(d-1)}^{\downarrow}} \binom{N}{\mathbf{r}^{\downarrow}} \frac{1}{\prod_{i \ge 1} \beta_i(\mathbf{r}^{\downarrow})!} [x, \mathbf{r}^{\downarrow}]_d = 1, \tag{1.15}$$

that is, for every $\mathbf{x}$,

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

represents a probability distribution on the space of random partitions of N.

For $|\alpha| > 0$, let $D_{|\alpha|,d}, \mathrm{DM}_{|\alpha|,N,d}$ denote the Dirichlet and Dirichlet multinomial distributions with symmetric parameter $(|\alpha|/d, \ldots, |\alpha|/d)$. Then

$$\begin{aligned}
&\mathrm{DM}^{\downarrow}_{|\alpha|,N,d}(\mathbf{r}^{\downarrow}) \\
&\quad = \mathbb{E}_{D^{\downarrow}_{|\alpha|,d}}\left\{\binom{N}{\mathbf{r}^{\downarrow}}\frac{1}{\prod_{i\ge 1}\beta_i(\mathbf{r}^{\downarrow})!}[\mathbf{x}^{\downarrow},\mathbf{r}^{\downarrow}]_d\right\} \qquad (1.16)\\
&\quad = d_{[k]}\frac{r!}{\prod_1^N j!^{\beta_j}\beta_j!}\cdot\frac{\prod_1^r(|\alpha|/d)_{(j)}^{\beta_j}}{(|\alpha|)_{(N)}} \\
&\quad \underset{d\to\infty}{\longrightarrow} \frac{r!}{\prod_1^r j^{\beta_j}\beta_j!}\cdot\frac{|\alpha|^k}{(|\alpha|)_{(r)}} := \mathrm{ESF}_{|\alpha|}(\mathbf{r}). \qquad (1.17)
\end{aligned}$$

Definition 1.3. *The limit distribution* $\mathrm{ESF}_{|\alpha|}(\mathbf{r})$ *in* (1.17) *is called the Ewens sampling formula with parameter* $|\alpha|$.

Poisson–Dirichlet point process [19]. Let $Y^{\infty} = (Y_1, Y_2, \ldots)$ be the sequence of points of a non-homogeneous point process with intensity measure

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

The probability generating functional is

$$\begin{aligned}
\mathcal{F}_{|\alpha|}(\xi) &= \mathbb{E}_{|\alpha|}\left(\exp\left\{\int \log\xi(y) N_{|\alpha|}(\mathrm{d}y)\right\}\right) \\
&= \exp\left\{|\alpha|\int_0^{\infty}\big(\xi(y)-1\big)y^{-1}\mathrm{e}^{-y}\,\mathrm{d}y\right\}
\end{aligned} \qquad (1.18)$$

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*

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

(ii) *The relationship between* D_α *and* $\mathrm{DM}_{\alpha,N}$ *is replicated by ESF, which arises as the (symmetric) moment formula for the PD distribution,*

$$\mathrm{ESF}_{|\alpha|,N}(r) = \mathbb{E}_{\mathrm{PD}_{|\alpha|}}\left\{\binom{r}{\mathbf{r}^\downarrow}\frac{1}{\prod_{i\geq 1}\beta_i(\mathbf{r}^\downarrow)!}[\mathbf{x},\mathbf{r}^\downarrow]\right\}, \qquad r \in N\Delta^\downarrow. \tag{1.19}$$

Proof. If $\mathbf{Y} = (Y_1,\ldots,Y_d)$ is a collection of d independent random variables with identical distribution Gamma$(|\alpha|/d, 1)$, then their sum $|\mathbf{Y}|$ is a Gamma$(|\alpha|)$ random variable independent of $\mathbf{Y}/|\mathbf{Y}|$, which has distribution $D_{\alpha|,d}$. The probability generating functional of $\mathbf{Y}$ is ([10])

$$\begin{aligned}\mathcal{F}_{|\alpha|,d}(\xi) &= \left(1+\int_0^\infty \bigl(\xi(y)-1\bigr)\frac{|\alpha|}{d}\frac{y^{|\alpha|/d-1}\mathrm{e}^{-y}}{\Gamma(|\alpha|/d+1)}\,\mathrm{d}y\right)^d \\ &\underset{d\to\infty}{\to} \mathcal{F}_{|\alpha|}(\xi),\end{aligned} \tag{1.20}$$

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 [10]. □

2. Polynomial kernels in the Dirichlet distribution

The aim of this section is to show that, for every fixed $d \in \mathbb{N}$ and $\alpha \in \mathbb{R}^d$, the orthogonal polynomial kernels with respect to D_α can be constructed from systems of two dependent Pólya urns sharing a fixed number of random elements in common.

Consider two Pólya urns U_1 and U_2 with identical initial composition α, and impose on them the constraint that the first m draws from U_1 are identical to the first m draws from U_2. For $M \leq N$ sample $M+m$ balls from U_1 and $N+m$ balls from U_2. At the end of the experiment, the probability of having observed frequencies $\mathbf{r}$ and $\mathbf{s}$, respectively, in the M unconstrained balls sampled from U_1 and in the N ones from U_2, is, by (1.14),

$$\begin{aligned}&\sum_{|\mathbf{l}|=m}\mathrm{DM}_{\alpha,m}(\mathbf{l})\,\mathrm{DM}_{\alpha+\mathbf{l},M}(\mathbf{r})\,\mathrm{DM}_{\alpha+\mathbf{l},N}(\mathbf{s}) \\ &\quad= \mathrm{DM}_{\alpha,M}(\mathbf{r})\,\mathrm{DM}_{\alpha,N}(\mathbf{s})\xi_m^{H,\alpha}(\mathbf{r},\mathbf{s}),\end{aligned} \tag{2.1}$$

where

$$\xi_m^{H,\alpha}(\mathbf{r},\mathbf{s}) = \sum_{|\mathbf{l}|=m}\frac{\mathrm{DM}_{\alpha+\mathbf{s},m}(\mathbf{l})\,\mathrm{DM}_{\alpha+\mathbf{r},m}(\mathbf{l})}{\mathrm{DM}_{\alpha,m}(\mathbf{l})}. \tag{2.2}$$

As $N, M \to \infty$, if we assume $N^{-1}\mathbf{s}\to\mathbf{x}$, $M^{-1}\mathbf{r}\to\mathbf{y}$, we find that this probability distribution tends to

$$D_\alpha(\mathrm{d}\mathbf{x})\,D_\alpha(\mathrm{d}\mathbf{y})\xi_m^\alpha(\mathbf{x},\mathbf{y}),$$

where

$$\xi_m^\alpha(\mathbf{x},\mathbf{y}) = \sum_{|\mathbf{l}|=m} \binom{m}{\mathbf{l}} \frac{|\alpha|_{(m)}}{\prod_1^d \alpha_{i\,(l_i)}} \prod_1^d (x_i y_i)^{l_i} \tag{2.3}$$

$$= \sum_{|\mathbf{l}|=m} \frac{\binom{m}{\mathbf{l}}\mathbf{x}^{\mathbf{l}}\binom{m}{\mathbf{l}}\mathbf{y}^{\mathbf{l}}}{\mathrm{DM}_{\alpha,|m|}(\mathbf{l})}. \tag{2.4}$$

Notice that, because Pólya sequences are exchangeable (i.e., their law is invariant under permutations of the sample coordinates), the same formula (2.4) would hold even if we only assumed that the sequences sampled from U_1 and U_2 have in common any m (and not necessarily the first m) coordinates.

2.1. Polynomial kernels for $d \geq 2$

We shall now prove the following:

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

$$Q_n^\alpha(\mathbf{x},\mathbf{y}) = \sum_{m=0}^{n} a_{nm}^{|\alpha|} \xi_m^\alpha(\mathbf{x},\mathbf{y}), \tag{2.5}$$

where

$$a_{nm}^{|\alpha|} = (|\alpha| + 2n - 1)(-1)^{n-m} \frac{(|\alpha| + m)_{(n-1)}}{m!(n-m)!} \tag{2.6}$$

form a lower-triangular, invertible system. An inverse relationship is

$$\xi_m^\alpha(\mathbf{x},\mathbf{y}) = 1 + \sum_{n=1}^{m} \frac{(m)_{[n]}}{(|\alpha| + |m|)_{(n)}} Q_n^\alpha(\mathbf{x},\mathbf{y}). \tag{2.7}$$

Remark 2.2. A first construction of the Kernel polynomials was given by [11]. We provide here a revised proof. Operators with a role analogous to the function ξ_m have, later on, appeared in different contexts, but with little emphasis on Pólya urns or on the probabilistic aspects ([25,26] are some examples). A closer, recent result is offered in [24] where a multiple integral representation for square-integrable functions with respect to Ferguson–Dirichlet random measures is derived in terms of Pólya urns.

Proof of Proposition 2.1. Let $\{Q_{\mathbf{n}}^\circ\}$ be a system of *orthonormal* polynomials with respect to D_α (i.e., such $\mathbb{E}(Q_{\mathbf{n}}^{\circ 2}) = 1$). We need to show that, for independent Dirichlet distributed vectors X, Y, if $n, k \leq m$, then

$$\mathbb{E}(\xi_m^\alpha(\mathbf{X},\mathbf{Y}) Q_{\mathbf{n}}^\circ(\mathbf{X}) Q_{\mathbf{k}}^\circ(\mathbf{Y})) = \delta_{\mathbf{nk}} \frac{(m)_{[n]}}{(|\alpha| + m)_{(n)}}. \tag{2.8}$$

If this is true, an expansion is therefore

$$\begin{aligned}\xi_m^\alpha(\mathbf{x},\mathbf{y}) &= 1+\sum_{n=1}^m \frac{(m)_{[n]}}{(|\alpha|+m)_{(n)}}\sum_{\{\mathbf{n}:|\mathbf{n}|=n\}} Q_{\mathbf{n}}^\circ(\mathbf{x})Q_{\mathbf{n}}^\circ(\mathbf{y}) \\ &= 1+\sum_{n=1}^m \frac{(m)_{[n]}}{(|\alpha|+m)_{(n)}} Q_n^\alpha(\mathbf{x},\mathbf{y}).\end{aligned} \tag{2.9}$$

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

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

gives (2.5) from (2.7). The inverse matrix is triangular with (m,n)th element

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

and the proof will be complete.

Proof of (2.8). Write

$$\mathbb{E}\left(\prod_1^{d-1} {X_i}^{n_i}\xi_m^\alpha(\mathbf{X},\mathbf{Y})\Big|\mathbf{Y}\right) = \sum_{\{\mathbf{l}:|\mathbf{l}|=m\}}\binom{m}{\mathbf{l}}\prod_1^d Y_i^{l_i}\frac{\prod_1^{d-1}(l_i+\alpha_i)_{(n_i)}}{(|\alpha|+m)_{(n)}}. \tag{2.10}$$

Expressing the last product in (2.10) as

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

for constants $b_{\mathbf{nk}}$, shows that

$$\mathbb{E}\left(\prod_1^{d-1} {X_i}^{n_i}\xi_m^\alpha(\mathbf{X},\mathbf{Y})\Big|\mathbf{Y}\right) = \frac{(m)_{[n]}}{(|\alpha|+m)_{(n)}}\prod_1^{d-1} Y_i^{n_i} + R_0(\mathbf{Y}). \tag{2.11}$$

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

$$\begin{aligned}\mathbb{E}(\xi_m^\alpha(\mathbf{X},\mathbf{Y})Q_{\mathbf{n}}^\circ(\mathbf{X})|\mathbf{Y}) &= \frac{(m)_{[n]}}{(|\alpha|+m)_{(n)}}\sum_{\{\mathbf{k}:|\mathbf{k}|=n\}} a_{\mathbf{nk}}\prod_1^{d-1} Y_i^{k_i} + R_1(\mathbf{Y}) \\ &= \frac{(m)_{[n]}}{(|\alpha|+m)_{(n)}} Q_{\mathbf{n}}^\circ(\mathbf{Y}) + R_2(\mathbf{Y}),\end{aligned} \tag{2.12}$$

where

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

are terms of leading degree n in $Q_{\mathbf{n}}^{\circ}(\mathbf{X})$ and $R_j(\mathbf{Y})$, $j = 0, 1, 2$, are polynomials of degree less than n in $\mathbf{Y}$. Thus if $n \le k \le m$,

$$\begin{aligned}\mathbb{E}(\xi_m^{\alpha}(\mathbf{X}, \mathbf{Y}) Q_{\mathbf{n}}^{\circ}(\mathbf{X}) Q_{\mathbf{k}}^{\circ}(\mathbf{Y})) &= \mathbb{E}\left(Q_{\mathbf{k}}^{\circ}(\mathbf{Y}) \left\{ \frac{(m)_{[n]}}{(|\alpha| + m)_{(n)}} Q_{\mathbf{n}}^{\circ}(\mathbf{Y}) + R_2(\mathbf{Y}) \right\} \right) \\ &= \frac{(m)_{[n]}}{(|\alpha| + m)_{(n)}} \delta_{\mathbf{nk}}.\end{aligned} \tag{2.13}$$

By symmetry, (2.13) holds for all $\mathbf{n}, \mathbf{k}$ such that $n, k \le m$. □

2.2. Some properties of the kernel polynomials

Particular cases

$$\begin{aligned}Q_0^{\alpha} &= 1, \\ Q_1^{\alpha} &= (|\alpha| + 1)(\xi_1 - 1) \\ &= (|\alpha| + 1)\left(|\alpha| \sum_{1}^{\mathrm{d}} x_i y_i / \alpha_i - 1 \right), \\ Q_2^{\alpha} &= \tfrac{1}{2}(|\alpha| + 3)\big((|\alpha| + 2)\xi_2 - 2(|\alpha| + 1)\xi_1 + |\alpha|\big),\end{aligned}$$

where

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

The jth coordinate kernel

A well-known property of Dirichlet measures is that, if $\mathbf{Y}$ is a Dirichlet(α) vector in $\Delta_{(d-1)}$, then its jth 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 $\mathbf{e}_j$ be the vector in $\mathbb{R}^d$ with every ith coordinate equal δ_{ij}, $i, j = 1, \dots, d$. Then

$$\xi_m^{\alpha}(\mathbf{y}, \mathbf{e}_j) = \frac{(|\alpha|)_{(m)}}{(\alpha_j)_{(m)}} y_j^m, \qquad m \in \mathbb{Z}_+, \mathbf{y} \in \Delta_{(d-1)}. \tag{2.14}$$

In particular,

$$\xi_m^{\alpha}(\mathbf{e}_j, \mathbf{e}_k) = \frac{(|\alpha|)_{(m)}}{(\alpha_j)_{(m)}} \delta_{jk}. \tag{2.15}$$

Therefore, for every d and $\alpha \in \mathbb{R}_+^d$, (2.14) implies

$$\begin{aligned} Q_n^\alpha(\mathbf{y}, \mathbf{e}_j) &= \sum_{m=0}^n a_{nm}^{|\alpha|} \xi_m^\alpha(\mathbf{e}_j, \mathbf{y}) = Q_n^{\alpha_j, |\alpha|-\alpha_j}(y_j, 1) \\ &= \zeta_n^{\alpha_j, |\alpha|-\alpha_j} R_n^{\alpha_j, |\alpha|-\alpha_j}(y_j), \qquad j = 1, \ldots, d, \mathbf{y} \in \Delta_{(d-1)}, \end{aligned} \tag{2.16}$$

where

$$\begin{aligned} R_n^{\alpha,\beta}(x) &= \frac{Q_n^{\alpha,\beta}(x,1)}{Q_n^{\alpha,\beta}(1,1)} \\ &= {}_2F_1\left(\begin{matrix} -n, n+\theta-1 \\ \beta \end{matrix} \middle| 1-x \right), \qquad n = 0, 1, 2, \ldots, \theta = \alpha + \beta, \end{aligned} \tag{2.17}$$

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

$$\frac{1}{\zeta_n^{\alpha,\beta}} := \mathbb{E}[R_n^{\alpha,\beta}(X)]^2.$$

In (2.18), ${}_pF_q$, $p, q \in \mathbb{N}$, denotes the Hypergeometric function (see [1] for basic properties).

Remark 2.3. For $\alpha, \beta \in \mathbb{R}_+$, let $\theta = \alpha + \beta$. It is known (e.g., [15], (3.25)) that

$$\frac{1}{\zeta_n^{\alpha,\beta}} = n! \frac{1}{(\theta + 2n - 1)(\theta)_{(n-1)}} \frac{(\alpha)_{(n)}}{(\beta)_{(n)}}. \tag{2.18}$$

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

$$\zeta_n^{\alpha_j, |\alpha|-\alpha_j} = Q_n^\alpha(\mathbf{e}_j, \mathbf{e}_j) = \sum_{m=0}^n a_{nm}^{|\alpha|} \frac{(|\alpha|)_{(m)}}{(\alpha_j)_{(m)}}. \tag{2.19}$$

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 $\mathbf{X}$ has distribution D_α, then $A\mathbf{X}$ has a $D_{A\alpha}$ distribution. Similarly, with some easy computation

$$\mathbb{E}\big(\xi_m^\alpha(\mathbf{X}, \mathbf{y}) | A\mathbf{X} = a\mathbf{x}\big) = \xi_m^{A\alpha}(A\mathbf{X}, A\mathbf{y}).$$

One has therefore the following:

Proposition 2.4. *A representation for Polynomial kernels in $D_{A\alpha}$ is*

$$Q_n^{A\alpha}(A\mathbf{x}, A\mathbf{y}) = \mathbb{E}[Q_n^\alpha(\mathbf{X}, \mathbf{y}) | A\mathbf{X} = A\mathbf{x}]. \tag{2.20}$$

Example 2.5. For any $\alpha \in \mathbb{R}^d$ and $k \le d$, suppose $A\mathbf{X} = (X_1 + \cdots + X_k, X_{k+1} + \cdots + X_d) = X'$. Then, denoting $\alpha' = \alpha_1 + \cdots + \alpha_k$ and $\beta' = \alpha_{k+1} + \cdots + \alpha_d$, one has

$$Q_n^{A\alpha}(x', y') = \zeta_n^{\alpha',\beta'} R_n^{\alpha',\beta'}(x') R_n^{\alpha',\beta'}(y') = \mathbb{E}[Q_n^\alpha(\mathbf{X}, \mathbf{y}) | X' = x'].$$

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. [15], Proposition 5.2. Let $\{Q_{\mathbf{n}}^\circ(\mathbf{x})\}$ be a orthonormal polynomial set on the Dirichlet, considered as functions of $(x_1, \dots, x_{d-1})$. Define, for $\mathbf{r} \in N\Delta_{(d-1)}$,

$$h_{\mathbf{n}}^\circ(\mathbf{r}; N) = \int Q_{\mathbf{n}}^\circ(\mathbf{x}) D_{\alpha+\mathbf{r}}(d\mathbf{x}), \tag{3.1}$$

then $\{h_{\mathbf{n}}^\circ\}$ is a system of multivariate orthogonal polynomials with respect to $\mathrm{DM}_{\alpha,N}$ with constant of orthogonality

$$\mathbb{E}_{\alpha,N}[h_{\mathbf{n}}^\circ(\mathbf{R}; N)]^2 = \frac{(N)_{[n]}}{(|\alpha| + N)_{(n)}}. \tag{3.2}$$

Note also that if $N \to \infty$ with $r_i/N \to x_i$, $i = 1, \dots, d$, then

$$\lim_{N\to\infty} h_{\mathbf{n}}^\circ(\mathbf{r}; N) = Q_{\mathbf{n}}^\circ(\mathbf{x}).$$

Proposition 3.1. *The Hahn kernel polynomials with respect to* $\mathrm{DM}_{\alpha,N}$ *are*

$$H_n^\alpha(\mathbf{r}, \mathbf{s}) = \frac{(|\alpha| + N)_{(n)}}{N_{[n]}} \int\int Q_n^\alpha(\mathbf{x}, \mathbf{y}) D_{\alpha+\mathbf{r}}(d\mathbf{x}) D_{\alpha+\mathbf{s}}(d\mathbf{y}) \tag{3.3}$$

for $\mathbf{r} = (r_1, \dots, r_d)$, $\mathbf{s} = (s_1, \dots, s_d)$, $|\mathbf{r}| = |\mathbf{s}| = N$ *fixed, and* $n = 0, 1, \dots, N$.

An explicit expression is

$$H_n^\alpha(\mathbf{r}, \mathbf{s}) = \frac{(|\alpha| + N)_{(n)}}{r_{[n]}} \cdot \sum_{m=0}^{n} a_{nm}^{|\alpha|} \xi_m^{H,\alpha}(\mathbf{r}, \mathbf{s}), \tag{3.4}$$

where $(a_{nm}^{|\alpha|})$ *is as in* (2.6) *and* $\xi_m^{H,\alpha}(\mathbf{r}, \mathbf{s})$ *is given by* (2.2).

Proof. The kernel sum is, by definition,

$$H_n^\alpha(\mathbf{r}, \mathbf{s}) = \frac{(|\alpha| + N)_{(n)}}{N_{[n]}} \sum_{\{\mathbf{n}:|\mathbf{n}|=n\}} h_{\mathbf{n}}^\circ(\mathbf{r}; N) h_{\mathbf{n}}^\circ(\mathbf{s}; N), \tag{3.5}$$

and from (3.3), (3.4) follows. The form of $\xi_m^{H,\alpha}$ is obtained by taking the expectation of $\xi_m^\alpha(\mathbf{X}, \mathbf{Y})$, appearing in the representation (2.5) of Q_n^α, with respect to the product measure $D_{\alpha+\mathbf{r}} D_{\alpha+\mathbf{s}}$. □

The first polynomial kernel is

$$H_1^{\alpha}(\mathbf{r},\mathbf{s}) = \frac{(|\alpha|+1)(|\alpha|+r)}{|\alpha|}\left(\frac{|\alpha|}{(|\alpha|+N)^2}\sum_{1}^{d}\frac{(\alpha_i+r_i)(\alpha_i+s_i)}{\alpha_i} - 1\right).$$

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 $N \in \mathbb{N}$ and $d \in \mathbb{N}$, denote $\mathbf{r}_{j,1} = N\mathbf{e}_j \in \mathbb{N}^d$, where $\mathbf{e}_j = (0,\dots,0,1,0,\dots,0)$ with 1 only at the jth coordinate.*

For every $\alpha \in \mathbb{N}^d$,

$$H_n^{\alpha}(\mathbf{s}, N\mathbf{e}_j) = \frac{1}{c_{N,n}^{|\alpha|}} h_n^{\circ(\alpha_j,|\alpha|-\alpha_j)}(s_j;N)\,h_n^{\circ(\alpha_j,|\alpha|-\alpha_j)}(N;N), \qquad |\mathbf{s}| = N, \tag{3.6}$$

where

$$c_{N,n}^{|\alpha|} := \frac{(N)_{[n]}}{(|\alpha|+N)_{(n)}} = \mathbb{E}\big[h_n^{\circ(\alpha,\beta)}(R;N)^2\big],$$

and $\{h_n^{\circ(\alpha_j,|\alpha|-\alpha_j)}\}$ are orthogonal polynomials with respect to $\mathrm{DM}_{(\alpha_j,|\alpha|-\alpha_j),N}$.

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

$$H_n^{\alpha}(\mathbf{s},\mathbf{r}) = \frac{1}{c_{N,n}^{|\alpha|}}\sum_{m=0}^{n} a_{nm}^{|\alpha|}\xi_m^{H,\alpha}(\mathbf{r},\mathbf{s})$$

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

$$h_n^{\circ(\alpha,\beta)}(k;N)\,h_n^{\circ(\alpha,\beta)}(j;N) = \sum_{m=0}^{n} a_{nm}^{|\alpha|}\xi_m^{H,\alpha,\beta}(k,j), \qquad k,j = 0,\dots,N.$$

Now, for $\mathbf{r},\mathbf{s} \in N\Delta_{(d-1)}$, rewrite $\xi_m^{H,\alpha}$ as

$$\begin{aligned}\xi_m^{H,\alpha}(\mathbf{s},\mathbf{r}) &= \sum_{|\mathbf{l}|=m}\frac{\mathrm{DM}_{\alpha+\mathbf{s},m}(\mathbf{l})\,\mathrm{DM}_{\alpha+\mathbf{r},m}(\mathbf{l})}{\mathrm{DM}_{\alpha,m}(\mathbf{l})}\\ &= \sum_{|\mathbf{l}|=m}\mathrm{DM}_{\alpha+\mathbf{s},m}(\mathbf{l})\frac{\mathrm{DM}_{\alpha+\mathbf{l},N}(\mathbf{r})}{\mathrm{DM}_{\alpha,N}(\mathbf{r})}.\end{aligned} \tag{3.7}$$

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

$$\mathrm{DM}_{\alpha,m}(\mathbf{l}) = \mathrm{DM}_{(\alpha_1,|\alpha|-\alpha_1),m}(l_1)\,\mathrm{DM}_{(\alpha_2,\dots,\alpha_d),m-l_1}(l_2,\dots,l_d),$$

then

$$\xi_m^{H,\alpha}(\mathbf{s}, Ne_1) = \sum_{l_1=0}^{m} \mathrm{DM}_{(\alpha_1+s_1,|\alpha|-\alpha_1+m-s_1),m}(l_1)\frac{\mathrm{DM}_{(\alpha_1+l_1,|\alpha|-\alpha_1+m-l_1),N}(N)}{\mathrm{DM}_{(\alpha_1,|\alpha|-\alpha_1),N}(N)}$$
$$\times \sum_{|\mathbf{u}|=m-l_1} \mathrm{DM}_{\alpha'+\mathbf{s}',m-l_1}(\mathbf{u})\frac{\mathrm{DM}_{\alpha+l,0}(0)}{\mathrm{DM}_{\alpha,0}(0)}$$
$$= \sum_{l_1=0}^{m} \mathrm{DM}_{\alpha+\mathbf{s},m}(l_1)\frac{\mathrm{DM}_{\alpha_1+l_1,N}(N)}{\mathrm{DM}_{\alpha,N}(N)} \sum_{|\mathbf{u}|=m-l_1} \mathrm{DM}_{\alpha'+\mathbf{s}',m-l_1}(\mathbf{u})$$
$$= \sum_{l_1=0}^{m} \mathrm{DM}_{\alpha+\mathbf{s},m}(l_1)\frac{\mathrm{DM}_{\alpha_1+l_1,N}(N)}{\mathrm{DM}_{\alpha,N}(N)} \tag{3.8}$$
$$= \xi_m^{H,\alpha_1,|\alpha|-\alpha_1}(s_1, N). \tag{3.9}$$

Then (3.6) follows immediately. □

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

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

$$h_n^{\alpha,\beta}(r; N) = {}_3F_2\left(\begin{matrix}-n, n+\theta-1, -r\\ \alpha, -N\end{matrix}\,\middle|\,1\right), \qquad n = 0, 1, \ldots, N, \tag{3.10}$$

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

$$\frac{1}{u_{N,n}^{\alpha,\beta}} := \mathbb{E}_{\alpha,\beta}[h_n^{\alpha,\beta}(R; N)]^2 = \frac{1}{\binom{N}{n}} \frac{(\theta+N)_{(n)}}{(\theta)_{(n-1)}} \frac{1}{\theta+2n-1} \frac{(\beta)_{(n)}}{(\alpha)_{(n)}}. \tag{3.11}$$

The following product formula was found by Gasper [8]:

$$h_n^{\alpha,\beta}(r; N)h_n^{\alpha,\beta}(s; N) = \frac{(-1)^n(\beta)_{(n)}}{(\alpha)_{(n)}} \sum_{l=0}^{n}\sum_{k=0}^{n-l} \frac{(-1)^{l+k} n_{[l+k]}(\theta+n-1)_{(l+k)} r_{[l]} s_{[l]} (N-r)_{[k]}(N-s)_{[k]}}{l!k!N_{[l+k]}N_{[l+k]}(\alpha)_{(l)}(\beta)_{(k)}}. \tag{3.12}$$

Thus

$$u_{N,n}^{\alpha,\beta} h_n^{\alpha,\beta}(r; N)h_n^{\alpha,\beta}(s; N)$$
$$= \frac{N_{[n]}}{(\theta+N)_{(n)}} \sum_{m=0}^{n} \frac{(-1)^{n-m}(\theta)_{(n-1)}(\theta+n-1)_{(m)}(\theta+2n-1)}{m!(n-m)!(\theta)_{(m)}} \chi_m^{H,\alpha,\beta}(r, s) \tag{3.13}$$
$$= \frac{N_{[n]}}{(\theta+N)_{(n)}} \sum_{m=0}^{n} a_{nm}^{\theta} \chi_m^{H,\alpha,\beta}(r, s),$$

where

$$\chi_m^{H,\alpha,\beta}(r,s) := \sum_{j=0}^{m} \frac{1}{\mathrm{DM}_{(\alpha,\beta),m}(j)} \left[\frac{\binom{m}{j} r_{[j]}(N-r)_{[m-j]}}{N_{[m]}}\right]\left[\frac{\binom{m}{j} s_{[j]}(N-s)_{[m-j]}}{N_{[m]}}\right]. \quad (3.14)$$

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* $r, s \in \{0, \dots, N\}$,

$$\xi_m^{H,\alpha,\beta}(r,s) = \sum_{l=0}^{m} b_{ml} \chi_l^{H,\alpha,\beta}(r,s), \quad (3.15)$$

where

$$b_{ml} = \sum_{n=l}^{m} \left(\frac{N_{[n]}}{(\theta+N)_{(n)}}\right)^2 \frac{m_{[n]}}{(\theta+m)_{(n)}} a_{nl}^{\theta}. \quad (3.16)$$

Proof. From (3.4),

$$u_{N,n}^{\alpha,\beta} h_n^{\alpha,\beta}(r;N) h_n^{\alpha,\beta}(s;N) = H_n^{\alpha,\beta}(r,s) = \frac{(\theta+N)_{(n)}}{N_{[n]}} \sum_{m=0}^{n} a_{nm}^{\theta} \xi_m^{H,\alpha,\beta}(r,s). \quad (3.17)$$

Since the array $A = (a_{nm}^{\theta})$ has inverse $C = A^{-1}$ with entries

$$c_{mn}^{\theta} = \left(\frac{m_{[n]}}{(\theta+m)_{(n)}}\right), \quad (3.18)$$

then equating (3.17) and (3.13) leads to

$$\begin{aligned}
\xi_m^{H,\alpha,\beta} &= \sum_{n=0}^{m} c_{mn}^{\theta} \frac{N_{[n]}}{(\theta+N)_{(n)}} H_n^{\alpha,\beta} \\
&= \sum_{n=0}^{m} c_{mn}^{\theta} \left(\frac{N_{[n]}}{(\theta+N)_{(n)}}\right)^2 \sum_{l=0}^{n} a_{nl}^{\theta} \chi_l^{H,\alpha,\beta} \\
&= \sum_{l=0}^{m} b_{ml} \chi_l^{H,\alpha,\beta}.
\end{aligned}$$

□

The following corollary is then straightforward.

Corollary 3.4.

$$\mathbb{E}[\xi_m^{H,\alpha,\beta} \chi_l^{H,\alpha,\beta}] = \mathbb{E}[\xi_l^{H,\alpha,\beta} \chi_m^{H,\alpha,\beta}] = \sum_{n=0}^{m \wedge l} \frac{m_{[l]} l_{[n]}}{(\theta+m)_{(n)} (\theta+l)_{(n)}}.$$

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

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

Gasper's product formula (3.12), or, rather, the representation (3.13), 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*:

$$H_n^\alpha(\mathbf{r},\mathbf{s}) = \frac{N_{[n]}}{(|\alpha|+N)_{(n)}} \sum_{m=0}^{n} a_{nm}^{|\alpha|} \chi_m^{H,\alpha}(\mathbf{r},\mathbf{s}), \qquad \mathbf{r},\mathbf{s} \in N\Delta_{(d-1)},\ n = 0,1,\ldots, \tag{3.19}$$

where

$$\chi_m^{H,\alpha}(\mathbf{r},\mathbf{s}) := \sum_{\mathbf{l}:|\mathbf{l}|=m} \frac{1}{\mathrm{DM}_{\alpha,m}(\mathbf{l})} \left(\frac{\binom{m}{\mathbf{l}} p_{\mathbf{l}}(\mathbf{r})}{N_{[m]}}\right)\left(\frac{\binom{m}{\mathbf{l}} p_{\mathbf{l}}(\mathbf{s})}{N_{[m]}}\right). \tag{3.20}$$

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

$$\chi_m^{H,\alpha}(\mathbf{r},\mathbf{s}) = \sum_{n=0}^{m} \frac{c_{mn}^{|\alpha|}}{c_{Nn}^{|\alpha|}} H_n^\alpha(\mathbf{r},\mathbf{s}),$$

where $c_{ij}^{|\alpha|}$ are given by (3.18) (independent of d!), then the proof follows by inversion.

Consider the orthonormal multivariate Jacobi polynomials $Q_n^\circ(\mathbf{x})$. The functions

$$h_{\mathbf{n}}^\circ(\mathbf{r};N) := \int_{\Delta_{(d-1)}} Q_{\mathbf{n}}^\circ(\mathbf{x}) D_{\alpha+\mathbf{r}}(\mathrm{d}\mathbf{x})$$

satisfy the identity

$$\mathbb{E}\left[h_{\mathbf{n}}^\circ(\mathbf{R};N)\binom{m}{\mathbf{l}} p_{\mathbf{l}}(\mathbf{R})\right] = N_{[m]} h_{\mathbf{n}}^\circ(\mathbf{l};m)\,\mathrm{DM}_{\alpha,m}(\mathbf{l}), \qquad \mathbf{l} \in m\Delta_{(d-1)}, \mathbf{n} \in \mathbb{Z}_+^d \tag{3.21}$$

([14], (5.71)).

Then for every fixed $\mathbf{s}$,

$$\mathbb{E}[\chi_m^{H,\alpha}(\mathbf{R},\mathbf{s}) h_{\mathbf{n}}^\circ(\mathbf{R};N)] = \sum_{l=m} \binom{m}{\mathbf{l}} \frac{p_{\mathbf{l}}(\mathbf{s})}{N_{[m]}} h_{\mathbf{n}}^\circ(\mathbf{l};m), \tag{3.22}$$

so, iterating the argument, we can write

$$\mathbb{E}[\chi_m^{H,\alpha}(\mathbf{R},\mathbf{S}) h_{\mathbf{n}}^\circ(\mathbf{R};N) h_{\mathbf{n}}^\circ(\mathbf{S};N)] = c_{mn}. \tag{3.23}$$

Now, by uniqueness of the polynomial kernel,

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

therefore

$$\chi_m^{H,\alpha}(\mathbf{r}, \mathbf{s}) = \sum_{n=0}^{m} \frac{c_{mn}^{|\alpha|}}{c_{Nn}^{|\alpha|}} H_n^\alpha(\mathbf{r}, \mathbf{s}),$$

and the proof is complete. □

The connection coefficients between $\xi_m^{H,\alpha}$ and ξ_m^α 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)

$$\xi_m^{H,\alpha,\beta}(\mathbf{r}, \mathbf{s}) = \sum_{l=0}^{m} b_{ml} \chi_l^{H,\alpha,\beta}(\mathbf{r}, \mathbf{s}), \tag{3.24}$$

where (b_{ml}) are given by (3.16).

(ii)

$$\mathbb{E}[\xi_m^{H,\alpha} \chi_l^{H,\alpha}] = \mathbb{E}[\xi_l^{H,\alpha} \chi_m^{H,\alpha}] = \sum_{n=0}^{m \wedge l} \frac{m_{[l]} l_{[n]}}{(|\alpha| + m)_{(n)} (|\alpha| + l)_{(n)}}, \qquad m, l = 0, 1, 2, \ldots.$$

3.2. Polynomial kernels on the hypergeometric distribution

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

$$\frac{\binom{c_1}{r_1} \cdots \binom{c_d}{r_d}}{\binom{|c|}{r}} \tag{3.25}$$

by replacing α by $-c$ in (3.4) and (2.2). Again a direct proof similar to that for $Q_n^\alpha(\mathbf{x}, \mathbf{y})$ would be possible.

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

From Dirichlet–Jacobi polynomial kernels we can also derive polynomial kernels orthogonal with respect to symmetrized Dirichlet measures. Let $D_{|\alpha|,d}$ be the Dirichlet distribution on d

points with symmetric parameters $(|\alpha|/d, \dots, |\alpha|/d)$, and $D^{\downarrow}_{|\alpha|,d}$ its ranked version. Denote with $Q_n^{(|\alpha|,d)}$ and $Q_n^{(|\alpha|,d)\downarrow}$ the corresponding n-kernels.

Proposition 4.1.

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

where summation is over all permutations σ *of* $1, \dots, d$. *The kernel polynomials have a similar form to* $Q_n^{(|\alpha|,d)}$, *but with* $\xi_m^{(|\alpha|,d)}$ *replaced by*

$$\xi_m^{(|\alpha|,d)\downarrow} = \sum_{\mathbf{l} \in m\Delta^{\downarrow}_{(d-1)}} \frac{m!|\theta|_{(m)}(d-k)!(\prod_1^m \beta_i(\mathbf{l})!)[\mathbf{x};\mathbf{l}][\mathbf{y};\mathbf{l}]}{d!\prod_1^m [j!(|\theta|/d)_{(j)}]^{\beta_j(\mathbf{l})}} \tag{4.1}$$

$$= \sum_{\mathbf{l} \in m\Delta^{\downarrow}_{(d-1)}} \frac{\sharp(\mathbf{l})[\mathbf{x};\mathbf{l}]\sharp(\mathbf{l})[\mathbf{y};\mathbf{l}]}{\mathrm{DM}^{\downarrow}_{|\alpha|,m,d}(\mathbf{l})}, \tag{4.2}$$

where

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

Proof. Note that

$$\begin{aligned}
Q_n^{(|\alpha|,d)\downarrow}(\mathbf{x},\mathbf{y}) &= \frac{1}{d!} \sum_{\sigma \in \mathcal{G}_d} Q_n^{(|\alpha|,d)}(\sigma\mathbf{x},\mathbf{y}) \\
&= \frac{1}{d!} \sum_{\sigma \in \mathcal{G}_d} \sum_{m \le n} a_{nm}^{|\alpha|} \xi_m^{(|\alpha|,d)}(\sigma\mathbf{x},\mathbf{y}) \\
&= d! \sum_{m \le n} a_{nm}^{|\alpha|} \frac{1}{(d!)^2} \sum_{\sigma \in \mathcal{G}_d} \sum_{|\mathbf{l}|=m} \frac{\binom{m}{\mathbf{l}}^2 (\sigma\mathbf{x})^{\mathbf{l}} \mathbf{y}^{\mathbf{l}}}{\mathrm{DM}_{|\alpha|,m,d}(\mathbf{l})} \\
&= \sum_{m \le n} a_{nm}^{|\alpha|} \frac{1}{(d!)^2} \sum_{\sigma,\tau \in \mathcal{G}_d} \sum_{|\mathbf{l}|=m} \frac{\binom{m}{\mathbf{l}}^2 (\sigma\tau\mathbf{x})^{\mathbf{l}} (\mathbf{y})^{\mathbf{l}}}{\mathrm{DM}_{|\alpha|,m}(\mathbf{l})} \\
&= \sum_{m \le n} a_{nm}^{|\alpha|} \frac{1}{(d!)^2} \sum_{\sigma,\tau \in \mathcal{G}_d} \sum_{|\mathbf{l}|=m} \frac{\binom{m}{\mathbf{l}}^2 (\sigma\mathbf{x})^{\mathbf{l}} (\tau\mathbf{y})^{\mathbf{l}}}{\mathrm{DM}_{|\alpha|,m,d}(\mathbf{l})} \qquad (4.3) \\
&= \frac{1}{(d!)^2} \sum_{\sigma,\tau \in \mathcal{G}_d} Q_n^{(|\alpha|,d)}(\sigma\mathbf{x}, \tau\mathbf{y}). \qquad (4.4)
\end{aligned}$$

Now,

$$\mathbb{E}_{D^{\downarrow}_{(|\alpha|,d)}}\big[Q_n^{(|\alpha|,d)\downarrow}(\mathbf{x},\mathbf{Y})Q_m^{(|\alpha|,d)\downarrow}(\mathbf{z},\mathbf{Y})\big] = \frac{1}{d!}\sum_{\sigma\in\mathcal{G}_d} Q_n^{(|\alpha|,d)}(\sigma\mathbf{x},\mathbf{z})\delta_{nm} = Q_n^{(|\alpha|,d)\downarrow}(\mathbf{x},\mathbf{z})\delta_{nm}, \tag{4.5}$$

and hence $Q_n^{(|\alpha|,d)\downarrow}$ is the n polynomial kernel with respect to $D^{\downarrow}_{(|\alpha|,d)}$. The second part of the theorem, involving identity (4.2), is just another way of rewriting (4.3). □

Remark 4.2. The first polynomial is $Q_1^{(|\alpha|,d)\downarrow} \equiv 0$.

4.1. Infinite-dimensional limit

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

$$\xi_m^{(|\alpha|,\infty)\downarrow} = |\alpha|_{(m)} \sum \frac{m!(\prod_1^m b_i!)[\mathbf{x};\mathbf{l}][\mathbf{y};\mathbf{l}]}{|\alpha|^k[0!1!]^{b_1}\cdots[(k-1)!k!]^{b_k}} \tag{4.6}$$

$$= \sum \frac{\sharp(\mathbf{l})[\mathbf{x};\mathbf{l}]\binom{m}{\mathbf{l}}\sharp(\mathbf{l})[\mathbf{y};\mathbf{l}]}{\mathrm{ESF}_{|\alpha|}(\mathbf{l})}. \tag{4.7}$$

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

$$Q_n^{(|\alpha|,\infty)\downarrow} = \sum_{m=0}^{n} a_{nm}^{|\alpha|}\xi_m^{(|\alpha|,\infty)\downarrow}. \tag{4.8}$$

The first polynomial is zero, and the second polynomial is

$$Q_2^{\infty} = (F_1-\mu)(F_2-\mu)/\sigma^2,$$

where

$$F_1 = \sum_1^{\infty} x_{(i)}^2, \qquad F_2 = \sum_1^{\infty} y_{(i)}^2,$$

and

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

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 $\mathrm{DM}^{\downarrow}_{|\alpha|,N,d}$ as $d \to \infty$. The proof of the following proposition can be obtained by the same arguments used to prove Proposition 4.1.

Proposition 4.4.

(i) *The polynomial kernels with respect to* $\mathrm{DM}^{\downarrow}_{|\alpha|,N,d}$ *are of the same form as* (3.4), *but with* $\xi_m^{H,(|\alpha|,d)}$ *replaced by*

$$\xi_m^{H,(|\alpha|,d)\downarrow} := (d!)^{-1} \sum_{\pi} \xi_m^{H,(|\alpha|,d)}(\pi(\mathbf{r}), \mathbf{s}). \tag{4.9}$$

(ii) *The kernel polynomials with respect to* $\mathrm{ESF}_{|\alpha|}$ *are derived by considering the limit form* $\xi_m^{H,|\alpha|\downarrow}$ *of* $\xi_m^{H,(|\alpha|,d)\downarrow}$. *This has the same form as* $\xi_m^{|\alpha|\downarrow}$ (4.7) *with* $[\mathbf{x};\mathbf{b}][\mathbf{y};\mathbf{b}]$ *replaced by* $[\mathbf{r};\mathbf{b}]'[\mathbf{s};\mathbf{b}]'$, *where*

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

and summation is over $\sum_1^m j b_j = m$, $\sum_1^m b_j = k$, $k = 1, \ldots, m$. *The kernel polynomials have the same form as* (3.4) *with* $\xi_m^{H,(|\alpha|,d)}$ *replaced by* $\xi_m^{H,|\alpha|\downarrow}$. *The first polynomial is identically zero under this symmetrization.*

5. Integral representation for Jacobi polynomial kernels

This section and Section 6 are 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 for $d = 2$ ([8,20] and [9]). 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. The relationship between our integral representation and a d-dimensional Jacobi product formula due to Koornwinder and Schwartz [21] is also explained (Section 5.3).

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

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

$$R_n^{\alpha,\beta}(x) = \frac{Q_n^{\alpha,\beta}(x,1)}{Q_n^{\alpha,\beta}(1,1)}. \tag{5.1}$$

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

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

via the transformation

$$R_n^{\alpha,\beta}(x) = \frac{P_n^{\beta-1,\alpha-1}(2x-1)}{P_n^{\beta-1,\alpha-1}(1)}. \tag{5.2}$$

The constant of orthogonality $\zeta_n^{(\alpha,\beta)}$ is given by (2.18).

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 [8].

Theorem 5.1 (Gasper [8]). *A necessary and sufficient condition for the equality*

$$\frac{P_n^{a,b}(x)}{P_n^{a,b}(1)}\frac{P_n^{a,b}(y)}{P_n^{a,b}(1)} = \int_{-1}^{1}\frac{P_n^{a,b}(z)}{P_n^{a,b}(1)}\widetilde{m}_{x,y;a,b}(\mathrm{d}z), \tag{5.3}$$

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

$$\frac{\mathrm{d}\widetilde{m}_{x,y;a,b}}{\mathrm{d}w_{a,b}}(z) = \sum_{n=0}^{\infty}\phi_n\frac{P_n^{a,b}(x)}{P_n^{a,b}(1)}\frac{P_n^{a,b}(y)}{P_n^{a,b}(1)}\frac{P_n^{a,b}(z)}{P_n^{a,b}(1)}, \tag{5.4}$$

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

An explicit formula for the density (5.4) is possible when $a \ge b > -1/2$.

$$\frac{P_n^{a,b}(x)}{P_n^{a,b}(1)}\frac{P_n^{a,b}(y)}{P_n^{a,b}(1)} = \int_0^1\int_0^{\pi}\frac{P_n^{a,b}(\psi)}{P_n^{a,b}(1)}\widetilde{m}_{a,b}(\mathrm{d}u,\mathrm{d}\omega), \tag{5.5}$$

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

$$\widetilde{m}_{a,b}(\mathrm{d}u,\mathrm{d}\omega) = \frac{2\Gamma(a+1)}{\sqrt{\pi}\Gamma(a-b)\Gamma(b+1/2)}(1-u^2)^{a-b-1}u^{2b+1}(\sin\omega)^{2b}\,\mathrm{d}u\,\mathrm{d}\omega. \tag{5.6}$$

See [20] for an analytic proof of this formula. Note that $\phi(1,1;u,\omega)=1$, so $\mathrm{d}\widetilde{m}_{a,b}(u,\omega)$ 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*

$$R_n^{\alpha,\beta}(x)R_n^{\alpha,\beta}(y) = \int_0^1 R_n^{\alpha,\beta}(z)m_{x,y;\alpha,\beta}(\mathrm{d}z) \tag{5.7}$$

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* $\mathrm{d}\widetilde{m}$ *is defined by* (5.4). *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*

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

where

$$K(x,y,z) = \sum_{n=0}^{\infty} \zeta_n^{\alpha,\beta} R_n^{\alpha,\beta}(x) R_n^{\alpha,\beta}(y) R_n^{\alpha,\beta}(z) \geq 0. \tag{5.8}$$

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 (5.5)–(5.6). Let $m_{\alpha\beta}(\mathrm{d}u, \mathrm{d}\omega) = \widetilde{m}_{\beta-1,\alpha-1}(\mathrm{d}u, \mathrm{d}\omega)$. Then

$$R_n^{\alpha,\beta}(x) R_n^{\alpha,\beta}(y) = \int_0^1 \int_0^{\pi} R_n^{\alpha,\beta}(\varphi) m_{\alpha\beta}(\mathrm{d}u, \mathrm{d}\omega), \tag{5.9}$$

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

$$\varphi(x,y;u,\omega) = xy + (1-x)(1-y) + 2u \cos\omega \sqrt{x(1-x)y(1-y)}. \tag{5.10}$$

In ϕ set $x \leftarrow 2x-1$, $y \leftarrow 2y-1$ to obtain (5.10).

5.2. Integral representation for $d > 2$

An extension of the product formula (5.7) 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* $j = 1, \dots, d$, $\alpha_j \leq \sum_{i=1}^{j-1} \alpha_i$ *and* $1/2 \leq \alpha_j$, *or* $\sum_{i=1}^{j} \alpha_i \geq 2$. *Then, for every* $\mathbf{x}, \mathbf{y} \in \Delta_{(d-1)}$ *and every integer* n,

$$Q_n^{\alpha}(\mathbf{x}, \mathbf{y}) = \mathbb{E}\big[Q_n^{\alpha_d, |\alpha|-\alpha_d}(Z_d, 1) | \mathbf{x}, \mathbf{y}\big], \tag{5.11}$$

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

$$Z_1 \equiv 1; \qquad Z_j = \Phi_j D_j Z_{j-1}, \qquad j = 2, \dots, d, \tag{5.12}$$

with

$$\begin{aligned} D_j &:= \frac{(1-x_j)(1-y_j)}{(1-X_j^*)(1-Y_j^*)}; \qquad X_j^* := \frac{x_j}{1 - x_j(1-\sqrt{Z}_{j-1})}; \\ Y_j^* &:= \frac{y_j}{1 - y_j(1-\sqrt{Z}_{j-1})}, \end{aligned} \tag{5.13}$$

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

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

where $\mathrm{d}m_{x,y;\alpha,\beta}$ *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 $\alpha_{(1)} + \alpha_{(2)} \geq 2$ to obtain the representation (5.11).

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* $m \in \mathbb{N}, d = 2, 3, \ldots$ *and* $\alpha \in \mathbb{R}^d$ *satisfying the assumptions of Proposition* 5.4,

$$\xi_m^\alpha(\mathbf{x}, \mathbf{y}) = \frac{|\alpha|_{(m)}}{(\alpha_d)_{(m)}} \mathbb{E}[Z_d^m | \mathbf{x}, \mathbf{y}], \tag{5.14}$$

where Z_d *is defined as in Proposition* 5.4.

Let $\theta = \alpha + \beta$. Assume the lemma is true. From (5.9) and (5.16) we know that, for every $n = 0, 1, \ldots$ and every $s \in [0,1]$,

$$Q_n^{\alpha,\beta}(s,1) = \sum_{m \leq n} a_{nm}^\theta \frac{(\theta)_{(m)}}{\alpha_{(m)}} s^m.$$

Thus from (5.14)

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

which is what is claimed in Proposition 5.4.

Now we proceed with the proof of the lemma.

Proof of Lemma 5.5. The proof is by induction.

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

$$\xi_m^{(\alpha,\beta)}(x,y) = \sum_{j=0}^m \binom{m}{j} \frac{(\alpha+\beta)_{(m)}}{(\alpha)_{(j)}(\beta)_{(m-j)}} (xy)^j [(1-x)(1-y)]^{m-j}. \tag{5.15}$$

Setting $y = 1$, the only positive addend in (5.15) is the one with $j = m$, so

$$\xi_m^{(\alpha,\beta)}(x,1) = \frac{(\alpha+\beta)_{(m)}}{(\alpha)_{(m)}} z^m. \tag{5.16}$$

Therefore, if $\theta=\alpha+\beta$, from (5.9) and (5.16), we conclude

$$\begin{aligned}\xi_m^{\alpha,\beta}(x,y) &= \sum_{j=0}^{m}\binom{m}{j}\frac{(\theta)_{(m)}}{(\alpha)_{(j)}(\beta)_{(m-j)}}(xy)^j[(1-x)(1-y)]^{m-j}\\ &= \frac{(\theta)_{(m)}}{(\alpha)_{(m)}}\int_{[0,1]} z^m m_{x,y;\alpha,\beta}(\mathrm{d}z).\end{aligned}\tag{5.17}$$

Thus the proposition is true for $d=2$.

To prove the result for any general $d>2$, consider

$$\begin{aligned}\xi_m^{\alpha}(\mathbf{x},\mathbf{y}) = \sum_{m_d=0}^{m}\binom{m}{m_d}(x_d y_d)^{m_d}[(1-x_d)(1-y_d)]^{m-m_d}\frac{(|\alpha|)_m}{(\alpha_d)_{(m_d)}(|\alpha|-\alpha_d)_{(m-m_d)}}\\ \times\sum_{\tilde{m}\in\mathbb{N}^{d-1}:|\tilde{m}|=m-m_d}\binom{m-m_d}{\tilde{m}}\frac{(|\alpha|-\alpha_d)_{(m-m_d)}}{\prod_{i=1}^{d-1}(\alpha_i)_{(\tilde{m}_i)}}\prod_{i=1}^{d-1}(\tilde{x}_i\tilde{y}_i)^{\tilde{m}_i},\end{aligned}\tag{5.18}$$

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

Now assume the proposition is true for $d-1$. Then the inner sum of (5.18) has a representation like (5.14), and we can write

$$\begin{aligned}\xi_m^{\alpha}(\mathbf{x},\mathbf{y}) = \sum_{m_d=0}^{m}\binom{m}{m_d}(x_d y_d)^{m_d}[(1-x_d)(1-y_d)]^{m-m_d}\\ \times\frac{(|\alpha|)_m}{(\alpha_d)_{(m_d)}(|\alpha|-\alpha_d)_{(m-m_d)}}\\ \times\frac{(|\alpha|-\alpha_d)_{(m-m_d)}}{(\alpha_{d-1})_{(m-m_d)}}\mathbb{E}[Z_{d-1}^{m-m_d}|\widetilde{\mathbf{x}},\widetilde{\mathbf{y}}],\end{aligned}\tag{5.19}$$

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

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

and define the random variable

$$D_d := \frac{(1-x_d)(1-y_d)}{(1-X_d^*)(1-Y_d^*)}.\tag{5.20}$$

Then simple algebra leads to rewriting equation (5.19) as

$$\xi_m^\alpha(\mathbf{x},\mathbf{y}) = \mathbb{E}\Bigg[\frac{|\alpha|_{(m)}(D_d Z_{d-1})^m}{(\alpha_{d-1}+\alpha_d)_{(m)}}\Bigg(\sum_{m_d=0}^{m}\binom{m}{m_d}\frac{(\alpha_{d-1}+\alpha_d)_{(m)}}{(\alpha_d)_{(m_d)}(\alpha_{d-1})_{(m-m_d)}} \times (X_d^* X_d^*)^{m_d}[(1-X_d^*)(1-Y_d^*)]^{m-m_d}\Bigg)\Bigg|\mathbf{x},\mathbf{y}\Bigg]. \tag{5.21}$$

Now the sum in (5.21) is of the form (5.15), 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 equality (5.17) to obtain

$$\begin{aligned}\xi_m^\alpha(\mathbf{x},\mathbf{y}) &= \mathbb{E}\bigg[\frac{(|\alpha|)_{(m)}}{(\alpha_d)_{(m)}}(D_d Z_{d-1})^m \mathbb{E}(\Phi_d^m | X_d^*, Y_d^*)\big|\mathbf{x},\mathbf{y}\bigg] \\ &= \frac{(|\alpha|)_{(m)}}{(\alpha_d)_{(m)}}\mathbb{E}[Z_d^m|\mathbf{x},\mathbf{y}]\end{aligned} \tag{5.22}$$

(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_α: $\alpha \in \mathbb{R}^d$, a product formula is proved in [21]. For every $\mathbf{x} \in \Delta_{(d-1)}$, $\alpha \in \mathbb{R}_+^d$ and $\mathbf{n} = (n_1, \dots, n_{d-1})$: $|\mathbf{n}| = n$, these polynomials can be written as

$$R_{\mathbf{n}}^\alpha(\mathbf{x}) = \prod_{j=1}^{d-1}\bigg[R_{n_j}^{\alpha_j, E_j+2N_j}\bigg(\frac{x_j}{1-\sum_{i=1}^{j-1}x_i}\bigg)\bigg]\bigg(1-\frac{x_j}{1-\sum_{i=1}^{j-1}x_i}\bigg)^{N_j}, \tag{5.23}$$

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 $R_n^\alpha(e_d) = 1$, where $e_d := (0, 0, \dots, 1) \in \mathbb{R}^d$. For an account of such polynomials see also [15].

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

$$R_{\mathbf{n}}^\alpha(\mathbf{x})R_n^\alpha(\mathbf{y}) = \int_{\Delta_{(d-1)}} R_{\mathbf{n}}^\alpha(\mathbf{z}) m_{\mathbf{x},\mathbf{y};\alpha}^*(\mathrm{d}\mathbf{z}). \tag{5.24}$$

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

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

Proof. From Proposition 5.4,

$$Q_n^\alpha(\mathbf{x},\mathbf{y}) = \zeta_n^{\alpha_d,|\alpha|-\alpha_d}\mathbb{E}\big(R_n^{\alpha_d,|\alpha|-\alpha_d}(Z_d)\mu_{\mathbf{x},\mathbf{y};\alpha}(Z_d)\big).$$

Now, by uniqueness,

$$\begin{aligned} Q_n^\alpha(\mathbf{x},\mathbf{y}) &= \sum_{|\mathbf{m}|=n} Q_{\mathbf{m}}^\alpha(\mathbf{x})Q_{\mathbf{m}}^\alpha(\mathbf{y}) \\ &= \sum_{|\mathbf{m}|=n} \zeta_{\mathbf{m}}^\alpha R_{\mathbf{m}}^\alpha(\mathbf{x})R_{\mathbf{m}}^\alpha(\mathbf{y}), \end{aligned} \tag{5.25}$$

where $\zeta_{\mathbf{n}}^\alpha := \mathbb{E}(R_{\mathbf{n}}^\alpha)^{-2}$.

So, by Theorem 5.6 and because $R_{\mathbf{n}}(e_d) = 1$,

$$\begin{aligned} Q_n^\alpha(\mathbf{x},\mathbf{y}) &= \int\bigg(\sum_{|\mathbf{m}|=n} \zeta_{\mathbf{m}}^\alpha R_{\mathbf{m}}^\alpha(\mathbf{z})\bigg)\,\mathrm{d}m^*_{\mathbf{x},\mathbf{y};\alpha}(\mathbf{z}) \\ &= \int Q_n^\alpha(\mathbf{z},e_d)\,\mathrm{d}m^*_{\mathbf{x},\mathbf{y};\alpha}(\mathbf{z}), \end{aligned} \tag{5.26}$$

where $Q_{\mathbf{n}}^\alpha$ are orthonormal polynomials. But we know that

$$Q_n^\alpha(\mathbf{z},e_d) = \zeta_n^{\alpha_d,|\alpha|-\alpha_d} R_n^{\alpha_d,|\alpha|-\alpha_d}(z_d),$$

so

$$\begin{aligned} Q_n^\alpha(\mathbf{x},\mathbf{y}) &= \zeta_n^{\alpha_d,|\alpha|-\alpha_d}\mathbb{E}\big(R_n^{\alpha_d,|\alpha|-\alpha_d}(Z_d)\mu_{x,y;\alpha}(Z_d)\big) \\ &= \zeta_n^{\alpha_d,|\alpha|-\alpha_d}\mathbb{E}\big(R_n^{\alpha_d,|\alpha|-\alpha_d}(Z_d)\mu^*_{x,y;\alpha}(Z_d)\big). \end{aligned} \tag{5.27}$$

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

$$\sum_{n=0}^\infty Q_n^\alpha(\mathbf{x},\mathbf{y})R_n^{\alpha_d,|\alpha|-\alpha_d}(z),$$

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 (3.1) for Hahn polynomial in the following version:

$$\widetilde{h}^{\alpha}_{\mathbf{n}}(\mathbf{r}; N) := \int R^{\alpha}_{\mathbf{n}}(\mathbf{x}) D_{\alpha+\mathbf{r}}(\mathrm{d}\mathbf{x}) = \frac{h^{0}_{\mathbf{n}}(\mathbf{r}; N)}{\sqrt{\zeta^{\alpha}_{\mathbf{n}}}}, \qquad \mathbf{r} \in N\Delta_{(d-1)}, \tag{6.1}$$

with the new coefficient of orthogonality

$$\frac{1}{\omega^{\alpha}_{\mathbf{n},N}} := \mathbb{E}[\widetilde{h}^{\alpha}_{\mathbf{n}}(\mathbf{R}; N)]^2 = \frac{N_{[n]}}{(|\alpha| + r)_{(n)}} \frac{1}{\zeta^{\alpha}_{\mathbf{n}}}. \tag{6.2}$$

Formula (6.1) is equivalent to

$$R^{\alpha}_{\mathbf{n}}(\mathbf{x}) = \frac{(|\alpha| + N)_{(n)}}{N_{[n]}} \sum_{|\mathbf{m}|=N} \widetilde{h}^{\alpha}_{n}(\mathbf{m}; N) \binom{N}{\mathbf{m}} \mathbf{x}^{\mathbf{m}}, \qquad \alpha \in \mathbb{R}^d, \mathbf{x} \in \Delta_{(d-1)} \tag{6.3}$$

(see [15], Section 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*

$$H^{\alpha}_{n}(\mathbf{r}, \mathbf{s}) = \omega^{\alpha_d, |\alpha|-\alpha_d}_{n,N} \frac{(|\alpha| + N)_{(n)}}{N_{[n]}} \mathbb{E}_{\mathbf{r},\mathbf{s}}\big[\widetilde{h}^{\alpha_d, |\alpha|-\alpha_d}_{n}(K; N)\big], \\ n \leq |\mathbf{r}| = |\mathbf{s}| = N, \alpha \in \mathbb{R}^d, \tag{6.4}$$

where the expectation is taken with respect to the measure

$$u_{\mathbf{r},\mathbf{s};\alpha}(k) := \int_{\Delta_{(d-1)}} \int_{\Delta_{(d-1)}} \mathbb{E}\left[\binom{r}{k} Z^k_d (1 - Z_d)^{r-k} \middle| \mathbf{x}, \mathbf{y}\right] D_{\alpha+\mathbf{r}}(\mathrm{d}\mathbf{x}) D_{\alpha+\mathbf{s}}(\mathrm{d}\mathbf{y}), \tag{6.5}$$

where Z_d, for every $\mathbf{x}, \mathbf{y}$, is the random variable defined recursively as in Proposition 5.4.

Proof. From (3.3),

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

Then (5.11) implies

$$H_n^\alpha(\mathbf{r},\mathbf{s}) = \frac{\zeta_n^{\alpha_d,|\alpha|-\alpha_d}(|\alpha|+N)_{(n)}}{N_{[n]}} \times \int_{\Delta_{(d-1)}}\int_{\Delta_{(d-1)}}\int_0^1 R_n^{\alpha_d,|\alpha|-\alpha_d}(z_d)m_{\mathbf{x},\mathbf{y};\alpha}(\mathrm{d}z_d)D_{\alpha+\mathbf{r}}(\mathrm{d}\mathbf{x})D_{\alpha+\mathbf{s}}(\mathrm{d}\mathbf{y}),$$

so, by (6.3),

$$H_n^\alpha(\mathbf{r},\mathbf{s}) = \zeta_n^{\alpha_d,|\alpha|-\alpha_d}\left(\frac{(|\alpha|+N)_{(n)}}{N_{[n]}}\right)^2 \times \sum_{k\le N}\widetilde{h}_n^{\alpha_d,|\alpha|-\alpha_d}(k;N)\int_{\Delta_{(d-1)}}\int_{\Delta_{(d-1)}}\int_0^1\binom{N}{k}z_d^k(1-z_d)^{N-k} \times m_{\mathbf{x},\mathbf{y};\alpha}(\mathrm{d}z_d)D_{\alpha+\mathbf{r}}(\mathrm{d}\mathbf{x})D_{\alpha+\mathbf{s}}(\mathrm{d}\mathbf{y}),$$

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 $\{\rho_{\mathbf{n}}\colon \mathbf{n}\in\mathbb{Z}_+^d\}$, which depend on $\mathbf{n}$ only via $|\mathbf{n}|$.

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 $\mathbf{x},\mathbf{y}\in\Delta_{(d-1)}$,*

$$p(\mathbf{x},\mathbf{y}) = \sum_{n=0}^\infty \rho_n Q_n^\alpha(\mathbf{x},\mathbf{y}) \ge 0. \tag{7.1}$$

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

$$p^H(\mathbf{r},\mathbf{s}) = \sum_{n=0}^\infty \rho_n H_n(\mathbf{r},\mathbf{s}) \ge 0. \tag{7.2}$$

7.1. Jacobi positivity from the integral representation

A consequence of the product formulae (5.7) and (5.9) 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 [3], Gasper [8]). *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*

$$\rho_n = \int R_n^{\alpha,\beta}(z)\nu_{\alpha,\beta}(z) \tag{7.3}$$

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

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

with

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

then

$$\nu(A) = \int_A u(x) D_{\alpha,\beta}(\mathrm{d}x) \tag{7.4}$$

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

We refer to [3,8] for the technicalities of the proof. To emphasize the key role played by (5.7), just observe that the positivity of ν and (7.3) entails the representation

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

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, that two different systems of Jacobi polynomials are connected by an integral formula if their parameters share the same total sum.

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

$$\int_0^1 R_n^{\alpha,\beta}\big(1 - (1 - x)z\big) D_{\beta,\mu}(\mathrm{d}z) = R_n^{\alpha-\mu,\beta+\mu}(x) \tag{7.5}$$

and

$$\int_0^1 R_n^{\alpha,\beta}(xz) D_{\alpha,\mu}(\mathrm{d}z) = \frac{\zeta_n^{\alpha+\mu,\beta-\mu}}{\zeta_n^{\alpha,\beta}} R_n^{a+\mu,b-\mu}(x). \tag{7.6}$$

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., [2] ff. formulae 7.392.3 and formulae 7.392.4).

Let us start with (7.6). The moments of a Beta(α,β) distribution on $[0,1]$ are, for every integer $m\le n=0,1,\dots$

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

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

$$\begin{aligned}\int_0^1 \zeta_n^{\alpha,\beta}R_n^{\alpha,\beta}(xz)D_{\alpha,\mu}(\mathrm{d}z)&=\int_0^1 Q_n^{\alpha,\beta}(xz,1)D_{\alpha,\mu}(\mathrm{d}z)\\&=\sum_{m\le n}a_{nm}\frac{(\alpha+\beta)_{(m)}}{(\alpha)_{(m)}}\int_0^1(xz)^mD_{\alpha,\mu}(\mathrm{d}z)\\&=\sum_{m\le n}a_{nm}\frac{(\alpha+\beta)_{(m)}}{(\alpha)_{(m)}}\frac{(\alpha)_{(m)}}{(\alpha+\mu)_{(m)}}x^m\\&=\zeta_n^{\alpha+\mu,\beta-\mu}R_n^{a+\mu,b-\mu}(x),\end{aligned}\tag{7.7}$$

and this proves (7.6).

To prove (7.5), simply remember (see, e.g., [15], Section 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 (7.6) to see that

$$\begin{aligned}\int_0^1\frac{R_n^{\beta,\alpha}((1-x)z)}{R_n^{\beta,\alpha}(0)}D_{\beta,\mu}(\mathrm{d}z)&=(-1)^n\frac{\alpha_{(n)}}{\beta_{(n)}}\frac{\zeta_n^{\beta+\mu,\alpha-\mu}}{\zeta_n^{\beta,\alpha}}R_n^{\beta+\mu,\alpha-\mu}(1-x)\\&=\zeta_n^{\alpha-\mu,\beta+\mu}(x),\end{aligned}\tag{7.8}$$

and the proof is complete. □

Lemma 7.3 completes Theorem 7.2:

Corollary 7.4. *Let $\alpha\le\beta$ with $\alpha+\beta\ge 2$. 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 $0\le\mu\le\beta$.*

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 (7.6) 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 (7.5) instead of (7.6). □

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, that is, all sequences $\rho_{\mathbf{n}} = \rho_{|\mathbf{n}|}$.

Proposition 7.5. *Let $\alpha \in R^d$ satisfy the same conditions as in Proposition 5.4. A sequence $\{\rho_n = \rho_n \colon n \in \mathbb{N}\}$ 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

$$Q_n^{\alpha,\beta}(x,y) = Q_n^{\beta,\alpha}(1-x,1-y), \tag{7.9}$$

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 $\alpha = (\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_d)$ satisfy the conditions of Proposition 5.4 (again, the decreasing order is assumed for simplicity) and let

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

If $\alpha_d > c_{|\alpha|}$ then Corollary 7.4, 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

$$0 \leq \int \left[\sum_{n=0}^{\infty} \rho_n Q_n^{\alpha_d,|\alpha|-\alpha_d}(z_d,1)\right] m_{\mathbf{x},\mathbf{y};\alpha}(\mathrm{d}z_d) = \sum_{n=0}^{\infty} \rho_n Q_n^{\alpha}(\mathbf{x},\mathbf{y}), \qquad x,y \in \Delta_{(d-1)}. \tag{7.10}$$

If $\alpha_d < c_{|\alpha|}$, then apply Corollary 7.4 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 (7.10), thus $\{\rho_n\}$ is positive-definite for D_α.

Necessity. For $I \subseteq \{1,\dots,d\}$, 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

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

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

$$\sum_{n=0}^\infty \rho_n Q_n^\alpha(\mathbf{x},\mathbf{y}) \geq 0$$

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 7.4 with $\mu = \pm(c|\alpha| - \alpha_I)$ (whichever is positive) to obtain, with the possible help of (7.9),

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

□

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. Thus we will identify a convex set of Jacobi positive-definite sequences satisfying the property (P1), as shown in Proposition 8.1. This is done by exploiting the probabilistic interpretation of the orthogonal polynomial kernels. Let us reconsider the function ξ_m^α. The bivariate measure

$$\mathcal{BD}_{\alpha,m}(\mathrm{d}\mathbf{x},\mathrm{d}\mathbf{y}) := \xi_m^\alpha(\mathbf{x},\mathbf{y})D_\alpha(\mathrm{d}\mathbf{x})D_\alpha(\mathrm{d}\mathbf{y}), \tag{8.1}$$

so $\xi_m^\alpha(\mathbf{x},\mathbf{y})$ has the interpretation as a (exchangeable) copula for the joint law of two vectors (X,Y), with identical Dirichlet marginal distribution, arising as the limit distribution of colors in two Pólya urns with m random draws in common. Such a joint law can be simulated via the following Gibbs sampling scheme:

(i) *Generate a vector* $\mathbf{X}$ *of* Dirichlet(α) *random frequencies on d points.*
(ii) *Conditional on the observed* $\mathbf{X} = \mathbf{x}$*, sample m i.i.d. observations with common law* $\mathbf{x}$.

(iii) *Given the vector* $\mathbf{l} \in m\Delta_{(d-1)}$, *which counts how many observations in the sample at step* (ii) *are equal to* $1, \dots, d$, *take* $\mathbf{Y}$ *as conditionally independent of* $\mathbf{X}$ *and with distribution* $D_{\alpha+\mathbf{l}}(\mathrm{d}\mathbf{y})$.

The bivariate measure $\mathcal{BD}_{\alpha,m}$ and its infinite-dimensional extension has found several applications in Bayesian statistics (e.g., by [23]), but no connections were made with orthogonal kernel and canonical correlation sequences. A recent important development of this direction is in [4].

Now, let us allow the number m of random draws in common to be a random number M, say, assume that the probability that the two Pólya urns have $M = m$ draws in common is d_m, for any probability distribution $\{d_m\colon m = 0, 1, 2, \dots\}$ on $\mathbb{N}$. Then we obtain a new joint distribution, with identical Dirichlet marginals and copula given by

$$\mathcal{BD}_{\alpha,d}(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y}) = \mathbb{E}[\mathcal{BD}_{\alpha,M}(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y})] = \sum_{m=0}^{\infty} d_m \xi_m^{\alpha}(\mathbf{x}, \mathbf{y}) D_\alpha(\mathrm{d}\mathbf{x}) D_\alpha(\mathrm{d}\mathbf{y}). \tag{8.2}$$

The probabilistic construction has just led us to prove the following:

Proposition 8.1. *Let* $\{d_m\colon m = 0, 1, \dots\}$ *be a probability measure on* $\{0, 1, 2, \dots\}$. *For every* $|\theta| \geq 0$, *the sequence*

$$\rho_n = \sum_{m \geq n} \frac{m_{[n]}}{(|\theta| + m)_{(n)}} d_m, \qquad n = 0, 1, 2, \dots, \tag{8.3}$$

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 (2.7) for the (positive) function ξ_m^α: we can rewrite (8.2) as

$$\begin{aligned} 0 &\leq \sum_{m=0}^{\infty} d_m \xi_m^{\alpha}(\mathbf{x}, \mathbf{y}) \\ &= \sum_{m=0}^{\infty} d_m \sum_{n \leq m} \frac{m_{[n]}}{(|\theta| + m)_{(n)}} Q_n^{\alpha}(\mathbf{x}, \mathbf{y}) \\ &= \sum_{n=0}^{\infty} \left[\sum_{m \geq n} \frac{m_{[n]}}{(|\theta| + m)_{(n)}} d_m \right] Q_m^{\alpha}(\mathbf{x}, \mathbf{y}) \\ &= \sum_{n=0}^{\infty} \rho_n Q_m^{\alpha}(\mathbf{x}, \mathbf{y}), \end{aligned} \tag{8.4}$$

and since (8.4) does not depend on the dimension of α, then the proposition is proved for D_α. □

Example 8.2. Take $d_m = \delta_{ml}$, the probability assigning full mass to l. The corresponding positive-definite sequence is

$$\rho_n(m) = \sum_{m\ge n} \frac{m_{[n]}}{(|\theta|+m)_{(n)}}\delta_{ml} = \frac{l_{[n]}}{(|\theta|+l)_{(n)}}\mathbb{I}(l \ge n). \tag{8.5}$$

and by Proposition 2.1,

$$\sum_{n=0}^{\infty} \rho_n(m) Q_n^\alpha(\mathbf{x},\mathbf{y}) = \sum_{n=0}^{l} \frac{l_{[n]}}{(|\theta|+l)_{(n)}} Q_n^\alpha(\mathbf{x},\mathbf{y}) = \xi_l^\alpha(\mathbf{x},\mathbf{y}) \ge 0. \tag{8.6}$$

Thus $\rho_n(m)$ forms the sequences of canonical correlations induced by the bivariate probability distribution $\xi_l^\alpha(\mathbf{x},\mathbf{y})D_\alpha(\mathbf{x})D_\alpha(\mathbf{y})$.

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

$$d_m(t) = \sum_{n\ge n} a_{mn}^{|\alpha|} \mathrm{e}^{-(1/2)n(n+|\alpha|-1)t}, \qquad m = 0, 1, 2 \ldots, \tag{8.7}$$

where $(a_{mn}^{|\alpha|})$ is the invertible triangular system (2.6) defining the polynomial kernels Q_n^α in Proposition 2.1. Since the coefficients of the inverse system are exactly of the form

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

then

$$\rho_n(t) = \mathrm{e}^{-(1/2)n(n+|\alpha|-1)t}$$

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 (8.7) is the so-called *coalescent lineage distribution* (see [12,13]), that is, 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 $A_{|\alpha|,d}$. More details on the connection between coalescent lineage distributions and Jacobi polynomials can be found in [14].

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

$$\mathbb{E}_0\big(Q_{\mathbf{n}}(\mathbf{Y})|\mathbf{X}=\mathbf{x}\big) = Q_{\mathbf{n}}(\mathbf{x})$$

so that, for every square-integrable function

$$f = \sum_{\mathbf{n}} c_{\mathbf{n}} Q_{\mathbf{n}},$$

we have

$$\mathbb{E}_0\big(f(\mathbf{Y})|\mathbf{X}=\mathbf{x}\big) = \sum_{\mathbf{n}} c_{\mathbf{n}} Q_{\mathbf{n}}(\mathbf{x}) = f(\mathbf{x});$$

that is, $\mathcal{BD}_{\alpha,\{0\}}$ is, in fact, the Dirac measure $\delta(\mathbf{y}-\mathbf{x})$.

In the latter case, $d_m(\infty) = \delta_{m0}$ so that $\rho_n(\infty) = 0$ for every $n > 1$ and $\mathbb{E}_0(Q_n(\mathbf{Y})|\mathbf{X}=\mathbf{x}) = \mathbb{E}[Q_n(\mathbf{Y})]$, implying that

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

that is, $\mathbf{X}, \mathbf{Y}$ 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 ξ_m^α with $\xi_m^{\downarrow|\theta|,\infty}$. We only need to observe that because the functions

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

forming the terms in $\xi_m^{\downarrow|\theta|,\infty}$ (see (4.2)), are probability measures on $m\Delta_\infty^\downarrow$, then the kernel

$$\xi_m^{\downarrow|\theta|,\infty}(\mathbf{x},\mathbf{y}) D_{|\theta|,\infty}^{\downarrow}(d\mathbf{y})$$

defines, for every $\mathbf{x}$, a proper transition probability function on $\Delta_\infty^\downarrow$, allowing for the Gibbs sampling interpretation as in Section 8, but are modified as follows:

(i) *Generate a point* $\mathbf{X}$ *in* $\Delta_\infty^\downarrow$ *with distribution* $\mathrm{PD}(|\theta|)$.

(ii) *Conditional on the observed* $\mathbf{X}=\mathbf{x}$*, sample a partition of* m *with distribution function* $\binom{m}{\mathbf{l}}\sharp(\mathbf{l})[\mathbf{x},\mathbf{l}]$.

(iii) *Conditionally on the vector* $\mathbf{l}$*, counting the cardinalities of the blocks in the partition obtained at step* (ii)*, take* Y *as stochastically independent of* X *and with distribution*

$$\frac{\binom{m}{\mathbf{l}}\sharp(\mathbf{l})[\mathbf{x},\mathbf{l}]\,\mathrm{PD}_\theta(d\mathbf{y})}{\mathrm{ESF}_{|\theta|}(\mathbf{l})}.$$

Thus the proof of the following statement is now obvious.

Proposition 8.5. *Let* $\{d_m\colon m=0,1,\ldots\}$ *be a probability measure on* $\{0,1,2,\ldots\}$*. For every* $|\theta|\ge 0$*, the sequence*

$$\rho_n = \sum_{m\ge n} \frac{m_{[n]}}{(|\theta|+m)_{(n)}} d_m, \qquad n = 0,1,2,\ldots, \tag{8.8}$$

is a positive-definite sequence for the Poisson–Dirichlet point process with parameter $|\theta|$.

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, JPDSs $\{\rho_n\}$ can be mapped to probability distributions $\{d_m\}$ on $\mathbb{Z}_+$, for every $m = 0, 1, \dots$, via the inversion

$$d_m(\rho) = \sum_{n=m}^{\infty} a_{nm}^{|\alpha|} \rho_n. \tag{9.1}$$

For this to happen we only need $d_m(\rho)$ to be non-negative for every m as it is easy to check that $\sum_m d_m(\rho) = 1$ always. In this section we give some sufficient conditions on ρ for $d_m(\rho)$ to be non-negative for every $m = 0, 1, \dots$, and an important counterexample showing that not all JPDSs 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

$$p_\rho(x, y) := \sum_{n=0}^{\infty} \rho_n Q_n^{\alpha,\beta}(x, y) \geq 0 \tag{9.2}$$

and, in particular,

$$p_\rho(x) := p_\rho(x, 1) \geq 0. \tag{9.3}$$

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

$$\begin{aligned} d_m(\rho) &= \sum_{v=0}^{\infty} a_{v+m,m}^{|\theta|} \rho_{v+m} \\ &= \frac{(|\theta| + m)_{(m)}}{m!} \sum_{v=0}^{\infty} a_{v0}^{|\theta|+2m} \rho_{v+m}, \qquad m = 0, 1, \dots. \end{aligned} \tag{9.4}$$

This follows from the general identity

$$a_{v+j,u+j}^{|\theta|} = a_{v,u}^{|\theta|+2j} \frac{u!}{(u + j)!} (|\theta| + u + j)_{(j)}. \tag{9.5}$$

Now consider the expansion of Jacobi polynomials. We know that

$$\begin{aligned} \zeta_n^{\alpha,\beta} R_n^{\alpha,\beta}(x) R_n^{\alpha,\beta}(y) &= Q_n^{\alpha,\beta}(x, y) \\ &= \sum_{m=0}^{n} a_{nm}^{|\theta|} \xi_m^{\alpha,\beta}(x, y). \end{aligned} \tag{9.6}$$

Since $R_n^{\alpha,\beta}(1) = 1$ and $\xi_m^{\alpha,\beta}(0,1) = \delta_{m0}$, then

$$\zeta_n^{\alpha,\beta} R_n^{\alpha,\beta}(0) = Q_n^{\alpha,\beta}(0,1) = a_{n0}^{|\theta|}. \tag{9.7}$$

Therefore (9.4) becomes

$$d_m(\rho) = \frac{(|\theta|+m)_{(m)}}{m!} \sum_{v=0}^{\infty} \zeta_v^{\alpha+m,\beta+m} R_v^{\alpha+m,\beta+m}(0)\rho_{v+m}, \qquad m = 0, 1, \ldots. \tag{9.8}$$

Now apply, for example, [16], (4.3.2), to deduce

$$\begin{aligned} &\frac{\mathrm{d}^m}{\mathrm{d}y^m}\big[D_{\alpha+m,\beta+m}(y) R_v^{\alpha+m,\beta+m}(y)\big] \\ &\quad = (-1)^m \frac{\theta_{(2m)}}{\alpha_{(m)}} R_{v+m}^{\alpha,\beta}(y) D_{\alpha,\beta}(y). \end{aligned} \tag{9.9}$$

For $m = 1$,

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

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

$$p_\rho^{(m)}(x) := \frac{\mathrm{d}^m}{\mathrm{d}x^m} p_\rho(x), \qquad m = 0, 1, \ldots.$$

It is easy to prove that

$$\rho_{v+m} = \frac{m!\alpha_{(m)}}{|\theta|_{(2m)}} \int_0^1 p_\rho^{(m)}(x) R_v^{\alpha+m,\beta+m}(x) D_{\alpha+m,\beta+m}(x)\,\mathrm{d}x, \tag{9.10}$$

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 $\{d_m(\rho)\colon m \in \mathbb{Z}_+\}$ is a probability mass function on $\mathbb{Z}_+$. Then its probability generating function (*p.g.f.*) must have all derivatives non-negative. For every $0 < \gamma < |\theta|$, the p.g.f. has the representation

$$\begin{aligned}\varphi(s) &= \sum_{m=0}^{\infty} d_m(\rho) s^m \\ &= \mathbb{E}_{\gamma,|\theta|-\gamma}\left[\sum_{m=0}^{\infty} d_m(\rho)\xi_m^{\gamma,|\theta|-\gamma}(sZ,1)\right] \\ &= \mathbb{E}_{\gamma,|\theta|-\gamma}\left[\sum_{m=0}^{\infty} \rho_n \zeta_n^{\gamma,|\theta|-\gamma} R_n^{\gamma,|\theta|-\gamma}(sZ)\right] \\ &= \mathbb{E}_{\gamma,|\theta|-\gamma}[p_\rho(sZ)],\end{aligned} \tag{9.11}$$

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

$$\frac{|\theta|_{(m)}}{\alpha_{(m)}} x^m = \xi_m^{\alpha,\beta}(x,1), \qquad \alpha,\beta > 0, \tag{9.12}$$

and the third equality comes from (9.6).

So, for every $k = 0, 1, \ldots,$

$$0 \le \frac{\mathrm{d}^k}{\mathrm{d}s^k}\varphi(s) = \mathbb{E}_{\gamma,|\theta|-\gamma}\left[Z^k p_\rho^{(k)}(sZ)\right] \tag{9.13}$$

for every $\gamma \in (0, |\theta|)$. Now, if we take the limit as $\gamma \to |\theta|$, $Z \to^d 1$ 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 lend 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*

$$d_m = \sum_{n\ge|m|} a_{nm}^{|\theta|} R_n^{\alpha,\beta}(x), \qquad m = 0, 1, 2, \ldots,$$

is not a probability measure.

Proof. Rewrite

$$\begin{aligned}\phi_x(s) &= \sum_{n=0}^{\infty} R_n^{\alpha,\beta}(x) \sum_{m=0}^{n} a_{nm}^{|\theta|} s^m \\ &= \mathbb{E} \sum_{n=0}^{\infty} \zeta_n^{\alpha,\beta} R_n^{\alpha,\beta}(x) R_n^{\alpha,\beta}(Ws),\end{aligned} \tag{9.14}$$

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

$$\frac{\mathrm{d}D_{\alpha,\beta}(y)}{\mathrm{d}y} \sum_{n=0}^{\infty} \zeta_n^{\alpha,\beta} R_n^{\alpha,\beta}(x) R_n^{\alpha,\beta}(y) = \delta_x(y),$$

that is, 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

$$\begin{aligned}q(s) &= \int_0^1 g(x)\phi_x(s) \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} \,\mathrm{d}x \\ &= \int_0^1 g(ws) \frac{w^{\alpha-1}(1-w)^{\beta-1}}{B(\alpha, \beta)} \,\mathrm{d}w\end{aligned} \tag{9.15}$$

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

10. 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, that is, 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 vice versa, but we could prove (P2) only for particular subclasses of positive-definite sequences. In Proposition 10.4 we prove that every JPDS is a limit of (P2) sequences. Later, in Proposition 10.8, 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 10.1. *For every d and $\alpha \in \mathbb{R}_+^d$, let $\rho = \{\rho_n\}$ be a α-JPDS. Then*

$$\rho_n \frac{N_{[n]}}{(|\alpha| + N)_{(n)}}, \qquad n = 0, 1, 2, \ldots, \tag{10.1}$$

is a positive-definite sequence for $\mathrm{DM}_{\alpha,N}$, *for every* $N = 1, 2, \ldots$.

Proof. From Proposition 3.1, if

$$\sum_{n=0}^{\infty} \rho_n Q_n^{\alpha}(\mathbf{x}, \mathbf{y}) \geq 0,$$

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

$$\sum_{n=0}^{\infty} \rho_n \int\int Q_n^{\alpha}(\mathbf{x}, \mathbf{y}) D_{\alpha+\mathbf{r}}(\mathrm{d}\mathbf{x}) D_{\alpha+\mathbf{s}}(\mathrm{d}\mathbf{y}) = \sum_{n=0}^{\infty} \rho_n \frac{N_{[n]}}{(|\alpha| + N)_{(n)}} H_n^{\alpha}(\mathbf{r}, \mathbf{s}) \geq 0.$$ □

Example 10.2. Consider the JPDS given in Example 8.3 from population genetics; $\rho_n(t) = \mathrm{e}^{-(1/2)tn(n+|\alpha|-1)}$, $t \geq 0$. The HPDS

$$\rho_n(t|N) = \frac{N_{[n]}}{(|\alpha| + N)_{(n)}} \mathrm{e}^{-(1/2)tn(n+|\alpha|-1)} \tag{10.2}$$

describes the survival function of the number of non-mutant surviving lineages at time t in the past, in a coalescent process with neutral mutation, starting with N surviving lineages at time 0 (see [13] for more details and references).

Two important HPDSs are given in the following lemma.

Lemma 10.3. *For every* d, *every* $m \leq N$ *and every* $\alpha \in \mathbb{R}_+^d$, *both sequences*

$$\left\{ \frac{m_{[n]}}{(|\alpha| + m)_{(n)}} \frac{(|\alpha| + N)_{(n)}}{N_{[n]}} \right\}_{n \in \mathbb{Z}_+} \tag{10.3}$$

and

$$\left\{ \frac{m_{[n]}}{(|\alpha| + m)_{(n)}} \right\}_{n \in \mathbb{Z}_+} \tag{10.4}$$

are α*-HPDSs for* $\mathrm{DM}_{\alpha,N}$.

Proof. From Proposition 3.5, by inverting (3.19) we know that, for $m = 0, \ldots, N$

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

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 10.1, 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 $(\widetilde{\rho}_n)$ 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 10.4. *For every d and $\alpha \in \mathbb{R}_+^d$, let $\rho = \{\rho_n\}$ be a α-JPDS. Then there exists a sequence $\{\rho_n^N \colon n \in \mathbb{Z}_+\}_{N=0}^\infty$, 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 $\mathrm{DM}_{(\alpha,\beta),N}$ and $\mathrm{DM}_{(\alpha,\beta),M}$ 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

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

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

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

Now take

$$p_\rho(x, y) = \sum_n \rho_n Q_n^{\alpha,\beta}(x, y) \geq 0$$

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

$$\begin{aligned}\rho_n &= \mathbb{E}[Q_n^{\alpha,\beta}(X,Y)p_\rho(X,Y)]\\ &= \mathbb{E}\Big[Q_n^{\alpha,\beta}(X,Y)\lim_{N\to\infty} B_{N,N}p_\rho(X,Y)\Big]\\ &= \lim_{N\to\infty}\mathbb{E}[Q_n^{\alpha,\beta}(X,Y)B_{N,N}p_\rho(X,Y)]\\ &= \lim_{N\to\infty}\rho_n^N,\end{aligned}$$

where

$$\rho_n^N := \mathbb{E}[Q_n^{\alpha,\beta}(X,Y)B_{N,N}p_\rho(X,Y)].$$

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

Now rewrite

$$\begin{aligned}\rho_n^N &= \int_0^1\int_0^1\sum_{i=1}^N\sum_{j=1}^N p_\rho\left(\frac{i}{N},\frac{j}{N}\right)Q_n^{\alpha,\beta}(x,y)\binom{N}{i}x^i(1-x)^{N-i}\\ &\qquad\times\binom{N}{j}y^j(1-y)^{N-j}D_\alpha(\mathrm{d}x)D_\alpha(\mathrm{d}y)\\ &= \sum_{i=1}^N\sum_{j=1}^N \mathrm{DM}_{\alpha,N}(i)\,\mathrm{DM}_{\alpha,N}(j)\,p_\rho\left(\frac{i}{N},\frac{j}{N}\right)\mathbb{E}[Q_n^{\alpha,\beta}(X,Y)|i,j]\\ &= \frac{N_{[n]}}{(\alpha+\beta+N)_{(n)}}\mathbb{E}\left[p_\rho\left(\frac{I}{N},\frac{J}{N}\right)H_n^{\alpha,\beta}(I,J)\right]\end{aligned} \tag{10.5}$$

for I, J are independent $\mathrm{DM}_{(\alpha,\beta),N}$ random variables. The last equality follows from (3.3). Since p_ρ is positive, from (10.5), it follows that

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

is, for every N, α-HPDS. But by Lemma 10.3, we can multiply every term of the sequence by the HPDS (10.4), 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 10.1 under the usual (extended) Gasper constraints on α.

Proposition 10.5. *If α satisfies the conditions of Proposition 5.4, let $\{\rho_n\}$ be α-HPDS for some integer N. Then both $\{\rho_n\}$ and (10.1) are positive definite for D_α.*

Proof. If

$$\sum_{n=0}^{\infty} \rho_n H_n^{\alpha}(\mathbf{r}, \mathbf{s}) \geq 0$$

for every $\mathbf{r}, \mathbf{s} \in N\Delta_{(d-1)}$, 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

$$\widetilde{h}_n^{\alpha_1, |\alpha|-\alpha_1}(r; N) = \int_0^1 \frac{Q_n^{\alpha_1, |\alpha|-\alpha_1}(x, 1)}{Q_n^{\alpha_1, |\alpha|-\alpha_1}(1, 1)} D_{\alpha+r}(\mathrm{d}x).$$

Then it is easy to prove that

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

and

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

(see also [15], (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,

$$\begin{aligned} \rho_n &= \mathbb{E}\big[f_N(R)\widetilde{h}_n^{\alpha_1, |\alpha|-\alpha_1}(R; N)\big] \\ &= \int_0^1 \phi_N(x) R_n(x) D_{\alpha}(\mathrm{d}x), \end{aligned} \tag{10.6}$$

where

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

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

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

is α-JPDS; thus (10.1) is JPDS. □

Remark 10.6. 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 (2.7) of ξ_m^α (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,

Corollary 10.7. *If $\{\rho_n\}$ is positive-definite for D_α, then (10.1) is positive-definite for both D_α and* DM_α.

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

We have seen in Proposition 10.5 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 (8.3), of a particular class of discrete probability measures.

Proposition 10.8. *For every N and $|\theta|>0$, let $\rho^{(N)}=\{\rho_n^{(N)}\colon n\in\mathbb{Z}_+\}$ be of the same form (8.3) for a probability mass function $d^{(N)}=\{d_m\colon m\in\mathbb{Z}_+\}$, 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 10.3, 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 $z\to 1$, resp.).

Now replace m with a random M with distribution given by $d^{(N)}$. Then

$$\begin{aligned} 0 &\le \mathbb{E}\left[\sum_{n=0}^{m}\frac{M_{[n]}}{(|\alpha|+M)_{(n)}}H_n^\alpha\right] \\ &= \sum_{n=0}^{N}\left(\sum_{m=n}^{N} d_m^{(N)}\frac{M_{[n]}}{(|\alpha|+M)_{(n)}}\right)H_n^\alpha, \end{aligned}$$

which proves the "Hahn" part of the claim. The "Jacobi" part is obviously proved by Proposition 8.3. □

Acknowledgements

Dario Spanò's research is partly supported by CRiSM, an EPSRC/HEFCE-funded grant. 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 [4,18]). Here we wish to thank them for their helpful comments.

References

[1] Abramowitz, M. and Stegun, I.A. (1964). *Handbook of Mathematical Functions with Formulas*, *Graphs*, *and Mathematical Tables*. *National Bureau of Standards Applied Mathematics Series* **55**. New York: Dover. MR0167642

[2] Andrews, G.E., Askey, R. and Roy, R. (1999). *Special Functions*. *Encyclopedia of Mathematics and Its Applications* **71**. Cambridge: Cambridge Univ. Press. MR1688958

[3] Bochner, S. (1954). Positive zonal functions on spheres. *Proc*. *Nat*. *Acad*. *Sci*. *U.S.A*. **40** 1141–1147. MR0068127

[4] Diaconis, P., Khare, K. and Saloff-Coste, L. (2008). Gibbs sampling, exponential families and orthogonal polynomials (with comments and a rejoinder by the authors). *Statist*. *Sci*. **23** 151–178. MR2446500

[5] Dunkl, C.F. and Xu, Y. (2001). *Orthogonal Polynomials of Several Variables*. *Encyclopedia of Mathematics and Its Applications* **81**. Cambridge: Cambridge Univ. Press. MR1827871

[6] Eagleson, G.K. (1964). Polynomial expansions of bivariate distributions. *Ann*. *Math*. *Statist*. **35** 1208–1215. MR0168055

[7] Eagleson, G.K. and Lancaster, H.O. (1967). The regression system of sums with random elements in common. *Austral*. *J*. *Statist*. **9** 119–125. MR0226761

[8] Gasper, G. (1972). Banach algebras for Jacobi series and positivity of a kernel. *Ann*. *of Math*. (2) **95** 261–280. MR0310536

[9] Gasper, G. (1973). Nonnegativity of a discrete Poisson kernel for the Hahn polynomials. *J*. *Math*. *Anal*. *Appl*. **42** 438–451. Collection of articles dedicated to Salomon Bochner. MR0320392

[10] Griffiths, R.C. (1979). On the distribution of allele frequencies in a diffusion model. *Theoret*. *Population Biol*. **15** 140–158. MR0528914

[11] Griffiths, R.C. (1979). A transition density expansion for a multi-allele diffusion model. *Adv. in Appl*. *Probab*. **11** 310–325. MR0526415

[12] Griffiths, R.C. (1980). Lines of descent in the diffusion approximation of neutral Wright–Fisher models. *Theoret*. *Population Biol*. **17** 37–50. MR0568666

[13] Griffiths, R.C. (2006). Coalescent lineage distributions. *Adv*. *in Appl*. *Probab*. **38** 405–429. MR2264950

[14] Griffiths, R.C. and Spanó, D. (2010). Diffusion processes and coalescent trees. In *Probability and Mathematical Genetics*. *London Mathematical Society Lecture Note Series* **378** 358–379. Cambridge: Cambridge Univ. Press. MR2744247

[15] Griffiths, R.C. and Spanò, D. (2011). Multivariate Jacobi and Laguerre polynomials, infinite-dimensional extensions, and their probabilistic connections with multivariate Hahn and Meixner polynomials. *Bernoulli* **17** 1095–1125.

[16] Ismail, M.E.H. (2005). *Classical and Quantum Orthogonal Polynomials in One Variable*. *Encyclopedia of Mathematics and Its Applications* **98**. Cambridge: Cambridge Univ. Press. With two chapters by Walter Van Assche, with a foreword by Richard A. Askey. MR2191786

[17] Karlin, S. and McGregor, J. (1975). 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) 261–288. New York: Academic Press. MR0406574

[18] Khare, K. and Zhou, H. (2009). Rates of convergence of some multivariate Markov chains with polynomial eigenfunctions. *Ann. Appl. Probab.* **19** 737–777. MR2521887

[19] Kingman, J.F.C., Taylor, S.J., Hawkes, A.G., Walker, A.M., Cox, D.R., Smith, A.F.M., Hill, B.M., Burville, P.J. and Leonard, T. (1975). Random discrete distribution. *J. Roy. Statist. Soc. Ser. B* **37** 1–22. 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. MR0368264

[20] Koornwinder, T. (1974). Jacobi polynomials. II. An analytic proof of the product formula. *SIAM J. Math. Anal.* **5** 125–137. MR0385198

[21] Koornwinder, T.H. and Schwartz, A.L. (1997). Product formulas and associated hypergroups for orthogonal polynomials on the simplex and on a parabolic biangle. *Constr. Approx.* **13** 537–567. MR1466065

[22] Lancaster, H.O. (1958). The structure of bivariate distributions. *Ann. Math. Statist.* **29** 719–736. MR0102150

[23] Muliere, P., Secchi, P. and Walker, S. (2005). Partially exchangeable processes indexed by the vertices of a k-tree constructed via reinforcement. *Stochastic Process. Appl.* **115** 661–677. MR2128635

[24] Peccati, G. (2008). Multiple integral representation for functionals of Dirichlet processes. *Bernoulli* **14** 91–124. MR2401655

[25] Rosengren, H. (1999). Multivariable orthogonal polynomials and coupling coefficients for discrete series representations. *SIAM J. Math. Anal.* **30** 232–272 (electronic). MR1664759

[26] Waldron, S. (2006). On the Bernstein–Bézier form of Jacobi polynomials on a simplex. *J. Approx. Theory* **140** 86–99. MR2226679

[27] Watterson, G.A. (1984). Lines of descent and the coalescent. *Theoret. Population Biol.* **26** 77–92. MR0760232

Received January 2011 and revised September 2011