Bernoulli **17**(3), 2011, 1095–1125
DOI: 10.3150/10-BEJ305

Multivariate Jacobi and Laguerre polynomials, infinite-dimensional extensions, and their probabilistic connections with multivariate Hahn and Meixner polynomials

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

[1]*Department of Statistics, 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*

Multivariate versions of classical orthogonal polynomials such as Jacobi, Hahn, Laguerre and Meixner are reviewed and their connection explored by adopting a probabilistic approach. Hahn and Meixner polynomials are interpreted as posterior mixtures of Jacobi and Laguerre polynomials, respectively. By using known properties of gamma point processes and related transformations, a new infinite-dimensional version of Jacobi polynomials is constructed with respect to the size-biased version of the Poisson–Dirichlet weight measure and to the law of the gamma point process from which it is derived.

Keywords: beta-Stacy; Dirichlet distribution; Hahn polynomials; Jacobi polynomials; Laguerre polynomials; Meixner polynomials; multivariate orthogonal polynomials; size-biased random discrete distributions

1. Introduction

In this paper we will review multivariate orthogonal polynomials, complete with respect to weight measures given by the Dirichlet and Dirichlet-multinomial probability distributions (denoted respectively as D_α or DM_α, $\alpha \in \mathbb{R}_+^d$), that is, polynomials $\{G_n : n \in \mathbb{N}^d\}$ satisfying

$$\int G_n G_m \,\mathrm{d}\mu = \frac{1}{c_m}\delta_{nm}, \qquad n, m \in \mathbb{N}^d. \tag{1.1}$$

The polynomials $\{G_n\}$ are known as multivariate Jacobi polynomials if (1.1) is satisfied with $\mu = D_\alpha$, and multivariate Hahn polynomials if $\mu = DM_\alpha$. Here c_m are positive constants. Completeness means that, for every function f with finite variance (under μ), there is an expansion

$$f(x) = \sum_{n \in \mathbb{N}^d} c_n a_n G_n(x), \tag{1.2}$$

where

$$a_n = \mathbb{E}[f(X)G_n(X)].$$

1350-7265

Systems of multivariate orthogonal polynomials are not unique, and a large number of characterizations of d-dimensional Jacobi and Hahn polynomials exist in literature. We will focus on a construction of Jacobi polynomials, based on a method originally proposed by Koornwinder [15] that has a strong probabilistic interpretation. Based on this, we will re-interpret the role of Jacobi polynomials in the construction of multivariate Hahn and several other well-known classes of multivariate orthogonal polynomials. In particular, we will (1) describe multivariate Hahn polynomials as *posterior* mixtures of Jacobi polynomials, in a sense which will become precise in Section 5; (2) construct, in Section 4, a new system of multiple Laguerre polynomials, orthogonal with respect to the product of several gamma probability distributions with identical scale parameters; (3) derive, in Section 6, a new class of multiple Meixner polynomials as posterior mixtures of the Laguerre polynomials mentioned in (2); (4) obtain polynomials in the multivariate hypergeometric distribution by taking the parameters in the Hahn polynomials to be negative; (5) obtain (Section 3.3) asymptotic results as the dimension $d \to \infty$ with $|\alpha| := \sum_{i=1}^{d} \alpha_i \to |\theta| > 0$, by considering size-biased Dirichlet measures.

Furthermore, we will see that an extensive application of Koornwinder's method leads directly to finding new systems of polynomials, orthogonal with respect to a wider family of distributions on the infinite simplex, known in Bayesian nonparametric statistics as the (discrete) beta-Stacy family [23], a popular member of which is the GEM distribution (so named after Griffiths, Engen and McCloskey who introduced it independently) and its two-parameter distribution.

The intricate relationship existing among all the mentioned systems of polynomials is traditionally described in terms of their analytic/algebraic expression as (multivariate) basic hypergeometric series (see, e.g., [5,7]). The main advantage of a probabilistic approach is that it re-expresses most relationships in terms of random variables, which may be more transparent to statisticians and probabilists. With this in mind we will begin the paper with an introductory summary (Section 2) of known facts from the theory of probability distributions. Section 3.1 is devoted to multivariate Jacobi polynomials, whose structure will be the building block for the subsequent sections: Multiple Laguerre in Section 4, Hahn in Section 5 and Meixner in Section 6.

It is worth observing that the posterior mixture representation of multivariate Hahn polynomials shown in Proposition 5.2 is obtained without imposing *a priori* any Bernstein–Bézier form to the Jacobi polynomials, and nevertheless it agrees with recent interpretations of Hahn polynomials as Bernstein coefficients of Jacobi polynomials in such a form [21,22], a result for which a new, more probabilistic proof is offered in Section 5.2.1. In particular, our approach will make more intuitive the link between the Bernstein–Bézier interpretation and the original formulation proposed decades ago by Karlin and McGregor [11]. In terms of applications, understanding such a link will complete Karlin and McGregor's analysis of some well-known d-type models in population genetics (Section 5.2.3). Our extensions of Sections 3.3 and 4.2 open for possible new infinite-dimensional versions of Karlin and McGregor's work.

Along the same lines one can view the Meixner polynomials obtained in Proposition 6.2 as re-scaled Bernstein coefficients of our multiple Laguerre polynomials, as shown in Section 6.1.

The original motivation for this study was to obtain some background material that can be used to characterize bivariate distributions, or transition functions, with fixed Dirichlet or Dirichlet-multinomial marginals, for which the following *canonical expansions* are possible:

$$p(\mathrm{d}x, \mathrm{d}y) = \left\{1 + \sum_{n \in \mathbb{Z}_+^d}^{\infty} c_n \rho_n G_n(x) G_n(y)\right\} D_\alpha(\mathrm{d}x) D_\alpha(\mathrm{d}y), \qquad x, y \in \Delta_{(d-1)},$$

for appropriate, positive-definite sequences $\rho_m : m \in \mathbb{N}^d$, called the *canonical correlation coefficients* of the model. Some results on such a problem are in [8] and [9]. Other possible applications in statistics are related to least-squares approximations and regression. An MCMC (Markov chain Monte Carlo)-Gibbs sampler use of orthogonal polynomials is explored, for example, in [3]; related applications are in [13]. In this paper, however, we will focus merely on the construction of the mentioned systems of polynomials.

2. Distributions on the discrete and continuous simplex

Throughout the paper we will denote by $|x|$ the total sum of all components of $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. We will also adopt the notation:

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

and

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

For example, the Dirichlet distribution $D_\alpha : \alpha \in \mathbb{R}_+^d$ will be written as

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

where $\underline{1} = (1, 1, \dots, 1)$ and, for $d = 2, 3, \dots,$ $\Delta_{(d-1)} = \{x \in \mathbb{R}_+^d : |x| = 1\}$.

2.1. Conditional independence in the Dirichlet distribution

2.1.1. *Gamma sums*

For every $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}_+^d$ and $\beta > 0$, let $Y = (Y_1, \dots, Y_d)$ be a collection of d-independent gamma random variables with parameter, respectively, (α_i, β). The distribution of Y is given by the product measure

$$\gamma_{\alpha,\beta}^d(\mathrm{d}y) = \frac{y^{\alpha - \underline{1}} \mathrm{e}^{-|y|/\beta}}{\Gamma(\alpha) \beta^{|\alpha|}} \mathbb{I}(y \in \mathbb{R}_+^d) \,\mathrm{d}y.$$

Consider the mapping

$$(Y_1, \ldots, Y_d) \longmapsto (|Y|, X_1, \ldots, X_{d-1}),$$

where

$$X_j := \frac{Y_j}{|Y|}, \qquad j = 1, \ldots, d-1,$$

and set $X_d = 1 - \sum_{i=1}^{d-1} X_i$. It is easy to rewrite

$$\gamma^d_{\alpha,\beta}(\mathrm{d}y) = \gamma^1_{|\alpha|,\beta}(\mathrm{d}|y|) D_\alpha(\mathrm{d}x),$$

that is: (i) $|Y| := \sum_{i=1}^d Y_i$ is a gamma$(|\alpha|, \beta)$ random variable, and (ii) X is independent of $|Y|$ and has Dirichlet distribution with parameter α.

2.1.2. *Dirichlet as a right-neutral distribution*

Let $X = (X_1, \ldots, X_d)$ be a random distribution on $\{1, \ldots, d\}$ with Dirichlet distribution $D_\alpha, \alpha \in \mathbb{R}^d_+$. Consider the random cumulative frequencies $S_j := \sum_{i=1}^j X_i,\ \ j = 1, \ldots, d-1$. Then the increments

$$B_j := \frac{X_j}{1 - S_{j-1}}, \qquad j = 1, \ldots, d-1, \tag{2.1}$$

are independent random variables, each with a beta distribution with parameters $(\alpha_j, |\alpha| - \sum_{i=1}^j \alpha_i)$. This property is also known as *right-neutrality* [4]. Notice that such a structure holds, with different parameters, for any reordering of the atoms of X.

2.2. Size-biased Dirichlet frequencies and limit distributions

One remarkable advantage of considering unordered versions of Dirichlet frequencies is that they admit sensible limits as the dimension d grows to infinity, whereas the original Dirichlet distribution is obviously bounded to finite dimensions. Two possible ways of unordering the Dirichlet atoms are equivalent: (1) Rearranging the frequencies in a size-biased random order; (2) Ranking them in order of magnitude. For Dirichlet measures, size-biased frequencies are much more mathematically treatable than the ranked ones.

2.2.1. *Size-biased order and the GEM distribution*

Let x be a point of $\Delta_{(d-1)}$. Then x induces a probability distribution on the group $\mathcal{G}_d$ of all permutations of $\{1, \ldots, d\}$:

$$\sigma_x(\pi) = \prod_{i=1}^{d-1} \frac{x_{\pi_i}}{1 - \sum_{j=1}^{i-1} x_{\pi_j}}, \qquad \pi \in \mathcal{G}_d.$$

Let $\alpha \in \mathbb{R}_+^d$. The *size-biased measure* on $\Delta_{(d-1)}$ induced by a Dirichlet distribution D_α is given by

$$\ddot{D}_\alpha(A) = \int \sigma_x(\pi : \pi x \in A) D_\alpha(\mathrm{d}x).$$

Note that $\widetilde{\sigma}_x\{y\} := \sigma_x(\pi : \pi x = y)$ is non-zero if and only if y is a permutation of x, and that

$$\widetilde{\sigma}_x\{y\} = \widetilde{\sigma}_{\pi x}\{y\} =: \widetilde{\sigma}\{y\} \qquad \forall \pi \in \mathcal{G},$$

hence the density of the size-biased measure is

$$\frac{\mathrm{d}\ddot{D}_\alpha}{\mathrm{d}y}(y) = \widetilde{\sigma}\{y\} \sum_{\pi \in \mathcal{G}_D} D_\alpha(d(\pi^{-1} y)).$$

In particular, if $\alpha = (|\theta|/d, \ldots, |\theta|/d)$ for some $|\theta| > 0$ (symmetric Dirichlet), then its size-biased measure is

$$\ddot{D}_{|\theta|,d}(\mathrm{d}x) = d! \prod_{i=1}^{d-1} \frac{x_i}{1 - \sum_{j=1}^{i-1} x_j} D_\alpha(\mathrm{d}x) \tag{2.2}$$

$$\propto \prod_{i=1}^{d-1} b_i^{|\theta|/d} (1 - b_i)^{((d-i)/d)\theta - 1} \,\mathrm{d}b_i, \tag{2.3}$$

where $b_i = x_i/(1 - \sum_{j=1}^{i-1} x_j),\ i = 1, \ldots, d-1$. So if $\ddot{X}^{(d)}$ has distribution $\ddot{D}_{|\theta|,d}$, then

$$\ddot{X}^{(d)} \overset{d}{=} \big(\ddot{B}_1^{(d)}, \ldots, \ddot{B}_{d-1}^{(d)}\big),$$

where $(\ddot{B}_i^{(d)})$ are $d-1$ independent beta random variables with parameters, respectively, $(|\theta|/d + 1, (d - i/d)\theta),\ i = 1, \ldots, d-1$.

The measure $\ddot{D}_{|\theta|,d}$ is, again, a right-neutral measure.

Now, let $d \to \infty$. Then $\ddot{D}_{|\theta|,d}$ converges to the law of a right-neutral sequence $\ddot{X}^\infty = (\ddot{X}_1, \ddot{X}_2, \ldots)$ such that

$$\ddot{X}_j \overset{\mathcal{D}}{=} \ddot{B}_j \prod_{i=1}^{j-1} (1 - \ddot{B}_i), \qquad j \geq 1, \tag{2.4}$$

for a sequence $\ddot{B} = (\ddot{B}_1, \ddot{B}_2, \ldots)$ of independent and identically distributed (i.i.d.) beta weights with parameter $(1, |\theta|)$ (here and in the following pages, $\mathcal{D}$ means "*in distribution*").

Definition 2.1. *The random sequence* $\ddot{X}^\infty$ *satisfying* (2.4) *for a sequence of* beta$(1, |\theta|)$ *weights is called the GEM distribution with parameter* $|\theta|$ (GEM$(|\theta|)$).

Poisson point process construction [14].

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

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

The probability generating functional is

$$\mathcal{F}_{|\theta|}(\xi) = \mathbb{E}_{|\theta|}\left(\exp\left\{\int \log \xi(y) N_{|\theta|}(\mathrm{d}y)\right\}\right) = \exp\left\{|\theta| \int_0^\infty \big(\xi(y) - 1\big) y^{-1} \mathrm{e}^{-y}\, \mathrm{d}y\right\} \tag{2.5}$$

for suitable functions $\xi : \mathbb{R} \to [0, 1]$. The GEM($|\theta|$) distribution can be redefined in terms of the same point process Y^∞: Reorder the jumps by their size-biased random order, that is, set

$$\ddot{Y}_1 = Y_{i1}$$

with probability $Y_{i1}/|Y^\infty|$ and

$$\mathbb{P}(\ddot{Y}_{k+1} = Y_{i,k+1} \,|\ddot{Y}_1, \ldots, \ddot{Y}_k) = \frac{Y_{i,k+1}}{|Y| - \sum_{j=1}^k \ddot{Y}_j}, \qquad k = 1, 2, \ldots.$$

Denote the vector of all the size-biased jumps by $\ddot{Y}^\infty$. Then $|\ddot{Y}^\infty| \stackrel{\mathcal{D}}{=} |Y^\infty|$ is a gamma(θ) random variable, independent of the normalized sequence

$$\ddot{X}^\infty := \frac{\ddot{Y}^\infty}{|\ddot{Y}^\infty|}$$

and $\ddot{X}^\infty$ has the GEM($|\theta|$) distribution.

To intuitively convince oneself of such a statement, just notice that the probability generating functional of $\gamma^d_{\alpha,1}$, for $\alpha = (|\theta|/d, \ldots, |\theta|/d)$, is [10]

$$\begin{aligned} \mathcal{F}_{|\theta|,d}(\xi) &= \left(\int_0^\infty \xi(y) \gamma_{|\theta|/d,1}(\mathrm{d}y)\right)^d \\ &= \left(1 + \int_0^\infty \big(\xi(y) - 1\big) \frac{|\theta|}{d} \frac{y^{|\theta|/d-1} \mathrm{e}^{-y}}{\Gamma(|\theta|/d + 1)}\, \mathrm{d}y\right)^d \\ &\underset{d\to\infty}{\to} \mathcal{F}_{|\theta|}(\xi), \end{aligned} \tag{2.6}$$

so a finite size-biased collection of d i.i.d., normalized gamma jumps has a GEM(θ) limit distribution, as $d \to \infty$.

2.2.2. *Beta-Stacy distributions*

The measures D_α, $\ddot{D}_{|\theta|,d}$, $\ddot{D}_{|\theta|}$ are all right-neutral distributions with independent beta parameters.

Definition 2.2. *For $d \leq \infty$, let $B_1^*, \ldots, B_{d-1}^*$ be a collection of mutually independent beta random variables with parameters $\{\alpha_i, \beta_i\}_{i=1}^d$ (if $d = \infty$, take an infinite sequence of such weights). A random discrete distribution $X \in \Delta_{(d-1)}$ is said to have a* beta-Stacy *law if $X_1 \overset{\mathcal{D}}{=} B_1^*$ and, for every $j \leq d-1$,*

$$1 - \sum_{i=1}^{j-1} X_i \overset{\mathcal{D}}{=} \prod_{i=1}^{j-1} (1 - B_i^*).$$

A notable example of infinite-dimensional beta-Stacy distribution is the two-parameter GEM(α, θ) distribution [18,19] *whereby, for every $j \leq d-1$, B_j^* is a* beta$(1-\sigma, \theta + j\sigma)$ *random variable, with either $\sigma \in [0, 1]$ and $\theta > -\sigma$ or $\sigma < 0$ and $\theta = |\sigma| m$ for some $m \in \mathbb{N}$.*

The two-parameter GEM distribution is the most general class of right-neutral distributions that is also invariant under size-biased permutation; other remarkable properties (it is regenerative and Gibbs*) make it one of the most studied models for generating consistent, exchangeable random partitions (see* [20] *and references therein).*

2.3. Sampling formulae

The multinomial-Dirichlet distribution can be obtained by mixing the parameter of a multinomial distribution with a Dirichlet mixing measure: If X has D_α distribution,

$$DM_\alpha(r; |r|) = \mathbb{E}\left[\binom{|r|}{r} X^r\right] = \binom{|r|}{r} \frac{\prod_{i=1}^d (\alpha_i)_{(r_i)}}{(|\alpha|)_{(|r|)}}, \tag{2.7}$$

where $(a)_{(x)} := \Gamma(a+x)/\Gamma(a)$ for $a > 0$.

2.3.1. *Partial right-neutrality*

For every $r \in \mathbb{N}^d$ and $\alpha \in \mathbb{R}_+^d$, denote as usual $R_j = \sum_{i=j+1}^d r_i$ and $A_j = \sum_{i=j+1}^d \alpha_i$. It is easy to see that

$$\begin{aligned} DM_\alpha(r; R) &= \prod_{j=1}^{d-1} \binom{R_{j-1}}{r_j} \int_0^1 z_j^{r_j} (1 - z_j)^{R_j} D_{\alpha_j, A_j}(\mathrm{d}z_j) \\ &= \prod_{j=1}^{d-1} DM_{\alpha_j, A_j}(r_j; R_{j-1}). \end{aligned} \tag{2.8}$$

In other words, for every $j = 1, \ldots, d-1$, r_j/R_j is *conditionally* independent of $r_1, \ldots, r_{j-1}$, given R_j. Such a property, a direct consequence of the Dirichlet, is responsible for our construction of multivariate Hahn polynomials.

2.3.2. *Negative binomial sums*

Another construction of DM_α is possible, based on negative binomial random sequences, which parallels the gamma construction of the Dirichlet measure of Section 2.1.1.

Let $NB_{|\alpha|,y}(k) : |\alpha| > 0$, denote the *negative binomial distribution* with probability mass function:

$$NB_{|\alpha|,p}(k) = \frac{(|\alpha|)_{(k)}}{k!} p^k (1-p)^{|\alpha|}, \qquad k = 0, 1, \dots. \tag{2.9}$$

With both parameters in $\mathbb{N}$, such a measure describes the distribution of the number of failures occurring in a sequence of i.i.d. Bernoulli experiments (with success probability $1 - p$), before the αth success.

Two features of $NB_{|\alpha|,p}$ will prove useful, in Section 6, to connect multiple Meixner polynomials to multivariate Hahn polynomials.

(1) Poisson–gamma mixtures:

$$\begin{aligned} NB_{|\alpha|,p}(k) &= \int_0^\infty Po_\lambda(k)\gamma_{|\alpha|,p/(1-p)}(\mathrm{d}\lambda), \\ Po_\lambda(k) &= \frac{\lambda^k e^{-\lambda}}{k!}, \qquad k = 0, 1, 2, \dots. \end{aligned} \tag{2.10}$$

(2) Normalized negative binomial vectors.

Consider any $\alpha \in \mathbb{R}_+^d$ and $p \in (0, 1)$. Let $R_1, \dots, R_d$ be independent negative binomial random variables with parameter (α_i, p), respectively, for $i = 1, \dots, d$. Then

(i) $|R| := \sum_{i=1}^d R_i$ has law $NB_{|\alpha|,p}$.
(ii) Conditional on $|R| = |r|$, the vector $R = (R_1, \dots, R_d)$ has a Dirichlet-multinomial distribution with parameter $(\alpha, |r|)$:

$$\prod_{i=1}^d NB_{\alpha_i,p}(r_i) = NB_{|\alpha|,p}(|r|)\, DM_\alpha(r; |r|). \tag{2.11}$$

2.3.3. *Hypergeometric distribution*

Consider the form of the probability mass function DM_α but now replace the parameter α with $-\varepsilon = (-\varepsilon_1, \dots, -\varepsilon_d)$ with $0 \le n_j \le \varepsilon_j$, $j = 1, \dots, d$. Then

$$DM_{-\varepsilon}(n) = \frac{|n|!}{n_1! \cdots n_d!} \frac{(-\varepsilon)_{(n)}}{(-|\varepsilon|)_{(|n|)}} = \frac{\prod_{i=1}^d \binom{\varepsilon_i}{n_i}}{\binom{|\varepsilon|}{|n|}} =: H_\varepsilon(n). \tag{2.12}$$

$H_\varepsilon(n)$ is known as the *multivariate hypergeometric distribution* with parameter ε.

The partial right-neutrality property of the Dirichlet-multinomial distribution is preserved for the hypergeometric law; however, the interpretation as a Dirichlet mixture of i.i.d. laws is lost as the Dirichlet (as well as the gamma and the beta) integral is not defined for negative parameters.

2.4. Conjugacy properties

The gamma and the Dirichlet distribution, and, similarly, the negative binomial and the Dirichlet-multinomial distributions, are entangled by yet another property known in Bayesian statistics as *conjugacy* with respect to sampling.

A statistical model can be described by a probability triplet $\{M, \mathcal{M}, l_\Lambda\}_{\Lambda \in E}$, where the likelihood function $l_\Lambda(x)$ depends on a random parameter Λ living in some probability space $(E, \mathcal{E}, \pi)$. The distribution π of Λ is called the *prior* measure of the model. The *posterior* measure of the model is any version $\pi_x(\cdot) = \pi(\cdot|X = x)$ of the conditional probability satisfying

$$\int_A \pi(B|X = x) \int l_\lambda(\mathrm{d}x)\pi(\mathrm{d}x) = \int_B l_\lambda(A)\pi(\mathrm{d}\lambda) \qquad \text{a.s. } \forall A \in \mathcal{M}, B \in \mathcal{E}. \tag{2.13}$$

Definition 2.3. *Let $\mathcal{C}$ be a family of prior measures for a statistical model with likelihood l_Λ. $\mathcal{C}$ is* conjugate *with respect to l_Λ if*

$$\pi \in \mathcal{C} \quad \Longrightarrow \quad \pi_x \in \mathcal{C} \qquad \forall x.$$

It is easy to check that both gamma and Dirichlet measures are conjugate classes of prior measures. Bayes' theorem shows us the role as *marginal distributions* played, respectively, by $NB_{\alpha,p}$ and DM_α.

Example 2.4. The class of gamma priors is conjugate with respect to $l_\lambda = Po_\lambda$ on $\{0, 1, 2, \ldots\}$. The posterior measure is

$$\pi_x(\mathrm{d}\lambda) = \frac{Po_\lambda(x)\gamma_{\alpha,\beta}(\mathrm{d}\lambda)}{NB_{\alpha,\beta/(1+\beta)}(x)} = \gamma_{\alpha+x,\beta/(1+\beta)}(\mathrm{d}\lambda). \tag{2.14}$$

Similarly, the class of multivariate gamma priors $\{\gamma^d_{\alpha,\beta} : \alpha \in \mathbb{R}^d, \beta > 0\}$ is conjugate with respect to $\{Po^d_\lambda(x), \lambda \in \mathbb{R}^d_+, x \in \mathbb{N}^d\}$.

Example 2.5. The class of beta priors $\{D_{\alpha,\beta} : (\alpha, \beta) \in \mathbb{R}^2_+\}$ is conjugate with respect to the binomial likelihood $l_\lambda = B_\lambda(\cdot)$ on $\{0, 1, 2, \ldots, |n|\}$, for any integer $|n|$. The posterior distribution is

$$\pi_r(\mathrm{d}\lambda) = \frac{B_\lambda(|r|, |n - r|)D_{\alpha,\beta}(\mathrm{d}\lambda)}{DM_{\alpha,\beta}(|r|; |n| - |r|)} = D_{\alpha+|r|,\beta+|n|-|r|}(\mathrm{d}\lambda). \tag{2.15}$$

Similarly, the class of Dirichlet measures is conjugate with respect to multinomial sampling.

3. Jacobi polynomials on the simplex

If X, Y are independent random variables, their distribution $W_{X,Y}$ is the product $W_X W_Y$ of their marginal distributions, and therefore orthogonal polynomials $Q_{n,k}(x, y)$ in $W_{X,Y}$ are simply obtained by products $P_n(x)R_k(y)$ of orthogonal polynomials with W_X and W_Y as weight measures, respectively.

The key idea for deriving multivariate polynomials with respect to Dirichlet measures on the simplex, and to all related distributions treated in the subsequent sections, exploits the several properties of *conditional* independence enjoyed by the increments of D_α, as pointed out in Section 2.1.1. A method for constructing orthogonal polynomials in the presence of a particular kind of conditional independence, where Y depends on X only through a polynomial $\rho(x)$ of the first-order, is illustrated by the following multidimensional modification of Koornwinder's method (see [15], Section 3.7.2).

Proposition 3.1. *For $l, d \in \mathbb{N}$, let (X, Y) be a random point of $\mathbb{R}^l \times \mathbb{R}^d$ with distribution W. Let $\rho : \mathbb{R}^l \to \mathbb{R}$ define polynomials on $\mathbb{R}^l$ of order at most* 1.

Assume that the random variable

$$Z := \frac{Y}{\rho(X)}$$

is independent of X. Denote with W_X and W_Z the marginal distributions of X and Z, respectively. Then a system of multivariate polynomials, orthogonal with respect to W, is given by

$$\begin{aligned} G_n(x, y) &= P^{(N_l)}_{(n_1,\ldots,n_l)}(x)(\rho(x))^{N_l} R_{(n_{l+1},\ldots,n_{l+d})}\left(\frac{y}{\rho(x)}\right), \\ &(x, y) \in \mathbb{R}^l \times \mathbb{R}^d, n \in \mathbb{N}^{l+d}, \end{aligned} \tag{3.1}$$

where $N_l = n_{l+1} + \cdots + n_{l+d}$, $\{P^{(|m|)}_k\}_{k\in\mathbb{R}^l}$ and $\{R_m\}_{m\in\mathbb{R}^d}$ are systems of orthogonal polynomials with weight measures given by $(\rho(x))^{2|m|} W_X$ and W_Z, respectively.

Proof. When $d = l = 1$ this proposition is essentially a probabilistic reformulation of Koornwinder's construction ([15], Section 3.7.2). The proof is similar for any l, d. That G_n is a polynomial of degree $|n|$ is evident as the denominator of the term of maximum degree in R simplifies with $(\rho(x))^{n_{l+1}+\cdots+n_{l+d}}$. To show orthogonality, note that the assumption of conditional independence implies that

$$W(\mathrm{d}x, \mathrm{d}y) = W_X(\mathrm{d}x) W_Z\left(\frac{1}{(\rho(x))^d}\,\mathrm{d}y\right).$$

Denote $b_n = \mathbb{E}[P_n^2]$ and $c_n = \mathbb{E}[R_n^2]$, $n = 0, 1, 2, \ldots$. For $k, r \in \mathbb{R}^l$ and $m, s \in \mathbb{R}^d$,

$$\begin{aligned} &\int G_{(k,m)}(x, y) G_{(r,s)}(x, y) W(\mathrm{d}x, \mathrm{d}y) \\ &= \int P_k^m(x) P_r^s(x)(\rho(x))^{m+s} W_X(\mathrm{d}x) \int R_m(z) R_s(z) W_Z(\mathrm{d}z) \\ &= \int P_k^m(x) P_r^m(x)(\rho(x))^{2m} W_X(\mathrm{d}x) c_m \delta_{ms} \\ &= b_k c_m \delta_{kr} \delta_{ms}. \end{aligned}$$

□

3.1. $d = 2$; Jacobi polynomials on [0, 1]

For $d = 2$, D_α reduces to the beta distribution, the weight measure of (shifted) Jacobi polynomials. These are functions of one variable living in $\Delta_1 \equiv [0, 1]$. It is convenient to recall some known properties of such polynomials. Consider the measure

$$\widetilde{w}_{a,b}(\mathrm{d}x) = (1-x)^a(1+x)^b\mathbb{I}\big(x \in (-1,1)\big)\,\mathrm{d}x, \qquad a, b > -1, \tag{3.2}$$

where $\mathbb{I}(A)$ is the indicator function, equal to 1 if A, and 0 otherwise. This is the weight measure of the Jacobi polynomials defined by

$$\widetilde{P}_n^{a,b}(x) := \frac{(a+1)_{(n)}}{n!}{}_2F_1\left(\begin{matrix}-n,\ n+a+b+1\\ a+1\end{matrix}\;\middle|\;\frac{1-x}{2}\right),$$

where ${}_pF_q$, $p, q \in \mathbb{N}$, denote the hypergeometric function (see [1] for basic properties).

The normalization constants are given by the relation

$$\int_{(-1,1)} \widetilde{P}_n^{a,b}(x)\widetilde{P}_m^{a,b}(x)\widetilde{w}_{a,b}(\mathrm{d}x) = \frac{2^{a+b+1}}{2n+a+b+1}\frac{\Gamma(n+a+1)\Gamma(n+b+1)}{n!\Gamma(n+a+b+1)}\delta_{mn}. \tag{3.3}$$

The Jacobi polynomials are known to be solution of the second-order partial differential equation

$$(1-x^2)y''(x) + [b-a-x(a+b+2)]y'(x) = -n(n+a+b+1)y(x). \tag{3.4}$$

By a simple shift of measure it is easy to see that, for $\alpha, \beta > 0$ and $\theta := \alpha + \beta$, the modified polynomials

$$P_n^{\alpha,\beta}(x) = \frac{n!}{(n+\theta-1)_{(n)}}\widetilde{P}_n^{\beta-1,\alpha-1}(2x-1), \qquad \alpha, \beta > 0, \tag{3.5}$$

are orthogonal with respect to the beta distribution on [0, 1], which can be written as

$$D_{\alpha,\beta}(\mathrm{d}x) = \frac{\widetilde{w}_{\beta-1,\alpha-1}(\mathrm{d}u)}{2^{\alpha+\beta-1}B(\alpha,\beta)}, \tag{3.6}$$

where $u = 2x - 1$.

Denote the standardized Jacobi polynomials with

$$\widetilde{R}_n^{a,b}(x) = \frac{\widetilde{P}_n^{a,b}(x)}{\widetilde{P}_n^{a,b}(1)} \quad \text{and} \quad R_n^{\alpha,\beta}(x) = \frac{P_n^{\alpha,\beta}(x)}{P_n^{\alpha,\beta}(1)}.$$

Obviously

$$R_n^{\alpha,\beta}(x) = \widetilde{R}_n^{(\beta-1,\alpha-1)}(2x-1). \tag{3.7}$$

By (3.3) the new constant of proportionality is

$$\begin{aligned}\frac{1}{\zeta_n^{(\alpha,\beta)}} &:= \int_0^1 [R_n^{\alpha,\beta}(x)]^2 D_{\alpha,\beta}(\mathrm{d}x) \\ &= \left(\frac{(\theta+n-1)_{(n)}}{(\beta)_{(n)}}\right)^2 \frac{n!\alpha_{(n)}(\beta)_{(n)}}{(\theta)_{(2n)}(\theta+n-1)_{(n)}} \\ &= n!\frac{1}{(\theta+2n-1)(\theta)_{(n-1)}}\frac{(\alpha)_{(n)}}{(\beta)_{(n)}}, \qquad n=0,1,\ldots.\end{aligned} \tag{3.8}$$

A symmetry relation is

$$R_n^{\alpha,\beta}(x) = \frac{R_n^{\beta,\alpha}(1-x)}{R_n^{\beta,\alpha}(0)}. \tag{3.9}$$

Note that, if $\{P_n^{*\alpha,\beta}(x)\}$ is a system of *orthonormal* polynomials with weight measure $D_{\alpha,\beta}$, then

$$\zeta_n^{(\alpha,\beta)} = [P_n^{*\alpha,\beta}(1)]^2. \tag{3.10}$$

3.2. $2 \le d < \infty$. Multivariate Jacobi polynomials on the simplex from right-neutrality

A system of multivariate polynomials with respect to a Dirichlet distribution on $d \le \infty$ points can be derived by using its right-neutrality property, via Proposition 3.1. Let $\mathbb{N}_{d,|m|} = \{n = (n_1,\ldots,n_d) \in \mathbb{N}^d : |n| = |m|\}$. For every $n \in \mathbb{N}_{d-1,|n|}$ and $\alpha \in \mathbb{R}_+^d$ denote $N_j = \sum_{i=j+1}^{d-1} n_i$ and $A_j = \sum_{i=j+1}^{d} \alpha_i$.

Proposition 3.2. *For $d < \infty$, a system of multivariate orthogonal polynomials on the Dirichlet distribution D_α is given by*

$$R_n^{\alpha}(x) = \prod_{j=1}^{d-1} R_{n_j}^{(\alpha_j, A_j+2N_j)}\left(\frac{x_j}{1-s_{j-1}}\right)(1-s_{j-1})^{N_j}, \qquad x \in \Delta_{(d-1)}, \tag{3.11}$$

where $s_j = \sum_{i=1}^{j} x_i$.

Notice that $R_n^\alpha(\mathbf{e}_d) = 1$, where $\mathbf{e}_j := (\delta_{ij} : i = 1,\ldots,d)$. A similar definition for polynomials in the Dirichlet distribution is proposed in [16], in terms of non-shifted Jacobi polynomials $\widetilde{R}_n$. For an alternative choice of basis, see [5].

Proof of Proposition 3.2. The polynomials in $R_n^{\alpha}(x)$ given in Proposition 3.2 admit a recursive definition as follows:

$$\begin{aligned} &R^{\alpha}_{n_1,\dots,n_{d-1}}(x_1,\dots,x_d) \\ &\quad = R^{(\alpha_1,A_1+2N_1)}_{n_1}(x_1)(1-x_1)^{N_1} R^{\alpha_2^*}_{n_2,\dots,n_{d-1}}\left(\frac{x_2}{1-x_1},\dots,\frac{x_d}{1-x_1}\right), \end{aligned} \tag{3.12}$$

where $\alpha_j^* = (\alpha_j,\dots,\alpha_d)$ $(j \le d-1)$; so Proposition 3.1 is used with $l=1$, $\rho(x)=1-x$ and inductively on d. The claim is a consequence of the neutral-to-the-right property and Proposition 3.1 – for consider the orthogonality of a term

$$\left(1-\frac{X_j}{1-S_{j-1}}\right)^{N_j} R^{\alpha_j,A_j+2N_j}_{n_j}\left(\frac{X_j}{1-S_{j-1}}\right) \tag{3.13}$$

in R_n^{α} with a similar term in R_m^{α} for some $m=(m_1,\dots,m_{d-1})$-polynomial. Assume without loss of generality that for some $j=1,\dots,d-1$, $m_k=n_k$ for $k=j+1,\dots,d-1$ and $m_j < n_j$. Then $N_j = M_j$ and, multiplying the product of (3.13) by the corresponding beta density $D_{\alpha_j,A_j}(\mathrm{d}B_j)/\mathrm{d}B_j$, where B_j is as in (2.1), gives

$$B_j^{\alpha_j-1}(1-B_j)^{A_j+2N_j-1} R^{\alpha_j,A_j+2N_j}_{n_j}(B_j) R^{\alpha_j,A_j+2N_j}_{m_j}(B_j). \tag{3.14}$$

Since R_{n_j} is orthogonal to polynomials of degree less than n_j on the weight measure D_{α_j,A_j+2N_j}, then the integral with respect to $\mathrm{d}B_j$ of the quantity (3.14) vanishes, which proves the orthogonality. □

The orthogonality constant for $\{R_n^{\alpha}\}$ can be easily derived as

$$\begin{aligned} \frac{1}{\zeta_n^{\alpha}} &:= \int_{\Delta_{(d-1)}} (R_n^{\alpha}(x))^2 D_{\alpha}(\mathrm{d}x) = \frac{1}{\prod_{j=1}^{d-1} \zeta_{n_j}^{\alpha_j,A_j+2N_j}} \\ &= \prod_{j=1}^{d-1} \frac{n_j!(\alpha_j)_{(n_j)}}{(A_{j-1}+N_j)_{(n_j-1)}(A_{j-1}+2N_{j-1}-1)(A_j+2N_j)_{(n_j)}}. \end{aligned} \tag{3.15}$$

Notice that the same construction shown in Proposition 3.2 could be similarly expressed in terms of the polynomials $\{P_{n_j}^{\alpha_j,A_j+2N_j}\}$ or $\{P^{\star\alpha_j,A_j+2N_j}\}$ instead of $\{R_{n_j}^{\alpha_j,A_j+2N_j}\}$, the only difference resulting in the orthogonality constants.

3.3. Multivariate Jacobi on beta-Stacy distributions

Random distributions of beta-Stacy type are all right-neutral. Orthogonal polynomials with respect to general beta-Stacy measures can be therefore constructed in very much the same way as in Proposition 3.2, with a similar proof.

Proposition 3.3. *Let $d \le \infty$ and $(\alpha, \beta) \in \mathbb{R}_+^d \times \mathbb{R}_+^d$. Let $\mu_{\alpha,\beta}$ be the distribution of a* beta-Stacy(α, β) *random point of $\Delta_{(d-1)}$. A system of orthogonal polynomials in $\mu_{\alpha,\beta}$ is given by*

$$R_n^{*(\alpha,\beta)}(x) = \prod_{j=1}^{d-1} R_{n_j}^{(\alpha_j,\beta_j+2N_j)}\left(\frac{x_j}{1-s_{j-1}}\right)(1-s_{j-1})^{N_j}, \qquad x \in \Delta_{(d-1)}, n \in \mathbb{N}^d. \tag{3.16}$$

The constant of orthogonality is given by

$$\begin{aligned} \frac{1}{\zeta_n^{\alpha,\beta}} &= \frac{1}{\prod_{j=1}^{d-1} \zeta_{n_j}^{\alpha_j,\beta_j+2N_j}} \\ &= \prod_{i=1}^{d-1} \frac{n_i!(\alpha_i)_{(n_i)}}{(\alpha_i+\beta_i+2N_{i-1}-1)(\alpha_i+\beta_i+2N_i)_{(n_i-1)}(\beta_i+2N_i)_{(n_i)}}. \end{aligned} \tag{3.17}$$

Example 3.4. We have seen that all size-biased Dirichlet measures are beta-Stacy. A system of orthogonal polynomials in $\ddot{D}_{|\theta|,d}$ is

$$\begin{aligned} &\ddot{R}_n^{(|\theta|,d)}(x) = \prod_{j=1}^{d-1} R_{n_j}^{(|\theta|/d+1,((d-j)/d)\theta+2N_j)}\left(\frac{x_j}{1-s_{j-1}}\right)(1-s_{j-1})^{N_j}, \\ &x \in \Delta_{(d-1)}, n \in \mathbb{N}^d. \end{aligned} \tag{3.18}$$

Example 3.5. As $d \to \infty$, $\ddot{D}_{|\theta|,d}$ converges to the so-called GEM(θ) distribution, that is, an infinite-dimensional beta-Stacy with all i.i.d. weights being beta random variables with parameter $(\alpha_j, \beta_j) = (1, \theta)$. Let $\ddot{D}_{|\theta|,\infty} = \lim_{d\to\infty} \ddot{D}_{|\theta|,d}$ denote the GEM distribution with parameter $|\theta|$. For $|\theta| > 0$, an orthogonal system with respect to the weight measure $\ddot{D}_{|\theta|,\infty}$ is given by the polynomials:

$$\begin{aligned} &\ddot{R}_n^{|\theta|}(x) = \prod_{j=1}^{\infty} R_{n_j}^{(1,\theta+2N_j)}\left(\frac{x_j}{1-s_{j-1}}\right)(1-s_{j-1})^{N_j}, \\ &x \in \Delta_\infty, n \in \mathbb{N}^\infty : |n| = 0, 1, \ldots. \end{aligned} \tag{3.19}$$

Example 3.6. For the two-parameter GEM(σ, θ) distribution, $\alpha_j = 1-\sigma$ and $\beta_j = \theta + j\sigma$. The polynomials are of the form

$$\begin{aligned} &\ddot{R}_n^{\sigma,\theta}(x) = \prod_{j=1}^{\infty} R_{n_j}^{(1-\sigma,\theta+j\sigma+2N_j)}\left(\frac{x_j}{1-s_{j-1}}\right)(1-s_{j-1})^{N_j}, \\ &x \in \Delta_\infty, n \in \mathbb{N}^\infty : |n| = 0, 1, \ldots. \end{aligned} \tag{3.20}$$

4. Multivariate Jacobi and multiple Laguerre polynomials

The Laguerre polynomials, defined by

$$L_{|n|}^{|\alpha|}(y) = \frac{(|\alpha|)_{(|n|)}}{|n|!} {}_1F_1(-|n|; |\alpha|; y), \qquad |\alpha| > 0, \tag{4.1}$$

are orthogonal to the gamma density $\gamma_{|\alpha|,1}$ with constant of orthogonality

$$\int_0^\infty \left[L_{|n|}^{|\alpha|}(y)\right]^2 \gamma_{|\alpha|}(\mathrm{d}y) = \frac{(|\alpha|)_{(|n|)}}{|n|!}. \tag{4.2}$$

(Note that the usual convention is to define Laguerre polynomials in terms of the parameter $|\alpha'| := |\alpha| - 1 > -1$. Here we prefer to use positive parameter for consistency with the parameters in the amma distribution.)

Remark 4.1. If Y is a gamma($|\alpha|$) random variable, then, for every scale parameter $\beta \in \mathbb{R}_+$, the distribution of $Z := \beta Y$ is $\gamma_{|\alpha|,\beta}(\mathrm{d}z)$. Thus the system

$$\left\{ L_n^{|\alpha|}\left(\frac{z}{\beta}\right)\right\}_{n=0,1,\dots}$$

is orthogonal with weight measure $\gamma_{|\alpha|,\beta}$.

Let $Y \in \mathbb{R}_+^d$ be a random vector with distribution $\gamma_{\alpha,\beta}^d$. By the stochastic independence of its coordinates, orthogonal polynomials of degree $|n|$ with the distribution of Y as weight measure are simply

$$L_n^{\alpha,\beta}(y) = \prod_{i=1}^d L_{n_i}^{\alpha_i}\left(\frac{y_i}{\beta}\right), \qquad y \in \mathbb{R}^d, n \in \mathbb{N}_n, \tag{4.3}$$

with constants of orthogonality of

$$\frac{1}{\varphi_n} = \mathbb{E}(L_n^\alpha(Y))^2 = \prod_{i=1}^d \frac{(\alpha_i)_{(n_i)}}{n_i!}. \tag{4.4}$$

Therefore, with the notation introduced in Section 2.1.1, because of the one-to-one mapping

$$(Y_1, \dots, Y_d) \mapsto (|Y|, X_1, \dots, X_d),$$

one can obtain an alternative system of orthogonal polynomials from $y_1, \dots, y_n$.

Proposition 4.2. *The polynomials defined by*

$$L_n^{\alpha,\beta*}(y) = L_{n_d}^{|\alpha|+2|n'|}\left(\frac{|y|}{\beta}\right)\left(\frac{|y|}{\beta}\right)^{|n'|} R_{n'}^\alpha\left(\frac{y}{|y|}\right), \qquad n \in \mathbb{N}^d, y \in \mathbb{R}^d, \tag{4.5}$$

with $n' = (n_1, \dots, n_{d-1})$ *and* R_m^α *defined by* (3.11), *are orthogonal with respect to* $\gamma^d_{\alpha,\beta}$.

Proof. The proof of (4.5) is straightforward and follows immediately from Proposition 3.1, with $l = 1$, $X = |Y|$ and $\rho(x) = x$ (remember that $|Y|$ is gamma with parameter $(|\alpha|, \beta)$). □

From now on we will only consider the case with $\beta = 1$, without much loss of generality. The constant of orthogonality of the resulting system $\{L_n^{\alpha *}\}$ is

$$\begin{aligned}
\frac{1}{\varphi_n^*} &:= \int_{\mathbb{R}^d} [L_n^{\alpha *}(y)]^2 \prod_{i=1}^d \gamma_{\alpha_i}(\mathrm{d}y_i) \\
&= \int_0^\infty \bigl[L_{n_d}^{|\alpha|+2(|n|-n_d)}(|y|)|y|^{|n|-n_d}\bigr]^2 \gamma_{|\alpha|}(\mathrm{d}|y|) \int_{\Delta_{(d-1)}} [R_{n'}^\alpha(x)]^2 D_\alpha(\mathrm{d}x) \\
&= \frac{(|\alpha|)_{(2|n'|)}}{\zeta_{n'}^\alpha} \int \bigl[L_{n_d}^{|\alpha|+2|n'|}(|y|)\bigr]^2 \gamma_{\alpha+2|n'|}(\mathrm{d}|y|) \\
&= \frac{1}{n_d!} \frac{((|\alpha|)_{(2|n'|)})^2}{\zeta_{n'}^\alpha},
\end{aligned} \tag{4.6}$$

where $\zeta_{n'}^\alpha$ is as in (3.15).

4.1. Connection coefficients

The two systems L_n^α and $L_n^{\alpha *}$ can be expressed as linear combinations of each other:

$$L_n^{\alpha *}(y) = \sum_{|m|=|n|} \varphi_m c_m^*(n) L_m^\alpha(y) \tag{4.7}$$

and

$$L_n^\alpha(y) = \sum_{|m|=|n|} \varphi_m^* c_m(n) L_m^{\alpha *}(y), \tag{4.8}$$

where

$$c_m^*(n)\delta_{|m||n|} = \mathbb{E}[L_n^{\alpha *}(y) L_m^\alpha(y)] = c_n(m)\delta_{|m||n|}.$$

For general m, n a representation for $c_m^*(n)$ can be derived in terms of a mixture of Lauricella functions of the first (A) type. Such functions are defined [17] as

$$F_A(|a|; b; c; z) = \sum_{m \in \mathbb{N}^d} \frac{1}{m_1! \cdots m_d!} \frac{|a|_{(|m|)} b_{(m)}}{c_{(m)}} z^m, \qquad a, b, c, z \in \mathbb{C}^d,$$

where $v_{(r)} := \prod_{i=1}^d (v_i)_{(r_i)}$ for every $v, r \in \mathbb{R}^d$.

Proposition 4.3. *For every $n \in \mathbb{N}^d$ denote $n' := (n_1, \dots, n_{d-1})$. A representation for the connection coefficients in (4.7) is*

$$c_m^*(n) = \delta_{mn} \frac{(|\alpha|)_{(|n|)}}{|n|!} DM_\alpha(m) \\ \times \sum_{j=0}^{|n|} d_j \int_{\Delta_{(d-1)}} R_{n'}^\alpha(t) F_A(|\alpha|; -m, -j; \alpha, |\alpha|; t, 1-|t|, 1) D_\alpha(\mathrm{d}t), \tag{4.9}$$

where

$$d_j := \sum_{i=0}^{|n'|} (-|n'|)_{(i)} \frac{(|\alpha|)_{(|n'|)}(|\alpha| + 2|n'|)_{(n_d)}}{i! n_d!} \\ \times F_A(|\alpha|; -i, -n_d, -j; |\alpha|, |\alpha| + 2i, |\alpha|; 1, 1, 1). \tag{4.10}$$

The proof relies on a beautiful representation due to Erdélyi [6]: for every $|a|, |z| \in \mathbb{R}$, $\alpha, k \in \mathbb{R}^d$ and $n \in \mathbb{N}^d$,

$$\prod_{j=1}^d L_{n_j}^{\alpha_j}(k_j |z|) = \sum_{s=0}^{|n|} \phi_s(|a|; \alpha; n; k) L_s^{|a|}(|z|), \tag{4.11}$$

where

$$\phi_s(|a|; \alpha; n; k) = F_A(|a|; -n, -s; \alpha, |a|; k, 1) \prod_{j=1}^d \frac{(\alpha_j)_{(n_j)}}{n_j!}.$$

The full proof of Proposition 4.3 involves tedious algebra that we omit here as not relevant for the general purposes of the paper.

Remark 4.4. A simplified representation of $c_m^*(n)$ in terms of Hahn polynomials will be given in Section 5.2.2.

Remark 4.5. Note that when $|n'| = 0$, $c_m^*(0, \dots, 0, n_d) = 1$, which agrees with the known identity

$$L_n^{\alpha+\beta}(x + y) = \sum_{j=0}^n L_j^\alpha(x) L_{n-j}^\beta(y), \qquad x, y \in \mathbb{R} \tag{4.12}$$

(see [2], formula (6.2.35), page 191), an identity with an obvious extension to the d-dimensional case.

Remark 4.6. It is immediate to verify that the coefficients $c_m^*(n)$ also satisfy

$$L_{|n-n'|}^{|\alpha|}(|\beta^{-1} y|) |\beta^{-1} y|^{|n'|} R_{n'}^\alpha\left(\frac{y}{|y|}\right) = \sum_{|m|=|n|} \varphi_m c_m^*(n) L_m^\alpha(|\beta^{-1} y|), \qquad \beta \in \mathbb{R}_+. \tag{4.13}$$

4.2. Size-biased multiple Laguerre

Let $Y^d = (Y_1, \dots, Y_d)$ be a collection of independent gamma random variables, each with parameters $(\theta/d, 1)$, $i = 1, \dots, d$. Let $\ddot{Y}^d$ be the same vector with the coordinates rearranged in size-biased random order. The proof of the following corollaries is, at this point, obvious from Proposition 4.2.

Corollary 4.7. *A system of polynomials, orthogonal with respect to the law of* $\ddot{Y}^d$, *is given by*

$$\ddot{L}^{|\theta|,d}_{(|m|,n')}(y) = L^{|\theta|+2|n'|}_{|m|}(|y|)(|y|)^{|n'|}\ddot{R}^{|\theta|,d}_{n'}\left(\frac{y}{|y|}\right), \tag{4.14}$$

$|m| \in \mathbb{N}$, $n' \in \mathbb{N}^d : |n'| \in \mathbb{N}$, *with* $\{\ddot{R}_n\}$ *as in* (3.18).

It is possible to derive an infinite-dimensional version of $\{L^{\alpha\star}_n\}$, orthogonal with respect to the law of the size-biased point process $\ddot{Y}^\infty$, obtained by Y^∞ of Section 2.2.1. Remember that $\ddot{X}^\infty := \ddot{Y}^\infty/|\ddot{Y}^\infty|$ has GEM($|\theta|$) distribution and it is independent of $|\ddot{Y}^\infty| \overset{\mathcal{D}}{=} |Y^\infty|$, which has a gamma($|\theta|$) law.

Corollary 4.8. *Let* $\ddot{\gamma}_{|\theta|}$ *be the probability distribution of the size-biased sequence* $\ddot{Y}^\infty$ *obtained by rearranging in size-biased random order the sequence* Y^∞ *of points of a Poisson process with generating functional* (2.5). *The polynomials defined by*

$$\ddot{L}^{|\theta|}_{(|m|,n')}(y) = L^{|\theta|+2|n'|}_{|m|}(|y|)(|y|)^{|n'|}\ddot{R}^{|\theta|}_{n'}\left(\frac{y}{|y|}\right) \tag{4.15}$$

for $|m| \in \mathbb{N}$, $n' \in \mathbb{N}^\infty : |n'| \in \mathbb{N}$, *with* $\{\ddot{R}_n\}$ *as in* (3.19), *are the limit, as* $d \to \infty$, *of the polynomials* $\{\ddot{L}^{|\theta|,d}_{(|m|,n')}\}$ *defined by* (4.14) *and form an orthogonal system with respect to* $\ddot{\gamma}_{|\theta|}$.

5. Multivariate Hahn polynomials

5.1. Hahn polynomials on {1, . . . ,N}

As for the Laguerre polynomials, we introduce the discrete Hahn polynomials on $\{1, \dots, N\}$ with parameters shifted by 1 to make the notation consistent with the standard probabilistic notation in the corresponding weight measure. The Hahn polynomials, orthogonal on $DM_{\alpha,\beta}(n; N)$, are defined as the hypergeometric series:

$$h^{\alpha,\beta}_n(r; N) = {}_3F_2\left(\begin{matrix} -n, n+\theta-1, -r \\ \alpha, -N \end{matrix} \,\middle|\, 1\right), \qquad n = 0, 1, \dots, N. \tag{5.1}$$

The orthogonality constants are given by

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

A special point value is ([12], formula (1.15))

$$h_n^{\alpha,\beta}(N;N) = (-1)^n \frac{(\beta)_{(n)}}{(\alpha)_{(n)}}. \tag{5.2}$$

Thus if we consider the normalization

$$q_n^{\alpha,\beta}(r;N) := \frac{h_n^{\alpha,\beta}(r;N)}{h_n^{\alpha,\beta}(N;N)},$$

then the new constant is, from (5.2),

$$\begin{aligned}\frac{1}{w_{N,n}^{\alpha,\beta}} &:= \mathbb{E}[q_n^{\alpha,\beta}(R;N)]^2 \\ &= \frac{1}{\binom{N}{n}} \frac{(\theta+N)_{(n)}}{(\theta)_{(n-1)}} \frac{1}{\theta+2n-1} \frac{(\alpha)_{(n)}}{(\beta)_{(n)}} \\ &= \left[\frac{(\theta+N)_{(n)}}{N_{[n]}}\right] \frac{1}{\zeta_n^{\alpha,\beta}},\end{aligned} \tag{5.3}$$

where ζ_n is the Jacobi orthogonality constant, given by (3.8).

A symmetry relation is

$$q_n^{\alpha,\beta}(r;N) = \frac{q_n^{\beta,\alpha}(N-r;N)}{q_n^{\beta,\alpha}(0;N)}. \tag{5.4}$$

A well-known relationship is in the limit:

$$\lim_{N\to\infty} h_n^{\alpha,\beta}(Nz;N) = \widetilde{R}_n^{\alpha-1,\beta-1}(1-2z), \qquad \alpha,\beta > 0 \tag{5.5}$$

(see [12]), where $\widetilde{R}_n^{a,b} = \widetilde{R}_n^{a,b}/\widetilde{R}_n^{a,b}(1)$ are standardized Jacobi polynomials orthogonal on $[-1,1]$ as defined in Section 3.1. Because of our definition (3.5), combining (3.9), (5.4) and (5.6) gives the equivalent limit: For every n,

$$\lim_{N\to\infty} q_n^{\alpha,\beta}(Nz;N) = R_n^{\alpha,\beta}(z), \qquad \alpha,\beta > 0. \tag{5.6}$$

Note that also

$$\lim_{N\to\infty} w_{N,n}^{\alpha,\beta} = \zeta_n^{\alpha,\beta}. \tag{5.7}$$

An inverse relation holds as well, which allows one to derive Hahn polynomials as a mixture of Jacobi polynomials. Denote by $B_x(r;N) = B_{x,1-x}(r,N-r)$ the binomial distribution.

Proposition 5.1. *The functions*

$$\tilde{q}_n^{\alpha,\beta}(r;N) := \int_0^1 R_n^{\alpha,\beta}(x)\frac{B_x(r;N)}{DM_{\alpha,\beta}(r;N)}D_{\alpha,\beta}(\mathrm{d}x) \tag{5.8}$$

$$= \int_0^1 R_n^{\alpha,\beta}(x)D_{\alpha+r,\beta+N-r}(\mathrm{d}x), \qquad n = 0, 1, \ldots, N, \tag{5.9}$$

form the Hahn system of orthogonal polynomials with $DM_{\alpha,\beta}$ *as the weight function, such that*

$$\tilde{q}_n^{\alpha,\beta}(r;N) = \frac{N_{[n]}}{(\theta+N)_{(n)}}q_n^{\alpha,\beta}(r;N). \tag{5.10}$$

The representation (5.9), in particular, shows a Bayesian interpretation of Hahn polynomials, as a *posterior* mixture of Jacobi polynomials evaluated on a random Bernoulli probability of success X, conditionally on having previously observed r successes out of N independent Bernoulli(X) trials, where X has a Beta(α,β) distribution on $\{0,\ldots,N\}$.

Proof of Proposition 5.1. The integral defined by (5.8) is a polynomial: Consider

$$\int_0^1 x^n(1-x)^m \frac{B_x(r;N)}{DM_{\alpha,\beta}(r;N)}D_{\alpha,\beta}(\mathrm{d}x) = \frac{(\alpha)_{(n+r)}(\beta)_{(N+m-r)}(\theta)_{(N)}}{(\alpha)_{(r)}(\beta)_{(N-r)}(\theta)_{(N+n+m)}}$$
$$= \frac{(\alpha+r)_{(n)}(\beta+N-r)_{(m)}}{(\theta+N)_{(n+m)}}.$$

The numerator is a polynomial in r of order $n+m$. Write

$$R_n^{\alpha,\beta}(x) = \sum_{j=1}^n c_j x^j,$$

then

$$\int_0^1 R_n^{\alpha,\beta}(x)\frac{B_x(r;N)}{DM_{\alpha,\beta}(r;N)}D_{\alpha,\beta}(\mathrm{d}x) = \sum_{j=1}^n \frac{c_j}{(\theta+N)_{(j)}}(\alpha+r)_{(j)}$$
$$= \sum_{j=1}^n \frac{c_j}{(\theta+N)_{(j)}}r_{[j]} + L, \tag{5.11}$$

where L is a polynomial in r of order less than n. Then $q_n^{\alpha,\beta}(r)$ is a polynomial of order n in r.

To show orthogonality it is sufficient to show that h_n are orthogonal with respect to polynomials of the basis formed by the falling factorials $\{r_{[l]}, l = 0, 1, \ldots\}$. For $l \le n$,

$$\begin{aligned}
&\sum_{r=0}^{n} DM_{\alpha,\beta}(r; N) r_{[l]} \tilde{q}_n^{\alpha,\beta}(r; N) \\
&= \frac{N!}{(N-l)!} \int_0^1 x^l R_n^{\alpha,\beta}(x) \left[\sum_{r=0}^{n} \binom{N-l}{r-l} x^{l-r}(1-x)^{N-r} \right] D_{\alpha,\beta}(\mathrm{d}x) \qquad (5.12) \\
&= N_{[l]} \int_0^1 x^l R_n^{\alpha,\beta}(x) D_{\alpha,\beta}(\mathrm{d}x).
\end{aligned}$$

The last integral is non-zero only if $l = n$, which proves the orthogonality of $q_n^{\alpha,\beta}(r; N)$.

Now consider that, in $R_n^{\alpha,\beta}(x)$, the leading coefficient c_n satisfies

$$\int_0^1 c_n x^n R_n^{\alpha,\beta}(x) D_{\alpha,\beta}(\mathrm{d}x) = \int_0^1 [R_n^{\alpha,\beta}(x)]^2 D_{\alpha,\beta}(\mathrm{d}x) = \frac{1}{\zeta_n^{\alpha,\beta}};$$

$$\begin{aligned}
\frac{1}{\omega_{N,n}^{\alpha,\beta}} &= \sum_{r=0}^{n} DM_{\alpha,\beta}(r; N) \tilde{q}_n^{\alpha,\beta}(r; N) \tilde{q}_n^{\alpha,\beta}(r; N) \\
&= \sum_{r=0}^{n} DM_{\alpha,\beta}(r; N) \left(\sum_{j=0}^{n} \frac{c_j}{(\theta+N)_{(j)}} r_{[j]} \right) \tilde{q}_n^{\alpha,\beta}(r; N) + L' \\
&= N_{[n]} \frac{c_n}{(\theta+N)_{(n)}} \int x^n R_n^{\alpha,\beta}(x) D_{\alpha,\beta}(\mathrm{d}x) \qquad (5.13) \\
&= \frac{N_{[n]}}{(\theta+N)_{(n)}} \frac{1}{\zeta_n^{\alpha,\beta}}.
\end{aligned}$$

That is,

$$\omega_{N,n}^{\alpha,\beta} = \left[\frac{(\theta+N)_{(n)}}{N_{[n]}} \right]^2 w_{N,n}^{\alpha,\beta} \qquad (5.14)$$

with $w_{N,n}^{\alpha,\beta}$ as in (5.2), and therefore the identity (5.10) follows, completing the proof. □

5.2. Multivariate polynomials on the Dirichlet-multinomial distribution

Multivariate polynomials orthogonal with respect to DM_α on the discrete d-dimensional simplex were first introduced by Karlin and McGregor [11] as eigenfunctions of the birth-and-death process with neutral mutation. Here we derive an alternative derivation as a posterior mixture of multivariate Jacobi polynomials, which extends Proposition 5.1 to a multivariate setting.

Proposition 5.2. *For every $\alpha \in \mathbb{R}^d$, a system of polynomials, orthogonal with respect to DM_α, is given by*

$$\tilde{q}_n^\alpha(r;|r|) = \int_{\Delta_{(d-1)}} R_n^\alpha(x)\frac{B_x(r)}{DM_\alpha(r)}D_\alpha(\mathrm{d}x) \tag{5.15}$$

$$= \int_{\Delta_{(d-1)}} R_n^\alpha(x)D_{\alpha+r}(\mathrm{d}x), \qquad |n|\le|r| \tag{5.16}$$

$$= \left(\frac{\prod_{j=1}^{d-1}(A_j+R_j+N_{j+1})_{(n_{j+1})}}{(|\alpha|+|r|)_{(N_1)}}\right)\prod_{j=1}^{d}\tilde{q}_{n_j}^{\alpha_j,A_j+2N_j}(r_j;R_{j-1}-N_j), \tag{5.17}$$

with constant of orthogonality given by

$$\frac{1}{\omega_n(\alpha;|r|)} := \mathbb{E}[\tilde{q}_n^\alpha(R;|r|)]^2 = \frac{|r|_{[n]}}{(|\alpha|+|r|)_{(n)}}\frac{1}{\zeta_n^\alpha}. \tag{5.18}$$

Proof. The identity between (5.15) and (5.16) is obvious from Section 2.4 and (5.17) follows from Proposition 5.1 and some simple algebra. For every $n \in \mathbb{N}^d$,

$$\begin{aligned}\int_{\Delta_{(d-1)}} x^n D_{\alpha+r}(\mathrm{d}x) = DM_{\alpha+r}(n) &= \prod_{i=1}^{d-1}\frac{(\alpha_i+r_i)_{(n_i)}(A_i+R_i)_{(N_i)}}{(A_{i-1}+R_{i-1})_{(N_{i-1})}}\\ &= \frac{\prod_{i=1}^{d}(\alpha_i+r_i)_{(n_i)}}{(|\alpha|+|r|)_{(|n|)}} = \frac{1}{(|\alpha|+|r|)_{(|n|)}}\prod_{i=1}^{d} r_{i[n_i]} + L,\end{aligned} \tag{5.19}$$

where L is a polynomial in r of order less than $|n|$. Therefore $\tilde{q}_n^\alpha(r;|r|)$ are polynomials of order $|n|$ in r.

To show that they are orthogonal, denote

$$p_l(r) := \prod_{i=1}^{d}(r_i)_{[l_i]}$$

and consider that, for every $l \in \mathbb{N}^d : |l| \le |n|$,

$$\begin{aligned}&\sum_{|m|=|r|} DM_\alpha(m;|r|)p_l(m)\tilde{q}_n^\alpha(m;|r|)\\ &\quad= \frac{|r|!}{(|r|-|l|)!}\int x^l R_n^\alpha(x)\left(\sum_{|m|=|r|}\binom{|r-l|}{m-l}x^{m-l}\right)D_\alpha(\mathrm{d}x)\\ &\quad= |r|_{[|l|]}\int x^l R_n^\alpha(x)D_\alpha(\mathrm{d}x),\end{aligned} \tag{5.20}$$

which, by orthogonality of R_n, is non-zero only if $|l| = |n|$. Since it is always possible to write, for appropriate coefficients c_{nm}

$$R_n^\alpha(x) = \sum_{|m|=|n|} c_{nm} x^m + C,$$

where C is a polynomial of order less than $|n|$ in x; then

$$\tilde{q}_s^\alpha(r; |r|) = \sum_{|m|=|s|} \frac{c_{sm}}{(|\alpha| + |r|)_{(|s|)}} p_m(r) + C'$$

and by (5.20)

$$\begin{aligned}
\mathbb{E}[\tilde{q}_s^\alpha(R; |r|)\tilde{q}_n^\alpha(R; |r|)] &= \sum_{|k|=|s|} \frac{c_{sk}}{(|\alpha| + |r|)_{(|s|)}} \mathbb{E}[p_k(R)\tilde{q}_n^\alpha(R; |r|)] + C'' \\
&= |r|_{[|n|]} \sum_{|k|=|r|} \frac{c_{sk}}{(|\alpha| + |r|)_{(|s|)}} \int x^k R_n^\alpha(x) D_\alpha(\mathrm{d}x) \\
&= \frac{|r|_{[|n|]}}{(|\alpha| + |r|)_{(|n|)}} \frac{1}{\zeta_n^\alpha} \delta_{sn}, \qquad |n| = |r|.
\end{aligned}$$

□

Remark 5.3. Note that the representation (5.17) holds also for negative parameters, so that, if we replace α with $-\varepsilon$ ($\varepsilon \in \mathbb{R}^d$) then (5.17) is a representation for polynomials with respect to the hypergeometric distribution (Section 2.3.3).

5.2.1. *Bernstein–Bézier coefficients of Jacobi polynomials*

As anticipated in the introduction, Proposition 5.2 gives a probabilistic proof of a recent result of [22], namely that Hahn polynomials are the Berstein–Bézier coefficients of the multivariate Jacobi polynomials. Remember that the Bernstein polynomials, when taken on the simplex, are essentially multinomial distributions $B_x(n) = \binom{|n|}{n} x^n$, seen as functions of x.

Corollary 5.4. *For every $d \in \mathbb{N}$, $\alpha \in \mathbb{R}^d$, $r \in \mathbb{N}^d$,*

$$R_r^\alpha(x) = \frac{(|\alpha| + |r|)_{(|n|)}}{|r|_{[|n|]}} \sum_{|m|=|r|} \tilde{q}_r^\alpha(m; |r|) B_x(m), \tag{5.21}$$

where $\omega_r(|\alpha|; |r|)$ is given by (5.18).

Proof. From Proposition 5.2,

$$DM_\alpha(m; |r|)\tilde{q}_r^\alpha(m; |r|) = \mathbb{E}[B_X(m) R_r^\alpha(X)]$$

so

$$B_x(m) = DM_\alpha(m; |m|) \sum_{|n|=0}^{|m|} \zeta_n^\alpha \tilde{q}_n^\alpha(m; |m|) R_n(x).$$

Hence

$$\sum_m \tilde{q}_r^\alpha(m;|r|)B_x(m)$$
$$= \sum_{|n|=0}^{|r|} \zeta_n^\alpha \left[\sum_{|m|=|r|} DM_\alpha(m;|r|)\tilde{q}_r^\alpha(m;|r|)\tilde{q}_n^\alpha(m;|r|)\right] R_n^\alpha(x) \tag{5.22}$$
$$= \sum_{|n|=0}^{|r|} \frac{\zeta_n^\alpha}{\omega_r(|\alpha|;|r|)}\delta_{rn} R_n^\alpha(x) = \frac{|r|_{[|n|]}}{(|\alpha|+|r|)_{(|n|)}} R_r^\alpha(x),$$

which completes the proof. □

Remark 5.5. By a similar argument it is easy to come back from (5.21) to (5.15).

5.2.2. *The connection coefficients of Proposition* 4.3

Consider again the connection coefficients $c_n^*(m)$ of Proposition 4.3 and their representations (4.9) and (4.10). An alternative representation can be given in terms of multivariate Hahn polynomials.

Corollary 5.6. *Let $c_n^*(m)$ be the connection coefficients between $L_n^{\alpha *}$ and L_m^α, as in Section* 4. *Then*

$$c_n^*(m) = \delta_{mn}\, b_{|n|,n_d}^{|\alpha|} DM_\alpha(m) \sum_{|r|=0}^{|n|} \frac{(-m)_{(r)}}{\prod_{l=1}^d r_l!} \tilde{q}_{n'}^\alpha(r;|r|), \tag{5.23}$$

where $n' = (n_1,\dots,n_d - 1)$,

$$b_{|n|,n_d}^{|\alpha|} = \frac{(|\alpha|)_{(|n|)}}{|n|!}\left[\sum_{j=0}^{|n|} \frac{d_j}{j!|\alpha|_{(j)}}\right]$$

and d_j is as in (4.10).

Proof. It is sufficient to use the explicit expression of the Lauricella function F_A in (4.9) to see that

$$\begin{aligned} c_m^*(n) &= \delta_{mn} \frac{(|\alpha|)_{(|n|)}}{|n|!} DM_\alpha(m) \left[\sum_{j=0}^{|n|} \frac{d_j}{j!|\alpha|_{(j)}}\right] \sum_{|r|=0}^{|n|} \frac{(-m)_{(r)}}{\prod_{l=1}^d r_l!} \int \frac{\binom{|r|}{r} t^r R_{n'}^\alpha(t)}{DM_\alpha(r)} D_\alpha(\mathrm{d}t) \\ &= \delta_{mn} b_{|n|,n_d}^{|\alpha|} DM_\alpha(m) \sum_{|r|=0}^{|n|} \frac{(-m)_{(r)}}{\prod_{l=1}^d r_l!} \tilde{q}_{n'}^\alpha(r;|r|). \end{aligned} \tag{5.24}$$

□

5.2.3. *Application*: *The d-types linear growth model*

The multivariate Hahn polynomials were first studied by Karlin and McGregor [11] to derive the transition density of the so-called d-type neutral Moran model of population genetics. This is, for any fixed $|r| \in \mathbb{N}$, a stochastic process $(N(t) : t \geq 0)$ living in the discrete simplex $\mathbb{N}_{d,|r|} = \{m \in \mathbb{N}^d : |m| = |r|\}$, with Dirichlet-multinomial stationary distribution, and whose generator has Hahn polynomials as eigenfunctions.

Karlin and McGregor's description of such eigenfunctions is structurally similar to our (5.17), up to some re-scaling and reordering of the variables.

In the same paper ([11], formula (6.2)), the functions (rewritten in our notation)

$$\psi(m) := \binom{|r|}{|m|} L_{|r|-|m|}^{|\alpha|+2|m|}(|y|)\tilde{q}_n^{\alpha}(m; |m|), \qquad m \in \mathbb{N}^d : |m| \leq |r|, |r| \in \mathbb{N},$$

were introduced to connect the d-type Moran model of reproduction to a d-type linear growth model with immigration rates proportional to $\alpha_1, \dots, \alpha_d$. The generator of the latter process has eigenfunctions that are the solution of the recursion

$$-|y|\psi(m) = \sum_{i=1}^{d} m_i[\psi(m - e_i) - \psi(m)] + \sum_{i=1}^{d}(m_i + \alpha_i)[\psi(m - e_i) - \psi(m)].$$

Note that, for every $z \in \mathbb{R}^d$ such that $|z| = |y|$, $\psi(m) = L^{\alpha}_{|r|-|m|,m}(y)$ is also a solution, hence so is $\psi(m) = L^{\alpha*}_{|r|-|m|,m}(z)$.

Reconsider now the system $L_n^{\alpha*}$ of multiple Laguerre polynomials. In view of our representation (5.16) of Hahn polynomials, it is easy to write

$$\psi(m) = \binom{|r|}{|m|} \frac{\Gamma(|\alpha|)}{\Gamma(\alpha)} \int_{\mathbb{R}^{d-1}} L^{\alpha*}_{|r|-|m|,m}(y) \frac{1}{|y|^{d-1}} y^{\alpha-1} \mathrm{d}y_1 \cdots \mathrm{d}y_{d-1},$$

which is identical to

$$\psi(m) = \binom{|r|}{|m|} L_{|r|-|m|}^{|\alpha|+2|m|}(|y|) \int_{\Delta_{d-1}} R_m^{\alpha}(x) D_{\alpha+m}(\mathrm{d}x).$$

Our representation in a sense completes Karlin and McGregor's analysis, in terms of eigenfunctions, of the relationship existing between the r-type linear growth model (product of independent Laguerre polynomials), the Moran model (multivariate Hahn) and its scaling limit, the d-type Wright–Fisher diffusion (multivariate Jacobi). In [11] the role of the latter was not very visible. The representation (5.16) shows how to map directly polynomial eigenfunctions of the scaling limit process (Jacobi) to polynomial eigenfunctions of its finite-size dual model (Hahn). In Karlin and McGregor's work this idea was present only implicitly (see their formula (3.8) and observation (3.10)), via their use of Laguerre products. Considering the system $\{L^{\alpha*}_{|r|-|m|,m}\}$ makes the connection between all the three processes more transparent.

6. Multivariate Hahn and multiple Meixner polynomials

The Meixner polynomials on $\{0, 1, 2, \dots\}$, defined by

$$M_n(k; \alpha, p) = {}_2F_1\left(\begin{matrix} -n, \ -k \\ \alpha \end{matrix} \,\middle|\, \frac{p-1}{p}\right), \qquad \alpha > 0, p \in (0, 1), \tag{6.1}$$

are orthogonal with respect to the negative binomial distribution $NB_{\alpha,p}$. The following representation of the Meixner polynomials comes from the interpretation of $NB_{\alpha,p}$ as a gamma mixture of Poisson likelihood (formula (2.10)).

Proposition 6.1. *For $\alpha \in \mathbb{R}_+$ and $p \in (0,1)$, a system of orthogonal polynomials with the negative binomial (α, p) distribution as weight measure is given by*

$$\widetilde{M}_n^{\alpha,p}(k) = \int_0^\infty \frac{Po_\lambda(k)}{NB_{\alpha,p}(k)} L_n^\alpha\left(\lambda\frac{1-p}{p}\right)\gamma_{\alpha,p/(1-p)}(\mathrm{d}\lambda) \tag{6.2}$$

$$= \int_0^\infty L_n^\alpha\left(\lambda\frac{1-p}{p}\right)\gamma_{\alpha+k,p}(\mathrm{d}\lambda), \qquad n = 0, 1, \ldots, \tag{6.3}$$

where L_n^α are Laguerre polynomials with parameter α.

Proof. For every n, consider that

$$\int_0^\infty \lambda^n \gamma_{\alpha+k,p}(\mathrm{d}\lambda) = \int_0^\infty \frac{\lambda^{\alpha+k+n-1}\mathrm{e}^{-\lambda/p}}{\Gamma(\alpha+k)p^{\alpha+k}}\,\mathrm{d}\lambda = (\alpha+k)_{(n)}p^n.$$

So every polynomial in Λ of order n is mapped to a polynomial in k of the same order.

To show orthogonality it is, again, sufficient to consider polynomials in the basis $\{r_{[k]} : k = 0, 1, \ldots\}$. Let $m \le n$.

$$\begin{aligned}
&\sum_{k=0}^\infty NB_{\alpha,p}(k)k_{[m]}\widetilde{M}_n^{\alpha,p}(k) \\
&\quad = \int_0^\infty L_n^\alpha\left(\lambda\frac{1-p}{p}\right)\left\{\sum_{k=0}^\infty \frac{(\alpha)_{(k)}}{k!}p^k(1-p)^\alpha k_{[m]}\frac{\lambda^{\alpha+k-1}\mathrm{e}^{-\lambda/p}}{\Gamma(\alpha+k)p^{\alpha+k}}\right\}\mathrm{d}\lambda \\
&\quad = \int_0^\infty L_n^\alpha\left(\lambda\frac{1-p}{p}\right)\left\{\sum_{k=0}^\infty k_{[m]}Po_\lambda(k)\right\}\gamma_{\alpha,p/(1-p)}(\mathrm{d}\lambda) \\
&\quad = \int_0^\infty L_n^\alpha\left(\lambda\frac{1-p}{p}\right)\lambda^m\gamma_{\alpha,p/(1-p)}(\mathrm{d}\lambda),
\end{aligned} \tag{6.4}$$

where the last line comes from the fact that, if K is a Poisson(λ) random variable, then

$$\mathbb{E}_\lambda(K_{[n]}) = \lambda^n, \qquad n = 0, 1, 2, \ldots.$$

Now, consider the change of measure induced by

$$z := \lambda\frac{1-p}{p}.$$

The last line of (6.4) reads

$$\left(\frac{p}{1-p}\right)^m \int_0^\infty L_n^\alpha(z) z^m \gamma_{\alpha,1}(\mathrm{d}z).$$

The integral vanishes for every $m < n$, and therefore the orthogonality is proved. □

From property (2) of the negative binomial distribution (Section 2.3.2), by using Propositions 6.1, 5.2 and 4.3, and Remark 4.6, it is possible to find the following alternative systems of multivariate Meixner polynomials, orthogonal with respect to $NB_{\alpha,p}^d(r)$.

Proposition 6.2. *Let $\alpha \in \mathbb{R}_+^d$ and $p \in (0,1)$.*

(i) *Two systems of multivariate orthogonal polynomials with weight measure $NB_{\alpha,p}^d(r)$ are:*

$$\widetilde{M}_n^{\alpha,p}(r) = \prod_{i=1}^d \tilde{M}_{n_i}^{\alpha_i,p}(r_i), \qquad n \in \mathbb{N}^d, \tag{6.5}$$

and

$$^*\widetilde{M}_n^{\alpha,p}(r) = (1-p)^{|n'|} \tilde{M}_{n_d}^{|\alpha|+2|n'|,p}(|r|-|n'|)(|\alpha+r|)_{(|n'|)}\, \tilde{q}_{n'}^\alpha(r;|r|), \qquad n \in \mathbb{N}^d, \tag{6.6}$$

where $n' = (n_1, \ldots, n_d - 1)$, $\{M_{n_i}^{\alpha_i,p}\}$ are Meixner polynomials as in Proposition 6.1 and $\tilde{q}_\alpha$ are multivariate Hahn polynomials defined by Proposition 5.2.

(ii) *A representation for these polynomials is:*

$$\begin{aligned}\widetilde{M}_n^{\alpha,p}(r) &= \int_{\mathbb{R}_+^d} \frac{Po_\lambda^d(r)}{NB_{\alpha,p}^d(r)} L_n^\alpha\left(\lambda\frac{1-p}{p}\right)\gamma_{\alpha,p/(1-p)}^d(\mathrm{d}\lambda) &(6.7)\\ &= \int_{\mathbb{R}_+^d} L_n^\alpha\left(\lambda\frac{1-p}{p}\right)\gamma_{\alpha+r,p}^d(\mathrm{d}\lambda) &(6.8)\end{aligned}$$

and

$$\begin{aligned}^*\widetilde{M}_n^{\alpha,p}(r) &= \int_{\mathbb{R}_+^d} \frac{Po_\lambda^d(r)}{NB_{\alpha,p}^d(r)} L_n^{\alpha *}\left(\lambda\frac{1-p}{p}\right)\gamma_{\alpha,p/(1-p)}^d(\mathrm{d}\lambda) &(6.9)\\ &= \int_{\mathbb{R}_+^d} L_n^{\alpha *}\left(\lambda\frac{1-p}{p}\right)\gamma_{\alpha+r,p}^d(\mathrm{d}\lambda), &(6.10)\end{aligned}$$

where $\{L_n^\alpha\}$ and $\{L_n^{\alpha *}\}$ are given by (4.3) and (4.5), and

$$\gamma_{\alpha,\beta}^d(\mathrm{d}z) := \prod_{i=1}^d \gamma_{\alpha_i,\beta}(\mathrm{d}z_i), \qquad \beta \in \mathbb{R}, z \in \mathbb{R}^d.$$

(iii) *The connection coefficients between* $\{\widetilde{M}_n^{\alpha,p}\}$ *and* ${}^*\widetilde{M}_n^{\alpha,p}$ *are given by*

$$\mathbb{E}[{}^*\widetilde{M}_n^{\alpha,p}(R)\widetilde{M}_m^{\alpha,p}(R)] = c_m^*(n), \tag{6.11}$$

where $c_m^*(n)$ *are as in* (4.9) *or* (5.23).

Proof. (6.5) is trivial and (6.7) and (6.8) follow from (6.2) and (6.3).

Now let us first prove (6.9) and (6.10). For every $z \in \mathbb{R}_+^d$, denote $x = z/|z|$. Consider that

$$\gamma_{\alpha,\beta}(\mathrm{d}z) = \gamma_{|\alpha|,\beta}(\mathrm{d}|z|)D_\alpha(\mathrm{d}x)$$

and that

$$Po_z^d(r) = Po_{|z|}(|r|)L_x(r).$$

Combining this with (2.11),

$$\begin{aligned}
&\int_{\mathbb{R}_+^d} \frac{Po_\lambda^d(r)}{NB_{\alpha,p}^d(r)} L_n^{\alpha*}\left(\lambda\frac{1-p}{p}\right)\gamma^d_{\alpha,p/(1-p)}(\mathrm{d}\lambda) \\
&\quad = \left(\int_{\mathbb{R}_+} \frac{Po_{|\lambda|}(|r|)}{NB_{|\alpha|,p}(|r|)} L_{n_d}^{|\alpha|+2|n'|}\left(|\lambda|\frac{1-p}{p}\right)\left[|\lambda|\frac{1-p}{p}\right]^{|n'|}\gamma_{|\alpha|,p/(1-p)}(\mathrm{d}|\lambda|)\right) \\
&\qquad \times \left(\int_{\Delta_{(d-1)}} \frac{L_x(r)}{DM_\alpha(r,|r|)} R_{n'}^\alpha(x)D_\alpha(\mathrm{d}x)\right).
\end{aligned} \tag{6.12}$$

From Proposition 5.2, the last integral in (6.12) is equal to $\tilde{q}_{n'}^\alpha(r;|r|)$.

The first integral can be rewritten as

$$\begin{aligned}
&\int_{\mathbb{R}_+} L_{n_d}^{|\alpha|+2|n'|}\left(|\lambda|\frac{1-p}{p}\right)\left[|\lambda|\frac{1-p}{p}\right]^{|n'|}\gamma_{|\alpha|+|r|,p/(1-p)}(\mathrm{d}|\lambda|) \\
&\quad = (1-p)^{|n'|}(|\alpha+r|)_{(|n'|)}\int_{\mathbb{R}_+} L_{n_d}^{|\alpha|+2|n'|}\left(|\lambda|\frac{1-p}{p}\right)\frac{|\lambda|^{|\alpha+r+n'|}\mathrm{e}^{-|\lambda|/p}}{\Gamma(|\alpha+r+n'|)p^{|\alpha+r+n'|}}\,\mathrm{d}|\lambda| \\
&\quad = (1-p)^{|n'|}(|\alpha+r|)_{(|n'|)}\widetilde{M}_{n_d}^{\alpha+2|n'|}(|r|-|n'|).
\end{aligned} \tag{6.13}$$

The last line in (6.13) is obtained from (6.3) by rewriting $|n'| = 2|n'| - |n'|$ in the mixing measure. Thus the identities (6.9) and (6.10) are proved.

To prove part (iii), simply use (4.7) with coefficients given by Proposition 4.3 to see that (6.7) and (6.8) and (6.9) and (6.10) imply

$$\begin{aligned}
{}^*\widetilde{M}_n^{\alpha,p}(r) &= \mathbb{E}_{\alpha+r,p}\left[L_n^{\alpha*}\left(\lambda\frac{1-p}{p}\right)\right] = \mathbb{E}_{\alpha+r,p}\left[\sum_{|m|=|n|} c_m^*(n)L_m^\alpha\left(\lambda\frac{1-p}{p}\right)\right] \\
&= \sum_{|m|=|n|} c_m^*(n)\mathbb{E}_{\alpha+r,p}\left[L_m^\alpha\left(\lambda\frac{1-p}{p}\right)\right] = \sum_{|m|=|n|} c_m^*(n)\widetilde{M}_m^{\alpha,p}(r).
\end{aligned}$$

This is equivalent to (6.11) because of the orthogonality of $\widetilde{M}_m^{\alpha,p}(R)$.

But (6.11) also implies that $\{{}^*\widetilde{M}_n^{\alpha,p}(r)\}$ is an orthogonal system with $NB^d_{\alpha,p}$ as weight measure since, for every polynomial $r_{[l]}$ of degree $|l| \le |n|$,

$$\sum_{r\in\mathbb{N}^d} NB^d_{\alpha,p}(r)^*\widetilde{M}_n^{\alpha,p}(r)r_{[l]} = \sum_{|m|=|n|} c_m^*(n)\left(\sum_{r\in\mathbb{N}^d} NB^d_{\alpha,p}(r)\widetilde{M}_m^{\alpha,p}(r)r_{[l]}\right).$$

The term between brackets is non-zero only for $|l| = |m| = |n|$, which implies orthogonality, so the proof of the proposition is now complete. □

6.1. The Bernstein–Bézier coefficients of the multiple Laguerre polynomials

The representation of Meixner polynomials given in Proposition 6.2 leads us, not surprisingly, to interpret these as the Bernstein–Bézier coefficients of the multiple Laguerre polynomials (for any choice of basis), up to proportionality constants. Note that, for products of Poisson distributions we can write

$$Po^d_\lambda(r) = \prod_{i=1}^d \frac{e^{-\lambda_i}\lambda_i^{r_i}}{r_i!} = \frac{e^{-|\lambda|}}{|\lambda|!}B_\lambda(r). \tag{6.14}$$

To simplify the notation, let (L_m, M_n) denote either $(L_m^\alpha, \widetilde{M}_m^{\alpha,p})$ or $(L_m^{\alpha*}, {}^*\widetilde{M}_m^{\alpha,p})$, for some $\alpha \in \mathbb{R}^d$ and $p \in (0,1)$. Let φ_n be either as in (4.4) or as in (4.6), consistently with the choice of L_n, and set $\rho_r(\alpha,p)^{-1} := E[M_r^2]$.

Corollary 6.3.

$$L_r\left(\lambda\frac{1-p}{p}\right) = \frac{\rho_r(\alpha,p)}{\varphi_r}\frac{e^{-|\lambda|}}{|\lambda|!}\sum_m M_r(m)B_\lambda(m). \tag{6.15}$$

Proof. The proof is along the same lines as for Corollary 5.4. From (6.7)–(6.9),

$$\mathbb{E}\left[L_n\left(Y\frac{1-p}{p}\right)Po^d_Y(m)\right] = M_n(m)NB^d_{\alpha,p}(m), \qquad n, m \in \mathbb{N}^d.$$

Then from (6.14),

$$B_\lambda(m) = |\lambda|!e^{|\lambda|}NB^d_{\alpha,p}(m)\sum_n \varphi_n M_n(m)L_n\left(Y\frac{1-p}{p}\right).$$

So for every $r \in \mathbb{N}^d$

$$\begin{aligned}\sum_m M_r(m) B_\lambda(m) &= |\lambda|!\mathrm{e}^{|\lambda|} \sum_n \varphi_n \left[\sum_m NB^d_{\alpha,p}(m) M_n(m) M_r(m)\right] L_n\left(Y\frac{1-p}{p}\right) \\ &= |\lambda|!\mathrm{e}^{|\lambda|} \sum_n L_n\left(Y\frac{1-p}{p}\right)\frac{\varphi_n}{\rho_r(\alpha,p)}\delta_{nr} \\ &= \frac{|\lambda|!\mathrm{e}^{|\lambda|}\varphi_r}{\rho_r(\alpha,p)} L_r\left(Y\frac{1-p}{p}\right),\end{aligned}$$

and the proof is complete. □

Acknowledgement

Dario Spanò's research is partly supported by CRiSM, an EPSRC-funded grant.

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**. Washington, DC: US Government Printing Office. 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] Diaconis, P., Khare, K. and Saloff-Coste, L. (2008). Gibbs sampling, exponential families and orthogonal polynomials. *Statist. Sci.* **23** 151–178. MR2446500

[4] Doksum, K. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probab.* **2** 183–201. MR0373081

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

[7] Exton, H. (1976). *Multiple Hypergeometric Functions and Applications*. Chichester: Ellis Horwood. MR0422713

[8] Griffiths, R.C. and Spanò, D. (2010). Diffusion processes and coalescent trees. In *Probability and Mathematical Genetics: Papers in Honour of Sir John Kingman* (N.H. Bingham and C.M. Goldie, eds.), Chapter 15, 358–380. *London Mathematical Society Lecture Notes Series*. Cambridge: Cambridge Univ. Press.

[9] Griffiths, R. and Spanò, D. (2010). n-kernel orthogonal polynomials on the Dirichlet, Dirichlet-multinomial, Poisson–Dirichlet and Ewens' Sampling Distributions, and positive-definite sequences. Available at http://arxiv.org/abs/1003.5131v1.

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

[11] 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

[12] Karlin, S. and McGregor, J.L. (1961). The Hahn polynomials, formulas and an application. *Scripta Math.* **26** 33–46. MR0138806

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

[14] 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 distributions. *J. Roy. Statist. Soc. Ser. B* **37** 1–22. MR0368264

[15] Koornwinder, T. (1975). Two-variable analogues of the classical orthogonal polynomials. In *Theory and Application of Special Functions (Proc. Advanced Sem., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1975)* 435–495. New York: Academic Press. MR0402146

[16] 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

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

[18] Pitman, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields* **102** 145–158. MR1337249

[19] Pitman, J. (1996). Random discrete distributions invariant under size-biased permutation. *Adv. in Appl. Probab.* **28** 525–539. MR1387889

[20] Pitman, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Mathematics* **1875**. Berlin: Springer. MR2245368

[21] Sauer, T. (2007). Jacobi polynomials in Bernstein form. *J. Comput. Appl. Math.* **199** 149–158. MR2267539

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

[23] Walker, S. and Muliere, P. (1997). Beta-Stacy processes and a generalization of the Pólya-urn scheme. *Ann. Statist.* **25** 1762–1780. MR1463574

Received November 2009 and revised July 2010