The Annals of Probability
2008, Vol. 36, No. 6, 2280–2310
DOI: 10.1214/07-AOP389

HOEFFDING DECOMPOSITIONS AND URN SEQUENCES

BY OMAR EL-DAKKAK AND GIOVANNI PECCATI

Université Paris VI, Université Paris Ouest and Université Paris VI

Let $\mathbf{X} = (X_1, X_2, \ldots)$ be a nondeterministic infinite exchangeable sequence with values in $\{0, 1\}$. We show that $\mathbf{X}$ is Hoeffding decomposable if, and only if, $\mathbf{X}$ is either an i.i.d. sequence or a Pólya sequence. This completes the results established in Peccati [*Ann. Probab.* **32** (2004) 1796–1829]. The proof uses several combinatorial implications of the correspondence between Hoeffding decomposability and weak independence. Our results must be compared with previous characterizations of i.i.d. and Pólya sequences given by Hill, Lane and Sudderth [*Ann. Probab.* **15** (1987) 1586–1592] and Diaconis and Ylvisaker [*Ann. Statist.* **7** (1979) 269–281]. The final section contains a partial characterization of Hoeffding decomposable sequences with values in a set with more than two elements.

1. Introduction. Let $\mathbf{X}_{[1,\infty)} = \{X_n : n \geq 1\}$ be an exchangeable sequence of random observations, with values in some finite set D. We say that $\mathbf{X}_{[1,\infty)}$ is *Hoeffding decomposable* if, for every $n \geq 2$, every symmetric statistic $T(X_1, \ldots, X_n)$ admits a unique representation as an orthogonal sum of uncorrelated U-statistics with degenerate kernels of increasing order. Hoeffding decompositions (also known as *ANOVA decompositions*) have been extensively studied for i.i.d. sequences and extractions without replacement from a finite population. Concerning i.i.d. sequences, the reader is referred to the seminal paper by Hoeffding (1948), where this technique is applied to normal approximations of U-statistics. In the subsequent years, Hoeffding-type decompositions have been further applied to different frameworks, such as linear rank statistics [Hajek (1968)], jackknife estimators [Karlin and Rinott (1982)], covariance analysis of symmetric statistics [Vitale (1992)], convergence of U-processes [Arcones and Giné (1993)], Edgeworth expansions [Bentkus, Götze and van Zwet (1997)], and tail estimates for U-statistics [Majór (2005)]. Concerning extractions without replacement from a finite population, the first analysis of Hoeffding decompositions has been developed by Zhao and Chen (1990). Bloznelis and Götze (2001, 2002) generalize these results, in order to characterize the asymptotic normality of symmetric statistics (when the size of the population diverges to infinity), and to obtain Edgeworth expansions. In Bloznelis (2005) Hoeffding-type decompositions are explicitly obtained for statistics depending on extractions without replacement from several distinct populations.

Received October 2006; revised August 2007.
AMS 2000 subject classifications. 60G09, 60G99.
Key words and phrases. Exchangeable sequences, Hoeffding decompositions, Pólya urns, weak independence.

In Peccati (2003, 2004, 2008), the second author of this paper has extended the theory of Hoeffding decompositions to the framework of general exchangeable random sequences. In Peccati (2004) 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 (1973) or Pitman (1996)]. The connection with Pólya urns is further exploited in Peccati (2008), where Hoeffding-type decompositions are used to establish several new properties of Dirichlet–Ferguson processes [see, e.g., Ferguson (1973) or James, Lijoi and Prünster (2006)]. Among the results obtained in Peccati (2008) via Hoeffding-type techniques we mention (i) the derivation of a chaotic representation property for Dirichlet–Ferguson processes; (ii) the extension of some Bayesian decision rules established in Ferguson (1973); (iii) a new probabilistic representation of Jacobi and generalized Jacobi polynomials, appearing in connection with the transition probabilities of Wright–Fisher diffusion processes of population genetics [see, e.g., Griffiths (1979)].

The aim of this paper is to complete the results established in Peccati (2004) in two directions. On the one hand, we shall prove that *a* (*nondeterministic*) *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.* As discussed in Section 6, this result links the seemingly unrelated notions of Hoeffding decomposable sequence and *urn process*, a concept studied, for example, in Hill, Lane and Sudderth (1987). On the other hand, in Section 7 we will develop a different approach to Hoeffding decomposability, in order to provide a partial characterization of Hoeffding decomposable exchangeable sequences with values in a finite set with more than two elements. This characterization is not as exhaustive as in the two-color case. However, we will be able to prove that Pólya urns are the only Hoeffding decomposable sequences among the class of exchangeable sequences whose directing measure is obtained by normalizing vectors of infinitely divisible (positive) random variables. This follows from some computations contained in James, Lijoi and Prünster (2006). We stress by now that, when specialized to the case of two-color sequences, certain results established in Section 7 (for instance, Theorem 10) may be used to deduce alternative proofs of some of the findings of the preceding sections. We also believe that it is crucial to keep the treatment of the two-color case separate, since the techniques used in this framework (which are quite difficult to reproduce in the general case) allow to give a new implicit combinatorial characterization of the system of predictive probabilities associated with two-color Pólya urns (see, e.g., Proposition 4 of Section 4), as well as to establish transparent connections with the classic results by Diaconis and Ylvisaker (1979) [see part (II) of Section 6].

Before stating our main theorem in the two-color case (see Section 3), we collect in Section 2 some basic definitions and facts concerning Hoeffding decompositions and exchangeable sequences. We focus on sequences with values in a finite set. The reader is referred to Peccati (2004) for any unexplained concept or notation, as well as for general statements concerning sequences with values in arbitrary Polish spaces.

2. Preliminaries. Let D be a finite set, and consider an infinite exchangeable sequence $\mathbf{X}_{[1,\infty)} = \{X_n : n \geq 1\}$ of D-valued random variables, defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$ such that $\mathscr{F} = \sigma(\mathbf{X}_{[1,\infty)})$. We recall that, according to the well-known *de Finetti theorem* [see, e.g., Aldous (1985)], the assumption of exchangeability is equivalent to saying that $\mathbf{X}_{[1,\infty)}$ is a *mixture* of i.i.d. sequences with values in D.

For every $n \geq 1$ and every $1 \leq u \leq n$, we write $[n] = \{1, \dots, n\}$ and $[u, n] = \{u, u+1, \dots, n\}$, and set $\mathbf{X}_{[u,n]} \triangleq (X_u, X_{u+1}, \dots, X_n)$ and $\mathbf{X}_{[n]} \triangleq \mathbf{X}_{[1,n]} = (X_1, X_2, \dots, X_n)$. For every $n \geq 2$, we define the sequence of spaces

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

generated by symmetric U-statistics of increasing order, as follows: $SU_0(\mathbf{X}_{[n]}) \triangleq \Re$ and, for $k = 1, \dots, n$, $SU_k(\mathbf{X}_{[n]})$ is the collection of all random variables of the type

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

where φ is a real-valued symmetric function from D^k to $\Re$. A random variable such as F in (1) is called a *U-statistic with symmetric kernel of order k*. It is easily seen that (since each $\mathbf{X}_{[k]}$ is an infinitely extendible exchangeable vector) the kernel φ appearing in (1) is unique, in the sense that if φ' is another symmetric kernel satisfying (1), then $\varphi(\mathbf{X}_{[k]}) = \varphi'(\mathbf{X}_{[k]})$, a.s.-$\mathbb{P}$. The following facts are immediately checked: (i) for every $k = 0, \dots, n$, $SU_k(\mathbf{X}_{[n]})$ is a vector space, (ii) $SU_{k-1}(\mathbf{X}_{[n]}) \subset SU_k(\mathbf{X}_{[n]})$, (iii) $SU_n(\mathbf{X}_{[n]}) = L_s(\mathbf{X}_{[n]})$, where (for $n \geq 1$) $L_s(\mathbf{X}_{[n]})$ is defined as the set of all random variables of the type $T(\mathbf{X}_{[n]}) = T(X_1, \dots, X_n)$, where T is a *symmetric function* from D^n to $\Re$. The class of all symmetric functions, from D^n to $\Re$, will be denoted by $\mathcal{S}(D^n)$. Note that $L_s(\mathbf{X}_{[n]})$ is a Hilbert space with respect to the inner product $\langle T_1, T_2 \rangle \triangleq \mathbb{E}[T_1(\mathbf{X}_{[n]}) T_2(\mathbf{X}_{[n]})]$, so that each $SU_k(\mathbf{X}_{[n]})$ is a closed subspace of $L_s(\mathbf{X}_{[n]})$. Finally, the sequence of *symmetric Hoeffding spaces* $\{SH_k(\mathbf{X}_{[n]}) : k = 0, \dots, n\}$ associated with $\mathbf{X}_{[n]}$ is defined as $SH_0(\mathbf{X}_{[n]}) \triangleq SU_0(\mathbf{X}_{[n]}) = \Re$, and

$$SH_k(\mathbf{X}_{[n]}) \triangleq SU_k(\mathbf{X}_{[n]}) \cap SU_{k-1}(\mathbf{X}_{[n]})^{\perp}, \qquad k = 1, \dots, n, \tag{2}$$

where all orthogonals (here and in the sequel) are taken in $L_s(\mathbf{X}_{[n]})$. Observe that $SH_k(\mathbf{X}_{[n]}) \subset SU_k(\mathbf{X}_{[n]})$ for every k, so that each $F \in SH_k(\mathbf{X}_{[n]})$ has necessarily the form (1) for some well-chosen symmetric kernel φ. Moreover, since

$SU_n(\mathbf{X}_{[n]}) = L_s(\mathbf{X}_{[n]})$, one has the following orthogonal decomposition:

$$L_s(\mathbf{X}_{[n]}) = \bigoplus_{k=0}^{n} SH_k(\mathbf{X}_{[n]}), \tag{3}$$

where "$\bigoplus$" stands for an orthogonal sum. In particular, (3) implies that every symmetric random variable $T(\mathbf{X}_{[n]}) \in L_s(\mathbf{X}_{[n]})$ admits a unique representation as a noncorrelated sum of $n+1$ terms, with the kth summand ($k = 0, \dots, n$) equal to an element of $SH_k(\mathbf{X}_{[n]})$.

The next definition, which is essentially borrowed from Peccati (2004), formalizes the notion of "Hoeffding decomposability" evoked at the beginning of the section.

DEFINITION A. The random sequence $\mathbf{X}_{[1,\infty)}$ is *Hoeffding decomposable* if, for every $n \ge 2$ and every $k = 1, \dots, n$, the following double implication holds: $F \in SH_k(\mathbf{X}_{[n]})$ if, and only if, the kernel φ appearing in its representation (1) satisfies the degeneracy condition

$$\mathbb{E}[\varphi(\mathbf{X}_{[k]}) \mid \mathbf{X}_{[2,k]}] = 0, \qquad \text{a.s.-}\mathbb{P}. \tag{4}$$

When a U-statistic F as in (1) is such that φ verifies (4), one says that F is a *completely degenerate* symmetric U-statistic of order k, and that φ is a *completely degenerate symmetric kernel* of order k.

For instance, when $k = 3$, one has $\mathbf{X}_{[2,k]} = (X_2, X_3)$, and condition (4) becomes $\mathbb{E}[\varphi(X_1, X_2, X_3) \mid X_2, X_3] = 0$. Of course, by exchangeability, (4) holds if, and only if, $\mathbb{E}[\varphi(\mathbf{X}_{[k]}) \mid \mathbf{X}_{[k-1]}] = 0$, a.s.-$\mathbb{P}$.

For every infinite nondeterministic exchangeable sequence $\mathbf{X}_{[1,\infty)}$ (not necessarily Hoeffding decomposable) and every $k \ge 1$, the class of all kernels $\varphi : D^k \mapsto \Re$, such that (4) is verified, is denoted by $\Xi_k(\mathbf{X}_{[1,\infty)})$. Observe that $\Xi_k(\mathbf{X}_{[1,\infty)})$, $k \ge 1$, is a vector space. Also, by definition, for every Hoeffding decomposable sequence $\mathbf{X}_{[1,\infty)}$ and for every $1 \le k \le n$, one has necessarily that $\dim SH_k(\mathbf{X}_{[n]}) = \dim \Xi_k(\mathbf{X}_{[1,\infty)})$.

It is well known [see, e.g., Hoeffding (1948), Hajek (1968) or Karlin and Rinott (1982)] that each i.i.d. sequence is decomposable in the sense of Definition A. In Peccati (2004), the second author established a complete characterization of Hoeffding decomposable sequences (with values in arbitrary Polish spaces), in terms of *weak independence*. To introduce this concept, we need some more notation. Fix $n \ge 2$, and consider a symmetric function $T \in \mathcal{S}(D^n)$. We define the function $[T]_{n,n-1}^{(n-1)}$ as the unique application from D^{n-1} to $\Re$ such that

$$[T]_{n,n-1}^{(n-1)}(\mathbf{X}_{[2,n]}) = \mathbb{E}(T(\mathbf{X}_{[n]}) \mid \mathbf{X}_{[2,n]}), \qquad \text{a.s.-}\mathbb{P}. \tag{5}$$

For instance, if $n = 2$, then $\mathbf{X}_{[2]} = (X_1, X_2)$, $\mathbf{X}_{[2,2]} = X_2$ and $[T]_{2,1}^{(1)}(X_2) = \mathbb{E}(T(X_1, X_2) \mid X_2)$. Note that the exchangeability assumption and the symmetry of T imply that the application $D^{n-1} \mapsto \Re : \mathbf{x} \mapsto [T]_{n,n-1}^{(n-1)}(\mathbf{x})$ is symmetric. Also, with this notation, $T \in \Xi_n(\mathbf{X}_{[1,\infty)})$ if, and only if, $[T]_{n,n-1}^{(n-1)}(\mathbf{X}_{[2,n]}) = 0$, a.s.-$\mathbb{P}$.

Analogously, for $u = 2, \dots, n$ we define the function $[T]_{n,n-1}^{(n-u)} : D^{n-1} \mapsto \Re$ through the relation

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

To understand our notation, observe that, for $u = 2, \dots, n$, the two sets $[n]$ and $[u+1, u+n-1]$ have exactly $n-u$ elements in common. For instance, if $n = 3$ and $u = 2$, then $[u+1, u+n-1] = \{3, 4\}$, and $[T]_{3,2}^{(1)}(X_3, X_4) = \mathbb{E}(T(X_1, X_2, X_3) \mid X_3, X_4)$. Again, exchangeability and symmetry yield that the function $\mathbf{x} \mapsto [T]_{n,n-1}^{(0)}(\mathbf{x})$ (corresponding to the case $u = n$) is symmetric on D^{n-1}. On the other hand, for $u = 2, \dots, n-1$, the application $(x_1, \dots, x_{n-1}) \mapsto [T]_{n,n-1}^{(n-u)}(x_1, \dots, x_{n-1})$ is (separately) symmetric in the variables $(x_1, \dots, x_{n-u})$ and $(x_{n-u+1}, \dots, x_{n-1})$, and not necessarily symmetric as a function on D^{n-1}.

From now on, the symbol $\mathfrak{S}_n$ ($n \geq 1$) stands for the group of permutations of the set $[n] = \{1, \dots, n\}$. Given a vector $\mathbf{x}_n = (x_1, \dots, x_n) \in D^n$ and a permutation $\pi \in \mathfrak{S}_n$, we denote by $\mathbf{x}_{\pi(n)}$ the action of π on $\mathbf{x}_n$, that is, $\mathbf{x}_{\pi(n)} = (x_{\pi(1)}, \dots, x_{\pi(n)})$. Given a function $f : D^n \to \Re$, we write $\widetilde{f}$ for its canonical symmetrization, that is, for every $\mathbf{x}_n \in D^n$

$$\widetilde{f}(\mathbf{x}_n) = \frac{1}{n!} \sum_{\pi \in \mathfrak{S}_n} f(\mathbf{x}_{\pi(n)}).$$

In particular, for $u = 2, \dots, n-1$, the symbol $\widetilde{[T]}_{n,n-1}^{(n-u)}$ indicates the symmetrization of the function $[T]_{n,n-1}^{(n-u)}$ defined above. Finally, for $u = 2, \dots, n$, we set

(7) $$\widetilde{\Xi}_{n,n-u}(\mathbf{X}_{[1,\infty)}) \triangleq \{T \in \mathcal{S}(D^n) : \widetilde{[T]}_{n,n-1}^{(n-u)}(\mathbf{X}_{[u+1,u+n-1]}) = 0, \text{ a.s.-}\mathbb{P}\}$$

[recall that $\mathcal{S}(D^n)$ denotes the class of symmetric functions on D^n]. Note that, by exchangeability, $\widetilde{[T]}_{n,n-1}^{(n-u)}(\mathbf{X}_{[u+1,u+n-1]}) = 0$, a.s.-$\mathbb{P}$, if, and only if, $\widetilde{[T]}_{n,n-1}^{(n-u)}(\mathbf{X}_{[n-1]}) = 0$, a.s.-$\mathbb{P}$. The following technical definition is taken from Peccati (2004).

DEFINITION B. The exchangeable sequence $\mathbf{X}_{[1,\infty)}$ is *weakly independent* if, for every $n \geq 2$,

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

In other words, $\mathbf{X}_{[1,\infty)}$ is weakly independent if, for every $n \geq 2$ and every $T \in \mathcal{S}(D^n)$, the following implication holds: if $[T]_{n,n-1}^{(n-1)}(\mathbf{X}_{[n-1]}) = 0$, then $\widetilde{[T]}_{n,n-1}^{(n-u)}(\mathbf{X}_{[n-1]}) = 0$ for every $u = 2, \dots, n$.

The next theorem, which is one of the main results of Peccati (2004), shows that the notions of weak independence and Hoeffding decomposability are equivalent for infinite exchangeable sequences.

THEOREM 0 [Peccati (2004), Theorem 6]. *Suppose that the infinite exchangeable sequence $\mathbf{X}_{[1,\infty)}$ is such that, for every $n \geq 2$,*

$$SH_k(\mathbf{X}_{[n]}) \neq \{0\}, \qquad \forall k = 1, \dots, n. \tag{9}$$

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

Condition (9) excludes, for instance, the case $X_n = X_1$, for each $n \geq 1$. Note that Theorem 0 also holds for exchangeable sequences with values in general Polish spaces. In Peccati (2004) Theorem 0 has been used to show the following two facts:

(F1) There are infinite exchangeable sequences which are Hoeffding decomposable and not i.i.d., as, for instance, the *generalized urn sequences* analyzed in Section 5 of Peccati (2004).

(F2) There exist infinite exchangeable sequences that *are not* Hoeffding decomposable. For instance, one can consider a $\{0, 1\}$-valued exchangeable sequence $\mathbf{X}^Y_{[1,\infty)}$ such that, conditioned on the realization of a random variable Y uniformly distributed on $(0, \varepsilon)$ $(0 < \varepsilon < 1)$, $\mathbf{X}^Y_{[1,\infty)}$ is composed of independent Bernoulli trials with random parameter Y. See Peccati (2004), pages 1807–1808, for more details.

Although the combination of Theorem 0, (F1) and (F2) gives several insights into the structure of Hoeffding decomposable sequences, the analysis contained in Peccati (2004) left open a crucial question: *can one characterize the laws of Hoeffding decomposable sequences, in terms of their de Finetti representation as mixtures of i.i.d. sequences?* In the following sections, we will provide a complete answer when $D = \{0, 1\}$, by proving that in this case the class of Hoeffding decomposable sequences contains exclusively i.i.d. and Pólya sequences. The extension of our results to spaces D with more than two elements is discussed in Section 7.

3. Main results for two-color sequences.

For the rest of this section, we will focus on the case $D = \{0, 1\}$. According to the de Finetti theorem, in this case the exchangeability of $\mathbf{X}_{[1,\infty)} = \{X_n : n \geq 1\}$ yields the existence of a (unique) probability measure γ on $[0, 1]$ such that, for every $n \geq 1$ and every vector $(j_1, \dots, j_n) \in \{0, 1\}^n$,

$$\mathbb{P}\{X_1 = j_1, \dots, X_n = j_n\} = \int_{[0,1]} \theta^{\sum_k j_k} (1 - \theta)^{n - \sum_k j_k} \gamma(d\theta). \tag{10}$$

The measure γ, appearing in (10), is called the *de Finetti measure* associated with $\mathbf{X}_{[1,\infty)}$. In what follows, we shall systematically suppose that $\mathbf{X}_{[1,\infty)}$ is *nondeterministic*, that is, that the support of the measure γ *is not contained* in $\{0\} \cup \{1\}$. The choice of the term "nondeterministic" is inspired by Hill, Lane and Sudderth (1987), where the adjective *deterministic* is used to describe exchangeable sequences whose de Finetti measure γ has support contained in $\{0\} \cup \{1\}$. We stress that a deterministic sequence, in the sense of Hill, Lane and Sudderth (1987), can actually be random. Define indeed $\gamma^* = p\delta_1 + (1-p)\delta_0$, where $p \in (0,1)$ and δ_a stands for the Dirac mass at a. Then, an exchangeable sequence $\mathbf{X}_{[1,\infty)}$ = $(X_1, X_2, \ldots)$ with de Finetti measure γ^* is such that: (i) $\mathbf{X}_{[1,\infty)}$ is deterministic in the sense of Hill, Lane and Sudderth (1987), (ii) $X_n = X_1$ for $n \geq 1$, and (iii) $\mathbb{P}(X_1 = 1) = p = 1 - \mathbb{P}(X_1 = 0)$. Moreover, when $D = \{0, 1\}$, condition (9) holds if, and only if, $\mathbf{X}_{[1,\infty)}$ is nondeterministic. To prove this last claim we only need to show that $SH_k(\mathbf{X}_{[n]}) = \{0\}$ for some n, k if, and only if, $\mathbf{X}_{[1,\infty)}$ is deterministic. Now, on the one hand, it is immediately seen that, if $\mathbf{X}_{[1,\infty)}$ is deterministic, then $SH_k(\mathbf{X}_{[n]}) = 0$ for any integers n, k such that $2 \leq k \leq n$. On the other hand, if $\mathbf{X}_{[1,\infty)}$ is nondeterministic, then, for every n, k such that $n \geq 2$ and $0 \leq k \leq n$, the vector space $SU_k(\mathbf{X}_{[n]})$ has exactly dimension $k+1$, so that [since $SU_{k-1}(\mathbf{X}_{[n]}) \subseteq SU_k(\mathbf{X}_{[n]})$] the space $SH_k(\mathbf{X}_{[n]}) = SU_k(\mathbf{X}_{[n]}) \cap SU_{k-1}(\mathbf{X}_{[n]})^{\perp}$ must necessarily have dimension 1.

DEFINITION C. The exchangeable sequence $\mathbf{X}_{[1,\infty)} = \{X_n : n \geq 1\}$ of $\{0, 1\}$-valued random variables is called a *two-color Pólya sequence* if there exist two real numbers $\alpha, \beta > 0$ such that

$$\gamma(d\theta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1}\, d\theta, \tag{11}$$

where γ is the de Finetti measure associated with $\mathbf{X}_{[1,\infty)}$ through formula (10), and

$$B(\alpha, \beta) = \int_0^1 \theta^{\alpha-1}(1-\theta)^{\beta-1}\, d\theta$$

is the usual Beta function. The numbers α and β are the *parameters* of the Pólya sequence $\mathbf{X}_{[1,\infty)}$. A random variable ξ, with values in $[0, 1]$ and with law γ as in (11), is called a *Beta random variable* of parameters α and β.

Classic references for the theory of Pólya sequences are Blackwell (1973) and Blackwell and MacQueen (1973) [see also Pitman (2006, 1996) for a state of the art review]. Thanks to Peccati (2004), Corollary 9, we already know that Pólya and i.i.d. sequences are Hoeffding decomposable. The next result, which is one of the main achievements of our paper, shows that those are the only exchangeable and Hoeffding decomposable sequences with values in $\{0, 1\}$. The proof is deferred to Section 5.

THEOREM 1. *Let* $\mathbf{X}_{[1,\infty)}$ *be a nondeterministic infinite exchangeable sequence of* $\{0,1\}$*-valued random variables. Then, the following two assertions are equivalent*:

1. $\mathbf{X}_{[1,\infty)}$ *is Hoeffding decomposable*;
2. $\mathbf{X}_{[1,\infty)}$ *is either an i.i.d. sequence or a two-color Pólya sequence.*

In Section 6 we will discuss some connections between Theorem 1 and the concept of *urn process,* as defined in Hill, Lane and Sudderth (1987).

REMARK. We state two projection formulae, concerning, respectively, i.i.d. and two-color Pólya sequences.

(I) Let $\mathbf{X}_{[1,\infty)}$ be an i.i.d. sequence with values in $\{0,1\}$, and fix $n \geq 2$ and $T \in L_s(\mathbf{X}_{[n]})$. Then, for $k = 1, \dots, n$, the projection of T on the kth Hoeffding space $SH_k(\mathbf{X}_{[n]})$, denoted by $\pi[T, SH_k]$, is

$$\pi[T, SH_k] = \sum_{a=1}^{k} (-1)^{k-a} \sum_{1 \leq j_1 < \cdots < j_a \leq n} [T - \mathbb{E}(T)]_{n,a}^{(a)}(X_{j_1}, \dots, X_{j_a}), \tag{12}$$

where, for $a = 1, \dots, k$ and $1 \leq j_1 < \cdots < j_a \leq n$,

$$[T - \mathbb{E}(T)]_{n,a}^{(a)}(X_{j_1}, \dots, X_{j_a}) = \mathbb{E}\big[T(\mathbf{X}_{[n]}) - \mathbb{E}(T) | X_{j_1}, \dots, X_{j_a}\big].$$

Formula (12) is classic [see, e.g., Hoeffding (1948) or Vitale (1992)], and can be easily deduced by an application of the inclusion–exclusion principle.

(II) Let $\mathbf{X}_{[1,\infty)}$ be a Pólya sequence of parameters $\alpha, \beta > 0$, and fix $n \geq 2$ and $T \in L_s(\mathbf{X}_{[n]})$. Then, for $k = 1, \dots, n$, the projection of T on the kth Hoeffding space associated with $\mathbf{X}_{[n]}$ is of the form

$$\pi[T, SH_k] = \sum_{a=1}^{k} \theta_n^{(k,a)} \sum_{1 \leq j_1 < \cdots < j_a \leq n} [T - \mathbb{E}(T)]_{n,a}^{(a)}(X_{j_1}, \dots, X_{j_a}).$$

The explicit formulae describing the real coefficients $\theta_n^{(k,a)}$ are given recursively in Peccati (2004), formula (24). For instance, when $n = 3$, then

$$\begin{cases} \theta_3^{(1,1)} = \dfrac{\alpha + \beta + 1}{\alpha + \beta + 2}, \\ \theta_3^{(2,1)} = -\dfrac{(\alpha + \beta + 1)(\alpha + \beta + 4)}{(\alpha + \beta + 3)(\alpha + \beta + 2)} - \dfrac{\alpha + \beta + 1}{\alpha + \beta + 2}, \\ \theta_3^{(2,2)} = \dfrac{\alpha + \beta + 4}{\alpha + \beta + 2}. \end{cases}$$

The rest of the paper is organized as follows: in Section 4 we collect several technical results, leading to a new characterization of Hoeffding decomposability for $\{0,1\}$-valued sequences in terms of conditional probabilities (see Proposition 4 below); the proof of Theorem 1 is contained in Section 5; in Section 6, a brief discussion is presented, relating Theorem 1 with several notions associated with $\{0, 1\}$-valued exchangeable sequences; Section 7 deals with Hoeffding decomposable sequences with values in a general finite set.

4. Ancillary lemmas. From now on, $\mathbf{X}_{[1,\infty)} = \{X_n : n \geq 1\}$ will be a nondeterministic exchangeable sequence with values in $D = \{0, 1\}$. For $n \geq 2$, we write $\mathcal{S}(\{0, 1\}^n)$ to indicate the vector space of symmetric functions on $\{0, 1\}^n$. By exchangeability, we have of course that

$$\mathbb{P}(\mathbf{X}_{[n]} = \mathbf{x}_n) = \mathbb{P}(\mathbf{X}_{[n]} = \mathbf{x}_{\pi(n)}) \qquad \forall n \geq 2,\ \forall \pi \in \mathfrak{S}_n,$$

yielding that, for $n \geq 2$, the value of the probability $\mathbb{P}(\mathbf{X}_{[n]} = \mathbf{x}_n)$ depends exclusively on n and on the number of zeros contained in the vector $\mathbf{x}_n$. For $n \geq 1$ and $j = 0, \dots, n$, we shall denote by $\mathbb{P}_n(0^{(j)})$ the *common value* taken by the quantity $\mathbb{P}(\mathbf{X}_{[n]} = \mathbf{x}_n)$ for all $\mathbf{x}_n = (x_1, \dots, x_n) \in \{0, 1\}^n$ such that $\mathbf{x}_n$ contains exactly j zeros. For instance, when $n = 3$ and $j = 1$, one has that $\mathbb{P}_3(0^{(1)}) = \mathbb{P}(\mathbf{X}_{[3]} = (0, 1, 1)) = \mathbb{P}(\mathbf{X}_{[3]} = (1, 0, 1)) = \mathbb{P}(\mathbf{X}_{[3]} = (1, 1, 0))$. Note that, since $\mathbf{X}_{[1,\infty)}$ is nondeterministic, $\mathbb{P}_n(0^{(j)}) > 0$ for every $n \geq 1$ and every $j = 0, \dots, n$. Analogously, for every $n \geq 2$, every $j = 0, \dots, n$, and every symmetric function $\varphi \in \mathcal{S}(\{0, 1\}^n)$, we will write $\varphi(0^{(j)})$ to indicate the common value taken by $\varphi(\mathbf{x}_n)$ for all $\mathbf{x}_n \in \{0, 1\}^n$ containing exactly j zeros.

The following result gives a complete characterization of the spaces

$$\Xi_n(\mathbf{X}_{[1,\infty)}), \qquad n \geq 2,$$

defined through relation (4) (note that, to define the spaces Ξ_n we do not need $\mathbf{X}_{[1,\infty)}$ to be Hoeffding decomposable).

LEMMA 2. *With the assumptions and notation of this section, the set* $\Xi_n(\mathbf{X}_{[1,\infty)})$ *is the one-dimensional vector space spanned by the symmetric kernel* $\varphi_n^{(0)} : \{0, 1\}^n \mapsto \Re$ *defined by*

$$\varphi_n^{(0)}(0^{(k)}) = (-1)^k \frac{\mathbb{P}_n(0^{(0)})}{\mathbb{P}_n(0^{(k)})}, \qquad k = 0, \dots, n. \tag{13}$$

PROOF. Consider $\varphi_n \in \Xi_n(\mathbf{X}_{[1,\infty)})$. By the definition of $\Xi_n(\mathbf{X}_{[1,\infty)})$, for any fixed $j = 0, \dots, n-1$ and any fixed $\mathbf{x}_{n-1} \in \{0, 1\}^{n-1}$ such that $\sum_{i=1}^{n-1}(1 - x_i) = j$, we have

$$\begin{aligned} 0 &= \mathbb{E}[\varphi_n(\mathbf{X}_{[n]}) \mid \mathbf{X}_{[2,n]} = \mathbf{x}_{n-1}] \\ &= \varphi_n(0^{(j+1)}) \frac{\mathbb{P}_n(0^{(j+1)})}{\mathbb{P}_{n-1}(0^{(j)})} + \varphi_n(0^{(j)}) \frac{\mathbb{P}_n(0^{(j)})}{\mathbb{P}_{n-1}(0^{(j)})}, \end{aligned}$$

and therefore $\varphi_n(0^{(j+1)}) = -(\mathbb{P}_n(0^{(j)})/\mathbb{P}_n(0^{(j+1)})) \times \varphi_n(0^{(j)})$. Arguing recursively on j, one has

$$\varphi_n(0^{(j+1)}) = (-1)^{j+1}\frac{\mathbb{P}_n(0^{(0)})}{\mathbb{P}_n(0^{(j+1)})}\varphi_n(0^{(0)}), \qquad j = 0, \dots, n-1, \tag{14}$$

showing that any symmetric kernel $\varphi_n \in \Xi_n(\mathbf{X}_{[1,\infty)})$ is completely determined by the quantity $\varphi_n(0^{(0)})$. Now define a kernel $\varphi_n^{(0)} \in \Xi_n(\mathbf{X}_{[1,\infty)})$ by using (14) and by setting $\varphi_n^{(0)}(0^{(0)}) = \mathbb{P}_n(0^{(0)})/\mathbb{P}_n(0^{(0)}) = 1$. It is easily seen that $\varphi_n^{(0)}$ must coincide with the function defined in (13). To conclude, consider another element φ_n of $\Xi_n(\mathbf{X}_{[1,\infty)})$. Since there exists a constant $K \in \Re$ such that $\varphi_n(0^{(0)}) = K = K\varphi_n^{(0)}(0^{(0)})$, and since φ_n has to satisfy (14), we deduce that $\varphi_n = K\varphi_n^{(0)}$, thus completing the proof. $\square$

The following result will prove very useful.

LEMMA 3. *Fix $m \geq 2$ and $v \in \{1, \dots, m-1\}$ and let the application*

$$f_{v,m-v} : \{0,1\}^m \mapsto \Re : (x_1, \dots, x_m) \mapsto f(x_1, \dots, x_m)$$

be separately symmetric in the variables $(x_1, \dots, x_v)$ and $(x_{v+1}, \dots, x_m)$ (and not necessarily symmetric as a function on $\{0,1\}^m$). Then, for any $\mathbf{x}_m = (x_1, \dots, x_m) \in \{0,1\}^m$ such that $\sum_{j=1}^m (1 - x_j) = z$ for some $z = 0, \dots, m$, the canonical symmetrization of $f_{v,m-v}$, computed at $\mathbf{x}_m$, reduces to

$$\widetilde{f}_{v,m-v}(\mathbf{x}_m) = \frac{\sum_{k=0\vee(z-(m-v))}^{z\wedge v} \binom{v}{k}\binom{m-v}{z-k} f_{v,m-v}(0^{(k)}, 0^{(z-k)})}{\sum_{k=0\vee(z-(m-v))}^{z\wedge v} \binom{v}{k}\binom{m-v}{z-k}}, \tag{15}$$

where $f_{v,m-v}(0^{(k)}, 0^{(z-k)})$ denotes the common value of $f_{v,m-v}(\mathbf{y}_m)$ when $\mathbf{y}_m = (y_1, \dots, y_m)$ is such that the vector $(y_1, \dots, y_v)$ contains exactly k zeros, and the vector $(y_{v+1}, \dots, y_m)$ contains exactly $(z-k)$ zeros.

As a consequence, $\widetilde{f}_{v,m-v}(\mathbf{x}_m) = 0$ for every $\mathbf{x}_m \in \{0,1\}^m$ if, and only if, for all $z = 0, \dots, m$,

$$\sum_{k=0\vee(z-(m-v))}^{z\wedge v} \binom{v}{k}\binom{m-v}{z-k} f_{v,m-v}(0^{(k)}, 0^{(z-k)}) = 0. \tag{16}$$

PROOF. Fix $\mathbf{x}_m \in \{0,1\}^m$ such that $\sum_{j=1}^m (1 - x_j) = z$ for some $z = 0, \dots, m$. Without loss of generality, we can assume

$$\mathbf{x}_m = (\underbrace{0, 0, \dots, 0}_{z \text{ times}}, \underbrace{1, 1, \dots, 1}_{m-z \text{ times}}).$$

Observe that, for all $k = \max\{0, z-(m-v)\}, \dots, \min\{z, v\}$, there are exactly $z!(m-z)!\binom{v}{k}\binom{m-v}{z-k}$ permutations $\pi \in \mathfrak{S}_m$ such that $\sum_{j=1}^v (1 - x_{\pi(j)}) = k$ and

$\sum_{j=v+1}^{m}(1-x_{\pi(j)})=z-k$. The set of all such permutations will be denoted by $\mathfrak{S}_m^{(k)}$. It is immediately seen that

$$\widetilde{f}_{v,m-v}(\mathbf{x}_m)=\frac{1}{m!}\sum_{k=0\vee(z-(m-v))}^{z\wedge v}\sum_{\pi\in\mathfrak{S}_m^{(k)}}f_{v,m-v}\big(0^{(k)},0^{(z-k)}\big)$$
$$=\frac{1}{m!}\sum_{k=0\vee(z-(m-v))}^{z\wedge v}f_{v,m-v}\big(0^{(k)},0^{(z-k)}\big)\times\operatorname{card}\big(\mathfrak{S}_m^{(k)}\big).$$

Formula (15) now follows by observing that

$$\frac{m!}{z!(m-z)!}=\binom{m}{z}=\sum_{k=0\vee(z-(m-v))}^{z\wedge v}\binom{v}{k}\binom{m-v}{z-k}.$$

The last assertion in the statement of this lemma is an easy consequence of (15). $\square$

We shall conclude the section by obtaining a full characterization of $\{0,1\}$-valued Hoeffding decomposable sequences (stated in Proposition 4 below).

To do this, recall that, for any symmetric $\varphi:\{0,1\}^n\mapsto\Re$, every $u=2,\dots,n$ and every $\mathbf{x}_{n-1}\in\{0,1\}^{n-1}$,

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

Observe that the function $[\varphi]_{n,n-1}^{(n-u)}:\{0,1\}^{n-1}\mapsto\Re$ clearly meets the symmetry properties of Lemma 3 with $m=n-1$ and $v=n-u$. Now fix $z\in\{0,\dots,n-1\}$, and suppose that $\mathbf{x}_{n-1}\in\{0,1\}^{n-1}$ is such that $\sum_{j=1}^{n-1}(1-x_j)=z$ and $\sum_{j=1}^{n-u}(1-x_j)=k$. Then,

$$[\varphi]_{n,n-1}^{(n-u)}(\mathbf{x}_{n-1})=\sum_{m=0}^{u}\binom{u}{m}\varphi\big(0^{(k+m)}\big)\frac{\mathbb{P}_{n-1+u}(0^{(z+m)})}{\mathbb{P}_{n-1}(0^{(z)})}. \tag{17}$$

By applying (16) in the case $m=n-1$ and $v=n-u$, we deduce that $\widetilde{[\varphi]}_{n,n-1}^{(n-u)}(0^{(z)})=0$ if, and only if,

$$\sum_{k=0\vee(z-(u-1))}^{z\wedge(n-u)}\binom{n-u}{k}\binom{u-1}{z-k}[\varphi]_{n,n-1}^{(n-u)}\big(0^{(k)},0^{(z-k)}\big)=0, \tag{18}$$

where the notation $\widetilde{[\varphi]}_{n,n-1}^{(n-u)}(0^{(z)})$ and $[\varphi]_{n,n-1}^{(n-u)}(0^{(k)},0^{(z-k)})$ has been introduced to indicate the value of $\widetilde{[\varphi]}_{n,n-1}^{(n-u)}(\mathbf{y}_{n-1})$ (resp., $[\varphi]_{n,n-1}^{(n-u)}(\mathbf{w}_{n-1})$), where $\mathbf{y}_{n-1}=(y_1,\dots,y_{n-1})\in\{0,1\}^{n-1}$ is any vector containing exactly z zeros [resp., $\mathbf{w}_{n-1}=(w_1,\dots,w_{n-1})\in\{0,1\}^{n-1}$ is any vector containing exactly k zeros in $(w_1,\dots,w_{n-u})$ and $z-k$ zeros in $(w_{n-u+1},\dots,w_{n-1})$].

Now recall that, by Theorem 0, $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable if, and only if, it is weakly independent, and that $\mathbf{X}_{[1,\infty)}$ is weakly independent if, and only if, for all $n \geq 2$ and for any $\varphi \in \Xi_n(\mathbf{X}_{[1,\infty)})$, one has $\varphi \in \widetilde{\Xi}_{n,u}(\mathbf{X}_{[1,\infty)})$ for all $u = 2, \dots, n$. By Lemma 2, we deduce that the sequence $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable if, and only if, for every $n \geq 2$ and every $u = 2, \dots, n$, $\varphi_n^{(0)} \in \widetilde{\Xi}_{n,u}(\mathbf{X}_{[1,\infty)})$, where $\varphi_n^{(0)}$ is defined in (13). By (18), this last relation is true if, and only if, for every $n \geq 2$, every $z = 0, \dots, n-1$ and every $u = 2, \dots, n$,

$$\sum_{k=0\vee(z-(u-1))}^{z\wedge(n-u)} \binom{n-u}{k}\binom{u-1}{z-k}[\varphi_n^{(0)}]_{n,n-1}^{(n-u)}(0^{(k)}, 0^{(z-k)}) = 0. \tag{19}$$

Substituting (13) and (17) in (18), we obtain that (19) is true if and only if

$$\begin{aligned} 0 = \frac{\mathbb{P}_n(0^{(0)})}{\mathbb{P}_{n-1}(0^{(z)})} &\sum_{k=0\vee(z-(u-1))}^{z\wedge(n-u)} (-1)^k \binom{n-u}{k}\binom{u-1}{z-k} \\ &\times \sum_{m=0}^{u} (-1)^m \binom{u}{m} \frac{\mathbb{P}_{n-1+u}(0^{(m+z)})}{\mathbb{P}_n(0^{(m+k)})}. \end{aligned} \tag{20}$$

Note that

$$\frac{\mathbb{P}_{n-1+u}(0^{(m+z)})}{\mathbb{P}_n(0^{(m+k)})} = \frac{1}{\binom{u-1}{z-k}} \mathbb{P}_{n+u-1}^{n}(0^{(m+z)} \mid 0^{(m+k)}), \tag{21}$$

where $\mathbb{P}_{n+u-1}^{n}(0^{(m+z)} \mid 0^{(m+k)})$ denotes the conditional probability that the vector $\mathbf{X}_{[n+u-1]}$ contains exactly $m+z$ zeros, given that the subvector $\mathbf{X}_{[n]}$ contains exactly $m+k$ zeros.

REMARK. For every $n \geq 1$, $0 \leq a \leq b$, every $v \geq 1$, the quantity $\mathbb{P}_{n+v}^{n}(0^{(b)} \mid 0^{(a)})$ is equal to

$$\mathbb{P}(\mathbf{X}_{[n+1,n+v]} \text{ contains exactly } b-a \text{ zeros} \mid \mathbf{X}_{[n]} \text{ contains exactly } a \text{ zeros}).$$

By plugging (21) into (20), we obtain the announced characterization of weak independence.

PROPOSITION 4. *Let $\mathbf{X}_{[1,\infty)}$ be a nondeterministic infinite sequence of exchangeable $\{0,1\}$-valued random variables. For $\mathbf{X}_{[1,\infty)}$ to be Hoeffding decomposable, it is necessary and sufficient that, for every $n \geq 2$, every $u = 2, \dots, n$ and every $z = 0, \dots, n-1$,*

$$\begin{aligned} 0 = &\sum_{k=0\vee(z-(u-1))}^{z\wedge(n-u)} (-1)^k \binom{n-u}{k} \\ &\times \sum_{m=0}^{u} (-1)^m \binom{u}{m} \mathbb{P}_{n+u-1}^{n}(0^{(m+z)} \mid 0^{(m+k)}). \end{aligned} \tag{22}$$

As shown in the next section, Proposition 4 is the key tool to prove Theorem 1.

5. Proof of Theorem 1. Here is an outline of the proof. We already know [thanks to Peccati (2004), Corollary 9] that, if $\mathbf{X}_{[1,\infty)}$ is either i.i.d. or Pólya, then it is also Hoeffding decomposable, thus proving the implication $2 \Rightarrow 1$. We shall therefore show that Hoeffding decomposability implies necessarily that $\mathbf{X}_{[1,\infty)}$ is either i.i.d. or Pólya. The proof of this last implication is divided in four steps. By using some easy remarks (Step 1) and Proposition 4, we will prove that (22) implies a universal relation linking the moments of the de Finetti measure γ underlying any Hoeffding decomposable exchangeable sequence (Step 2). After a discussion concerning the moments of Beta random variables (Step 3), we conclude the proof in Step 4.

STEP 1. We start with an easy remark. Define the two functions

$$f(x,y,z) = 2x^2z - xy^2 - x^2y, \tag{23}$$

$$g(x,y,z) = zx - 2y^2 + yz. \tag{24}$$

Then, the set

$$S \triangleq \{(x,y,z) : 0 < x < y < z \leq 1\} \tag{25}$$

does not contain any solution of the system

$$\begin{cases} f(x,y,z) = 0, \\ g(x,y,z) = 0. \end{cases} \tag{26}$$

We stress that this system can actually be solved. For instance, any vector (x, y, z) such that $x = y = z$ is a solution of (26).

STEP 2. Let $\mathbf{X}_{[1,\infty)} = \{X_n : n \geq 1\}$ be a nondeterministic exchangeable sequence with values in $\{0, 1\}$, and let γ be the de Finetti measure uniquely associated with $\mathbf{X}_{[1,\infty)}$ through formula (10). We denote by

$$\mu_n = \mu_n(\gamma) = \int_{[0,1]} \theta^n \gamma(d\theta), \qquad n \geq 0, \tag{27}$$

the sequence of moments of γ (the dependence on γ is dropped when there is no risk of confusion). We shall prove the following statement: *if* $\mathbf{X}_{[1,\infty)}$ *is Hoeffding decomposable, then*

$$\mu_{n+1} g(\mu_n, \mu_{n-1}, \mu_{n-2}) = f(\mu_n, \mu_{n-1}, \mu_{n-2}), \qquad n \geq 2, \tag{28}$$

where f and g are, respectively, defined by (23) *and* (24).

To prove (28), first recall that, due to Proposition 4, if $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable, then formula (22) must hold for every $n \geq 2$, every $u = 2, \ldots, n$ and every $z = 0, \dots, n-1$. In particular, it has to hold true for $u = 2$, that is, for all $n \ge 2$ and all $z = 0, \dots, n-1$, one must have that

$$\sum_{k=0\vee(z-1)}^{z\wedge(n-2)} (-1)^k \binom{n-2}{k} \sum_{m=0}^{2} (-1)^m \binom{2}{m} \mathbb{P}_{n+1}^{n}(0^{(m+z)} \mid 0^{(m+k)}) = 0, \tag{29}$$

for every $n \ge 2$ and every $z = 0, \dots, n-1$. Specializing formula (29) to the case $z = 0$, one obtains

$$\mathbb{P}_{n+1}^{n}(0^{(2)} \mid 0^{(2)}) - 2\mathbb{P}_{n+1}^{n}(0^{(1)} \mid 0^{(1)}) + \mathbb{P}_{n+1}^{n}(0^{(0)} \mid 0^{(0)}) = 0. \tag{30}$$

When specialized to the case $z = n-1$, relation (29) is equivalent to

$$\mathbb{P}_{n+1}^{n}(0^{(n)} \mid 0^{(n)}) - 2\mathbb{P}_{n+1}^{n}(0^{(n-1)} \mid 0^{(n-1)}) + \mathbb{P}_{n+1}^{n}(0^{(n-2)} \mid 0^{(n-2)}) = 0, \tag{31}$$

where we have used the fact that, by additivity,

$$\mathbb{P}_{n+1}^{n}(0^{(j+1)} \mid 0^{(j)}) = 1 - \mathbb{P}_{n+1}^{n}(0^{(j)} \mid 0^{(j)}), \qquad j = 0, \dots, n. \tag{32}$$

Analogously, for $1 \le z \le n-2$, (29) becomes

$$\begin{aligned} 0 = \binom{n-2}{z-1} &[\mathbb{P}_{n+1}^{n}(0^{(z+2)} \mid 0^{(z+1)}) \\ &\quad - 2\mathbb{P}_{n+1}^{n}(0^{(z+1)} \mid 0^{(z)}) + \mathbb{P}_{n+1}^{n}(0^{(z)} \mid 0^{(z-1)})] \\ - \binom{n-2}{z} &[\mathbb{P}_{n+1}^{n}(0^{(z+2)} \mid 0^{(z+2)}) \\ &\quad - 2\mathbb{P}_{n+1}^{n}(0^{(z+1)} \mid 0^{(z+1)}) + \mathbb{P}_{n+1}^{n}(0^{(z)} \mid 0^{(z)})]. \end{aligned} \tag{33}$$

Again by (32), relation (33) is equivalent to

$$\begin{aligned} 0 = -\binom{n-2}{z-1} &[\mathbb{P}_{n+1}^{n}(0^{(z+1)} \mid 0^{(z+1)}) \\ &\quad - 2\mathbb{P}_{n+1}^{n}(0^{(z)} \mid 0^{(z)}) + \mathbb{P}_{n+1}^{n}(0^{(z-1)} \mid 0^{(z-1)})] \\ - \binom{n-2}{z} &[\mathbb{P}_{n+1}^{n}(0^{(z+2)} \mid 0^{(z+2)}) \\ &\quad - 2\mathbb{P}_{n+1}^{n}(0^{(z+1)} \mid 0^{(z+1)}) + \mathbb{P}_{n+1}^{n}(0^{(z)} \mid 0^{(z)})]. \end{aligned} \tag{34}$$

Combining (30), (31) and (34), one deduces that (29) holds true if, and only if, for all $p = 0, \dots, n-2$,

$$\mathbb{P}_{n+1}^{n}(0^{(p+2)} \mid 0^{(p+2)}) - 2\mathbb{P}_{n+1}^{n}(0^{(p+1)} \mid 0^{(p+1)}) + \mathbb{P}_{n+1}^{n}(0^{(p)} \mid 0^{(p)}) = 0. \tag{35}$$

Now, for $p = 0, \dots, n$, one has that

$$\begin{aligned} &\mathbb{P}^n_{n+1}(0^{(p)} \mid 0^{(p)}) \\ &\quad = \mathbb{P}(X_{n+1} = 1 | \mathbf{X}_{[n]} \text{ contains exactly } p \text{ zeros}) \\ &\quad = \frac{\int_0^1 \theta^{n+1-p}(1-\theta)^p \gamma(d\theta)}{\int_0^1 \theta^{n-p}(1-\theta)^p \gamma(d\theta)} = \frac{\sum_{k=0}^p (-1)^k \binom{p}{k} \int_0^1 \theta^{n+1+k-p} \gamma(d\theta)}{\sum_{k=0}^p (-1)^k \binom{p}{k} \int_0^1 \theta^{n+k-p} \gamma(d\theta)} \\ &\quad = \frac{\sum_{k=0}^p (-1)^k \binom{p}{k} \mu_{n+1+k-p}}{\sum_{k=0}^p (-1)^k \binom{p}{k} \mu_{n+k-p}}, \end{aligned} \tag{36}$$

where μ_j denotes the jth moment of γ, as given in (27). Now let Δ_p denote the (forward) difference operator of order p, given by $\Delta_0 f(n) = f(n)$, $\Delta_1 f(n) = f(n+1) - f(n)$ and

$$\Delta_p = \underbrace{\Delta_1 \circ \cdots \circ \Delta_1}_{p \text{ times}}.$$

By a simple recursion on p one sees immediately that the quantity in (36) equals indeed $\frac{\Delta_p \mu_{n+1-p}}{\Delta_p \mu_{n-p}}$. Since (35) must hold for $p = 0$, we deduce that

$$\frac{\Delta_2 \mu_{n-1}}{\Delta_2 \mu_{n-2}} - 2\frac{\Delta_1 \mu_n}{\Delta_1 \mu_{n-1}} + \frac{\mu_{n+1}}{\mu_n} = 0$$

and straightforward calculations yield relation (28).

REMARK. Suppose that $\mathbf{X}_{[1,\infty)}$ is exchangeable and nondeterministic, and define μ_n, $n \ge 0$, via (27). Then, we have that $\mu_{n+1} \in (0, 1)$ for every $n \ge 0$, and that, for every $n \ge 2$, $(\mu_n, \mu_{n-1}, \mu_{n-2}) \in S$, where S is defined as in (25). As a consequence, the conclusions of Step 1 and (28) imply that, if $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable, then $f(\mu_n, \mu_{n-1}, \mu_{n-2}) \neq 0$ and $g(\mu_n, \mu_{n-1}, \mu_{n-2}) \neq 0$ for every $n \ge 2$. Therefore,

$$\mu_{n+1} = \frac{f(\mu_n, \mu_{n-1}, \mu_{n-2})}{g(\mu_n, \mu_{n-1}, \mu_{n-2})}. \tag{37}$$

STEP 3. We claim that, for any $(c_1, c_2) \in (0, 1)^2$ such that $c_1^2 < c_2 < c_1$, there exists a unique pair $(\alpha^*, \beta^*) \in (0, +\infty) \times (0, +\infty)$ such that

$$\mathbb{E}[\xi] = c_1 \quad \text{and} \quad \mathbb{E}[\xi^2] = c_2,$$

where ξ is a Beta random variable of parameters α^* and β^*. To check this, just observe that, if ξ is Beta of parameters α and β, then

$$\mathbb{E}(\xi) = \frac{\alpha}{\alpha + \beta} \quad \text{and} \quad \mathbb{E}(\xi^2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)},$$

and that, for every fixed $(c_1, c_2) \in (0,1)^2$ such that $c_1^2 < c_2 < c_1$, the system

$$\begin{cases} \dfrac{\alpha}{\alpha+\beta} = c_1, \\ \dfrac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} = c_2, \end{cases} \tag{38}$$

admits a unique solution $(\alpha^*, \beta^*) \in (0, +\infty) \times (0, +\infty)$, namely,

$$\begin{cases} \alpha^* = \dfrac{c_1(c_1 - c_2)}{c_2 - c_1^2}, \\ \beta^* = \dfrac{(1-c_1)(c_1-c_2)}{c_2 - c_1^2}. \end{cases} \tag{39}$$

We are now in a position to conclude the proof of the implication $1 \Rightarrow 2$ in the statement of Theorem 1.

STEP 4. Let $\mathbf{X}_{[1,\infty)}$ be a nondeterministic exchangeable sequence, denote by γ its de Finetti measure and by $\{\mu_n(\gamma): n \geq 0\}$ the sequence of moments appearing in (27). We suppose that $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable. There are only two possible cases: either $\mu_1(\gamma)^2 = \mu_2(\gamma)$, or $\mu_1(\gamma)^2 < \mu_2(\gamma)$. If $\mu_1(\gamma)^2 = \mu_2(\gamma)$, then necessarily $\gamma = \delta_x$ for some $x \in (0,1)$, and therefore $\mathbf{X}_{[1,\infty)}$ is a sequence of i.i.d. Bernoulli trials with common parameter equal to x. If $\mu_1(\gamma)^2 < \mu_2(\gamma)$, then, thanks to the results contained in Step 3 (note that $\mu_2(\gamma) < \mu_1(\gamma)$, since $\mathbf{X}_{[1,\infty)}$ is nondeterministic), there exists a unique pair $(\alpha^*, \beta^*) \in (0, +\infty) \times (0, +\infty)$ such that

$$\mu_1(\gamma) = \mathbb{E}(\xi) = \frac{1}{B(\alpha^*, \beta^*)} \int_0^1 \theta \theta^{\alpha^*-1}(1-\theta)^{\beta^*-1}\, d\theta, \tag{40}$$

$$\mu_2(\gamma) = \mathbb{E}(\xi^2) = \frac{1}{B(\alpha^*, \beta^*)} \int_0^1 \theta^2 \theta^{\alpha^*-1}(1-\theta)^{\beta^*-1}\, d\theta, \tag{41}$$

where ξ stands for a Beta random variable of parameters α^* and β^*. Moreover, (37) and the fact that Pólya sequences are Hoeffding decomposable imply that, for any $n \geq 2$,

$$\mu_{n+1}(\gamma) = \frac{f(\mu_n(\gamma), \mu_{n-1}(\gamma), \mu_{n-2}(\gamma))}{g(\mu_n(\gamma), \mu_{n-1}(\gamma), \mu_{n-2}(\gamma))},$$

$$\mathbb{E}(\xi^{n+1}) = \frac{f(\mathbb{E}(\xi^n), \mathbb{E}(\xi^{n-1}), \mathbb{E}(\xi^{n-2}))}{g(\mathbb{E}(\xi^n), \mathbb{E}(\xi^{n-1}), \mathbb{E}(\xi^{n-2}))},$$

where f and g are given by (23) and (24). As (40) and (41) are in order, we deduce that, for every $n \geq 1$,

$$\mu_n(\gamma) = \mathbb{E}(\xi^n) = \frac{1}{B(\alpha^*, \beta^*)} \int_0^1 \theta^n \theta^{\alpha^*-1}(1-\theta)^{\beta^*-1}\, d\theta. \tag{42}$$

Since probability measures on [0, 1] are determined by their moments, the combination of (40), (41) and (42) gives

$$\gamma(d\theta) = \frac{1}{B(\alpha^*, \beta^*)} \theta^{\alpha^*-1}(1-\theta)^{\beta^*-1}\,d\theta,$$

implying that $\mathbf{X}_{[1,\infty)}$ is a two-color Pólya sequence of parameters α^* and β^*. This concludes the proof of Theorem 1.

6. Further remarks on the two-color case. (I) With the terminology of Hill, Lane and Sudderth (1987), a random sequence $\mathbf{X}_{[1,\infty)} = \{X_n : n \geq 1\}$, with values in $\{0, 1\}$, is called an *urn process* if there exist a measurable function $f : [0, 1] \mapsto [0, 1]$ and positive natural numbers $r, b > 0$, such that, for every $n \geq 1$,

$$\mathbb{P}(X_{n+1} = 1 \mid X_1, \dots, X_n) = f\left(\frac{r + X_1 + \cdots + X_n}{r + b + n}\right). \tag{43}$$

According to Theorem 1 in Hill, Lane and Sudderth (1987), the only exchangeable and nondeterministic urn processes are i.i.d. and Pólya sequences with integer parameters (for which f is, resp., constant and equal to the identity map). This yields immediately the following consequence of Theorem 1, showing that the two (seemingly unrelated) notions of urn process and Hoeffding decomposable sequence are in many cases equivalent. The proof can be achieved by using the calculations performed in Step 4.

COROLLARY 5. *Let $\mathbf{X}_{[1,\infty)} = \{X_n : n \geq 1\}$ be a $\{0, 1\}$-valued infinite exchangeable nondeterministic sequence such that*

$$\mathbb{P}(X_1 = 1) = c_1 \quad \textit{and} \quad \mathbb{P}(X_1 = X_2 = 1) = c_2, \tag{44}$$

for some constants c_1 and c_2 such that $0 < c_1^2 < c_2 < c_1 < 1$. If the system (38) admits integer solutions, then $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable if, and only if, it is an urn process.

In general, a sequence $\mathbf{X}_{[1,\infty)}$ verifying (44) is Hoeffding decomposable if, and only if, it is a Pólya sequence with parameters α^* and β^* given by (39).

(II) The arguments rehearsed in the proof of Theorem 1 provide an alternative proof of Therorem 5 in Diaconis and Ylvisaker (1979). Indeed, Theorem 5 in that reference can be translated in the language of the present paper as follows. *Suppose that $\mathbf{X}_{[1,\infty)}$ is a nondeterministic exchangeable sequence associated with a de Finetti measure γ whose predictive probabilities are such that there exist numbers a_n and b_n satisfying*

$$\mathbb{P}^n_{n+1}(0^{(n-p)} \mid 0^{(n-p)}) = a_n p + b_n, \qquad p = 0, \dots, n. \tag{45}$$

Then, either γ is a Dirac mass concentrated at some $x \in (0,1)$, or γ is a Beta distribution. If $\gamma = \delta_x$, then $a_n = 0$ and $b_n = x$; if γ is Beta, then there exist $a > 0, b > 0$ with $a + b < 1$, such that

$$a_n = \frac{a}{1 + a(n-1)}, \qquad b_n = \frac{b}{1 + a(n-1)}. \tag{46}$$

To see how Diaconis and Ylvisaker's result can be recovered by using our techniques, observe that if (45) holds, then, by setting $A_n = -a_n$ and $B_n = na_n + b_n$, one has that

$$\mathbb{P}_{n+1}^{n}(0^{(p)} \mid 0^{(p)}) = A_n p + B_n, \qquad p = 0, \ldots, n. \tag{47}$$

Now, if (47) is true, it is immediately seen that (35) must also hold, and one deduces from the discussion contained in the previous section that γ must be Beta or Dirac. Conversely, if γ is either Dirac at some $x \in (0,1)$ or Beta, then the difference equation (35) holds, and one must conclude that there exist numbers A_n and B_n such that

$$\mathbb{P}_{n+1}^{n}(0^{(p)} \mid 0^{(p)}) = A_n p + B_n, \qquad p = 0, \ldots, n. \tag{48}$$

Specializing (48) to $p = 0$ and $p = 1$ one gets, respectively,

$$B_n = \mathbb{P}_{n+1}^{n}(0^{(0)} \mid 0^{(0)}) \quad \text{and} \quad A_n = \mathbb{P}_{n+1}^{n}(0^{(1)} \mid 0^{(1)}) - \mathbb{P}_{n+1}^{n}(0^{(0)} \mid 0^{(0)}).$$

If $\gamma = \delta_x$, then $B_n = x$ and $A_n = 0$. If γ is Beta, then necessarily $A_n < 0$. Indeed, the function

$$f(p) = \mathbb{P}_{n+1}^{n}(0^{(p)} \mid 0^{(p)}) = \mathbb{P}(X_{n+1} = 1 | \mathbf{X}_{[n]} \text{ contains exactly } p \text{ zeros})$$

is strictly decreasing [cf. Hill, Lane and Sudderth (1987), Lemma 1]. Now, if one sets $b_n = nA_n + B_n$ and $a_n = -A_n$, then one obtains exactly relation (45) with nonnegative a_n and b_n, verifying relation (46) in the case where γ is Beta.

7. Hoeffding decompositions and m-color sequences. In this section we deal with Hoeffding decomposable sequences with values in a set with more than two elements. One of our main findings (see Theorem 10 of Section 7.1) is that, under some additional conditions, there exists a universal recurrence relation analogous to formula (37), implying that the law of an exchangeable and Hoeffding decomposable sequence is completely determined by the mean vector and the covariance matrix of its de Finetti measure. The results of Section 7.1 are used in Section 7.2 to prove that Pólya urns are the only Hoeffding decomposable sequences whose de Finetti measure can be obtained as the law of a normalized vector of independent and infinitely divisible random variables. This provides a partial generalization of Theorem 1.

7.1. *General recurrence relations.* Fix an integer $m \geq 2$, and let $D = \{d_1, \ldots, d_m\}$ be a finite set of m elements. In what follows, we will denote by

$\mathbf{X}_{[1,\infty)} = (X_1, \dots, X_n, \dots)$ an infinite exchangeable sequence of random variables with values in D. We also denote by $\prod_D$ the class of all probability measures on D; the elements of $\prod_D$ are written $p = \{p\{d_i\} : i = 1, \dots, m\}$, where $p\{d_i\}$ indicates the p-probability of d_i. According to the de Finetti theorem, the assumption of exchangeability yields the existence of a unique probability measure γ on $\prod_D$ (called, as before, the *de Finetti measure* associated with $\mathbf{X}_{[1,\infty)}$) such that, for every $(x_1, \dots, x_n) \in D^n$,

$$\mathbb{P}\big[\mathbf{X}_{[n]} = (x_1, \dots, x_n)\big] = \int_{\prod_D} \prod_{j=1}^{n} p\{x_j\}\gamma(dp) \tag{49}$$

[recall the notation $\mathbf{X}_{[n]} = \mathbf{X}_{[1,n]} = (X_1, \dots, X_n)$]. We will also assume that the following "nondegeneracy" condition is satisfied: for every $n \geq 1$ and every vector $\mathbf{x}_n = (x_1, \dots, x_n) \in D^n$, one has that

$$\mathbb{P}\big[\mathbf{X}_{[n]} = (x_1, \dots, x_n)\big] > 0. \tag{50}$$

When $m = 2$ condition (50) is verified if and only if $\mathbf{X}_{[1,\infty)}$ is nondeterministic in the sense of Section 3. Note, however, that, in the case $m > 2$, condition (50) rules out exchangeable sequences that are highly nontrivial. As an example, set $D = \{0, 1, 2\}$ and consider a sequence $\mathbf{X}^*_{[1,\infty)}$ such that its de Finetti measure γ^* verifies $\gamma^*(p\{0\} = 1/2) = 1$ and $\gamma^*(p\{1\} = 1/2) = 1/2 = \gamma^*(p\{1\} = 0)$. Then, $\mathbf{X}^*_{[1,\infty)}$ does not verify (50), since any realization of $\mathbf{X}^*_{[1,\infty)}$ a.s. contains either zeros and ones (with no twos), or zeros and twos (with no ones).

The collection of vector spaces $\{SU_k(\mathbf{X}_{[n]}) : n \geq 2, k = 0, \dots, n\}$, associated with $\mathbf{X}_{[1,\infty)}$, is defined as in Section 2. In particular, for $1 \leq k \leq n$, $SU_k(\mathbf{X}_{[n]})$ is generated by random variables of the type (1). The family of Hoeffding spaces $\{SH_k(\mathbf{X}_{[n,\infty]}) : n \geq 2, k = 0, \dots, n\}$ is defined by formula (2). For $k \geq 1$, the symbol $\Xi_k(\mathbf{X}_{[1,\infty)})$ indicates the linear space of all symmetric kernels φ on D^k such that the degeneracy condition (4) is verified.

In what follows, for $k \geq 1$ and for a real-valued symmetric function $\varphi \in \mathcal{S}(D^k)$ [$\mathcal{S}(D^1)$ is just the class of real-valued functions on D], we shall use the following shorthand notation: for every $n > k$,

$$\sigma^n(\varphi)\big(\mathbf{X}_{[n]}\big) = \sum_{1 \leq j_1 < \cdots < j_k \leq n} \varphi(X_{j_1}, \dots, X_{j_k}), \tag{51}$$

so that, for example, the random variable $F(\mathbf{X}_{[n]})$ appearing in (1) can be rewritten as $F(\mathbf{X}_{[n]}) = \sigma^n(\varphi)(\mathbf{X}_{[n]})$. Observe that the application $\varphi \mapsto \sigma^n(\varphi)$ yields a one-to-one linear mapping from $\mathcal{S}(D^k)$ onto $SU_k(\mathbf{X}_{[n]})$. We will also need the following "composition rule": for every $k \geq 2$, every $n > k$ and every $\varphi \in \mathcal{S}(D^{k-1})$,

$$\begin{aligned}\sigma^n(\sigma^k(\varphi))\big(\mathbf{X}_{[n]}\big) &= \sum_{1 \leq j_1 < \cdots < j_k \leq n} \sum_{\substack{\{i_1, \dots, i_{k-1}\} \subset \\ \{j_1, \dots, j_k\}}} \varphi(X_{i_1}, \dots, X_{i_{k-1}}) \\ &= (n - k + 1)\sigma^n(\varphi)\big(\mathbf{X}_{[n]}\big).\end{aligned} \tag{52}$$

For every $k \geq 1$, we denote by $\mathcal{N}(k,m)$ the class of *weak m-compositions* of k. This means that $\mathcal{N}(k,m)$ is the set of all vectors of the type $\mathbf{n}_m = (n_1, \dots, n_m)$, where the numbers n_i, $i = 1, \dots, m$, are *nonnegative* integers such that $n_1 + \cdots + n_m = k$. For instance, the vectors $(1,0,5)$ and $(2,2,2)$ are two elements of $\mathcal{N}(6,3)$. It is well known [see, e.g., Stanley (1997), page 15] that $\mathcal{N}(k,m)$ contains exactly $\binom{k+m-1}{m-1}$ elements. For every $k \geq 1$ and every $\mathbf{n}_m = (n_1, \dots, n_m) \in \mathcal{N}(k,m)$, we use the notation

$$\binom{k}{\mathbf{n}_m} = \binom{k}{n_1 n_2, \dots, n_m} = \frac{k!}{n_1! \cdots n_m!},$$

where $\binom{k}{n_1 n_2, \dots, n_m}$ is the usual multinomial symbol. To every $k \geq 1$ and every $\mathbf{n}_m = (n_1, \dots, n_m) \in \mathcal{N}(k,m)$ we associate the set

$$C(k, \mathbf{n}_m) = \{(x_1, \dots, x_k) \in D^k : \text{exactly } n_i \text{ of the } x_j\text{'s equal } d_i,\ i = 1, \dots, m\}. \tag{53}$$

In other words, a vector $\mathbf{x}_k \in D^k$ is an element of $C(k, \mathbf{n}_m)$ if, and only if, exactly n_i of the components of $\mathbf{x}_k$ are equal to d_i, for every $i = 1, \dots, m$. For instance, if $m = k = 3$, $\mathbf{n}_3' = (2,0,1) \in \mathcal{N}(3,3)$ and $\mathbf{n}_3'' = (1,1,1) \in \mathcal{N}(3,3)$, then

$$\begin{aligned} C(3, \mathbf{n}_3') &= \{(d_1, d_1, d_3), (d_1, d_3, d_1), (d_3, d_1, d_1)\}, \\ C(3, \mathbf{n}_3'') &= \{(d_{\pi(1)}, d_{\pi(2)}, d_{\pi(3)}) : \pi \in \mathfrak{S}_3\}. \end{aligned}$$

The following facts (A)–(D) can be immediately checked: (A) for every $k \geq 1$, the collection $\{C(k, \mathbf{n}_m) : \mathbf{n}_m \in \mathcal{N}(k,m)\}$ is a partition of D^k; (B) for every $k \geq 1$ and every $\mathbf{n}_m \in \mathcal{N}(k,m)$, the indicator function $\mathbf{x}_k \mapsto \mathbf{1}_{C(k,\mathbf{n}_m)}(\mathbf{x}_k)$ is an element of $\mathcal{S}(D^k)$; (C) for every $k \geq 1$, the collection $\{\mathbf{1}_{C(k,\mathbf{n}_m)} : \mathbf{n}_m \in \mathcal{N}(k,m)\}$ is a basis of the vector space $\mathcal{S}(D^k)$; (D) since (50) is in order, for every $k \geq 1$ and every $\mathbf{n}_m = (n_1, \dots, n_m) \in \mathcal{N}(k,m)$,

$$\begin{aligned} &\mathbb{P}[\mathbf{X}_{[k]} \in C(k, \mathbf{n}_m)] \\ &\quad = \binom{k}{\mathbf{n}_m} \mathbb{P}[\mathbf{X}_{[k]} = (\underbrace{d_1, \dots, d_1}_{n_1 \text{ times}}, \dots, \underbrace{d_m, \dots, d_m}_{n_m \text{ times}})] \in (0,1). \end{aligned} \tag{54}$$

The equality in (54) is just a consequence of exchangeability. Note also that the probability appearing on the right-hand side of such an equality could be rewritten (with a notation analogous to the one adopted in Section 4) as $\mathbb{P}_k[d_1^{(n_1)}, \dots, d_m^{(n_m)}]$, where $\mathbb{P}_k[d_1^{(n_1)}, \dots, d_m^{(n_m)}]$ stands for the *constant value* taken by the application $\mathbf{x}_k \mapsto \mathbb{P}[\mathbf{X}_{[k]} = \mathbf{x}_k]$ on the set $C(k, \mathbf{n}_m)$. Point (B) above implies that the space $\mathcal{S}(D^k)$ has exactly dimension $\binom{k+m-1}{m-1}$. By combining the above described facts (A)–(D), one immediately deduces the following result.

PROPOSITION 6. *Let the above assumptions and notations prevail* [*in particular* (50), *and therefore* (54), *hold*]. *Then, for every* $k \geq 1$ *and every* $n \geq k$:

1. *The set* $\{\sigma^n(\mathbf{1}_{C(k,\mathbf{n}_m)})(\mathbf{X}_{[n]}) : \mathbf{n}_m \in \mathcal{N}(k,m)\}$ *is a basis of the vector space* $SU_k(\mathbf{X}_{[n]})$, *which has therefore dimension* $\binom{k+m-1}{m-1}$.
2. *The vector spaces* $SH_k(\mathbf{X}_{[n]})$ *and* $\Xi_k(\mathbf{X}_{[1,\infty)})$ *have dimension* $\binom{k+m-1}{m-1} - \binom{k+m-2}{m-1}$.

Note that point 1 in the statement of Proposition 6 is a consequence of (54), of the relation $\dim \mathcal{S}(D^k) = \binom{k+m-1}{m-1}$, and of the fact that $SU_k(\mathbf{X}_{[n]})$ is the collection of all U-statistics of the form $\sigma^n(\varphi)(\mathbf{X}_{[n]})$, where $\varphi \in \mathcal{S}(D^k)$. Since $\binom{k+m-1}{m-1} - \binom{k+m-2}{m-1} > 0$, point 2 ensures that condition (9) is satisfied, so that Theorem 0 can be applied in our framework. Observe also that Proposition 6 is consistent with the discussion contained in Section 2. In particular, if $m = 2$, then $\dim SU_k(\mathbf{X}_{[n]}) = k+1$, and $\dim SH_k(\mathbf{X}_{[1,\infty)}) = \dim \Xi_k(\mathbf{X}_{[n]}) = 1$, $k = 1, \dots, n$. As anticipated, we shall now obtain a class of *necessary conditions* for the sequence $\mathbf{X}_{[1,\infty)}$ to be Hoeffding decomposable. To this end, suppose that $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable and recall that, by Theorem 0, $\mathbf{X}_{[1,\infty)}$ is also weakly independent in the sense of Definition B. Now, for $k \geq 2$ consider a function $\eta \in \mathcal{S}(D^k)$, and observe that point 1 in Proposition 6 implies that the projection of the symmetric statistic $\eta(\mathbf{X}_{[k]})$ on the space $SU_{k-1}(\mathbf{X}_{[k]})$ has necessarily the form of a linear combination of the elements of the basis $\{\sigma^k(\mathbf{1}_{C(k-1,\mathbf{n}_m)})(\mathbf{X}_{[k]}) : \mathbf{n}_m \in \mathcal{N}(k-1,m)\}$, namely

$$\begin{aligned}\pi[\eta, SU_{k-1}](\mathbf{X}_{[k]}) &= \pi^k_{k-1}[\eta](\mathbf{X}_{[k]}) \\ &\triangleq \sum_{\mathbf{n}_m \in \mathcal{N}(k-1,m)} z_\gamma(\eta, \mathbf{n}_m)\sigma^k(\mathbf{1}_{C(k-1,\mathbf{n}_m)})(\mathbf{X}_{[k]}),\end{aligned} \tag{55}$$

where we used the notation introduced in (53), and where $\{z_\gamma(\eta, \mathbf{n}_m)\}$ is a (uniquely defined) collection of real coefficients which of course depend on γ (or, equivalently, on the law of $\mathbf{X}_{[1,\infty)}$). Since $\eta \in \mathcal{S}(D^k)$, one has that $\sigma^n(\eta)(\mathbf{X}_{[n]}) \in SU_k(\mathbf{X}_{[n]})$ for every $n \geq k$. Moreover, since the function $\mathbf{x}_k \mapsto \pi^k_{k-1}[\eta](\mathbf{x}_k)$ is symmetric on D^k, and since (by construction)

$$\mathbb{E}[\eta(\mathbf{X}_{[k]}) - \pi^k_{k-1}[\eta](\mathbf{X}_{[k]}) \mid \mathbf{X}_{[2,k]}] = 0, \tag{56}$$

the Hoeffding decomposability of $\mathbf{X}_{[1,\infty)}$ yields that, for every $n \geq k$, the random variable

$$\begin{aligned}&\sum_{1 \leq j_1 < \cdots < j_k \leq n} \{\eta(X_{j_1}, \dots, X_{j_k}) - \pi^k_{k-1}[\eta](X_{j_1}, \dots, X_{j_k})\} \\ &\quad = \sigma^n(\eta)(\mathbf{X}_{[n]}) - \sigma^n(\pi^k_{k-1}[\eta])(\mathbf{X}_{[n]})\end{aligned}$$

is an element of $SH_k(\mathbf{X}_{[n]})$, implying that

$$\sigma^n(\pi_{k-1}^k[\eta])(\mathbf{X}_{[n]}) = \sum_{\mathbf{n}_m \in \mathcal{N}(k-1,m)} z_\gamma(\eta, \mathbf{n}_m)(n-k+1)\sigma^n(\mathbf{1}_{C(k-1,\mathbf{n}_m)})(\mathbf{X}_{[n]})$$

[where we have used (52)] is the projection of $\sigma^n(\eta)(\mathbf{X}_{[n]})$ on $SU_{k-1}(\mathbf{X}_{[n]})$.

For integers $a, b \geq 1$ and kernels $\xi \in \mathcal{S}(D^a)$, $\psi \in \mathcal{S}(D^b)$, we shall write:

$$\begin{aligned} A_\gamma(\xi, \psi) &\triangleq \mathbb{E}[\xi(X_1, \dots, X_a)\psi(X_{a+1}, \dots, X_{a+b})] \\ &= \int_{\prod_D} \left(\int_{D^a} \xi \, dp^{\otimes a}\right)\left(\int_{D^b} \psi \, dp^{\otimes b}\right)\gamma(dp), \end{aligned}$$

where the de Finetti measure γ is defined in (49), and $p^{\otimes l}$, $l \geq 2$, indicates the lth product measure induced by p on D^l (with $p^{\otimes 1} = p$). We claim that the collection

$$(57) \quad \mathfrak{A}_\gamma \triangleq \{A_\gamma(\mathbf{1}_{C(k,\mathbf{n}'_m)}, \mathbf{1}_{C(j,\mathbf{n}''_m)}) : k, j \geq 1, \mathbf{n}'_m \in \mathcal{N}(k,m), \mathbf{n}''_m \in \mathcal{N}(j,m)\}$$

completely characterizes γ. The basic idea to prove this last statement is that one can represent all probabilities $\mathbb{P}[\mathbf{X}_{[n]} = (x_1, \dots, x_n)]$ as linear combinations of elements of $\mathfrak{A}_\gamma$, through an appropriate use of formula (49). In the following lemma we establish a more precise result. To this end, define, for every fixed $j, k \geq 1$,

$$(58) \quad \mathfrak{A}_\gamma(k, j) \triangleq \{A_\gamma(\mathbf{1}_{C(k,\mathbf{n}'_m)}, \mathbf{1}_{C(j,\mathbf{n}''_m)}) : \mathbf{n}'_m \in \mathcal{N}(k,m), \mathbf{n}''_m \in \mathcal{N}(j,m)\}.$$

LEMMA 7. *For every fixed $j, k \geq 1$, the class $\mathfrak{A}_\gamma(k, j)$ completely determines the family of probabilities*

$$(59) \qquad \mathbb{P}[\mathbf{X}_{[k+j]} = (x_1, \dots, x_{k+j})], \qquad (x_1, \dots, x_{k+j}) \in D^{k+j},$$

via the relation

$$(60) \quad \mathbb{P}[\mathbf{X}_{[k+j]} = (x_1, \dots, x_{k+j})] = \binom{k}{\mathbf{n}'_m}^{-1}\binom{j}{\mathbf{n}''_m}^{-1} A_\gamma(\mathbf{1}_{C(k,\mathbf{n}'_m)}, \mathbf{1}_{C(j,\mathbf{n}''_m)}),$$

where $\mathbf{n}'_m \in \mathcal{N}(k,m)$ and $\mathbf{n}''_m \in \mathcal{N}(j,m)$ are such that $(x_1, \dots, x_k) \in C(k, \mathbf{n}'_m)$ and $(x_{k+1}, \dots, x_{k+j}) \in C(j, \mathbf{n}''_m)$.

PROOF. Define the sets $C(k, \mathbf{n}'_m)$ and $C(j, \mathbf{n}''_m)$ as in the statement. By exchangeability,

$$\begin{aligned} &\mathbb{P}[\mathbf{X}_{[k]} \in C(k, \mathbf{n}'_m), \mathbf{X}_{[k+1,k+j]} \in C(j, \mathbf{n}''_m)] \\ &\quad = \binom{k}{\mathbf{n}'_m}\binom{j}{\mathbf{n}''_m}\mathbb{P}[\mathbf{X}_{[k]} = (x_1, \dots, x_k), \mathbf{X}_{[k+1,k+j]} = (x_{k+1}, \dots, x_{k+j})], \end{aligned}$$

and the conclusion is obtained from the relation

$$\mathbb{P}[\mathbf{X}_{[k]} \in C(k, \mathbf{n}'_m), \mathbf{X}_{[k+1,k+j]} \in C(j, \mathbf{n}''_m)] = A_\gamma(\mathbf{1}_{C(k,\mathbf{n}'_m)}, \mathbf{1}_{C(j,\mathbf{n}''_m)}). \qquad \square$$

In what follows, we shall prove that, whenever $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable, the class $\mathfrak{A}_\gamma$ defined in (57) is completely determined by the family $\mathfrak{A}_\gamma(1,1)$ defined in (58). This will be done by establishing a recursive relation on the classes $\mathfrak{A}_\gamma(k,j)$. This relation plays a role which is analogous to the recursive formula (37), that we extensively used in Section 5. To do this, observe that, for every $k \geq 2$ and every $\eta \in \mathcal{S}(D^k)$, the symmetric function $\mathbf{x}_k \mapsto \eta(\mathbf{x}_k) - \pi^k_{k-1}[\eta](\mathbf{x}_k) \in \mathcal{S}(D^k)$, defined via (55), verifies the degeneracy condition (56). Since $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable, and therefore weakly independent, we deduce that, a.s.-$\mathbb{P}$,

$$\mathbb{E}[\eta(\mathbf{X}_{[k]}) - \pi^k_{k-1}[\eta](\mathbf{X}_{[k]}) \mid X_{k+1}, \ldots, X_{2k-1}] = 0.$$

Hence, for every $\mathbf{n}'_m \in \mathcal{N}(k-1,m)$,

$$\begin{aligned} & A_\gamma(\eta, \mathbf{1}_{C(k-1,\mathbf{n}'_m)}) - A_\gamma(\pi^k_{k-1}[\eta], \mathbf{1}_{C(k-1,\mathbf{n}'_m)}) \\ &\quad = A_\gamma(\{\eta - \pi^k_{k-1}[\eta]\}, \mathbf{1}_{C(k-1,\mathbf{n}'_m)}) \\ &\quad = \mathbb{E}[\{\eta(\mathbf{X}_{[k]}) - \pi^k_{k-1}[\eta](\mathbf{X}_{[k]})\}\mathbf{1}_{C(k-1,\mathbf{n}'_m)}(X_{k+1}, \ldots, X_{2k-1})] = 0. \end{aligned} \tag{61}$$

By specializing (61) to the case $\eta = \mathbf{1}_{C(k,\mathbf{n}^*_m)}$ for some fixed $\mathbf{n}^*_m \in \mathcal{N}(k,m)$, and by using the explicit form of $\pi^k_{k-1}[\eta]$ given in (55), we obtain the first part of the following result. The second part is obtained in an analogous way, by first writing the projection $\pi[\mathbf{1}_{C(2k-1,\mathbf{n}^{**})}, SU_{2k-2}](\mathbf{X}_{[2k-1]})$, of $\mathbf{1}_{C(2k-1,\mathbf{n}^{**})}(\mathbf{X}_{[2k-1]})$ on $SU_{2k-2}(\mathbf{X}_{[2k-1]})$ [by means of an appropriate modification of formula (55)], and then by using the fact that, by weak independence,

$$\mathbb{E}[\mathbf{1}_{C(2k-1,\mathbf{n}^{**})}(\mathbf{X}_{[2k-1]}) - \pi[\mathbf{1}_{C(2k-1,\mathbf{n}^{**})}, SU_{2k-2}](\mathbf{X}_{[2k-1]}) \mid X_{2k}] = 0.$$

PROPOSITION 8. *For every $k \geq 2$, every $\mathbf{n}'_m \in \mathcal{N}(k-1,m)$ and every $\mathbf{n}^*_m \in \mathcal{N}(k,m)$,*

$$\begin{aligned} & A_\gamma(\mathbf{1}_{C(k,\mathbf{n}^*_m)}, \mathbf{1}_{C(k-1,\mathbf{n}'_m)}) \\ &\quad = \sum_{\mathbf{n}_m \in \mathcal{N}(k-1,m)} z_\gamma(\mathbf{1}_{C(k,\mathbf{n}^*_m)}, \mathbf{n}_m) k A_\gamma(\mathbf{1}_{C(k-1,\mathbf{n}_m)}, \mathbf{1}_{C(k-1,\mathbf{n}'_m)}). \end{aligned} \tag{62}$$

For every $k \geq 2$, every $\mathbf{n}''_m \in \mathcal{N}(1,m)$ and every $\mathbf{n}^{**}_m \in \mathcal{N}(2k-1,m)$,

$$\begin{aligned} & A_\gamma(\mathbf{1}_{C(2k-1,\mathbf{n}^{**}_m)}, \mathbf{1}_{C(1,\mathbf{n}''_m)}) \\ &\quad = \sum_{\mathbf{n}_m \in \mathcal{N}(2k-2,m)} z_\gamma(\mathbf{1}_{C(2k-1,\mathbf{n}^{**}_m)}, \mathbf{n}_m)(2k-1) \\ &\qquad \times A_\gamma(\mathbf{1}_{C(2k-2,\mathbf{n}_m)}, \mathbf{1}_{C(1,\mathbf{n}''_m)}). \end{aligned} \tag{63}$$

Observe that a consequence of the first part of Proposition 8 is that, for every fixed $\mathbf{n}^*_m$, the matrix

$$\{A_\gamma(\mathbf{1}_{C(k-1,\mathbf{n}_m)}, \mathbf{1}_{C(k-1,\mathbf{n}'_m)}), A_\gamma(\mathbf{1}_{C(k,\mathbf{n}^*_m)}, \mathbf{1}_{C(k-1,\mathbf{n}'_m)})\},$$

with the columns indexed by $\mathbf{n}'_m \in \mathcal{N}(k-1,m)$ and $\mathbf{n}^*_m$ and the rows indexed by $\mathbf{n}'_m \in \mathcal{N}(k-1,m)$, has rank at most equal to $\binom{m+k-2}{m-1}$. The following result is one of the keys of this section.

PROPOSITION 9. *There exists a universal class of deterministic functions*

$$\{F^{(k)}_{\mathbf{n}^*_m,\mathbf{n}_m} : k \geq 2, \mathbf{n}^*_m \in \mathcal{N}(k,m), \mathbf{n}_m \in \mathcal{N}(k-1,m)\},$$

such that, for every Hoeffding decomposable exchangeable sequence $\mathbf{X}_{[1,\infty)}$ *verifying* (50) (*with de Finetti measure* γ) *and for every* $k \geq 2$, *the following two properties hold*: (I) *the coefficients* $\{z_\gamma(\mathbf{1}_{C(k,\mathbf{n}^*_m)}, \mathbf{n}_m) : \mathbf{n}_m \in \mathcal{N}(k-1,m)\}$ *appearing in* (62) *admit the representation*

$$z_\gamma(\mathbf{1}_{C(k,\mathbf{n}^*_m)}, \mathbf{n}_m) = F^{(k)}_{\mathbf{n}^*_m,\mathbf{n}_m}(\mathfrak{A}_\gamma(a,b) : a, b \geq 1, a+b \leq k),$$

and (II) *the coefficients* $\{z_\gamma(\mathbf{1}_{C(2k-1,\mathbf{n}^{**}_m)}, \mathbf{n}_m) : \mathbf{n}_m \in \mathcal{N}(2k-2,m)\}$ *appearing in* (63) *admit the representation*

$$z_\gamma(\mathbf{1}_{C(2k-1,\mathbf{n}^{**}_m)}, \mathbf{n}_m) = F^{(2k-1)}_{\mathbf{n}^{**}_m,\mathbf{n}_m}(\mathfrak{A}_\gamma(a,b) : a, b \geq 1, a+b \leq 2k-1).$$

PROOF. Fix a sequence $\mathbf{X}_{[1,\infty)}$ verifying (50) and with de Finetti measure γ. We will only prove point (I) in the statement [the proof of point (II) is analogous]. Recall that the real-valued coefficients $\{z_\gamma(\mathbf{1}_{C(k,\mathbf{n}^*_m)}, \mathbf{n}_m) : \mathbf{n}_m \in \mathcal{N}(k-1,m)\}$ are those determining the projection

$$\pi^k_{k-1}[\mathbf{1}_{C(k,\mathbf{n}^*_m)}](\mathbf{X}_{[k]}) = \sum_{\mathbf{n}_m \in \mathcal{N}(k-1,m)} z_\gamma(\mathbf{1}_{C(k,\mathbf{n}^*_m)}, \mathbf{n}_m)\sigma^k(\mathbf{1}_{C(k-1,\mathbf{n}_m)})(\mathbf{X}_{[k]})$$

of the symmetric statistic $\mathbf{1}_{C(k,\mathbf{n}^*_m)}(\mathbf{X}_{[k]})$ on the space $SU_{k-1}(\mathbf{X}_{[k]})$, expressed as a linear combination of the elements of the (not necessarily orthonormal) basis $\{\sigma^k(\mathbf{1}_{C(k-1,\mathbf{n}_m)}) : \mathbf{n}_m \in \mathcal{N}(k-1,m)\}$. It follows that such coefficients can be computed by implementing the following procedure:

- Use a Gram–Schmidt procedure to obtain from $\{\sigma^k(\mathbf{1}_{C(k-1,\mathbf{n}_m)})\}$ an orthonormal basis $\{a(j) : j = 1, \ldots, \binom{m+k-2}{m-1}\}$ of $SU_{k-1}(\mathbf{X}_{[k]})$.
- Write

$$\pi^k_{k-1}[\mathbf{1}_{C(k,\mathbf{n}^*_m)}] = \sum_{j=1}^{\binom{m+k-2}{m-1}} a(j)\mathbb{E}[\mathbf{1}_{C(k,\mathbf{n}^*_m)} \times a(j)]. \tag{64}$$

- Compute the $\{z_\gamma(\mathbf{1}_{C(k,\mathbf{n}^*_m)}, \mathbf{n}_m)\}$ by plugging into (64) the expression of each $a(j)$ in terms of linear combinations of the functions $\sigma^k(\mathbf{1}_{C(k-1,\mathbf{n}_m)})$.

We therefore deduce (by exchangeability) that each $z_\gamma(\mathbf{1}_{C(k,\mathbf{n}^*_m)}, \mathbf{n}_m)$ can be expressed as a function *not depending on* γ (and therefore not depending on the law of $\mathbf{X}_{[1,\infty)}$) of expectations the type

$$\mathbb{E}\big[\mathbf{1}_{C(k-1,\mathbf{n}_m)}\big(\mathbf{X}_{[k-1]}\big)\big], \qquad \mathbb{E}\big[\mathbf{1}_{C(k-1,\mathbf{n}_m)}\big(\mathbf{X}_{[k-1]}\big)\mathbf{1}_{C(k-1,\mathbf{n}'_m)}\big(\mathbf{X}_{[2,k]}\big)\big],$$
$$\mathbb{E}\big[\mathbf{1}_{C(k,\mathbf{n}^*_m)}\big(\mathbf{X}_{[k]}\big)\mathbf{1}_{C(k-1,\mathbf{n}_m)}\big(\mathbf{X}_{[k-1]}\big)\big],$$

where $\mathbf{n}_m, \mathbf{n}'_m \in \mathcal{N}(k-1,m)$ [recall that $\mathbf{n}^*_m \in \mathcal{N}(k,m)$]. As a consequence, the result in the statement is proved, once it is shown that there exist universal functions (not depending on the law of $\mathbf{X}_{[1,\infty)}$)

$$\big\{\Phi^{(k-1)}_{\mathbf{n}_m} : k \geq 2, \mathbf{n}_m \in \mathcal{N}(k-1,m)\big\},$$
$$\big\{\Psi^{(k-1)}_{\mathbf{n}_m,\mathbf{n}'_m} : k \geq 2, \mathbf{n}_m \in \mathcal{N}(k-1,m), \mathbf{n}'_m \in \mathcal{N}(k-1,m)\big\},$$
$$\big\{\Sigma^{(k,k-1)}_{\mathbf{n}_m,\mathbf{n}^*_\mathbf{m}} : k \geq 2, \mathbf{n}_m \in \mathcal{N}(k-1,m), \mathbf{n}^*_m \in \mathcal{N}(k,m)\big\},$$

such that, for every $k \geq 2$, and every $\mathbf{n}_m, \mathbf{n}'_m \in \mathcal{N}(k-1,m)$, $\mathbf{n}^*_m \in \mathcal{N}(k,m)$

$$\begin{aligned}&\mathbb{E}\big[\mathbf{1}_{C(k-1,\mathbf{n}_m)}\big(\mathbf{X}_{[k-1]}\big)\big]\\ &\quad = \Phi^{(k-1)}_{\mathbf{n}_m}\big(\mathfrak{A}_\gamma(a,b) : a,b \geq 1, a+b \leq k\big),\end{aligned}\tag{65}$$

$$\begin{aligned}&\mathbb{E}\big[\mathbf{1}_{C(k-1,\mathbf{n}_m)}\big(\mathbf{X}_{[k-1]}\big)\mathbf{1}_{C(k-1,\mathbf{n}'_m)}\big(\mathbf{X}_{[2,k]}\big)\big]\\ &\quad = \Psi^{(k-1)}_{\mathbf{n}_m,\mathbf{n}'_m}\big(\mathfrak{A}_\gamma(a,b) : a,b \geq 1, a+b \leq k\big),\end{aligned}\tag{66}$$

$$\begin{aligned}&\mathbb{E}\big[\mathbf{1}_{C(k,\mathbf{n}^*_m)}\big(\mathbf{X}_{[k]}\big)\mathbf{1}_{C(k-1,\mathbf{n}_m)}\big(\mathbf{X}_{[k-1]}\big)\big]\\ &\quad = \Sigma^{(k,k-1)}_{\mathbf{n}_m,\mathbf{n}^*_\mathbf{m}}\big(\mathfrak{A}_\gamma(a,b) : a,b \geq 1, a+b \leq k\big).\end{aligned}\tag{67}$$

For $\mathbf{n}_m \in \mathcal{N}(k-1,m)$ one has that

$$\begin{aligned}&\mathbb{E}\big[\mathbf{1}_{C(k-1,\mathbf{n}_m)}\big(\mathbf{X}_{[k-1]}\big)\big]\\ &\quad = \sum_{\mathbf{n}'_m \in \mathcal{N}(1,m)} \mathbb{P}\big[\mathbf{X}_{[k-1]} \in C(k-1,\mathbf{n}_m), X_1 \in C(1,\mathbf{n}'_m)\big]\\ &\quad = \sum_{\mathbf{n}'_m \in \mathcal{N}(1,m)} A_\gamma\big(\mathbf{1}_{C(k-1,\mathbf{n}_m)}, \mathbf{1}_{C(1,\mathbf{n}'_m)}\big),\end{aligned}$$

so that (65) is proved. Since

$$\mathbb{E}\big[\mathbf{1}_{C(1,\mathbf{n}_m)}(X_1)\mathbf{1}_{C(1,\mathbf{n}'_m)}(X_2)\big] = A_\gamma\big(\mathbf{1}_{C(1,\mathbf{n}_m)}, \mathbf{1}_{C(1,\mathbf{n}'_m)}\big),$$

we need to prove (66) only for $k \geq 3$. For $k \geq 3$, given $\mathbf{n}_m, \mathbf{n}'_m \in \mathcal{N}(k-1,m)$, we say that $\mathbf{n}_m$ and $\mathbf{n}'_m$ are *compatible*, if there exists $\mathbf{n}^0_m = (n^0_1, \dots, n^0_m) \in \mathcal{N}(k-2,m)$, as well as $i, j \in \{1,\dots,m\}$ such that

$$\mathbf{n}_m = (n^0_1, \dots, n^0_i + 1, \dots, n^0_m), \qquad \mathbf{n}'_m = (n^0_1, \dots, n^0_j + 1, \dots, n^0_m).\tag{68}$$

If $\mathbf{n}_m$ and $\mathbf{n}'_m$ are compatible in the sense of (68), we write $\mathbf{n}_m^i$ and $\mathbf{n}_m^{i,j}$, respectively, for the vector $(n_1^i, \dots, n_m^i) \in \mathcal{N}(2, m)$ such that $n_i^i = 2$ and $n_a^i = 0$ for $a \neq i$, and for the vector $(n_1^{i,j}, \dots, n_m^{i,j}) \in \mathcal{N}(2, m)$ such that $n_i^{i,j} = n_j^{i,j} = 1$ and $n_a^{i,j} = 0$ for $a \neq i, j$. Then,

$$
\begin{aligned}
&\mathbb{E}\big[\mathbf{1}_{C(k-1,\mathbf{n}_m)}(\mathbf{X}_{[k-1]})\mathbf{1}_{C(k-1,\mathbf{n}'_m)}(\mathbf{X}_{[2,k]})\big] \\
&\quad = \mathbb{P}\big[\mathbf{X}_{[2,k]} \in C(k-1, \mathbf{n}'_m), \mathbf{X}_{[1,k-1]} \in C(k-1, \mathbf{n}_m)\big] \\
&\quad = \begin{cases} 0, & \text{if } \mathbf{n}_m, \mathbf{n}'_m \text{ are not compatible,} \\ A_\gamma\big(\mathbf{1}_{C(k-2,\mathbf{n}_m^0)}, \mathbf{1}_{C(2,\mathbf{n}_m^i)}\big), & \text{if } \mathbf{n}_m, \mathbf{n}'_m \text{ are compatible and } i = j, \\ A_\gamma\big(\mathbf{1}_{C(k-2,\mathbf{n}_m^0)}, \mathbf{1}_{C(2,\mathbf{n}_m^{i,j})}\big), & \text{if } \mathbf{n}_m, \mathbf{n}'_m \text{ are compatible and } i \neq j. \end{cases}
\end{aligned}
$$

This proves (66). To prove (67), we will use the following notation: for $\mathbf{n}_m = (n_1, \dots, n_m) \in \mathcal{N}(k-1, m)$ and $\mathbf{n}_m^* = (n_1^*, \dots, n_m^*) \in \mathcal{N}(k, m)$, we write $\mathbf{n}_m \leq \mathbf{n}_m^* \in \mathcal{N}(k, m)$ whenever $\mathbf{n}_m^*$ is obtained by adding 1 to one of the components of $\mathbf{n}_m$, that is, whenever there exists $i = 1, \dots, m$ such that $n_i = n_i^* - 1$ and $n_a = n_a^*$ for every $a \neq i$. Now write $\mathbf{n}_m^{1,i}$ to indicate the element of $\mathcal{N}(1, m)$ such that the ith component of $\mathbf{n}_m^{1,i}$ equals 1 (and all the other components are zero). Then, one proves immediately that

$$
\begin{aligned}
&\mathbb{E}\big[\mathbf{1}_{C(k,\mathbf{n}_m^*)}(\mathbf{X}_{[k]})\mathbf{1}_{C(k-1,\mathbf{n}_m)}(\mathbf{X}_{[k-1]})\big] \\
&\quad = \begin{cases} 0, & \text{if } \mathbf{n}_m \not\leq \mathbf{n}_m^*, \\ A_\gamma\big(\mathbf{1}_{C(k-1,\mathbf{n}_m)}, \mathbf{1}_{C(1,\mathbf{n}_m^{1,i})}\big), & \text{if } \mathbf{n}_m \leq \mathbf{n}_m^*. \end{cases}
\end{aligned}
$$

This concludes the proof of the proposition. $\square$

Now consider an exchangeable sequence $\mathbf{X}_{[1,\infty)}$, with de Finetti measure γ and satisfying (50), and suppose that $\mathbf{X}_{[1,\infty)}$ is Hoeffding decomposable. The combination of Proposition 8 and Proposition 9 implies that there exists a universal (i.e., not depending on γ) recursive relation, according to which the following properties hold for every $k \geq 2$: (i) the elements of the class $\{\mathfrak{A}_\gamma(i, j) : i + j \leq 2k-1\}$ can be expressed in terms of the class $\{\mathfrak{A}_\gamma(i, j) : i + j \leq 2k-2\}$, and (ii) the elements of the class $\{\mathfrak{A}_\gamma(i, j) : i + j \leq 2k\}$ can be expressed in terms of the class $\{\mathfrak{A}_\gamma(i, j) : i + j \leq 2k-1\}$. Since the set $\mathfrak{A}_\gamma$ [as defined in (57)] determines the law of $\mathbf{X}_{[1,\infty)}$ (thanks to Lemma 7), one deduces immediately the following result.

THEOREM 10. *Let $\mathbf{X}_{[1,\infty)}$ be an exchangeable and Hoeffding decomposable sequence with values in $D = \{d_1, \dots, d_m\}$, verifying* (50) *and with de Finetti measure γ. Then, the law of $\mathbf{X}_{[1,\infty)}$ is completely determined by the class $\mathfrak{A}_\gamma(1, 1)$* [*as defined in* (58) *for $j = k = 1$*]. *The fact that the law of $\mathbf{X}_{[1,\infty)}$ is completely determined by $\mathfrak{A}_\gamma(1, 1)$ must be interpreted in the following sense: suppose that $\mathbf{X}^*_{[1,\infty)}$ is another exchangeable and Hoeffding decomposable sequence, verifying* (50) *and with de Finetti measure γ^*; then, the equality $\mathfrak{A}_{\gamma^*}(1, 1) = \mathfrak{A}_\gamma(1, 1)$*

implies necessarily that $\gamma = \gamma^*$, *and therefore that* $\mathbf{X}^*_{[1,\infty)}$ *has the same law as* $\mathbf{X}_{[1,\infty)}$.

It is easily seen that the class $\mathfrak{A}_\gamma(1,1)$ can be always expressed in terms of the mean vector $M^\gamma = (M_1^\gamma, \dots, M_m^\gamma)$ and the covariance matrix $V^\gamma = \{V^\gamma(i,j) : 1 \le i, j \le m\}$ of a random probability measure $p = \{p\{d_1\}, \dots, p\{d_m\}\}$ with law γ; these objects are defined as

$$M_i^\gamma = \int p\{d_i\}\,d\gamma \quad \text{and} \quad V^\gamma(i,j) = \int p\{d_i\}p\{d_j\}\,d\gamma,$$

where the integration $d\gamma$ is implicitly performed with respect to the marginal laws of $p\{d_i\}$ and $(p\{d_i\}, p\{d_j\})$ induced by γ. Theorem 10 can therefore be rephrased by saying that the law of a Hoeffding decomposable sequence is completely determined by the means, the variances and the covariances associated with its de Finetti measure. As already evoked in the Introduction, this last conclusion is equivalent, in the case $m = 2$, to some of the findings contained in the "Step 2" of the proof of Theorem 1 (see Section 5), where the moments of the de Finetti measure associated with a $\{0,1\}$-valued Hoeffding decomposable sequence were shown to be uniquely determined (via a recurrence relation) by its first and second moment. In particular, it is not difficult to obtain an alternative proof of Theorem 1, by combining Theorem 10 with the results contained in the "Step 3" of Section 5.

REMARK. Even in the case $m = 2$, and unlike formula (37) of Section 5, Theorem 10 only ensures that the law of a Hoeffding decomposable sequence is determined by the quantities M^γ and V^γ, but does not give any explicit representation of the recursive relation linking the moments of such a sequence.

In the following section we will discuss the extent to which Theorem 10 can be used to characterize the class of Hoeffding decomposable sequences with values in D.

7.2. *Pólya urns, normalized random measures and Theorem* 10. Let ν be a positive measure on $\mathbb{R}^+$ such that $\int_{\mathbb{R}^+} \min(1,x)\nu(dx) < +\infty$ and $\nu(\mathbb{R}^+) = +\infty$. We shall consider a vector of strictly positive numbers $\alpha = (\alpha_1, \dots, \alpha_m) \in (0,+\infty)^m$, as well as a vector of *independent and infinitely divisible* random variables $(\xi_1, \dots, \xi_m)$ with the following property: for every $i = 1, \dots, m$ and every $\lambda > 0$

$$\mathbb{E}[\exp(-\lambda\xi_i)] = \exp[-\alpha_i \psi(\lambda)],$$

where $\psi(\lambda) = \int_{\mathbb{R}^+}(1 - e^{-\lambda x})\nu(dx)$. It is easily seen that our assumptions imply that $\mathbb{P}\{\xi_i > 0\} = 1$ for every $i = 1, \dots, m$. It follows that, for any choice of α and ν as above, the collection of random variables

$$p^{\alpha,\nu}\{d_i\} \triangleq \frac{\xi_i}{\xi_0}, \qquad i = 1, \dots, m, \tag{69}$$

where $\xi_0 = \sum_{i=1}^m \xi_i$, is a well-defined random probability on $D = \{d_1, \dots, d_m\}$. The probability defined in (69) is a special case of the *normalized homogeneous random measures with independent increments* (normalized HRMI) studied, for example, in James, Lijoi and Prünster (2006). In particular the probability $p^{\alpha,\nu}$ can always be obtained by an appropriate time-change and renormalization of a subordinator (i.e., an increasing Lévy process) with no drift. We refer the reader to Pitman (1996) for a discussion of the relations between normalized HRMI and species sampling models, and to Regazzini (1978), Regazzini, Lijoi and Prünster (2003) and James, Lijoi and Prünster (2006) for a description of the role of normalized HRMI in Bayesian nonparametric statistics.

We will also need the following generalization of Definition C of Section 3.

DEFINITION D. Fix $m \geq 2$ and denote by $\Sigma_{m-1} = \{(\theta_1, \dots, \theta_{m-1}) \in [0,1]^{m-1} : \sum_{i=1}^{m-1} \theta_i \leq 1\}$ the simplex of order $m-1$. Let $\mathbf{X}_{[1,\infty)}$ be an exchangeable sequence with de Finetti measure γ. Then, $\mathbf{X}_{[1,\infty)}$ is said to be an (*m-color*) *Pólya sequence with values in* $D = \{d_1, \dots, d_m\}$, if there exists a vector of strictly positive numbers $\alpha = (\alpha_1, \dots, \alpha_m) \in (0, \infty)^m$ such that, for every Borel set $C \subset \Sigma_{m-1}$,

$$\begin{aligned} &\gamma[(p\{d_1\}, \dots, p\{d_{m-1}\}) \in C] \\ &\quad = \frac{1}{B(\alpha)} \int_C \left(\prod_{i=1}^{m-1} \theta_i^{\alpha_i - 1} \right) \left(1 - \sum_{i=1}^{m-1} \theta_i \right)^{\alpha_m - 1} d\theta_1 \cdots d\theta_{m-1}, \end{aligned} \tag{70}$$

where $B(\alpha) = \prod_{i=1}^m \Gamma(\alpha_i)/\Gamma(\sum_{i=1}^m \alpha_i)$, and $\Gamma(\cdot)$ is the usual Gamma function. Note that (70) completely determines the distribution of the vector $(p\{d_1\}, \dots, p\{d_m\})$, since $p\{d_m\} = 1 - \sum_{i=1}^{m-1} p\{d_i\}$ by definition. A random probability measure $p = \{p\{d_1\}, \dots, p\{d_m\}\}$ such that $(p\{d_1\}, \dots, p\{d_{m-1}\})$ has the law γ given in (70) is said to have a *Dirichlet distribution* of parameters $\alpha_1, \dots, \alpha_m$. Note that, in the case $m = 2$ and $D = \{0, 1\}$, the just given definition of an m-color Pólya sequence coincides with that of a two-color Pólya sequence with values in $\{0, 1\}$ and parameters α_1, α_2, as provided in Definition C.

The following well-known result shows that Dirichlet random measures are indeed a special case of normalized HRMI with finite support [as the one defined in (69)], obtained by considering normalized vectors of independent Gamma random variables [see, e.g., James, Lijoi and Prünster (2006)].

PROPOSITION 11. *A random probability* $p = \{p\{d_1\}, \dots, p\{d_m\}\}$ *has a Dirichlet distribution with parameter* α *if, and only if,* $p \overset{\text{law}}{=} p^{\alpha,\nu}$, *where* $p^{\alpha,\nu}$ *is the random probability defined in* (69) *for* $\nu(dx) = x^{-1}e^{-x}\,dx$.

Now let $p = \{p\{d_1\}, \dots, p\{d_m\}\}$ have a Dirichlet distribution of parameters $\alpha_1, \dots, \alpha_m$ and write $\alpha_0 = \sum_{i=1}^m \alpha_i$ and $\mu_i = \alpha_i/\alpha_0$. The following classic relations [see, e.g., James, Lijoi and Prünster (2006)] provide the explicit expressions of the mean and of the covariance matrix of p:

$$\mathbb{E}(p\{d_i\}) = \mu_i, \qquad \mathbf{Var}\,(p\{d_i\}) = \frac{\mu_i(1-\mu_i)}{\alpha_0+1}, \qquad i = 1, \dots, m, \tag{71}$$

$$\mathbf{Cov}\,(p\{d_i\}, p\{d_j\}) = \frac{-\mu_i\mu_j}{\alpha_0+1}, \qquad 1 \le i \ne j \le n. \tag{72}$$

Now observe that, due again to Corollary 9 in Peccati (2004), every m-color Pólya sequence in the sense of Definition D and every i.i.d. sequence with values in D is Hoeffding decomposable. In the light of this result, one would be tempted to use Theorem 10 to deduce (as we did in the proof of Theorem 1) that every Hoeffding decomposable sequence with values in D is either i.i.d. or Pólya, by first showing that the means and covariances of any de Finetti measure verifying (50) can be "replicated" by those of an appropriate Dirichlet or Dirac distribution. However, it is not difficult to see that this last claim is not true, as for $m \ge 3$ there are examples of random probability measures whose associated exchangeable sequence verifies (50), and whose distribution neither is Dirac nor is compatible with (71) and (72) for *any* choice of $\alpha_1, \dots, \alpha_m > 0$. For instance, for $m \ge 3$, any random probability $p = \{p\{d_1\}, \dots, p\{d_m\}\}$ such that there exists one parameter $p\{d_i\}$ which is deterministic and in $(0, 1)$ (the others being random and nonzero) verifies (50) and has a covariance matrix which is not compatible with the second equality in (71), since in this last relation only strictly positive variances are allowed. Another example of a nonreplicable covariance structure is obtained by considering a random probability $p = \{p\{d_1\}, \dots, p\{d_m\}\}$ such that $p\{d_1\}$ is uniform on $(0, 1/4)$ and $p\{d_1\} \overset{\text{a.s.}}{=} p\{d_2\}$; indeed, in this case $\mathbf{Cov}\,(p\{d_1\}, p\{d_2\}) = \mathbf{Var}\,(p\{d_1\}) > 0$, whereas (72) only allows for negative covariances.

Nonetheless, the next result shows that m-color Pólya sequences are the only Hoeffding decomposable sequences among those having a de Finetti measure equal to the law of an object such as $p^{\alpha,\nu}$ in (69). This is the announced partial generalization of Theorem 1.

THEOREM 12. *Let $\mathbf{X}_{[1,\infty)}$ be a D-valued exchangeable and Hoeffding decomposable sequence with de Finetti measure γ. Suppose that γ is equal to the law of the random probability $p^{\alpha,\nu}$ for some measure ν and some $\alpha = (\alpha_1, \dots, \alpha_m) \in (0, +\infty)^m$. Then, there exists $\alpha^* = (\alpha_1^*, \dots, \alpha_m^*) \in (0, +\infty)^m$ such that γ equals the law of p^{α^*,ν^*}, where $\nu^*(dx) = x^{-1}e^{-x}\,dx$. This implies that $\mathbf{X}_{[1,\infty)}$ is an m-color Pólya sequence of parameters $\alpha_1^*, \dots, \alpha_m^*$.*

PROOF. According to Proposition 1 in James, Lijoi and Prünster (2006), for $p^{\alpha,\nu}$ as in the statement, there always exists a constant $\mathcal{I}(\alpha, \nu) \in (0, 1)$ such that, by setting $\alpha_0 = \Sigma\alpha_i$,

$$\mathbb{E}(p^{\alpha,\nu}\{d_i\}) = \frac{\alpha_i}{\alpha_0}, \qquad \mathbf{Var}\,[p^{\alpha,\nu}\{d_i\}] = \frac{\alpha_i}{\alpha_0}\left(1 - \frac{\alpha_i}{\alpha_0}\right)\mathcal{I}(\alpha,\nu),$$
$$\mathbf{Cov}\,\{p^{\alpha,\nu}\{d_i\}, p^{\alpha,\nu}\{d_j\}\} = -\frac{\alpha_i}{\alpha_0}\frac{\alpha_j}{\alpha_0} \times \mathcal{I}(\alpha,\nu), \qquad i \neq j.$$

It follows from (71) and (72) that $\mathfrak{A}_\gamma(1,1) = \mathfrak{A}_{\gamma^*}(1,1)$, where γ^* is the law of a Dirichlet probability measure with parameters

$$\alpha_i^* = (\alpha_i/\alpha_0)\big(\mathcal{I}(\alpha,\nu)^{-1} - 1\big).$$

The conclusion is obtained by applying Theorem 10. □

Acknowledgments. We wish to express our sincere gratitude to an anonymous Associate Editor for inspiring the approach to Hoeffding decompositions adopted in Section 7. We are also indebted to an anonymous referee for several insightful remarks. We wish to thank Igor Prünster for a number of fundamental discussions as this work progressed.

REFERENCES

ALDOUS, D. J. (1985). Exchangeability and related topics. In *École d'Été de Probabilités de Saint-Flour XIII—1983. Lecture Notes in Math.* **1117** 1–198. Springer, Berlin. MR883646

ARCONES, M. A. and GINÉ, E. (1993). Limit theorems for U-processes. *Ann. Probab.* **21** 1494–1542. MR1235426

BENTKUS, V., GÖTZE, F. and VAN ZWET, W. R. (1997). An Edgeworth expansion for symmetric statistics. *Ann. Statist.* **25** 851–896. MR1439326

BLACKWELL, D. (1973). Discreteness of Ferguson selections. *Ann. Statist.* **1** 356–358. MR0348905

BLACKWELL, D. and MACQUEEN, J. B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353–355. MR0362614

BLOZNELIS, M. (2005). Orthogonal decomposition of symmetric functions defined on random permutations. *Combin. Probab. Comput.* **14** 249–268. MR2138112

BLOZNELIS, M. and GÖTZE, F. (2001). Orthogonal decomposition of finite population statistics and its applications to distributional asymptotics. *Ann. Statist.* **29** 899–917. MR1865345

BLOZNELIS, M. and GÖTZE, F. (2002). An Edgeworth expansion for symmetric finite population statistics. *Ann. Probab.* **30** 1238–1265. MR1920107

DIACONIS, P. and YLVISAKER, D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7** 269–281. MR520238

FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230. MR0350949

GRIFFITHS, R. C. (1979). A transition density expansion for a multi-allele diffusion model. *Adv. in Appl. Probab.* **11** 310–325. MR526415

HÁJEK, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Statist.* **39** 325–346. MR0222988

HILL, B. M., LANE, D. and SUDDERTH, W. (1987). Exchangeable urn processes. *Ann. Probab.* **15** 1586–1592. MR905350

HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325. MR0026294

KARLIN, S. and RINOTT, Y. (1982). Applications of ANOVA type decompositions for comparisons of conditional variance statistics including jackknife estimates. *Ann. Statist.* **10** 485–501. MR653524

JAMES, L. F., LIJOI, A. and PRÜNSTER, I. (2006). Conjugacy as a distinctive feature of the Dirichlet process. *Scand. J. Statist.* **33** 105–120. MR2255112

MAJOR, P. (2005). Tail behaviour of multiple random integrals and U-statistics. *Probab. Surv.* **2** 448–505 (electronic). MR2203678

PECCATI, G. (2003). Hoeffding decompositions for exchangeable sequences and chaotic representation of functionals of Dirichlet processes. *C. R. Math. Acad. Sci. Paris* **336** 845–850. MR1990026

PECCATI, G. (2004). Hoeffding-ANOVA decompositions for symmetric statistics of exchangeable observations. *Ann. Probab.* **32** 1796–1829. MR2073178

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

PITMAN, J. (1996). Some developments of the Blackwell–MacQueen urn scheme. In *Statistics, Probability and Game Theory*. *IMS Lecture Notes Monogr. Ser.* **30** 245–267. IMS, Hayward, CA. MR1481784

PITMAN, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Springer, Berlin. MR2245368

REGAZZINI, E. (1978). Intorno ad alcune questioni relative alla definizione del premio secondo la teoria della credibilità. *Giornale dell'Istituto Italiano degli Attuari* **41** 77–89.

REGAZZINI, E., LIJOI, A. and PRÜNSTER, I. (2003). Distributional results for means of normalized random measures with independent increments. *Ann. Statist.* **31** 560–585. Dedicated to the memory of Herbert E. Robbins. MR1983542

STANLEY, R. P. (1997). *Enumerative Combinatorics*. **1**. *Cambridge Studies in Advanced Mathematics* **49**. Cambridge Univ. Press, Cambridge. With a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original. MR1442260

VITALE, R. A. (1992). Covariances of symmetric statistics. *J. Multivariate Anal.* **41** 14–26. MR1156678

ZHAO, L. C. and CHEN, X. R. (1990). Normal approximation for finite-population U-statistics. *Acta Math. Appl. Sinica* (*English Ser.*) **6** 263–272. MR1078067

LABORATOIRE DE STATISTIQUE
THÉORIQUE ET APPLIQUÉE
UNIVERSITÉ PARIS VI
175 RUE DU CHEVALERET
75013 PARIS
FRANCE
E-MAIL: omar.eldakkak@gmail.com

LABORATOIRE DE STATISTIQUE
THÉORIQUE ET APPLIQUÉE
UNIVERSITÉ PARIS VI
175, RUE DU CHEVALERET
75013 PARIS
AND
EQUIPE MODAL′ X
UNIVERSITÉ PARIS OUEST - NANTERRE LA DÉFENSE
92000 NANTERRE
FRANCE
E-MAIL: giovanni.peccati@gmail.com