Reinforced Urn Processes Based on Pitman Sequences.

D. Spanò [a,*], P. Muliere [b],

[a] *University of Oxford, Department of Statistics, 1 South Parks Road, Oxford OX1 3TG, UK.*

[b] *Università "L. Bocconi", Istituto Metodi Quantitativi, v.le Isonzo 25, 20135 Milano, Italy.*

Abstract

We construct a generalized version of Muliere *et al* (2000) Reinforced Urn Processes, with two major differences: (i) the state space is allowed to be continuous and (ii) the urns involved in the reinforcement device are those introduced by Pitman (1995) and (1996). Such urns do not necessarily generate exchangeable sequences. We study the dynamic and asymptotic properties of such a process. In particular we derive a representation theorem which hints for a weak notion of mixture of Markov chain distributions. We study the connection with some known priors in Bayesian Nonparametrics.

Key words: Reinforced processes, partial exchangeability, mixtures of Markov chains, urn schemes

1 Introduction.

Pólya's urn is the most popular sampling scheme used to generate exchangeable sequences, but it can be a basis to generate non-exchangeable processes as well. Muliere *et al.* (2000) introduced in the context of Bayesian nonparametrics the so-called Reinforced Urn Processes (RUP), i.e. reinforced random walks on a countable state space, whose points are associated to discrete Pólya's urns, and stressed the connection with many nonparametric prior measures.

* Corresponding author.
Email addresses: spano@stats.ox.ac.uk (D. Spanò), pietro.muliere@uni-bocconi.it (P. Muliere).

A RUP can be described by the following example. Suppose an individual (or particle) moves around in a space S made of a countable number of possible positions, that we call “islands”. At each given island, the particle remembers only the sequence of past migrations performed from *that* island, and tend to repeat them, with a rate equal to the number of times they were performed in the past. In this setting the sequence of all migrations from every fixed island, is a Pólya (exchangeable) sequence, whereas the whole process turns out to be Markov-exchangeable: under some recurrence conditions, such a process turns out to be a mixture of Markov chain distributions (see Diaconis and Freedman (1981), Zaman (1984)).

In this paper we introduce a model where we are interested not only in the island the particle moves to, but also in the exact point within that island where the particle chooses to “rest” until the next migration. We construct an extended RUP process suitable for this problem when it is assumed that:

(i) the set of all possible exact locations in each island is uncountable;

(ii) before migrating, the particle remembers, as in the original RUP, only the migrations performed in the past from the current island, say i; however in the new model the particle will consider any two migrations from island i to island j as distinct if they lead to two different exact locations in j;

(iii) the memory of the particle is also allowed to follow an updating rule which is not necessarily Pólya and may differ from island to island: the updating rule itself becomes a parameter of the model.

To this purpose, we construct a process on a continuous state space partitioned in a countable number of classes, and we will associate to each class an urn model with a sampling scheme of the type introduced by Pitman (1996); for this reason we will call such a process P-RUP. Pitman’s sampling schemes generate sequences on continuous state spaces which are not necessarily exchangeable, but which include Blackwell-MacQueen (1973)’s extension of Pólya’s scheme as a particular case. When all Pitman urns used in the process are exchangeable, and suitable recurrence conditions are satisfied, then the P-RUP is a concrete example of *Markov-exchangeable sequence with atomic kernel* i.e. it admits a representation as mixture of Markov chain laws (Fortini *et al* (2002)).

The choice of Pitman’s sampling schemes is motivated by the fact that they are useful for sampling in contexts, like Population Genetics, where the conditional probability, given the past observations, that the next item is of a new distinct type, does not depend on the labels (the colors) used to mark the previous observations. As stressed by Zabell (1992), when sampling from an urn of Pitman’s type, we don’t need to know in advance the set of colors that will actually be selected. Consistently P-RUPs, are an appealing tool to deal with contexts of similar nature, where the observations may not be exchangeable. In the case of the migrating particle, P-RUPs are a possible tool to predict the colonization of a new distinct island and/or the settlement into a new distinct location, where no Euclidean distance between locations is necessarily assumed.

The paper is structured as follows. In Section 2 we recall Muliere *et al.* RUPs; in Section 3, after a review of the basic traits of Pitman's schemes, we define Reinforced Urn Processes based on Pitman's sequences (P-RUP); in section 4 the connection with the discrete RUPs is pointed out; explicit formulas for the discovery of new distinct islands and locations are given in section 5; in section 6 we discuss the types of recurrence for a P-RUP and we find a limit theorem for recurrent P-RUPs; section 6 is devoted to studying the support of recurrent P-RUP; in section 7 we stress the connections with Bayesian nonparametrics and with known partition structures (e.g. 2-parameter GEM distribution), section 8 has some concluding remarks.

2 Reinforced Urn Processes.

Muliere *et al.* (2000) introduced Reinforced Urn Processes (RUP) as reinforced random walks on a countable state space of independent Pólya urns. The definition of the process depends on four elements:

1. A countable state space of "islands ", S;

2. A finite set of colors E, with cardinality $k \geq 1$.

3. An urn composition U which maps S into the set of k-tuples of nonnegative real numbers whose sum is a strictly positive number.

4. A law of motion $q : S \times E \to S$ (for every $x \in S$ we denote it by $c \mapsto q_x(c)$).

A Pólya urn is associated to every point $x \in S$, whose initial composition is given by $U(x) = \{\alpha_x\{1\}, ..., \alpha_x\{l\}\}$. The law of motion q is assumed to be such that, to every $x, y \in S$, there is at most one color $c(x,y) \in E$ such that $q_x(c(x,y)) = y$. The RUP $X = (X_n)_{n \geq 0}$ is defined recursively, as follows: fix a starting point $X_0 = x_0$. For every $n \geq 1$, if $X_{n-1} = x \in S$, then a ball is randomly picked from $U(x)$, and its color $c \in E$ will indicate where X will go next, according to the function q, i.e. $X_n = q_x(c)$. Meanwhile the ball is returned in $U(x)$ along with another ball of color c (Pólya urn scheme). We will write $X \in RUP(S, E, U, q)$ with initial state x_0.

An equivalent definition of RUP, suitable for our purpose, is given in terms of their predictive distributions and the array of *successor states*. Consider, for every $m \geq 1$ and $x \in S$, the times $\tau_{x,m}$ of the m -th visit of X to the island $x \in S$, as:

$$\tau_{x,1} := \inf \{n : X_n = x\}$$

and, for $m \geq 2$

$$\tau_{x,m} := \inf \{n > \tau_{x,m-1} : X_n = x\} . \tag{1}$$

The m-th successor state of x is the next destination of the particle after its m-th visit to x:

$$V_{x,m} := X_{\tau_{x,m}+1} I (\tau_{x,m} < \infty) + \eta I (\tau_{x,m} = \infty) \tag{2}$$

for some conventional $\eta \notin S$.

Definition 1 *A random sequence X, starting from x_0, is a RUP with parameters (S, E, q, U) if it satisfies, for every $x, y \in S$:*

$$\mathbb{P} (X_{n+1} = y | X_{n-1} = x_{n-1}, X_n = x) = \frac{\tilde{\alpha}_x \{y\}}{\theta_x + n_x} + \frac{1}{\theta_x + n_x} \sum_{j=1}^{n_x} \delta_{V_{x,j}} \{y\} \tag{3}$$

where $\tilde{\alpha}_x \{y\} = \alpha_x \{q_x^{-1} (y)\}$, $\theta_x = \sum_{j=1}^{l} \tilde{\alpha}_x \{j\}$ and $n_x = \sum_{i=1}^{n-1} I (X_i = x)$.

In this definition the space of colors E is no longer essential if we replace the urn parameter U with $\tilde{U} = \{\tilde{U}(x) : x \in S\}$ where $\tilde{U}(x) = (\tilde{\alpha}_x\{y\} \geq 0 : y \in S)$ where, for every $x \in S$, $\tilde{\alpha}_x$ is positive only on finitely many points of S. We can therefore write $X \in RUP(S, \tilde{U})$.

If a $X \in RUP(S, \tilde{U})$ satisfies, for every $x \in S$

$$\mathbb{P} \left(\bigcap_{m>1} \{\tau_{x,m} < \infty\} | \tau_{x,1} < \infty \right) = 1, \tag{4}$$

that is, if X is *strongly recurrent*, then (Diaconis-Freedman (1981)) there exists an a.s. unique random kernel π on $S^* \times S^*$, where $S^* = S \cup \{\eta\}$, such that $X|\pi$ is a Markov chain with transition matrix π.

The random kernel $\pi = (\pi (x, \cdot) : x \in S^*)$ turns out to be made of independent rows given by

$$\pi (x, \cdot) = \lim_{n_x \to \infty} \frac{1}{n_x} \sum_{j=1}^{n_x} \delta_{V_{x,j}} (\cdot),$$

and therefore, by Pólya's properties, $\pi (x, \cdot)$ has a discrete Dirichlet with parameter $U (x)$, for every x such that $\tau_{x,1} < \infty$, whereas $\pi (x, \cdot) = \delta_\eta$ for all other x's (Muliere *et al.* (2000), 2.11); the necessity of augmenting S with η is exploited in Fortini *et al.* (2002)).

3 Reinforced Urn Processes based on Pitman's sequences.

In this section we extend the notion of reinforced urn process in two directions:

(1) we allow the spaces S and E to be measurable complete, separable metric spaces.

(2) instead of discrete Pólya's urn scheme, we use Pitman's scheme to model the reinforcement of the process.

As we said in the introduction, the choice of Pitman's family of prediction rules stems from the fact that minimal assumptions are made on the state space, which is uncountable and does not even need to be ordered. Sampling new items and labelling them are independent matters. Pitman sampling schemes lead to random distributions where the sizes of the atoms are independent of their locations. This makes Pitman's schemes a very general class of nonparametric models. We first recall how Pitman's model is described, and then we will construct extended RUPs.

For each $n \geq 1$, let $X_{(n)} = (X_1, \ldots, X_n)$ be an S-valued sample. To predict X_{n+1} under exchangeability, we would only need the information carried by the empirical distribution function induced by the sample:

$$F(X_{(n)}) = n^{-1} \sum_{i=1}^{n} \delta_{X_i}(\cdot).$$

Pitman's scheme is defined in terms of similar prediction rules, where the empirical measure, sufficient for prediction, is given by a more informative function:

$$T_n(X_{(n)}) = (K_n, L_n, \mathbf{N}_n)$$

where $K_n \leq n$ is the number of distinct values observed in $X_{(n)}$; $L_n = (\bar{X}_1, ...\bar{X}_k)$ is the set of such values ranked in their (random) order of appearance, and $\mathbf{N}_n = (N_1, ..., N_{K_n})$ is the vector of their absolute frequencies where $\sum_{j=1}^{K_n} N_j = n$.

Definition 2 *Let ν be a diffuse probability measure on $(S, \mathcal{S})$. A Pitman sequence is an S-valued exchangeable sequence $X = (X_1, X_2, ...)$ with predictive distribution given by:*

$$\mathbb{P}(X_1 \in \cdot) = \nu(\cdot) \tag{5}$$

and for $n \geq 1$,

$$\mathbb{P}(X_{n+1} \in \cdot | X_{(n)} = x_{(n)}) = f^* \left(T_n(x_{(n)}), \cdot \right) \tag{6}$$

where, if $T_n(x_{(n)}) = (k, \{y_1, \ldots, y_k\}, \mathbf{n}_n)$,

$$f^*(T_n(x_n), \cdot) = \sum_{j=1}^{k} \lambda_j(\mathbf{n}_n)\delta_{y_j}(\cdot) + \left(1 - \sum_{j=1}^{k} \lambda_j(\mathbf{n}_n)\right) \nu(\cdot), \tag{7}$$

for some non-negative function $\lambda = \{\lambda_j(\mathbf{n}_n) \in [0,1] : j = 1, \ldots, k+1\}$ *such that*
$\sum_{j=1}^{k} \lambda_j(\mathbf{n}_n) \leq 1$.

Pitman's prediction rules depend on two parameters: (λ, ν). The role of ν is of a sort of a random paintbrush, which assigns a distinct color to every new distinct item appearing in the sample. The diffuseness of ν makes sure that the colors are a.s. distinct. On the other hand, the function λ determines alone the probability that the next observation will be equal to the j-th distinct observed item or else that it will be of a new type. As shown by Pitman (1995), if $\mathbf{n}_n = (n_1, \ldots, n_k)$,

$$\lambda_j(\mathbf{n}_n) = \mathbb{P}(\mathbf{N}_{n+1} = \mathbf{n}_n + \mathbf{e}_j | \mathbf{N}_n = \mathbf{n}_n) = \frac{p_\lambda(\mathbf{n}_n + \mathbf{e}_j)}{p_\lambda(\mathbf{n}_n)}, \tag{8}$$

$(j = 1, \ldots, k+1)$ for some non-negative function p_λ, such that

$$\sum_{j=1}^{k+1} p_\lambda(\mathbf{n}_n + \mathbf{e}_j) = p_\lambda(\mathbf{n}_n)$$

where $\mathbf{e}_j = (\delta_{ij} : i = 1, \ldots, k+1)$ and δ_{ij} is the Kronecker delta. A de Finetti-style representation theorem for p_λ is proved in Pitman (1995). There are random a.s. limit relative frequencies

$$(P_1, P_2, \ldots) = \lim_{n \to \infty} n^{-1}\mathbf{N}_n$$

such that $\sum_j P_j \leq 1$, and their distribution μ_λ is such that

$$p_\lambda(\mathbf{n}_n) = \int \prod_{j=1}^{k} P_j^{n_j - 1}(1 - \sum_{i=1}^{j-1} P_i) d\mu_\lambda \tag{9}$$

The function p_λ is called the *partially exchangeable partition probability function (PEPF)* relative to λ (or to μ_λ).

The predictive distribution (7) converges in total variation, as n goes to infinity, to a random limit distribution of the form

$$F^{(\lambda,\nu)}(\cdot) = \sum_{j \geq 1} P_j \delta_{\bar{Y}_j}(\cdot) + \left(1 - \sum_{j \geq 1} P_j\right) \nu(\cdot) \tag{10}$$

where: $\mathbf{P} = (P_j : j \geq 1)$ has the distribution μ_λ, independent of the $\bar{Y}_j$'s, and the $\bar{Y}_j$'s are iid with common law ν (see Pitman (1996), prop. 14).

If a Pitman sequence is exchangeable, then by de Finetti's theorem it is iid $(F^{(\lambda,\nu)})$, conditional on $F^{(\lambda,\nu)}$. We call $F^{(\lambda,\nu)}$ a *Pitman random distribution with parameters* (λ, ν). In Pitman-Hansen (2000) a necessary and sufficient condition on λ to make a Pitman sequence exchangeable is that, for every $\mathbf{n}_n$, $p_\lambda(\mathbf{n}_n)$ is a symmetric function of its arguments and it is called *exchangeable partition probability function (EPPF)*. Examples of EPPFs can be found e.g. in Pitman (2002) and Gnedin-Pitman (2005).

We are now ready to extend Reinforced Urn Processes, by using Pitman's prediction rules instead of discrete Pólya urns; this will lead to a reinforced random walk on a continuous state space. Denote with $\mathcal{Z}$ the space of all possible parameters (λ, ν) for Pitman's sequences. The parameters of a Pitman RUP are 5:

1. A Polish state space $(S, \mathcal{S})$, endowed with its Borel σ-field.

2. A continuous spectrum of colours, represented by any Polish set $(E, \mathcal{E})$.

3. *A countable partition* $A = (A_0, A_1, ...)$ of S.

4. A function $U : \mathbb{N}_0 \to \mathcal{Z}$ which associates, to every class $A_i \in A$ $(i \geq 0)$, a pair $U_i = (\lambda^i, \nu^i)$ of Pitman's parameters. Thus $U = \{\lambda^i, \nu^i\}_{i\geq 0}$.

5. A law of motion $q : \mathbb{N}_0 \times E \to S$.

We assume that, for every $i = 0, 1, 2, ...$, $q_i(\cdot) = q(i, \cdot)$ is such that, if $S^{(i)} := q_i(E)$, then $q_i(\cdot) : E \to S^{(i)}$ is measurable and one-to-one and onto. Thus the induced measure $\tilde{\nu}^i := \nu^i \circ q_i^{-1}$ is also diffuse on S.

In the migrating particle example, any point $x \in S$ represents an exact location; any class $A_i \in A$ a distinct island. The process behaves like the discrete RUP; however we will need to introduce some constraints on the reinforcement device: for every $i \geq 0$ and $x, y, z \in S$, if $x, y \in A_i$, the difference between the two observed transitions $x \to z$ and $y \to z$ will be considered as irrelevant to predict future observations, because both starting points belong to the same "island". However, we will still distinguish between any two transitions with different arrival points (e.g. between $z \to x$ and $z \to y$).

The new random walk $X = \{X_n\}_{n\geq 1} \in S^\infty$ is defined recursively as follows: fix $X_0 = x_0 \in A_0$. For all $n \geq 1$, $i \geq 0$, if $X_n \in A_i$, then a ball is drawn from U_i according to the Pitman rule (λ^i, ν^i), and its color c shows where X will go next, according to $q_i(c)$. The ball can be of an old (i.e. already sampled from U_i) color, or a new distinct one, randomly chosen according to ν^i. For such a process we will write $X \in P - RUP(S, E, A, U, q)$.

In order to give a formal definition of P-RUPs, the notion of successor states will be redefined in terms of a different sequence of stopping times. Let

$$\tilde{\tau}_{j,1} = \inf\{n : X_n \in A_j\}$$

and for $m \geq 2$,

$$\tilde{\tau}_{j,m} = \inf\{n \geq \tilde{\tau}_{j,m-1} : X_n \in A_j\}.$$

In words, $\tilde{\tau}_{j,m}$ is the time of the $m-th$ entrance of X in A_j. Fix some $\eta \notin S$. For every $j = 0, 1, ...,$ define the new successor states as

$$V_{j,m} = X_{\tilde{\tau}_{j,m}+1} I(\tilde{\tau}_{j,m} < \infty) + \eta I(\tilde{\tau}_{j,m} = \infty).$$

For every n and i, let $\tilde{n}_j = \sum_{i=1}^{n} I(X_i \in A_j)$ and denote

$$T_n^j(X_{(n)}) = T_{\tilde{n}_j}(V_{j,\tilde{n}_j}) = \{K_n^j, L_n^j, N_n^j\} \tag{11}$$

with the function T_n defined as before.

Definition 3 *A random sequence $X = \in S^\infty$ is a Pitman-Reinforced Urn process (P-RUP) with parameters (S, A, E, U, q), starting from $X_0 \in A_0$ if, for every n and i*

$$\mathbb{P}(X_{n+1} \in B | x_1, \ldots, x_{n-1}, x_n \in A_i) = f^*\left(T_n^i(x_{(n)}), B\right), \tag{12}$$

where, if $T_n^i(x_n) = (k(i), \{\bar{v}_{i,1}, \ldots, \bar{v}_{i,k(i)}\}, \mathbf{n}_{\tilde{n}_i}\}$,

$$f^*(T_n^i(x_n), B) = \sum_{j=1}^{k^i} \lambda_j^i(\mathbf{n}_{\tilde{n}_i}) \delta_{\bar{v}_{i,j}}(B) + \left(1 - \sum_{j=1}^{k^i} \lambda_j^i(\mathbf{n}_{\tilde{n}_i})\right) \tilde{\nu}^i(B), \qquad B \in \mathcal{S}.$$

Remark 1. Due to the previous definition, we can study, without loss of generality, the properties of P-RUPs just by looking at the case when $E = S$ and q_i is $\forall i$ the identity map from S to itself. In this case, we will write that X is a $P - RUP\ (S, A, U)$.

Let $X \in P - RUP(S, A, U)$ such that, for every $(\lambda, \nu) \in U$, λ satisfies (8) for a p_λ symmetric in its coordinates. Then all sequences of "successor states" V_i are exchangeable, and therefore the process is Markov-exchangeable (Fortini *et al.* (2002)) so that, if the process satisfies a certain property of class-recurrence then its distribution is that of a Markov chain in S with a random transition matrix made of independent rows (see section 5).

4 The relation between Reinforced Urn Processes and Pitman's Reinforced Urn processes.

The original RUPs can be recovered by a discrete filtering of an appropriate P-RUP, where each urn involved is of a Blackwell-MacQueen type.

Proposition 4 *Assume that* X *is a P-RUP*(S, U, q)*, satisfying the following conditions:*

(i) $U = \{\lambda^i, \nu^i\}$ *is such that, for every* $i \geq 0$, $\lambda^i = \lambda^{\theta_i}$ *, where, if* $T_n^i(X_{(n)}) = (k, L_n^i, (n_{i1}, \ldots, n_{ik}))$*,*

$$\lambda_j^{\theta_i}(n_{i1}, \ldots, n_{ik}) = \frac{n_{ij}}{\tilde{n}_i + \theta_i} \qquad (j = 1, \ldots, k).$$

(ii) For every $i = 1, 2, \ldots,$ $\nu^i(A_j) > 0$ *for finitely many* j*'s.*

For a countable state space $Z = \{z_0, z_1, ...\}$ *define a measurable function* $f_A : S \to Z$ *such that, for every* $i \geq 0$*,*

$$f_A(x) = z_i \Longleftrightarrow x \in A_i \tag{13}$$

Then the process $Y = f_A(X)$ *is a discrete RUP* (Z, U_A)*, where, for each* $z_i \in Z$ $U_A(z_i) = \{\theta_i \nu^i \circ f_A^{-1}\{z\} : z \in Z\}$*.*

Proof. Let $\tilde{m}_i = \sup\{n : \tilde{\tau}_{i,n} < \infty\}$. Condition (i) implies that, for every i, the first $\tilde{m}_i$ coordinates of $V_i(X)$ follow a Blackwell-MacQueen prediction rule. Remember that, if W is a Blackwell-MacQueen sequence on $(S, \mathcal{S})$, a well known property is: for every $B \in \mathcal{S}$

$$\mathbb{P}(W_{n+1} \in B | W_n = w_n) = \phi(\sum_{m=1}^{n} I(w_m \in B)) \tag{14}$$

that is, the above probability depends only on how many observations already felt in B (see Zabell (1992)). Property (14) implies that, for every partition A of S, for every $C \in Z$

$$\mathbb{P}(f_A(W_{n+1}) \in C | W_{(n)} = w_{(n)}) = \mathbb{P}\left(f_A(W_{n+1}) \in C | f_A(W_{(n)}) = f_A(w_{(n)})\right)$$

and $f_A(W)$ is a discrete Pólya sequence on Z. As a consequence, $V_i^* := f_A(V_i)$ is the sequence of successor states of Y relative to z_i, so that Y satisfies the prediction rule (3) and this makes Y a $RUP(Z, \tilde{U}^*)$ where $\tilde{U}_z^*\{x\} = \theta_z \nu^z\{x\}$ for each $z, x \in Z$. ∎

Remark 2. The equality (4) actually characterizes Blackwell-MacQueen prediction rule among Pitman's sequences. For a discrete filtering of general P-RUP, where there is a parameter λ^i different from λ^{θ_i}, the information carried by the counts

$$\sum_{m=1}^{n} I\left(f_A\left(v_{i,m}\right) \in C\right), \qquad C \in \mathcal{S}$$

is no longer sufficient to predict $f_A(V_{i,n+1})$. In general, a discrete version of a Pitman sequence X would obey to a prediction rule of the form:

$$\mathbb{P}\left(f_A(X_{n+1}) \in B | f_A(X_{(n)}) = f_A(x_{(n)}), \gamma(T_n(x_{(n)}), B)\right) \tag{15}$$

for some additional measurable function γ.

For instance, if λ^i is given by Pitman's 2-parameter model

$$\lambda^i(n_1, \ldots, n_k) = \frac{n_i - \alpha k}{n + \theta} \qquad i = 1, \ldots, k, \tag{16}$$

for some $\theta > 0$ and some α satisfying either $0 \leq \alpha < 1$ or $\theta = -\alpha$, then it is easy to see (Spanò (2003)) that $\gamma\left(T_n\left(x_{(n)}\right), B\right)$ is the number of *distinct* values of $x_{(n)}$ appeared in B:

$$\gamma\left(T_n\left(x_{(n)}\right), B\right) = \sum_{j=1}^{K_n} I\left(\bar{x}_{,j} \in B\right).$$

It is still an interesting open problem how to classify Pitman's sequences (and hence also P-RUPs) according to the function $t\left(v_{i,n}, B\right)$ which is minimally sufficient to satisfy (15) .

5 Discovering new islands and new exact locations.

Pitman's sequences were entirely defined in terms of two conditional probabilities, given the frequencies of the past observed types ranked in their order of appearance: respectively, that the next observation is either of the j-th type, or of a new one. In the same spirit, we can describe the law of motion of a *P-RUP*, emphasizing its exclusive dependence on the islands and locations actually observed in the past, ranked in their order of appearance. To this extent we introduce a sort of island-analog of the triplet $T_n(X_{(n)})$. Note that all the results in this section do not need to assume exchangeability of the Pitman sequences involved, i.e. U may contain any choice of $\lambda^i : i = 1, 2, \ldots$.

Given A, for every sample $X_{(n)} \in S^n$ starting from $x_0 \in A_0$ let $H_n \leq K_n \leq n$ be the number of distinct islands visited by $X_{(n)}$, and let $C_n = (\bar{A}_1, ..., \bar{A}_{H_n})$ be the list of such islands in their order of appearance. A class-analog of $T_n(X_(n))$ is given by

$$R(X_{(n)}) = (H_n, C_n, \bar{N}_n,), \tag{17}$$

where $\bar{N}_n(X) = (\bar{N}_{n1}, .., \bar{N}_{nH_n})$, with

$$\bar{N}_{nj} = \sum_{i=1}^{n} I(X_i \in \bar{A}_j).$$

$j = 1, ..., H_n$. As $n \to \infty$, the random limits $C := \lim_n C_n$ and $H := \lim_n H_n(X) \leq \infty$ are well defined. Notice that the set $C = \bigcup_{j\geq 1}\{A_j : \tilde{\tau}_{j,1} < \infty\}$ may not cover the whole A. For the rest of this section, if not differently indicated, we will use ν^i, λ_i as the parameters associated to $\bar{A}_i$ instead of A_i, and consistently we will index V_i, T_n^i and so on.

5.1 The probability of migrating to new distinct locations.

The next proposition shows that in a P-RUP, after each visit to the the island A_i, the probability of next moving to the point $\bar{v}_{i,j}$ depends on the number of times the location $\bar{v}_{i,j}$ was "selected"in the past from the urn associated to i. Moreover, while in Pitman's sequences, each new distinct location was colored at random according to a diffuse measure ν, in a P-RUP there are countably many "paintbrushes"($\nu^i : i = 1, 2, \dots$) available. We see that the choice of which ν^i to use is determined by the last sample observation.

Proposition 5 *Let X be a P-RUP. Suppose that $X_{(n)} = x_{(n)}$ is such that $x_n \in \bar{A}_i$, $L_n(x_{(n)}) = (\bar{x}_1, ..., \bar{x}_k)$, $L_n^i(x_{(n)}) = L_{\tilde{n}_i}(v_i) = (\bar{v}_{i,j} : j = 1, .., k^i)$ and $N_n^i(x_{(n)}) = \mathbf{n}_n^i$. Then the probability of migrating to the l-th distinct location ($l = 1, ..., k$) is*

$$\mathbb{P}(X_{n+1} = \bar{x}_l | X_{(n)} = x_{(n)}) = \sum_{j=1}^{K_n^i} \lambda_j^i(\mathbf{n}_n^i)\delta_{\bar{v}_{i,j}}(\{\bar{x}_j\}). \tag{18}$$

The probability of discovering a new distinct location is given by

$$\begin{aligned}\mathbb{P}(K_{n+1} = k+1 | X_{(n)} = x_{(n)}) &= \mathbb{P}(X_{n+1} \notin L_n | X_{(n)} = x_{(n)}) \\ &= 1 - \sum_{j=1}^{K_n^i} \lambda_j^i(\mathbf{n}_n^i). \end{aligned} \tag{19}$$

The conditional distribution of the next value, given that it is a distinct value is

$$\mathbb{P}(X_{n+1} \in \cdot | K_{n+1} = K_n + 1, x_n \in \bar{A}_i) = \nu^i(\cdot). \tag{20}$$

Proof. The part (18) is a trivial consequence of the definition of P-RUP. Since ν_i is diffuse, then

$$0 = \nu^i(L_n(x_{(n)})) = \nu^i(L_{\tilde{n}_i}(v_i)). \tag{21}$$

Now $L_n^i \subset L_n$; then (18) follows, as well as

$$\sum_{j=1}^{k^i} \lambda_j^i(n_n^i)\delta_{\bar{v}_{i,j}}(L_n^c) = 0,$$

so that

$$\begin{aligned}
\mathbb{P}(X_{n+1} \notin L_n(X_n)|x_{(n)}) &= \sum_{j=1}^{k^i} \lambda_j^i(\mathbf{n}_n^i)\delta_{\bar{v}_{i,j}}(L_n^c) + \left(1 - \sum_{j=1}^{k^i} \lambda_j^i(\mathbf{n}_n^i)\right)\nu^i(L_n^c) \\
&= 0 + \left(1 - \sum_{j=1}^{k^i} \lambda_j^i(\mathbf{n}_n^i)\right)\nu^i(S) \\
&= \left(1 - \sum_{j=1}^{k^i} \lambda_j^i(\mathbf{n}_n^i)\right).
\end{aligned}$$

The third part follow just by considering that, for every $B \in \mathcal{S}, L_n \in S^{k^i}$,

$$\begin{aligned}
\mathbb{P}(X_{n+1} \in B \cap L_n^c|x_n) &= \left(1 - \sum_{j=1}^{k^i} \lambda_j^i(\mathbf{n}_n^i)\right)\nu^i(B \cap L_n^c) \\
&= \left(1 - \sum_{j=1}^{k^i} \lambda_j^i(\mathbf{n}_n^i)\right)\nu^i(B)
\end{aligned}$$

by diffuseness, and since $\{X_{n+1} \notin L_n\} = \{K_{n+1} = K_n + 1\}$, then

$$\mathbb{P}(X_{n+1} \in B|K_{n+1} = K_n + 1, x_n \in \bar{A}_i) = \nu^i(B).$$

■

5.2 The probability of new distinct islands.

A self-evident consequence of Proposition 5 is of some interest because it shows that, whilst the probability of visiting a new distinct *location* only depends on the parameter λ^i of the urn U_i, the probability of moving to a new distinct *island* depends also on the labeling measure ν_i. This makes the choice of the latter not so conventional as in the original Pitman sequences. Formally, let X be a $P - RUP(S, \tilde{U})$ starting from $x_0 \in A_0$ and let $L_n(x_{(n)}) = (\bar{x}_1, ..., \bar{x}_k)$ and

$C_n(x_{(n)}) = (\bar{A}_1, ..., \bar{A}_{H_n})$. The probability of discovering a new distinct island is

$$\mathbb{P}(X_{n+1} \notin C_n | X_{(n)} = x_{(n)}, x_n \in \bar{A}_i) = \left(1 - \sum_{j=1}^{k^i} \lambda_j^i(\mathbf{n}_n^i)\right) \nu^i(S \cup C_n^c)$$

The next corollary says that, conditionally on $X_{(n)}$, if we know that X_{n+1} is a point in a new island, then X_{n+2} will be a further new distinct location with probability one.

Corollary 6 *Let X be a $P - RUP(S, \tilde{U})$ and consider the number K_n of all distinct locations observed up to $X_{(n)}$. Then*

$$\mathbb{P}\{K_{n+2} = K_n + 2 | H_{n+1} = H_n + 1\} = 1$$

Proof. By construction,

$$K_n = |L_n| = |\bigcup_{j=1}^{H_n} L_n^i|, \tag{22}$$

where $|B|$ is the cardinality of B, Since $|L_n^i| \geq 1$, then

$$\{H_{n+1} = H_n + 1\} \subset \{K_{n+1} = K_n + 1\}$$

Now

$$\{H_{n+1} = H_n + 1\} = \{X_{n+1} \notin C_n\} = \{X_{n+2} = X_{\tau_{(H_n+1)}+1}\}$$

which implies that $L_{n+2} = L_{n+1} \bigcup \{X_{\tau_{(H_n+1)}+1}\}$, but by (12) and (13),

$$\mathbb{P}\{X_{\tau_{H_n+1}+1} \in \cdot | X_{\tau_{H_n+1}}\} = \nu^{H_n+1}(\cdot), \tag{23}$$

and since ν^{H_n+1} is diffuse, then

$$\nu^{H_n+1}(L_{\tau_{H_n+1}}) = 0$$

so by (22), if $X_{n+1} \notin C_n$

$$K_{n+2} = |L_{n+2}| = |L_{n+1} + 1| = |L_n + 2| = K_{n+2}.$$

■

Actually the statement of corollary 6 can be extended: from every two distinct islands, the particle will move almost surely to two distinct exact locations (although these might both be in the same island). Let X be a P-RUP.

Proposition 7 *For every $m, l \leq n$ suppose that $X_m = X_{\tilde{\tau}_{i,h_i}}$ and $X_l = X_{\tilde{\tau}_{j,h_j}}$ with $h_i \leq \tilde{n}_i$, $h_j \leq \tilde{n}_j$ and $i, j \leq H_n$. If $i \neq j$, then $X_{m+1} \neq X_{l+1}$ a.s.$\mathbb{P}$.*

The proof is an immediate consequence of the property of mutual singularity of any two Pitman random distributions, stated in the following lemma.

Lemma 8 *Let ν_1 and ν_2 be two diffuse measures on (S, S). Then two Pitman random distributions $F_{(\lambda,\nu_1)}$ and $F_{(\lambda,\nu_2)}$ put positive mass on two a.s. distinct distinct sets of points.of S.*

Proof. Let $Y_1, ..., Y_n$ be an independent S-valued sequence of random elements, where for some $m < n$, $Y_1, ..., Y_m$ are iid with common law ν_1, and $Y_{m+1}, ..., Y_n$ are iid according to ν_2, with both ν_1, ν_2 diffuse. Then by diffuseness $\mathbb{P}(Y_1 \neq ... \neq Y_n) = 1$. Consequently, if $X_1, ..., X_{2n}$ are independent, with $X_1, ..., X_n | F_{(\lambda,\nu_1)}$ iid with common law $F_{(\lambda,\nu_1)}$, and $X_{n+1}, ..., X_{2n} | F_{(\lambda,\nu_2)}$ iid with common law $F_{(\lambda,\nu_2)}$, then

$$\mathbb{P}\left\{L_n(X_1, ..., X_n) \neq L_n(X_{n+1}, ..., X_{2n})\right\} = 1.$$

If we denote

$$L := \lim_{n\to\infty} L_n(X_1, ..., X_n) = \{\bar{X}_1, \bar{X}_2, \ldots\}$$

and

$$L' := \lim_{n\to\infty} L_n(X_{n+1}, ..., X_{2n}) = \left\{\bar{Y}_1, \bar{Y}_2, ...\right\}.$$

it follows that, a.s. $\mathbb{P}$,

$$\sum_{j\geq 1} P_j \delta_{\bar{X}_j}(L') + \left(1 - \sum_{j\geq 1} P_j\right) \nu_1(L') = \sum_{j\geq 1} P_j \delta_{\bar{Y}_j}(L) + \left(1 - \sum_{j\geq 1} P_j\right) \nu_2(L) = 0,$$

and the proof is complete. ■

Remark 3. Lemma 8 is interesting also because it shows that the mutual singularity of Dirichlet processes is a property holding for the whole class of Pitman's random distributions. Notice that the distribution of sizes of the atoms, $(P_j)_{j\geq 1}$ can be the same for both $F_{(\lambda,\nu_1)}$ and $F_{(\lambda,\nu_2)}$.

6 Recurrence of the process.

In this section we discuss the role of two different notions of recurrence; the first one (class-recurrence) is needed for a de Finetti-style representation of P-RUPs as it ensures the convergence of an appropriate empirical measure to a random transition kernel, the second one (state-recurrence) can be used in applications.

6.1 The role of class-recurrence.

In the discrete case, strong recurrence was defined as the property that, if a state is visited once, then it will be visited infinitely often with probability one. We may call such property *state-recurrence.* For P-RUPs we need a weaker recurrence condition, which we call *class-recurrence*, namely: if a class is visited once by X, then it will be visited again infinitely often, almost surely.

Definition 9 *Let $A = (A_0, A_1, \ldots)$ be a countable partition of S. A random sequence $X \in S^\infty$ is class-recurrent with respect to A if, for every $i \geq 0$,*

$$\mathbb{P}\left(\bigcap_{m>1} \{\tilde{\tau}_{i,m} < \infty\} \,|\tilde{\tau}_{i,1} < \infty \right) = 1. \tag{24}$$

Remark 4 Of course, state recurrence implies class recurrence whereas the converse is not true. In general, if a sequence X is class recurrent with respect to A, then it is class recurrent with respect to any coarser partition $B \supseteq A$.

For the law of a P-RUP with exchangeable V- rows, class-recurrence is a sufficient condition in order to achieve a representation as mixture of Markov chain distributions, as shown in the next proposition. We extend the state space to $S^* = S \bigcup \{\eta\}$ $(\eta \notin S)$. For each sample realization $x_{(n)}$ consider the *empirical transition matrix* $m(x_{(n)}) = \{m_n(i, \cdot) : i = 1, 2, \ldots\}$ given by

$$m_n(i, B) = \frac{1}{\tilde{n}_i} \sum_{j=1}^{\tilde{n}_i} \delta_{V_{i,j}}(B), \qquad i = 1, 2, \ldots; , B \in \mathcal{S}. \tag{25}$$

Proposition 10 *Let X be a P-RUP (S^*, A, U), starting from $x_0 \in A_0$, and let $C(X) = (A_{i_1}, A_{i_2}, \ldots) \subset A$ be the sequence of classes of A visited by X in order of appearance. Assume that $U = \{\lambda_i, \nu^i\}_{i \geq 0}$ is such that, for every i, λ_i satisfies (8) for a symmetric function g_i. If X is class-recurrent, then*

$$m_n(X) \underset{n \to \infty}{\overset{\mathcal{D}}{\to}} \pi = \{\pi(i, \cdot) : i \geq 0\}$$

where π is a random atomic kernel on $\mathbb{N} \times \mathcal{S}$ such that:

(i) $\pi(i, \cdot)$ is independent of $\pi(j, \cdot)$ for $i \neq j$;

(ii) for every $j \in I_C$

$$\pi(j, \cdot) \overset{d}{=} \bar{F}^j(\cdot) \tag{26}$$

where $\bar{F}^j$ is a Pitman random distribution, with parameters $(\lambda^{i_j}, \nu^{i_j})$;

(iii) for every $j \notin I_C$,

$$\pi(j, \cdot) = \delta_\eta.$$

Moreover, $X|(\pi, C)$ is a Markov chain on S with transition atomic kernel π that is, for every n:

$$\mathbb{P}(X_{n+1} \in \cdot | X_{(n)} = x_{(n)}, x_n \in \bar{A}_j; \pi) = \pi(j, \cdot). \quad (27)$$

Proof. Under recurrence, for every $i \neq j$ the sequences $\bar{V}_i$ and $\bar{V}_j$ are independent, infinite exchangeable sequences. In fact, for every n and i, the value of $V_{i,\tilde{n}_i+1}$ depends on the set $\{T_n^j(X_{(n)}) : j \neq i\}$, only through the event $\{\tilde{\tau}_{i,\tilde{n}_i} < \infty\}$ which, given C, has probability $\mathbf{I}(i \in I_C)$. Moreover, because of the assumption on U, the empirical distribution of every $\bar{V}_j$ converges a.s. to a Pitman random distribution $\bar{F}^i$. Together with the independence, this proves (26). The rest of the proof (Markov property) is a consequence of Theorem 4 in Fortini *et al.* (2002). ■

Let $X \in P - RUP(S, A, \tilde{U})$ starting from $X_0 \in A_0$ and for each n, consider $C_n(X_{(n)}) = (A_{i_1}, \ldots, A_{i_{H_n}})$. If for every n

$$\nu^{i_n}(A_0) > 0$$

then X is obviously class-recurrent. Class-recurrence of a P-RUP can be characterized by means of the measures $\nu^i, i = 0, 1, \ldots$. This is done in the next Proposition, which show once again that the choice of the diffuse measures $(\nu^i)_{i \geq 0}$ is not immaterial in that it is the only parameter that determines the convergence of the process. For every ordered collection $C \subseteq A$ of subsets of A, define $I_C = (i_1, i_2, \ldots) \subseteq \mathbb{N}$ if $C = (A_{i_1}, A_{i_2}, \ldots)$

Proposition 11 *Let $X \in P - RUP(S, A, U)$ start from A_0. X is class-recurrent if and only if, almost surely, there exists a set $C \subseteq A$ such that $A_0 \in C$ and ,*

$$\begin{array}{ll} (i)\ \forall i \in I_C \backslash \{0\}, & \nu^i(A_i) < 1; \\ (ii)\ \forall i \in I_C, & \nu^i(C) = 1; \\ (iii)\ \forall i \in I_C \backslash \{0\}, & \exists j_i \in I_C,\ j_i \neq i,\ s.t.\ \nu^{J_i}(A_i) > 0. \end{array}$$

The key point for the proof is in the following Lemma.

Lemma 12 *Let X be a Pitman sequence with parameters (λ, ν). For every set $B \in \mathcal{S}$ such that $\nu(B) > 0$, then $X_n \in B$ for infinitely many n.*

Proof. Assume $\nu(B) > 0$ and define for every n, $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ (with $\mathcal{F}_0 = \{\emptyset, S\}$) and $E_n = \{X_n \in B\}$. By (Levy's extension of) Borel-Cantelli Lemma, E_n is true infinitely often if and only if

$$\sum_{n \geq 1} E(I_{E_n} | \mathcal{F}_{n-1}) = \infty.$$

But in our case,

$$E(I_{E_n}|\mathcal{F}_{n-1}) \to \sum_j P_j \delta_{Y_j}(B) + (1 - \sum_j P_j)\nu(B) > 0, \ n \to \infty$$

thus the series diverges and the proof is complete. ■

Now we can prove the proposition.

Proof. If X is a class-recurrent P-RUP, consider $C(X) = \lim_n C_n(X_{(n)})$. Then $A_0 \in C(X)$ by construction. Now, if $\nu^0(A_0) = 1$, X will never exit from A_0, i.e. $\mathbb{P}(C(X) = A_0) = 1$ thus the proposition is true for the degenerate case $I_C = 0$. If $\nu^0(A_0) < 1$ then by Lemma 12 $C(X)$ contains a.s. at least another set A_l, say. If $\nu_l(A_l) = 1$, then, once entered, X will never exit from A_l, but this imply that A_0 is not visited infinitely often, which is a contradiction. Thus (i) is proved. (ii) follows by the way $C(X)$ is defined: if $\nu^i(C(X)) < 1$, then a.s. X_n will enter in some $A_l \notin C(X)$ which is a contradiction. Part (iii) trivially applies to $C = C(X)$ as, if $\nu^{j_i}(A_i) = 0$ for all $j_i \in I_C(X)$, then $\tilde{\tau}_{i,1} = \infty$ but that contradict that $i \in I_C(X)$.

Conversely, assume $(i) - (iii)$ for some set C including A_0. (iii) implies that C is a connected set i.e. there is a path between any two visited islands in C, that X will perform with positive probability. Lemma 12 implies that, thanks to (i) and, every such path will be performed infinitely often and, by (ii), X will never exit from C. ■

6.2 A representation for class-recurrent Pitman-reinforced urn processes.

When in a P-RUP not all the V-sequences are exchangeable, the process is no longer mixture of markov chains; nonetheless a weaker representation theorem can be obtained.

6.2.1 Limit distribution for Pitman sequences

A partial exchangeable random partition of $\mathbb{N}$ can be viewed as the partition induced on $\mathbb{N}$ by a Pitman sequence $X = (X_1, X_2, \ldots)$ via the equivalence rule: $i \sim j \leftrightarrow X_i = X_j$, with the equivalence classes ranked by their least elements. In Pitman (1995) a representation theorem is proved for partially exchangeable random partitions. Such a representation theorem can be trivially rephrased in terms of Pitman sequences on a general state space S and so we state it without proof: the only difference is that now (7) implies that iid(ν) random positions are attached to distinct atoms.

Proposition 13 *(Pitman (1995, th. 6)). Let X be a sequence in S with pa-*

rameter (λ, ν)*, and let* $X_{(n)}$ *be the restriction of* $X_{(n)}$ *to its first* n *coordinates. Denote with* $K_n = K(X_{(n)})$ *the number of distinct values observed in* $X_{(n)}$*, and with* $L_n = (\bar{X}_1, \ldots, \bar{X}_k)$ *their locations. The following conditions are equivalent:*

(i) X *is a Pitman sequence;*

(ii) There exists a sequence of random variables $P = (P_1, P_2, \ldots)$ *such that* $\sum_j P_j \leq 1$ *and a sequence of iid*(ν) *random variables* $Y = (Y_1, Y_2, \ldots)$*, independent of* P*, such that, conditional on* $K_\infty = k^* \leq \infty$*, for every* $0 \leq n < \infty$*,* $k \leq n$*, and* $l_n = \bar{x}_1, \ldots, \bar{x}_k$

$$\mathbb{P}(X_{n+1} \in B \bigcap L_n^c | K_n = k, L_n = l_n; P, Y) = (1 - \sum_{j=1}^{k} P_j)\delta_{Y_{k+1}}(B) \qquad B \in \mathcal{S}; \tag{28}$$

$$\mathbb{P}(X_{n+1} \in B \bigcap L_n | K_n = k, L_n = l_n; P) = \sum_{j=1}^{k} P_j \delta_{\bar{x}_1}(B), \qquad B \in \mathcal{S} \tag{29}$$

almost surely.

Let $W_{(n)} = W(X_{(n)}) = (1 = W_1 < \ldots < W_{K_n} \leq n)$ be the random positions in $1, \ldots, n$ at which $(K(X_{(i)}) : i = 1, \ldots, n)$ jumps. Then $W_{(n)}$ is the collection of the least elements of the partition $\Pi_n = (\Pi_{n1}, \ldots, \Pi_{nK_n})$ induced by $X_{(n)}$ on $\{1, \ldots, n\}$. The likelihood of the first n observations of a Pitman sequence, conditional on the limit (P, Y) can be derived by a repeated application of proposition 13. For every $x_{(n)}$ such that $K(x_{(n)}) = k, W(x_{(n)}) = w_1, \ldots, w_k$,

$$\begin{aligned}\mathbb{P}\left(X_{(n)} \in dx_{(n)} | P, Y\right) &= \prod_{j=1}^{k} (1 - S_{j-1}) \delta_{Y_j}(dx_{w_j}) \prod_{i_j = w_j + 1}^{w_{j+1}-1} \left(\sum_{m=1}^{j} P_m \delta_{x_{wm}}(dx_{i_j}) \right) \\ &= \prod_{j=1}^{k} (1 - S_{j-1}) \delta_{Y_j}(dx_{w_j}) \prod_{i \in \pi_{nj}} P_j \delta_{dx_{w_j}}(dx_i) \end{aligned} \tag{30}$$

Averaging with respect to the Y's, from (30)

$$\mathbb{P}(\{\Pi_n = \pi_{n1}, \ldots, \pi_{nk}\} | P) = \prod_{j=1}^{k} P_j^{|\pi_{nj}|-1} (1 - S_{j-1}).$$

which shows the connection with the characterization of PEPFs given by (9): From Proposition 13, the unconditional probability of predicting a type already sampled is:

$$\mathbb{P}(X_{n+1} \in B \bigcap L_n | K_n = k; P, Y) = \sum_{j=1}^{k} P_j \delta_{Y_j}(B) \qquad j = 1, \ldots, k; \qquad (31)$$

Thus, for each collection of Borel sets $A_i \in \mathcal{S}$, $(i = 1, 2, \ldots)$,

$$\mathbb{P}\left(X_{W_k+1} \in A_1, \ldots, X_{W_{k+1}-1} \in A_{W_{k+1}-W_k-2} | P, Y\right) = \prod_{m=1}^{W_{k+1}-W_k-2} F_{(k)}(A_m), \qquad (32)$$

where

$$F_{(k)}(\ \cdot\ ) = \sum_{j=1}^{k} P_j \delta_{Y_j}(\ \cdot\ ), \qquad (33)$$

that is: between any two consecutive jumps in K_n the observations in a Pitman sequence are exchangeable. Notice that the property (32) is consistent with Pitman (1995)'s Lemma 14 on partially exchangeable random partitions.

6.2.2 Mixtures of "Partially Markov"Processes.

From proposition 13 we are now able to give a representation for general class-recurrent P-RUPs. Let X be a sequence in $S^* = S \bigcup \{\eta\}$ and let $A = (A_m : m = 0, 1, \ldots)$ be a countable partition of S. For every n and $i = 0, 1, 2, \ldots$, consider the empirical measure $Q_n = \{Q_n(i, \cdot) : i \in I_{C_n}\}$ given by

$$Q_n(i, B) = T_n^i(X_{(n)}) = \{K_n^i, L_n^i, N_n^j\}.$$

Proposition 14 *The following conditions are equivalent:*

(i) X is a class-recurrent $P - RUP(S^*, A, U)$ starting from A_0;

(ii) Let $\psi_n^* = (L_n^i, K_n^i : i \in I_{C_n})$ and set $\psi^* := \lim_{n\to\infty} \psi_n^*$. There exists a set $C \subseteq A$ satisfying (i)-(iii) of proposition 11 and an a.s. unique independent sequence $Q = \{Q_i : i \in I_C\}$ of random triplets $Q_i = (K^i, P^i, Y^i)$, such that, for each $i \in I_C$, conditional on K^i, $(K^i \leq \infty)$

(a) $P^i = (P_1^i, P_2^i, \ldots) \in [0, 1]^\infty$ is independent of Y^i, with distribution μ_{λ^i} such that $\mu_{\lambda^i}(\sum_{j=1} P_j^i \leq 1) = 1$;

(b) $Y^i = (Y_1^i, Y_2^i, \ldots) \in S^\infty$ is a sequence of iid(ν^i) random variables;

(c) Conditional on Q and on ψ_n^*, for $n = 0, 1, \ldots$ the law of X is described by:

$$\mathbb{P}(X_{n+1} \in \cdot | X_{(n)} = x_{(n)}; Q) = \chi_Q(x_n, \psi_{(n)}^*; \cdot) \qquad (34)$$

where

$$\chi_Q(x_n, \psi^*_{(n)}; \cdot) = \sum_{i \in C} \delta_{x_n}(A_i) \sum_{m=1}^{k_i^*} \mathbf{I}(k_n^i = m) \left[F^i_{(m)}(l^i_n; \cdot) + \left(1 - \sum_{j=1}^{m} P^i_j\right) \delta_{Y^i_{m+1}}(\cdot) \right]$$
$$+ \delta_\eta \sum_{i \notin I_C} \delta_{x_n}(A_i) \qquad (35)$$

where

$$F^i_{(m)}(l^i_n; \cdot) = \sum_{j=1}^{m} P^i_j \delta_{\bar{v}_{i,j}}(\cdot).$$

(d) $Q_n \xrightarrow{d} Q$ *a.s.* $(\mathbb{P})$.

Proof. The proof is similar to Fortini *et al.* (2002)'s proof of theorem 4 for markov-exchangeable sequences. Since for each pair l, m, $V_{l,m} = X_{\tilde{\tau}_{l,m}+1}$ for some integer $\tilde{\tau}_{l,m}$, then the event $\{X_1 \in dx_1, \ldots, X_n \in dx_n\}$ can be expressed as $\bigcap_{l_1}^{H_n} \bigcap_{m=1}^{\tilde{n}_l} \{V_{l,m} \in dx_{\phi(l,m)}\}$ for some $\phi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. We can rewrite

$$\bigcap_{m=1}^{\tilde{n}_l} \{V_{l,m} \in dx_{\phi(l,m)}\} = \{\Pi(V_{l,(\tilde{n}_l)}) = \pi^{(l)}_{\tilde{n}_l,1}, \ldots \pi^{(l)}_{\tilde{n}_l,k^l_{\tilde{n}_l}}\} \bigcap \{L^l_n = \{\bar{v}_{l,1}, \ldots, \bar{v}_{l,k^l,\tilde{n}_l}\}\} \qquad (36)$$

where $\bar{v}_{l,j}$ is the j-th distinct value appearing in $x_{\phi(l,1)}, \ldots, x_{\phi(l,\tilde{n}_l)}$. To see that (i) implies (ii), consider that, by class-recurrence, for every $i \in I_C$ V_i is an infinite sequence so that, by Proposition 13 there is a.s. a triplet pair (K^i, P^i, Y^i), with P^i independent of Y^i (given K^i) and Y^i $iid(\nu^i)$. Let Q be the collection of such triplets. Since P_i are the limit frequencies of V_i in order of appearance, then (d) is true. Given Q, the event (36) has probability

$$\prod_{j=1}^{k} (1 - S^{(l)}_{j-1}) \delta_{\bar{V}_{u,j}}(dv_{w_{l,j}}) \prod_{m \in \pi^{(l)}_{nj}} P^{(l)}_j \delta_{dv_{l,w_j}}(dv_{l,m}),$$

which implies that: (i), given Q V_i is independent of V_j for every $i, j \in I_C : i \neq j$ and therefore *(d)* is true; (ii) by proposition 13, the predictive distribution of X is just (34)-(35).

The converse is straightforward as taking the expectation over Q of a process satisfying (34)-(35) returns (12). ■

From (31) we know that a Pitman sequence is not exchangeable apart from between successive occurrence of new distinct values. Similarly, a P-RUP is not a mixture of Markov chain distributions, apart from those periods while no new locations are discovered by the process.

Corollary 15 *Let X be a class-recurrent P-RUP(S^*, A, U). Let E_{mn} denote the event "no new locations are visited between X_m and X_n, $(n < m)$". Conditional on Q and E_{mn}, $X_{n+1}, \ldots, X_{m-1}$ is a Markov chain with transition* *probability given by*

$$\chi_Q(x,\psi^*_{(n)};\cdot|E_{mn}) = \sum_{i\in I_C}\sum_{j=1}^{k_i^*}\delta_x(A_i)\mathbf{I}(k^i_m=j)F^i_{(j)}(\cdot)+\delta_\eta\sum_{i\notin I_C}\delta_{x_n}(A_i). \quad (37)$$

where $F^i_{(j)}(\cdot)=\sum_{l=1}^{j}P^i_l\delta_{Y^i_l}(\cdot)$.

The proof follows immediately from proposition 14 and from (32)

6.3 The role of state-recurrence.

Although only class-recurrence is necessary for the law of a P-RUP to be represented as a mixture of Markov chain distributions, in some applications one may wonder when not only the islands, but also the visited exact locations are strongly recurrent (state-recurrence). We want to provide necessary and sufficient conditions for state-recurrence of a P-RUP. When a P-RUP is Markov-exchangeable, a characterization is possible from a property of exchangeable Pitman sequences (Pitman (1996)). Let Y be an exchangeable Pitman sequence with parameters (λ,ν). Define the *size-biased pick* of Y as

$$\bar{P}_1(Y)=\lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}I(Y_n=\bar{Y}_1).$$

If λ is such that

$$\mu_\lambda(P_1(Y)>0)=1, \quad (38)$$

then the predictive distribution of Y converges, in total variation, to an a.s. discrete distribution:

$$F^{(\lambda,\nu)}(\cdot)=\sum_{j\geq 1}\bar{P}_j\delta_{\bar{Y}_j}(\cdot)$$

where $(\bar{P}_j)_{j\geq 1}$ and $(\bar{Y}_j)_{j\geq 1}$ satisfy the same properties $F^{(\lambda,\nu)}$ as in (10). In other words, (38) implies that

$$(1-\sum_{j\geq 1}\bar{P}_j)=0$$

almost surely, which means that every point of S, if sampled once, is sampled infinitely often. The application of such a property to P-RUPa is immediate:

Proposition 16 *A* $P-RUP(S,\tilde{U})$ X *is state-recurrent if and only if it is class-recurrent and the sequence of successor states* $V_i=(V_{i,1},V_{i,2},\ldots)$ *has a.s. positive size-biased pick for every* $i\geq 0$ *such that* $\tilde{\tau}_{i,1}>0$.

Proof. If X is a state-recurrent P-RUP then by definition, for each $i \in I_C$, the state $\bar{V}_{i,1}$ (which has distribution ν^i) will occur infinitely often, and therefore $P_1^i > 0$ almost surely. The converse is trivial because, if Q is such that $P_1^i > 0$ for each $i \in I_{C(X)}$ almost surely, then V_i is, conditionally on Q an $iid(F^I)$ sequence where $F^i = \sum_j P_j^i \delta_{Y_j^i}$ and, given Y^i, each state Y_j will be visited infinitely often. ■

Remark 5. From the distinction between class-recurrence and state-recurrence, we may derive a refined version of P-RUP under some additional constraints on the parameters of the urns. In fact, we have seen in the previous sections that the limit transition kernel π is a function of the set of the islands actually observed by a $P - RUP$ X. Consider a subset A_0 of the state space S and denote $S^0 = S \setminus A_0$. To A_0 we associate as usual a Pitman urn U_0 with parameters (λ_0, ν_0). To every single point $x \in S^0$, we now associate a Pitman urn U_x with parameters λ_x, ν_x such that: (i) (38) is satisfied for every x and (ii) $\nu^x(A_0) > 0$ for every $x \in S^0$. Define the times $\tau_{x,j}$ and the successor states $V_{x,j}$ as in (1) and (2). Define a process X such that $X_0 \in A_0$ and, for every $n > 1$,

$$\mathbb{P}(X_{n+1} \in B | X_1 = x_1, \ldots, X_{n-1} = z) = f^*(T_n^z(x_{(n)}), B) \tag{39}$$

where f^* is defined as in (12) but with the new times $\tau_{x,j}$ and successor states $V_{x,j}$.

Such a process behaves exactly like a P-RUP but it identifies every new distinct point as a new distinct island, so it is a P-RUP on S where $A = A_0 \cup \bigcup_{x \in S} x \notin A_0$ is now uncountable, however the process is state-recurrent and therefore, with probability 1, each sequence V_i will select only at most countably many points. Thus $C = C(X)$ will be still countable with probability 1.

7 Considerations on the support.

A recurrent P-RUP is conditionally distributed as a Markov chain, given the limit random transition kernel $\tilde{\pi}$ made of independent Pitman's distributions. We may be interested in specifying which possible transition kernels on S^*, are possible values of $\tilde{\pi}$. This is easily determined once we know the support of a Pitman distribution. In fact, we saw that, under recurrence, the rows of $\tilde{\pi}$ are Pitman independent random distributions. In this section we show that the support of a Pitman random distribution is determined similarly to the Dirichlet case, a result which is of independent interest also from a traditional (i.e. exchangeable) nonparametric Bayesian perspective.

We recall that the *support* of any probability measure Q defined on some space $(A, \mathcal{A})$ is the smallest closed set $J_Q \subset \mathcal{A}$ such that $Q\left(J_Q^c\right) = 0$. We are interested in studying the *weak support*, then here we use the assumption that S is a Polish space. In fact, when S is Polish, the space $\mathbb{M}_S$ of all probability measures defined on S, is metrizable under the topology of the weak convergence. Under such topology, let $\mathcal{M}_S$ be the Borel σ- field generated by $\mathbb{M}_S$. Now every Pitman's random distributions $F_{(\lambda,\nu)}$ will be viewed as a random element $F_{(\lambda,\nu)} : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{M}_S, \mathcal{M}_S)$. We will call its probability distribution, $\pi_{(\lambda,\nu)} := \mathbb{P} \circ F_{(\lambda,\nu)}^{-1}$, the *Pitman prior induced by* (λ, ν).

Proposition 17 *Let F be a random distribution with a Pitman prior $\pi_{(\lambda,\nu)}$. Denote with $\bar{\mu}_{\lambda,n}$ the joint distribution of the first n coordinates $\left(\bar{P}_1, ..., \bar{P}_n\right)$ of $\bar{P}$ $(n \geq 1)$. Assume that $\bar{\mu}_{\lambda,n}$ is continuous on the set*

$$\bar{\Delta}_{(n)} := \left\{x \in \bar{\Delta} : \forall j > n, x_j = 0\right\}$$

Then the support of $\pi_{(\lambda,\nu)}$ is given by

$$J_{(\lambda,\nu)}^* := \{P \in (\mathbb{M}_S, \mathcal{M}_S) : J_P \subset J_\nu\}.$$

Proof. Let $\hat{\partial}$ and $\mathcal{S}^\infty$ be the Borel σ -field (under the topology of the weak convergence) of $\bar{\Delta}$ and S^∞, respectively. Since, for every $A \times B \in \hat{\partial} \times \mathcal{S}^\infty$, $\pi_{(\lambda,\nu)}(A \times B) = 0$ if $B \cap S_\nu^\infty = \emptyset$, i.e. if $\bar{y} = (\bar{y}_1, \bar{y}_2, ...) \in B$, then for every $j \geq 1$

$$y_j \notin J_\nu.$$

Thus

$$\pi_{(\lambda,\nu)}\left(F \in J_{(\lambda,\nu)}^*\right) = 1.$$

Now, for some integer m, consider a collection $(A_1, ..., A_m)$ of m disjoint open sets of $\mathbb{R}$, and define $B_j := A_j \cap J_\nu$ $(j = 1, ..., m)$. Consider the subset $U \subset \{1, ..., m\}$ given by $U = \{j \leq m : B_j \neq \emptyset\}$. Then $U = \{i_1, ..., i_{k_U}\}$ for some $k_U \leq m$. We denote $\bar{B}_l = B_{i_l}$, for $l = 1, ..., k_U$.

We want to prove that, for every $\delta > 0$,

$$\pi_{(\lambda,\nu)}\left(\left|F\left(\bar{B}_j\right) - P_0\left(\bar{B}_j\right)\right| < \delta : j = 1, ..., k_U\right) > 0, \tag{40}$$

that is, that $\pi_{(\lambda,\nu)}$ actually gives a positive mass in any open neighborhood of P_0.

To prove it, first notice that, for every $j = 1, ..., k_U$, and $n \geq 1$

$$\nu(Y_n \in B_j) > 0,$$

and also

$$\nu\left(Y_n \notin B_j\right) > 0.$$

so that, for every $l \leq k < \infty$,

$$\nu\left(\bigcap_{n=1}^{l}\left\{Y_n \in B_j\right\}, \bigcap_{n=l+1}^{k}\left\{Y_n \notin B_j\right\}\right) = \prod_{n=1}^{l}\nu\left(Y_n \in B_j\right)\prod_{n=l+1}^{k}\nu\left(Y_n \notin B_j\right) > 0. \tag{41}$$

Now, by assumption, for every $n \geq 1$, $\sum_{i=1}^{n}\bar{P}_i$ is a continuous random variable in $[0, 1]$. Therefore, for every $\varepsilon, \delta, \eta \in (0, 1)$, such that $\varepsilon + \delta < 1$, there exists a $h, 1 \leq h < \infty$ such that

$$\bar{\mu}_\lambda\left(\sum_{i>h}\bar{P}_i < 1 - \varepsilon \pm \delta\right) > \eta$$

which implies that,

$$\bar{\mu}_\lambda\left(\sum_{i\leq h}\bar{P}_i < \varepsilon \pm \delta\right) > \eta$$

Consequently, by independence of atoms and locations, for every $\varepsilon > 0$,

$$\bar{\mu}_\lambda\left(\sum_{i\leq l}P_i + \sum_{i\geq h}P_i < \varepsilon \pm \delta\right) > 0$$

and this, together with, with $k = h$, implies that

$$\pi_{(\lambda,\nu)}\left(F\left(\bar{B}_j\right) < \varepsilon \pm \delta\right) > 0.$$

Thus, setting $\varepsilon = P_0\left(B_j\right)$,we obtain

$$\pi_{(\lambda,\nu)}\left(\left|F\left(\bar{B}_j\right) - P_0\left(B_j\right)\right| < \varepsilon\right) > 0.$$

Analog reasoning apply for the joint distribution of $\left(F\left(\bar{B}_i\right) : i = 1, ..., m\right)$: since we assumed continuity of $\bar{\mu}_{\lambda,n}$ on its support, then for every $\varepsilon_1, ..., \varepsilon_m$

$$\bar{\mu}_\lambda\left(\sum_{i\leq l}P_i\delta_{Y_i}\left(B_1\right) < \varepsilon_1 \pm \delta, \sum_{i\leq l}P_i\delta_{Y_i}\left(B_2\right) < \varepsilon_2 \pm \delta, ..., \sum_{i\leq l}P_i\delta_{Y_i}\left(B_m\right) < \varepsilon_m \pm \delta\right) > 0$$

then (40) holds true, and the theorem is proved. ■

The next corollary tells us how the support of a P-RUP looks like.

Corollary 18 *Let A be a countable partition of S, and set $A^*A\cup\{\eta\}$. Let $\mathcal{K}$ be the space of all atomic kernels on S, indexed by the classes of A. Endow*

$\mathcal{K}$ with the Borel σ-field κ, under the topology of the coordinate-wise weak convergence.

Let $\pi : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathcal{K}, \kappa)$ be the limit kernel of a class-recurrent P-RUP, and $\phi = \mathbb{P} \circ \pi^{-1}$ its distribution. The support of ϕ is given by the set of all $\pi = \{\pi(i, \cdot)\}_{i \in \mathbb{N}}$ for which there exists a subset $B \subset \mathbb{N}$ for which:

$$J_{\pi(i,\cdot)} \subset J_{\nu_i} \subseteq S; \tag{42}$$

$\forall i \in B$;

$$J_{\pi(j,\cdot)} = \delta_\eta \tag{43}$$

$\forall\, j \notin B$; and

$$\prod_{i \in B} \nu_i \left(\bigcup_{j \neq i} J_{\nu_j} \right) > 0; \tag{44}$$

$$\nu_i \left(\bigcup_{j \in B^c} J_{\nu_j} \right) = 0 \tag{45}$$

Proof. The support of every random row of π is the support of a Pitman prior, thus proposition 17 implies (42) for all recurrent classes. All non-recurrent classes are trivial sequences of $\eta' s$ which lead to (43). The last condition (44) is necessary for B to be actually the index set of all and only the recurrent classes. ■

8 Prior measures derived from Pitman-Reinforced Urn process.

Discrete Reinforced Urn Processes are used in Bayesian Nonparametrics to generate and characterize known classes of prior measures, such as the Ferguson-Dirichlet process, the Beta-Stacy process and, in general all Neutral to the Right (NTR) processes. We can use a discrete filtering of a P-RUP to construct particular instances of NTR processes.

Assume X is a class-recurrent P-RUP (S, A, U) starting from A_0, with all $(\lambda^i, \nu^i) \in U$ satisfying, (8) for some function g_i (respectively) symmetric on its arguments, and such that

$$\tilde{\nu}^0 (A_1) = 1$$

and, for every $i > 0$

$$\nu^i (A_{i+1}) = 1 - \tilde{\nu}^i (A_0) \in (0, 1)\,.$$

Consider a discrete filtering f_A as seen in section 4, such that $f_A : S \to \mathbb{N}_0$ and $f_A(A_i) = i$. Then the discrete process

$$Y = f_A(X)$$

is a reinforced random walk which, from every integer i, can only reach either $i+1$ or 0. Now, consider the sequence of "waiting times"spent by X between the $(j-1)$-th and the j-th entrance in A_0:

$$T = \{T_j = (\tilde{\tau}_{0,j} - \tilde{\tau}_{0,j-1}) : j = 1, 2, ...\},$$

where $\tilde{\tau}_{0,0} := 0$.

Being each sequence $V_{j,n}$ exchangeable in all its coordinates until η appears, the same arguments as in Muliere *et al.* (2000) (sec. 4) prove that the sequence T is exchangeable, and its directing measure is a NTR process. The particular type of NTR can be specified just by choosing the set of parameters $U = \{\lambda^i, \nu^i\}$ in the underlying P-RUP. For example, the following instances of priors can be derived:

(i) If we set $\lambda^i = \lambda^\theta$ as in Proposition 4, then T is a Beta-Stacy process on $\mathbb{N}_0$, with parameters $\{\theta^i\nu_i(A_0), \theta^i\nu_i(A_{i+1})\}_{i\geq 0}$.

(ii) If we assume (ii) and, for every i, we set $\{\theta^i\nu_i(A_0), \theta^i\nu_i(A_{i+1})\}_{i\geq 0} = \{\theta^1\nu_1(A_0), a + i\theta^1\nu_1(A_1)\}$, for some $a > 0$, then T is a two-parameter GEM with parameters $(\theta^1\nu_1(A_0), a)$ such that either $0 \leq 1 - \theta^1\nu_1(A_0) < 1$ and $a + \theta^1\nu_1(A_0) > 1$, or $1 < \theta^1\nu_1(A_0)$ and $a = -m(1 - \theta^1\nu_1(A_0))$ for some integer m.

(iii) If, in particular, we let $\theta^i = \theta > 0 \; \forall i$, and

$$\nu_i(A_{i+1}) = a + b[\nu_{i-1}(A_i)]$$

for an appropriate choice of $a, b \in \mathbb{R}$, then T turns out to be a GEM with parameter $\theta^0\nu_0$.

It would be interesting to understand the behavior of T when the V_i-sequences are not necessarily exchangeable. In this case, from Corollary 15, the P-RUP is Markov-exchangeable only between jumps in K_n. Therefore, for $k \in \mathbb{N}$ and $j = 1, 2, \ldots$ define

$$T_j^{(k)} = (\tau_{0,j} - \tau_{0,j-1})\mathbf{I}(K_{\tau_{0,j}} = K_{\tau_{0,j-1}} = k).$$

Then, each sequence

$$T^{(k)} = \{T_j^{(k)}\mathbf{I}(T_j^{(k)} > 0), j = 1, 2, \ldots\}$$

is *finitely exchangeable* (unless there exists some m such that $K_n = k$ for every $n \geq m$ almost surely), hence the array

$$T^* = \left(T^{(k)} : k \geq 0\right)$$

is (finitely) partially exchangeable, in the meaning of de Finetti (1938). For finitely exchangeable and partially sequences unfortunately there is not an equivalent, de Finetti-style representation theorem. Some approximations can nonetheless be found in Aldous (1985) and Eaton (1989). One should also note that T^* is a smaller collection than the exchangeable T, as T^* does not include those T_j's for which a jump in K_n occurred between $\tilde{\tau}_{j-1}$ and $\tilde{\tau}_j$.

9 Conclusions.

Pitman's prediction rules are an extension of Blackwell-MacQueen urn schemes, which in turn generalize Pólya's urn to general state spaces. In this paper we have carried out a parallel extension for Reinforced Urn Processes. P-RUPs are an interesting generalization of RUP not only for the state space where they are defined, but also because allow the updating rule to vary from urn to urn. P-RUP are not necessarily Markov-exchangeable: we have seen that in general they are only "locally"Markov exchangeable, in those segments, of random length, where the number of distinct visited locations does not increase. However P-RUP is not, strictly speaking, a wider class than the Markov-exchangeable family: the latter as shown by Fortini *et al* (2002), allows for exchangeable V-sequences with possibly dependent updating rules. It would be interesting to find a meaningful, similar construction with dependent Pitman urns. One possible way is to observe that Markov-exchangeable processes can be defined as those processes X for which $m_n(X)$ is, for every n, a sufficient statistic. Similarly it would be interesting to generalize the study of P-RUPs to those processes for which a sufficient statistic is $(C_n(X), Q_n(X))$ as defined above. Another appealing generalization is finding a P-RUP analog to Reinforced Urn Processes in a continuous time setting (Muliere *et al.* (2003)).

A P-RUP generates exchangeable zero-blocks only is it is assumed that all urns involved generate exchangeable sequences. In this case, P-RUP are a tool to generate particular cases of neutral to the right processes. In this paper we have only partial results about zero-blocks for general Pitman Rule. In particular, is is an interesting open problem to find a de Finetti-style representation theorem for zero-blocks generated by a P-RUP.

10 Acknowledgements

The authors wish to thank R.V. Ramamoorthi for many useful conversations and Stefano Favaro for interesting comments on exchangeable random partitions.

References

[1] D. J. Aldous. Exchangeability and related topics. In *École d'été de probabilités de Saint-Flour, XIII—1983*, volume 1117 of *Lecture Notes in Math.*, pages 1–198. Springer, Berlin, 1985.

[2] D. Blackwell and J. B. MacQueen. Ferguson distributions via Pólya urn schemes. *Ann. Statist.*, 1:353–355, 1973.

[3] B. de Finetti. Sur la condition d' "équivalence partielle ". *Actualités Scientifiques et Industrielles*, 739, 1938.

[4] P. Diaconis and D. Freedman. de Finetti's theorem for Markov chains. *Ann. Probab.*, 8(1):115–130, 1980.

[5] M. L. Eaton. *Group invariance applications in statistics.* NSF-CBMS Regional Conference Series in Probability and Statistics, 1. Institute of Mathematical Statistics, Hayward, CA, 1989.

[6] S. Fortini, L. Ladelli, G. Petris, and E. Regazzini. On mixtures of distributions of Markov chains. *Stochastic Process. Appl.*, 100:147–165, 2002.

[7] A. Gnedin and J. Pitman. Exchangeable Gibbs partitions and Stirling triangles. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 325(Teor. Predst. Din. Sist. Komb. i Algoritm), 2005.

[8] B. Hansen and J. Pitman. Prediction rules for exchangeable sequences related to species sampling. *Statist. Probab. Lett.*, 46(3):251–256, 2000.

[9] P. Muliere, P. Secchi, and S. G. Walker. Urn schemes and reinforced random walks. *Stochastic Process. Appl.*, 88(1):59–78, 2000.

[10] P. Muliere, P. Secchi, and S. G. Walker. Reinforced random processes in continuous time. *Stochastic Process. Appl.*, 104(1):117–130, 2003.

[11] J. Pitman. Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields*, 102(2):145–158, 1995.

[12] J. Pitman. Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, probability and game theory*, volume 30 of *IMS Lecture Notes Monogr. Ser.*, pages 245–267. Inst. Math. Statist., Hayward, CA, 1996.

[13] D. Spanó. *Alternative Series Representation for Random Discrete Distributions, and Applications to Statistics.* PhD thesis, Universitá degli Studi di Pavia, Italy, July 2003. Unpublished PhD Thesis.

[14] S. L. Zabell. Predicting the unpredictable. *Synthese*, 90, 1992.

[15] S. L. Zabell. Characterizing Markov exchangeable sequences. *J. Theoret. Probab.*, 8(1):175–178, 1995.

[16] A. Zaman. Urn models for Markov exchangeability. *Ann. Probab.*, 12(1):223–229, 1984.