A stochastic geometry characterization of Pitman-Yor processes

Roman Gambelin

Master's thesis
A stochastic geometry characterization of Pitman-Yor processes

Roman Gambelin
Under the direction of Bartłomiej Błaszczyszyn

September, 2020

Abstract

In this master's thesis, we give a new integral characterization of Pitman-Yor processes. It is inspired by a similar characterization for Dirichlet processes given by G. Last in [12]. The proof makes use of classical point processes theory arguments and is based on a key result found by T. Lehéricy in his 2015 master's thesis [14]. In addition, we give (Appendice C) an application of this integral equation to the computation of the moments of the random mass of a Borel set given by a Pitman-Yor process.

Contents

Acknowledgements

I would like to thank Bartłomiej Błaszczyszyn for giving me this project and for his guidance throughout its completion. I'm especially thankful of the time he took for our hebdomadary calls on the subject. His advices and suggestions on how solving the problem have been a precious help.

Introduction

Pitman-Yor (abreviated PY) processes are a class of random discrete probability measures. First introduced by M. Perman in [16], they are named after Jim Pitman and Marc Yor, following their review of the subject in [19]. Their study appears to be of significant importance as they naturally occur in different and seemingly unconnected domains: Markov processes excursion theory [16], combinatorics [2] and non-parametric Bayesian statistics [15]. This latter, in particular, concentrates most of the existing literature. Indeed, the applications are numerous as statisticians found them to be good prior distributions in various fields: natural language processing [21], image analysis [22] and populations genetics [7] to name a few.

In [12], G. Last gives an integral characterization of Dirichlet processes. Those, first introduced by T. Ferguson in [10], are in fact a special case of PY processes. The goal of this master's thesis is to give a similar characterization for the general case, using stochastic geometry arguments and formulation.
Stochastic geometry originally denotes the study of random patterns in Euclidean spaces. At its core lies the concept of random measures [3] and, more specifically, point processes. Because of their ubiquity, PY processes admit several definitions. While most rely on physically meaningful processes [13] (the names speak for themselves: "stick-breaking" process, "chinese restaurant" process, etc.), one can construct them via renormalized Poisson point processes [17]. This gives the perfect opportunity to study those processes with the classical stochastic geometry toolbox.

The paper, which is organized around the proof of this new characterization, is divided in 4 sections. The first one serves as an extended introduction and presents Dirichlet processes, Last's work and the main result (which is to be proven). The second is about the "stick-breaking" process construction. It defines the essential notions of size-biased permutations and random allocation models and gives a characterization (in those terms) of PY processes, which will be a key intermediary result. The goal of Section 3 is to translate the arguments of Section 2 in a point process formulation. As such, it is divided into two parts: the first one presents the renormalized Poisson process construction while the second one adapts the characterization of Section 2 in those terms. The final section (4) is the actual proof of the main result. Since it is a characterization, it is also divided in two parts: one for each implication.
In addition, we give in Appendice C a direct application of the main result to the computation of the moments of the mass of a Borel set given by a PY process.

Nota Bene: if one is not familiar with point processes, one might want to have a look at the reminder on the subject (Appendice A) before reading the paper.

Notations

- If $\mathbb{A}$ and $\mathbb{B}$ are two sets, we note $\mathbb{A}^{\mathbb{B}}$ the set of $\mathbb{A}$-valued families indexed by $\mathbb{B}$.
- If $\mathbb{X}$ is a topological space, then $\mathbb{M}(\mathbb{X})$ denotes the set of Radon measures on $\mathbb{X}$. Equivalently, $\mathbb{M}^d$ stands for (Radon) discrete measures, $\mathbb{M}_1$ for probability measures and $\mathbb{M}_1^d$ for discrete probability measures.
- If μ is a measure on a standard Borel space $\mathbb{X}$, we write $Supp(\mu)$ the support of μ.
- If $\mathbb{X}$ is a measurable space and $x \in \mathbb{X}$, δ_x denotes the Dirac measure centered on x. Moreover, for $x := (x_1, \dots, x_n) \in \mathbb{X}^n$, we write $\tilde{\delta}_x := \sum_{i=1}^n \delta_{x_i}$.
- If $n \in \mathbb{N}^*$, $[\![n]\!]$ denotes the set of relative integers between 1 and n (in the large sense, i.e. $1, n \in [\![n]\!]$).
- If P is a logical proposition, [P] denotes the Iverson bracket of P. That is, [P]=1 if P is true, [P]=0 otherwise.

Sets

- $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space with integral operator $\mathbf{E}$ associated to $\mathbf{P}$. In this paper, all mentioned random variables will be assumed to be defined on Ω. Some results about the existence of random variables related to almost sure equalities may require a suitable extension of $(\Omega, \mathcal{F}, \mathbf{P})$: this will always be implicitly assumed.
- $\mathbb{X}$ is a Polish space endowed with its Borel σ-algebra $\mathcal{X}$. The spaces $\mathbb{M}(\mathbb{X})$ and $\mathbb{M}^d(\mathbb{X})$ are endowed with the coarsest topology that makes the evaluation maps continuous and its corresponding Borel σ-algebra.
- We write $\Xi := \{(\alpha, \theta) \in [0, 1) \times \mathbb{R} : \theta > -\alpha\}$. Effectively, this set will serve as a parameter space.

1 Dirichlet processes

1.1 The one parameter process

Definition 1.1.1 (Dirichlet process)**.** *A $\mathbb{M}_1(\mathbb{X})$-valued random variable ζ is called a Dirichlet process with concentration parameter $\theta \in \mathbb{R}_+^*$ and base measure $\nu \in \mathbb{M}_1(\mathbb{X})$ if, for all $k \in \mathbb{N}^* \backslash \{1\}$ and any (nontrivial) partition $B_1, \dots, B_k \in \mathcal{X}$ of $\mathbb{X}$, the random vector $(\zeta(B_1), \dots, \zeta(B_k))$ follows a Dirichlet distribution with parameters $(\theta\nu(B_1), \dots, \theta\nu(B_k))$.*

The concentration parameter is sometimes ignored in the definition as it can be fully deduced by the total mass of the base measure (if one allows the latter to be any finite measure). This distinction has several justifications that will be highlighted in the next results.

Proposition 1.1.2. *Let ζ be a Dirichlet process with concentration parameter $\theta \in \mathbb{R}_+^*$ and base measure $\nu \in \mathbb{M}_1(\mathbb{X})$. Then, for any $B \in \mathcal{X}$, we have:*

$$\mathbf{E}\left[\zeta(B)\right] = \nu(B)$$

Proof. Let $B \in \mathcal{X}$ and B^c its complement in $\mathbb{X}$. By definition of ζ, we have:

$$(\zeta(B), \zeta(B^c)) \sim Dir(\theta\nu(B), \theta\nu(B^c))$$

Then, the formula for the expectation of a Dirichlet distributed random variable gives:

$$\mathbf{E}\left[\zeta(B)\right] = \frac{\theta\nu(B)}{\theta\nu(B) + \theta\nu(B^c)} = \nu(B)$$

□

Proposition 1.1.2 states that the base measure of a Dirichlet process serves as its expectation. In particular, it is independent of its concentration parameter. As given in Proposition 1.1.3, the concentration parameter has the role of an inverse variance (hence the name).

Proposition 1.1.3. *Let $(\zeta_\theta)_{\theta \in \mathbb{R}_+^*}$ be a family of $\mathbb{M}_1(\mathbb{X})$-valued random variables such that, for $\theta \in \mathbb{R}_+^*$, ζ_θ is a Dirichlet process with concentration parameter θ and base measure $\nu \in \mathbb{M}_1(\mathbb{X})$. Then, for $B \in \mathcal{X}$, we have:*

$$Var(\zeta_\theta(B)) = \frac{\nu(B)(1-\nu(B))}{1+\theta}$$

In particular:

$$\zeta_\theta(B) \xrightarrow[\theta \to \infty]{L^2} \nu(B)$$

Proof. The proof is similar to the one of Proposition 1.1.2 as it follows immediately from the variance of a Dirichlet distributed random variable. □

It can be noted that, for a fixed concentration parameter, the variance is maximal when a Borel set $B \in \mathcal{X}$ with probability $1/2$ is measured. This is because of the constraint of ν being finite (the precise value $1/2$ is due to the total mass of 1): the mass of the complementary B^c is totally determined by the one of B and so they play a symmetric role in the variance.

The following result is due to Günter Last and is the main motivation for this master's thesis.

Theorem 1.1.4 (Last, 2019)**.** *Let ν be a probability measure on $\mathbb{X}$ such that there exists $B \in \mathcal{X}$ with $\nu(B) \in (0,1)\backslash\{1/2\}$ and G a probability measure on $[0,1]$ satisfying $b := \int_{[0,1]} uG(du) \in (0,1)$. Moreover, assume that ζ is a random measure on $\mathbb{X}$, such that, for all measurable functions $f : \mathbb{M}(\mathbb{X}) \times \mathbb{X} \to \mathbb{R}_+$, the following equation holds true:*

$$\mathbf{E}\left[\int_{\mathbb{X}} f(\zeta, x)\zeta(dx)\right] = \mathbf{E}\left[\int_{\mathbb{X}}\int_{[0,1]} f((1-u)\zeta + u\delta_x, x)G(du)\nu(dx)\right] \tag{1}$$

Then, ζ is a Dirichlet process with base measure ν and concentration parameter $\theta := (1-b)/b$ and G is a Beta law with shape parameters $(1, \theta)$.

Proof. We refer directly to Last's paper [12] (sections 1-4). □

In his paper, Last also proves that the reverse of Theorem 1.1.4 is true as well. That is, equation (1) gives a characterization of Dirichlet processes. Although it may seem very abstract at this stage, its idea relies on fundamental properties of the Dirichlet process, some of which are given in the next results.

Proposition 1.1.5. *Let ζ be a Dirichlet process. Then, $\mathbf{P}$-a.s. $\zeta \in \mathbb{M}_1^d(\mathbb{X})$.*

Proof. See [4]. □

Proposition 1.1.5 is essential as it states that the Dirichlet process is ($\mathbf{P}$-a.s.) discrete. As such, one can make use of the extensive theory of point processes to formulate and study its key properties. More precisely, one can decompose a Dirichlet process ζ in the following way:

$$\zeta = \sum_{X \in Supp(\zeta)} \zeta(\{X\})\delta_X$$

Remark. *In the jargon of point processes, Proposition 1.1.2 translates as the base measure of a Dirichlet process being its intensity measure.*

In fact, a stronger result is proved.

Proposition 1.1.6. *Let ζ be a Dirichlet process with base measure $\nu \in \mathbb{M}_1(\mathbb{X})$. Then, there exists a sequence $((P_n, X_n))_{n\in\mathbb{N}^*}$ of $((0,1)\times\mathbb{X})$-valued random variables such that the elements of $(X_n)_{n\in\mathbb{N}^*}$ are i.i.d. of law ν independent of $(P_n)_{n\in\mathbb{N}^*}$, ζ admits the following decomposition $\mathbf{P}$-a.s.:*

$$\zeta = \sum_{n\in\mathbb{N}^*} P_n \delta_{X_n}$$

and the law of $(P_n)_{n\in\mathbb{N}^*}$ is unique for all Dirichlet process with same concentration parameter.

Proof. It follows from an explicit construction of the process via the gamma process. Full details can be found in [8] (Chapters 2 & 3 until Theorem 3.2.2). □

Remark. *The random sequence $(P_n)_{n\in\mathbb{N}^*}$ (in Proposition 1.1.6) arranged in decreasing order is said to have the Kingman (or 1 parameter Poisson-Dirichlet) distribution (over the set of decreasing positive sequences with sum 1). It is named after J. Kingman who first studied the process in [11]. It is worth stating that the sequence $(P_n)_{n\in\mathbb{N}^*}$ can be recovered from ζ is and only if ν is diffuse.*

This way of decomposing the process into two independent random sequences is very useful. Proposition 1.1.6 states that the distribution of the sequence $(P_n)_{n\in\mathbb{N}^*}$ is fully governed by the concentration parameter. Consequently, all the properties of the process that do not depend on a specific choice of base measure are described by the so-called (parameter dependent) Kingman distribution. This reduces a large part of the study of Dirichlet processes to the one of the Kingman distribution.

As a consequence of the reverse of Theorem 1.1.4, Last also gives the following result. As we will see in Section 2, it is one of those properties which arise directly from the specificities of the Kingman distribution.

Notation. *For $\mu \in \mathbb{M}(\mathbb{X})$ and $x \in \mathbb{X}$, we write*

$$\mu^{[x]} := \frac{\mu - \mu(\{x\})\delta_x}{1 - \mu(\{x\})}$$

if $\mu(\{x\}) < 1$, $\mu^{[x]} :\equiv 0$ otherwise.

Corollary 1.1.7. *Let ζ be a Dirichlet process with diffuse base measure and let τ be a $\mathbb{X}$-valued random variable such that, $\mathbf{P}$-a.s. and for all $B \in \mathcal{X}$, we have:*

$$\mathbf{P}(\tau \in B|\zeta) = \zeta(B)$$

Then, ζ and $\zeta^{[\tau]}$ have the same law. Moreover, $\zeta^{[\tau]}$, $\zeta(\{\tau\})$ and τ are independent.

Proof. Let ζ and τ be as defined in Proposition 1.1.7. By Theorem 1.1.4 and for all $f : \mathbb{M}(\mathbb{X}) \to \mathbb{R}_+$ measurable, we have:

$$\begin{aligned}
\mathbf{E}\left[f(\zeta^{[\tau]})\right] &= \mathbf{E}\left[\mathbf{E}\left[f(\zeta^{[\tau]}) \mid \zeta\right]\right] \\
&= \mathbf{E}\left[\int_{\mathbb{X}} f(\zeta^{[x]})\zeta(dx)\right] \\
&= \mathbf{E}\left[\int_{\mathbb{X}}\int_{[0,1]} f\left(((1-u)\zeta + u\delta_x)^{[x]}\right) G(du)\nu(dx)\right] \\
&= \mathbf{E}\left[\int_{\mathbb{X}}\int_{[0,1]} f\left(\frac{(1-u)\zeta + u\delta_x - (1-u)\zeta(\{x\})\delta_x - u\delta_x}{1-(1-u)\zeta(\{x\}) - u}\right) G(du)\nu(dx)\right]
\end{aligned}$$

Now, since ν is diffuse and ζ is $\mathbf{P}$-a.s. discrete, this gives:

$$\begin{aligned}\mathbf{E}\left[f(\zeta^{[\tau]})\right] &= \mathbf{E}\left[\int_{\mathbb{X}}\int_{[0,1]} f\left(\frac{(1-u)\zeta + u\delta_x - u\delta_x}{1-u}\right) G(du)\nu(dx)\right] \\ &= \mathbf{E}\left[\int_{\mathbb{X}}\int_{[0,1]} f(\zeta) G(du)\nu(dx)\right] \\ &= \mathbf{E}\left[f(\zeta)\right] \quad (G \in \mathbb{M}_1([0,1]) \text{ and } \nu \in \mathbb{M}_1(\mathbb{X}))\end{aligned}$$

That is, ζ and $\zeta^{[\tau]}$ have the same law. The independence of $\zeta^{[\tau]}$, $\zeta(\{\tau\})$ and τ can be shown in the same way, using the trick that $\mathbf{E}\left[f(\tau)|\zeta\right] = \int_{\mathbb{X}} f(x)\zeta(dx)$ in distribution for any $f : \mathbb{X} \to \mathbb{R}_+$ measurable. □

Remark. *The conditional expectation trick mentioned at the end of the proof of Corollary 1.1.7 is very important as it will be implicitly used all along the paper.*

It is worth mentioning that what makes Corollary 1.1.7 interesting is the discreteness of the Dirichlet process. Indeed, if not (more precisely, if it was $\mathbf{P}$-a.s. diffuse), the result would be trivially true. Moreover, since the process and its renormalized version have the same law, the corollary can be applied recursively and we deduce that the independence properties remain true if we remove any finite number of its atoms (and renormalize the process consequently).
This leads us to the goal of this paper. We want to give a similar characterization to the one given in Theorem 1.1.4 to a broader class of processes that share those same recursive independence properties.

1.2 The Pitman-Yor generalization

PY processes are a class of discrete random probability measures (in our case, on $\mathbb{X}$) that are parameterized by a base measure $\nu \in \mathbb{M}_1(\mathbb{X})$ and two real parameters $(\alpha, \theta) \in \Xi$. As suggested by the subsection's name, they are a generalization of Dirichlet processes. Indeed, a PY process with base measure ν and parameters $(0, \theta)$ is a Dirichlet process with base measure ν and concentration parameter θ. In the general case ($\alpha > 0$) θ plays the exact same role than in the Dirichlet case (and as such, is still called concentration parameter). That is, the convergence property stated in Proposition 1.1.3 holds true (see C.2). The other parameter, α, is called the dispersion parameter: while a sequence with the Kingman distribution has an exponential tail, the atoms of a PY process with diffuse base measure and positive dispersion parameter arranged in decreasing order follow asymptotically a power law. This behaviour is particularly useful, for example, to model natural language data, as the distribution of word frequencies in a ranked frequency table appear to follow a power law (see Zipf's law).

Although we cannot give a formal (parameter dependent) definition of PY processes at this stage (one will be given in Definition 2.2.2), we can state in what sense those with a diffuse base measure generalize Dirichlet processes: PY processes (with a diffuse base measure) are the random discrete probability measures satisfying the recursive independence properties deduced from Corollary 1.1.7 for Dirichlet processes (with diffuse base measure) and discussed at the end of Subsection 1.1. In this view, what makes a PY process ζ with diffuse base measure and positive dispersion parameter different from a Dirichlet process is that its renormalized process ("$\zeta^{[\tau]}$", as defined in Corollary 1.1.7) does not have the same distribution. More precisely, if ζ has real parameters $(\alpha, \theta) \in \Xi$, $\zeta^{[\tau]}$ is a PY process with same base measure but real parameters $(\alpha, \theta + \alpha)$ (see Corollary 2.2.9). As we will see in Subsection 2.2, a very similar result can be obtained for "marked versions" of PY processes with general base measures.

The next theorem gives an integral characterization of the PY processes with a diffuse base measure. It is the main result of this paper and its proof will be the end goal of the three subsequent sections. Although it is clearly inspired by Theorem 1.1.4 of G. Last, it relies on totally different arguments.

Theorem 1.2.1. *Let ζ be a random element of $\mathbb{M}_1^d(\mathbb{X})$ with $\mathbf{P}$-a.s. infinite support and such that, for any $f : \mathbb{M}^d(\mathbb{X}) \times \mathbb{X} \to \mathbb{R}_+$, the following equation holds true:*

$$\mathbf{E}\left[\int_{\mathbb{X}} f(\zeta, x)\zeta(dx)\right] = \mathbf{E}\left[\int_{\mathbb{X}}\int_{\mathbb{R}} L(\zeta) f((1-u)\zeta + u\delta_x, x) G(du)\nu(dx)\right] \tag{2}$$

for some $G \in \mathbb{M}(\mathbb{R})$, $\nu \in \mathbb{M}(\mathbb{X})$ diffuse and $L : \mathbb{M}^d(\mathbb{X}) \to \mathbb{R}_+$ such that, $\mathbf{P}$-a.s. and for $(G \otimes \nu)$-a.a. $(u, v) \in \mathbb{R} \times \mathbb{X}$, $L((1-u)\zeta + \delta_x) = (1-u)^\alpha L(\zeta)$ for some $\alpha \in [0, 1]$.
Then ζ is a PY process with base measure $\nu/\nu(\mathbb{X})$ and parameters (α, θ), where θ is given by

$$\theta := (1-\alpha)\left(\nu(\mathbb{X})\mathbf{E}\left[L(\zeta)\right]\int_{\mathbb{R}} uG(du)\right)^{-1} - 1$$

Moreover:
- $\nu(\mathbb{X})\mathbf{E}\left[L(\zeta)\right] G$ is the Beta law with parameters $(1-\alpha, \alpha+\theta)$.
- If $\alpha > 0$, $L(\zeta) = \lim_{n\to+\infty} n\left(P_{(n)}\right)^\alpha$ where $\left(P_{(n)}\right)_{n\in\mathbb{N}^*}$ are the weights of ζ arranged in decreasing order.
Finally, the reverse result is true. That is, a PY process with parameters $(\alpha', \theta') \in \Xi$ and base measure $\nu'/\nu'(\mathbb{X})$ for some $\nu' \in \mathbb{M}(\mathbb{X})$ finite satisfies equation (2) with α and ν replaced by α' and ν'.

Proof. See Section 4. □

Remark. *One can observe that, if we set $L \equiv 1$ in Theorem 1.2.1, then Theorem 1.1.7 applies directly. It is, however, not a real generalization of the latter as we require ν to be diffuse and ζ to be $\mathbf{P}$-a.s. discrete with infinite support.*

2 The stick-breaking viewpoint

2.1 Size-biased permutations and random allocation models

Definition 2.1.1 (Random probability vector)**.** *We call random probability vector any $[0,1]^{\mathbb{N}^*}$-valued random variable $P := (P_n)_{n\in\mathbb{N}^*}$ such that, $\mathbf{P}$-a.s., we have:*

$$\sum_{n\in\mathbb{N}^*} P_n = 1$$

Remark. *The terminology "random probability vector" is not common as the name used in the literature is "random probability". We chose to add the word vector in order to differentiate it from random probability measures on $\mathbb{X}$ (which we are dealing with in this paper).*

Given a random probability vector $P := (P_n)_{n\in\mathbb{N}}$, one can associate a unique point process Φ on $(0,1]$ defined by:

$$\Phi := \sum_{\substack{n\in\mathbb{N}^* \\ P_n>0}} \delta_{P_n}$$

This is possible because P almost surely does not have any accumulation point in $(0,1]$ (otherwise its series would diverge) and so Φ is Radon. Reciprocally, one can always obtain a random probability vector from a point process on $(0,1]$ with ($\mathbf{P}$-a.s.) total mass 1 and $\mathbb{N}^*$-valued weights. Indeed, taking its atoms in decreasing order (with potential multiplicities - if an atom has mass greater than 1 - and / or the addition of infinitely many 0's if the process has finite support) leads to a measurable numbering, that is, a random probability vector which associated point process is $\mathbf{P}$-a.s. equal to Φ.

Notation. *We write K and S for the following measurable functions:*

$$\begin{aligned} K : [0,1]^{\mathbb{N}^*} &\to \mathbb{N}\cup\{+\infty\} & \qquad S : [0,1]^{\mathbb{N}^*} &\to \mathbb{N}\cup\{+\infty\} \\ (P_i)_{i\in\mathbb{N}^*} &\mapsto |\{i\in\mathbb{N}^* : P_i > 0\}| & (P_i)_{i\in\mathbb{N}^*} &\mapsto \sup\{i\in\mathbb{N}^* : P_i > 0\} \end{aligned}$$

Definition 2.1.2 (Size-biased permutation)**.** *Let $P := (P_n)_{n\in\mathbb{N}^*}$ and $\tilde{P} := (\tilde{P}_n)_{n\in\mathbb{N}^*}$ be random probability vectors with the same associated point process. We call $\tilde{P}$ a size-biased permutation of P if, for all $i_1 \in \mathbb{N}^*$,*

$$\mathbf{P}(\tilde{P}_1 = P_{i_1} | P) = P_{i_1}$$

and, for all $k, i_k \in \mathbb{N}^*$ such that $k \le K(P)$, we have:

$$\mathbf{P}\left(\tilde{P}_k = P_{i_k} \,\middle|\, P, \tilde{P}_1, \dots, \tilde{P}_{k-1}\right) = \frac{P_{i_k}}{1-\sum_{i=1}^{k-1}\tilde{P}_i}\left[|\{i\in\mathbb{N}^* : P_i = P_{i_k}\}| > |\{i\in[\![k-1]\!] : \tilde{P}_i = P_{i_k}\}|\right]$$

Remark. *In the previous definition, in case the positive components of P are all different (i.e. the point process associated to P is simple), the random variable in Iverson bracket notation can be removed from the definition.*

Size-biased permutations are sometimes defined by the following equivalent formula:

$$\mathbf{P}\left(\tilde{P}_1 = P_{i_1}, \dots, \tilde{P}_k = P_{i_k} \,\middle|\, P\right) = \prod_{j=1}^{k}\left(1-\sum_{l=1}^{j-1} P_{i_l}\right)^{-1} A_j P_{i_j} \tag{3}$$

where $A_j := \left|\{n\in\mathbb{N}^* : P_n = P_{i_j}\}\right| - \left|\{l\in\mathbb{N}^* : (u<j)\wedge(P_{n_u} = P_{n_j})\}\right|$ for $j\in[\![k]\!]$.

Remark. *As before, the A_j's can be ignored in the definition whenever the point process associated to P is simple.*

Definition 2.1.2 has the advantage (over the one using Equation (3)) to be very intuitive. Indeed, it describes the process of re-indexing a random probability vector in a "self-biased" way. More precisely, let $P := (P_n)_{n\in\mathbb{N}^*}$ be a random probability vector and imagine an infinite deck of cards indexed by $\mathbb{N}^*$. For each $n\in\mathbb{N}^*$, write P_n (the value) on the n^{th} card. The two following ways of sampling give rise to a size-biased permutation of P with the same law.
Sampling without replacement: Select a card from the deck according to the probabilities written on them. Note its index (which is a random variable) N_1. Then write V_1 the value of the N_1^{th} card and remove the latter from the deck. Replace each value v written on a remaining card by v/V_1. Repeat this process from the beginning infinitely many times or until no cards with a positive value written on remain in the deck. This gives rise to two random

sequences: $(V_k)_{k\in[\![K(P)]\!]}$ and $(N_k)_{k\in[\![K(P)]\!]}$. The random probability vector $\tilde{P} := (\tilde{P}_n)_{n\in\mathbb{N}*}$ such that $\tilde{P}_k = P_{N_k}$ if $k \leq K(P)$, $\tilde{P}_k = 0$ otherwise, is a size-biased permutation of P.
Sampling with replacement and repetition rejection: This method is fundamentally the same that the first one as the distribution of the picked cards is identical. As before, we start by selecting a card from the deck according according to the values written on and noting its index N_1. We replace the card in the deck and repeat this operation until the index of the picked card is different than N_1, we note it N_2. Repeating those operations infinitely many times or until all cards with a positive value written on are picked, we obtain a sequence $N_1, \ldots, N_{K(P)}$. The random probability vector $\tilde{P}$ constructed, as in the first method, with the sequence $N_1, \ldots, N_{K(P)}$ is a size-biased permutation of P.

From both methods, we notice that, if $\tilde{P}$ is a size-biased permutation of some random probability vector P, then necessary we have $K(\tilde{P}) = S(\tilde{P}) = K(P)$ **P**-a.s.. Another important observation is that the law of a size-biased permutation does not depend on the original numbering of the random probability vector. This has two implications: the distribution of a size-biased permutation is unique for all measurable numberings of the same point process (associated to a random probability vector) and, as a consequence, taking the size-biased permutation of a random probability vector is an idempotent operation (in terms of law).

Let us define the following, related, invariance property.

Definition 2.1.3 (Invariance by sized-biased permutation)**.** *A random probability vector is said to be invariant by sized-biased permutation (ISBP) if it has the same law than its size-biased permutation.*

As mentioned previously, any size-biased permutation of a random probability vector is an ISBP random probability vector. We can thus define the unique (in law) ISBP numbering of a point process associated to a random probability vector.

Definition 2.1.4 (Residual allocations)**.** *Let $P := (P_n)_{n\in\mathbb{N}^*}$ be a random probability vector. For $n \in \mathbb{N}$, we define the n^{th} residual allocation of P as the following random variable:*

$$V_n := \begin{cases} \frac{P_n}{1-\sum_{i=1}^{n-1} P_i}, & \text{if } n \leq S(P) \\ 1, & \text{otherwise} \end{cases}$$

One can observe that the residual allocations occur naturally in the definition of a size-biased permutation. This is directly due to its underlying interpretation as a picking process without replacement. Indeed, for a random probability vector representing a division of a population into countably infinitely many sub-populations (with potential empty sub-populations, i.e. some components of P are equal to 0), its n^{th} residual allocation represents the proportion of the n^{th} sub-population in the remaining total population after removing the $n-1$ first sub-populations. The analogy with the card picking process is immediate if we replace the proportions of sub-populations by the probabilities for a card to be picked.
One can check that, for a random probability vector $(P_n)_{n\in\mathbb{N}^*}$ with family of residual allocations $(V_n)_{n\in\mathbb{N}^*}$, the following formulas hold true for any $n \in \mathbb{N}^*$:

$$1 - \sum_{i=1}^{n} P_i = \prod_{i=1}^{n}(1 - V_i) \qquad\qquad P_n = V_n \prod_{i=1}^{n-1}(1 - V_i)$$

Definition 2.1.5 (Random allocation model)**.** *We say that a random probability vector is a random allocation model (RAM) if its residual allocations are independent.*

Random allocation models are sometimes called stick-breaking models. Indeed, one can imagine a stick of length 1 that we break in two in order to remove a random proportion $V_1 \in [0, 1]$ of it. We then repeat this process on the remaining piece with a new random proportion V_2 independent of V_1 and so on. After repeating those operations an infinite number of times (or until the whole remaining stick is set apart, i.e. $V_n = 1$ for some $n \in \mathbb{N}^*$), the sequence (ordered with respect to time) of the lengths of the removed pieces of stick (with infinitely many 0's added at the end if the sequence is finite) is a random allocation model.

2.2 Griffiths-Engen-McCloskey characterizations

Definition 2.2.1 (Griffiths-Engen-McCloskey)**.** *We call Griffiths-Engen-McClosckey (GEM) with parameters $(\alpha, \theta) \in \Xi$ a RAM which n^{th} residual allocation follows the Beta law with parameters $(1 - \alpha, \theta + n\alpha)$ for each $n \in \mathbb{N}^*$.*

We can now properly define PY processes.

Definition 2.2.2 (PY process)**.** *We call PY process with base measure $\nu \in \mathbb{M}(\mathbb{X})$ and parameters $(\alpha, \theta) \in \Xi$ a $\mathbb{M}_1^d(\mathbb{X})$-valued random variable ζ such that,* **P***-a.s., we have:*

$$\zeta = \sum_{n\in\mathbb{N}^*} P_n \delta_{X_n}$$

for some GEM $P := (P_n)_{n\in\mathbb{N}^*}$ with parameters (α,θ) and sequence $(X_n)_{n\in\mathbb{N}^*}$ of i.i.d. ν-distributed random variables independent of P.

Definition 2.2.3. *We call Poisson-Dirichlet process with parameters $(\alpha,\theta)\in\Xi$ a GEM with the same parameters but ranked in decreasing order.*

Remark. *The term "Poisson" in Poisson-Dirichlet is due to a construction of these processes via specific Poisson processes. This will be the subject of Section 3. The Dirichlet part comes directly from the next proposition (2.2.4).*

Proposition 2.2.4. *Let $\theta\in\mathbb{R}_+^*$. The Poisson-Dirichlet distribution with parameters $(0,\theta)$ is the Kingman distribution with parameter θ.*

Proof. See [20]. □

From Propositions 1.1.6 and 2.2.4, one immediatly sees that a PY process with base measure $\nu\in\mathbb{M}_1(\mathbb{X})$ and parameters $(0,\theta)$ for some $\theta\in\mathbb{R}_+^*$ is a Dirichlet process with base measure ν and concentration parameter θ. In fact, this is what is directly proved by Sethuraman in [20]. As such, PY processes can be rightfully considered as a generalization of Dirichlet processes.

Remark. *Initially, the term "Poisson-Dirichlet processes" denoted only the Kingman case. The two-parameter processes are hence sometimes referred to as generalized Poisson-Dirichlet processes.*

As mentioned in the next result by Pitman, GEMs are the only ISBP RAMs with **P**-a.s. infinitely many positive components.

Theorem 2.2.5 (Pitman, 1992)**.** *Let $P := (P_n)_{n\in\mathbb{N}^*}$ be an ISBP RAM with associated sequence of residual allocations $(V_n)_{n\in\mathbb{N}^*}$ and such that $P_1<1$ **P**-a.s..*
*Then, if $K(P)=+\infty$ **P**-a.s.:*
(i) There exists $(\alpha,\theta)\in\Xi$ such that P is a GEM with parameters (α,θ).
*Otherwise, there exists $m\in\mathbb{N}^*$ such that $K(P)=m$ **P**-a.s. and we are in one of the following cases:*
(ii) There exists $\beta>0$ such that, for each $n\in[\![m]\!]$, V_n follows the Beta law with parameters $(1+\beta, m\beta-n\beta)$.
*(iii) $V_n := 1/(m-n+1)$ **P**-a.s., that is, P is constant on its support.*
(iv) If $m=2$, the probability measure F on $(0,1)$ defined by

$$F(dv) = (1-v)\mathbf{P}(V_1\in dv)/\mathbf{E}[1-V_1]$$

is symmetric with respect to $1/2$.
Reciprocally, all of the above enumerate cases (for arbitrary parameters $(\alpha,\beta)\in\Xi$) are ISBP RAMs.

Proof. See Theorem 2 of [18]. □

Theorem 2.2.5 is a first step toward the result we are aiming to prove as it gives a classification of ISBP RAMs and a characterization of the distribution of the weights of a PY process with a diffuse base measure. Indeed, recall that the ISBP numbering of a point process associated to a random probability vector (in our case, the weights of the PY process) is unique in law. As a consequence, the distribution of the residual allocations of the ISBP numbering is also unique and one can check if the process is a PY by studying their independence. The base measure is required to be diffuse because, otherwise, one could have (in fact **P**-a.s.) several weights associated to the same Dirac measure in the decomposition given in Definition 2.2.2 (which would make them indistinguishable).
Since our main tool for this study is the use of point processes, it makes sense to reformulate the definiton of an ISBP RAM in relative terms. The following result (Proposition 2.2.6), directly taken from the 2015 master's thesis of Thomas Lehéricy, fulfills this role. First, some notation is to be introduced.

Notation. *For a point process Φ on a subset of $\mathbb{R}_+^*$ and a constant $a\in\mathbb{R}_+^*$, we write:*

$$a\cdot\Phi := \sum_{U\in Supp(\Phi)} \Phi(\{U\})\delta_{aU}$$

Notation. *For a point process Φ associated to a random probability vector and a corresponding ISBP numbering $P := (P_n)_{n\in\mathbb{N}^*}$, we will write, for each $n\in\mathbb{N}^*$:*

$$\Phi^{[n]} := \left(1-\sum_{i=1}^n P_i\right)^{-1}\cdot\left(\Phi-\sum_{i=1}^n\delta_{P_i}\right)$$

In what follows, unless precised otherwise, when the notation $\Phi^{[n]}$ is used with respect to a point process Φ and an associated ISBP numbering P has already been mentioned, we will always assume that $\Phi^{[n]}$ is obtained from P. (This has to be assume since, although $\Phi^{[n]}$ is unique in law (for each $n\in\mathbb{N}^*$) by the uniqueness (in law again) of P, one can have have an ISBP numbering P' not related to $\Phi^{[n]}$ given Φ, or even independent of it.)

Proposition 2.2.6 (Lehéricy 2015). *Let P be an ISBP random probability vector with associated point process Φ and family of residual allocations $(V_n)_{n\in\mathbb{N}^*}$. The following properties are equivalent:*
(i) For all $n \in \mathbb{N}^*$, $V_1, \dots, V_n$ and $\Phi^{[n]}$ are all independent.
(ii) For all $n \in \mathbb{N}^*$, V_n and $\Phi^{[n]}$ are independent.
(iii) P is a RAM.

Proof. Let us prove by induction that $(ii) \implies (i)$. It is clear that the two properties are equal for $n = 1$. Let then $m \in \mathbb{N}^*$ and assume (i) holds true for $n \leq m$. By construction, the law of V_{m+1} given $(V_1, \dots, V_m, \Phi^{[m]})$ depends only on $\Phi^{[m]}$, and, by the induction hypothesis, $\Phi^{[m]}$ is independent of $(V_1, \dots, V_m)$. Therefore $(V_{m+1}, \Phi^{[m+1]})$ is independent of $(V_1, \dots, V_m)$. Now, by (ii), V_{m+1} and $\Phi^{[m+1]}$ are independent. We deduce (i) for $n = m + 1$.
Let us prove that $(iii) \implies (i)$. Let $n \in \mathbb{N}^*$. For all $k \in \mathbb{N}^*$ such that $k > n$, let us define:

$$Y_k^n := \frac{P_k}{\prod_{i=1}^n (1 - V_i)} = V_k \prod_{i=n+1}^{k-1} (1 - V_i)$$

By definition of a RAM, the family $(V_k)_{k\in\mathbb{N}^*}$ is independent. In particular, $(V_1, \dots, V_n)$ is independent of $(V_k)_{k>n}$. Hence, the right hand side of the above equation makes it clear that the family $(Y_k^n)_{k>n}$ is independent of $(V_1, \dots, V_n)$. Now, after recalling that $1 - \sum_{i=1}^n P_i = \prod_{i=1}^n (1 - V_i)$, we have:

$$\Phi^{[n]} = \left(\prod_{i=1}^n (1 - V_i)\right)^{-1} \cdot \left(\Phi - \sum_{k=1}^n \delta_{P_k}\right) = \sum_{k>n} \delta_{\left(\prod_{i=1}^n (1-V_i)\right)^{-1} P_k} = \sum_{k>n} \delta_{Y_k^n}$$

We deduce that $\Phi^{[n]}$ is independent of $(V_1, \dots, V_n)$ and, as a consequence, that (i) holds true.
Finally, (ii) follows immediately (i) and the independence of $(V_n)_{n\in\mathbb{N}}$ in (i) is the definition of a RAM (hence $(i) \implies (iii)$). □

Property (ii) of Proposition 2.2.6 gives a method to prove that the ISBP numbering of a point process associated to a random probability vector is a RAM. By Theorem 2.2.5, it is consequently a method to prove that (the atoms arranged in decreasing order of) a point process on $(0, 1]$ follows a Poisson-Dirichlet distribution. Ideally, one would like to have a similar technique for PY processes.
Before looking for such an equivalence, one can state the following result on the law of the renormalized processes.

Proposition 2.2.7. *Let $P = (P_n)_{n\in\mathbb{N}^*}$ be a GEM with parameters $(\alpha, \theta) \in \Xi$ and Φ its associated point process. Then, for each $n \in \mathbb{N}^*$, $\Phi^{[n]}$ is a point process associated to a GEM with parameters $(\alpha, \theta + n\alpha)$ and is independent of $(P_1, \dots, P_n)$.*

Proof. The first step is to show that $\Phi^{[1]}$ is associated to an ISBP RAM. Let $(V_n)_{n\in\mathbb{N}^*}$ be the family of residual allocations of P. By definition, we have:

$$\begin{aligned}\Phi^{[1]} &:= (1 - P_1)^{-1} \cdot (\Phi - \delta_{P_1}) \\ &= (1 - V_1)^{-1} \cdot \left(\sum_{n\in\mathbb{N}*\setminus\{1\}} \delta_{P_n}\right)\end{aligned}$$

Since $P_n = V_n \prod_{k=1}^{n-1} (1 - V_k)$ for all $n \in \mathbb{N}^*$, this gives:

$$\Phi^{[1]} = \sum_{n\in\mathbb{N}^*} \delta_{V_{n+1} \prod_{k=2}^n (1-V_k)}$$

By independence of $(V_n)_{n\in\mathbb{N}^*}$ we deduce that the process $P' := (P'_n)_{n\in\mathbb{N}^*} = (V_{n+1} \prod_{k=2}(1 - V_k))_{n\in\mathbb{N}^*}$ (which is a measurable numbering of $\Phi^{[1]}$) is a GEM with parameters (α', θ') for some $(\alpha', \theta') \in \Xi$. It remains to identify (α', θ'). Let $(V'_n)_{n\in\mathbb{N}^*}$ be the family of residual allocations of P'. By definition, V'_1 follows the Beta law with parameters $(1 - \alpha', \theta' + \alpha')$ and as a consequence:

$$\mathbf{E}[V'_1] = \frac{1 - \alpha'}{1 + \theta'}$$

On the other hand, V_2 follows the Beta law with parameters $(1 - \alpha, \theta + 2\alpha)$ and we have $V'_1 = V_2$. Equaling the expectations gives:

$$\frac{1 - \alpha'}{1 + \theta'} = \frac{1 - \alpha}{1 + \theta + \alpha}$$

Repeating the same argument for the equality $V'_2 = V_3$ gives:

$$\frac{1 - \alpha'}{1 + \theta' + \alpha'} = \frac{1 - \alpha}{1 + \theta + 2\alpha}$$

Solving the two equation system for (α', θ'), we obtain $\alpha' = \alpha$ and $\theta' = \theta + \alpha$. By induction, we deduce that, for each $n \in \mathbb{N}^*$, any ISBP numbering P^n of $\Phi^{[n]}$ is a GEM with parameters $(\alpha, \theta + n\alpha)$. The independence property follows from the independence of the residual allocations and from the fact that the n^{th} component of a random probability vector is a measurable function of its n^{th} first residual allocations. □

From the definition of the PY process (2.2.2), one can see that such a process can be fully represented by a point process associated to a GEM independently marked with i.i.d. random variables distributed accordingly to its base measure. More precisely, let ζ be a PY process with base measure $\nu \in \mathbb{M}_1(\mathbb{X})$ and admitting the following decomposition $\mathbf{P}$-a.s:

$$\zeta = \sum_{n \in \mathbb{N}^*} P_n \delta_{X_n}$$

where $P := (P_n)_{n \in \mathbb{N}*}$ is a GEM and X_n are i.i.d ν-distributed random variables independent of P. Then, ζ is a measurable function of the simple point process ξ on $(0,1] \times \mathbb{X}$ defined by:

$$\xi := \sum_{n \in \mathbb{N}^*} \delta_{(P_n, X_n)}$$

Moreover, and as mentioned previously, the function which maps ξ to ζ defined by

$$\mathbb{M}^d((0,1] \times \mathbb{X}) \to \mathbb{M}^d(\mathbb{X})$$
$$\left(\mu : f \mapsto \int_{(0,1]\times\mathbb{X}} f(u,x)\mu(d(u,x))\right) \mapsto \left(\mu : f \mapsto \int_{(0,1]\times\mathbb{X}} uf(x)\mu(d(u,x))\right)$$

is bijective when restricted to the possible values of ξ if and only if ν is diffuse (in the above definition Radon measures are seen as functionals).

The next result (Theorem 2.2.8), also due to Thomas Lehéricy, gives a generalization of Proposition 2.2.6 to marked processes. We first have to adapt the previous notation.

Notation. *For a point process ξ on $(0,1] \times \mathbb{X}$ and a constant $a \in \mathbb{R}_+^*$, we write:*

$$a \cdot \xi := \sum_{(U,X) \in Supp(\xi)} \xi(\{(U,X)\}) \delta_{(aU,X)}$$

Notation. *Let $((P_n, X_n))_{n\in\mathbb{N}^*}$ be an ISBP numbering (in the sense that the elements of the sequence are permutated by a size-biased permutation accordingly to their first component) of a point process ξ on $(0,1] \times \mathbb{X}$ such that $\int_{(0,1]\times\mathbb{X}} u\xi(d(u,x)) = 1$ $\mathbf{P}$-a.s.. Then, for each $n \in \mathbb{N}^*$, we write:*

$$\xi^{[n]} := \left(1 - \sum_{i=1}^{n} P_i\right)^{-1} \cdot \left(\xi - \sum_{i=1}^{n} \delta_{(P_i, X_i)}\right)$$

As before, if an ISBP numbering $((P'_n, X'_n))_{n\in\mathbb{N}^*}$ of ξ has already been mentioned, we will always assume (unless precised otherwise) that $\xi^{[n]}$ is taken from $((P'_n, X'_n))_{n\in\mathbb{N}^*}$.

Theorem 2.2.8 (Lehéricy, 2015)**.** *Let $((P_n, X_n))_{n\in\mathbb{X}^*}$ be an ISBP numbering of a point process ξ on $(0,1] \times \mathbb{X}$ such that $\int_{(0,1]\times\mathbb{X}} u\xi(d(u,x)) = 1$ $\mathbf{P}$-a.s.. Moreover, let $(V_n)_{n\in\mathbb{N}^*}$ be the family of residual allocations associated to the random probability vector $(P_n)_{n\in\mathbb{N}^*}$. Then, the following properties are equivalent:*
(i) For all $n \in \mathbb{N}^*$, $(V_1, X_1), \dots, (V_n, X_n)$ and $\xi^{[n]}$ are independent.
(ii) For all $n \in \mathbb{N}^*$, (V_n, X_n) and $\xi^{[n]}$ are independent.
(iii) $(P_n)_{n\in\mathbb{N}}$ is a RAM and there exists $\nu \in \mathbb{M}_1(\mathbb{X})$ such that the family $(X_n)_{n\in\mathbb{N}}$ is i.i.d. independent of $(P_n)_{n\in\mathbb{N}}$ and, for each $n \in \mathbb{N}^*$, X_n has law ν.

Proof. See Appendice B. □

Theorem 2.2.8 gives a sufficient condition for a point process on $(0,1] \times \mathbb{X}$ to represent a PY process on $\mathbb{X}$. Since the natural marked process associated to a PY process (the combination of weights and atoms) satisfies the properties of Theorem 2.2.8 whenever its base measure is diffuse, we can deduce a characterization of such processes. As such, it is a key intermediary result in our way for proving Theorem 1.2.1. This necessary condition of the base measure to be diffuse in order to recover the distribution of the GEM associated to a PY process is the main reason why it is set as an assumption in Theorem 1.2.1.
The goal of the next section (Section 3) is to translate the characterization deduced from Theorem 2.2.8 exclusively into point processes vocabulary. That is, to get rid of the notion of ISBP random probability vectors and RAMs. As a first step toward this transition to random measures and to close this subsection, we explicit Proposition 2.2.7 in term of PY processes.

Corollary 2.2.9. *Let ζ be a PY process with parameters $(\alpha, \theta) \in \Xi$ and diffuse base measure $\nu \in \mathbb{M}_1(\mathbb{X})$. Moreover, let τ be an $\mathbb{X}$ valued random variable such that:*

$$\mathbf{P}(\tau \in B|\zeta) = \zeta(B)$$

for all $B \in \mathcal{X}$. Then, the renormalized process $\zeta^{[\tau]}$ is a PY process with parameters $(\alpha, \theta + \alpha)$ and base measure ν. Moreover, τ, $\zeta^{[\tau]}$ $\zeta(\{\tau\})$ are independent.

Proof. The law of $\zeta^{[\tau]}$ and its independence with $\zeta(\{\tau\})$ are a direct consequence of Proposition 2.2.7 (the diffuseness of ν allows one to recover the distribution of the GEM). By the construction of a PY process as an independently marked process with i.i.d. random variables, we deduce the two others independence properties. □

Remark. *Corollary 2.2.9 is the extension of Corollary 1.1.7 to general PY processes.*

3 The normalized Poisson process viewpoint

As in Section 2, we first aim to introduce a new construction of the Poisson-Dirichlet processes.

3.1 Poisson-Dirichlet processes

The following construction is due to Pitman, Perman and Yor and was presented in [17]. In the case where one of the two parameters is null, the point process associated to a Poisson-Dirichlet process can be represented by a normalized Poisson point process. This latter (un-normalized) represents the jumps of pure jump subordinator (in the case with a null dispersion parameter, we are back to the construction via a Gamma process, already mentioned for the proof of Proposition 1.1.6). For the general case (the two parameters are non-zero), one has to add a parameter dependent density to one of the two simpler cases.

Null concentration parameter

Let $\alpha \in (0, 1)$ and let Ψ_α be a Poisson point process on $\mathbb{R}_+^*$ with intensity Λ_α given by:

$$\Lambda_\alpha(dt) = t^{-1-\alpha}dt$$

We recall that, by definition of a point process, the atoms of Ψ_α have $\mathbf{P}$-a.s. no accumulation point on $\mathbb{R}_+^*$. Moreover, it has $\mathbf{P}$-a.s. only a finite number of atoms past ε for any $\varepsilon \in \mathbb{R}_+^*$. Indeed, for $\varepsilon \in \mathbb{R}_+^*$, we have:

$$\mathbf{E}\left[\Psi_\alpha([\varepsilon, +\infty))\right] = \Lambda_\alpha([\varepsilon, +\infty)) = \int_\varepsilon^{+\infty} t^{-1-\alpha}dt < +\infty$$

On the other hand, since $\Lambda_\alpha((0, \varepsilon])) = +\infty$ for any $\varepsilon \in \mathbb{R}_+^*$, 0 is $\mathbf{P}$-a.s. an accumulation point in $\mathbb{R}_+$. Put in more formal terms, Ψ_α is a point process on $\mathbb{R}_+^*$ with $\mathbf{P}$-a.s. infinite support but bounded atoms. As such, one can give a measurable numbering (a random sequence whose elements are the atoms of the process) of Ψ by arranging its atoms in decreasing order. As a point process on a subset of an Euclidean space, it is possible to do some standard operations, such as the sum of its atoms.

Notation. *Given a point process Φ on a subset A of $\mathbb{R}_+^*$, we write $\Sigma(\Phi)$ the sum of its atoms. That is:*

$$\Sigma(\Phi) := \int_A t\Phi(dt)$$

In our case, we have:

$$\mathbf{E}\left[\int_{(0,1]} t\Psi_\alpha(dt)\right] = \int_0^1 t^{-\alpha}dt < +\infty$$

Hence, Ψ_α has $\mathbf{P}$-a.s. finite sum on $(0, 1]$. Since, Ψ_α has $\mathbf{P}$-a.s. finitely many atoms greater than 1, we deduce that, $\mathbf{P}$-a.s.:

$$\Sigma(\Psi_\alpha) < +\infty$$

Remark. *Although $\Sigma(\Psi_\alpha)$ is $\mathbf{P}$-a.s. finite, it is not integrable. Indeed, we have:*

$$\mathbf{E}\left[\Sigma(\Psi_\alpha)\right] \geq \mathbf{E}\left[\Psi_\alpha([1, \infty))\right] = \int_1^{+\infty} t^{-\alpha}dt = +\infty$$

Using the previous property, one can define the normalized process Φ_α given by:

$$\Phi_\alpha := \frac{\Psi_\alpha}{\Sigma(\Psi_\alpha)}$$

Φ_α is a point process on $(0,1]$ of sum $\Sigma(\Phi_\alpha) = 1$ **P**-a.s. and, as such, can be associated with a random probability vector. In fact, the next result connects it with the Poisson-Dirichlet process.

Proposition 3.1.1. *The atoms of Φ_α arranged in decreasing order have the Poisson-Dirichlet distribution with parameters $(\alpha, 0)$.*

Null discount parameter ("Kingman's" case)

Let now $\theta \in \mathbb{R}_+^*$ and Ψ^θ be a Poisson point process on $\mathbb{R}_+*$ with density Λ^θ (the upper script θ in Ψ^θ is not to be confused with the θ^{th} - assuming $\theta \in \mathbb{N}$ - moment of Ψ) given by:

$$\Lambda^\theta(dt) := \frac{\theta}{t} e^{-t} dt$$

As before, one can check that, **P**-a.s., 0 is the only accumulation point of the atoms of Ψ^θ in $\mathbb{R}_+$ and $\Sigma(\Psi^\theta) < \infty$. We can then define the normalized process Φ^θ given by:

$$\Phi^\theta := \frac{\Psi^\theta}{\Sigma(\Psi^\theta)}$$

Proposition 3.1.2. *The atoms of Φ^θ arranged in decreasing order have the Poisson-Dirichlet distribution with parameters $(0, \theta)$.*

In this case, one can state the following additional result.

Proposition 3.1.3. *$\Sigma(\Psi^\theta)$ and Φ^θ are independent.*

Proof. See Theorem 2.2 of [9]. □

General case

Before giving an alternative definition for the general case, we must first introduce a measurable function of the point processes, which stands for the "L" in Equation (2).

Notation. *For a point process Φ on $\mathbb{R}_+^*$ (or any interval of the form $(0, a]$ with $a \in \mathbb{R}_+^*$) with **P**-a.s. infinite support and bounded atoms (this implies that 0 is the only accumulation point of its atoms and that there are only finitely many points greater than a certain positive quantity) and $\alpha' \in (0,1)$, we define $L_{\alpha'}(\Phi)$ as the following random variable (provided the limit exists **P**-a.s.):*

$$L_{\alpha'}(\Phi) := \lim_{n \to +\infty} n \left(Y_{(n)}\right)^{\alpha'}$$

where $\left(\left(Y_{(n)}\right)_{n\in\mathbb{N}}\right.$ is the atoms of Φ arranged in decreasing order.

Let $\left(Y_{(n)}\right)_{n\in\mathbb{N}^*}$ be the atoms of Ψ_α arranged in decreasing order. By the displacement theorem for Poisson point processes, we know that the sequence $\left(Y_{(n)}^{-\alpha}\right)_{n\in\mathbb{N}^*}$ has the distribution of the atoms (arranged in increasing order) of a Poisson point process with Lebesgue intensity on $\mathbb{R}_+^*$. That is, the increments of $\left(Y_{(n)}^{-\alpha}\right)_{n\in\mathbb{N}^*}$ are i.i.d. exponential random variables with parameter 1. By the strong law of large numbers, we deduce that the following convergence happens **P**-a.s.:

$$\frac{Y_{(n)}^{-\alpha}}{n} \underset{n\to\infty}{\longrightarrow} 1$$

Hence, $L_\alpha(\Psi_\alpha)$ and $L_\alpha(\Phi_\alpha)$ are well defined and we have the following equalities **P**-a.s.:

$$L_\alpha(\Psi_\alpha) = 1$$
$$L_\alpha(\Phi_\alpha) = \Sigma(\Psi_\alpha)^{-\alpha}$$

We now assume that $(\alpha, \theta) \in \Xi$ (without any other restricton). Let $\Phi_{(\alpha,\theta)}$ be a point process on $(0,1]$ which distribution is defined, for any $f : \mathbb{M}^d((0,1]) \to \mathbb{R}_+$ measurable, by:

$$\mathbf{E}\left[f(\Phi_{(\alpha,\theta)})\right] := \begin{cases} \mathbf{E}\left[f(\Phi_\alpha)\right], & \text{if } \theta = 0 \\ \mathbf{E}\left[f(\Phi^\theta)\right], & \text{if } \alpha = 0 \\ \mathbf{E}\left[L_\alpha(\Phi_\alpha)^{\theta/\alpha}\right]^{-1} \mathbf{E}\left[L_\alpha(\Phi_\alpha)^{\theta/\alpha} f(\Phi_\alpha)\right], & \text{otherwise} \end{cases} \tag{4}$$

One can observe that, because of the relation $L_\alpha(\Phi_\alpha)^{\theta/\alpha} = \Sigma(\Psi_\alpha)^{-\theta}$ **P**-a.s. (provided $\alpha \neq 0$), the case $\theta = 0$ in (4) is a special case of the third case and hence can be removed from the above definition.

Proposition 3.1.4. *The atoms of $\Phi_{(\alpha,\theta)}$ arranged in decreasing order follow the Poisson-Dirichlet distribution with parameters (α,θ).*

In [18], Pitman and Yor proved the additional result that the function which maps the parameters (α,θ) to the corresponding Poisson-Dirichlet distribution (in the space of distributions on $[0,1]^{\mathbb{N}^*}$ equipped with the topology of weak convergence of finite dimensional distributions) is continuous on Ξ.

3.2 Slivnyak-like theorem for marked processes

We here aim to give a necessary and sufficient condition for (the elements of) a sequence of point processes on $(0,1]\times\mathbb{X}$ to be PY processes. More precisely, we want to translate the characterization given by Theorem 2.2.8 in terms of the Palm distributions of the processes, just as the Slivnyak theorem does it for Poisson point processes (hence the "Slivnyak-like"). This new characterization is stated by the following two corollaries. It has to be cited that they are a straightforward extension to marked processes of two results given in [14].

The first corollary (3.2.1) gives the Palm distributions of the marked version of a PY process.

Corollary 3.2.1. *Let $(\alpha,\theta)\in\Xi$ and $(\xi_n)_{n\in\mathbb{N}}$ a sequence of point processes on $(0,1]\times\mathbb{X}$ such that, for each $n\in\mathbb{N}$, ξ_n admits the following decomposition $\mathbf{P}$-a.s.:*

$$\xi_n=\sum_{k\in\mathbb{N}^*}\delta_{(P_k^n,X_k^n)}$$

where $P^n:=(P_k^n)_{k\in\mathbb{N}^*}$ is a GEM with parameters $(\alpha,\beta+n\alpha)$ and $(X_k^n)_{k\in\mathbb{N}^*}$ is a sequence of i.i.d. μ_n-distributed random variables independent of P^n, with μ_n being the intensity of ξ_n.
Then, for all $n\in\mathbb{N}$, $f:\mathbb{M}^d((0,1]\times\mathbb{X})\to\mathbb{R}_+$ and μ_n-a.a. $(u,x)\in(0,1]\times\mathbb{X}$, we have:

$$\int_{\mathbb{M}^d((0,1]\times\mathbb{X})}f(\phi)\mathcal{P}^{n,(1)}_{(u,x)}(d\phi)=\mathbf{E}\left[f((1-u)\cdot\xi_{n+1})\right]$$

where $\mathcal{P}^{n,(1)}_{(u,x)}$ is the first order reduced Palm distribution of ξ_n at (u,x).

Proof. Let $n\in\mathbb{N}^*$. We write $(P^n_{(k)})$ the sequence P^n in decreasing order and $u\in(0,1]$, we define:

$$T(u)=\mathbf{P}(P_1^n=u|\xi_n)$$

We recall that, by the CLM formula and for all $f:\mathbb{M}^d((0,1]\times\mathbb{X})\times(0,1]\times\mathbb{X}\to\mathbb{R}_+$ measurable, we have:

$$\mathbf{E}\left[\int_{(0,1]\times\mathbb{X}}f(\xi_n-\delta_{(u,x)},u,x)\xi_n(d(u,x))\right]=\int_{(0,1]\times\mathbb{X}}\int_{\mathbb{M}^d((0,1]\times\mathbb{X})}f(\phi,u,x)P^{n,(1)}_{(u,x)}(d\phi)\mu_n(d(u,x))$$

On the other hand, we have:

$$\begin{aligned}\mathbf{E}\left[\int_{(0,1]\times\mathbb{X}}f(\xi_n-\delta_{(u,x)},u,x)\xi_n(d(u,x))\right]&=\mathbf{E}\left[\int_{(0,1]\times\mathbb{X}}T(u)T(u)^{-1}f(\xi_n-\delta_{(u,x)},u,x)\xi_n(d(u,x))\right]\\&=\mathbf{E}\left[T(P_1)^{-1}f(\xi_n-\delta_{(P_1,X_1)},P_1,X_1)\right]\\&=\mathbf{E}\left[T(P_1)^{-1}f\left((1-P_1)\cdot\xi_n^{[1]},P_1,X_1\right)\right]\end{aligned}$$

where in the second inequality we have used the conditional expectation trick mentioned in the remark after the proof of Corollary 1.1.7. Now, since $\xi_n^{[1]}$ is independent of (P_1,X_1) by Theorem 2.2.8 and, by Corollary 2.2.9, has the same law as ξ_{n+1}, that we can assume (without loss of generality) independent of ξ_n (and so (P_1,X_1)), we have:

$$\begin{aligned}\mathbf{E}\left[\int_{(0,1]\times\mathbb{X}}f(\xi_n-\delta_{(u,x)},u,x)\xi_n(d(u,x))\right]&=\mathbf{E}\left[T(P_1)^{-1}f\left((1-P_1)\cdot\xi_{n+1},P_1,X_1\right)\right]\\&=\mathbf{E}\left[\int_{(0,1]\times\mathbb{X}}T(u)T(u)^{-1}f((1-u)\cdot\xi_{n+1},u,x)\xi_n(d(u,x))\right]\\&=\int_{(0,1]\times\mathbb{X}}\mathbf{E}\left[f((1-u)\cdot\xi_{n+1},u,x)\right]\mu_n(d(u,x))\end{aligned}$$

Equaling the right terms of the last and first equations of the proof yields:

$$\int_{(0,1]\times\mathbb{X}}\int_{\mathbb{M}^d((0,1]\times\mathbb{X})}f(\phi,u,x)P^{n,(1)}_{(u,x)}(d\phi)\mu_n(d(u,x))=\int_{(0,1]\times\mathbb{X}}\mathbf{E}\left[f((1-u)\cdot\xi_{n+1},u,x)\right]\mu_n(d(u,x))$$

Taking $f(\phi,u,x)=h(\phi)\mathbb{1}_\Gamma((u,x))$ for some $h:M^d((0,1]\times\mathbb{X})\to\mathbb{R}_+$ measurable and Borel set $\Gamma\subset(0,1]\times\mathbb{X}$, we deduce (by density) the desired result. □

Remark. *As it is made clear in its proof, Corollary 3.2.1 relates the Palm distributions of (the marked version of) a PY process ξ with the ones of its renormalized processes $\xi^{[n]}$ ($n \in \mathbb{N}^*$).*

The next corollary (3.2.2) is the exact reverse of Corollary 3.2.1 and ends the characterization.

Corollary 3.2.2. *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of point processes on $(0,1]\times\mathbb{X}$ with $\mathbf{P}$-a.s. infinite support and such that, for each $n \in \mathbb{N}$:*

$$\int_{(0,1]\times\mathbb{X}} u\xi_n(d(u,x)) = 1 \quad \mathbf{P}\text{-a.s.} \tag{5}$$

and

$$\int_{\mathbb{M}^d((0,1]\times\mathbb{X})} f(\phi)\mathcal{P}^{n,(1)}_{(u,x)}(d\phi) = \mathbf{E}\left[f\left((1-u)\cdot\xi_{n+1}\right)\right] \tag{6}$$

for all $f : \mathbb{M}^d((0,1]\times\mathbb{X}) \to \mathbb{R}_+$ and μ_n-a.a. $(u,x) \in (0,1]\times\mathbb{X}$, where μ_n is the intensity of ξ_n and $\mathcal{P}^{n,(k)}_{(u,x)}$ its first order Palm distribution at (u,x).
Then, there exist $(\alpha,\theta) \in \Xi$ such that, for all $n \in \mathbb{N}$ and any ISBP numbering $((P^n, X^n_k))_{k\in\mathbb{N}^*}$ of ξ_n, $P^n := (P^n_k)_{k\in\mathbb{N}^*}$ is a GEM with parameters $(\alpha,\theta+n\alpha)$ and $(X^n_k)_{k\in\mathbb{N}}$ is a sequence of i.i.d. distributed random variables independent of P^n. Moreover, for all $f : \mathbb{M}^d((0,1]\times\mathbb{X}) \to \mathbb{R}_+$ and $\mu^{(k)}_n$-a.a. $((u_1,x_1),\dots,(u_k,x_k)) \in ((0,1]\times\mathbb{X})^k$, we have:

$$\int_{\mathbb{M}^d((0,1]\times\mathbb{X})} f(\phi)\mathcal{P}^{n,(k)}_{((u_1,x_1),\dots,(u_k,x_k))}(d\phi) = \mathbf{E}\left[f\left(\left(1-\sum_{i=1}^k u_i\right)\cdot\xi_{n+k}\right)\right] \tag{7}$$

where $\mu^{(k)}_n$ is the k^{th} factorial moment of ξ_n and $\mathcal{P}^{n,(k)}_{((u_1,x_1),\dots,(u_k,x_k))}$ its k^{th} order reduced Palm distribution at $((u_1,x_1),\dots,(u_k,x_k))$.

Proof. Let $(\xi_n)_{n\in\mathbb{N}}$ be a family of point processes on $(0,1]\times\mathbb{X}$ such that, for each $n \in \mathbb{N}$, ξ_n satisfies Equations (5) and (6). Let $n \in \mathbb{N}$.
We are first going to prove by induction that (7) holds true for every $k \in \mathbb{N}^*$. For $k = 1$, Equations (6) and (7) are exactly the same. Let $k \in \mathbb{N}^*$ and let us assume that (7) holds for $i \in [\![k]\!]$. For $l \in \mathbb{N}^*$, we will write $(u,x)_l$ to denote an element of $((0,1]\times\mathbb{X})^l$. Moreover, for $(u,x)_l := ((u_1,x_1),\dots,(u_l,x_l)) \in ((0,1]\times\mathbb{X})^l$, we write $\tilde{u}_l := \sum_{i=1}^l u_i$ and for all $u := (u_1,\dots,u_l) \in (0,1]^l$ and $i \in [\![l]\!]$ we write $\tilde{u}_i = \sum_{j=1}^i u_j$. For $f : ((0,1]\times\mathbb{X})^k \to \mathbb{R}_+$ measurable, alternating between the CLM formula and the induction hypothesis gives:

$$\begin{aligned}
(*) &:= \mathbf{E}\left[\int_{((0,1]\times\mathbb{X})^{k+1}} f\left(\xi_n - \tilde{\delta}_{(u,x)_{k+1}}, (u,x)_{k+1}\right)\xi^{(k+1)}_n(d(u,x)_{k+1})\right] \\
&= \mathbf{E}\left[\int_{((0,1]\times\mathbb{X})^k}\int_{(0,1]\times\mathbb{X}} f\left(\xi_n - \tilde{\delta}_{(u,x)_{k+1}}, (u,x)_{k+1}\right)\left(\xi_n - \tilde{\delta}_{(u,x)_k}\right)(d(u_{k+1},x_{k+1}))\xi^{(k)}_n(d(u,x)_k)\right] \\
&= \int_{((0,1]\times\mathbb{X})^k}\int_{\mathbb{M}^d((0,1]\times\mathbb{X})}\int_{(0,1]\times\mathbb{X}} f\left(\phi - \delta_{(u_{k+1},x_{k+1})}, (u,x)_{k+1}\right)\phi(d(u_{k+1},x_{k+1}))\mathcal{P}^{n,(k)}_{(u,x)_k}\mu^{(k)}_n(d(u,x)_k) \\
&= \int_{((0,1]\times\mathbb{X})^k} \mathbf{E}\left[\int_{(0,1]\times\mathbb{X}} f\left((1-\tilde{u}_k)\cdot\xi_{n+k} - \delta_{(u_{k+1},x_{k+1})}, (u,x)_{k+1}\right)\left((1-\tilde{u}_k)\cdot\xi_{n+k}\right)(d(u_{k+1},x_{k+1}))\right]\mu^{(k)}_n(d(u,x)_k) \\
&= \int_{((0,1]\times\mathbb{X})^k} \mathbf{E}\left[\int_{(0,1]\times\mathbb{X}} f\left((1-\tilde{u}_k)\cdot\xi_{n+k} - \delta_{((1-\tilde{u}_k)u_{k+1},x_{k+1})}, (u,x)_k, (1-\tilde{u}_k)u_{k+1}, x_{k+1}\right)\xi_{n+k}(d(u_{k+1},x_{k+1}))\right]\mu^{(k)}_n(d(u,x)_k) \\
&= \int_{((0,1]\times\mathbb{X})^k} \mathbf{E}\left[\int_{(0,1]\times\mathbb{X}} f\left((1-\tilde{u}_k)\cdot\left(\xi_{n+k} - \delta_{(u_{k+1},x_{k+1})}\right), (u,x)_k, (1-\tilde{u}_k)u_{k+1}, x_{k+1}\right)\xi_{n+k}(d(u_{k+1},x_{k+1}))\right]\mu^{(k)}_n(d(u,x)_k) \\
&= \int_{((0,1]\times\mathbb{X})^k}\int_{(0,1]\times\mathbb{X}}\int_{\mathbb{M}^d((0,1]\times\mathbb{X})} f\left((1-\tilde{u}_k)\cdot\phi, (u,x)_k, (1-\tilde{u}_k)u_{k+1}, x_{k+1}\right)\mathcal{P}^{n+k,(1)}_{u_{k+1},x_{k+1}}\mu_{n+k}(d(u_{k+1},x_{k+1}))\mu^{(k)}_n(d(u,x)_k) \\
&= \int_{((0,1]\times\mathbb{X})^k}\int_{(0,1]\times\mathbb{X}} \mathbf{E}\left[f\left((1-u_{k+1})(1-\tilde{u}_l)\cdot\xi_{n+k+1}, (u,x)_k, (1-\tilde{u}_k)u_{k+1}, x_{k+1}\right)\right]\mu_{n+k}(d(u_{k+1},x_{k+1}))\mu^{(k)}_n(d(u,x)_k)
\end{aligned}$$

On the other hand, by the CLM formula again, we have:

$$(*) = \int_{((0,1]\times\mathbb{X})^{k+1}}\int_{\mathbb{M}^d((0,1]\times\mathbb{X})} f(\phi,(u,x)_k,u_{k+1},x_{k+1})\mathcal{P}^{n,(k+1)}_{(u,x)_{k+1}}(d\phi)\mu^{(k+1)}_n(d(u,x)_{k+1})$$

Taking f constant on $\mathbb{M}^d((0,1]\times\mathbb{X})$ yields:

$$\begin{aligned}
&\int_{((0,1]\times\mathbb{X})^{k+1}} f((u,x)_k,u_{k+1},x_{k+1})\mu^{(k+1)}_n(d(u,x)_{k+1}) \\
&= \int_{((0,1]\times\mathbb{X})^k}\int_{(0,1]\times\mathbb{X}} f\left((u,x)_k,(1-\tilde{u}_k)u_{k+1},x_{k+1}\right)\mu_{n+k}(d(u_{k+1},x_{k+1}))\mu^{(k)}_n(d(u,x)_k)
\end{aligned}$$

We deduce:

$$(*) = \int_{((0,1]\times\mathbb{X})^{k+1}} \mathbf{E}\left[f\left((1-(1-\tilde{u}_k)^{-1}u_{k+1})(1-\tilde{u}_l)\cdot\xi_{n+k+1},(u,x)_k,u_{k+1},x_{k+1}\right)\right]\mu_n^{(k+1)}(d(u,x)_{k+1})$$
$$= \int_{((0,1]\times\mathbb{X})^{k+1}} \mathbf{E}\left[f\left((1-\tilde{u}_{k+1})\cdot\xi_{n+k+1},(u,x)_{k+1}\right)\right]\mu_n^{(k+1)}(d(u,x)_{k+1})$$

By identification, we obtain Equation (7) for $k+1$.
Let now $((P_k^n,X_k^n))_{k\in\mathbb{N}^*}$ be an ISBP numbering of ξ_n and $(V_k^n)_{k\in\mathbb{N}^*}$ the sequence of residual allocations associated to $P^n := (P_k^n)_{k\in\mathbb{N}^*}$. We want to prove that P^n is a GEM and that $(X_k^n)_{k\in\mathbb{N}^*}$ is a sequence of i.i.d. random variables independent of P^n. Again, let $k\in\mathbb{N}^*$. We define the function T_k^n as follows:

$$T_k : \{u\in(0,1]^k : \tilde{u}_i<1\,\forall i\in[\![k-1]\!]\}\to\mathbb{R}_+^*$$
$$(u_1,\dots,u_k)\mapsto\prod_{i=1}^k\left(1-\sum_{j=1}^{i-1}u_j\right)^{-1}u_i$$

For all $f:(0,1]\times\mathbb{X}\to\mathbb{R}_+$ and $g:\mathbb{M}^d((0,1]\times\mathbb{X})\to\mathbb{R}_+$ measurable, we have:

$$(\star) := \mathbf{E}\left[f(V_k^n,X_k^n)g(\xi_n^{[k]})\right]$$
$$= \mathbf{E}\left[f\left(\left(1-\sum_{i=1}^{k-1}P_i^n\right)^{-1}P_k^n,X_k^n\right)g\left(\left(1-\sum_{i=1}^{k}P_i^n\right)^{-1}\cdot\left(\xi_n-\sum_{i=1}^k\delta_{(P_i^n,X_i^n)}\right)\right)\right]$$

Forcing a conditional expectation into the expectation yields:

$$(\star) = \mathbf{E}\left[\int_{((0,1]\times\mathbb{X})^k} T_k(u)f\left((1-\tilde{u}_{k-1})^{-1}u_k,x_k\right)g\left((1-\tilde{u}_k)^{-1}\cdot(\xi_n-\tilde{\delta}_{(u,x)_k}\right)\xi_n^{(k)}(d(u,x)_k)\right]$$

By the CLM formula, this gives:

$$(\star) = \int_{((0,1]\times\mathbb{X})^k}\int_{\mathbb{M}^d((0,1]\times\mathbb{X})} T_k(u)f\left((1-\tilde{u}_{k-1})^{-1}u_k,x_k\right)g\left((1-\tilde{u}_k)^{-1}\cdot\phi\right)\mathcal{P}_{(u,x)_k}^{n,(k)}(d\phi)\mu_n^{(k)}(d(u,x)_k)$$

Applying Equation (7):

$$(\star) = \int_{((0,1]\times\mathbb{X})^k}\mathbf{E}\left[T_k(u)f\left((1-\tilde{u}_{k-1})^{-1}u_k,x_k\right)g\left(\xi_{n+k}\right)\right]\mu_n^{(k)}(d(u,x)_k)$$
$$= \mathbf{E}\left[\int_{((0,1]\times\mathbb{X})^k} T_k(u)f\left((1-\tilde{u}_{k-1})^{-1}u_k,x_k\right)g\left(\xi_{n+k}\right)\xi_n^{(k)}(d(u,x)_k)\right]$$

Now, (without loss of generality again) one can assume ξ_{n+k} and ξ_n to be independent. Hence, we obtain:

$$(\star) = \mathbf{E}\left[\int_{((0,1]\times\mathbb{X})^k} T_k(u)f\left((1-\tilde{u}_{k-1})^{-1}u_k,x_k\right)\xi_n^{(k)}(d(u,x)_k)\right]\mathbf{E}\left[g(\xi_{n+k})\right]$$
$$= \mathbf{E}\left[f(V_k^n,X_k^n)\right]\mathbf{E}\left[g(\xi_{n+k})\right]$$

Moreover, repeating the same arguments, we have:

$$\mathbf{E}\left[g(\xi_n^{[k]})\right] = \mathbf{E}\left[g\left(\left(1-\sum_{i=}^k U_i^n\right)^{-1}\cdot\left(\xi_n-\sum_{i=1}^k\delta_{(U_i^n,X_i^n)}\right)\right)\right]$$
$$= \mathbf{E}\left[\int_{((0,1]\times\mathbb{X})^k} T_k(u)g\left((1-\tilde{u}_k)^{-1}\cdot\left(\xi_n-\tilde{\delta}_{(u,x)_k}\right)\right)\xi_n^{(k)}(d(u,x)_k)\right]$$
$$= \int_{((0,1]\times\mathbb{X})^k}\int_{\mathbb{M}^d((0,1]\times\mathbb{X}} T_k(u)g\left((1-\tilde{u}_k)^{-1}\cdot\phi_n\right)\mathcal{P}_{(u,x)_k}^{n,(k)}(d\phi)\mu_n^{(k)}(d(u,x)_k)$$
$$= \int_{((0,1]\times\mathbb{X})^k}\mathbf{E}\left[T_k(u)g(\xi_{n+k})\right]\mu_n^{(k)}(d(u,x)_k)$$
$$= \mathbf{E}\left[\int_{((0,1]\times\mathbb{X})^k} T_k(u)\xi_n^{(k)}(d(u,x)_k)\right]\mathbf{E}\left[g(\xi_{n+k})\right]$$
$$= \mathbf{E}\left[g(\xi_{n+k})\right]$$

Therefore, (V_k^n, X_k^n) and ξ_k^n are independent. By Theorem 2.2.8, we deduce that P^n is an ISBP RAM and that there exists a probability Q_n on $\mathbb{X}$ such that $(X_k^n)_{k\in\mathbb{N}^*}$ is a sequence of i.i.d. Q_n-distributed random variables independent of P^n. Moreover, since ξ_n has $\mathbf{P}$-a.s. infinite support, $K(P^n) = +\infty$ $\mathbf{P}$-a.s. and, by Theorem 2.2.5, there exist $(\alpha, \theta) \in \Xi$ such that P^n is a GEM with parameters (α, θ). Since, for $k \in \mathbb{N}^*$, we have proven that ξ_{n+k} has the same law than $\xi_n^{[k]}$, we deduce, by a recursive application of Corollary 2.2.9, that ξ_{n+k} is a GEM with parameters $(\alpha, \theta + k\alpha)$. □

4 Proof of the characterization

This section is dedicated to the proof of this paper's main result, Theorem 1.2.1. The first subsection proves the implication where one starts with a PY process, the second one its reverse. Although the first one is denoted as the "easy one" because it relies solely on existing results, the proof of the second one is quite straightforward as we have already done most of the work in Sections 2 and 3.

4.1 The easy direction

Let ζ be PY process with parameters $(\alpha, \theta) \in \Xi$ and base measure ν. We want to prove that ζ satisfies Equation (2) for some appropriate L and G.

Notations

In this subsection, we keep the same notation than in Subsection 3.1. In particular, for $(\alpha', \theta') \in \Xi$, $\Lambda_{\alpha'}$ and $\Lambda^{\theta'}$ are measures on $\mathbb{R}_+^*$ given by:

$$\Lambda_{\alpha'}(dt) := t^{-1-\alpha'}dt \qquad\qquad \Lambda^{\theta'}(dt) := \frac{\theta'}{t}e^{-t}dt$$

Moreover, we extend the notations the notations Σ and $L_{\alpha'}$ (for $\alpha' \in (0,1)$) to $\mathbb{M}^d(I \times \mathbb{X})$ (where I is $\mathbb{R}_+^*$ or any interval of the form $(0, a]$ with $a \in \mathbb{R}_+^*$), in the following sense. For two point processes ξ and Φ on $I \times \mathbb{X}$ and I (respectively) defined by

$$\xi := \sum_{n\in\mathbb{N}^*} \delta_{(U_n, X_n)} \qquad\qquad \Phi := \sum_{n\in\mathbb{N}^*} \delta_{U_n}$$

for some sequence $((U_n, X_n))_{n\in\mathbb{N}^*}$ of $(I \times \mathbb{X})$-valued random variables such that $\sum_{n\in\mathbb{N}^*} P_n$ is $\mathbf{P}$-a.s. finite, we write:

$$\Sigma(\xi) := \Sigma(\Phi) := \sum_{n\in\mathbb{N}^*} U_n$$

Equivalently, for such processes and for $\alpha' \in (0,1)$, we write:

$$L_{\alpha'}(\xi) := L_{\alpha'}(\Phi) := \lim_{n\to\infty} nP_{(n)}^{\alpha'}$$

where $(P_{(n)})_{n\in\mathbb{N}^*}$ is the sequence $(P_n)_{n\in\mathbb{N}^*}$ arranged in decreasing order, provided the limit exists.

Appropriate construction of the process

For each $(\alpha', \theta') \in \Xi$, let $\xi_{\alpha'}$ and $\xi^{\theta'}$ be point processes on $\mathbb{R}_+^* \times \mathbb{X}$ with respective intensities $\Lambda_{\alpha'} \otimes \nu$ and $\Lambda^{\theta'} \otimes \nu$, where $\nu \in \mathbb{M}_1(\mathbb{X})$ is diffuse. Moreover, for all $(\alpha', \theta') \in \Xi$, define the point process $\xi_{(\alpha',\theta')}$ on $\mathbb{X}$ by its distribution:

$$\mathbf{E}\left[f(\xi_{(\alpha',\theta')})\right] = \begin{cases} \mathbf{E}\left[f(\xi_{\alpha'})\right], & \text{if } \theta' = 0 \\ \mathbf{E}\left[f(\xi^{\theta'})\right], & \text{if } \alpha' = 0 \\ \mathbf{E}\left[\Sigma(\xi_{\alpha'})^{-\theta'}\right]^{-1} \mathbf{E}\left[\Sigma(\xi_{\alpha'})^{-\theta'} f(\xi_{\alpha'})\right], & \text{otherwise} \end{cases}$$

for all $f : \mathbb{M}^d(\mathbb{X}) \to \mathbb{R}_+$ measurable. For all (α', θ'), let now define the point process $\Gamma_{(\alpha',\theta')}$ on $\mathbb{X}$ given by:

$$\int_{\mathbb{X}} f(x)\Gamma_{(\alpha',\theta')}(dx) := \int_{(0,1]\times\mathbb{X}} uf(x)\xi_{(\alpha',\theta')}(d(u,x))$$

for all $f : \mathbb{M}^d(\mathbb{X}) \to \mathbb{R}_+$ measurable. By Proposition 3.1.4, we know that, for all $(\alpha', \theta') \in \Xi$, $\Gamma_{(\alpha',\theta')}(\mathbb{X})^{-1}\Gamma_{(\alpha',\theta')}$ is a PY process with base measure ν and parameters (α', θ').

Computations

If $\alpha = 0$, then, by Theorem 1.1.4, Equation (2) holds true with $L \equiv 1$ and $G = Beta(1, \theta)$. Let thus assume that $\alpha > 0$. For $(\alpha', \theta') \in \mathbb{R}_+^*$, we will write:

$$c_{(\alpha',\theta')} := \mathbf{E}\left[\Gamma_{(\alpha',0)}(\mathbb{X})^{-\theta'}\right]^{-1} = \mathbf{E}\left[L_{\alpha'}(\xi_{\alpha'})^{\theta'/\alpha'}\right]^{-1}$$

Without loss of generality, we can assume that $\zeta = \Gamma_{(\alpha,\theta)}(\mathbb{X})^{-1}\Gamma_{(\alpha,\theta)}$ $\mathbf{P}$-a.s. and, for the sake of a non-unreadable notation, we will write:

$$\Gamma := \Gamma_{(\alpha,0)}$$

For all measurable $f : \mathbb{M}^d(\mathbb{X}) \times \mathbb{X} \to \mathbb{R}_+$, we then have:

$$\begin{aligned}(*) &= \mathbf{E}\left[\int_{\mathbb{X}} f(\zeta, x)\zeta(dx)\right] \\ &= c_{(\alpha,\theta)}\mathbf{E}\left[\Gamma(\mathbb{X})^{-\theta}\int_{\mathbb{X}} f\left(\Gamma(\mathbb{X})^{-1}\Gamma, x\right)\Gamma(\mathbb{X})^{-1}\Gamma(dx)\right] \\ &= c_{(\alpha,\theta)}\mathbf{E}\left[\int_{\mathbb{R}_+^*\times\mathbb{X}} \Gamma(\mathbb{X})^{-1-\theta} r f\left(\Gamma(\mathbb{X})^{-1}\Gamma, x\right)\xi_\alpha(d(r,x))\right]\end{aligned}$$

Now, Γ is a measurable function of ξ_α. Hence, by the CLM formula and Slivnyak theorem (the real one), we obtain:

$$(*) = c_{(\alpha,\theta)}\int_{\mathbb{R}_+^*\times\mathbb{X}} \mathbf{E}\left[(\Gamma + r\delta_x)(\mathbb{X})^{-1-\theta} r f\left((\Gamma + r\delta_x)(\mathbb{X})^{-1}(\Gamma + r\delta_x), x\right)\right](\Lambda_\alpha \otimes \nu)(d(r,x))$$

By Tonelli's and developing, this gives:

$$(*) = \mathbf{E}\left[\int_{\mathbb{X}}\int_{\mathbb{R}_+^*} (\Gamma(\mathbb{X}) + r)^{-1-\theta} f\left((\Gamma(\mathbb{X}) + r)(\Gamma + u\delta_x), x\right) r^{-\alpha} dr\nu(dx)\right]$$

Proceeding to the change of variable $u := (\Gamma(\mathbb{X}) + r)^{-1} r$ inside the integral yields:

$$\begin{aligned}(*) &= c_{(\alpha,\theta)}\mathbf{E}\left[\int_{\mathbb{X}}\int_{(0,1)} \frac{(1-u)^{1+\theta}}{\Gamma(\mathbb{X})^{1+\theta}} f\left(\frac{(1-u)}{\Gamma(\mathbb{X})}\Gamma + u\delta_x, x\right)\frac{(1-u)^\alpha}{u^\alpha\Gamma(\mathbb{X})^\alpha}\frac{\Gamma(\mathbb{X})}{(1-u)^2} du\nu(dx)\right] \\ &= c_{(\alpha,\theta)}\mathbf{E}\left[\int_{\mathbb{X}}\int_{(0,1)} \Gamma(\mathbb{X})^{-\theta-\alpha}(1-u)^{\alpha+\theta-1} f\left((1-u)\zeta + u\delta_x, x\right) u^{-\alpha} du\nu(dx)\right]\end{aligned}$$

By Tonelli's again:

$$\begin{aligned}(*) &= \int_{\mathbb{X}}\int_{(0,1)} c_{(\alpha,\theta)}(1-u)^{\alpha+\theta-1}u^{-\alpha}\mathbf{E}\left[\Gamma(\mathbb{X})^{-(\alpha+\theta)} f((1-u)\zeta + u\delta_x, x)\right] du\nu(dx) \qquad (8) \\ &= \int_{\mathbb{X}}\int_{(0,1)} (1-u)^{\alpha+\theta-1}u^{-\alpha}\mathbf{E}\left[\Gamma(\mathbb{X})^{-\alpha} f((1-u)\zeta + u\delta_x, x)\right] du\nu(dx)\end{aligned}$$

Writing $G(du) := (1-u)^{\alpha+\theta-1}u^{-\alpha}\mathbb{1}_{(0,1)}(du)$ and $L(\zeta) := L_\alpha(\Sigma(\xi_\alpha)^{-1}\cdot\xi_\alpha)$, we finally obtain:

$$(*) = \mathbf{E}\left[\int_{\mathbb{X}}\int_{\mathbb{R}} L(\zeta) f((1-u)\zeta + u\delta_x, x) G(du)\nu(dx)\right]$$

4.2 The other one

Let ζ be a $\mathbb{M}_1^d(\mathbb{X})$-valued random variable with $\mathbf{P}$-a.s. infinite support and such that, for all $f : \mathbb{M}^d(\mathbb{X}) \times \mathbb{X} \to \mathbb{R}_+$ measurable, we have:

$$\mathbf{E}\left[\int_{\mathbb{X}} f(\zeta, x)\zeta(dx)\right] = \mathbf{E}\left[\int_{\mathbb{X}}\int_{\mathbb{R}} L(\zeta) f((1-u)\zeta + u\delta_x, x) G(du)\nu(dx)\right] \qquad (2)$$

where $G \in \mathbb{M}(\mathbb{R})$ is finite, $\nu \in \mathbb{M}(\mathbb{X})$ is finite (in fact finiteness of G and ν is deduced from the equation) and diffuse and $L : \mathbb{M}^d(\mathbb{R}) \to \mathbb{R}_+$ is measurable and such that, for $(G \otimes \nu)$-a.a. $(u, x) \in \mathbb{R} \times \mathbb{X}$, we have $\mathbf{P}$-a.s.

$$L((1-u)\zeta + u\delta_x) = (1-u)^\alpha L(\zeta)$$

for some $\alpha \in [0, 1)$.
We want to prove that ζ is a PY process with base measure $\nu/\nu(\mathbb{X})$ and parameters (α, θ) where

$$\theta = \left(\nu(\mathbb{X})\mathbf{E}\left[L(\zeta)\right]\int_{(0,1]} uG(du)\right)^{-1}(1-\alpha) - 1$$

and that $\nu(\mathbb{X})\mathbf{E}\left[L(\zeta)\right]G$ is the Beta law with parameters $(1-\alpha, \theta+\alpha)$.

Notations

We define the function ι as follows:

$$\begin{aligned}\iota : \mathbb{M}^d(\mathbb{X}) &\to \mathbb{M}^d(\mathbb{R}_+^* \times \mathbb{X}) \\ \phi &\mapsto \sum_{x \in Supp(\phi)} \delta_{(\phi(\{x\}),x)}\end{aligned}$$

One can check that ι is bijective. Indeed, for $\phi \in \mathbb{M}^d(\mathbb{X})$ and $B \in \mathcal{X}$, we have:

$$\int_{\mathbb{R}_+^* \times B} u\iota(\phi)(d(u,x)) = \phi(B)$$

We will write ι^* the inverse of ι.
Moreover, for $\phi \in \mathbb{M}(I \times \mathbb{X})$ (where I is $\mathbb{R}_+^*$ or an interval of the form $(0,a]$ with $a \in \mathbb{R}_+^*$) and $x \in \mathbb{X}$, we will write:

$$h(\phi, x) := \int_{I \times \{x\}} u\phi(d(u,x))$$

Reformulation of the hypotheses

Consider the point process ξ on $(0,1] \times \mathbb{X}$ defined by:

$$\xi := \iota(\zeta)$$

It is clear that, **P**-a.s.:

$$\int_{(0,1]\times\mathbb{X}} u\xi(d(u,x)) = 1$$

Now, Equation (2) can be rewritten:

$$\mathbf{E}\left[\int_{(0,1]\times\mathbb{X}} uf(\iota^*(\xi),u)\xi(d(u,x))\right] = \mathbf{E}\left[\int_{\mathbb{X}}\int_{\mathbb{R}} \tilde{L}(\xi)f(\iota^*((1-u)\cdot\xi + \delta_{(u,x)}),x)G(du)\nu(dx)\right]$$

with $\tilde{L} := L \circ \iota^*$. For $(G \otimes \nu)$-a.a. $(u,x) \in (0,1] \times \mathbb{X}$, we have $\tilde{L}((1-u)\cdot\xi + \delta_{(u,x)}) = (1-u)^\alpha \tilde{L}(\xi)$. Since ι is bijective, this is equivalent to:

$$\mathbf{E}\left[\int_{(0,1]\times\mathbb{X}} uf(\xi,u)\xi(d(u,x))\right] = \mathbf{E}\left[\int_{\mathbb{X}}\int_{\mathbb{R}} \tilde{L}(\xi)f((1-u)\cdot\xi + \delta_{(u,x)},x)G(du)\nu(dx)\right] \tag{9}$$

for all $f : \mathbb{M}^d(\mathbb{R}\times\mathbb{X})\times\mathbb{X} \to \mathbb{R}_+$ measurable.

Application of the Slivnyak-like theorem

Applying Equation (9) to $f(\xi,u) := \tilde{L}(\xi)$ (for G-a.a. $u \in \mathbb{R}$) yields:

$$\mathbf{E}\left[\tilde{L}(\xi)\right] = \nu(\mathbb{X})\int_{\mathbb{R}} (1-u)^\alpha G(du)\mathbf{E}\left[\tilde{L}(\xi)^2\right]$$

We deduce that $\tilde{L}(\xi)$ has finite second moment. By induction, all its moments are finite. Let $(\xi_n)_{n\in\mathbb{N}}$ be a family of independent point processes on $(0,1]\times\mathbb{X}$ which distributions are given by:

$$\mathbf{E}\left[g(\xi_n)\right] = \frac{\mathbf{E}\left[\tilde{L}(\xi)^n g(\xi)\right]}{\mathbf{E}\left[\tilde{L}(\xi)^n\right]}$$

for all $n \in \mathbb{N}$ and $g : \mathbb{M}^d((0,1]\times\mathbb{X}) \to \mathbb{R}_+$ measurable. Let $n \in \mathbb{N}$. For all $f : \mathbb{M}^d(\mathbb{R}\times\mathbb{X})\times\mathbb{R}\times\mathbb{X} \to \mathbb{R}_+$ measurable, we have:

$$\begin{aligned}(*) &:= \mathbf{E}\left[\int_{(0,1]\times\mathbb{X}} f(\xi_n - \delta_{(u,x)},u,x)\xi_n(d(u,x))\right] \\ &= \frac{\mathbf{E}\left[\int_{(0,1]\times\mathbb{X}} \tilde{L}(\xi)^n f(\xi_n - \delta_{(u,x)},u,x)\xi_n(d(u,x))\right]}{\mathbf{E}\left[\tilde{L}(\xi)^n\right]}\end{aligned}$$

In order to apply Equation (9), we need to express f solely in terms of x and ξ. By definition of ξ, $\mathbf{P}$-a.s. for ξ-a.a. $(u,x) \in \mathbb{R}_+^* \times \mathbb{X}$, we have $h(\xi, x) = u$. Hence:

$$(*) = \frac{\mathbf{E}\left[\int_{(0,1]\times\mathbb{X}} uh(\xi,x)^{-1}\tilde{L}(\xi)^n f(\xi - \delta_{(h(\Phi,x),x)}, h(\xi,x), x)\xi(d(u,x))\right]}{\mathbf{E}\left[\tilde{L}(\xi)^n\right]}$$

Now, for all $(u,x) \in \mathbb{R}_+^* \times \mathbb{X}$, we have $\mathbf{P}$-a.s.:

$$\begin{aligned} h((1-u)\cdot\xi + \delta_{(u,x)}, x) &= \int_{\mathbb{R}\times\{x\}} v((1-u)\cdot\xi + \delta_{(u,x)})(d(v,z)) \\ &= \int_{\mathbb{R}\times\{x\}} (1-u)v\xi(d(v,z)) + \int_{\mathbb{R}\times\{x\}} v\delta_{(u,x)}(d(v,z)) \\ &= (1-u)h(\xi,x)\mathbb{1}_{Supp(\zeta)}(x) + u \end{aligned}$$

Since ζ is $\mathbf{P}$-a.s. discrete and ν is diffuse, we have $\nu(Supp(\zeta)) = 0$ $\mathbf{P}$-a.s.. Therefore, for $(G\otimes\nu)$-a.a. $(u,x) \in \mathbb{R}\times\mathbb{X}$, we have $\mathbf{P}$-a.s.:

$$h((1-u)\cdot\xi + \delta_{(u,x)}, x) = u$$

We deduce:

$$\begin{aligned} (*) &= \frac{\mathbf{E}\left[\int_{\mathbb{X}}\int_{\mathbb{R}} \tilde{L}(\xi)\tilde{L}((1-u)\cdot\xi + \delta_{(u,x)})^n f((1-u)\cdot\xi, u, x)G(du)\nu(dx)\right]}{\mathbf{E}\left[L(\xi)^n\right]} \\ &= \frac{\mathbf{E}\left[\int_{\mathbb{X}}\int_{\mathbb{R}} \tilde{L}(\xi)^{n+1}(1-u)^{n\alpha} f((1-u)\cdot\xi, u, x)G(du)\nu(dx)\right]}{\mathbf{E}\left[\tilde{L}(\xi)^n\right]} \end{aligned}$$

Taking f constant on $\mathbb{M}^d(\mathbb{R}\times\mathbb{X})$ gives us the mean measure μ_n of ξ_n:

$$\int_{(0,1]\times\mathbb{X}} f(u,x)\mu_n(d(u,x)) = \frac{\mathbf{E}\left[\tilde{L}(\xi)^{n+1}\right]}{\mathbf{E}\left[\tilde{L}(\xi)^n\right]} \int_{\mathbb{X}}\int_{\mathbb{R}} (1-u)^{n\alpha} f(u,x)G(du)\nu(dx) \tag{10}$$

Using Tonelli's and the CLM formula gives us:

$$\int_{\mathbb{M}^d((0,1]\times\mathbb{X})} g(\phi)\mathcal{P}^{n,(1)}_{(u,x)}(d\phi) = \frac{\mathbf{E}\left[\tilde{L}(\xi)^{n+1} g((1-u)\cdot\xi)\right]}{\mathbf{E}\left[\tilde{L}(\xi)^{n+1}\right]} = \mathbf{E}\left[g((1-u)\cdot\xi_{n+1})\right]$$

for μ_n-a.a. $(u,x) \in \mathbb{R}_+^* \times \mathbb{X}$, where $\mathcal{P}^{n,(1)}_{(u,x)}$ is the first order reduced Palm distribution of ξ_n at (u,x).

Applying Corollary 3.2.2, we deduce that there exist $(\alpha', \theta') \in \Xi$ such that, for all $n \in \mathbb{N}$ and any $ISBP$ numbering $((P_k^n, X_k^n))_{k\in\mathbb{N}^*}$ of ξ_n, $P^n := (P_k^n)_{k\in\mathbb{N}^*}$ is a GEM with parameters $(\alpha', \theta' + n\alpha')$ and there exists a probability Q_n on $\mathbb{X}$ such that $(X_k^n)_{k\in\mathbb{N}^*}$ is a sequence of i.i.d. Q_n-distributed random variables independent of P^n. In particular, observing that ξ_0 and ξ have the same law, we deduce that ζ is a PY process with parameters (α', θ') and base measure Q_0.

Identification of the parameters

First of all, by definition and for all $n \in \mathbb{N}$, the intensity measure μ_n of ξ_n has support $(0,1]\times\mathbb{X}$. By Equation (10), we deduce that G has support $(0,1]$. Moreover, applying Equation (2) to $f \equiv 1$ gives us:

$$1 = \mathbf{E}\left[L(\zeta)\right] G((0,1])\nu(\mathbb{X})$$

We deduce that $\nu(\mathbb{X})\mathbf{E}\left[L(\zeta)\right] G$ is a probability on $(0,1]$. Now, the intensity measure μ of ζ is given by:

$$\mathbf{E}\left[\zeta(B)\right] = \mathbf{E}\left[L(\zeta)\right] G((0,1])\nu(B) = \nu(\mathbb{X})^{-1}\nu(B)$$

Since the intensity measure of a PY process is its base measure, we deduce that $Q_0 = \nu(\mathbb{X})^{-1}\nu$.
As in Proposition 2.2.9, let τ be a random element of $\mathbb{X}$ such that:

$$\mathbf{P}(\tau \in B|\zeta) = \zeta(B)$$

for all $B \in \mathcal{X}$. τ can be identified with the second component of the first element of an ISBP numbering of ξ. As such, from the definition of a PY process (and of a GEM), we know that $\zeta(\{\tau\})$ follows the law $Beta(1-\alpha', \theta'+\alpha')$. Moreover, using Equation (2) again for all $f : (0,1] \to \mathbb{R}_+$ measurable yields:

$$\begin{aligned}
\mathbf{E}\left[f(\zeta(\{\tau\}))\right] &= \mathbf{E}\left[\int_{\mathbb{X}} f(\zeta(\{x\}))\zeta(dx)\right] \\
&= \mathbf{E}\left[\int_{\mathbb{X}}\int_{(0,1]} L(\zeta) f\left(((1-u)\zeta + u\delta_x)\left(\{x\}\right)\right) G(du)\nu(dx)\right] \\
&= \mathbf{E}\left[L(\zeta)\right]\nu(\mathbb{X})\int_{(0,1]} f(u)G(du)
\end{aligned}$$

We deduce that $\nu(\mathbb{X})\mathbf{E}\left[L(\zeta)\right]G = Beta(1-\alpha', \theta'+\alpha')$. Computing its first moment gives:

$$\nu(\mathbb{X})\mathbf{E}\left[L(\zeta)\right]\int_{(0,1]} uG(du) = \frac{1-\alpha'}{1+\theta'}$$

That is:

$$\theta' = \left(\nu(\mathbb{X})\mathbf{E}\left[L(\zeta)\right]\int_{(0,1]} uG(du)\right)^{-1}(1-\alpha') - 1$$

It remains to identify α' and L. By the reverse of Theorem 1.2.1, already proven in Subsection 4.1, we know that equation (2) is satisfied for some $L' : \mathbb{M}^d(\mathbb{X}) \to \mathbb{R}_+$ measurable such that $\mathbf{P}$-a.s. and for $(G\otimes\nu)$-a.a. $(u,x) \in (0,1]\times\mathbb{X}$, $L'((1-u)\zeta + u\delta_x) = (1-u)^{\alpha'}L'(\zeta)$. That is, for all $f : \mathbb{M}^d(\mathbb{X}) \times \mathbb{X}\mathbb{R}_+$ measurable, we have:

$$\mathbf{E}\left[\int_{\mathbb{X}} f(\zeta, x)\zeta(dx)\right] = \mathbf{E}\left[\int_{\mathbb{X}}\int_{(0,1]} L'(\zeta) f((1-u)\zeta + u\delta_x, x)G(du)\nu(dx)\right] \tag{11}$$

Equaling Equations (2) and (11) with $f(\phi, x) = g((1-h(\phi,x))^{-1}\phi - h(\phi,x)\delta_x)$ for some $g : \mathbb{M}^d(\mathbb{X}) \to \mathbb{R}_+$ measurable and all $(\phi, x) \in \mathbb{M}^d(\mathbb{X}) \times \mathbb{X}$ yields:

$$\mathbf{E}\left[L(\zeta)g(\zeta)\right] = \mathbf{E}\left[L'(\zeta)g(\zeta)\right]$$

By density, we deduce that $L(\zeta) = L'(\zeta)$ $\mathbf{P}$-a.s. and so $\alpha' = \alpha$.

Appendices

A Reminder on point processes

We here give some basic definitions and results about point processes and, more specifically, Poisson point processes. For an excellent and illustrated introduction to the subject, we recommend [5] (Chapters 7-12).

General theory

We recall that the space of discrete Radon measures $\mathbb{M}^d(\mathbb{X})$ on $\mathbb{X}$ can be equipped with the coarsest topology that makes the evaluation maps $\mathbb{M}^d(\mathbb{X}) \to \mathbb{R}_+, \mu \mapsto \mu(A)$ (for $A \in \mathcal{X}$) continuous. Under this topology, we write $\mathcal{M}$ the Borel σ-algebra of $\mathbb{M}^d(\mathbb{X})$.

Definition A.1 (Point process)**.** *We call point process on $\mathbb{X}$ any random variable $\Phi : \Omega \to \mathbb{M}^d(\mathbb{X})$.*

Remark. *The respective masses of the atoms of a point process are referred to as its weights.*

Although this definition is not universal (they are sometimes defined as random collections of points), it is the most frequent in the literature and by far the most mathematically fruitful. Its main advantage is to allow one to use topological measure theory to study their properties. In the same idea that the points collection, the weights of point processes (as random discrete measures) are often required to be positive integers. We decided, in this paper, not to restrict the definition to this particular case as discrete random measures are a central element (of the paper) and it is sometimes more convenient to simply refer to them as point processes.

Remark. *A point process which every atom has mass 1 is called simple. In most cases, point processes are assumed to be simple as one can always recover a simple one by treating weights as marks (i.e. if the original process is on $\mathbb{X}$, one can obtain a simple process on $\mathbb{R}_+^* \times \mathbb{X}$).*

Definition A.2 (Intensity measure)**.** *Given a point process Φ on $\mathbb{X}$, we call intensity measure (or mean measure) of Φ the unique measure μ on $\mathbb{X}$ such that, for any $B \in \mathcal{X}$, we have:*

$$\mu(B) = \mathbf{E}\left[\Phi(B)\right]$$

Theorem A.3 (Campbell)**.** *Let Φ be a point process on $\mathbb{X}$ with intensity measure $\mu \in \mathbb{M}(\mathbb{X})$ Then, for every $f : \mathbb{X} \to \mathbb{R}_+$ measurable, we have:*

$$\mathbf{E}\left[\int_{\mathbb{X}} f(x)\Phi(dx)\right] = \int_{\mathbb{X}} f(x)\mu(dx)$$

Proof. See Theorem 7.1.5 of [5]. □

Let $n \in \mathbb{N}^*$ a positive integer and equip $\mathbb{X}^n$ with its natural product σ-algebra $\mathcal{X}^{\otimes n}$.

Definition A.4 (Factorial powers)**.** *Given a point process Φ on $\mathbb{X}$ with ($\mathbf{P}$-a.s.) at least n atoms, we call n^{th} factorial power of Φ the point process $\Phi^{(n)}$ on $\mathbb{X}^n$ defined by:*

$$\Phi^{(n)} := \sum_{\substack{(x_1,\dots,x_n)\in(Supp(\Phi))^n \\ \forall i\neq j : x_i \neq x_j}} \left(\prod_{i=1}^{n} \Phi(\{x_i\})\right) \delta_{(x_1,\dots,x_n)}$$

Remark. *The first power of a point process is the point process itself.*

Definition A.5 (Factorial moment measures)**.** *Given a point process Φ on $\mathbb{X}$ with ($\mathbf{P}$-a.s.) at least n atoms, we call n^{th} factorial moment measure of Φ the unique measure $\mu^{(n)}$ on $\mathbb{X}^n$ such that, for any measurable function $f : \mathbb{X}^n \to \mathbb{R}_+$, we have:*

$$\mathbf{E}\left[\int_{\mathbb{X}^n} f(x)\Phi^{(n)}(dx)\right] = \int_{\mathbb{X}^n} f(x)\mu^{(n)}(dx)$$

where $\Phi^{(n)}$ denotes the n^{th} factorial power of Φ.

Remark. *The first moment measure of a point process is its intensity measure.*

Definition A.6 (Reduced Campbell measures)**.** *Given a point process Φ on $\mathbb{X}$ with ($\mathbf{P}$-a.s.) at least n atoms, we call n^{th} order reduced Campbell measure of Φ the unique measure $C^{(n)}$ on $\left(\mathbb{X}^n \times \mathbb{M}^d(\mathbb{X}), \mathcal{X}^{\otimes n} \otimes \mathcal{M}\right)$ such that, for any rectangle $B \times \Gamma \in \mathcal{X}^{\otimes n} \otimes \mathcal{M}$, we have:*

$$C^{(n)}(B \times \Gamma) = \mathbf{E}\left[\int_{\mathbb{X}^n} \mathbb{1}_B(x)\mathbb{1}_\Gamma(\Phi - \tilde{\delta}_x)\Phi^{(n)}(dx)\right]$$

Definition A.7 (Reduced Palm distributions)**.** *Let Φ be a point process on $\mathbb{X}$ with ($\mathbf{P}$-a.s.) at least n atoms and let $C^{(n)}$ and $\mu^{(n)}$ be respectively its n^{th} order reduced Campbell measure and n^{th} factorial moment. For $\Gamma \in \mathcal{M}$ and $x \in Supp(\mu^{(n)})$, we note $\mathcal{P}_x^{(n)}(\Gamma)$ the image of x under the Radon-Nikodym derivative of the measure $\mathcal{X}^{\otimes n} \to \mathbb{R}_+, B \mapsto C^{(n)}(B \times \Gamma)$ by $\mu^{(n)}$. For $\mu^{(n)}$-a.a. $x \in \mathbb{X}^n$, we call n^{th} order reduced Palm distribution of Φ at x the probability measure $\mathcal{M} \to [0,1], \Gamma \mapsto \mathcal{P}_x^{(n)}(\Gamma)$.*

Remark. *Reduced Palm distributions are to be understood as size-biased distributions. Indeed, if $x := (x_1, \dots, x_n) \in \mathbb{X}^n$ is a vector of fixed atoms of a point process Φ on $\mathbb{X}$ (meaning $\mathbf{P}(\Phi(\{x_i\}) > 0) > 0$ for each $i \in [\![n]\!]$), we have, for $\Gamma \in \mathcal{M}$:*

$$\mathcal{P}_x^{(n)}(\Gamma) = \frac{\mathbf{E}\left[\mathbb{1}_\Gamma(\Phi - \tilde{\delta}_x)\prod_{i=1}^n \Phi(\{x_i\})\right]}{\mathbf{E}\left[\prod_{i=1}^n \Phi(\{x_i\})\right]}$$

In case Φ is simple, it is consistent with the distribution of $\Phi - \tilde{\delta}_x$ conditioned on $x_1, \dots, x_n$ being atoms. As previously mentioned, this latter interpretation is the most common one.

Theorem A.8 (Campbell-Little-Mecke)**.** *Let Φ be a point process on $\mathbb{X}$ with ($\mathbf{P}$-a.s.) at least n atoms and denote respectively by $(\mathcal{P}_x^{(n)})_{x \in Supp(\mu^{(n)})}$ and $\mu^{(n)}$ its family of n^{th} order reduced Palm distributions and its n^{th} factorial moment. The following holds true for any measurable $f : \mathbb{X}^n \times \mathbb{M}^d(\mathbb{X}) \to \mathbb{R}_+$:*

$$\mathbf{E}\left[\int_{\mathbb{X}^n} f(x, \Phi - \tilde{\delta}_x)\Phi^{(n)}(dx)\right] = \int_{\mathbb{X}^n}\int_{\mathbb{M}^d(\mathbb{X})} f(x,\phi)\,\mathcal{P}_x^{(n)}(d\phi)\mu^{(n)}(dx)$$

Proof. Follows directly from the definition of Palm distributions and classical measure theoretic arguments. □

Poisson processes

Poisson point processes are one of the most simple and studied class of point processes. One of the reason of its importance in probability theory is its connection with Markov processes (and more specifically Lévy processes).

Definition A.9 (Poisson point process)**.** *Let Φ be a point process on $\mathbb{X}$ with intensity measure μ. We say that Φ is a Poisson point process (PPP) if the following two conditions are satisfied:*
- For every $k \in \mathbb{N}^*$ and $B_1, \dots, B_k \in \mathcal{X}$ pairwise disjoint, $\Phi(B_1), \dots, \Phi(B_k)$ are mutually independent.
- For every $B \in \mathcal{X}$, $\Phi(B)$ follows a Poisson distribution with parameter $\mu(B)$.

In fact, one can show that, if a simple point process with no fixed atom satisfies the independence property of Definition A.9, it is a PPP.

Theorem A.10 (Slivnyak)**.** *A point process Φ is a PPP if and only if its distribution and its reduced first order Palm distributions are equal.*

Proof. See Theorem 8.1.8 of [5]. □

Definition A.10.1 (Random transformation)**.** *Let $\mathbb{X}'$ be a Polish space equipped with its Borel σ-algebra $\mathcal{X}'$ and $p : \mathbb{X} \times \mathcal{X}' \to [0,1]$ a random probability kernel. The random transformation Φ^p of a point process Φ on $\mathbb{X}$ by p is the point process on $\mathbb{X}'$ given by:*

$$\Phi^p = \sum_{X \in Supp(\Phi)} \Phi(\{X\})\delta_{Y_X}$$

where, for $X \in Supp(\Phi)$, Y_X is a random variable which distribution given Φ is given by $\mathbf{P}(Y_X \in B' | \Phi) = p(X, B')$ for every $B' \in \mathcal{X}'$.

Theorem A.10.2. *The transformation of a Poisson point process on $\mathbb{X}$ with intensity $\Lambda \in \mathbb{M}(\mathbb{X})$ by a random probability kernel $p : \mathbb{X} \times \mathcal{X}' \to [0,1]$ (for some $(\mathbb{X}', \mathcal{X}')$ as in Definition A.10.1) is a Poisson point process on $\mathbb{X}'$ with intensity $\Lambda' \in \mathbb{M}(\mathbb{X}')$ given by:*

$$\Lambda'(B') = \int_{\mathbb{X}} p(x, B')\Lambda(dx)$$

for all $B' \in \mathcal{X}$.

B Proof of Theorem 2.2.8

We hereby present the proof of Theorem 2.2.8 due to Thomas Lehéricy in his 2015 master's thesis and as presented in the latter. As it is quite a long one, it will be divided in different parts.

First, all implications but one are almost immediate:
$(i) \implies (ii)$ is trivial.
$(ii) \implies (i)$ can be proven in the exact same way (by induction) than in Proposition 2.2.6.
$(iii) \implies (i)$: if P is a RAM then $(V_n)_{n\in\mathbb{N}^*}$ is independent (we remind that $K(P)$ is $\mathbf{P}$-a.s. constant by Theorem 2.2.5). Since $(X_n)_{n\in\mathbb{N}^*}$ is a sequence of i.i.d. random variables independent of P (and, as a consequence, of $(V_n)_{n\in\mathbb{N}^*}$), we deduce (i).

The difficulty relies on the proof of $(i) \implies (iii)$. Before undertaking such work, we have to introduce a lemma (also due to T. Lehéricy), which is the subject of the next "subsection".

Invariance under finite random permutations

Notation. *For $n \in \mathbb{N}^*$, we denote $\mathfrak{S}_n$ the group of permutations on $[\![n]\!]$.*

Lemma B.1. *Let $U := (U_n)_{n\in\mathbb{N}^*} := ((P_n, X_n))_{n\in\mathbb{N}^*}$ be an ISBP numbering of a point process on $(0,1]\times\mathbb{X}$ ($\sum_{n\in\mathbb{N}^*} P_n = 1$ $\mathbf{P}$-a.s.) such that $K(P)$ (with $P := (P_n)_{n\in\mathbb{N}^*}$) is $\mathbf{P}$-a.s. constant. Let $k \in \mathbb{N}^*$ such that $2 \le k \le K(P)$, $\mathfrak{S}$ a subgroup of $\mathfrak{S}_n$ and $U' := (U'_n)_{n\in\mathbb{N}^*}$ a $([0,1]\times\mathbb{X})^{\mathbb{N}^*}$-valued random variable such that, given U, $U'_n = U_n$ $\mathbf{P}$-a.s. for all $n \in \mathbb{N}^*$ with $n > k$, and, for all $\sigma \in \mathfrak{S}_k$, we have:*

$$\mathbf{P}\left((U'_1,\dots,U'_k) = (U_{\sigma(1)},\dots,U_{\sigma(k)})|P\right) = T(\sigma)$$

where

$$g(\sigma) := \prod_{i=1}^{k} \frac{A_{\sigma(i)}P_{\sigma(i)}}{1-\sum_{j=1}^{i-1}P_{\sigma(j)}}$$
$$T(\sigma) := \frac{\mathbb{1}_{\mathfrak{S}}(\sigma)g(\sigma)}{\sum_{\sigma'\in\mathfrak{S}} g(\sigma')}$$

with $A_i = |\{n \in \mathbb{N}^* : (n \ge i) \wedge (P_n = P_i)\}|$.
Then, U' is ISBP and has the same law than U. In particular, if $\sigma := (k, k-1)$ (the transposition of k wih $k-1$) and $\mathfrak{S} := \{\sigma, 1_{\mathfrak{S}_k}\}$ (with $1_{\mathfrak{S}_k}$ the identity element of $\mathfrak{S}_k$), then for all $f : [0,1] \to \mathbb{R}_+$ measurable, we have:

$$\mathbf{E}\left[f(U_1,\dots,U_{k-2},U_{k-1},U_k)T(\sigma)\right] = \mathbf{E}\left[f(U_1,\dots,U_{k-2},U_k,U_{k-1})T(\sigma)\left[\sum_{i=1}^{k-1}P_i < 1\right]\right]$$

with

$$T(\sigma) = \frac{\left(1-P_k-\sum_{i=1}^{k-2}P_i\right)^{-1}A_{k-1}}{\left(1-P_k-\sum_{i=1}^{k-2}P_i\right)^{-1}A_{k-1} + \left(1-P_{k-1}-\sum_{i=1}^{k-2}\right)^{-1}A_k}$$

Proof. Without loss of generality, we will assume that U is a probability vector (we ignore the second components of its elements).
Let $k \in \mathbb{N}^*$ with $k \ge 2$, $\mathfrak{S}$ a subgroup of $\mathfrak{S}_k$ and $\hat{U} := (\hat{U}_n)_{n\in\mathbb{N}^*} := (U_{(n)})_{n\in\mathbb{N}^*}$ (the sequence U arranged in decreasing order). We have:

$$\mathbf{P}\left((U'_1,\dots,U'_k) = (\hat{U}_{\sigma(1)},\dots,\hat{U}_{\sigma(k)})|\hat{U}\right) = \sum_{\gamma\in\mathfrak{S}}\mathbf{P}\left((U'_1,\dots,U'_k) = (U_{\gamma(1)},\dots,U_{\gamma(k)}),(U_1,\dots,U_k) = (\hat{U}_{\sigma\gamma^{-1}(1)},\dots,\hat{U}_{\sigma\gamma^{-1}(k)})|\hat{U}\right)$$

Since $\hat{U}$ is a measurable function of U, we have:

$$\mathbf{P}\left((U'_1,\dots,U'_k) = (U_{\gamma(1)},\dots,U_{\gamma(k)})|\hat{U},U\right) = \mathbf{P}\left((U'_1,\dots,U'_k) = (U_{\gamma(1)},\dots,U_{\gamma(k)})|U\right) = T(\sigma)$$

Therefore:

$$\begin{aligned}\mathbf{P}\left((U'_1,\dots,U'_k) = (\hat{U}_{\sigma(1)},\dots,\hat{U}_{\sigma(k)})|\hat{U}\right) &= \sum_{\gamma\in\mathfrak{S}}\mathbf{E}\left[\frac{g(\gamma)}{\sum_{\gamma'\in\mathfrak{S}}g(\gamma')}\left[(U_1,\dots,U_k) = (\hat{U}_{\sigma\gamma^{-1}(1)},\dots,\hat{U}_{\sigma\gamma^{-1}(k)})\right]\,\middle|\,\hat{U}\right]\\ &= \sum_{\gamma\in\mathfrak{S}}\frac{\hat{g}(\sigma)}{\sum_{\gamma'\in\mathfrak{S}}\hat{g}(\sigma\gamma^{-1}\gamma')}\mathbf{P}\left((U_1,\dots,U_k) = (\hat{U}_{\sigma\nu^{-1}(1)},\dots,\hat{U}_{\sigma\nu^{-1}(k)})|\hat{U}\right)\\ &= \hat{g}(\sigma)\sum_{\gamma\in\mathfrak{S}}\frac{\hat{g}(\sigma\gamma^{-1})}{\sum_{\gamma'\in\mathfrak{S}}\hat{g}(\sigma\gamma^{-1}\gamma')}\end{aligned}$$

where $\hat{g}$ is obained by replacing U with $\hat{U}$ in the definition of g. Now, since $\mathfrak{S}$ is a group, the sum (on γ) in the last equation is equal to 1. That is:

$$\mathbf{P}\left((U'_1,\dots,U'_k)=(\hat{U}_{\sigma(1)},\dots,\hat{U}_{\sigma(k)})|\hat{U}\right)=\hat{g}(\sigma)$$

This equation proves that U' is a size-biased permutation of $\hat{U}$ (and, as such, is ISBP). Since U is itself a size-biased permutation of $\hat{U}$ (because it is ISBP), we deduce that U' has same law than U. □

Proof of (i) $\implies$ (iii)

This is the delicate part: it will itself be divided into several steps.

Notation. *For $n\in\mathbb{N}^*$, we will write $\mathcal{V}_n$ the law of V_n and $\mathcal{Q}_n(v)$ the law of X_n conditionally to the event $V_n=v$ (for $v\in[0,1]$). That is, for any $f:[0,1]\times\mathbb{X}\to\mathbb{R}_+$ measurable and $n\in\mathbb{N}^*$, we have:*

$$\mathbf{E}\left[f(V_n,X_n)\right]=\int_{[0,1]}\int_{\mathbb{X}}f(v,x)\mathcal{Q}_n(v)(dx)\mathcal{V}_n(dv)$$

$\mathcal{Q}_n(v)$ is well defined for all $n\in\mathbb{N}$ and $\mathcal{V}_n$-a.a. $v\in[0,1]$ because $\mathbb{X}$ is a Polish space endowed with its Borel σ-algebra.

Step 1

The goal of this first part is to give an equation connecting $\mathcal{Q}_n$ and $\mathcal{Q}_{n+1}$.
For $n\in\mathbb{N}^*$, $f:[0,1]\to\mathbb{R}_+$ measurable and $B\in\mathcal{X}$, let us apply the second part of Lemma B.1 to g, defined as follows:

$$\begin{aligned}g:[0,1]^n&\to\mathbb{R}_+\\(p_1,\dots,p_n)&\mapsto f\left(\left(1-\sum_{i=1}^{n-2}P_i\right)^{-1}P_{n-1},\left(1-\sum_{i=1}^{n-1}P_i\right)^{-1}P_{n-1}\right)\end{aligned}$$

We have:

$$\mathbf{E}\left[T(\sigma)f(V_n,V_{n+1})\mathbb{1}_B(X_n)\right]=\mathbf{E}\left[T(\sigma)f((1-V_n)V_{n+1},(1-(1-V_n)V_{n+1})^{-1}V_n)\mathbb{1}_B(X_{n+1})\right]$$

Since (by hypothesis), (V_n,X_n) and (V_{n+1},X_{n+1}) are independent, this gives:

$$\begin{aligned}&\int_{[0,1]}\int_{[0,1]}T(v_n,v_{n+1})f(v_n,v_{n+1})Q_n(v_n)(B)\mathcal{V}_n(dv_n)\mathcal{V}_{n+1}(dv_{n+1})\\&=\int_{[0,1]}\int_{[0,1]}T(v_n,v_{n+1})g((1-v_n)v_{n+1},(1-(1-v_n)v_{n+1})^{-1}v_n)\mathcal{Q}_{n+1}(v_{n+1})(B)\mathcal{V}_n(dv_n)\mathcal{V}_{n+1}(dv_{n+1})\end{aligned}$$

where

$$T(v_n,v_{n+1}):=\int_{[0,1]^{n-1}}T(\sigma)\left(\bigotimes_{i=1}^{n-1}\mathcal{V}_i\right)(d(v_1,\dots,v_{n-1}))$$

In particular, for all $g:[0,1]^2\to\mathbb{R}_+$ measurable and if $B=\mathbb{X}$, we have:

$$\begin{aligned}&\int_{[0,1]}\int_{[0,1]}T(v_n,v_{n+1})g(v_n,v_{n+1})\mathcal{V}_n(dv_n)\mathcal{V}_{n+1}(dv_{n+1})\\&=\int_{[0,1]}\int_{[0,1]}T(v_n,v_{n+1})g((1-v_n)v_{n+1},(1-(1-v_n)v_{n+1})^{-1}v_n)\mathcal{V}_n(dv_n)\mathcal{V}_{n+1}(dv_{n+1})\end{aligned}$$

Applying this last equation to $g(v,w):=f(v,w)Q_n(v)(B)$ (for $(\mathcal{V}_n\otimes\mathcal{V}_{n+1})$-a.a. $(v,w)\in[0,1]^2$) gives:

$$\begin{aligned}&\int_{[0,1]}\int_{[0,1]}T(v_n,v_{n+1})f(v_n,v_{n+1})Q_n(v_n)(B)\mathcal{V}_n(dv_n)\mathcal{V}_{n+1}(dv_{n+1})\\&=\int_{[0,1]}\int_{[0,1]}T(v_n,v_{n+1})f((1-v_n)v_{n+1},(1-(1-v_n)v_{n+1})^{-1}v_n)\mathcal{Q}_n((1-v_n)v_{n+1})(B)\mathcal{V}_n(dv_n)\mathcal{V}_{n+1}(dv_{n+1})\end{aligned}$$

Taking $f:(u,v)\mapsto h((1-u)v,(1-(1-u)v)^{-1}u)/T(u,v)$ (which is well defined under the integral sign since $T(x,y)>0$ $(\mathcal{V}_n\otimes\mathcal{V})$-a.e.) for some $h:[0,1]^2\to\mathbb{R}_+$ measurable and combining equations yields:

$$\begin{aligned}&\int_{[0,1]^2}h(v_n,v_{n+1})\mathcal{Q}_{n+1}(v_{n+1})(B)(\mathcal{V}_n\otimes\mathcal{V}_{n+1})(d(v_n,v_{n+1}))\\&=\int_{[0,1]^2}h(v_n,v_{n+1})Q_n((1-v_n)v_{n+1})(B)(\mathcal{V}_n\otimes\mathcal{V}_{n+1})(d(v_n,v_{n+1}))\end{aligned}$$

By density, we deduce that for all $A\in\mathcal{X}$ and $(\mathcal{V}_n\otimes\mathcal{V}_{n+1})$-a.a. $(v_n,v_{n+1})\in[0,1]^2$:

$$\mathcal{Q}_n((1-v_n)v_{n+1})(A)=\mathcal{Q}_{n+1}(v_{n+1})(A)\tag{12}$$

Step 2

We now want to exploit Equation (12) in order to show that $\mathcal{Q}_n$ is constant.
Let $(B_k)_{k\in\mathbb{N}^*} \in \mathcal{X}^{\mathbb{N}^*}$ such that $(B_k)_{k\in\mathbb{N}^*}$ generates $\mathcal{X}$ and is stable by finite intersections. Such a sequence exists because $\mathbb{X}$ is Polish. If $K(P) = 1$ **P**-a.s. then there is nothing to prove. Let us assume that $K(P) > 1$ **P**-a.s.. There exists $V \subset [0,1]^2$ such that $\mathcal{V}_n \otimes \mathcal{V}_{n+1}(V) = 1$ and Equation (12) holds true for all $(v_n, v_{n+1}) \in V$ and all B_k (and by the monotone class theorem, to all $A \in \mathcal{X}$).
P is an ISBP RAM. Theorem 2.2.5 states, as a consequence, that $\mathcal{V}_m$ can take only 2 forms for $m \geq 2$: a Dirac measure or a measure absolutely continuous with respect to the Lebesgue measure such that the Lebesgue measure on $[0,1]$ is itself absolutely continuous with respect to $\mathcal{V}_m$ (it is in fact a Beta law but those two properties are sufficient). Let us treat both cases separately.

Step 2 - Case 1

Let us here assume that $\mathcal{V}_{n+1} = \delta_w$ for some $w \in [0,1]$. This implies that, for all $k \in \mathbb{N}^*$, we have **P**-a.s.:

$$\mathcal{Q}_n((1-V_n)V_{n+1})(B_k) = \mathcal{Q}_{n+1}(w)(B_k) =: \mathcal{Q}(B_k) \tag{13}$$

From the equations of Step 1 (with $f : (v,w) \mapsto h(v)$ for some $h : [0,1] \to \mathbb{R}_+$ measurable and $B = B_k$), we have:

$$\begin{aligned}
\int_{[0,1]} T(v_n,w)h(v_n)\mathcal{Q}_n(v_n)(B_k)\mathcal{V}_n(dv_n) &= \int_{[0,1]}\int_{[0,1]} T(v_n,v_{n+1})h(v_n)\mathcal{Q}_n(v_n)(B_k)\mathcal{V}_n(dv_n)\mathcal{V}_{n+1}(dv_{n+1}) \\
&= \int_{[0,1]}\int_{[0,1]} T(v_n,v_{n+1})h((1-v_n)v_{n+1})\mathcal{Q}_n((1-v_n)v_{n+1})(B_k)\mathcal{V}_n(dv_n)\mathcal{V}_{n+1}(dv_{n+1}) \\
&= \int_{[0,1]}\int_{[0,1]} T(v_n,v_{n+1})h((1-v_n)v_{n+1})\mathcal{Q}(B_k)\mathcal{V}_n(dv_n)\mathcal{V}_{n+1}(dv_{n+1}) \\
&= \int_{[0,1]}\int_{[0,1]} T(v_n,v_{n+1})h(v_n)\mathcal{Q}(B_k)\mathcal{V}_n(dv_n)\mathcal{V}_{n+1}(dv_{n+1}) \\
&= \int_{[0,1]} T(v_n,w)h(v_n)\mathcal{V}_n(dv_n)\mathcal{Q}(B_k)
\end{aligned}$$

Since $T(V_n, V_{n+1}) > 0$ **P**-a.s., we deduce (by density) that, for $\mathcal{V}_n$-a.a. $v \in [0,1]$, $\mathcal{Q}_n(v)(B_k) = \mathcal{Q}(B_k)$. Naturally, this is also true for $\mathcal{Q}_{n+1}$ since V_{n+1} is **P**-a.s. constant. It implies that X_n and X_{n+1} are independent of V_n and V_{n+1} respectively.

Step 2 - Case 2

Let us now assume that $\mathcal{V}_{n+1} \ll \mathcal{L}$ and $\mathcal{L} \ll \mathcal{V}_{n+1}$ (where $\mathcal{L}$ is the Lebesgue measure on $[0,1]$). We then have $\mathcal{V}_n \otimes \mathcal{V}_{n+1}(V) = 1 = \mathcal{L}^{\otimes 2}(V)$. There exists $W \subset (0,1]$ with Lebesgue measure 1 such that, for all $w \in W$, there exists $V_w \in (0,1]$ of Lebesgue measure 1 with

$$\mathcal{Q}_n((1-v)w, \cdot) = \mathcal{Q}_{n+1}(w, .)$$

for all $v \in V_w$. That is, there exists $N_w \subset (0,w]$ Lebesgue-negligible such that, for all $x \in (0,w]\backslash N_w$:

$$\mathcal{Q}_n(x, \cdot) = \mathcal{Q}_{n+1}(w, \cdot)$$

Let $w_0 \in W$. For all $w \in W$ and all $x \in [0, \min\{w, w_0\}]\backslash(N_w \cup N_{w_0})$:

$$\mathcal{Q}(\cdot) := \mathcal{Q}_{n+1}(w_0, \cdot) = \mathcal{Q}_n(x, \cdot) = \mathcal{Q}_{n+1}(w, \cdot)$$

This latter set is non-empty because it has positive Lebesgue measure. Hence, $\mathcal{Q}_{n+1}$ is constant with respect to the Lebesgue measure, and to $\mathcal{V}_{n+1}$ by absolute continuity. We can then find an increasing sequence $(w_n)_{n\in\mathbb{N}^*} \in W^{\mathbb{N}^*}$ converging to 1. Writing $N := \cup_{k\in\mathbb{N}^*} N_{w_k}$ (always negligible), with have, for all $x \in (0,1]\backslash N$:

$$\mathcal{Q}(\cdot) = \mathcal{Q}_n(x, \cdot)$$

Hence the constance of $\mathcal{Q}_n$ with respect to $\mathcal{L}$ and $\mathcal{V}_n$.

Conclusion

By (i), the $((V_n, X_n))_{n\in[\![K(P)]\!]}$ is an independent sequence. We have shown that, for all $n \in [\![K(P)]\!]$, X_n is independent of V_n and with law $\mathcal{Q}$ (since the conditional law of X_n given V_n does not depend on V_n). This shows that the elements of $(X_n)_{n\in\mathbb{N}^*}$ are i.i.d. $\mathcal{Q}$-distributed and independent of P, which is the desired result.

C Moments of the random mass of a Borel set

Let $(\zeta_{(\alpha,\theta)})_{(\alpha,\theta)\in\Xi}$ be an independent family of PY processes such that, for $(\alpha,\theta)\in\Xi$, $\zeta_{(\alpha,\theta)}$ has parameters (α,θ) and base measure $\nu\in\mathbb{M}_1(\mathbb{X})$. We are interested in the moments of $\zeta_{(\alpha,\theta)}(B)$ for some $(\alpha,\theta)\in\Xi$ and $B\in\mathcal{X}$. During our review of the existing literature, we didn't find an exact "closed-form" (finite combination of products and sums of real numbers) formula. It is however mentioned in [1] that the moments of such a random variable have been computed in an unpublished PhD thesis ([6]) in 1999. Although the next result (Theorem C.1) may not be original, it is worth mentioning as the hereby presented proof relies on a direct application of Theorem 1.2.1.

Notation. *For $n\in\mathbb{N}^*$ and $k\in\mathbb{N}$ such that $k<n$, we write:*

$$\Delta_k^n := \{(i_1,\dots,i_k)\in\mathbb{N}^k : n>i_1>\cdots>i_k\}$$

Theorem C.1. *Let ζ be a PY process with parameters $(\alpha,\theta)\in\Xi$ and base measure $\nu\in\mathbb{M}_1(\mathbb{X})$ and let $B\in\mathcal{X}$. Then, for all $n\in\mathbb{N}^*$, we have:*

$$\mathbf{E}\left[\zeta(B)^n\right] = \sum_{i=1}^{n} A_i^n \nu(B)^i$$

where, for all $i\in[\![n]\!]$,

$$A_i^n := \begin{cases} a_{(n-1,0,1)}, & \text{if } i=1 \\ \sum_{i=1}^{n-1} a_{(n-1,k,1)} a_{(k-1,0,2)} & \text{if } i=2 \\ \sum_{(k_1,\dots,k_{i-1})\in\Delta_{i-1}^n} a_{(n-1,k_1,1)} a_{(k_{i-1}-1,0,i)} \prod_{j=1}^{i-2} a_{(k_j,k_{j+1},j+1)} & \text{otherwise} \end{cases}$$

with

$$a_{(n',k,l)} := \binom{n'}{k} \sum_{m=0}^{k} \binom{k}{m} (-1)^m \prod_{r=0}^{n'+m-k-1} \frac{1-\alpha+r}{1+\theta+(l-1)\alpha+r}$$

for $(n',k,l)\in\mathbb{N}^3$.

Proof. First, one needs to introduce some random variables (independent of ζ). Let $(W_l)_{l\in\mathbb{N}}$ be an independent sequence of random variables independent of $(\zeta_{(\alpha',\theta')})_{(\alpha',\theta')\in\Xi}$ and such that, for $l\in\mathbb{N}^*$, W_l has law $Beta(1-\alpha,\theta+l\alpha)$. For $(\alpha',\theta')\in\Xi$ and $n\in\mathbb{N}$, we write:

$$E_{(\alpha',\theta')}^n = \mathbf{E}\left[\zeta_{(\alpha',\theta')}(B)^n\right]$$

Clearly, $E_{(\alpha',\theta')}^0 = 1$ and $E_{(\alpha',\theta')}^1 = \nu(B)$ for all $(\alpha',\theta')\in\Xi$. Let $n\in\mathbb{N}^*$. We want to compute $E_{(\alpha,\theta)}^{n+1}$. We have:

$$E_{(\alpha,\theta)}^{n+1} = \mathbf{E}\left[\zeta(B)^{n+1}\right] = \mathbf{E}\left[\int_{\mathbb{X}} \zeta(B)^n \mathbb{1}_B(x)\zeta(dx)\right]$$

By Equation (2) of Theorem 1.2.1, this gives:

$$E_{(\alpha,\theta)}^{n+1} = \mathbf{E}\left[\int_{\mathbb{X}}\int_{(0,1)} L(\zeta)\left((1-u)\zeta+u\delta_x\right)(B)^n \mathbb{1}_B(x) G(du)\nu(dx)\right]$$

By the same result, we know that $\mathbf{E}\left[L(\zeta)\right] G$ is the Beta law with parameters $(\alpha,\theta+\alpha)$. Hence, resuming from Equation (8) of Subsection 4.1 and using the construction given by Proposition 3.1.4, we deduce:

$$\begin{aligned} E_{(\alpha,\theta)}^{n+1} &= \mathbf{E}\left[\int_{\mathbb{X}} \left((1-W_1)\zeta_{(\alpha,\theta+\alpha)} + W_1\delta_x\right)(B)^n \mathbb{1}_B(x)\nu(dx)\right] \\ &= \mathbf{E}\left[\int_{\mathbb{X}} \left((1-W_1)\zeta_{(\alpha,\theta+\alpha)}(B) + W_1\mathbb{1}_B(x)\right)^n \mathbb{1}_B(x)\nu(dx)\right] \end{aligned}$$

Using the binomial formula, we obtain:

$$E_{(\alpha,\theta)}^{n+1} = \mathbf{E}\left[\int_{\mathbb{X}} \sum_{k=0}^{n} \binom{n}{k} (1-W_1)^k \zeta_{(\alpha,\theta+\alpha)}(B)^k W_1^{n-k} \mathbb{1}_B(x)\nu(dx)\right]$$

By independence of W_1 and ζ, this gives:

$$E_{(\alpha,\theta)}^{n+1} = \nu(B) \sum_{k=0}^{n} \binom{n}{k} \mathbf{E}\left[(1-W_1)^k W_1^{n-k}\right] E_{(\alpha,\theta+\alpha)}^k$$

One can see that $E^{n+1}_{(\alpha,\theta)}$ is now expressed in terms of $\left(E^k_{(\alpha,\theta+\alpha)}\right)_{k\in[\![n]\!]}$. Before continuing, let us define $a_{(n',k,l)}$ for $(n,k,l)\in\mathbb{N}^3$ with $n\geq k$ as follows:

$$a_{(n',k,l)} := \binom{n'}{k}\mathbf{E}\left[(1-W_l)^k W_l^{n'-k}\right]$$

Using the binomial formula again, we have:

$$\begin{aligned} a_{(n',k,l)} &= \binom{n'}{k}\sum_{m=0}^{k}\binom{k}{m}(-1)^m\mathbf{E}\left[W_l^{n'+m-k}\right] \\ &= \binom{n'}{k}\sum_{m=0}^{k}\binom{k}{m}(-1)^m\prod_{r=0}^{n'+m-k-1}\frac{1-\alpha+r}{1+\theta+(l-1)\alpha+r} \end{aligned}$$

Now, by the same arguments and for any $n'\in\mathbb{N}^*$ and $l\in\mathbb{N}$, we have:

$$E^{n'}_{(\alpha,\theta+l\alpha)} = \nu(B)\sum_{k=0}^{n'-1} a_{(n',k,l)}E^k_{(\alpha,\theta+(l+1)\alpha)}$$

At this stage, one can see that $E^{n+1}_{(\alpha,\theta)}$ is a $n+1$ degree polynom of $\nu(B)$ which coefficents are sums of products of $a_{(n',k,l)}$'s. Developing the equation leads to the desired result. □

A first observation about Theorem C.1 is that, for any $\mathcal{B}$ in $\mathcal{X}$, $(\alpha,\theta)\in\Xi$ and $n\in\mathbb{N}$, $\mathbf{E}\left[\zeta_{(\alpha,\theta)}(B)^n\right]$ can be expressed as a n degree polynomial of $\nu(B)$. Moreover, its coefficients do not depend on ν or B. As such, it appears to be very useful for numerical computations. Indeed, given a process with fixed parameters, one can pre-compute the coefficients A^i's or $a_{(n,k,l)}$'s up to a certain desired degree so that the computation of the moments of the mass of a Borel set are then reduced to a simple polynomial evaluation.

We hereby give two direct applications of Theorem C.1.

Corollary C.2. *For all $(\alpha,\theta)\in\Xi$ and $B\in\mathcal{X}$, we have:*

$$Var(\zeta_{(\alpha,\theta)}(B)) = \frac{1-\alpha}{1+\theta}\left(\nu(B)-\nu(B)^2\right)$$

In particular:

$$\zeta_{(\alpha,\theta)}(B)\xrightarrow[\theta\to\infty]{L^2}\nu(B)$$

Proof. From Theorem C.1, we have:

$$\begin{aligned} Var\left(\zeta(B)\right) &= E^2_{(\alpha,\theta)} - \left(E^1_{(\alpha,\theta)}\right)^2 \\ &= \left(\nu(B)A_1^2+\nu(B)^2A_2^2\right) - \left(\nu(B)A_1^1\right)^2 \\ &= \nu(B)a_{(1,0,1)} + \nu(B)^2\left(a_{(1,1,1)}a_{(0,0,2)}-1\right) \\ &= \nu(B)\mathbb{E}\left[W_1\right] + \nu(B)^2\left(\mathbb{E}\left[(1-W_1)\right]-1\right) \\ &= \frac{1-\alpha}{1+\theta}\left(\nu(B)-\nu(B)^2\right) \end{aligned}$$

□

The result obtained in Corollary C.2 is consistent with what is given in [1] (Equation (1.2) of the introduction) and in Proposition 1.1.3. We observe that the concentration parameter plays the same role than in the specific Dirichlet case. As an immediate consequence of Corollary C.2, we have the following result.

Corollary C.3. *Let $B_1,\ldots,B_k\in\mathcal{X}$ ($k\in\mathbb{N}^*$). Then, for all $\alpha\in(0,1)$ we have:*

$$\mathbf{E}\left[\prod_{i=1}^{k}\zeta_{(\alpha,\theta)}(B_i)\right]\underset{\theta\to+\infty}{\longrightarrow}\prod_{i=1}^{n}\nu(B_i)$$

In particular, for $B\in\mathcal{X}$ and $n\in\mathbb{N}^*$:

$$\mathbf{E}\left[\zeta_{(\alpha,\theta)}(B)^n\right]\underset{\theta\to+\infty}{\longrightarrow}\nu(B)^n$$

Proof. It follows directly from Corollary C.2 and the recursive application of Slutsky's lemma. □

From Corollary C.3, we deduce that, for any $(\alpha,\theta)\in\Xi$ and $B\in\mathcal{X}$, the convergence in expectation of $f(\zeta_{(\alpha,\theta)}(B))$ towards $f(\nu(B))$ (when θ goes to $+\infty$) happens for any measurable analytical function $f:\mathbb{R}_+\to\mathbb{R}$.

References

[1] Luai Al Labadi and Mahmoud Zarepour. "On simulations from the two-parameter Poisson-Dirichlet process and the normalized inverse-Gaussian process". In: *Sankhya A* 76.1 (2014), pp. 158–176.

[2] Richard Arratia, Andrew D Barbour, and Simon Tavaré. *Logarithmic combinatorial structures: a probabilistic approach*. Vol. 1. European Mathematical Society, 2003.

[3] François Baccelli, Bartłomiej Błaszczyszyn, and Mohamed Karray. *Random Measures, Point Processes, and Stochastic Geometry*. 2020.

[4] David Blackwell. "Discreteness of Ferguson selections". In: *The Annals of Statistics* 1.2 (1973), pp. 356–358.

[5] Bartlomiej Blaszczyszyn. "Lecture Notes on Random Geometric Models — Random Graphs, Point Processes and Stochastic Geometry". Doctoral. Lecture. France, Dec. 2017. URL: https://hal.inria.fr/cel-01654766.

[6] MA Carlton. "Applications of the Two-Parameter Poisson-Dirichlet Distribution Ph. D". PhD thesis. thesis, Dept. Math. Univ. California, Los Angeles, 1999.

[7] Lloyd T Elliott et al. "Modeling population structure under hierarchical Dirichlet processes". In: *Bayesian Analysis* 14.2 (2019), pp. 313–339.

[8] Richard Emilion. "Non-parametric Bayesian statistics". In: *Courses of the mathematics master's, Université d'Orléans* (2007-2008).

[9] Shui Feng. *The Poisson-Dirichlet distribution and related topics: models and asymptotic behaviors*. Springer Science & Business Media, 2010.

[10] Thomas S Ferguson. "A Bayesian analysis of some nonparametric problems". In: *The annals of statistics* (1973), pp. 209–230.

[11] John FC Kingman. "Random discrete distributions". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 37.1 (1975), pp. 1–15.

[12] Günter Last. "An integral characterization of the Dirichlet process". In: *Journal of Theoretical Probability* (2019), pp. 1–13.

[13] Caroline Lawless and Julyan Arbel. "A simple proof of Pitman–Yor's Chinese restaurant process from its stick-breaking representation". In: *Dependence Modeling* 7.1 (2019), pp. 45–52.

[14] Thomas Lehéricy. "Une propriété caractéristique des processus de Poisson-Dirichlet". In: *Mémoire du M2 Probabilités & Modèles Aléatoires, Université Paris 6* (2015).

[15] Kar Wai Lim et al. "Nonparametric Bayesian topic modelling with the hierarchical Pitman–Yor processes". In: *International Journal of Approximate Reasoning* 78 (2016), pp. 172–191.

[16] Mihael Perman. "Random discrete distributions derived from subordinators." In: (1992).

[17] Mihael Perman, Jim Pitman, and Marc Yor. "Size-biased sampling of Poisson point processes and excursions". In: *Probability Theory and Related Fields* 92.1 (1992), pp. 21–39.

[18] Jim Pitman. "Random discrete distributions invariant under size-biased permutation". In: *Advances in Applied Probability* (1996), pp. 525–539.

[19] Jim Pitman and Marc Yor. "The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator". In: *The Annals of Probability* (1997), pp. 855–900.

[20] Jayaram Sethuraman. "A constructive definition of Dirichlet priors". In: *Statistica sinica* (1994), pp. 639–650.

[21] Narges Sharif-Razavian and Andreas Zollmann. "An overview of nonparametric bayesian models and applications to natural language processing". In: *Science* (2008), pp. 71–93.

[22] Erik B Sudderth and Michael I Jordan. "Shared segmentation of natural scenes using dependent Pitman-Yor processes". In: *Advances in neural information processing systems*. 2009, pp. 1585–1592.