Accepted Manuscript

Exchangeable Hoeffding decompositions over finite sets: A combinatorial characterization and counterexamples

Omar El-Dakkak, Giovanni Peccati, Igor Prünster

Please cite this article as: O. El-Dakkak, G. Peccati, I. Prünster, Exchangeable Hoeffding decompositions over finite sets: A combinatorial characterization and counterexamples, *Journal of Multivariate Analysis* (2014), http://dx.doi.org/10.1016/j.jmva.2014.04.012

Exchangeable Hoeffding decompositions over finite sets: a combinatorial characterization and counterexamples

Omar El-Dakkak[a], Giovanni Peccati[b], Igor Prünster[c]

[a]*Equipe MODAL'X, Université Paris-Ouest, Nanterre La Défense, France*
[b]*Mathematics Research Unit, University of Luxembourg, Luxembourg*
[c]*Department of Economics and Statistics and Collegio Carlo Alberto, University of Torino, Italy*

Abstract

We study Hoeffding decomposable exchangeable sequences with values in a finite set $D = \{d_1, \ldots, d_K\}$. We provide a new combinatorial characterization of Hoeffding decomposability and use this result to show that, for every $K \geq 3$, there exists a class of neither Pólya nor i.i.d. D-valued exchangeable sequences that are Hoeffding decomposable.

Keywords: Exchangeability, Hoeffding decomposition, Pólya urn, urn sequence, weak independence
2000 MSC: 60G09, 60G99

1. Introduction and framework

1.1. Overview

Let $\mathbf{X}_{[1,\infty)} := \{X_n : n \geq 1\}$ be a sequence of random variables (r.v.) with values in some Polish space. We say that $\mathbf{X}_{[1,\infty)}$ is *Hoeffding-decomposable* if every square-integrable symmetric statistic of any n-subvector of $\mathbf{X}_{[1,\infty)}$, for every $n \geq 2$, can be uniquely represented as an orthogonal sum of n U-statistics with degenerate kernels of increasing order. The classic notion of 'degeneracy' that is needed in this context is formally introduced in formula (1.8) below.

Since their discovery in the landmark paper by Hoeffding [12], Hoeffding decompositions in the case of i.i.d. sequences have been successfully applied in a variety of frameworks, e.g.: linear rank statistics (Hajek [11]), jack-knife estimators (Karlin and Rinott [15]), covariance analysis of symmetric statistics (Vitale [24]), convergence of U-processes (Arcones and Giné [2]),

asymptotic problems in geometric probability (Avram and Bertsimas [3]), Edgeworth expansions (Bentkus, Götze and van Zwet [4]), and tail estimates for U-statistics (Major [18]). See also Koroljuk and Borovskich [16] and references therein.

Outside the i.i.d. framework, Hoeffding decompositions have been notably applied to study sampling without replacement from finite populations. The first analysis in this direction can be found in Zhao and Chen [25]. Bloznelis and Götze [7, 8] generalized these results in order to characterize the asymptotic normality of symmetric statistics based on sampling without replacement (when the size of the population diverges to infinity), as well as to obtain explicit Edgeworth expansions. In Bloznelis [6], Hoeffding-type decompositions are explicitly computed for statistics depending on extractions without replacement from several distinct populations.

In Peccati [19, 20, 21] the theory of Hoeffding decompositions was extended to the framework of general exchangeable (infinitely extendible) random sequences. In Peccati [20] it was shown that the class of Hoeffding decomposable exchangeable sequences coincides with the collection of weakly independent sequences, and that the class of weakly independent (and, therefore, Hoeffding decomposable) sequences contains the family of generalized Pólya urn sequences (see, e.g., Blackwell and MacQueen [5] or Pitman [22]). The connection with Pólya urns was further exploited in Peccati [21], where Hoeffding-type decompositions were used in order to establish several new spectral properties of Ferguson-Dirichlet processes (Ferguson [10]), such as for instance a chaotic representation property.

In El-Dakkak and Peccati [9] the results established in Peccati [20] were enriched and completed in two directions. On the one hand, it was proved that a (non deterministic) infinite exchangeable sequence with values in $\{0, 1\}$ is Hoeffding decomposable if and only if it is either a Pólya sequence or i.i.d.. This result connects *de facto* the seemingly unrelated notions of a Hoeffding decomposable sequence and of an *urn process*, a concept thoroughly studied in Hill, Lane and Sudderth [13]. For the sake of completeness, it is worth recalling that, following [13], an exchangeable sequence $\mathbf{X}_{[1,\infty)}$ will be termed *deterministic* if $\mathbb{P}[X_k = X_1, \forall k \geq 2] = 1$. On the other hand, and using different techniques, a partial characterization of Hoeffding decomposable exchangeable sequences with values in a finite set with more than two elements was obtained. While not being as exhaustive as the one in the two-color case, this characterization was used to prove that Pólya urns are the only Hoeffding decomposable sequences within a large class of exchange-

able sequences. Such a family of exchangeable sequences is defined in terms of their directing (or de Finetti) measure, which is obtained by normalizing vectors of infinitely divisible (positive) independent r.v. (see Regazzini, Lijoi and Prünster [23] and James, Lijoi and Prünster [14]). See Lijoi and Prünster [17] for an overview of their use in Bayesian Statistics.

Therefore, the analysis carried out in El-Dakkak and Peccati [9] left the following question unanswered:

> **Problem A:** Are Pólya and i.i.d. sequences the only infinite non deterministic Hoeffding decomposable sequences with values in a finite set with ≥ 3 elements?

We shall give a negative answer to Problem A. This is surprising given the above mentioned positive characterization might somehow lead to conjecture the opposite and, hence, makes the present result even more remarkable. In fact, the negative answer is obtained by explicitly building a class of *neither Pólya nor i.i.d.* yet Hoeffding decomposable exchangeable sequences with values in a finite set with strictly more than two elements. A precise statement is given in Theorem 2.3. Interestingly, this class turns out to be a generalization of a counterexample appearing in Hill, Lane and Sudderth [13, p.1591], showing that, unlike in the two-color case (see Theorem 1.5), there exist non-deterministic exchangeable 3-color urn processes that are neither Pólya nor i.i.d. sequences. Our fundamental tool is a new combinatorial characterization, stated in Theorem 3.4, of the system of predictive probabilities associated with Hoeffding-decomposable exchangeable sequences taking values in an arbitrary finite set. This characterization, which is of independent interest, represents a generalization of a crucial combinatorial statement proved in El-Dakkak and Peccati [9, Proposition 4].

In what follows, we recall some well-known facts concerning key notions and concepts to be used in the sequel: exchangeability (Section 1.2), Hoeffding decomposability (Section 1.3), weak independence (Section 1.4) and urn processes (Section 1.5). Comprehensive accounts on these can be found in, e.g., [1, 20, 9, 13].

Remark 1.1. Every exchangeable sequence $\{X_1, X_2, ...\}$ considered in this paper is assumed to take values in some finite set D. In particular, every r.v. of the type $F = \rho(X_1, ..., X_n)$, with $n \geq 1$ and ρ any real-valued function on D^n, is automatically bounded.

3

1.2. Exchangeability

For every $n \geq 2$, we denote by $\mathfrak{S}_n$ the group of all permutations of the set $[n] = \{1, ..., n\}$. A vector $(X_1, ..., X_n)$ of D-valued r.v. is said to be *exchangeable* if, for all $\mathbf{x}_n = (x_1, ..., x_n) \in D^n$ and all $\pi \in \mathfrak{S}_n$,

$$\mathbb{P}(X_1 = x_1, ..., X_n = x_n) = \mathbb{P}\left(X_1 = x_{\pi(1)}, ..., X_n = x_{\pi(n)}\right).$$

A D-valued infinite sequence $\mathbf{X}_{[1,\infty)}$ is exchangeable if every n-subvector of $\mathbf{X}_{[1,\infty)}$ is exchangeable. Let Π_D denote the set of all probability measures on D. By the de Finetti representation theorem (see, e.g, [1]), an infinite sequence $\mathbf{X}_{[1,\infty)}$ with values in $D = \{d_1, ..., d_K\}$ is exchangeable if and only if there exists a unique probability measure γ on Π_D (called *directing* or the *de Finetti measure* associated with the sequence $\mathbf{X}_{[1,\infty)}$) such that, for all $n \geq 2$ and all $(x_1, ..., x_n) \in D^n$,

$$\mathbb{P}(X_1 = x_1, ..., X_n = x_n) = \int_{\Pi_D} \prod_{j=1}^{n} p\{x_j\} \gamma(\mathrm{d}p), \qquad (1.1)$$

where the elements of Π_D are written in the form $p := \{p\{d_i\} : i = 1, ..., K\}$. In other words, the de Finetti representation theorem states that a sequence of r.v. is exchangeable if and only if it is a mixture of i.i.d. random sequences with values in D.

Any probability measure $p \in \Pi_D$ can be parameterized in terms of the simplex

$$\Theta_{K-1} := \left\{ (\theta_1, ..., \theta_{K-1}) : \theta_h \geq 0, \ h = 1, ..., K-1 \text{ and } \sum_{h=1}^{K-1} \theta_h \leq 1 \right\},$$

by setting $p\{d_1\} = \theta_1, ..., p\{d_{K-1}\} = \theta_{K-1}$ and $p\{d_K\} = 1 - \sum_{h=1}^{K-1} \theta_h$. Thus, the representation in (1.1) can be rewritten as

$$\mathbb{P}(X_1 = x_1, ..., X_n = x_n) = \int_{\Theta_{K-1}} \left(\Pi_{j=1}^{K-1} \theta_j^{i_j} \right) \left(1 - \Sigma_{h=1}^{K-1} \theta_h \right)^{i_K} \gamma(\mathrm{d}\theta_1, ..., \mathrm{d}\theta_{K-1}), \qquad (1.2)$$

where (with an abuse of notation) we have identified γ with its image through the canonical bijection between Π_D and Θ_{K-1}, and $i_j := \sum_{v=1}^{n} \mathbf{1}(x_v = d_j)$, $j = 1, ..., K$. Clearly, when $K = 2$, (1.2) becomes

$$\mathbb{P}(X_1 = x_1, ..., X_n = x_n) = \int_{[0,1]} \theta^i (1 - \theta)^{n-i} \gamma(\mathrm{d}\theta) \qquad (1.3)$$

4

with $i = \sum_{v=1}^{n} \mathbf{1}\, (x_v = d_1)\,.$

If there exists a vector $\alpha = (\alpha_1, ..., \alpha_K) \in (0, +\infty)^K$ of strictly positive numbers such that

$$\gamma(\mathrm{d}\theta_1, \ldots, \mathrm{d}\theta_{K-1}) = \frac{1}{B\,(\alpha)} \left(\Pi_{j=1}^{K-1} \theta_j^{\alpha_j - 1} \right) \left(1 - \Sigma_{h=1}^{K-1} \theta_h \right)^{\alpha_K - 1} \mathrm{d}\theta_1 \cdots\ \mathrm{d}\theta_{K-1}, \tag{1.4}$$

where $B\,(\alpha) := \Pi_{j=1}^{K} \Gamma\,(\alpha_j) / \Gamma\left(\Sigma_{j=1}^{K} \alpha_j \right),$ and $\Gamma\,(\cdot)$ stands for the gamma function, we say that γ is a *Dirichlet probability measure* and that $\mathbf{X}_{[1,\infty)}$ is a *K-color Pólya sequence* with parameter α. Specializing (1.4) to the case $K = 2$, one sees immediately that the measure γ in (1.3) becomes a beta distribution with parameters α_1, α_2. It follows that an exchangeable sequence $\mathbf{X}_{[1,\infty)}$ is a two-color Pólya sequence if and only if its de Finetti measure is a beta distribution.

1.3. Hoeffding decomposability

Let us first introduce some notation. For all $n \geq 1$ and all $1 \leq u \leq n$, we write $[n] = \{1, ..., n\}$ and $[u, n] = \{u, u + 1, ..., n\}$, and set $\mathbf{X}_{[n]} := (X_1, X_2, ..., X_n)$ and $\mathbf{X}_{[u,n]} := (X_u, X_{u+1}, ..., X_n)$. As in [9] define, for all $n \geq 2$, the sequence of spaces

$$\left\{ SU_k\left(\mathbf{X}_{[n]} \right) : k = 0, ..., n \right\},$$

generated by symmetric U-statistics of increasing order, as follows: $SU_0\left(\mathbf{X}_{[n]} \right) := \mathfrak{R}$ and, for all $k = 1, ..., n$, $SU_k\left(\mathbf{X}_{[n]} \right)$ is the collection of all r.v. of the type

$$F\left(\mathbf{X}_{[n]} \right) = \sum_{1 \leq j_1 < \cdots < j_k \leq n} \varphi\, (X_{j_1}, ..., X_{j_k}), \tag{1.5}$$

where φ is a real-valued symmetric function on D^k. Any r.v. F as in (1.5) is called a *U-statistic with symmetric kernel of order k*. The collection $\left\{ SU_k\left(\mathbf{X}_{[n]} \right) \right\}$ is an increasing sequence of vector spaces such that $SU_n\left(\mathbf{X}_{[n]} \right) = L_s\left(\mathbf{X}_{[n]} \right)$, where $L_s\left(\mathbf{X}_{[n]} \right)$ is defined as the Hilbert space of all symmetric statistics $T\left(\mathbf{X}_{[n]} \right)$ w.r.t. the inner product $\langle T_1, T_2 \rangle := \mathbb{E}\left(T_1\left(\mathbf{X}_{[n]} \right) T_2\left(\mathbf{X}_{[n]} \right) \right).$ Hence, one can meaningfully define the sequence of *symmetric Hoeffding spaces* associated with $\mathbf{X}_{[n]}$, denoted by

$$\left\{ SH_k\left(\mathbf{X}_{[n]} \right) : k = 0, ..., n \right\},$$

5

as follows: $SH_0\left(\mathbf{X}_{[n]}\right) := SU_0\left(\mathbf{X}_{[n]}\right) = \mathfrak{R}$, and

$$SH_k\left(\mathbf{X}_{[n]}\right) := SU_k\left(\mathbf{X}_{[n]}\right) \cap SU_{k-1}\left(\mathbf{X}_{[n]}\right)^{\perp}, \qquad k = 1, ..., n,$$

where all orthogonals (here and in the sequel) are taken in $L_s\left(\mathbf{X}_{[n]}\right)$. The following representation is therefore at hand

$$L_s\left(\mathbf{X}_{[n]}\right) = \bigoplus_{k=0}^{n} SH_k\left(\mathbf{X}_{[n]}\right),$$

where '$\oplus$" stands for an orthogonal sum. This fact implies that, for all $n \geq 2$, every symmetric statistic $T\left(\mathbf{X}_{[n]}\right)$ admits an almost-surely unique representation of the type:

$$T\left(\mathbf{X}_{[n]}\right) = \mathbb{E}\left(T\right) + \sum_{k=1}^{n} F_k\left(\mathbf{X}_{[n]}\right), \tag{1.6}$$

where the F_k's are uncorrelated U-statistics such that, for all $k = 1, ..., n$, there exists a symmetric kernel φ_k of order k satisfying

$$F_k\left(\mathbf{X}_{[n]}\right) = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \varphi_k\left(X_{i_1}, ..., X_{i_k}\right). \tag{1.7}$$

We are now in a position to recall the definition of Hoeffding decomposability for exchangeable sequences given in Peccati [20].

Definition 1.2. The exchangeable sequence $\mathbf{X}_{[1,\infty)}$ is said to be *Hoeffding decomposable* if, for all $n \geq 2$ and all $k = 1, ..., n$, the following double implication holds: $F_k \in SH_k\left(\mathbf{X}_{[n]}\right)$ if and only if the kernel φ_k appearing in its representation (1.6)–(1.7) is *completely degenerate*, that is

$$\mathbb{E}\left(\varphi_k\left(\mathbf{X}_{[k]}\right) \mid \mathbf{X}_{[2,k]}\right) = 0, \qquad a.s.\text{-}\mathbb{P}. \tag{1.8}$$

1.4. Weak independence

Fix $n \geq 2$ and let $\mathcal{S}\left(D^n\right)$ be the class of all symmetric real-valued functions on D^n. Fix $\varphi \in \mathcal{S}\left(D^n\right)$ and define the functions $[\varphi]_{n,n-1}^{(n-1)} : D^{n-1} \to \mathfrak{R}$ and $[\varphi]_{n,n-1}^{(n-u)} : D^{n-1} \to \mathfrak{R}$, $u = 2, ..., n$ as the unique mappings such that

$$[\varphi]_{n,n-1}^{(n-1)}\left(\mathbf{X}_{[2,n]}\right) = E\left(\varphi\left(\mathbf{X}_{[n]}\right) \mid \mathbf{X}_{[2,n]}\right), \qquad a.s.\text{-}\mathbb{P}, \tag{1.9}$$

6

and

$$[\varphi]_{n,n-1}^{(n-u)}\left(\mathbf{X}_{[u+1,u+n-1]}\right) = E\left(\varphi\left(\mathbf{X}_{[n]}\right) \mid \mathbf{X}_{[u+1,u+n-1]}\right), \qquad a.s.\text{-}\mathbb{P}, \quad (1.10)$$

respectively. Exchangeability and symmetry imply that $D^{n-1} \to \mathfrak{R} : \mathbf{x} \mapsto [\varphi]_{n,n-1}^{(n-1)}(\mathbf{x})$ and $D^{n-1} \to \mathfrak{R} : \mathbf{x} \mapsto [\varphi]_{n,n-1}^{(0)}(\mathbf{x})$ (the latter being the function appearing in (1.10) written for $u = n$) are symmetric functions whereas, for $u = 2, ..., n-1$, the function $D^{n-1} \ni (x_1, ..., x_{n-1}) \mapsto [\varphi]_{n,n-1}^{(n-u)}(x_1, ..., x_{n-1})$ is separately symmetric in the variables $(x_1, ..., x_{n-u})$ and $(x_{n-u+1}, ..., x_{n-1})$, and not necessarily as a function on D^{n-1}. Recall that, given a function $f : D^n \to \mathfrak{R}$, the canonical symmetrization of f, denoted by $\tilde{f}$, is given by

$$\tilde{f}(\mathbf{x}_n) = \frac{1}{n!} \sum_{\pi \in \mathfrak{S}_n} f\left(\mathbf{x}_{\pi(n)}\right), \ \mathbf{x}_n \in D^n.$$

Now define the sequence of vector spaces

$$\Xi_n\left(\mathbf{X}_{[1,\infty)}\right) := \left\{\varphi \in \mathcal{S}(D^n) : [\varphi]_{n,n-1}^{(n-1)}\left(\mathbf{X}_{[2,n]}\right) = 0\right\}, \qquad n \geq 2,$$

and the array of spaces

$$\tilde{\Xi}_{n,n-u}\left(\mathbf{X}_{[1,\infty)}\right) := \left\{\varphi \in \mathcal{S}(D^n) : \widetilde{[\varphi]}_{n,n-1}^{(n-u)}\left(\mathbf{X}_{[u+1,u+n-1]}\right) = 0\right\}, \ u = 2, ..., n, \ n \geq 2,$$

where $\widetilde{[\varphi]}_{n,n-1}^{(n-u)}$ is the canonical symmetrization of $[\varphi]_{n,n-1}^{(n-u)}$.

Definition 1.3. We say that the sequence $\mathbf{X}_{[1,\infty)}$ is *weakly independent* if, for every $n \geq 2$,

$$\Xi_n\left(\mathbf{X}_{[1,\infty)}\right) \subset \bigcap_{u=2}^{n} \tilde{\Xi}_{n,n-u}\left(\mathbf{X}_{[1,\infty)}\right).$$

In other words, weak independence occurs if, for every $n \geq 2$ and every $\varphi \in \mathcal{S}(D^n)$, the relation $[\varphi]_{n,n-1}^{(n-1)}\left(\mathbf{X}_{[2,n]}\right) = 0$ necessarily implies that $\widetilde{[\varphi]}_{n,n-1}^{(n-u)}\left(\mathbf{X}_{[u+1,u+n-1]}\right) = 0$ for all $u = 2, ..., n$. For instance, when $n = 2$ weak independence yields the following implication for every symmetric φ on D^2:

$$\mathbb{E}\left(\varphi(X_1, X_2) \mid X_2\right) = 0 \quad \Rightarrow \quad \mathbb{E}\left(\varphi(X_1, X_2) \mid X_3\right) = 0.$$

Parts I and II of the following statement contain some of the main findings in Peccati [20] and El-Dakkak and Peccati [9], respectively.

7

Theorem 1.4 (Peccati [20]; El-Dakkak and Peccati [9]). *Let* $\mathbf{X}_{[1,\infty)}$ *be an exchangeable sequence of r.v. with values in the finite set D.*

(I) *Assume that*

$$SH_k\left(\mathbf{X}_{[n]}\right) \neq \{0\}, \qquad \forall k = 1,...,n, \ \forall n \geq 2. \qquad (1.11)$$

Then $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable if and only if it is weakly independent.

(II) *Assume that $D = [2]$ and that $\mathbf{X}_{[1,\infty)}$ is non deterministic (so that (1.11) is automatically satisfied). Then, $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable if and only if $\mathbf{X}_{[1,\infty)}$ is either a Pólya sequence or an i.i.d. sequence.*

1.5. Urn processes and a result by Hill, Lane and Sudderth

Let $\mathbf{X}_{[1,\infty)} := \{X_n : n \geq 1\}$ be a sequence of $\{0,1\}$-valued r.v.. Roughly speaking, $\mathbf{X}_{[1,\infty)}$ is a *two-color urn process* if its probabilistic structure can be represented by successive drawings from an urn with changing composition. More precisely, consider an urn containing r red balls and b black balls, $r, b \in \{1, 2, ...\}$, and let $Y_0 := r/(r+b)$ denote the initial proportion of red balls in the urn. Suppose that a red ball is added with probability $f(Y_0)$ and that a black ball is added with probability $1 - f(Y_0)$, where f denotes a function from the unit interval into itself, and let Y_1 be the new proportion of red balls in the urn. Now, iterate the procedure to generate a sequence $(Y_0, Y_1, Y_2, ...)$. For all $n \geq 1$, let X_n denote the indicator of the event that the n-th ball added is red. The process $\mathbf{X}_{[1,\infty)} := \{X_n : n \geq 1\}$ constructed in this manner is called a *two-color urn process with initial composition (r, b) and urn function f*. It is immediately seen that, for all $n \geq 1$,

$$\mathbb{P}\left(X_{n+1} = 1 \mid X_1, \ldots, X_n\right) = f(Y_n).$$

In other words, two-color urn processes are characterized by the fact that the conditional probability that, at stage $n+1$, a red ball is added depends uniquely on the proportion of red balls at stage n, via the function f.

A two-color urn process is said to be exchangeable if the sequence $\mathbf{X}_{[1,\infty)}$ is exchangeable. In particular, if $\mathbf{X}_{[1,\infty)}$ is a two-color urn process with initial composition (r, b), and the identity map as urn function then (a) $\mathbf{X}_{[1,\infty)}$ is exchangeable and (b) the de Finetti measure of $\mathbf{X}_{[1,\infty)}$ is a beta distribution

8

with parameters r and b. In this case, $\mathbf{X}_{[1,\infty)}$ is termed *two-color Pólya urn process*. Similarly, a two-color urn process, $\mathbf{X}_{[1,\infty)}$, with constant urn function, identically equal to Y_0 is (a) exchangeable and (b) has de Finetti measure equal to a point mass at Y_0. In other words, $\mathbf{X}_{[1,\infty)}$ is a sequence of i.i.d. Bernoulli trials with parameter Y_0. Finally, a two-color urn process, $\mathbf{X}_{[1,\infty)}$, with urn function

$$f(x) = p\mathbf{1}_{\{Y_0\}}(x) + \mathbf{1}_{(Y_0,1]}(x), \qquad p \in (0,1),$$

is (a) exchangeable and (b) has de Finetti measure $\gamma = p\delta_{\{1\}} + (1-p)\delta_{\{0\}}$. In this case, $\mathbf{X}_{[1,\infty)}$ is termed *deterministic urn process*. Note that such processes are characterized by the fact that the support of their de Finetti measure is $\{0\} \cup \{1\}$.

The following statement is the main result of Hill, Lane and Sudderth [13]: it shows that the three classes described above are the only two-color exchangeable urn processes.

Theorem 1.5 (Hill, Lane and Sudderth [13]). *Let $\mathbf{X}_{[1,\infty)}$ be an exchangeable non deterministic urn process with values in $\{0,1\}$. Then, $\mathbf{X}_{[1,\infty)}$ is either a two-color Pólya urn process or an i.i.d. Bernoulli sequence.*

Remark 1.6. In the parlance of the present article, a distinction is made between *Pólya sequences* and *Pólya urn processes*, the latter being a proper subset of the former: in fact, according to our definitions, a Pólya urn process is a Pólya sequence with de Finetti measure given by a beta distribution whose parameters are integer-valued.

We now turn to the definition of multicolor urn processes. Consider an urn containing balls of K colors, $K \in \{3,4,\ldots\}$, and suppose that it contains exactly r_i balls of color d_i, respectively, $r_i \in \{1,2,\ldots\}$, $i = 1,\ldots,K$. Let $Y_0 := (Y_{0,1},\ldots,Y_{0,K})$ be the vector of initial proportions of balls of each color in the urn, where $Y_{0,i} := \frac{r_i}{\sum_{k=1}^{K} r_k}$ denotes the proportion of balls of color d_i, $i = 1,\ldots,K$. Suppose that, at stage 1, a ball is added and that it is of color d_j with probability $f_j(Y_0)$, $j = 1,\ldots,K$, where the f_j's are $[0,1]$-valued functions defined on the simplex

$$S_K := \left\{ \mathbf{y} = (y_1,\ldots,y_K) : \sum_{k=1}^{K} y_k = 1, \ y_k \geq 0, \ k = 1,\ldots,K \right\}, \quad (1.12)$$

9

such that, for all $\mathbf{y} \in S_K$, $\sum_{j=1}^{K} f_j(\mathbf{y}) = 1$. Let $Y_1 := (Y_{1,1}, \ldots, Y_{1,K})$ be the new composition of the urn and iterate the process to generate a sequence $(Y_0, Y_1, Y_2, \ldots)$. For all $n \geq 1$, let X_n be the $\{d_1, \ldots, d_K\}$-valued r.v. such that $X_n = d_j$ if and only if the ball added, at stage n, is of color d_j. The process $\mathbf{X}_{[1,\infty)} := \{X_n : n \geq 1\}$ obtained in this manner is called a K-color urn process with initial composition $(r_1, \ldots, r_K)$ and urn function $f = (f_1, \ldots, f_K)$, and we have, for all $n \geq 1$ and all $j \in \{1, \ldots, K\}$,

$$\mathbb{P}\left(X_{n+1} = d_j \mid X_1, \ldots, X_n\right) = f_j(Y_n).$$

A K-color urn process with initial composition $(r_1, \ldots, r_K)$ and an urn function given by the identity map is (a) exchangeable and (b) has de Finetti measure given by a Dirichlet distribution of parameters $r_1, \ldots, r_K$. Such an urn process is called K-color Pólya urn process with initial composition $(r_1, \ldots, r_K)$. Once more, the class of K-color Pólya urn processes is a proper subset of the class of K-color Pólya sequences. The following example, taken from Hill, Lane and Sudderth [13, p. 1591], shows that a neat result such as Theorem 1.5 cannot hold for exchangeable urn processes with values in sets with strictly more than two elements.

Example 1.7. An urn contains three balls, 1 red, 1 black and 1 green. At each stage, a ball is drawn. If the ball is red, it is replaced and another red ball is added. If the ball is black or green, it is replaced, and a green or black ball is added, depending whether a fair coin falls head or tails. Attaching the labels 1, 2, 3, respectively, to the colors red, black and green, one sees immediately that the sequence $\{X_n : n \geq 1\}$, defined as $X_n = j$ ($j = 1, 2, 3$) according to whether the nth ball added to the urn is of color j, is an exchangeable urn process with urn function given by $f_1(\mathbf{y}) = y_1$ and $f_2(\mathbf{y}) = f_3(\mathbf{y}) = (y_2 + y_3)/2$. In particular, $\{X_n\}$ is not a Pólya urn process.

The main achievement of the present paper is the proof that a generalization of the previous example provides examples of Hoeffding decomposable exchangeable sequences that are neither Pólya nor i.i.d..

1.6. Outline

Section 2 contains a discussion and the statement of our main result: Theorem 2.3. Section 3 contains the main combinatorial tools and the novel combinatorial characterization that are needed throughout the present paper, whereas the proof of Theorem 2.3 is provided in Section 4.

2. A remarkable class of exchangeable sequences

To achieve the announced negative result we introduce a remarkable class of exchangeable sequences. As will be clear from its definition, this class generalizes the exchangeable sequence introduced by Hill, Lane and Sudderth [13] recalled in Example 1.7. Let $K \geq 3$, and let $\mathbf{X}_{[1,\infty)}$ be an exchangeable sequence with values in $D = \{d_1, \ldots, d_K\}$, whose de Finetti measure γ is such that

$$\gamma(\mathrm{d}\theta_1, \ldots, \mathrm{d}\theta_{K-1}) =$$
$$\frac{1}{B(\pi, \nu)} \delta_{\alpha_1(1-\theta_1)}(\mathrm{d}\theta_2) \cdots \delta_{\alpha_{K-2}(1-\theta_1)}(\mathrm{d}\theta_{K-1}) \theta_1^{\pi-1}(1-\theta_1)^{\nu-1}\mathrm{d}\theta_1, \quad (2.13)$$

where $\pi, \nu > 0$ and $\alpha_1, \ldots, \alpha_{K-2} > 0$ are such that $\sum_{i=1}^{K-2} \alpha_i < 1$.

Remark 2.1. Equation (2.13) defines the de Finetti measure of an exchangeable sequence that is neither i.i.d. nor Pólya. In the sequel, we refer to any such sequence as a

$$HLS_K(\pi, \nu, \alpha_1, \ldots, \alpha_{K-2}) \ exchangeable \ sequence,$$

(or, simply, a HLS_K-exchangeable sequence, if the parameters need not be specified in a given context) with the acronym pointing to Hill, Lane and Sudderth [13]. In particular, as deduced from the discussion below, the case $HLS_3(1, 2, \frac{1}{2})$ corresponds to the 3-color urn sequence described in Example 1.7.

When π and ν are integer-valued, for any fixed K, all $HLS_K(\pi, \nu, \alpha_1, \ldots, \alpha_{K-2})$ exchangeable sequences are non Pólya exchangeable urn processes. To see this, it suffices to notice that any such sequence can be generated by means of an urn with initial composition $(\pi, \nu_1, \ldots, \nu_{K-1})$, where the integers ν_i are such that $\nu = \sum_{i=1}^{K-1} \nu_i$, and with urn function $f = (f_1, \ldots, f_K)$ given by $f_1(\mathbf{y}) = y_1$, $f_j(\mathbf{y}) = \alpha_{j-1} \sum_{i=2}^{K} y_i$ $(j = 2, ..., K-1)$, and $f_K(\mathbf{y}) = (1-\alpha) \sum_{i=2}^{K} y_i$, with $\alpha = \sum_{i=1}^{K-2} \alpha_i$. An $HLS_K(\pi, \nu, \alpha_1, \ldots, \alpha_{K-2})$ exchangeable urn process (i.e. with integer-valued π and ν) has consequently the following interpretation: suppose an urn contains initially π balls of color d_1 and ν_{i-1} balls of color d_i, $i = 2, \ldots, K$, with $\sum_{i=1}^{K-1} \nu_i = \nu$. The following random experiment is run at each stage: a ball is drawn, if it is of color d_1, it is replaced along with another of the same color. If the ball drawn is of color

11

d_i, $i = 2, \ldots, K$, it is replaced along with a ball of color d_j, with probability t_j, where $t_j = \alpha_{j-1}$, if $j = 2, \ldots, K - 1$ and $t_j = 1 - \alpha = 1 - \sum_{s=1}^{K-2} \alpha_s$, if $j = K$.

Remark 2.2. In the above described explicit realization of a sequence of the type $HLS_K(\pi, \nu, \alpha_1, \ldots, \alpha_{K-2})$, the initial decomposition of the index ν into integers ν_i, $i = 2, ..., K$, is immaterial.

Let $\mathbf{X}_{[1,\infty)}$ be an $HLS_K(\pi, \nu, \alpha_1, \ldots, \alpha_{K-2})$ exchangeable sequence. For any $\mathbf{x}_n = (x_1, \ldots, x_n) \in \{d_1, \ldots, d_K\}^n$, containing exactly z_i coordinates equal to d_i, $i = 1, \ldots, K - 1$, setting $\theta := \sum_{i=1}^{K-1} \theta_i$ and $z := \sum_{i=1}^{K-1} z_i$, one has that

$$
\begin{aligned}
\mathbb{P}(\mathbf{X}_{[n]} = \mathbf{x}_n) &= \frac{1}{B(\pi, \nu)} \int_{\Theta_{K-1}} (1 - \theta)^{n-z} \Big(\prod_{i=1}^{K-1} \theta_i^{z_i} \Big) \\
&\quad \times \delta_{\alpha_1(1-\theta_1)}(\mathrm{d}\theta_2) \cdots \delta_{\alpha_{K-2}(1-\theta_1)}(\mathrm{d}\theta_{K-1}) \theta_1^{\pi-1}(1 - \theta_1)^{\nu-1} \mathrm{d}\theta_1 \\
&= \frac{(1 - \alpha)^{n-z} \prod_{i=2}^{K-1} \alpha_{i-1}^{z_i}}{B(\pi, \nu)} \int_0^1 \theta_1^{z_1 + \pi - 1} (1 - \theta_1)^{n - z_1 + \nu - 1} \mathrm{d}\theta_1 \\
&= \Big[(1 - \alpha)^{n-z} \prod_{i=1}^{K-2} \alpha_i^{z_{i+1}} \Big] \frac{B(z_1 + \pi, n - z_1 + \nu)}{B(\pi, \nu)}. \qquad (2.14)
\end{aligned}
$$

The following result provides a negative answer to Problem A stated in Section 1.1, and is one of the main achievement of the present paper. In particular, it shows that a naive generalization of Theorem 1.4-(II) cannot be achieved for sets containing strictly more than 2 elements.

Theorem 2.3. *For any $K \geq 3$ and choice of the parameters $\pi, \nu > 0$ and $\alpha_1, \ldots, \alpha_{K-2} > 0$ with $\sum_{i=1}^{K-2} \alpha_i < 1$, the corresponding $HLS_K(\pi, \nu, \alpha_1, \ldots, \alpha_{K-2})$ exchangeable sequence is Hoeffding decomposable while being neither i.i.d. nor Pólya.*

Section 3 contains a combinatorial characterization of Hoeffding decomposability in the framework of exchangeable sequences taking values in a finite set with $K \geq 3$ elements. Such a result will be our main tool in the proof of Theorem 2.3, as detailed in the subsequent Section 4.

12

3. A combinatorial characterization of Hoeffding-decomposability on finite spaces

3.1. Framework

Let $\mathbf{X}_{[1,\infty)} := \{X_n : n \geq 1\}$ be a sequence of exchangeable r.v. with values in $D = \{d_1, ..., d_K\}$, $K \geq 3$. Let γ be the de Finetti measure associated with $\mathbf{X}_{[1,\infty)}$. Throughout this section, we will systematically assume that $\mathbf{X}_{[1,\infty)}$ is such that

$$\mathbb{P}(\mathbf{X}_{[n]} = \mathbf{x}_n) > 0. \qquad \forall \mathbf{x}_n \in D^n, \quad \forall n \geq 1. \tag{3.15}$$

In the sequel, we will adopt the following notation: let $\mathcal{N}(n, K)$ denote the set of weak K-compositions of n, that is the collection of all vectors $\mathbf{i}_K = (i_1, ..., i_K) \in \mathbb{N}^K$ such that $\sum_{j=1}^{K} i_j = n$. For each $n \geq 1$ and each $\mathbf{i}_K \in \mathcal{N}(n, K)$, define the set

$$C(n, \mathbf{i}_K) := \left\{ \mathbf{x}_n \in D^n : \sum_{h=1}^{n} \mathbf{1}(x_h = d_j) = i_j, \; j = 1, ..., K \right\}.$$

By exchangeability of $\mathbf{X}_{[1,\infty)}$ and symmetry of all $\varphi \in \Xi_n(\mathbf{X}_{[1,\infty)})$, for all $n \geq 2$ and all $\mathbf{i}_K \in \mathcal{N}(n, K)$, the functions $\mathbf{x}_n \mapsto \mathbb{P}(\mathbf{X}_{[n]} = \mathbf{x}_n)$ and $\mathbf{x}_n \mapsto \varphi(\mathbf{x}_n)$ are constant on $C(n, \mathbf{i}_K)$. The constant values taken by each of these functions is indicated, respectively, $\mathbb{P}_n(i_1, ..., i_{K-1})$ and $\varphi_n(i_1, ..., i_{K-1})$. Note that the omission of the last coordinate of the vector $\mathbf{i}_K$ comes from the fact that its value is completely determined by those of the previous $K - 1$ coordinates.

Remark 3.1 (On multinomial coefficients). Consider integers $m \geq 1$ and $b_1, ..., b_k \geq 0$ such that $\sum b_i \leq m$. In what follows we adopt the notation $\binom{m}{b_1, \cdots, b_k}$ in order to indicate the multinomial coefficient $m![b_1! \cdots b_k!(m - \sum b_i)!]^{-1}$. We shall also use the following special "star notation":

$$\binom{m}{b_1, \cdots, b_k}_* = \binom{m}{b_1}_* \binom{m - b_1}{b_2}_* \cdots \binom{m - (b_1 + \cdots + b_{k-1})}{b_k}_* \tag{3.16}$$

where

$$\binom{a}{b}_* = \binom{a}{b} \mathbf{1}_{\{0,...,a\}}(b) \tag{3.17}$$

and $\binom{a}{b}$ is the usual binomial coefficient. Note that $\binom{m}{b_1, \cdots, b_k} = \binom{m}{b_1, \cdots, b_k}_*$, whenever the binomial coefficients on the RHS of (3.16) are all different from zero.

13

3.2. Two technical lemmas

Our first technical result concerns the structure of the spaces $\Xi_n\left(\mathbf{X}_{[1,\infty)}\right)$ introduced in Section 1.4.

Lemma 3.2. *If* $\mathbf{X}_{[1,\infty)}$ *is an exchangeable random sequence satisfying (3.15). Then the vector space* $\Xi_n\left(\mathbf{X}_{[1,\infty)}\right)$ *is the* $\left[\binom{n+K-1}{K-1} - \binom{n+K-2}{K-1}\right]$-*dimensional vector space spanned by the symmetric kernels* $\varphi_n^{\mathbf{m}_{K-2}}$, $\mathbf{m}_{K-2} = (m_1, \ldots, m_{K-2}) \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$, *such that, for each* $\mathbf{m}_{K-2} = (m_1, \ldots, m_{K-2}) \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$, *and each* $\mathbf{i}_K = (i_1, \ldots, i_{K-1}, i_K) \in \mathcal{N}(n, K)$,

$$\varphi_n^{\mathbf{m}_{K-2}}(i_1, ..., i_{K-1}) = (-1)^{i_1} \binom{i_1}{m_1 - i_2 \cdots m_{K-2} - i_{K-1}}_* \frac{\mathbb{P}_n(0, m_1, ..., m_{K-2})}{\mathbb{P}_n(i_1, ..., i_{K-1})}. \tag{3.18}$$

Proof. The fact that

$$\dim(\Xi_n(\mathbf{X}_{[1,\infty)})) = \binom{n+K-1}{K-1} - \binom{n+K-2}{K-1}$$

follows from [9, Proposition 6]. In order to prove the rest of the statement, we will show that the collection

$$\Phi_n := \left\{ \varphi_n^{\mathbf{m}_{K-2}} \colon \mathbf{m}_{K-2} \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2) \right\},$$

is indeed a basis of the vector space $\Xi_n(\mathbf{X}_{[1,\infty)})$. To do this, we will first show that, for each $\mathbf{i}_K \in \{\mathbf{i}_K = (i_1, \ldots, i_K) \in \mathcal{N}(n, K) : i_1 \geq 1\}$, there exists a *linear* mapping $f_{\mathbf{i}_K} : \mathfrak{R}^A \to \mathfrak{R}$, where $A = \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$, such that, for all $\varphi_n \in \Xi_n(\mathbf{X}_{[1,\infty)})$,

$$\varphi_n(i_1, \ldots, i_{K-1}) = f_{\mathbf{i}_K}\left(\varphi_n(0, \mathbf{m}_{K-2}) : \mathbf{m}_{K-2} \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)\right). \tag{3.19}$$

Once the explicit representation (3.19) will be at hand, the characterization of Φ_n as a basis will be deduced from the fact that, for $\mathbf{m}_{K-2}$ in $\bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$, the functions $\varphi_n^{\mathbf{m}_{K-2}}$ appearing in the statement verify the relation

$$\varphi_n^{\mathbf{m}_{K-2}}(i_1, \ldots, i_K) = f_{\mathbf{i}_K}\left(\mathbf{1}_{\{\mathbf{m}_{K-2}\}}(\mathbf{y}_{K-2}) : \mathbf{y}_{K-2} \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)\right) \tag{3.20}$$

14

Let $\varphi_n \in \Xi_n \left(\mathbf{X}_{[1,\infty)} \right)$. It turns out that, for all $\mathbf{i}_K \in \{ \mathbf{i}_K = (i_1, \ldots, i_K) \in \mathcal{N}(n, K) : i_1 \geq 1 \}$,

$$\varphi_n (i_1, \ldots, i_{K-1}) = -\frac{1}{\mathbb{P}_n (i_1, \ldots, i_{K-1})} \sum_{j_1=2}^{K} \varphi_n \cdot \mathbb{P}_n \left(\mu_1^{j_1} (i_1, \ldots, i_{K-1}) \right), \quad (3.21)$$

where $\varphi_n \cdot \mathbb{P}_n (\cdot) := \varphi_n (\cdot) \mathbb{P}_n (\cdot)$, for $1 \leq l < p \leq K - 1$,

$$\mu_l^p (i_1, \ldots, i_{K-1}) := (i_1, \ldots, i_{l-1}, i_l - 1, i_{l+1}, \ldots, i_{p-1}, i_p + 1, i_{p+1}, \ldots, i_{K-1}),$$

and, for $1 \leq l \leq K - 1$,

$$\mu_l^K (i_1, \ldots, i_{K-1}) := (i_1, \ldots, i_{l-1}, i_l - 1, i_{l+1}, \ldots, i_{K-1}).$$

Before proving formula (3.21), let us clarify our notation with some simple examples and remarks. For $1 \leq l < p \leq K - 1$, the action of the operator μ_l^p consists in subtracting 1 from the l-th coordinate of the vector $(i_1, \ldots, i_{K-1})$ an adding 1 to the p-th coordinate. For example:

$$\mu_2^4 (2, 7, 5, 9, 4) = (2, 6, 5, 10, 4).$$

On the other hand, when $1 \leq l \leq K-1$ and $p = K$, the action of the operator μ_l^K consists in just subtracting 1 from the l-th coordinate. This is consistent with our conventions since we are omitting the K-th coordinate of the vectors $(i_1, \ldots, i_K)$; in other words the the 1 subtracted from the l-th coordinate is actually added to the last coordinate whose value we are omitting since it is completely determined by the values of the previous ones.

We shall now prove formula (3.21). Fix $n \geq 2$ and $\varphi_n \in \Xi_n \left(\mathbf{X}_{[1,\infty)} \right)$. By the definition of $\Xi_n \left(\mathbf{X}_{[1,\infty)} \right)$, we must have

$$E \left(\varphi_n \left(\mathbf{X}_{[n]} \right) \mid \mathbf{X}_{[2,n]} \right) = 0.$$

Then, for any arbitrarily fixed $\mathbf{x}_{n-1} \in D^{n-1}$,

$$\sum_{i=1}^{K} \varphi_n (d_i, x_2, \ldots, x_{n-1}) \frac{\mathbb{P}_n \left(X_1 = d_i, \mathbf{X}_{[2,n]} = \mathbf{x}_{n-1} \right)}{\mathbb{P}_{n-1} \left(\mathbf{X}_{[2,n]} = \mathbf{x}_{n-1} \right)} = 0.$$

Suppose $\mathbf{x}_{n-1} \in C(n-1, \mathbf{h}_K)$, for some $\mathbf{h}_K \in \mathcal{N}(n-1, K)$. Then, by (3.15), the just-stated formula is equivalent to

$$\varphi_n \cdot \mathbb{P}_n (h_1 + 1, h_2, \ldots, h_{K-1}) + \sum_{j_1=2}^{K} \varphi_n \cdot \mathbb{P}_n \left(\mu_1^{j_1} (h_1 + 1, h_2, \ldots, h_{K-1}) \right) = 0,$$

15

thus proving that (3.21) holds for the vector $\mathbf{i}_K^* = (i_1, ..., i_K) := (h_1 + 1, h_2, ..., h_K)$. Clearly, $\mathbf{i}_K^* \in \{\mathbf{i}_K = (i_1, \ldots, i_K) \in \mathcal{N}(n, K) : i_1 \geq 1\}$. To see that (3.21) holds for all $\mathbf{i}_K \in \{\mathbf{i}_K = (i_1, \ldots, i_K) \in \mathcal{N}(n, K) : i_1 \geq 1\}$, observe that for any such $\mathbf{i}_K$, there exists $\mathbf{h}_K \in \mathcal{N}(n - 1, K)$ such that $\mathbf{i}_K = (h_1 + 1, ..., h_K)$. This proves also that

$$card\left(\{\mathbf{i}_K = (i_1, ..., i_K) \in \mathcal{N}(n, K) : i_1 \geq 1\}\right) = card\left(\mathcal{N}(n - 1, K)\right). \quad (3.22)$$

Now, the recursion in (3.21) yields

$$\varphi_n(i_1, ..., i_{K-1}) = \frac{(-1)^{i_1}}{\mathbb{P}_n(i_1, ..., i_{K-1})} \sum_{j_1=2}^{K} \cdots \sum_{j_{i_1}=2}^{K} \varphi_n \cdot \mathbb{P}_n\left(\mu_1^{j_{i_1}} \circ \cdots \circ \mu_1^{j_1}(i_1, ..., i_{K-1})\right), \quad (3.23)$$

where the operator $\mu_1^{j_{i_1}} \circ \cdots \circ \mu_1^{j_1}$ denotes the successive iteration of operators $\mu_1^{j_1}, ..., \mu_1^{j_{i_1}}$. For instance,

$$\mu_1^2 \circ \mu_1^3 \circ \mu_1^2 \circ \mu_1^4(4, 7, 5, 4, 9) = \mu_1^2\left(\mu_1^3\left(\mu_1^2\left(\mu_1^4(4, 7, 5, 4, 9)\right)\right)\right) = (0, 9, 6, 5, 9).$$

To see this, fix $j_1 \in \{2, ..., K\}$ and apply (3.21) to $\varphi_n\left(\mu_1^{j_1}(i_1, ..., i_{K-1})\right)$ to obtain

$$\varphi_n\left(\mu_1^{j_1}(i_1, ..., i_{K-1})\right) = -\frac{1}{\mathbb{P}_n\left(\mu_1^{j_1}(i_1, ..., i_{K-1})\right)} \sum_{j_2=2}^{K} \varphi_n \cdot \mathbb{P}_n\left(\mu_1^{j_2} \circ \mu_1^{j_1}(i_1, ..., i_{K-1})\right).$$

Do that for all $j_1 \in \{2, ..., K\}$ and plug in (3.21) to obtain

$$\varphi_n(i_1, ..., i_{K-1}) = \frac{1}{\mathbb{P}_n(i_1, ..., i_{K-1})} \sum_{j_1=1}^{K} \sum_{j_2=2}^{K} \varphi_n \cdot \mathbb{P}_n\left(\mu_1^{j_2} \circ \mu_1^{j_1}(i_1, ..., i_{K-1})\right).$$

Iterating the process i_1 times gives (3.23).

Next, observe that the term $\varphi_n \cdot \mathbb{P}_n\left(\mu_1^{j_{i_1}} \circ \cdots \circ \mu_1^{j_1}(i_1, ..., i_{K-1})\right)$ is certainly of the form

$$\varphi_n \cdot \mathbb{P}_n(0, i_2 + b_1, i_3 + b_2, ..., i_{K-1} + b_{K-2}), \quad (3.24)$$

where $b_v = \sum_{t=1}^{i_1} \mathbf{1}(j_t = v + 1)$, $v = 1, ..., K - 2$ is the number of 1's subtracted from the first coordinate of $(i_1, ..., i_{K-1})$ and added to coordinate $v + 1$. Note that b_{K-1} (i.e. the number of 1's subtracted from the first

16

coordinate and added to the last) is completely determined by the vector $\mathbf{b}_{K-2} = (b_1, ..., b_{K-2})$. Clearly, the operator $\mu_1^{j_{i_1}} \circ \cdots \circ \mu_1^{j_1}$ is commutative in the sense that

$$\mu_1^{j_{i_1}} \circ \cdots \circ \mu_1^{j_1} = \mu_1^{j_{\sigma(i_1)}} \circ \cdots \circ \mu_1^{j_{\sigma(1)}},$$

for any permutation σ of $(1, ..., i_1)$. It follows that, for any fixed $(b_1, ..., b_{K-2}) \in \bigcup_{a=0}^{i_1} \mathcal{N}(a, K-2)$ the term (3.24) occurs exactly $\binom{i_1}{b_1, b_2 ..., b_{K-2}}$ times in the sum described in (3.23). Consequently, (3.23) can be rewritten as follows:

$$\varphi_n(i_1, ..., i_{K-1}) = \frac{(-1)^{i_1}}{\mathbb{P}_n(i_1, ..., i_{K-1})}$$
$$\times \sum_{(+)} \binom{i_1}{b_1, ..., b_{K-2}} \varphi_n \cdot \mathbb{P}_n(0, i_2 + b_1, i_3 + b_2 ...., i_{K-1} + b_{K-2}), \quad (3.25)$$

where the sum $(+)$ is extended to all vectors $\mathbf{b}_{K-2} = (b_1, \ldots, b_{K-2}) \in \bigcup_{a=0}^{i_1} \mathcal{N}(a, K-2)$. Set $m_p := b_p + i_{p+1}$, $p = 1, \ldots, K-2$, and rewrite (3.25) as

$$\varphi_n(i_1, \ldots, i_{K-1}) = \frac{(-1)^{i_1}}{\mathbb{P}_n(i_1, \ldots, i_{K-1})}$$
$$\times \sum_{(\sharp)} \binom{i_1}{m_1 - i_2, \ldots, m_{K-2} - i_{K-1}} \varphi_n \cdot \mathbb{P}_n(0, m_1, \ldots, m_{K-2}), \quad (3.26)$$

where the sum $(\sharp)$ is extended to all vectors $(m_1, \ldots, m_{K-2}) \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$, such that

$$
\begin{aligned}
m_1 &\in \{i_2, \ldots, i_1 + i_2\}, \\
m_2 &\in \{i_3, \ldots, (i_1 - m_1) + i_2 + i_3\} \\
m_3 &\in \{i_4, \ldots, (i_1 - m_1) + (i_2 - m_2) + i_3 + i_4\} \\
&\vdots \\
m_{K-2} &\in \left\{ i_{K-1}, \ldots, \sum_{v=1}^{K-3}(i_v - m_v) + i_{K-2} + i_{K-1} \right\}.
\end{aligned}
$$

It is immediately seen that the multinomial coefficient in (3.26) is always well defined. It follows that (3.26) can be rewritten, using the convention defined

17

in (3.16), as

$$\varphi_n(i_1, \ldots, i_{K-1}) = \frac{(-1)^{i_1}}{\mathbb{P}_n(i_1, \ldots, i_{K-1})}$$
$$\times \sum_{(=)} \binom{i_1}{m_1 - i_2, \ldots, m_{K-2} - i_{K-1}}_* \varphi_n \cdot \mathbb{P}_n(0, m_1, \ldots, m_{K-2}), \quad (3.27)$$

where the sum $(=)$ is extended to all vectors $\mathbf{m}_{K-2} = (m_1, \ldots, m_{K-2}) \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$. Since equality (3.27) holds for any $\varphi_n \in \Xi_n(\mathbf{X}_{[1,\infty)})$ and any $\mathbf{i}_K \in \{\mathbf{i}_K = (i_1, \ldots, i_k) \in \mathcal{N}(n, K) : i \geq 1\}$, (3.19) is proved. The claim of the present step of the proof follows, now, immediately from (3.20).   ∎

Next, an adaptation of the arguments rehearsed in the proof of Lemma 3 in [9] yields the following statement about symmetrizations.

Lemma 3.3. *Fix $m \geq 2$, $v \in \{1, ..., m-1\}$, $D = \{d_1, ..., d_K\}$ and let the application*

$$f_{v,m-v} : D^m \mapsto \mathfrak{R} : (x_1, ..., x_m) \mapsto f_{v,m-v}(x_1, ..., x_m),$$

be separately symmetric in the variables $(x_1, ..., x_v)$ and $(x_{v+1}, ..., x_m)$ (and not necessarily as a function on D^m). Then, for any $\mathbf{x}_m \in C(m, \mathbf{z}_K)$ for some $\mathbf{z}_K = (z_1, ..., z_K) \in \mathcal{N}(m, K)$, the canonical symmetrization of f, denoted by $\tilde{f}$, reduces to

$$\tilde{f}(\mathbf{x}_m) = \sum_{(*)} \binom{v}{k_1, k_2, ..., k_{K-1}} \binom{m-v}{z_1 - k_1, z_2 - k_2, ..., z_{K-1} - k_{K-1}}$$
$$\times \frac{f_{v,m-v}((k_1, ..., k_{K-1}), (z_1 - k_1, ..., z_K - k_K))}{\sum_{(*)} \binom{v}{k_1, k_2, ..., k_{K-1}} \binom{m-v}{z_1 - k_1, z_2 - k_2, ..., z_{K-1} - k_{K-1}}}, \quad (3.28)$$

where the sums $(*)$ are extended to all the vectors $(k_1, ..., k_{K-1})$, in sequel referred to as $(m, v, \mathbf{z}_K)$-coherent vectors, with

$$k_1 \in \{0 \vee [z_1 - (m-v)], ..., z_1 \wedge v\},$$
$$k_2 \in \{0 \vee [z_2 - (m-v) - (z_1 - k_1)], ..., z_2 \wedge (v - k_1)\},$$
$$\vdots$$
$$k_{K-1} \in \{0 \vee [z_{K-1} - (m-v) - \Sigma_1^{K-2}(z_p - k_p)], ..., z_{K-1} \wedge (v - \Sigma_1^{K-2} k_p)\}$$

18

where $f_{v,m-v}\left(\left(k_1,...,k_{K-1}\right),\left(z_1-k_1,...,z_K-k_K\right)\right)$ denotes the common value of $f_{v,m-v}\left(\mathbf{y}_m\right)$ when $\mathbf{y}_m=\left(y_1,...,y_m\right)$ is such that the subvector $\left(y_1,...,y_v\right)$ contains exactly k_i coordinates equal to d_i, $i=1,...,K$, and the subvector $\left(y_{v+1},...,y_m\right)$ contains exactly $\left(z_i-k_i\right)$ coordinates equal to d_i, $i=1,...,K$.

As a consequence, $\tilde{f}_{v,m-v}\left(\mathbf{x}_m\right)=0$ for every $\mathbf{x}_m\in D^m$ if, and only if, for all $\mathbf{z}_K=\left(z_1,...,z_K\right)\in\mathcal{N}\left(m,K\right),$

$$
\sum_{(*)}\binom{v}{k_1,k_2,...,k_{K-1}}\binom{m-v}{z_1-k_1,z_2-k_2,...,z_{K-1}-k_{K-1}}
$$
$$
\times f_{v,m-v}\left(\left(k_1,...,k_{K-1}\right),\left(z_1-k_1,...,z_K-k_K\right)\right)=0. \quad (3.29)
$$

3.3. The characterization

We are now ready to prove the announced full characterization of D-valued Hoeffding decomposable exchangeable sequences satisfying (3.15), where $D=\{d_1,...,d_K\}$. To this end, recall that, for every symmetric $\varphi:D^n\to\mathfrak{R}$, every $u=2,...,n$ and every $\mathbf{x}_{n-1}=\left(x_1,...,x_{n-1}\right)\in D^{n-1}$,

$$
\left[\varphi\right]_{n,n-1}^{(n-u)}\left(\mathbf{x}_{n-1}\right)=\mathbb{E}\left(\varphi\left(\mathbf{X}_{[n]}\right)\mid\mathbf{X}_{[u+1,u+n-1]}=\mathbf{x}_{n-1}\right).
$$

Observe that the function $\left[\varphi\right]_{n,n-1}^{(n-u)}:D^{n-1}\to\mathfrak{R}$ clearly meets the symmetry properties of Lemma 3.3 with $m=n-1$ and $v=n-u$. Now, fix $\mathbf{z}_K=\left(z_1,...,z_K\right)\in\mathcal{N}\left(n-1,K\right)$ and suppose $\mathbf{x}_{n-1}=\left(x_1,...,x_{n-1}\right)\in C\left(n-1,\mathbf{z}_K\right)$ is such that $\sum_{t=1}^{n-u}\mathbf{1}\left(x_t=d_p\right)=k_p,$ $p=1,...K-1$. Then

$$
\left[\varphi\right]_{n,n-1}^{(n-u)}\left(\mathbf{x}_{n-1}\right)=\sum_{q_1=0}^{u}\sum_{q_2=0}^{u-q_1}\cdots\sum_{q_{K-1}=1}^{u-\Sigma_1^{K-2}q_j}\binom{u}{q_1,...,q_{K-1}}
$$
$$
\times\varphi_n\left(k_1+q_1,...,k_{K-1}+q_{K-1}\right)\frac{\mathbb{P}_{n+u-1}\left(z_1+q_1,...,z_{K-1}+q_{K-1}\right)}{\mathbb{P}_{n-1}\left(z_1,...,z_{K-1}\right)}. \quad (3.30)
$$

By applying (3.29) in the case $m=n-1$, we deduce that $\widetilde{\left[\varphi\right]}_{n,n-1}^{(n-u)}\left(\mathbf{x}_{n-1}\right)=0$ if and only if

$$
\sum_{(*)}\binom{n-u}{k_1,k_2,...,k_{K-1}}\binom{u-1}{z_1-k_1,z_2-k_2,...,z_{K-1}-k_{K-1}}\times
$$
$$
\times\left[\varphi\right]_{n,n-1}^{(n-u)}\left(\left(k_1,...,k_{K-1}\right),\left(z_1-k_1,...,z_K-k_K\right)\right)=0, \quad (3.31)
$$

19

where the sum $(*)$ runs over all $(n-1, n-u, \mathbf{z}_K)$-coherent vectors $(k_1, ..., k_{K-1})$, i.e.

$$
\begin{aligned}
k_1 &\in \left\{ 0 \vee [z_1 - (u-1)], ..., z_1 \wedge (n-u) \right\}, \\
k_2 &\in \left\{ 0 \vee [z_2 - (u-1) - (z_1 - k_1)], ..., z_2 \wedge (n-u-k_1) \right\}, \\
&\vdots \\
k_{K-1} &\in \left\{ 0 \vee \left[z_{K-1} - (u-1) - \Sigma_1^{K-2}(z_p - k_p) \right], ..., z_{K-1} \wedge \left(n - u - \Sigma_1^{K-2} k_p \right) \right\},
\end{aligned}
$$

and the notation $[\varphi]_{n,n-1}^{(n-u)}((k_1, ..., k_{K-1}), (z_1 - k_1, ..., z_K - k_K))$ indicates the common value of $[\varphi]_{n,n-1}^{(n-u)}(\mathbf{x}_{n-1})$, for all $\mathbf{x}_{n-1}(x_1, ..., x_{n-1}) \in C(n-1, \mathbf{z}_K)$ such that $\sum_{t=1}^{n-u} \mathbf{1}(x_t = d_p) = k_p$, $p = 1, ..., K-1$.

Now recall that, by Theorem 1.4-(I), Hoeffding decomposability and weak independence are equivalent provided condition (1.11) is verified. The fact that such condition is verified in our case is a consequence of [9, Proposition 6, Point 2]. Moreover, $\mathbf{X}_{[1,\infty)}$ is weakly independent if, and only if, for all $n \geq 2$ and all $\varphi_n \in \Xi_n\left(\mathbf{X}_{[1,\infty)}\right)$, one has $\varphi_n \in \tilde{\Xi}_{n,n-u}\left(\mathbf{X}_{[1,\infty)}\right)$, for all $u = 2, ..., n$. By Lemma 3.2, this implies that for every $\mathbf{m}_{K-2} \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$, the corresponding basis function $\varphi_n^{\mathbf{m}_{K-2}}$ belongs to $\tilde{\Xi}_{n,n-u}\left(\mathbf{X}_{[1,\infty)}\right)$. On the other hand since any $\varphi_n \in \Xi_n\left(\mathbf{X}_{[1,\infty)}\right)$ is a linear combination of the basis functions $\varphi_n^{\mathbf{m}_{K-2}}$, we deduce that weak independence occurs if, and only if, for all $\mathbf{m}_{K-2} \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$, $\varphi_n^{\mathbf{m}_{K-2}} \in \tilde{\Xi}_{n,n-u}\left(\mathbf{X}_{[1,\infty)}\right)$. In other words, weak independence occurs if, and only if, for all $n \geq 2$, all $u = 2, ..., n$, all $\mathbf{z}_K = (z_1, ..., z_K) \in \mathcal{N}(n-1, K)$, and all $\mathbf{m}_{K-2} \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$

$$
\begin{aligned}
\sum_{(*)} &\binom{n-u}{k_1, k_2, ..., k_{K-1}} \binom{u-1}{z_1 - k_1, z_2 - k_2, ..., z_{K-1} - k_{K-1}} \\
&\times [\varphi_n^{\mathbf{m}_{K-2}}]_{n,n-1}^{(n-u)}((k_1, ..., k_{K-1}), (z_1 - k_1, ..., z_K - k_K)) = 0, \quad (3.32)
\end{aligned}
$$

the sum $(*)$ being, as before, extended to all $(n-1, n-u, \mathbf{z}_K)$-coherent vectors $(k_1, ..., k_{K-1})$. Substituting (3.18) and (3.30) in (3.31), one has that (3.32) is true if, and only if, all $\mathbf{m}_{K-2} \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$, for the following

20

quantity equals 0:

$$\frac{\mathbb{P}_n\left(0, m_1, ..., m_{K-2}\right)}{\mathbb{P}_{n-1}\left(z_1, ..., z_{K-1}\right)} \sum_{(*)} (-1)^{k_1} \binom{n-u}{k_1, k_2, ..., k_{K-1}}$$

$$\times \binom{u-1}{z_1 - k_1, z_2 - k_2, ..., z_{K-1} - k_{K-1}} \sum_{(**)} (-1)^{q_1} \binom{u}{q_1, ..., q_{K-1}} \qquad (3.33)$$

$$\times \binom{k_1 + q_1}{m_1 - k_2 - q_2 \cdots m_{K-2} - k_{K-1} - q_{K-1}}_* \frac{\mathbb{P}_{n+u-1}\left(z_1 + q_1, ..., z_{K-1} + q_{K-1}\right)}{\mathbb{P}_n\left(k_1 + q_1, ..., z_{K-1} + k_{K-1}\right)},$$

where the sum $(*)$ is over all $(n-1, n-u, \mathbf{z}_K)$-coherent vectors $(k_1, ..., k_{K-1})$ and the sum $(**)$ is w.r.t. all $\mathbf{q}_{K-1} = (q_1, ..., q_{K-1}) \in \bigcup_{a=0}^{u} \mathcal{N}(a, K-1)$. Note that

$$\frac{\mathbb{P}_{n+u-1}\left(z_1 + q_1, ..., z_{K-1} + q_{K-1}\right)}{\mathbb{P}_n\left(k_1 + q_1, ..., z_{K-1} + k_{K-1}\right)}$$
$$= \frac{\mathbb{P}_{n+u-1}^n\left(z_1 + q_1, ..., z_{K-1} + q_{K-1} \mid k_1 + q_1, ..., z_{K-1} + k_{K-1}\right)}{\binom{u-1}{z_1 - k_1, z_2 - k_2, ..., z_{K-1} - k_{K-1}}}, \quad (3.34)$$

where $\mathbb{P}_{n+u-1}^n\left(z_1 + q_1, ..., z_{K-1} + q_{K-1} \mid k_1 + q_1, ..., z_{K-1} + k_{K-1}\right)$ denotes the conditional probability that the vector $\mathbf{X}_{[n+u-1]}$ contains exactly $z_p + q_p$ coordinates equal to d_p, $p = 1, ..., K-1$, given that the subvector $\mathbf{X}_{[n]}$ contains exactly $k_p + q_p$ coordinates equal to d_p, $p = 1, ..., K-1$. Plugging (3.34) into (3.33), one deduces immediately the announced characterization of weak independence.

Theorem 3.4. *Let* $\mathbf{X}_{[1,\infty)}$ *be an infinite sequence of exchangeable D-valued r.v. satisfying* (3.15). *For the sequence to be Hoeffding-decomposable, it is necessary and sufficient that, for every* $n \geq 2$, *every* $u = 2, ..., n$, *every* $\mathbf{z}_K = (z_1, ..., z_K) \in \mathcal{N}(n-1, K)$ *and every* $\mathbf{m}_{K-2} \in \bigcup_{a=0}^{n} \mathcal{N}(a, K-2)$, *the following quantity equals* 0:

$$\sum_{(*)} (-1)^{k_1} \binom{n-u}{k_1, k_2, ..., k_{K-1}} \times \qquad (3.35)$$

$$\times \sum_{(**)} (-1)^{q_1} \binom{u}{q_1, ..., q_{K-1}} \binom{k_1 + q_1}{m_1 - k_2 - q_2 \cdots m_{K-2} - k_{K-1} - q_{K-1}}_*$$

21

$$\times \, \mathbb{P}^n_{n+u-1} \left(z_1 + q_1, ..., z_{K-1} + q_{K-1} \mid k_1 + q_1, ..., q_{K-1} + k_{K-1} \right)$$

where the sums $(*)$ *and* $(**)$ *run over, respectively, all* $(n-1, n-u, \mathbf{z}_K)$*-coherent vectors* $(k_1, ..., k_{K-1})$ *and all* $\mathbf{q}_{K-1} = (q_1, ..., q_{K-1}) \in \bigcup_{a=0}^{u} \mathcal{N}(a, K-1)$ *.*

In the next section, we will use the content of Theorem 3.4 specialized to the case $K = 3$.

4. Proof of Theorem 2.3

We start by stating a result that is easily deduced from the proof of the Theorem 1 in [9].

Lemma 4.1. *For all* $\pi, \nu > 0$, *all* $n \geq 2$, *all* $u = 2, \ldots, n$, *all* $z = 0, \ldots, n-1$ *and all* $k \in \{\max\{0, z - (u-1)\}, \ldots, \min\{z, n-u\}\}\}$,

$$\sum_{q=0}^{u} (-1)^q \binom{u}{q} \frac{B(\pi + z + q, \nu + n + u - 1 - z - q)}{B(\pi + k + q, \nu + n - k - q)} = 0.$$

To minimize the notational burden, we shall restrict ourselves to the case $K = 3$. The proof for general K carries out exactly in the same way. Let $\mathbf{X}_{[1,\infty)}$ be a $HLS_3(\pi, \nu, \alpha)$- exchangeable sequence with values in $D = \{d_1, d_2, d_3\}$, where $\pi, \nu > 0$ and $0 < \alpha < 1$. By (2.14), the following facts are in order : (a) $\mathbf{X}_{[1,\infty)}$ satisfies (3.15), (b) $\mathbf{X}_{[1,\infty)}$ is neither i.i.d. nor a K-color Pólya sequence and (c) if $\mathbf{x}_n \in D^n$ contains exactly z_1 coordinates equal to d_1 and z_2 coordinates equal to d_2, then, in the language of the present paper,

$$\mathbb{P}(\mathbf{X}_{[n]} = \mathbf{x}_n) = \mathbb{P}_n(z_1, z_2) = \alpha^{z_2}(1-\alpha)^{n-z_1-z_2} \frac{B(z_1 + \pi, n - z_1 + \nu)}{B(\pi, \nu)}. \quad (4.36)$$

Recall that, by Theorem 3.4, an exchangeable sequence with values in $D = \{d_1, d_2, d_3\}$ is Hoeffding-decomposable if and only if, for all $n \geq 2$, all $u = 2, \ldots, n$, all $m = 0, \ldots, n$ and all $(z_1, z_2) \in S(z_1, z_2) := \{(z_1, z_2) \in \{0, \ldots, n-1\}^2 \colon z_1 + z_2 \leq n-1\}$, one has

$$0 = \sum_{k_1 \in A(z_1, u)} \sum_{k_2 \in A_{z_1, k_1}(z_2, u)} (-1)^{k_1} \binom{n-u}{k_1 \ \ k_2} \binom{u-1}{z_1 - k_1 \ \ z_2 - k_2}$$

22

$$\times \sum_{q_1=0}^{u} \sum_{q_2=0}^{u-q_1} (-1)^{q_1} \binom{u}{q_1 \ q_2} \binom{k_1+q_1}{m-k_2-q_2}_* \frac{\mathbb{P}_{n+u-1}(z_1+q_1, z_2+q_2)}{\mathbb{P}_n(k_1+q_1, k_2+q_2)},$$

where the notation $\binom{a}{b}_*$ is defined in (3.17) and

$$A(z_1, u) = \{\max\{0, z_1-(u-1)\}, \ldots, \min\{z_1, n-u\}\}$$
$$A_{z_1,k_1}(z_2, u) = \{\max\{0, z_2-(u-1)-(z_1-k_1)\}, \ldots, \min\{z_1, n-u-k_1\}\}.$$

It follows that $\mathbf{X}_{[1,\infty)}$ is Hoeffding-decomposable if and only if for all $n \geq 2$, all $u = 2, \ldots, n$, all $m = 0, \ldots, n$ and all $(z_1, z_2) \in S(z_1, z_2) := \{(z_1, z_2) \in \{0, \ldots, n-1\}^2 : z_1 + z_2 \leq n-1\}$, one has

$$0 = \sum_{k_1 \in A(z_1,u)} \sum_{k_2 \in A_{z_1,k_1}(z_2,u)} (-1)^{k_1} \binom{n-u}{k_1 \ k_2} \binom{u-1}{z_1-k_1 \ z_2-k_2}$$
$$\times C \sum_{q_1=0}^{u} \sum_{q_2=0}^{u-q_1} (-1)^{q_1} \binom{u}{q_1 \ q_2} \binom{k_1+q_1}{m-k_2-q_2}_*$$
$$\times \frac{B(\pi+z_1+q_1, \nu+n+u-1-z_1-q_1)}{B(\pi+k_1+q_1, \nu+n-k_1-q_1)},$$

where

$$C = C(\alpha, u, z_1, z_2, k_1, k_2) = \alpha^{z_2-k_2}(1-\alpha)^{(u-1)-(z_1-k_1)-(z_2-k_2)},$$

clearly does not depend on q_1 and q_2.

We will, in fact, show an even stronger fact. More precisely, we will show that, for all $n \geq 2$, all $u = 2, \ldots, n$, all $m = 0, \ldots, n$, all $(z_1, z_2) \in S(z_1, z_2) := \{(z_1, z_2) \in \{0, \ldots, n-1\}^2 : z_1 + z_2 \leq n-1\}$, all $k_1 \in A(z_1, u)$ and all $k_2 \in A_{z_1,k_1}(z_2, u)$, one has

$$\sigma(n, u, m, z_1, k_1, k_2) = 0,$$

where

$$\sigma(n, u, m, z_1, k_1, k_2) := \sum_{q_1=0}^{u} \sum_{q_2=0}^{u-q_1} (-1)^{q_1} \binom{u}{q_1 \ q_2} \binom{k_1+q_1}{m-k_2-q_2}_*$$
$$\times \frac{B(\pi+z_1+q_1, \nu+n+u-1-z_1-q_1)}{B(\pi+k_1+q_1, \nu+n-k_1-q_1)}.$$

23

Towards this aim, we first show that

$$\sigma(n, u, m, z_1, k_1, k_2) = \quad\quad\quad\quad\quad\quad (4.37)$$

$$= \sum_{q_1=0}^{u} (-1)^{q_1} \binom{u}{q_1} \frac{B(\pi + z_1 + q_1, \nu + n + u - 1 - z_1 - q_1)}{B(\pi + k_1 + q_1, \nu + n - k_1 - q_1)}$$

$$\times \sum_{q_2=0}^{u-q_1} \binom{u - q_1}{q_2} \binom{k_1 + q_1}{m - k_2 - q_2}_*$$

$$= \binom{k_1 + u}{m - k_2}_* \sum_{q_1=0}^{u} (-1)^{q_1} \binom{u}{q_1} \frac{B(\pi + z_1 + q_1, \nu + n + u - 1 - z_1 - q_1)}{B(\pi + k_1 + q_1, \nu + n - k_1 - q_1)}.$$

In other words, we show that

$$\tilde{\sigma}_{q_1}(m, u, k_1, k_2) = \sum_{q_2=0}^{u-q_1} \binom{u - q_1}{q_2} \binom{k_1 + q_1}{m - k_2 - q_2}_* = \binom{k_1 + u}{m - k_2}_*, \quad (4.38)$$

i.e. that $\tilde{\sigma}_{q_1}(m, u, k_1, k_2)$ does actually not depend on q_1. For reading convenience, set $u - q_1 = i$ and $m - k_2 = j$ and rewrite $\tilde{\sigma}_{q_1}(m, u, k_1, k_2)$ as

$$\tilde{\sigma}_i(j, u, k_1) = \sum_{q_2=0}^{i} \binom{i}{q_2} \binom{k_1 + u - i}{j - q_2}_*.$$

To see that

$$\tilde{\sigma}_i(j, u, k_1) = \binom{k_1 + u}{j}_*,$$

start by fixing $i \in \{0, \ldots, u\}$. If $j < 0$ (i.e. if $m < k_2$), then the equality is trivial. If $0 \le j \le i$, then, by definition of $\binom{k_1 + q_1}{j - q_2}_*$, we have

$$\tilde{\sigma}_i(j, u, k_1) = \tilde{\sigma}_j(j, u, k_1),$$

and the result follows by a direct application of the Vandermonde formula. If $j \ge i$, then a direct application of the classical Pascal's triangle gives

$$\binom{k_1 + u}{j}_* = \binom{k_1 + u - 1}{j}_* + \binom{k + u - 1}{j - 1}_*$$

$$= \binom{k_1 + u - 2}{j}_* + 2\binom{k + u - 2}{j - 1}_* + \binom{k_1 + u - 2}{j - 2}_*$$

24

$$= \binom{k_1 + u - 3}{j}_* + 3\binom{k + u - 3}{j - 1}_* + 3\binom{k_1 + u - 3}{j - 2}_* + \binom{k_1 + u - 3}{j - 3}_*$$
$$= \cdots$$
$$= \tilde{\sigma}_i(j, u, k_1).$$

Now that (4.38) is in order, fix, arbitrarily, n, u, z_1, k_1. For all k_2 and m such that $m \in \{0, \ldots, k_2 - 1\} \cup \{k_1 + u + 1, \ldots, n\}$, we have, by definition of $\binom{k_1 + u}{m - k_2}_*$, that $\sigma(n, u, m, z_1, k_1, k_2) = 0$ as desired. The fact that this is still the case for all $m \in \{k_2, \ldots, k_1 + u\}$, follows from (4.37) and Lemma 4.1. The proof of Theorem 2.3 is complete.

5. Concluding remarks

A Bayesian reader will immediately recognize several quantities and structures scattered throughout the paper. To some extent this is not surprising given we are dealing with exchangeable sequences but still the pivotal role played by conditional distributions of the exchangeable r.v., interpreted as predictive distributions by the Bayesian, is quite striking. This hints towards the worthiness of further investigating the connections between Hoeffding decomposability and Bayesian inference. Given the impact Hoeffding decomposability had in the classical framework, one could hope for significant implications and applications also in the Bayesian context. Of particular interest is the nonparametric case, where Hoeffding decomposability could lead to significant insights on the infinite-dimensional objects one is required to deal with. To this end a full characterization of Hoeffding decomposability for the case of exchangeable sequences taking values in a arbitrary Polish space is fundamental. In fact, from [20] we already know that the Ferguson-Dirichlet process is indeed Hoeffding decomposable. Is it the only one? Or does such a property hold also for some other class of nonparametric processes generalizing the Ferguson-Dirichlet process, among the many present in the literature (see [17])? On the one hand the results of the present paper seem to hint that Hoeffding decomposability is not unique to the Ferguson-Dirichlet process. But on the other hand, in passing from the finite to the infinite case, the consistency conditions for the finite-dimensional distributions might restrict the class of Hoeffding-decomposable models significantly. This important open question constitutes the object of future research.

References

References

[1] D.J. Aldous (1983). Exchangeability and related topics. *École d'été de Probabilités de Saint-Flour XIII*, **LNM** 1117, Springer, New York

[2] M.A. Arcones and E. Giné (1993). Limit theorems for U-processes. *The Annals of Probability* **21**, 14941542.

[3] F. Avram and D. Bertsimas (1993). On central limit theorems in geometric probabilities. *The Annals of Applied Probability* **3**, 1033-1046.

[4] V. Bentkus, F. Götze and W.R. van Zwet (1997). An Edgeworth expansion for symmetric statistics. *Annals of Statistics* **25**, 851896.

[5] D. Blackwell and J. MacQueen (1973). Ferguson distribution via Pólya urn schemes. *Annals of Statistics* **1** (2), 353-355

[6] M. Bloznelis (2005). Orthogonal decomposition of symmetric functions defined on random permutations. *Combinatorics, Probability and Computing*, **14**, 249-268

[7] M. Bloznelis and F. Götze (2001). Orthogonal decomposition of finite population statistics and its applications to distributional asymptotics. *The Annals of Statistics* **29**, 353-365.

[8] M. Bloznelis and F. Götze (2002). An Edgeworth expansion for finite population statistics. *The Annals of Probability*, **30**, 1238-1265

[9] O. El-Dakkak and G. Peccati (2008). Hoeffding decompositions and urn processes. *The Annals of Probability* **36**(6), 2280-2310.

[10] T.S. Ferguson (1973). A Bayesian analysis of some nonparametric problems. *The Annals of Statistics* **1**, 209–230.

[11] J. Hajek (1968). Asymptotic normality of simple linear rank statistics under alternatives. *The Annals of Mathematical Statistics* **39** 325346.

[12] W. Hoeffding (1948). A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics* **19**, 293325.

[13] B.H. Hill, D. Lane and W. Sudderth (1987). Exchangeable urn processes. *The Annals of Probability*, **15**(4), 1586-1592.

[14] L.F. James, A. Lijoi and I. Prünster (2006). Conjugacy as a distinctive feature of the Dirichlet process. *Scandinavian Journal of Statistics* **33**, 105–120.

[15] S. Karlin et Y. Rinott (1982). Applications of ANOVA type decompositions for comparisons of conditional variance statistics including jack-knife estimates. *The Annals of Statistics*, **10** (2), 485-501.

[16] V.S. Koroljuk and Yu. V. Borovskich (1994). *Theory of U-Statistics.* Kluwer Academic Publishers, London.

[17] A. Lijoi and I. Prünster (2010). Models beyond the Dirichlet process. In *Bayesian Nonparametrics* (Hjort, N.L., Holmes, C.C. Müller, P., Walker, S.G. Eds.), pp. 80–136. Cambridge University Press, Cambridge.

[18] P. Major (2005). Tail behaviour of multiple random integrals and U-statistics. *Probability Surveys* **2**, 448505.

[19] G. Peccati (2003). Hoeffding decompositions for exchangeable sequences and chaotic representation for functionals of Dirichlet processes. *C.R.A.S. - Mathématiques*, Vol. **336**/10, 845-850.

[20] G. Peccati (2004). Hoeffding-ANOVA decompositions for symmetric statistics of exchangeable observations. *The Annals of Probability*, **32** (3A), 1796-1829.

[21] G. Peccati (2008). Multiple integral representation for functionals of Dirichlet processes. *Bernoulli* **14(1)**, 91-124

[22] J. Pitman (1996). Some developments of the Blackwell-MacQueen urn scheme. In: *Statistics, Probability and Game Theory: papers in honor of David Blackwell.* Volume 30 of Lecture Notes-Monograph Series. Institute of Mathematical Statistics, Hayward, California.

27

[23] E. Regazzini, A. Lijoi and I. Prünster (2003). Distributional results for means of random measures with independent increments. *The Annals of Statistics*, **31**,560-585

[24] R.A. Vitale (1992). Covariances of symmetric statistics. *Journal of Multivariate Analysis*, **41**, 14-26

[25] L. Zhao and X. Chen (1990). Normal approximation for finite-population U-statistics. *Acta Mathematicae Applicatae Sinica* **6** (3), 263-272