Article

Mixture of Species Sampling Models

Federico Bassetti [†] and Lucia Ladelli *,[†]

Department of Mathematics, Politecnico of Milano, 20133 Milano, Italy; federico.bassetti@polimi.it
* Correspondence: lucia.ladelli@polimi.it
† These authors contributed equally to this work.

Abstract: We introduce mixtures of species sampling sequences (mSSS) and discuss how these sequences are related to various types of Bayesian models. As a particular case, we recover species sampling sequences with general (not necessarily diffuse) base measures. These models include some "spike-and-slab" non-parametric priors recently introduced to provide sparsity. Furthermore, we show how mSSS arise while considering hierarchical species sampling random probabilities (e.g., the hierarchical Dirichlet process). Extending previous results, we prove that mSSS are obtained by assigning the values of an exchangeable sequence to the classes of a latent exchangeable random partition. Using this representation, we give an explicit expression of the Exchangeable Partition Probability Function of the partition generated by an mSSS. Some special cases are discussed in detail—in particular, species sampling sequences with general base measures and a mixture of species sampling sequences with Gibbs-type latent partition. Finally, we give explicit expressions of the predictive distributions of an mSSS.

Keywords: species sampling models; exchangeable random partitions; exchangeable sequences; predictive distributions

1. Introduction

Discrete random measures have been widely used in Bayesian nonparametrics. Noteworthy examples of such random measures are the Dirichlet process [1], the Pitman–Yor process [2,3], (homogeneous) normalized random measures with independent increments (see, e.g., [4–7]), Poisson–Kingman random measures [8] and stick-breaking priors [9]. All the previous random measures are of the form

$$P = \sum_{j \geq 1} p_j^{\downarrow} \delta_{Z_j}, \tag{1}$$

where $(Z_j)_{j \geq 1}$ are i.i.d. random variables taking values in a Polish space $(\mathbb{X}, \mathcal{X})$ with common distribution H, and $(p_j^{\downarrow})_{j \geq 1}$ are random positive weights in $[0, 1]$, independent of $(Z_j)_{j \geq 1}$, such that $p_1^{\downarrow} \geq p_2^{\downarrow} \geq p_3^{\downarrow} \geq \ldots$.

With a few exceptions—see, e.g., [1,4,10–14]—the *base measure* H of a random measure P in (1) is usually assumed to be diffuse, since this simplifies the derivation of various analytical results.

The diffuseness of H is assumed also to define the so-called *species sampling sequences* [15], exchangeable sequences whose directing measure is a discrete random probability of type (1). In this case, the diffuseness of H is motivated by the interpretation of species sampling sequences as sequences describing a sampling mechanism in discovering species from an unknown population. In this context, the Z_js are the possible infinite different species, and the diffuseness of H ensures that there is no redundancy in this description.

On the other hand, from a Bayesian point of view, the diffuseness of H is not always reasonable and there are situations in which a discrete (or mixed) H is indeed natural. For example, recent interest in species sampling models with a spike-and-slab base measure

emerged in [16–21] in order to induce sparsity and facilitate variable selection. Other models, which are implicitly related to species sampling sequences with non-diffused base measures, are mixtures of Dirichlet processes [10] and hierarchical random measures; see, e.g., [22–25].

The combinatorial structure of species sampling sequences derived from random measure (1) with general H have been recently studied in [14].

In this paper, we discuss some relevant properties of species sampling sequences with general base measures, as well as some further generalizations, namely mixtures of species sampling sequences with general base measures (mSSS).

An mSSS is an exchangeable sequence whose directing random measure is of type (1), where $(Z_n)_{n\geq 1}$ is a sequence of exchangeable random variables and $(p_n^\downarrow)_{n\geq 1}$ are random positive weights in $[0,1]$ with $p_1^\downarrow \geq p_2^\downarrow \geq p_3^\downarrow \geq \ldots$, independent of $(Z_n)_{n\geq 1}$.

The core of the results that we prove in this paper is that all the mSSS can be obtained by assigning the values of an exchangeable sequence to the classes of a latent exchangeable random partition. We summarize the results of Section 3 in the next statement.

The following are equivalent:

1. $\xi = (\xi_n)_{n\geq 1}$ *is an mSSS;*
2. *with probability one* $(\xi_n)_{n\geq 1} = (Z_{I_n})_{n\geq 1}$, *where* $(I_n)_{n\geq 1}$ *is a sequence of integer-valued random variables independent of the Zs such that, conditionally on* $p^\downarrow := (p_1^\downarrow, p_2^\downarrow, \ldots)$, *the* I_n *are independent and* $\mathbb{P}\{I_n = i | p^\downarrow\} = p_i^\downarrow$.
3. *with probability one* $(\xi_n)_{n\geq 1} := (Z'_{\mathscr{C}_n(\Pi)})_{n\geq 1}$, *where* $(Z'_n)_{n\geq 1}$ *is an exchangeable sequence with the same law of* $(Z_n)_{n\geq 1}$, Π *is an exchangeable partition, independent of* $(Z'_n)_{n\geq 1}$, *obtained by sampling from* $(p_n^\downarrow)_{n\geq 1}$, *and* $\mathscr{C}_n(\Pi)$ *is the index of the block in* Π *containing* n.

The partition Π obtained from $p^\downarrow = (p_1^\downarrow, p_2^\downarrow, \ldots)$ is the so-called paint-box process associated with $p^\downarrow$. In general, this partition, called the latent partition, does not coincide with the partition induced by the $(\xi_n)_{n\geq 1}$. Note that also the sequence $(Z'_n)_{n\geq 1}$ is latent, in the sense that it cannot be obtained if only $(\xi_n)_{n\geq 1}$ is known. On the other hand, combining the information contained in $(Z'_n)_{n\geq 1}$ and in Π, one obtains complete knowledge of $(\xi_n)_{n\geq 1}$, and, in particular, of its clustering behavior. This last observation is essential for the development of all the other results presented in our paper.

The rest of the paper is organized as follows. Section 2 reviews some important results on species sampling models and exchangeable random partitions. Section 3 introduces mixtures of species sampling sequences and discusses how these sequences are related to various types of Bayesian models. In the same section, the stochastic representations for mixtures of species sampling sequences sketched above are proven. In Section 4, we provide an explicit expression of the *Exchangeable Partition Probability Function* (EPPF) of the partition generated by such sequences. This result is achieved considering two EPPFs arising from a suitable latent partition structure. Some special cases are further detailed. Finally, Section 5 deals with the predictive distributions of mixtures of species sampling sequences.

2. Background Materials

In this section, we briefly review some basic notions of exchangeable random partitions and species sampling models.

2.1. Exchangeable Random Partitions

A partition π_n of $[n] := \{1, \ldots, n\}$ is an unordered collection $\{\pi_{1,n}, \ldots, \pi_{k,n}\}$ of disjoint non-empty subsets (blocks) of $\{1, \ldots, n\}$ such that $\cup_{j=1}^k \pi_{j,n} = [n]$. A partition $\pi_n = \{\pi_{1,n}, \pi_{2,n}, \ldots, \pi_{k,n}\}$ has $|\pi_n| := k$ blocks (with $1 \leq |\pi_n| \leq n$) and $|\pi_{c,n}|$, with $c = 1, \ldots, k$, is the number of elements of the block c. We denote by $\mathcal{P}_n$ the collection of all partitions of $[n]$ and, given a partition, we list its blocks in ascending order of their smallest element, i.e., *in order of their appearance*. For instance, we write $[(1, 3), (2, 4), (5)]$ and not $[(2, 4), (3, 1), (5)]$.

A sequence of random partitions, $\Pi = (\Pi_n)_{n\geq 1}$, defined on a common probability space, is called a *random partition of* $\mathbb{N}$ if, for each n, the random variable Π_n takes values in $\mathcal{P}_n$ and, for $m < n$, the restriction of Π_n to $\mathcal{P}_m$ is Π_m (*consistency property*).

In order to define an *exchangeable* random partition, given a permutation ρ of $[n]$ and a partition π_n in $\mathcal{P}_n$, we denote by $\rho(\pi_n)$ the partition with blocks $\{\rho(j) : j \in \pi_{i,n}\}$ for $i = 1, \ldots, |\pi_n|$. A random partition of $\mathbb{N}$ is said to be *exchangeable* if Π_n has the same distribution of $\rho(\Pi_n)$ for every n and every permutation ρ of $[n]$. In other words, its law is invariant under the action of all permutations (acting on Π_n in the natural way).

The law of any exchangeable random partition on $\mathbb{N}$ is completely characterized by its *Exchangeable Partition Probability Function* (EPPF); in other words, there exists a unique symmetric function $\mathfrak{q}$ on the integers such that, for any partition π_n in $\mathcal{P}_n$,

$$\mathbb{P}\{\Pi_n = \pi_n\} = \mathfrak{q}(|\pi_{1,n}|, \ldots, |\pi_{k,n}|) \tag{2}$$

where k is the number of blocks in π_n. In the following, we shall write $\Pi \sim \mathfrak{q}$ to denote an exchangeable partition of $\mathbb{N}$ with EPPF $\mathfrak{q}$. Note that an EPPF is indeed a family of symmetric functions $\mathfrak{q}_k^n(\cdot)$ defined on $\mathcal{C}_{n,k} = \{(n_1, \ldots, n_k) \in \mathbb{N}^k : \sum_{i=1}^k n_i = n\}$. To simplify the notation, we write $\mathfrak{q}$ instead of $\mathfrak{q}_k^n$. Alternatively, one can assume that $\mathfrak{q}$ is a function on $\cup_{n\in\mathbb{N}} \cup_{k=1}^n \mathcal{C}_{n,k}$. See [26].

Given a sequence of random variables $X = (X_j)_{j\geq 1}$ taking values in some measurable space, the random partition $\Pi^*(X)$ induced by X is defined as the random partition obtained by the equivalence classes under the random equivalence relation $i(\omega) \sim j(\omega)$ if and only if $X_i(\omega) = X_j(\omega)$. One can check that a partition induced by an exchangeable random sequence is an exchangeable random partition.

Recall that, by de Finetti's theorem, a sequence $X = (X_n)_{n\geq 1}$ taking values in a Polish space $(\mathbb{X}, \mathcal{X})$ is exchangeable if and only if the X_ns, given some random probability measure Q on $\mathcal{X}$, are conditionally independent with common distribution Q. Moreover, the random probability Q, known as the *directing random measure of* X, is the almost sure limit (with respect to weak convergence) of the empirical process $\frac{1}{n}\sum_{i=1}^n \delta_{X_i}$.

Based on de Finetti's theorem, Kingman's correspondence theorem sets up a one-to-one map between the law of an exchangeable random partition on $\mathbb{N}$ (i.e., its EPPF) and the law of random ranked weights $p^\downarrow = (p_j^\downarrow)_{j\geq 1}$ satisfying $1 \geq p_1^\downarrow \geq p_2^\downarrow \geq \cdots \geq 0$ and $\sum_j p_j^\downarrow \leq 1$ (with probability one). To state the theorem, recall that a partition Π is said to be generated by a (possibly random) $p^\downarrow$, if it is generated by a sequence $(I_n)_{n\geq 1}$ of integer-valued random variables that are conditionally independent given $p^\downarrow$ with conditional distribution

$$\mathbb{P}\{I_n = i | p^\downarrow\} := \begin{cases} 1 - \sum_{j\geq 1} p_j^\downarrow & \text{if } i = -n \\ p_i^\downarrow & \text{if } i \geq 1, \end{cases} \tag{3}$$

Note that $1 - \sum_{j\geq 1} p_j^\downarrow$ is the magnitude of the so-called "dust" component; indeed, each I_n sampled from this part, i.e., $I_n = -n$, contributes to a singleton n in the partition Π. A consequence is that if $\sum_{j\geq 1} p_j^\downarrow = 1$ a.s., the partition Π has no singleton. The partition $\Pi^*(I)$ is sometimes referred to as the $p^\downarrow$-paintbox process; see [27].

Let $\nabla := \{p_j^\downarrow \in [0,1] : p_1^\downarrow \geq p_2^\downarrow \geq \ldots, \sum_{j\geq 1} p_j^\downarrow \leq 1\}$. We are now ready to state Kingman's theorem.

Theorem 1 ([28]). *Given any exchangeable random partition Π with EPPF $\mathfrak{q}$, denote by $\Pi_{j,n}^\downarrow$ the blocks of the partition rearranged in decreasing order with respect to the number of elements in the blocks of Π_n. Then,*

$$\lim_n \left(\frac{|\Pi_{j,n}^\downarrow|}{n}\right)_{j\geq 1} = (p_j^\downarrow)_{j\geq 1} \quad a.s. \tag{4}$$

for some random $p^{\downarrow} = (p_j^{\downarrow})_{j \geq 1}$ taking values in ∇. Moreover, on a possibly enlarged probability space, there is a sequence of integer-valued random variables $I = (I_n)_{n \geq 1}$, conditionally independent given $p^{\downarrow}$, such that (3) holds and the partition induced by I is equal to Π a.s.

Kingman's theorem is usually stated in a slightly weaker form (see, e.g., Theorem 2.2 in [26]) and the equality between $\Pi^*(I)$ and Π is given in law. The previous "almost sure" version can be easily derived by inspecting the proof of Kingman's theorem given in [29].

A consequence of the previous theorem is that $\mathfrak{q} \mapsto \mathrm{Law}(p^{\downarrow})$ for $p^{\downarrow}$ in (4) defines a bijection from the set of the EPPF and the laws on ∇.

If $p^{\downarrow}$ is proper, i.e., $\sum_{j \geq 1} p_j^{\downarrow} = 1$ a.s., then Kingman's correspondence between $p^{\downarrow}$ and the EPPF $\mathfrak{q}$ can be made explicit by

$$\mathfrak{q}(n_1, \ldots, n_k) = \sum_{(j_1, \ldots, j_k) \in \mathbb{N}_k} \mathbb{E}\Big[\prod_{i=1}^{k} (p_{j_i}^{\downarrow})^{n_i} \Big]. \tag{5}$$

where $\mathbb{N}_k$ is the set of all ordered k-tuples of distinct positive integers. See Chapter 2 [26].

Given an EPPF $\mathfrak{q}$, one deduces the corresponding sequence of predictive distributions, which is the sequence of conditional distributions

$$P\{\Pi_{n+1} = \pi_{n+1} | \Pi_n = \pi_n\}$$

when $\Pi \sim \mathfrak{q}$. Starting with $\Pi_1 = \{1\}$, given $\Pi_n = \pi_n$ (with $|\pi_n| = k$), the conditional probability of adding a new block (containing $n+1$) to Π_n is

$$\nu_n(\pi_n) = \nu_n(|\pi_{1,n}|, \ldots, |\pi_{k,n}|) := \frac{\mathfrak{q}(|\pi_{1,n}|, \ldots, |\pi_{k,n}|, 1)}{\mathfrak{q}(|\pi_{1,n}|, \ldots, |\pi_{k,n}|)}; \tag{6}$$

while the conditional probability of adding $n+1$ to the ℓ-th block of Π_n (for $\ell = 1, \ldots, k$) is

$$\omega_{n,\ell}(\pi_n) = \omega_{n,\ell}(|\pi_{1,n}|, \ldots, |\pi_{k,n}|) := \frac{\mathfrak{q}(|\pi_{1,n}|, \ldots, |\pi_{\ell,n}| + 1, \ldots, |\pi_{k,n}|)}{\mathfrak{q}(|\pi_{1,n}|, \ldots, |\pi_{k,n}|)}. \tag{7}$$

2.2. Species Sampling Models

A *species sampling random probability* (SSrp) is a random probability of the form

$$P = \sum_{j \geq 1} p_j \delta_{Z_j} + (1 - \sum_{j \geq 1} p_j) H \tag{8}$$

where $(Z_j)_{j \geq 1}$ are i.i.d. random variables taking values in a Polish space $(\mathbb{X}, \mathcal{X})$ with common distribution H, and $(p_j)_{j \geq 1}$ are random positive weights in $[0, 1]$, independent of $(Z_j)_{j \geq 1}$, such that $\sum_{j \geq 1} p_j \leq 1$ with probability one. These random probability measures are also known as *Type III* random probability measures; see [30].

Given the SSrp in (8), let $(p_j^{\downarrow})_{j \geq 1}$ be the ranked sequence obtained from $(p_j)_{j \geq 1}$ rearranging the p_js in decreasing order. One can always write

$$P = \sum_{j \geq 1} p_j^{\downarrow} \delta_{\tilde{Z}_j} + (1 - \sum_{j \geq 1} p_j^{\downarrow}) H \tag{9}$$

where $(\tilde{Z}_j)_{j \geq 1}$ is a suitable random reordering of the original sequence $(Z_j)_{j \geq 1}$. It is easy to check that $(\tilde{Z}_j)_{j \geq 1}$ are i.i.d. random variables with law H independent of $(p_j^{\downarrow})_{j \geq 1}$. Hence, H and the EPPF $\mathfrak{q}$ associated via Kingman's correspondence with $(p_j^{\downarrow})_{j \geq 1}$ completely characterize the law of P, from now on denoted by $SSrp(\mathfrak{q}, H)$.

SSrp with H diffuse are also characterized as directing random measures of a particular type of exchangeable sequences, known as species sampling sequences. Let $\mathfrak{q}$ be an EPPF and H a diffuse probability measure on a Polish space $\mathbb{X}$. An exchangeable sequence $\xi := (\xi_n)_n$ taking values in $\mathbb{X}$ is a *species sampling sequence, SSS*$(\mathfrak{q}, H)$, if the law of $(\xi_n)_n$ is characterized by the predictive system:

- (PS1) $\mathbb{P}\{\xi_1 \in dx\} = H(dx)$;
- (PS2) *the conditional distribution of* ξ_{n+1} *given* $(\xi_1, \dots, \xi_n)$ *is*

$$\mathbb{P}\{\xi_{n+1} \in dx | \xi_1, \dots, \xi_n\} = \sum_{c=1}^{K} \omega_{n,c} \delta_{\xi_c^*}(dx) + \nu_n H(dx),$$

where $(\xi_1^*, \dots, \xi_K^*)$ *is the sequence of distinct observations in order of appearance,* $\omega_{n,c} = \omega_{n,c}(|\Pi_{1,n}|, \dots, |\Pi_{K,n}|)$, $\nu_n = \nu_n(|\Pi_{1,n}|, \dots, |\Pi_{K,n}|)$, $K = |\Pi_n|$, Π_n *is the random partition induced by* $(\xi_1, \dots, \xi_n)$ *and* $\omega_{n,c}$ *and* ν_n *are related to the* $\mathfrak{q}$ *by* (6) *and* (7).

We summarize here some results proven in [15].

Proposition 1 ([15]). *Let H be a diffuse probability measure; then, an exchangeable sequence $(\xi_n)_n$ is characterized by* (PS1)–(PS2) *if and only if its directing random measure is an* $SSrp(\mathfrak{q}, H)$.

As noted in [29], the partition induced by any exchangeable sequence taking values in $\mathbb{X}$ with directing measure $\tilde{\mu}$ depends only on the sequence $\tilde{\mu}(\tilde{x}_j)$, where $\tilde{x}_j$ are the random atoms forming the discrete component of $\tilde{\mu}$ and ordered in such a way that $\tilde{\mu}(\tilde{x}_1) \geq \tilde{\mu}(\tilde{x}_2) \geq \dots$. Combining this observation with the previous proposition, one can see that, when H is diffuse and ξ is an $SSS(\mathfrak{q}, H)$, the partition $\Pi^*(\xi)$ is equal (a.s.) to $\Pi^*(I)$ (where I is defined as in Kingman's theorem) and $\Pi^*(\xi)$ has EPPF $\mathfrak{q}$. Note that [29] defines the $p^{\downarrow}$-paintbox process as any random partition $\Pi^*(\xi)$ where ξ is an exchangeable sequence with directing random measure (9) and H is a diffuse measure.

One can show (see the proof of Proposition 13 in [15]) that an $SSS(\mathfrak{q}, H)$ can be obtained by assigning the values of an i.i.d. sequence $(Z_n)_n$ with distribution H to the classes of an independent exchangeable random partition with EPPF $\mathfrak{q}$. More formally, for a random partition Π, let $\mathscr{C}_n(\Pi)$ be the random index denoting the block containing n, i.e.,

$$\mathscr{C}_n(\Pi) = c \text{ if } n \in \Pi_{c,n}$$

or equivalently if $n \in \Pi_{c,j}$ for some (and hence all) $j \geq n$. If $Z' = (Z'_n)_n$ is an i.i.d. sequence with law H (diffuse), Π is an exchangeable partition with $\Pi \sim \mathfrak{q}$, and Z' and Π are stochastically independent, then

$$(\xi_n)_{n \geq 1} := (Z'_{\mathscr{C}_n(\Pi)})_{n \geq 1} \tag{10}$$

is an $SSS(\mathfrak{q}, H)$. Note that the Z'_ns appearing in (10) are not the same Z_ns of (8), although they have the same law.

It is worth mentioning that the original characterization given in [15] of species sampling sequences is stronger than the one summarized here. Indeed, the original definition of SSS is given using a slightly weaker predictive assumption. For details, see Proposition 13 and the discussion following Proposition 11 in [15].

In summary, when H is diffuse, one can build a species sampling sequence $(\xi_n)_n$ by one of the following equivalent constructions:

- using the system of predictive distributions (PS1)–(PS2);
- sampling (conditionally) i.i.d. variables from (8);
- combining an i.i.d. sequence from H with an exchangeable random partition by (10).

3. Mixture of Species Sampling Models

We now discuss some possible generalizations of the notion of species sampling sequences and we show that the three constructions presented above are no more equivalent in this setting.

3.1. Definitions and Relation to Other Models

Exchangeable sequences sampled from an *SSrm* with a general base measure, also known as *generalized species sampling sequences (gSSS)*, have been introduced and studied in [14,25].

Definition 1 (*gSSS*($\mathfrak{q}, H$)). *$(\xi_n)_{n \geq 1}$ is a gSSS($\mathfrak{q}, H$) if it is an exchangeable sequence with directing random measure P, where $P \sim SSrp(\mathfrak{q}, H)$, H being any measure on $(\mathbb{X}, \mathcal{X})$ (not necessarily diffuse).*

Clearly, a $gSSS(\mathfrak{q}, H)$ with H diffuse is an $SSS(\mathfrak{q}, H)$. On the contrary, if ξ is a $gSSS(\mathfrak{q}, H)$ with H non-diffuse, (PS1)–(PS2) are no longer true. Moreover, the EPPF of the random partition induced by ξ with H non-diffuse is not $\mathfrak{q}$. The relation between the partition induced by ξ and $\mathfrak{q}$ has been studied in [14].

In [25], the definition of $gSSS(\mathfrak{q}, H)$ with H not necessarily diffuse was motivated by an interest in defining the class of the so-called *hierarchical species sampling models*. If $\xi_1, \xi_2, \dots$ are exchangeable random variables with a directing random measure of hierarchical type, one has that

$$\xi_n | P_1, P_0 \overset{i.i.d.}{\sim} P_1 \quad n \geq 1$$
$$P_1 | P_0 \sim SSrp(\mathfrak{q}, P_0)$$
$$P_0 \sim SSrp(\mathfrak{q}_0, H_0).$$

In order to understand why the general definition of $gSSS(\mathfrak{q}, H)$ is useful in this context, note that, even if H_0 is diffuse and $\mathfrak{q}_0$ is proper (i.e., the $p^\downarrow$ associated with $\mathfrak{q}_0$ by Kingman's correspondence are proper), the conditional distribution of $[\xi_n]_{n \geq 1}$ given P_0 is not an SSS, since P_0 is a.s. a purely atomic probability measure on $\mathcal{X}$. Moreover, assuming that $\mathfrak{q}$ is proper, we can write

$$P_1 = \sum_j p_{j1} \delta_{Z_j}$$

where Z_j are conditionally i.i.d. with common distribution P_0, given P_0, and $(p_{j1})_j$ are associated by Kingman's correspondence with the EPPF $\mathfrak{q}$. In other words, in this case, $(\xi_n)_{n \geq 1}$ are exchangeable with directing random measure $P_1 = \sum_j p_{j1} \delta_{Z_j}$, where $(p_{j1})_j$ and $(Z_j)_j$ are independent and $(Z_j)_j$ are exchangeable with directing measure $P_0 \sim SSrp(\mathfrak{q}_0, H_0)$.

The previous observation suggests a further generalization of species sampling sequences.

Definition 2 (*mSSS*). *We say that $(\xi_n)_{n \geq 1}$ is a mixture of species sampling sequences (mSSS) if it is an exchangeable sequence with directing random measure*

$$P = \sum_{j \geq 1} p_j^\downarrow \delta_{Z_j} + (1 - \sum_{j \geq 1} p_j^\downarrow) \tilde{H} \tag{11}$$

where $Z := (Z_n)_{n \geq 1}$ is an exchangeable sequence with directing random measure $\tilde{H}$, $p^\downarrow$ a sequence of random weight in ∇ with EPPF $\mathfrak{q}$ such that $P\{\sum_{j \geq 1} p_j^\downarrow > 0\} > 0$, $(Z, \tilde{H})$ and $p^\downarrow$ are stochastically independent.

First of all, note that $gSSS(\mathfrak{q}, H)$ is a particular case of Definition 2, obtained from a deterministic $\tilde{H} = H$. Moreover, Definition 2 can be seen as a mixture of $gSSS$. Indeed, if $\xi = (\xi_n)_{n \geq 1}$ is as in Definition 2 and $\tilde{H}$ is the directing random measure of $(Z_n)_n$, then the conditional distribution of ξ given $\tilde{H}$ is a $gSSS(\mathfrak{q}, \tilde{H})$. This motivates the name "mixture of species sampling sequences".

It is worth noticing that one can also consider more general mixtures of SSS. The most general mixture one can take into consideration leads to a random probability measure of the form (11), where $Z := (Z_n)_{n \geq 1}$ are exchangeable random variables with directing random measure $\tilde{H}$, $p^{\downarrow}$ is a sequence of random weight in ∇ such that $P\{\sum_{j \geq 1} p_j^{\downarrow} > 0\} > 0$, where $[Z, \tilde{H}]$, and $p^{\downarrow}$ are not necessarily stochastically independent.

As an example of this more general situation, we describe the so-called mixtures of Dirichlet processes as defined in [10]. Recall that a Dirichlet process $P \sim \mathcal{D}ir(\alpha)$ is defined as a random probability measure characterized by the system of finite n-dimensional distributions

$$\mathbb{P}\{(P(A_1), \ldots, P(A_n)) \in \cdot\} = Dir\left(\cdot \, ; \alpha(A_1), \ldots, \alpha(A_n)\right) \quad \forall n \geq 1, \forall A_i \in \mathcal{X}$$

where $Dir(\cdot \, ; a_1, \ldots, a_n)$ is the Dirichlet measure (on the $n-1$ simplex) of parameters $a_1, \ldots, a_n$ and α is a finite σ-additive measure on $\mathcal{X}$. It is well known that a Dirichlet process is an $SSrp(\mathfrak{q}, H)$ for $H(\cdot) = \alpha(\cdot)/\alpha(\mathbb{X})$ and

$$\mathfrak{q}(n_1, \ldots, n_k) = \frac{\alpha(\mathbb{X})^k}{(\alpha(\mathbb{X}))_n} \prod_{c=1}^{k} (n_c - 1)!, \tag{12}$$

where $(x)_n = x(x+1) \ldots (x+n-1)$ is the rising factorial (or Pochhammer polynomial); see [2,31]. A mixture of Dirichlet processes is defined in [10] as a random probability measure P characterized by the n-dimensional distributions

$$\mathbb{P}\{(P(A_1), \ldots, P(A_n)) \in \cdot\} = \int_U Dir\left(\cdot \, ; \alpha_u(A_1), \ldots, \alpha_u(A_n)\right) Q(du) \tag{13}$$

where, now, $(u, A) \mapsto \alpha_u(A)$ is a kernel measure on $U \times \mathcal{X}$ (in particular, $A \mapsto \alpha_u(A)$ is a finte σ-additive measure on $\mathcal{X}$ for every $u \in U$), $(U, \mathcal{U})$ is a (Borel) regular space (e.g., a Polish space) and Q is a probability measure on $\mathcal{U}$.

Using the fact that a Dirichlet process is the $SSrp$ described above, one can prove that any mixture of Dirichlet processes has a representation of the form (11), where $((Z_n)_{n \geq 1} \tilde{H})$ and $p^{\downarrow}$ are stochastically dependent. More precisely, the joint law of $(\tilde{H}, (Z_n)_{n \geq 1}, p^{\downarrow})$ is characterized by the law of the (augmented) random element

$$(\tilde{H}, (Z_n)_{n \geq 1}, p^{\downarrow}, \tilde{u})$$

given by the following:

- $\tilde{u}$ is a random variable taking values in U with law Q;
- $\tilde{H}(\cdot) := \alpha_{\tilde{u}}(\cdot)/\alpha_{\tilde{u}}(\mathbb{X})$;
- $(Z_n)_{n \geq 1}$ are exchangeable random variables with directing random measure $\tilde{H}$;
- $p^{\downarrow}$ is sequence of random weight in ∇ such that $P\{\sum_{j \geq 1} p_j^{\downarrow} = 1\} = 1$ and the conditional distribution of $p^{\downarrow}$ given $\tilde{u}$ depends only on $\alpha_{\tilde{u}}(\mathbb{X})$. In particular, the (conditional) EPPF associated with the law of $p^{\downarrow}$ given $\tilde{u}$ has the form

$$\mathfrak{q}(n_1, \ldots, n_k | \tilde{u}) := \frac{\alpha_{\tilde{u}}(\mathbb{X})^k}{(\alpha_{\tilde{u}}(\mathbb{X}))_n} \prod_{c=1}^{k} (n_c - 1)! \tag{14}$$

Note that the marginal EPPF of the $p^\downarrow$, obtained by integrating (14) with respect to the law of $\tilde{u}$, is

$$\mathfrak{q}(n_1, \ldots, n_k) = \prod_{c=1}^{k} (n_c - 1)! \int_U \frac{\alpha_u(\mathbb{X})^k}{(\alpha_u(\mathbb{X}))_n} Q(du). \tag{15}$$

Without further assumptions, a mixture of Dirichlet processes is a mixture of SSrp with $p^\downarrow$ and $\tilde{H}(\cdot)$ possibly dependent. Nevertheless, with this representation at hand, one can easily deduce that if $(\xi_n)_{n \geq 1}$ is sampled from a mixture of Dirichlet processes under the additional hypothesis that Q is such that $\alpha_{\tilde{u}}(\mathbb{X})$ and $\alpha_{\tilde{u}}(\cdot)/\alpha_{\tilde{u}}(\mathbb{X})$ are independent, then $(\xi_n)_{n \geq 1}$ satisfies Definition 2, with $\tilde{H} = \alpha_{\tilde{u}}(\cdot)/\alpha_{\tilde{u}}(\mathbb{X})$ and $\mathfrak{q}$ given by (15).

In the rest of the paper, we focus our attention on mSSS for which $[Z, \tilde{H}]$ and $p^\downarrow$ are independent.

3.2. Representation Theorems for mSSS

In this section, we give two alternative representations for exchangeable sequences as in Definition 2, which generalize Proposition 1 in [14].

Proposition 2. *An exchangeable sequence $\xi = (\xi_n)_{n \geq 1}$ is an mSSS as in* Definition 2 *if and only if*

$$\xi_n = Z_{I_n} \quad a.s.$$

where $Z^+ = (Z_n)_{n \geq 1}$, $\tilde{H}$ and $p^\downarrow$ are as in Definition 2, $Z^- = (Z_n)_{n \leq -1}$ *are further conditionally (given $\tilde{H}$) i.i.d. random variables with conditional distribution $\tilde{H}$, and $(I_n)_{n \geq 1}$ is a sequence of integer-valued random variables independent of the Zs and $\tilde{H}$, such that, conditionally on $p^\downarrow$, the I_n are independent and (3) holds. All the random elements are defined on a possibly enlarged probability space.*

Proof. Let $\sigma_2 = [Z^+, \tilde{H}, p^\downarrow]$, where Z^+, $\tilde{H}$, $p^\downarrow$ are defined as in Definition 2 (mSSS). Set $\alpha = 1 - \sum_{j \geq 1} p_j^\downarrow$. On a possibly enlarged probability space, let $(Z')^- = (Z'_n)_{n \leq -1}$ be a sequence of random variables conditionally i.i.d. given $\tilde{H}$ with conditional distribution $\tilde{H}$ and let $I' = (I'_n)_{n \geq 1}$ be a sequence of integer-valued random variables conditionally independent given $p^\downarrow$ with conditional distribution (3) with I'_n in place of I_n. One can also assume that I' and $Z = [Z^+, (Z')^-]$ are independent given $[p^\downarrow, \tilde{H}]$; see Lemma A1 in the Appendix A. Set $\tau_1 = [I', (Z')^-]$ and define

$$(\xi'_n)_{n \geq 1} = \phi(\tau_1, \sigma_2) := (Z_{I'_n} \mathbf{1}\{I'_n \geq 1\} + Z'_{-n} \mathbf{1}\{I'_n = -n\})_{n \geq 1}.$$

Let us show that the law of $\xi' := (\xi'_n)_{n \geq 1}$ given σ_2 is the same as the law of ξ given σ_2. Take n Borel sets $A_1, \ldots, A_n$ and non-zero integer numbers $i_1, \ldots, i_n$. One has

$$\mathbb{P}\left\{ \xi'_1 \in A_1, \ldots, \xi'_n \in A_n, I'_1 = i_1, \ldots, I'_n = i_n \middle| \tilde{H}, p^\downarrow, Z \right\}$$

$$= \prod_{j=1}^{n} \left[\delta_{Z_{i_j}}(A_j) p_{i_j}^\downarrow \mathbf{1}\{i_j > 0\} + \alpha \delta_{Z'_{-j}}(A_j) \mathbf{1}\{i_j = -j\} \right].$$

Conditionally on $\tilde{H}$, the $(Z'_n)_{n \leq -1}$ are i.i.d. with law $\tilde{H}$ so that

$$\mathbb{P}\Big\{\xi_1' \in A_1, \dots, \xi_n' \in A_n, I_1' = i_1, \dots, I_n' = i_n \Big| \tilde{H}, p^\downarrow, Z^+\Big\}$$

$$= \mathbb{E}\Big[\mathbb{P}\Big\{\xi_1' \in A_1, \dots, \xi_n' \in A_n, I_1' = i_1, \dots, I_n' = i_n \Big| \tilde{H}, p^\downarrow, Z\Big\} \Big| \tilde{H}, p^\downarrow, Z^+\Big]$$

$$= \mathbb{E}\Big[\prod_{j=1}^n \Big[\delta_{Z_{i_j}}(A_j) p_{i_j}^\downarrow \mathbf{1}\{i_j > 0\} + \alpha \delta_{Z_{-j}'}(A_j)\mathbf{1}\{i_j = -j\}\Big] \Big| \tilde{H}, p^\downarrow, Z^+\Big]$$

$$= \prod_{j=1}^n \mathbb{E}\Big[\delta_{Z_{i_j}}(A_j) p_{i_j}^\downarrow \mathbf{1}\{i_j > 0\} + \alpha \delta_{Z_{-j}'}(A_j)\mathbf{1}\{i_j = -j\} \Big| \tilde{H}, p^\downarrow, Z^+\Big]$$

$$= \prod_{j=1}^n \Big[\delta_{Z_{i_j}}(A_j) p_{i_j}^\downarrow \mathbf{1}\{i_j > 0\} + \alpha \tilde{H}(A_j)\mathbf{1}\{i_j = -j\}\Big].$$

Marginalizing with respect to $i_1, \dots, i_n$,

$$\mathbb{P}\Big\{\xi_1' \in A_1, \dots, \xi_n' \in A_n \Big| \tilde{H}, p^\downarrow, Z^+\Big\} = \prod_{j=1}^n P(A_j).$$

Recalling that $P = \sum_{j \geq 1} p_j^\downarrow \delta_{Z_j} + \alpha \tilde{H}$,

$$\mathbb{P}\big\{\xi_1' \in A_1, \dots, \xi_n' \in A_n | P\big\} = \prod_{j=1}^n P(A_j)$$

almost surely. Since $\mathbb{X}$ is Polish, we have proven that, given P, $(\xi_n')_{n \geq 1}$ are i.i.d. with common distribution P. In particular, we have proven that $\xi' := (\xi_n')_{n \geq 1}$ given σ_2 is the same as the law of ξ given σ_2. This concludes the proof of the "if part", since, by the previous argument, any sequence of the form (ξ') is of type (mSSS). To complete the proof, it remains to conclude the "only if part". Setting $\sigma_1 = \xi$, we have proven that the conditional distribution of σ_1 given σ_2 is the same as the conditional distribution of $\phi(\tau_1, \sigma_2)$ given σ_2. At this stage, Lemma A3 in the Appendix A yields that there is $\tau = [(I_n)_{n \geq 1}, (Z_n)_{n \leq -1}]$ such that $(\xi_n)_n = \phi(\tau, \sigma_2)$ a.s., i.e., $(\xi_n)_n = (Z_{I_n})_n$ a.s. In addition, $\mathcal{L}(\tau, \sigma_2) = \mathcal{L}(\tau_1, \sigma_2)$; hence, the $(Z_n)_{n \leq -1}$ are conditionally i.i.d. given $\tilde{H}$ and the I_ns are conditionally independent given $[Z^+, Z^-, \tilde{H}, p^\downarrow]$ with the conditional distribution defined by (3). $\square$

Proposition 3. *An exchangeable sequence $\xi = (\xi_n)_{n \geq 1}$ is an mSSS as defined in* Definition 2 *if and only if*

$$(\xi_n)_{n \geq 1} := (Z'_{\mathscr{C}_n(\Pi)})_{n \geq 1} \quad a.s. \tag{16}$$

where $Z' := (Z'_n)_{n \geq 1}$ is an exchangeable sequence with the same law of Z, Π is an exchangeable partition with EPPF $\mathfrak{q}$ and Π and Z' are independent.

Remark 1. *Note that the Z'_ns appearing in* (16) *are not the same Z_ns appearing in* Definition 2, *although they have the same law.*

Proof of Proposition 3. If ξ is mSSS, then, by Proposition 2, we know that $\xi = (Z_{I_n})_{n \geq 1}$. Let $\Pi = \Pi^*(I)$ be the partition induced by $(I_n)_{n \geq 1}$; then, Π has EPPF $\mathfrak{q}$ by Kingman's theorem 1. Denote by $I_1^* = I_1, I_2^*, \dots, I_K^*$ (with $K \leq +\infty$) the distinct values of $(I_n)_{n \geq 1}$ in order of appearance, and set

$$Z'_n = Z_{I_n^*} \quad n = 1, \dots, K.$$

When $K < +\infty$, set $I_{K+j}^* = D + j$, where $D = \max\{i : i = I_n^* \text{ for } n \leq K\}$, and define the remaining Z'_m for $m > K$ accordingly as $Z'_m = Z_{I_m^*}$. Let $\{i_1, \dots, i_M\}$ be integers in $\mathbb{Z} \setminus \{0\}$,

and denote the distinct values in $(i_1, \ldots, i_M)$ in order of appearance by $(i_1^*, \ldots, i_m^*)$. Let $A_1, \ldots, A_n$ be measurable sets in $\mathcal{X}$, if $n > m$, then

$$
\begin{aligned}
& \mathbb{P}\{Z_1' \in A_1, \ldots, Z_n' \in A_n, I_1 = i_1, \ldots, I_M = i_M\} \\
& \quad = \sum_{\ell_1, \ldots, \ell_{n-m}} \mathbb{P}\Big\{Z_{i_1^*} \in A_1, \ldots, Z_{i_m^*} \in A_m, Z_{\ell_1} \in A_{m+1} \ldots, \\
& \qquad Z_{\ell_{n-m}} \in A_n, I_1 = i_1, \ldots, I_M = i_M, I_{m+1}^* = \ell_1, \ldots, I_n^* = \ell_{n-m}\Big\}
\end{aligned}
$$

where the sum runs over all the non-zero distinct integers $\ell_1, \ldots, \ell_{n-m}$ different from $i_1^*, \ldots, i_m^*$. Since I^* is a function of I and I and Z are independent, it follows that

$$
\begin{aligned}
& \mathbb{P}\Big\{Z_{i_1^*} \in A_1, \ldots, Z_{i_m^*} \in A_m, Z_{\ell_1} \in A_{m+1} \ldots, \\
& \qquad Z_{\ell_{n-m}} \in A_n, I_1 = i_1, \ldots, I_M = i_M, I_{m+1}^* = \ell_1, \ldots, I_n^* = \ell_{n-m}\Big\} \\
& = \mathbb{P}\{Z_{i_1^*} \in A_1, \ldots, Z_{i_m^*} \in A_m, Z_{\ell_1} \in A_{m+1}, \ldots, Z_{\ell_{n-m}} \in A_n\} \\
& \qquad \mathbb{P}\{I_1 = i_1, \ldots, I_M = i_M, I_{m+1}^* = \ell_1, \ldots, I_n^* = \ell_{n-m}\} \\
& = \mathbb{P}\{Z_1 \in A_1, \ldots, Z_n \in A_n\}\mathbb{P}\{I_1 = i_1, \ldots, I_M = i_M, I_{m+1}^* = \ell_1, \ldots, I_n^* = \ell_{n-m}\}
\end{aligned}
$$

where the second equality follows by exchangeability. Summing in ℓ, one obtains

$$
\begin{aligned}
& \mathbb{P}\{Z_1' \in A_1, \ldots, Z_n' \in A_n, I_1 = i_1, \ldots, I_M = i_M\} \\
& \qquad = \mathbb{P}\{Z_1 \in A_1, \ldots, Z_n \in A\}\mathbb{P}\{I_1 = i_1, \ldots, I_M = i_M\}.
\end{aligned}
$$

For $m \geq n$, the sum is not needed and the same result follows. This shows that $(Z_n')_n$ is an exchangeable sequence with the same law of Z, and $(Z_n')_n$ and $(I_n)_{n \geq 1}$ are independent. To conclude, note that, with probability one, $I_{\mathscr{C}_n(\Pi)}^* = I_n$, and hence

$$
\xi_n = Z_{I_n} = Z_{I_{\mathscr{C}_n(\Pi)}^*} = Z_{\mathscr{C}_n(\Pi)}'.
$$

Conversely, let us assume that $\xi_n = Z_{\mathscr{C}_n(\Pi)}'$ and let $(p_j^{\downarrow})_{j \geq 1}$ be the weights obtained from Π by (4). Let $I_1, I_2, \ldots$ be the integer-valued random variables appearing in Theorem 1 such that $\Pi = \Pi^*(I)$ a.s. It follows that $\mathscr{C}_n(\Pi) = \mathscr{C}_n(\Pi^*(I))$ and $I_{\mathscr{C}_n(\Pi)}^* = I_n$, where the I_n^* are defined as above for $n \leq K$. Setting

$$
Z_m := \begin{cases} Z_k' & \text{if } I_k^* = m \\ Z_m'' & \text{if } I_k^* \neq m \,\forall\, k, \end{cases}
$$

with $Z_m'', m \in \mathbb{Z}$ conditionally i.i.d. given $\tilde{H}$ with law $\tilde{H}$, independent of everything else. Arguing as above, one can check that the $(Z_m)_{m \in \mathbb{Z}, m \neq 0}$ are exchangeable random variables with the same law of Z' independent of $(I, p^{\downarrow})$. To conclude, note that, in particular,

$$
Z_{I_n} = Z_{I_{\mathscr{C}_n(\Pi)}^*} = Z_{\mathscr{C}_n(\Pi)}' \quad a.s..
$$

The conclusion follows by Proposition 2. $\quad\square$

A simple consequence of the previous proposition is the following.

Corollary 1. *Let $(\xi_n)_{n \geq 1}$ be an mSSS as defined in* Definition 2. *For every $A_1, \ldots, A_n$ Borel set in $\mathbb{X}$,*

$$
\mathbb{P}\{\xi_1 \in A_1, \ldots, \xi_n \in A_n\} = \sum_{\pi_n \in \mathcal{P}_n} \mathfrak{q}(|\pi_{1,n}|, \ldots, |\pi_{k,n}|)\mathbb{E}\Big[\prod_{c=1}^{|\pi_n|} \tilde{H}(\cap_{j \in \pi_{c,n}} A_j)\Big].
$$

4. Random Partitions Induced by *mSSS*

Let $\tilde{\Pi} = \Pi^*(\xi)$ be the random partition induced by an exchangeable sequence ξ defined as in Definition 2, and let $\Pi^{(0)} := \Pi^*(Z')$ be the partition induced by the corresponding exchangeable sequence $(Z'_n)_n$ (see Proposition 3). Finally, let Π be the partition with EPPF q appearing in Proposition 3. As already observed, if Z' is an i.i.d. sequence from a diffuse H, then $\Pi^{(0)} = [(1),(2),(3),\dots]$ a.s. and hence $\Pi^*(\xi) = \Pi$. The same result follows if Z' is exchangeable without ties (see Corollary 2). When $\Pi^{(0)}$ is not the trivial partition, it is clear by construction that different blocks in Π can merge in the final clustering configuration (i.e., $\Pi^*(\xi)$). In other words, two observations can share the same value because either they belong to the same block in the latent partition Π or they are in different blocks but they share the same value (from Z'). This simple observation leads us to write the EPPF of the random partition $\Pi^*(\xi)$ using the EPPF of $\Pi^{(0)}$ and of Π.

4.1. Explicit Expression of the EPPF

If $\tilde{\pi}_n = \{\tilde{\pi}_{1,n}\dots,\tilde{\pi}_{k,n}\}$ is a partition of $[n]$ with $|\tilde{\pi}_{i,n}| = n_i$ $(i = 1,\dots,k)$ and $\boldsymbol{n} = (n_1,\dots,n_k)$, we can easily describe all the partitions π_n more finely than $\tilde{\pi}_n$, which are compatible with $\tilde{\pi}_n$ in the merging process described above. To do this, first of all, note that any block $\tilde{\pi}_{i,n}$ can arise from the union of $1 \le m_i \le n_i$ blocks in the latent partition. Hence, given $\boldsymbol{n} = (n_1,\dots,n_k)$, where $n = \sum_{i=1}^k n_i$, we define the set

$$\mathcal{M}(\boldsymbol{n}) = \left\{ \boldsymbol{m} = (m_1,\dots,m_k) \in \mathbb{N}^k : 1 \le m_i \le n_i \right\}.$$

See Figure 1 for an example. Once a specific configuration $\boldsymbol{m}$ in $\mathcal{M}(\boldsymbol{n})$ is considered, the m_i blocks of the latent partition contributing to the block $\tilde{\pi}_{i,n}$, are characterized by the sufficient statistics $\boldsymbol{\lambda}_i = (\lambda_{i1},\dots,\lambda_{in_i}) \in \mathbb{N}^{n_i}$, where λ_{ij} is the number of blocks of j elements among the m_i blocks above. This leads, for $\boldsymbol{m}$ in $\mathcal{M}(\boldsymbol{n})$, to the definition of

$$\Lambda(\boldsymbol{n},\boldsymbol{m}) := \left\{ \begin{array}{c} \boldsymbol{\lambda} = [\boldsymbol{\lambda}_1,\dots,\boldsymbol{\lambda}_k] \text{ where } \boldsymbol{\lambda}_i = (\lambda_{i1},\dots,\lambda_{in_i}) \in \mathbb{N}^{n_i} : \\ \sum_{j=1}^{n_i} j\lambda_{ij} = n_i, \sum_{j=1}^{n_i} \lambda_{ij} = m_i \ \text{ for } i = 1,\dots,k \end{array} \right\}.$$

In summary, the set of partitions $\tilde{\pi}_n$, which are compatible with $\tilde{\pi}_n$ in the merging process described above, can be written as

$$\mathcal{P}_{\tilde{\pi}_n} := \cup_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n})} \cup_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{n},\boldsymbol{m})} \mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda}) \tag{17}$$

where $\mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda})$ is the set of all the partitions in $\mathcal{P}_n$ with $m_1 + \dots + m_k =: |\boldsymbol{m}|$ blocks such that

- it is possible to determine k subset containing $m_1,\dots,m_k$ of these blocks;
- the union of the blocks in the i-th subset coincides with the i-th block of $\tilde{\pi}_n$ for $i = 1,\dots,k$;
- in the i-th block, there are λ_{ij} blocks with j elements, for $j = 1,\dots,n_i$.

Given the EPPF q, let

$$\bar{\mathsf{q}}(\boldsymbol{\lambda}) := \mathsf{q}(n_{11},\dots,n_{1m_1},n_{21},\dots,n_{km_k}),$$

where $(n_{11},\dots,n_{1m_1},\dots,n_{km_k})$ is any sequence of integer numbers such that $\sum_{c=1}^{m_i} n_{ic} = \sum_j j\lambda_{ij}$ for every i and $\#\{c : n_{ic} = j\} = \lambda_{ij}$ for every i and j. Note that since the value of $\mathsf{q}(n_{11},\dots,n_{1m_1},n_{21},\dots,n_{km_k})$ depends only on the statistics $\boldsymbol{\lambda}$, $\bar{\mathsf{q}}(\boldsymbol{\lambda})$ is well-defined. See, e.g., [26].
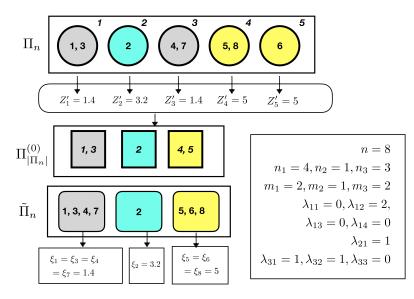
Figure 1. Pictorial representation of the latent partition structure of an mSSS. In the example, the partition induced by $(\xi_1, \ldots, \xi_n)$ for $n = 8$ is $\tilde{\Pi}_n = \{[1, 3, 4, 7], [2], [5, 6, 8]\}$, and it is represented using rounded squares (left bottom). Circles at the top left represent a compatible latent partition, namely $\Pi_n = \{[1, 3], [2], [4, 7], [5, 8], [6]\}$. The partition on $\{1, \ldots, 5\}$ induced by the latent Z'_n, i.e., $\Pi^{(0)}_{|\Pi_n|} = \{[1, 3], [2], [4, 5]\}$, is represented with squares in the middle of the figure. Combining Π_n and $\Pi^{(0)}_{|\Pi_n|}$, one obtains $\tilde{\Pi}_n$. The statistics $\boldsymbol{n}$, $\boldsymbol{m}$ and $\boldsymbol{\lambda}$ corresponding to this particular configuration are shown in the box at the bottom right.

Finally, recall that the cardinality of $\mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda})$ is

$$c(\boldsymbol{\lambda}) := \prod_{i=1}^{k} \frac{(\sum_j j\lambda_{ij})!}{\prod_{j=1}^{n_i} \lambda_{ij}! (j!)^{\lambda_{ij}}} = \prod_{i=1}^{k} \frac{n_i!}{\prod_{j=1}^{n_i} \lambda_{ij}! (j!)^{\lambda_{ij}}},$$

See Equation (39) in [15].

Proposition 4. *Let $\xi = (\xi_n)_{n \geq 1}$ be an (mSSS). Denote by $\tilde{\Pi} = \Pi^*(\xi)$ the random partition induced by ξ and by $\mathfrak{q}^{(0)}$ the EPPF of the partition induced by $(Z'_n)_{n \geq 1}$. If $\tilde{\pi}_n = \{\tilde{\pi}_{1,n} \ldots, \tilde{\pi}_{k,n}\}$ is a partition of $[n]$ with $|\tilde{\pi}_{i,n}| = n_i$ $(i = 1, \ldots, k)$ and $\boldsymbol{n} = (n_1, \ldots, n_k)$, then*

$$\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n\} = \sum_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n})} \mathfrak{q}^{(0)}(\boldsymbol{m}) \sum_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{n}, \boldsymbol{m})} c(\boldsymbol{\lambda}) \bar{\mathfrak{q}}(\boldsymbol{\lambda}). \tag{18}$$

Proof. Start by writing

$$\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n\} = \mathbb{P}(\cup_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n})} \cup_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{n}, \boldsymbol{m})} \cup_{\pi_n \in \mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda})} \{\Pi_n = \pi_n, \tilde{\Pi}_n = \tilde{\pi}_n\}), \tag{19}$$

which gives

$$\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n\} = \sum_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n})} \sum_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{n}, \boldsymbol{m})} \sum_{\pi_n \in \mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda})} \mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n | \Pi_n = \pi_n\} P\{\Pi_n = \pi_n\} \tag{20}$$

Whenever $\pi_n \in \mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda})$,

$$\mathbb{P}\{\Pi_n = \pi_n\} = \bar{\mathfrak{q}}(\boldsymbol{\lambda}).$$

Therefore, we can write (20) as

$$\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n\} = \sum_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n})} \sum_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{n}, \boldsymbol{m})} \sum_{\pi_n \in \mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda})} P\{\tilde{\Pi}_n = \tilde{\pi}_n | \Pi_n = \pi_n\} \bar{\mathfrak{q}}(\boldsymbol{\lambda}). \tag{21}$$

Define now the function $M_{\tilde{\pi}_n, \pi_n} : \{1, \ldots, |\boldsymbol{m}|\} \to \{1, \ldots, |\tilde{\pi}_n|\}$ as $M(j) = i$ if $\pi_{j,n}$ is in the i-th subset of blocks, i.e., if $\pi_{j,n} \subset \tilde{\pi}_{i,n}$. Recalling that k is the number of blocks in $\tilde{\pi}_n$, define now a partition $\pi(M_{\tilde{\pi}_n, \pi_n})$ on $\{1, \ldots, |\boldsymbol{m}|\}$ with k block where the i-th block is

$$\{j : M_{\tilde{\pi}_n, \pi_n}(j) = i\}.$$

Recalling that $\Pi^{(0)}$ is the partition induced by the Z's, one has

$$\{\tilde{\Pi}_n = \tilde{\pi}_n, \Pi_n = \pi_n\} = \{\Pi^{(0)}_{|\boldsymbol{m}|} = \pi(M_{\tilde{\pi}_n, \pi_n}), \Pi_n = \pi_n\}$$

which gives

$$P\{\tilde{\Pi}_n = \tilde{\pi}_n | \Pi_n = \pi_n\} = P\{\Pi^{(0)}_{|\boldsymbol{m}|} = \pi(M_{\tilde{\Pi}_n, \pi_n}) | \Pi_n = \pi_n\}$$
$$= \mathbb{P}\{\Pi^{(0)}_{|\boldsymbol{m}|} = \pi(M_{\tilde{\Pi}_n, \pi_n})\}$$

since $\Pi^{(0)}$ and Π are independent. To conclude, note that the vector of the cardinalities of the blocks in $\pi(M_{\tilde{\pi}_n, \pi_n})$ is $\boldsymbol{m}$; hence, if $\mathfrak{q}^{(0)}$ is the EPPF of $\Pi^{(0)}$, one has $P\{\Pi^{(0)}_{|\boldsymbol{m}|} = \pi(M_{\tilde{\pi}_n, \pi_n})\} = \mathfrak{q}^{(0)}(\boldsymbol{m})$. Since the cardinality of $\mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda})$ is $c(\boldsymbol{\lambda})$, one obtains the thesis. $\square$

Corollary 2. *Let $\xi = (\xi_n)_n$ be defined according to (mSSS). If $\mathbb{P}\{Z'_1 = Z'_2\} = 0$, then $\Pi^*(\xi) = \Pi$ with probability one.*

Proof. If $\mathbb{P}\{Z'_1 = Z'_2\} = 0$, by exchangeability, $\mathbb{P}\{Z'_{i_1} = Z'_{i_2} = \cdots = Z'_{i_k}\} \leq \mathbb{P}\{Z'_1 = Z'_2\} = 0$. Hence, the Z'_is are distinct with probability one. Since $(\xi_1, \ldots, \xi_n) = (Z'_{\mathscr{C}_1(\Pi)}, \ldots, Z'_{\mathscr{C}_n(\Pi)})$ by Proposition 3, it follows that $\tilde{\Pi}_n = \Pi_n$. $\square$

Remark 2. *Note that, as a special case, we recover the fact that if ξ is a gSSS$(\mathfrak{q}, H)$ with H diffuse (i.e., it is a SSS$(\mathfrak{q}, H)$), then the random partition induced by ξ is a.s. Π.*

Remark 3. *The fact that the EPPF of $\tilde{\Pi}_n$ is $\mathfrak{q}$ when $\mathbb{P}\{Z'_1 = Z'_2\} = 0$ can be deduced from (18). Indeed, if $\mathbb{P}\{Z'_1 = Z'_2\} = 0$, then the partition induced by Z' is a.s. the trivial partition $[(1), (2), (3), \ldots]$, so that $\mathfrak{q}^{(0)}(\boldsymbol{m}) = 0$ for every $\boldsymbol{m} \neq (1, 1, \ldots, 1)$. For $\boldsymbol{m} = (1, 1, \ldots, 1)$, $\tilde{\pi}_n = \{\tilde{\pi}_{1,n} \ldots, \tilde{\pi}_{k,n}\}$ with $|\tilde{\pi}_{i,n}| = n_i$ $(i = 1, \ldots, k)$, and $\boldsymbol{n} = (n_1, \ldots, n_k)$, the set $\Lambda(\boldsymbol{n}, \boldsymbol{m})$ reduces to the singleton $\boldsymbol{\lambda}^{(1)} := [\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_k]$, where $\boldsymbol{\lambda}_i = (0, 0, \ldots, 1)$ with $\boldsymbol{\lambda}_i$ of length n_i. Hence, $\bar{\mathfrak{q}}(\boldsymbol{\lambda}^{(1)}) = \mathfrak{q}(\boldsymbol{n})$ and (18) gives $\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n\} = \bar{\mathfrak{q}}(\boldsymbol{\lambda}^{(1)}) = \mathfrak{q}(\boldsymbol{n})$.*

4.2. EPPF When Π Is of Gibbs Type

An important class of exchangeable random partitions is that of Gibbs-type partitions, introduced in [32] and characterized by the EPPF

$$\mathfrak{q}(n_1, \ldots, n_k) := V_{n,k} \prod_{j=1}^{k} (1 - \sigma)_{n_j - 1}, \tag{22}$$

where $(x)_n = x(x+1) \ldots (x+n-1)$, $\sigma < 1$ and $V_{n,k}$ are positive real numbers such that $V_{1,1} = 1$ and

$$(n - \sigma k)V_{n+1,k} + V_{n+1,k+1} = V_{n,k}, \quad n \geq 1, \ 1 \leq k \leq n.$$

A noteworthy example of Gibbs-type EPPF is the so-called Pitman–Yor two-parameter family. It is defined by

$$\mathfrak{q}(n_1, \ldots, n_k) := \frac{\prod_{i=1}^{k-1}(\theta + i\sigma)}{(\theta + 1)_{n-1}} \prod_{c=1}^{k} (1 - \sigma)_{n_c - 1}, \tag{23}$$

where $0 \leq \sigma < 1$ and $\theta > -\sigma$; or $\sigma < 0$ and $\theta = |\sigma|m$ for some integer m; see [2,31].

In order to state the next result, we recall that

$$\sum_{\substack{(\lambda_1,\dots,\lambda_n) \\ \sum_{j=1}^n j\lambda_j=n, \sum_{j=1}^n \lambda_j=k}} \prod_{j=1}^n [(1-\sigma)_{j-1}]^{\lambda_j} \frac{n!}{\prod_{j=1}^n \lambda_i!(j!)^{\lambda_j}} = S_\sigma(n,k) \tag{24}$$

where $S_\sigma(n,k)$ is the generalized Stirling number of the first kind; see (3.12) in [26]. In the same book, various equivalent definitions of generalized Stirling numbers are presented.

Corollary 3. *Let $\tilde{\Pi} = \Pi^*(\xi)$ be defined as in* Proposition 4 *with* q *of Gibbs type defined in* (22). *If $\tilde{\pi}_n = \{\tilde{\pi}_{1,n}\dots,\tilde{\pi}_{k,n}\}$ is a partition of $[n]$ with $|\tilde{\pi}_{i,n}| = n_i$ ($i=1,\dots,k$) and $\boldsymbol{n} = (n_1,\dots,n_k)$, then*

$$\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n\} = \sum_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n})} \mathsf{q}^{(0)}(\boldsymbol{m}) V_{n,|\boldsymbol{m}|} \prod_{i=1}^k S_\sigma(n_i, m_i).$$

Proof. Combining Proposition 4 with (22), one obtains

$$\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n\} = \sum_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n})} \mathsf{q}^{(0)}(\boldsymbol{m}) V_{n,|\boldsymbol{m}|} \sum_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{n},\boldsymbol{m})} \prod_{i=1}^k \prod_{j=1}^{n_i} [(1-\sigma)_{j-1}]^{\lambda_{i,j}} \frac{n_i!}{\prod_{j=1}^{n_i} \lambda_{ij}!(j!)^{\lambda_{ij}}}$$

$$= \sum_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n})} \mathsf{q}^{(0)}(\boldsymbol{m}) V_{n,|\boldsymbol{m}|}$$

$$\times \prod_{i=1}^k \sum_{\substack{(\lambda_{i1},\dots,\lambda_{in_i}) \\ \sum_{j=1}^{n_i} j\lambda_{ij}=n_i, \sum_{j=1}^{n_i} \lambda_{ij}=m_i}} \prod_{j=1}^{n_i} [(1-\sigma)_{j-1}]^{\lambda_{i,j}} \frac{n_i!}{\prod_{j=1}^{n_i} \lambda_{ij}!(j!)^{\lambda_{ij}}}$$

$$= \sum_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n})} \mathsf{q}^{(0)}(\boldsymbol{m}) V_{n,|\boldsymbol{m}|} \prod_{i=1}^k S_\sigma(n_i, m_i).$$

□

4.3. The EPPF of a $gSSS(\mathsf{q}, H)$

As a special case, we now consider the partition induced by a $gSSS(\mathsf{q}, H)$ with general base measure H. For the rest of the section, it is useful to decompose H as

$$H(dx) = \sum_{i \geq 1} a_i \delta_{\bar{x}_i}(dx) + (1-a)H^c(dx) \tag{25}$$

where $\mathbb{X}_0 := \{\bar{x}_1, \bar{x}_2, \dots\}$ is the collection of points with positive H probability, $a_i = H(\bar{x}_i)$, $a = H(\mathbb{X}_0) \in [0,1]$ and $H^c(\cdot) = H(\cdot \cap \mathbb{X}_0^c)/H(\mathbb{X}_0^c)$ is a diffuse probability measure on $\mathbb{X}$. The sum is assumed taken over $i \in \{1, \dots, |\mathbb{X}_0|\}$.

We now describe $\mathsf{q}^{(0)}$, i.e., the EPPF of the partition induced by $(Z'_n)_{n \geq 1}$. Let $\boldsymbol{m}$ in $\mathcal{M}(\boldsymbol{n})$, where $\boldsymbol{n} = (n_1, \dots, n_k)$, and assume that the realization of $\Pi_{|\boldsymbol{m}|}^{(0)}$ has k blocks of cardinality $m_1, \dots, m_k$. Set $z_i = 0$ if the Z'_n corresponding to the i-th block of $\Pi_{|\boldsymbol{m}|}^{(0)}$ comes from the diffuse component H^c, while $z_k = \ell$ if it is equal to $\bar{x}_\ell$. Since the blocks in $\Pi^{(0)}$ need to be associated with different values of the Z'_n, one has that necessarily $z_i = z_j = 0$ if $z_i = z_j$ for $i \neq j$. In this case, the block is a singleton, which is $m_i = m_j = 1$. On the other hand, if $m_i \geq 2$, i.e., a merging occurred, necessarily, $z_i > 0$. Note that it is also possible that $m_i = 1$ but $z_i > 0$. This motivates the definition of the set

$$\mathcal{Z}(\boldsymbol{m}) := \Big\{ (z_1, \dots, z_k) \in \{0, 1, \dots, |\mathbb{X}_0|\}^k : \text{if } z_i = z_j \text{ for } i \neq j \text{ then } z_i = z_j = 0$$

$$\text{and } m_i = m_j = 1; \text{ if } m_i \geq 2 \text{ then } z_i > 0 \Big\}$$

for m in $\mathcal{M}(n)$ where $n = (n_1, \ldots, n_k)$. The probability of obtaining, in an i.i.d. sample of length $|m|$ from H, exactly k ordered blocks with cardinality $m_1, \ldots, m_k$, such that observations in each block are equal and observations in distinct blocks are different, is

$$H^{\#}(m) := \sum_{(z_1, \ldots, z_k) \in \mathcal{Z}(m)} (1-a)^{\#\{j : z_j = 0\}} \prod_{j : z_j > 0} a_{z_j}^{m_j}.$$

By exchangeability, $H^{\#}(m)$ turns out to be $\mathfrak{q}^{(0)}(m)$. Note also that if $a = 1$, $H^{\#}(m)$ reduces to

$$\sum_{(z_1, \ldots, z_k)} \prod_{j=1}^{k} a_{z_j}^{m_j}$$

where $z_1, \ldots, z_k$ runs over all distinct positive integers (less than or equal to $|\mathbb{X}_0|$ if $\mathbb{X}_0$ is finite), which is nothing else (5) for deterministic weights.

To rewrite $H^{\#}(m)$ in a different way, given $m = (m_1, \ldots, m_k)$ in $\mathcal{M}(n)$, let m^* be the vector containing all the elements $m_i > 1$ and let r be its length, with possibly $r = 0$ if $m = (1, 1, \ldots, 1)$, and define for $\ell \geq 0$

$$A_{m,\ell} = \sum_{j_1 \neq \cdots \neq j_{r+\ell}} a_{j_1}^{m_1^*} \ldots a_{j_r}^{m_r^*} a_{j_{r+1}} \ldots a_{j_{r+\ell}}$$

with the convention that $A_{m,0} = 1$ when $r = 0$. A simple combinatorial argument shows that

$$H^{\#}(m) = \sum_{\ell=0}^{k-r} (1-a)^{k-\ell-r} \binom{k-r}{\ell} A_{m,\ell}.$$

Proposition 4 gives immediately the next proposition.

Proposition 5. *Let ξ be a $gSSS(\mathfrak{q}, H)$ and let $\tilde{\Pi} = \Pi^*(\xi)$ be the random partition induced by ξ. If $\tilde{\pi}_n = [\tilde{\pi}_{1,n} \ldots, \tilde{\pi}_{k,n}]$ is a partition of $[n]$ with $|\tilde{\pi}_{i,n}| = n_i$ $(i = 1, \ldots, k)$ and $n = (n_1, \ldots, n_k)$, then*

$$\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n\} = \sum_{m \in \mathcal{M}(n)} H^{\#}(m) \sum_{\lambda \in \Lambda(n,m)} c(\lambda) \bar{\mathfrak{q}}(\lambda).$$

Remark 4. *Once again, if H is diffuse, then $H^{\#}(m) = 0$ for every $m \neq (1, 1, \ldots, 1)$. Hence, the above formula reduces to the familiar*

$$\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n\} = \mathfrak{q}(|\tilde{\pi}_{n,1}|, \ldots, |\tilde{\pi}_{n,k}|) = \mathbb{P}\{\Pi_n = \tilde{\pi}_n\}.$$

4.4. EPPF for gSSS with Spike-and-Slab Base Measure

We now consider the special case of *gSSS* with a spike-and-slab base measure. A spike-and-slab measure is defined as

$$H(dx) = a\delta_{x_0}(dx) + (1-a)H^c(dx) \tag{26}$$

where $a \in (0, 1)$, x_0 is a point of $\mathbb{X}$ and H^c is a diffuse measure on $\mathbb{X}$. This type of measure has been used as a base measure in the Dirichlet process by [16–20] and in the Pitman–Yor process by [21].

Here, we deduce by Proposition 5 the explicit form of the EPPF of the random partition induced by a sequence sampled from a species sampling random probability with such a base measure.

Proposition 6. *Let H be as in* (26), *$\tilde{\Pi}$ be the random partition induced by a $gSSS(\mathfrak{q}, H)$ and Π be an exchangeable random partition with EPPF $\mathfrak{q}$. If $\pi_n = \{\pi_{1,n} \ldots, \pi_{k,n}\}$ is a partition of $[n]$ with $|\pi_{i,n}| = n_i$ $(i = 1, \ldots, k)$, then*

$$\mathbb{P}\{\tilde{\Pi}_n = \pi_n\} = (1-a)^k \mathfrak{q}(n_1, \ldots, n_k)$$
$$+ (1-a)^{k-1} \sum_{i=1}^{k} \mathfrak{q}(n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_k) \sum_{r=1}^{n_i} a^r q_n(r|n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_k) \tag{27}$$

where, conditionally on the fact that Π_{n-n_i} has $k-1$ blocks with sizes $n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_k$, the probability that Π_n has $k-1+r$ blocks is denoted by $q_n(r|n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_k)$. If, in addition, $\mathfrak{q}$ is of Gibbs type (22), *then*

$$\mathbb{P}\{\tilde{\Pi}_n = \pi_n\} = (1-a)^k V_{n,k} \prod_{j=1}^{k} (1-\sigma)_{n_j-1}$$
$$+ (1-a)^{k-1} \sum_{i=1}^{k} \prod_{j=1, j\neq i}^{k} (1-\sigma)_{n_j-1} \sum_{r=1}^{n_i} a^r V_{n,k-1+r} S_\sigma(n_i, r).$$

Proof. In this case, $H^{\#}(\boldsymbol{m}) = 0$ if $m_i \geq 2$ and $m_j \geq 2$ for some $i \neq j$ because H has only one atom. Moreover, $H^{\#}(\boldsymbol{m})$ is clearly symmetric and

$$H^{\#}(1, 1, 1, \ldots, 1) = (1-a)^k + k(1-a)^{k-1} a$$

$$H^{\#}(m, 1, \ldots, 1) = a^m (1-a)^{k-1} \qquad \text{for } m > 1.$$

By Proposition 5,

$$\mathbb{P}\{\tilde{\Pi}_n = \pi_n\} = [(1-a)^k + k(1-a)^{k-1} a]\mathfrak{q}(n_1, \ldots, n_k)$$
$$+ (1-a)^{k-1} \sum_{i=1}^{k} \sum_{m_i=2}^{n_i} a^{m_i} \sum_{\lambda \in \Lambda(\boldsymbol{m})} c(\boldsymbol{\lambda}) \bar{\mathfrak{q}}(\boldsymbol{\lambda})$$
$$= [(1-a)^k + k(1-a)^{k-1}]\mathfrak{q}(n_1, \ldots, n_k) +$$
$$+ (1-a)^{k-1} \sum_{i=1}^{k} \sum_{r=2}^{n_i} a^r \sum_{\substack{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{m}) \text{ for } \boldsymbol{m}: \\ m_i=r, \, m_j=1, j\neq i}} c(\boldsymbol{\lambda}) \mathfrak{q}(n_1, \ldots, n_{i-1}, \mathbf{n}_r^{(i)}, n_{i+1}, \ldots, n_k)$$
$$= (1-a)^k \mathfrak{q}(n_1, \ldots, n_k)$$
$$+ (1-a)^{k-1} \sum_{i=1}^{k} \sum_{r=1}^{n_i} a^r \sum_{\substack{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{m}) \text{ for } \boldsymbol{m}: \\ m_i=r, \, m_j=1, j\neq i}} c(\boldsymbol{\lambda}) \mathfrak{q}(n_1, \ldots, n_{i-1}, \mathbf{n}_r^{(i)}, n_{i+1}, \ldots, n_k)$$

where $\mathbf{n}_r^{(i)}$ is any vector of r positive integers with sum n_i such that λ_{ij} of them are equal to j. In view of the definition of $c(\boldsymbol{\lambda})$, formula (27) is immediately obtained.

If $\mathfrak{q}$ is of Gibbs type, taking into account (24), then

$$q_n(r|n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_k) = \frac{V_{n,k-1+r}}{V_{n-n_i,k-1}} S_\sigma(n_i, r)$$

and the second part of the thesis follows by simple algebra. $\quad\square$

Applying Proposition 6 to the Pitman–Yor EPPF defined in (23), one immediately recovers the results stated in Theorem 1 and Corollary 1 of [21].

5. Predictive Distributions

In this section, we provide some expressions for the predictive distributions of mixtures of species sampling sequences.

5.1. Some General Results

Let ξ be as in Definition 2 and let $(Z'_n)_n$ and Π_n be the sequence of exchangeable random variables and the exchangeable random partition appearing in Proposition 3. Let

$$\mathcal{G}_n = \sigma(Z'_1, \ldots, Z'_{|\Pi_n|}, \Pi_n)$$

be the σ-field generated by $(Z'_1, \ldots, Z'_{|\Pi_n|}, \Pi_n)$. By Proposition 3, one has $\xi_n = Z'_{\mathscr{C}_n(\Pi)}$ a.s.; hence, ξ_n is $\mathcal{G}_n$ measurable. Note that, in general, $\sigma(\xi_1, \ldots, \xi_n)$ can be strictly contained in $\mathcal{G}_n$. Set $\Xi_n := |\Pi_n|$ and, for any $k \geq 1$, let $\mathcal{K}_{k+1}(\cdot|\cdot)$ be a kernel corresponding to the conditional distribution of Z'_{k+1} given $Z'_1, \ldots, Z'_k$ (i.e., the $k+1$-predictive distribution of the exchangeable sequence Z'). Finally, recall that $\tilde{\Pi} = \Pi^*(\xi)$ is the partition induced by ξ and define $\xi^*_{1:\tilde{K}_n} = (\xi^*_1, \ldots, \xi^*_{\tilde{K}_n})$ as the distinct values in order of appearance of $\xi_{1:n} := (\xi_1, \ldots, \xi_n)$ with $\tilde{K}_n = |\tilde{\Pi}_n|$.

Proposition 7. *Let ξ as in Definition 2. Then,*

$$\mathbb{P}\{\xi_{n+1} \in \cdot|\mathcal{G}_n\} = \sum_{\ell=1}^{\Xi_n} \omega_{n,\ell}(\Pi_n)\delta_{Z'_\ell}(\cdot) + \nu_n(\Pi_n)\mathcal{K}_{\Xi_n+1}(\cdot|Z'_1, \ldots, Z'_{\Xi_n}) \tag{28}$$

where ν_n and $\omega_{n,\ell}$ are defined by (6) and (7). If $P\{Z'_1 = Z'_2\} = 0$, then

$$\begin{aligned}
\mathbb{P}\{\xi_{n+1} \in \cdot|\xi_1, \ldots, \xi_n\} &= \mathbb{P}\{\xi_{n+1} \in \cdot|\xi^*_1, \ldots, \xi^*_{\tilde{K}_n}, \tilde{\Pi}_n\} \\
&= \sum_{\ell=1}^{\tilde{K}_n} \omega_{n,\ell}(\tilde{\Pi}_n)\delta_{\xi^*_\ell}(\cdot) + \nu_n(\tilde{\Pi}_n)\mathcal{K}_{\tilde{K}_n+1}(\cdot|\xi^*_1, \ldots, \xi^*_{\tilde{K}_n}).
\end{aligned} \tag{29}$$

Proof. Set

$$E^*_{new} = \{\Xi_{n+1} = \Xi_n + 1\}.$$

Since $\xi_n = Z'_{\mathscr{C}_n(\Pi)}$, one can write

$$\begin{aligned}
\mathbb{P}\{\xi_{n+1} \in A|\mathcal{G}_n\} &= \sum_{\ell=1}^{\Xi_n} \mathbb{P}\{\xi_{n+1} \in A, n+1 \in \Pi_{\ell,n}|\mathcal{G}_n\} + \mathbb{P}\{\xi_{n+1} \in A, E^*_{new}|\mathcal{G}_n\} \\
&= \sum_{\ell=1}^{\Xi_n} \mathbb{P}\{Z'_\ell \in A, n+1 \in \Pi_{\ell,n}|\mathcal{G}_n\} + \mathbb{P}\{Z'_{\Xi_{n+1}} \in A, E^*_{new}|\mathcal{G}_n\} \\
&= \sum_{\ell=1}^{\Xi_n} \delta_{Z'_\ell}(A)\mathbb{P}\{n+1 \in \Pi_{\ell,n}|\mathcal{G}_n\} + \mathbb{P}\{Z'_{\Xi_{n+1}} \in A, E^*_{new}|\mathcal{G}_n\}
\end{aligned}$$

Now, since Π and $(Z'_n)_n$ are independent, it follows that $\mathbb{P}\{n+1 \in \Pi_{\ell,n}|\mathcal{G}_n\} = \mathbb{P}\{n+1 \in \Pi_{\ell,n}|\Pi_n\} = \omega_{n,\ell}(\Pi_n)$ and also

$$\begin{aligned}
\mathbb{P}\{Z'_{\Xi_{n+1}} &\in A, E^*_{new}|\mathcal{G}_n\} \\
&= \mathbb{P}\{Z'_{\Xi_{n+1}} \in A|Z'_1, \ldots, Z'_{\Xi_n}\}\mathbb{P}\{E^*_{new}|\Pi_n\} \\
&= \mathcal{K}_{\Xi_n+1}(A|Z'_1, \ldots, Z'_{\Xi_n})\nu_n(\Pi_n).
\end{aligned}$$

Combining all the claims, one obtains (28). The second part of the proof follows since, if $P\{Z'_1 = Z'_2\} = 0$, the Z'_is are distinct with probability one. Since $(\xi_1, \ldots, \xi_n) = (Z'_{\mathscr{C}_1(\Pi)}, \ldots, Z'_{\mathscr{C}_n(\Pi)})$, it follows that $\tilde{\Pi}_n = \Pi_n$, $\Xi_n = \tilde{K}_n$ and $(\xi^*_1, \ldots, \xi^*_{\tilde{K}_n}) = (Z'_1, \ldots, Z'_{\Xi_n})$ with probability one and $\mathcal{G}_n = \sigma(\xi_1, \ldots, \xi_n)$. Hence, (29) follows from (28). □

Remark 5. *Note that* (29) *can be also derived as follows.* $\mathbb{P}\{Z_1' = Z_2'\} = 0$ *is equivalent to the fact that $\tilde{H}$ is almost sure diffuse. Hence, conditionally on $\tilde{H}$, we have a $SSS(\mathfrak{q}, \tilde{H})$; then, by (PS2) in Section* 2.2, *one has*

$$\mathbb{P}\{\xi_{n+1} \in \cdot | \xi_1, \ldots, \xi_n, \tilde{H}\} = \sum_{\ell=1}^{\tilde{K}_n} \omega_{n,\ell}(\tilde{\Pi}_n)\delta_{\xi_\ell^*}(\cdot) + \nu_n(\tilde{\Pi}_n)\tilde{H}(dx). \tag{30}$$

Taking the conditional expectation of the previous equation, given $\xi_1, \ldots, \xi_n$, we obtain

$$\mathbb{P}\{\xi_{n+1} \in A | \xi_1, \ldots, \xi_n\} = \sum_{\ell=1}^{\tilde{K}_n} \omega_{n,\ell}(\tilde{\Pi}_n)\delta_{\xi_\ell^*}(A) + \nu_n(\tilde{\Pi}_n)\mathbb{E}[\tilde{H}(A)|\xi_1, \ldots, \xi_n] \tag{31}$$

and the thesis follows since one can check (arguing as in the proof of the proposition) that

$$\mathbb{E}[\tilde{H}(A)|\xi_1, \ldots, \xi_n] = \mathbb{E}[\tilde{H}(A)|Z_1', \ldots, Z_{\tilde{K}_n}'] = \mathcal{K}_{\tilde{K}_n+1}(A|Z_1', \ldots, Z_{\tilde{K}_n}').$$

Assume now that the random variables Z_j' are defined on $\mathbb{X}$ by a Bayesian model with likelihood $f(z_j|u)$ and prior $Q(u)$, where f is a density with respect to a dominating measure λ and Q is a probability measure defined on a Polish space U (the space of parameters). In other words,

$$\mathbb{P}\{Z_1' \in A_1, \ldots, Z_k' \in A_k\} = \int_U \left(\int_{A_1 \times A_2 \cdots \times A_k} \prod_{j=1}^k f(z_j|u)\lambda(dz_1) \ldots \lambda(dz_k) \right) Q(du).$$

Note that this means that $\tilde{H}(A) = \int_A f(z|\tilde{u})\lambda(dz)$, where $\tilde{u} \sim Q$. Bayes' theorem (see, e.g., Theorem 1.31 in [33]) gives

$$\mathbb{P}\{Z_{k+1}' \in dz_{k+1}|Z_1', \ldots, Z_k'\} = \left(\int_U f(z_{k+1}|u)Q(du|Z_1', \ldots, Z_k') \right) \lambda(dz_{k+1})$$

where $Q(du|Z_1, \ldots, Z_k)$ is the usual posterior distribution, which is

$$Q(du|Z_1', \ldots, Z_k') := \frac{\prod_{j=1}^k f(Z_j'|u)Q(du)}{\int_U \prod_{j=1}^k f(Z_j'|v)Q(dv)}.$$

If λ is a diffuse measure, one obtains $\mathbb{P}\{Z_1' = Z_2'\} = 0$. Hence, (29) in Proposition 7 applies and one has

$$\mathbb{P}\{\xi_{n+1} \in dx|\xi_1, \ldots, \xi_n\} = \sum_{\ell=1}^{\tilde{K}_n} \omega_{n,\ell}(\tilde{\Pi}_n)\delta_{\xi_\ell^*}(dx) + \nu_n(\tilde{\Pi}_n)\left(\int_U f(x|u)Q(du|\xi_1^*, \ldots, \xi_{\tilde{K}_n}^*) \right) dx. \tag{32}$$

For example, one can apply this result to a mixture of Dirichlet processes in the sense of [10], as briefly described at the end of Section 3.1. Assume that $\alpha_{\tilde{u}}(\mathbb{X})$ and $\tilde{H}(\cdot) = \alpha_{\tilde{u}}(\cdot)/\alpha_{\tilde{u}}(\mathbb{X})$ are independent and that $\alpha_u(A)/\alpha_u(\mathbb{X}) = \int_Z f(z|u)\lambda(dz)$ for a suitable dominating diffuse measure λ.

Under these hypotheses, a sample $(\xi_n)_{n\geq 1}$ from a mixture of Dirichlet processes is an mSSS with $\mathfrak{q}$ described in (15) and, in addition, $\mathbb{P}\{Z_1' = Z_2'\} = 0$. Combining (15) with (6) and (7), one obtains

$$\omega_{n,\ell}(\tilde{\Pi}_n) = |\tilde{\Pi}_{n,\ell}| \frac{\int_U \frac{\alpha_u(\mathbb{X})^{\tilde{K}_n}}{(\alpha_u(\mathbb{X}))_{n+1}} Q(du)}{\int_U \frac{\alpha_u(\mathbb{X})^{\tilde{K}_n}}{(\alpha_u(\mathbb{X}))_n} Q(du)}$$

and

$$\nu_n(\tilde{\Pi}_n) = \frac{\int_U \frac{\alpha_u(\mathbb{X})^{\tilde{K}_n+1}}{(\alpha_u(\mathbb{X}))_{n+1}} Q(du)}{\int_U \frac{\alpha_u(\mathbb{X})^{\tilde{K}_n}}{(\alpha_u(\mathbb{X}))_n} Q(du)}$$

Hence, the predictive distribution of ξ_{n+1} given $(\xi_1, \ldots, \xi_n)$ is (33) for $\omega_{n,\ell}(\tilde{\Pi}_n)$ and $\nu_n(\tilde{\Pi}_n)$ given above.

Note that the same result can be deduced by combining Lemma 1 and Corollary 3.2' in [10].

Example 1 (Species Sampling NIG). *Let Z'_n be defined as a mixture of normal random variables with Normal-Inverse-Gamma prior. In other words, given $\mu_0 \in \mathbb{R}, k_0 > 0, \alpha_0 > 0, \beta_0 > 0$,*

$$Z_n|\tilde{\mu}, \tilde{\sigma}^2 \overset{i.i.d.}{\sim} \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$$
$$\tilde{\mu}|\tilde{\sigma}^2 \sim \mathcal{N}(\mu_0, \tilde{\sigma}^2/k_0)$$
$$\tilde{\sigma}^2 \sim \mathrm{In}\Gamma(\alpha_0, \beta_0)$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes a normal distribution of mean μ and variance σ^2 and $\mathrm{In}\Gamma(\alpha, \beta)$ is the inverse gamma distribution with shape α and scale β. Let $\mathcal{T}_\nu(\cdot|\mu, \sigma^2)$ be the density of a Student-T distribution with ν degrees of freedom and (μ, σ) position/scale parameters, i.e.,

$$\mathcal{T}_\nu(x|\mu, \sigma^2) := \frac{1}{\sqrt{\sigma^2}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{1}{\nu\sigma^2}(x-\mu)^2\right)^{\frac{-\nu+1}{2}}.$$

It is well known that, under these assumptions, $\mathcal{K}_{k+1}(A|z_1, \ldots, z_k)$ has density $\mathcal{T}_{2\alpha_k}(z|\mu_k, \sigma_k^2)$, where the parameters are updated

$$\mu_k = \frac{k_0\mu_0 + k\bar{z}_n}{k_0 + n} \qquad \bar{z}_k = \frac{1}{k}\sum_{j=1}^k z_j \quad \alpha_k = \alpha_0 + k/2,$$

$$\sigma_k^2 = \frac{\left(\beta_0 + \frac{1}{2}\sum_{j=1}^k (z_j - \bar{z}_k)^2 + \frac{kk_0(\bar{z}_k-\mu_0)^2}{2(n+k_0)}\right)(k_0 + k + 1)}{(\alpha_0 + k/2)(k_0 + k)}$$

Thus, in this case, if $z_1, \ldots, z_k$ are distinct real numbers and $\pi_n = [\pi_{1,n}, \ldots, \pi_{k,n}]$, one has

$$\mathbb{P}\{\xi_{n+1} \in dx | \xi_1^* = z_1, \ldots, \xi_k^* = z_k, \Pi^*(\xi) = \pi_n\} = \sum_{\ell=1}^k \omega_{n,\ell}(\pi_n)\delta_{\xi_\ell^*}(dx) \qquad (33)$$
$$+ \nu_n(\pi_n)\mathcal{T}_{2\alpha_k}(x|\mu_n, \sigma_k^2)dx.$$

We show an application of (33) to a true dataset by choosing $\omega_{n,\ell}$ and ν_n according to a Pitman–Yor two-parameter family; see (23). The data are the relative changes in reported larcenies between 1991 and 1995 (relative to 1991) for the 90 most populous US counties, taken from Section 2.1 of [34]. We apply our models to both the raw data and the rounded data (approximated to the second digit) in order to obtain ties in the ξs. In the evaluation of the predictive CDFs, we fix $\mu_0 = 0, \alpha_0 = 0.1, \beta_0 = 0.1$ and $k_0 = 0.1$. In Figure 2, we report the empirical CDF of the rounded data (solid line), the predictive CDF obtained from (33) (dotted line) and the predictive CDF of a Pitman–Yor species sampling sequence (see PS2) with $H = \mathcal{T}_{2\alpha_0}(\cdot|\mu_0, \sigma_0^2), \sigma_0^2 = \beta_0(k_0 + 1)/\alpha_0 k_0$ (dashed line). Similar plots are reported in Figure 3, with raw data in place of the rounded data. Note that in all the various settings, the influence of the hyper-parameters (θ, σ) is stronger in the CDF of the simple Pitman–Yor species sampling model with respect to the corresponding predictive CDF derived from (33).
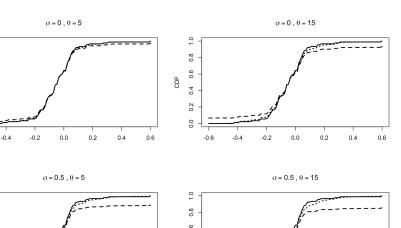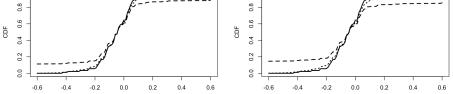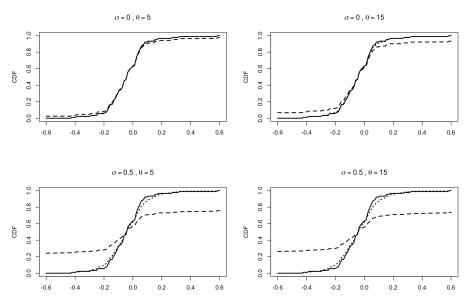
Figure 2. Predictive CDFs for the relative changes in larcenies between 1991 and 1995 (relative to 1991) for the 90 most populous US counties; data taken from Section 2.1 of [34]. Data have been rounded to the second decimal. Here, $n = 90$ and $k = 36$. Solid line: empirical CDF. Dotted line: predictive CDF from (33). Dashed line: predictive CDF from PS2 with $H = \mathcal{T}_{2\alpha_0}(\cdot|\mu_0, \sigma_0^2)$, $\sigma_0^2 = \beta_0(k_0 + 1)/\alpha_0 k_0$. Different plots correspond to different values of θ and σ. In all the plots, the predictive CDFs are evaluated with $\mu_0 = 0$, $\alpha_0 = 0.1$, $\beta_0 = 0.1$ and $k_0 = 0.1$.

Figure 3. Predictive CDFs for the relative changes in larcenies between 1991 and 1995 (relative to 1991) for the 90 most populous US counties; data taken from Section 2.1 of [34]. Raw data, without rounding. Here, $n = 90$ and $k = 36$. Solid line: empirical CDF. Dotted line: predictive CDF from (33). Dashed line: predictive CDF from PS2 with $H = \mathcal{T}_{2\alpha_0}(\cdot|\mu_0, \sigma_0^2)$, $\sigma_0^2 = \beta_0(k_0 + 1)/\alpha_0 k_0$. Different plots correspond to different values of θ and σ. In all the plots, the predictive CDFs are evaluated with $\mu_0 = 0$, $\alpha_0 = 0.1$, $\beta_0 = 0.1$ and $k_0 = 0.1$.

5.2. Predictive Distributions for gSSS

We now deduce an explicit form for the predictive distribution of a *gSSS* with general base measure H given in (25).

Recall that we denote by $\tilde{\Pi}_n$ the partition induced by $\xi_{1:n}$, with $\tilde{K}_n = |\tilde{\Pi}_n|$, and by Π_n the latent partition appearing in Proposition 3. We also set

$$\zeta_i = \begin{cases} \ell & \text{if } \xi_i^* = \bar{x}_\ell \\ 0 & \text{if } \xi_i^* \in \mathbb{X}_0^c. \end{cases}$$

The variable ζ_i is a discrete random variable that takes value 0 if ξ_i^* comes from the diffuse component of H.

Let $\mathcal{P}_{\tilde{\pi}_n, z_{1:k}} \subset \mathcal{P}_{\tilde{\pi}_n}$ be the set of all the possible configurations of Π_n that are compatible with the observed partition $\tilde{\Pi}_n = \tilde{\pi}_n$ and the additional information given by $\zeta_{1:\tilde{K}_n} = z_{1:k}$, $\tilde{K}_n = k$. In order to describe this set, observe that if $z_i > 0$, then the block $\tilde{\pi}_{i,n}$ may arise from the union of more blocks in π_n, while, if $z_i = 0$, then $\tilde{\pi}_{i,n} = \pi_{\phi(i),n}$ for some ϕ. Note that it may happen that $\phi(i) \neq i$.

Recalling that the elements in $\boldsymbol{m} = (m_1, \ldots, m_k)$ in (17) are used to describe the numbers of sub-blocks into which the blocks of $\tilde{\pi}_n$ have been divided to form the latent partition π_n, it turns out that the set $\mathcal{P}_{\tilde{\pi}_n, z_{1:k}}$ has the additional constraint $m_i = 1$ whenever $z_i = 0$. These considerations yield that, starting from $\tilde{\pi}_n$ and $z_{1:k}$, the set of admissible $\boldsymbol{m}$ can also be described by resorting to the definition of $\mathcal{Z}(\boldsymbol{m})$ as follows:

$$\mathcal{M}(\boldsymbol{n}, z_{1:k}) := \left\{ \boldsymbol{m} \in \mathcal{M}(\boldsymbol{n}) : z_{1:k} \in \mathcal{Z}(\boldsymbol{m}) \right\}$$
$$= \left\{ \boldsymbol{m} \in \mathbb{N}^k : m_i = 1 \text{ if } z_i = 0, 1 \leq m_i \leq n_i \text{ if } z_i > 0 \right\}.$$

With this position, one has

$$\begin{aligned}
\mathcal{P}_{\tilde{\pi}_n, z_{1:k}} &= \cup_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n}) : z_{1:k} \in \mathcal{Z}(\boldsymbol{m})} \cup_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{n}, \boldsymbol{m})} \mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda}) \\
&= \cup_{\boldsymbol{m} \in \mathcal{M}(\boldsymbol{n}, z_{1:k})} \cup_{\boldsymbol{\lambda} \in \Lambda(\boldsymbol{n}, \boldsymbol{m})} \mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda}),
\end{aligned} \tag{34}$$

where $\Lambda(\boldsymbol{n}, \boldsymbol{m})$ and $\mathcal{P}_{\tilde{\pi}_n}(\boldsymbol{\lambda})$ have been defined in Section 4.1.

For any $\boldsymbol{m}$ in $\mathcal{M}(\boldsymbol{n}, z_{1:k})$ and any $\boldsymbol{\lambda}$ in $\Lambda(\boldsymbol{n}, \boldsymbol{m})$, we define

$$\boldsymbol{\lambda}^{new} := [\lambda_1, \ldots, \lambda_k, 1].$$

In other words, $\boldsymbol{\lambda}^{new}$ corresponds to the configuration obtained from $\boldsymbol{\lambda}$ by adding one new element as a new block. In the following, let $\tilde{\mathbf{N}} = (|\tilde{\Pi}_{1,n}|, \ldots, |\tilde{\Pi}_{\tilde{K}_n, n}|)$, and let $\tilde{\mathbf{N}}^{i+}$ be obtained from $\tilde{\mathbf{N}}$ by adding 1 to its i-th component.

Proposition 8. *Let $(\xi_n)_{n \geq 1}$ be a $gSSS(\mathfrak{q}, H)$. Then, for any A in $\mathcal{X}$,*

$$P\{\xi_{n+1} \in A | \xi_{1:n}\} = \frac{1}{\mathcal{Z}_n} \left(\sum_{i=1}^{\tilde{K}_n} w_i \delta_{\xi_i^*}(A) + w_0 \bar{H}_n(A) \right) \qquad a.s.$$

where

$$\bar{H}_n(A) := \left[\sum_{\ell : \bar{x}_\ell \notin \xi_{1:\tilde{K}_n}^*} a_\ell \delta_{\bar{x}_\ell}(A) + (1-a) H^c(A) \right] H(\mathbb{X} \setminus \xi_{1:\tilde{K}_n}^*)^{-1},$$

$$w_i := \sum_{\boldsymbol{m} \in \mathcal{M}(\tilde{\mathbf{N}}^{i+}, \zeta_{1:\tilde{K}_n})} \prod_{j : \zeta_j > 0} a_{\zeta_j}^{m_j} \sum_{\boldsymbol{\lambda} \in \Lambda(\tilde{\mathbf{N}}^{i+}, \boldsymbol{m})} c(\boldsymbol{\lambda}) \bar{\mathfrak{q}}(\boldsymbol{\lambda}),$$

$$w_0 := H(\mathbb{X} \setminus \xi_{1:\tilde{K}_n}^*) \sum_{\boldsymbol{m} \in \mathcal{M}(\tilde{\mathbf{N}}, \zeta_{1:\tilde{K}_n})} \prod_{j : \zeta_j > 0} a_{\zeta_j}^{m_j} \sum_{\boldsymbol{\lambda} \in \Lambda(\tilde{\mathbf{N}}, \boldsymbol{m})} c(\boldsymbol{\lambda}) \bar{\mathfrak{q}}(\boldsymbol{\lambda}^{new}),$$

$$\mathcal{Z}_n := \sum_{\boldsymbol{m} \in \mathcal{M}(\tilde{\mathbf{N}}, \zeta_{1:\tilde{K}_n})} \prod_{j : \zeta_j > 0} a_{\zeta_j}^{m_j} \sum_{\boldsymbol{\lambda} \in \Lambda(\tilde{\mathbf{N}}, \boldsymbol{m})} c(\boldsymbol{\lambda}) \bar{\mathfrak{q}}(\boldsymbol{\lambda}).$$

Proof. We start by defining the following events for $i = 1, \ldots, \tilde{K}_n$:

$$E_i = \{\xi_{n+1} = \xi_i^*\}, \quad E_{new} = \{\xi_{n+1} \notin \xi_{1:\tilde{K}_n}^*\}.$$

Since conditioning on $\xi_{1:n}$ is equivalent to the condition on $[\xi_{1:\tilde{K}_n}^*, \tilde{\Pi}_n]$, one can write

$$P\{\xi_{n+1} \in A | \xi_{1:n}\} = \sum_{i=1}^{\tilde{K}_n} P\{\xi_{n+1} \in A, E_i | \xi_{1:\tilde{K}_n}^*, \tilde{\Pi}_n\} + P\{\xi_{n+1} \in A, E_{new} | \xi_{1:\tilde{K}_n}^*, \tilde{\Pi}_n\}$$

Now, set

$$E_{new}^* := \{\mathscr{C}_{n+1}(\Pi) = |\Pi_n| + 1\}$$

and

$$E_i^* = \{|\mathscr{C}_{n+1}(\Pi)| \leq |\Pi_n| \text{ and } \Pi_{\mathscr{C}_{n+1}(\Pi),n} \subset \tilde{\Pi}_{i,n}\}.$$

On $\zeta_i = 0$, one has (up to zero probability sets)

$$\{\xi_{n+1} \in A\} \cap E_i = \{\xi_i^* \in A\} \cap E_i^*$$

while, on $\zeta_i > 0$ (up to zero probability sets),

$$\{\xi_{n+1} \in A\} \cap E_i = \left(\{\xi_i^* \in A\} \cap E_i^*\right) \cup \left(\{\xi_i^* \in A\} \cap \{Z'_{|\Pi_n|+1} = \bar{x}_{\zeta_i}\} \cap E_{new}^*\right).$$

Note that (up to zero probability sets)

$$\{Z'_{|\Pi_n|+1} = \bar{x}_{\zeta_i}\} \cap E_{new}^* \cap \{\zeta_i = 0\} = \emptyset.$$

Hence,

$$\begin{aligned} &\mathbb{P}\{\xi_{n+1} \in A, E_i | \xi_{1:\tilde{K}_n}^*, \tilde{\Pi}_n\} \\ &= \delta_{\xi_i^*}(A)\mathbb{P}\{E_i^* \cup (\{Z'_{|\Pi_n|+1} = \bar{x}_{\zeta_i}\} \cap E_{new}^*) | \tilde{\Pi}_n, \xi_{1:\tilde{K}_n}^*\}. \end{aligned} \tag{35}$$

Similarly, using that $E_{new} \subset E_{new}^*$, one obtains

$$\begin{aligned} \mathbb{P}\{\xi_{n+1} \in A, E_{new} | \xi_{1:n}\} &= \mathbb{P}\{\xi_{n+1} \in A, E_{new}, E_{new}^* | \tilde{\Pi}_n, \xi_{1:\tilde{K}_n}^*\} \\ &= \mathbb{P}\{\xi_{n+1} \in A, E_{new} | \xi_{1:\tilde{K}_n}^*, E_{new}^*\}\mathbb{P}\{E_{new}^* | \tilde{\Pi}_n, \xi_{1:\tilde{K}_n}^*\} \\ &= H_n(A)\mathbb{P}\{E_{new}^* | \tilde{\Pi}_n, \xi_{1:\tilde{K}_n}^*\} \end{aligned} \tag{36}$$

where

$$H_n(A) = \sum_{\ell:\bar{x}_\ell \notin \xi_{1:\tilde{K}_n}^*} \bar{a}_\ell \delta_{\bar{x}_\ell}(A) + (1-a)H^c(A).$$

At this stage, note that, by construction,

$$\mathcal{L}(\xi_{1:\tilde{K}_n}^* | \Pi_n, Z'_{|\Pi_n|+1}, \zeta_{1:\tilde{K}_n}, \tilde{\Pi}_n, \Pi_{n+1}) = \mathcal{L}(\xi_{1:\tilde{K}_n}^* | \zeta_{1:\tilde{K}_n})$$

where $\mathcal{L}(\xi_{1:\tilde{K}_n}^* | \zeta_{1:\tilde{K}_n})$ is characterized by

$$P(\xi_1^* \in A_1, \ldots, \xi_{\tilde{K}_n}^* \in A_{\tilde{K}_n} | \zeta_{1:\tilde{K}_n}) = \prod_{i=1}^{\tilde{K}_n} \left(H^c(A_i)\mathbf{1}\{\zeta_i = 0\} + \delta_{\bar{x}_{\zeta_i}}(A_i)\mathbf{1}\{\zeta_i > 0\}\right),$$

and then

$$\mathcal{L}(Z'_{|\Pi_n|+1}, \zeta_{1:\tilde{K}_n}, \tilde{\Pi}_n, \Pi_{n+1}, \xi_{1:\tilde{K}_n}^* | \Pi_n) = \mathcal{L}(Z'_{|\Pi_n|+1}, \zeta_{1:\tilde{K}_n}, \tilde{\Pi}_n, \Pi_{n+1} | \Pi_n)\mathcal{L}(\xi_{1:\tilde{K}_n}^* | \zeta_{1:\tilde{K}_n}).$$

Hence,

$$
\begin{aligned}
\mathcal{L}(\Pi_n, \Pi_{n+1}, Z'_{|\Pi_n|+1}, & \xi^*_{1:\tilde{K}_n}, \zeta_{1:\tilde{K}_n}, \tilde{\Pi}_n) \\
&= \mathcal{L}(\Pi_n)\mathcal{L}(Z'_{|\Pi_n|+1}, \zeta_{1:\tilde{K}_n}, \tilde{\Pi}_n, \Pi_{n+1}|\Pi_n)\mathcal{L}(\xi^*_{1:\tilde{K}_n}|\zeta_{1:\tilde{K}_n})
\end{aligned} \tag{37}
$$

which shows, in particular, that $[\Pi_n, \Pi_{n+1}, Z'_{|\Pi_n|+1}, \tilde{\Pi}_n]$ and $\xi^*_{1:\tilde{K}_n}$ are conditionally independent given $\zeta_{1:\tilde{K}_n}$. Since E_i^*, E_{new}^* and $\{Z'_{|\Pi_n|+1} = \bar{x}_{\zeta_i}\} \cap E_{new}^*$ depend logically only on $\Pi_{n+1}, Z'_{|\Pi_n|+1}, \tilde{\Pi}_n, \zeta_{1:\tilde{K}_n}$, one obtains

$$
\begin{aligned}
\mathbb{P}\{E_i^*|\tilde{\Pi}_n, \xi^*_{1:\tilde{K}_n}\} &= \mathbb{P}\{E_i^*|\tilde{\Pi}_n, \zeta_{1:\tilde{K}_n}\}, \\
\mathbb{P}\{(Z'_{|\Pi_n|+1} = \bar{x}_{\zeta_i}) \cap E_{new}^*|\tilde{\Pi}_n, \xi^*_{1:\tilde{K}_n}\} &= \mathbb{P}\{(Z'_{|\Pi_n|+1} = \bar{x}_{\zeta_i}) \cap E_{new}^*|\tilde{\Pi}_n, \zeta_{1:\tilde{K}_n}\}
\end{aligned} \tag{38}
$$

and, finally,

$$
\mathbb{P}\{E_{new}^*|\tilde{\Pi}_n, \xi^*_{1:\tilde{K}_n}\} = \mathbb{P}\{E_{new}^*|\tilde{\Pi}_n, \zeta_{1:\tilde{K}_n}\}. \tag{39}
$$

Since $[\tilde{\Pi}_n, \zeta_{1:\tilde{K}_n}, \tilde{K}_n]$ are discrete random variables, we use the elementary definition of the conditional probability of events to evaluate the conditional distributions (38) and (39). Specifically, assume that $\tilde{K}_n = k$, $[\tilde{\Pi}_n, \zeta_{1:\tilde{K}_n}] = [\tilde{\pi}_n, z_{1:k}]$, $\tilde{\mathbf{N}} = \boldsymbol{n}$, and, for a given event E, write

$$
\mathbb{P}\{E|\tilde{\Pi}_n = \tilde{\pi}_n, \zeta_{1:\tilde{K}_n} = z_{1:k}\} = \frac{\mathbb{P}\{E, \tilde{\Pi}_n = \tilde{\pi}_n, \zeta_{1:\tilde{K}_n} = z_{1:k}\}}{\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n, \zeta_{1:\tilde{K}_n} = z_{1:k}\}}. \tag{40}
$$

As for the denominator in (40), letting $\mathcal{M}_n^* = \mathcal{M}(\boldsymbol{n}, z_{1:k})$ and $J = \#\{i : z_i > 0\}$, using (34), one obtains

$$
\begin{aligned}
\mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n, & \zeta_{1:\tilde{K}_n} = z_{1:k}\} \\
&= \sum_{\pi_n \in \mathcal{P}_{\tilde{\pi}_n, z_{1:k}}} \mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n, \zeta_{1:\tilde{K}_n} = z_{1:k}, \Pi_n = \pi_n\} \\
&= \sum_{\boldsymbol{m} \in \mathcal{M}_n^*} \sum_{\lambda \in \Lambda(\boldsymbol{n}, \boldsymbol{m})} \sum_{\pi_n \in \mathcal{P}_{\tilde{\pi}_n}(\lambda)} \mathbb{P}\{\tilde{\Pi}_n = \tilde{\pi}_n, \zeta_{1:n} = z_{1:k}, \Pi_n = \pi_n\} \\
&= (1-a)^{k-J} \sum_{\boldsymbol{m} \in \mathcal{M}_n^*} \prod_{j:z_j>0} a_{z_j}^{m_j} \sum_{\lambda \in \Lambda(\boldsymbol{n}, \boldsymbol{m})} \sum_{\pi_n \in \mathcal{P}_{\tilde{\pi}_n}(\lambda)} \bar{\mathsf{q}}(\lambda) \\
&= (1-a)^{k-J} \sum_{\boldsymbol{m} \in \mathcal{M}_n^*} \prod_{j:z_j>0} a_{z_j}^{m_j} \sum_{\lambda \in \Lambda(\boldsymbol{n}, \boldsymbol{m})} c(\lambda)\bar{\mathsf{q}}(\lambda).
\end{aligned}
$$

As for the numerators in (40), when $E = E_{new}^*$, we start from

$$
\begin{aligned}
P\{E_{new}^*, & \tilde{\Pi}_n = \tilde{\pi}_n, \zeta_{1:\tilde{K}_n} = z_{1:k}, \Pi_n = \pi_n\} \\
&= P\{E_{new}^*|\Pi_n = \pi_n\}P\{\tilde{\Pi}_n = \tilde{\pi}_n, \zeta_{1:\tilde{K}_n} = z_{1:k}, \Pi_n = \pi_n\} \\
&= P\{E_{new}^*|\Pi_n = \pi_n\}(1-a)^{k-J} \prod_{j:z_j>0} a_{z_j}^{m_j} \bar{\mathsf{q}}(\lambda) \\
&= (1-a)^{k-J} \prod_{j:z_j>0} a_{z_j}^{m_j} \bar{\mathsf{q}}(\lambda^{new}).
\end{aligned}
$$

where, in the last equality, we used that for $\pi_n \in \mathcal{P}_{\tilde{\pi}_n}(\lambda)$, one has

$$
P\{E_{new}^*|\Pi_n = \pi_n\} = \frac{\bar{\mathsf{q}}(\lambda^{new})}{\bar{\mathsf{q}}(\lambda)}.
$$

Taking the sum over $\mathcal{P}_{\tilde{\pi}_n, z_{1:k}}$ gives

$$
P\{E_{new}^*, \tilde{\Pi}_n = \tilde{\pi}_n, \zeta_{1:\tilde{K}_n} = z_{1:k}\} = (1-a)^{k-J} \sum_{\boldsymbol{m} \in \mathcal{M}_n^*} \prod_{j:z_j>0} a_{z_j}^{m_j} \sum_{\lambda \in \Lambda(\boldsymbol{n}, \boldsymbol{m})} c(\lambda)\bar{\mathsf{q}}(\lambda^{new}).
$$

Combining these with (39) and (40) and recalling that $\mathcal{M}_n^* = \mathcal{M}(n, z_{1:k})$, one obtains

$$\mathbb{P}\{E_{new}^* | \tilde{\Pi}_n, \xi_{1:\tilde{K}_n}^*\} = \frac{\sum_{m \in \mathcal{M}(n, z_{1:k})} \prod_{j:z_j>0} a_{z_j}^{m_j} \sum_{\lambda \in \Lambda(n,m)} c(\lambda)\bar{q}(\lambda^{new})}{\sum_{m \in \mathcal{M}(n, z_{1:k})} \prod_{j:z_j>0} a_{z_j}^{m_j} \sum_{\lambda \in \Lambda(n,m)} c(\lambda)\bar{q}(\lambda)}.$$

Finally, it remains to consider (40) when $E = E_i^* \cup (\{Z'_{|\Pi_n|+1} = \bar{x}_{\zeta_i}\} \cap E_{new}^*)$. Now, observe that

$$\left(E_i^* \cup (\{Z'_{|\Pi_n|+1} = \bar{x}_{\zeta_i}\} \cap E_{new}^*)\right) \cap \{\tilde{\Pi}_n = \tilde{\pi}_n, \zeta_{1:\tilde{K}_n} = z_{1:k}\}$$
$$= \{\tilde{\Pi}_{n+1} = \tilde{\pi}_n^{i+}, \zeta_{1:\tilde{K}_n} = z_{1:k}\} = \{\tilde{\Pi}_{n+1} = \tilde{\pi}_n^{i+}, \zeta_{1:\tilde{K}_{n+1}} = z_{1:k}\}$$

where $\tilde{\pi}_n^{i+}$ denote the partition in $\mathcal{P}_{n+1}$ obtained from $\tilde{\pi}_n$ by adding $n+1$ to the i-th block of $\tilde{\pi}_n$. Note that, for the second equality, we used that, on $\tilde{\Pi}_{n+1} = \tilde{\pi}_n^{i+}$, one has $\tilde{K}_{n+1} = \tilde{K}_n = k$.

Hence, using (34) with $\tilde{\pi}_n^{i+}$ in place of $\tilde{\pi}_n$, one concludes that

$$\mathbb{P}\{\tilde{\Pi}_{n+1} = \tilde{\pi}_n^{i+}, \zeta_{1:\tilde{K}_{n+1}} = z_{1:k}\}$$
$$= \sum_{\substack{m \in \mathcal{M}(n^{i+}, z_{1:k}) \\ \lambda \in \Lambda(n,m)}} \sum_{\pi_{n+1} \in \mathcal{P}_{\tilde{\pi}_{n+1}}(\lambda)} P\{\tilde{\Pi}_{n+1} = \tilde{\pi}_n^{i+}, \Pi_{n+1} = \pi_{n+1}, \zeta_{1:\tilde{K}_{n+1}} = z_{1:k}\}$$
$$= (1-a)^{k-J} \sum_{m \in \mathcal{M}(n^{i+}, z_{1:k})} \prod_{j:z_j>0} a_{z_j}^{m_j} \sum_{\lambda \in \Lambda(n,m)} c(\lambda)\bar{q}(\lambda)$$

where $n^{i+} = (n_1, \ldots, n_i+1, \ldots, n_k)$. Hence, by (38)–(40), one can write

$$\mathbb{P}\{E_i^* \cup (\{Z'_{|\Pi_n|+1} = \bar{x}_{\zeta_i}\} \cap E_{new}^*) | \tilde{\Pi}_n, \xi_{1:\tilde{K}_n}^*\}$$
$$= \frac{\sum_{m \in \mathcal{M}(n^{i+}, z_{1:k})} \prod_{j:z_j>0} a_{z_j}^{m_j} \sum_{\lambda \in \Lambda(n,m)} c(\lambda)\bar{q}(\lambda)}{\sum_{m \in \mathcal{M}(n, z_{1:k})} \prod_{j:z_j>0} a_{z_j}^{m_j} \sum_{\lambda \in \Lambda(n,m)} c(\lambda)\bar{q}(\lambda)}.$$

Combining these results, one obtains the thesis. $\quad\square$

6. Conclusions and Discussion

We have defined a new class of exchangeable sequences, called mixtures of species sampling sequences (mSSS). We have shown that these sequences include various well-known Bayesian nonparametric models. In particular, the observations of many nonparametric hierarchical models (e.g., hierarchical Dirichlet process, hierarchical Pitman–Yor process and, more generally, hierarchical species sampling models [22–25]) are mSSS. We have shown that also observations sampled from a mixture of Dirichlet processes [10] are mSSS, under some additional assumptions. Our general class also includes species sampling sequences with a general (not necessarily diffuse) base measure, which have been used in various applications, e.g., in the case of "spike-and-slab"-type nonparametric priors [16–21].

We believe that our general framework sheds light on the common structure of all the above-mentioned models, leading to a possible unified treatment of some of their important features. Our techniques provides unified proofs for various results that, up to now, have been proven with ad hoc methods.

We have proven that all the mSSS are obtained by assigning the values of an exchangeable sequence to the classes of a latent exchangeable random partition. This representation is proven in the strong sense of an almost sure equality (see Section 3) and leads to the simple and clear derivation of an explicit expression for the EPPF of an mSSS. We believe that our general proof simplifies the derivation of the EPPF of many of the above-mentioned particular cases. Moreover, our results show that the clustering and the predictive structure

of various well-known models do not depend on the relation between these models and completely random measures, but are essentially a consequence of the simple combinatorial structure of these sequences. Many important differences between well-known models (such as mixtures of Dirichlet and hierarchical Dirichlet) can be explained easily by simple differences in the latent partition and the corresponding latent exchangeable sequence.

We stress that a clear understanding of the clustering structure of mSSS is fundamental for practical purposes, since these models are typically used to cluster observations. Moreover, we hope that the explicit expression for EPPFs in our general framework can lead to the development of new MCMC algorithms for sampling from the posterior distribution.

Finally, we believe that some of the results we have proven here for mSSS can be extended to the more general case of partially exchangeable arrays. In this direction, for future works, a possible generalization of mSSS is to consider partially exchangeable arrays with a mixture of species sampling random probability measures as directing measures.

Author Contributions: F.B.: Methodology, simulation, writing and editing. L.L.: Methodology, writing and editing. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

In what follows, $\mathcal{L}(X)$ denotes the law of a random element X. For ease of reference, we state here Lemma 5.9 and Corollary 5.11 of [35].

Lemma A1 (Extension 1)**.** *Fix a probability kernel K between two measurable spaces S and T, and let σ be a random element defined on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in S. Then, there exists a random element η in T, defined on some extension of the original probability space Ω, such that $\mathbb{P}[\eta \in \cdot | \sigma] = K(\cdot | \sigma)$ a.s. and, moreover, η is conditionally independent given σ from any other random element on Ω.*

Lemma A2 (Extension 2)**.** *Fix two Borel spaces S and T, a measurable mapping $f : T \to S$ and some random elements σ in S and $\tilde{\eta}$ in T with $\mathcal{L}(\sigma) = \mathcal{L}(f(\tilde{\eta}))$. Then, there is a random element η defined on some extension of the original probability space, such that $\mathcal{L}(\eta) = \mathcal{L}(\tilde{\eta})$ and $\sigma = f(\eta)$ a.s.*

We need the following variant of the previous result.

Lemma A3 (Extension 3)**.** *Fix three Borel spaces S_1, S_2 and T_1, a measurable mapping $\phi : T_1 \times S_2 \to S_1$ and some random elements $\sigma = (\sigma_1, \sigma_2)$ in $S_1 \times S_2$ and τ_1 in T_1, all defined on a probability space $(\Omega, \mathcal{F}, P)$. Assume that the conditional law of σ_1 given σ_2 is the same as the conditional law of $\phi(\tau_1, \sigma_2)$ given σ_2 (P-almost surely). Then, there is a random element τ defined on some extension of the original probability space $(\Omega, \mathcal{F}, P)$ taking values in T_1 such that*

- *$\sigma_1 = \phi(\tau, \sigma_2)$ a.s.*
- *$\mathcal{L}(\tau_1, \sigma_2) = \mathcal{L}(\tau, \sigma_2)$.*

Proof. Define $f : T_1 \times S_2 =: T \to S_1 \times S_2 =: S$ by $f(a, b) = (\phi(a, b), b)$, set $\tilde{\eta} = (\tau_1, \sigma_2)$ and $\sigma = (\sigma_1, \sigma_2)$. By hypothesizing, it is clear that $\mathcal{L}(f(\tilde{\eta})) = \mathcal{L}((\phi(\tau_1, \sigma_2), \sigma_2)) = \mathcal{L}(\sigma_1, \sigma_2) = $

$\mathcal{L}(\sigma)$. Thus, by Lemma A2, one has that, on an enlargement of $(\Omega, \mathcal{F}, P)$, there exists $\eta := (\tau, \sigma_2^*)$ such that $\mathcal{L}(\eta) = \mathcal{L}(\tilde{\eta})$ and $(\phi(\tau, \sigma_2^*), \sigma_2^*) = f(\eta) = \sigma = (\sigma_1, \sigma_2)$ a.s. Hence, $\sigma_2^* = \sigma_2$ a.s. but also $\phi(\tau, \sigma_2) = \phi(\tau, \sigma_2^*) = \sigma_1$ a.s. It remains to show the second part of the thesis. Since $(\tau, \sigma_2) = (\tau, \sigma_2^*) = \eta$ a.s. and $\mathcal{L}(\eta) = \mathcal{L}(\tilde{\eta})$, where $\tilde{\eta} = (\tau_1, \sigma_2)$, it follows that $\mathcal{L}(\tau, \sigma_2) = \mathcal{L}(\tau_1, \sigma_2)$. $\square$

References

1. Ferguson, T.S. A Bayesian analysis of some nonparametric problems. *Ann. Stat.* **1973**, *1*, 209–230. [CrossRef]
2. Pitman, J.; Yor, M. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **1997**, *25*, 855–900. [CrossRef]
3. Perman, M.; Pitman, J.; Yor, M. Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Relat. Fields* **1992**, *92*, 21–39. [CrossRef]
4. Regazzini, E.; Lijoi, A.; Prünster, I. Distributional results for means of normalized random measures with independent increments. *Ann. Stat.* **2003**, *31*, 560–585. [CrossRef]
5. James, L.F.; Lijoi, A.; Prünster, I. Posterior analysis for normalized random measures with independent increments. *Scand. J. Stat.* **2009**, *36*, 76–97. [CrossRef]
6. Lijoi, A.; Prünster, I. Models beyond the Dirichlet process. In *Bayesian Nonparametrics*; Hjort, N.L., Holmes, C., Müller, P., Walker, S., Eds.; Cambridge University Press: New York, NY, USA, 2010.
7. De Blasi, P.; Favaro, S.; Lijoi, A.; Mena, R.H.; Prunster, I.; Ruggiero, M. Are Gibbs-Type Priors the Most Natural Generalization of the Dirichlet Process? *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 212–229. [CrossRef]
8. Pitman, J. Poisson-Kingman partitions. In *Statistics and Science: A Festschrift for Terry Speed*; IMS Lecture Notes Monograph Series; Institute of Mathematical Statistics: Beachwood, OH, USA, 2003; Volume 40, pp. 1–34.
9. Ishwaran, H.; James, L.F. Gibbs sampling methods for stick-breaking priors. *J. Am. Stat. Assoc.* **2001**, *96*, 161–173. [CrossRef]
10. Antoniak, C.E. Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Stat.* **1974**, *2*, 1152–1174. [CrossRef]
11. Cifarelli, D.M.; Regazzini, E. Distribution functions of means of a Dirichlet process. *Ann. Stat.* **1990**, *18*, 429–442. [CrossRef]
12. Sangalli, L.M. Some developments of the normalized random measures with independent increments. *Sankhyā* **2006**, *68*, 461–487.
13. Broderick, T.; Wilson, A.C.; Jordan, M.I. Posteriors, conjugacy, and exponential families for completely random measures. *Bernoulli* **2018**, *24*, 3181–3221. [CrossRef]
14. Bassetti, F.; Ladelli, L. Asymptotic number of clusters for species sampling sequences with non-diffuse base measure. *Stat. Probab. Lett.* **2020**, *162*, 108749. [CrossRef]
15. Pitman, J. Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory*; IMS Lecture Notes Monograph Series; Institute of Mathematical Statistics: Hayward, CA, USA, 1996; Volume 30, pp. 245–267. [CrossRef]
16. Dunson, D.B.; Herring, A.H.; Engel, S.M. Bayesian selection and clustering of polymorphisms in functionally related genes. *J. Am. Stat. Assoc.* **2008**, *103*, 534–546. [CrossRef]
17. Kim, S.; Dahl, D.B.; Vannucci, M. Spiked Dirichlet process prior for Bayesian multiple hypothesis testing in random effects models. *Bayesian Anal.* **2009**, *4*, 707–732. [CrossRef] [PubMed]
18. Suarez, A.J.; Ghosal, S. Bayesian Clustering of Functional Data Using Local Features. *Bayesian Anal.* **2016**, *11*, 71–98. [CrossRef]
19. Cui, K.; Cui, W. Spike-and-Slab Dirichlet Process Mixture Models. *Spike Slab Dirichlet Process. Mix. Model.* **2012**, *2*, 512–518. [CrossRef]
20. Barcella, W.; De Iorio, M.; Baioa, G.; Malone-Leeb, J. Variable selection in covariate dependent random partition models: An application to urinary tract infection. *Stat. Med.* **2016**, *35*, 1373–13892. [CrossRef]
21. Canale, A.; Lijoi, A.; Nipoti, B.; Prünster, I. On the Pitman–Yor process with spike and slab base measure. *Biometrika* **2017**, *104*, 681–697. [CrossRef]
22. Teh, Y.; Jordan, M.I. Hierarchical Bayesian nonparametric models with applications. In *Bayesian Nonparametrics*; Hjort, N.L., Holmes, C., Müller, P., Walker, S., Eds.; Cambridge University Press: New York, NY, USA, 2010.
23. Teh, Y.W.; Jordan, M.I.; Beal, M.J.; Blei, D.M. Hierarchical Dirichlet processes. *J. Am. Stat. Assoc.* **2006**, *101*, 1566–1581. [CrossRef]
24. Camerlenghi, F.; Lijoi, A.; Orbanz, P.; Pruenster, I. Distribution theory for hierarchical processes. *Ann. Stat.* **2019**, *1*, 67–92. [CrossRef]
25. Bassetti, F.; Casarin, R.; Rossini, L. Hierarchical Species Sampling Models. *Bayesian Anal.* **2020**, *15*, 809–838. [CrossRef]
26. Pitman, J. *Combinatorial Stochastic Processes*; Lectures from the 32nd Summer School on Probability Theory Held in Saint-Flour, 7–24 July 2002, with a Foreword by Jean Picard; Lecture Notes in Mathematics; Springer: Berlin, Germany, 2006; Volume 1875.
27. Crane, H. The ubiquitous Ewens sampling formula. *Stat. Sci.* **2016**, *31*, 1–19. [CrossRef]
28. Kingman, J.F.C. The representation of partition structures. *J. Lond. Math. Soc.* **1978**, *18*, 374–380. [CrossRef]
29. Aldous, D.J. Exchangeability and related topics. In *École d'été de Probabilités de Saint-Flour, XIII—1983*; Lecture Notes in Mathematics; Springer: Berlin, Germany, 1985; Volume 1117, pp. 1–198. [CrossRef]
30. Kallenberg, O. Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrscheinlichkeitstheorie Und Verw. Geb.* **1973**, *27*, 23–36. [CrossRef]
31. Pitman, J. Exchangeable and partially exchangeable random partitions. *Probab. Theory Relat. Fields* **1995**, *102*, 145–158. [CrossRef]

32. Gnedin, A.; Pitman, J. Exchangeable Gibbs partitions and Stirling triangles. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **2005**, *325*, 83–102, 244–245. [CrossRef]
33. Schervish, M.J. *Theory of Statistics*; Springer Series in Statistics; Springer: New York, NY, USA, 1995. [CrossRef]
34. Marin, J.M.; Robert, C.P. *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*; Springer Texts in Statistics; Springer: New York, NY, USA, 2007; pp. xiv+255.
35. Kallenberg, O. *Foundations of Modern Probability*, 3rd ed.; Probability Theory and Stochastic Modelling; Springer: New York, NY, USA, 2021; Volume 99.