BRANCHING PROCESS REPRESENTATION OF POISSONIZED CHINESE-RESTAURANT PROCESS [OCRP($\alpha, 0$)]

BRANCHING PROCESS REPRESENTATION OF POISSONIZED CHINESE-RESTAURANT PROCESS [OCRP($\alpha, 0$)]

BY

SOUMYAJYOTI KUNDU, B.Math

A Thesis Submitted to the School of Graduate Studies in the Partial Fulfillment of the Requirements for the Degree Master of Science

Master of Science (2024)
(Department of Mathematics & Statistics)

McMaster University
Hamilton, Ontario, Canada

TITLE:	Branching Process Representation of Poissonized Chinese-Restaurant Process [oCRP$(\alpha, 0)$]
AUTHOR:	Soumyajyoti Kundu
SUPERVISOR:	Dr. Noah Forman
NUMBER OF PAGES:	vi, 46

Abstract

The Chinese Restaurant Process (CRP) is a stochastic process on partitions. One of its importance lies in Markov chain Monte Carlo algorithm for Bayesian non parametric clustering. This thesis is built in the realm of a special type of CRP called Poissonized up-down CRP. Inspired by Roger's work in [RW22] to recover CRPs from a continuous-time stochastic process called a Lévy process, we study a branching process construction that we show is equivalent to Poissonized up down CRP. This study touches on discrete trees, continuous trees namely chronological trees, Jumping Chronological Contour Process (JCCP) and Skewer process. In the course of this study we explored interesting identities involving conditional exponential distribution.

Acknowledgements

I am indebted to my advisor Dr. Noah Forman for his continuous support, his mathematical wisdom, and his invaluable editing advice on this article. I am fortunate to have had the opportunity to work under him. This thesis would not have been possible without him. I would like to thank my committee members, Dr. Benjamin Bolker and Dr. Shui Feng.

I would also like to thank my friends at Mac for the non mathematical discussions in between.

Table of Contents

List of Figures

Declaration of Academic Achievement

I, Soumyajyoti Kundu, affirm that the thesis titled "Branching Process Representation of Poissonized Chinese Restaurant Process oCRP$(\alpha, 0)$" and its contents are my own, with support provided from my advisor Dr. Noah Forman.

Chapter 1

Introduction

1.1 Basic ingredients and branching process construction

This thesis takes place in the realm of Birth and Death chain constructions, an active field in Probability Theory that emerges from Branching Processes. Imagine a group of herds of bison originating from a single herd and undergoing the following transition scheme: birth, increasing the bison population of a herd by 1; death, decreasing the bison population of a herd by 1; and branching, where a bison branches off to form a new herd. This is the discrete analogue of the setup studied by Forman et.al. in [For+20a; For+20b].

We introduce a few basic notations that we will be following throughout this thesis. Let

$$\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n \tag{1.1.1}$$

with the convention $\mathbb{N}^0 = \{\emptyset\}$. The herds will be indexed by $\mathcal{U}$. We view $\mathcal{U}$ as an infinite tree rooted at $\emptyset$ and if $u, v \in \mathcal{U}$ and u is a prefix of v, then u is an ancestor of v. For example, $(3,1)$ is an ancestor of $(3,1,4,4,1)$. We define the generation or genealogical height of $u = (u^1, u^2, \cdots, u^n)$ as $|u| = n$. If $v = (v^1, v^2, \cdots, v^m)$, then the concatenation of u and v is defined as $uv = (u^1, \cdots, u^n, v^1, \cdots, v^m)$. The mapping $\pi : \mathcal{U} \setminus \{\emptyset\} \to \mathcal{U}$ is defined by $\pi(u^1, u^2, \cdots, u^n) = (u^1, u^2, \cdots, u^{n-1})$. We refer to $\pi(u)$ as the mother of u.

Definition 1.1 (Continuous time Markov chain)**.** A stochastic process $(X_t)_{t \geq 0}$ with discrete state space $\mathcal{S}$ is called a continuous time Markov chain if for all

$t \geq 0$, $s \geq 0$, $i \in \mathcal{S}$ and $j \in \mathcal{S}$ we have the Markov property

$$\mathbb{P}\big(X(s+t) = j \big| X(s) = i, \{X(u) : 0 \leq u < s\}\big) = \mathbb{P}\big(X(s+t) = j \big| X(s) = i\big) \\ = \mathbb{P}_{ij}(t).$$

For each $t \geq 0$ there is a transition matrix

$$P(t) = \big(\mathbb{P}_{ij}(t)\big).$$

Proposition 1.2. *Let $(X_t)_{t\geq 0}$ be a continuous time Markov chain with state space $\mathcal{S}$. Consider $x \in \mathcal{S}$. Given $X(0) = x$, define T_x by*

$$T_x := \inf \{t > 0 : \; X(t) \neq x\}.$$

Then there exists a scalar $\lambda_x > 0$ such that $T_x \sim \mathrm{Exp}(\lambda_x)$. Let

$$Q_{xy} = \lambda_x \mathbb{P}\left(\{X(T_x = y\} \,\middle|\, \{X(0) = x\}\right).$$

Then the matrix $Q = (Q)_{x,y\in\mathcal{S}}$ satisfies

$$P'(t) = QP(t).$$

We say Q is the intensity matrix of $(X_t)_{t\geq 0}$.

Definition 1.3 (Q_α Markov chain)**.** A continuous time Markov chain $(X_t)_{t\geq 0}$ is called a Q_α Markov chain if the non-zero off diagonal entries of the intensity matrix Q are given by $q(m, m+1) = m - \alpha$ and $q(m, m-1) = m$ for all $m \geq 1$. For example, if Q is a 3×3 matrix then

$$Q = \begin{pmatrix} 0 & 1-\alpha & 0 \\ 1 & 0 & 2-\alpha \\ 0 & 2 & 0 \end{pmatrix}.$$

We now present a construction of the process that is the main subject of this thesis. We proceed in two stages: first, we describe the changing population in a single generic bison herd, then we use this as an ingredient to describe a family tree of herds branching off of each other. **Throughout this document $\alpha \in (0,1)$ is fixed.** We construct a single continuous time Markov chain $(Z_t, K_t)_{t\geq 0}$ on $\mathbb{N}^2$ with

$$Z_0 = n \in \mathbb{N}, \; K_0 = 0.$$

- We will view Z_t as the number of bison at time t and $K(t)$ as the number of herds that branched off of upto time t. We will construct a branching process $(Z_t)_{t\geq 0}$ that is eventually absorbed at 0.

- Let $(T_k)_{k\geq 1}$ denote the sequence of random times when (Z, K) makes a transition. Given $(Z_t, K_t)_{t\in[0,\ T_n]}$, $(T^{n+1} - T^n)$ is conditionally Exp $(2Z_{T^n})$. In the notation of Proposition 1.2, $\lambda_{(n,j)} = 2n$.

- K and Z stay constant on each interval $[T_{k+1},\ T_k)$, then at T_{k+1} one of the following transitions takes place.

  Birth: $Z_{T_{k+1}} = Z_{T_k}+1$ and $K_{T_{k+1}} = K_{T_k}$ with probability $\dfrac{1}{2}\left(1 - \dfrac{\alpha}{Z_{T_k}}\right)$.

  Death: $Z_{T_{k+1}} = Z_{T_k} - 1$ and $K_{T_{k+1}} = K_{T_k}$ with probability $\dfrac{1}{2}$.

  Branch: $Z_{T_{k+1}} = Z_{T_k}$ and $K_{T_{k+1}} = K_{T_k} + 1$ with probability $\left(\dfrac{\alpha}{2Z_{T_k}}\right)$.

We now construct our full process with multiple herds. Let $\left(\left(\hat{Z}_t^{(w)},\ \hat{K}_t^{(w)}\right)_{t\geq 0},\ w \in \mathcal{U}\right)$ be independent copies of the previously constructed chain with

$$\hat{Z}_0^{(\emptyset)} = n \in \mathbb{N},\ \hat{Z}_0^{(w)} = 1 \text{ for } w \neq \emptyset \text{ and } \hat{K}_0^{(w)} = 0 \text{ for all } w.$$

We will use these chains as ingredients in our construction. Define the absorption time of each herd w by

$$\zeta_w := \inf\left\{t \geq 0 :\ \hat{Z}_t^{(w)} = 0\right\}. \tag{1.1.2}$$

Recursively we define $\mathcal{T}_0 := \{\emptyset\}$ and

$$\mathcal{T}_n := \bigcup_{w\in\mathcal{T}_{n-1}} \left\{(wi) \in \mathbb{N}^{n-1} \times \mathbb{N} :\ 1 \leq i \leq \hat{K}_{\zeta_w}^{(w)}\right\}. \tag{1.1.3}$$

Define

$$\mathcal{T} := \bigcup_{n=0}^{\infty} \mathcal{T}_n. \tag{1.1.4}$$

Note $\mathcal{T}$ denotes the set of all herds w that arise in the process.

Define $\beta(\emptyset)$ to be the birth time of herd $\emptyset$. We will denote by $\beta(w)$ the time when herd w first appears in our process branching off of its parent herd $\pi(w)$. For $wi \in \mathcal{T}_n$ (so $w \in \mathcal{T}_{n-1}$), recursively we define

$$\beta(wi) := \beta(w) + \inf\left\{t > 0 :\ \hat{K}_t^{(w)} \geq i\right\}. \tag{1.1.5}$$

Adopt the convention $\beta(w) = \infty$ for $w \in \mathcal{U} \setminus \mathcal{T}$. Define

$$\left(Z^{(w)}(t),\ K^{(w)}(t)\right) := \begin{cases} (0,0) & \text{if } t < \beta(w) \\ \left(\hat{Z}^{(w)}\left(t-\beta(w)\right),\ \hat{K}^{(w)}\left(t-\beta(w)\right)\right) & \text{if } t \geq \beta(w) \end{cases} \tag{1.1.6}$$

Our interest is in the process

$$\left(\left(Z_t^{(w)},\ K_t^{(w)}\right)_{t\geq 0},\ w \in \mathcal{T}\right).$$

Let

$$W(t) := \left\{w \in \mathcal{U} :\ Z_t^{(w)} > 0\right\} \tag{1.1.7}$$

oCRP$(\alpha, 0)$ refers to the ordered Chinese Restaurant process (CRP), a Markov chain in the state space $\mathcal{C} := \{(n_1, n_2, \cdots, n_k) : k \geq 0, n_1, \cdots, n_k \geq 1\}$ consists of finite length of integer compositions and $n_1, n_2, \cdots, n_k$ are interpreted as the number of customers at an ordered list of k tables in a restaurant. A new transition state is reached each time a customer joins a table or opens up a new table or leaves a table. Relevant literature about CRPs can be found in the textbook Combinatorial Stochastic Processes [Pit06] and [PW09].

1.2 Main theorem and some results of this thesis

In section 2.1.2, we will define a Poissonized up-down ordered CRP as a continuous time Markov chain in which customers leave and join tables. In chapter 2, we define a total order $<_{\mathcal{U},r}$ on $\mathcal{U}$.

Theorem 1.4. *Let $\varphi_t : \{1, 2, \cdots, \#W(t)\} \to W(t)$ denote the order preserving bijection such that $\varphi_t(1) <_{\mathcal{U},r} \varphi_t(2) <_{\mathcal{U},r} \cdots <_{\mathcal{U},r} \varphi_t(\#W(t))$. The process $\Pi_t = \left(Z^{\varphi_t(j)}(t),\ 1 \leq j \leq \#W(t)\right)_{t\geq 0}$ is a Poissonized up-down ordered* CRP$(\alpha, 0)$.

The main work of Roger's PhD Thesis ([RW22]) was to recover a Ray-Knight theorem which recovers CRPs from the heights of Lévy processes with jumps, attempting to determine the scaling limits of both the Lévy process and the integer-valued paths that denote its jumps, with the goal of aligning with the framework introduced by Forman et al. [For+20a; For+18; For+20b].

The following lemmas may or may not be novel, but their usefulness lies in their contribution to understanding the branching process of the bison herds through the insights of the associated Jumping Chronological Contour Process

(JCCP) which has been discussed in chapter 2. Our goal of understanding JCCPs inspired us to ask interesting questions about Exponential random variables.

Lemma 1.5 (Lemma 3.3)**.** *Consider a time T which is independent of the Poisson process $(N(t), t \geq 0)$ of rate λ. Let, $S_0 = 0 < S_1 < S_2 < \cdots$ denote the arrival times of the Poisson process. Let,*

$$Y := \begin{cases} S_1 & \text{if } S_1 \geq T \\ T - S_{N(T)} & \text{if } S_1 < T \end{cases}$$

Then, $Y \sim \mathrm{Exp}(\lambda)$. (Note that $S_{N(T)}$ denotes the last arrival time before T).

A simplified explanation of Lemma 1.2 is that if you have a random time, T, which is independent with respect to the Poisson process, then the gap between T and the last arrival before T follows an Exp(λ) unless $T \leq S_1$. The interesting fact about this lemma is that the inter-arrival times of the Poisson process $(N(t))_{t\geq 0}$ follow Exp(λ) but the time gap between T and the last arrival before T is shorter than the inter arrival time $S_{N(T)+1} - S_{N(T)}$ and still follows Exp(λ).

Lemma 1.6 (Lemma 3.5)**.** *Let $(T_j)_{j\geq 1}$ be a positive sequence of real numbers with $\sum_{j=1}^{\infty} T_j = \infty$, and $(X_j)_{j\geq 1}$ be IID $\mathrm{Exp}(\alpha)$. Define $J = \min\{j \in \mathbb{N} : X_j \leq T_j\}$. Let $Y = \sum_{j=1}^{J-1} T_j + X_J$, and $Y = 0$ if $J = \infty$. Then Y follows $\mathrm{Exp}(\alpha)$.*

Lemma 1.7 (Lemma 3.6)**.** *Consider a Poisson process $(J_1(t))_{t\geq 0}$ with rate α. Let τ be a stopping time of $(J_1(t))_{t\geq 0}$. Let K denote the last arrival before τ, and S_K be its arrival time. Let, $D_j = S_j - S_{j-1}\ \forall\ j \geq 1$ be the inter-arrival time of $(J_1(t))_{t\geq 0}$. Let $(Y_j)_{j\geq 0}$ be IID $\mathrm{Exp}(\alpha)$ and be independent of the Poisson process. Let*

$$\hat{D}_j = \begin{cases} D_j & \textit{if } 1 \leq j \leq K \\ \tau - S_K + Y_1 & \textit{if } j = K+1 \\ Y_{j-K} & \textit{if } j \geq K+2 \end{cases}$$

Then, $(\hat{D}_j)_{j\geq 0}$ are IID $\mathrm{Exp}(\alpha)$.

Proposition 1.8 (Proposition 3.7)**.** *Let $(J(t))_{t\geq 0}$ be a Poisson process of rate α and T be an independent time w.r.t the process. Then $S_{J(T)+1} - T \sim \mathrm{Exp}(\alpha)$.*

It states that for an independent time T w.r.t the Poisson process, the time gap between the next arrival after T follows Exp(α).

Theorem 1.9. *$\left(Z_t\right)_{t\geq 0}$ is a Q_α Markov chain. [Refer to Definition 1.3].*

The proof of the above theorem centers on the fact that the probability of a non-branch event is $(\frac{2m-\alpha}{2m})$ at each Poissonian event.

1.3 Overview

In Chapter 2, we lay out most of the definitions. We start by discussing a few types of Chinese Restaurant processes, then move on to the Lévy process and results associated with it. Next we explore chronological trees, beginning with discrete ones and moving towards $\mathbb{R}$-trees. Finally, we discuss the concept of the Jumping Chronological Contour Process (JCCP) and its relation to chronological trees. We end by stating Lambert's theorem.

Chapter 3 focuses on dependent findings of the Exp(α) based on the JCCP of the bison herds. The findings of this chapter are mainly based on theorem 1.8, as previously mentioned. The results of the chapter are closely linked to our theorem, but in probabilistic models, conditioning on an event disrupts the entire framework. We conclude by illustrating an example where the inter-arrival of a Poisson process time no longer follows Exp(α) when conditioned by an event.

In Chapter 4, we briefly discuss the constructions of birth and death chains and provide the proof of the main results.

Chapter 2

Basis for the thesis

In this chapter, we begin by introducing fundamental definitions and concepts essential for understanding the thesis. Section 2.1 introduces CRP models that will be taken into account. Section 2.2 outlines the concepts and results related to Lévy processes, specifically within the context of this thesis. Then we move on to the discussion about chronological trees in Section 2.3. We begin by discussing a discrete tree, then a few ways of coding the tree, the most important of which is the height function, where the elements of the tree are considered in lexicographical order. Then we state the CMJ process before defining a chronological tree and its genealogy, and finally, in Section 2.4, the contour process obtained from a chronological tree is termed JCCP.

2.1 Chinese Restaurant Process $\left[\mathrm{CRP}(\alpha, \theta)\right]$

The Chinese restaurant process is analogous to seating customers at tables in a restaurant. The Chinese restaurant process is closely connected to Dirichlet processes and Pólya's urn scheme by some means of exchangeability. A minor objective of this thesis is to study a continuous time up-down oCRP(α, θ) where $0 \leq \alpha \leq 1$ and $\theta \geq 0$, discussed in subsection 2.1.1. This CRPs evolve from the Dubins-Pitman two-parameter CRP [Pit] with the additional ordered CRP of Pitman-Winkel [PW09]. We will also discuss in subsection 2.1.2 the discrete time CRP(α, θ) which is an exchangeable partition-valued Markov chain introduced by Pitman, which will be called an unordered CRP(α, θ) since this doesn't specify the location of the new tables. Finally, we will end our discussion on CRPs by defining the ordered CRP(α, θ) as introduced by Pitman-Winkel.

2.1.1 Discrete-time Chinese Restaurants

We will start by stating an unordered CRP(α, θ) which is an exchangeable partition valued Markov chain studied by Pitman. This model of unordered CRPs begins with a fixed number of customer, say n, and a fixed number of tables, say k, and the transition occur when the $(n+1)^{th}$ customer joins or opens up a new table.

Definition 2.1 (Exchangeable partitions, [Pit])**.** A random partition of Π_n of $[n]$ is said to be exchangeable if for any partition $\{A_1, A_2, \cdots, A_k\}$ of $[n]$, $\mathbb{P}\big(\{\Pi_n = \{A_1, A_2, \cdots, A_k\}\}\big) = p\big(|A_1|, |A_2|, \cdots, |A_k|\big)$, where $p\big(n_1, n_2, \cdots, n_k\big)$ is a symmetric function of compositions of $[n]$ termed as the exchangeable partition probability function.

Imagine each subset of a partition as a table and each element of one of these subsets as a customer. The following transitions and rates are equipped with this model with the following bounds $0 \leq \alpha \leq 1$ and $\theta > -\alpha$ or $\alpha < 0$ and $\theta = -m\alpha$ where $m \in \mathbb{Z}^+$:

- $\mathbb{P}\big(\{(n+1)-$th customer joins $i-$th table where $1 \leq i \leq k\}\big) \propto (n_i - \alpha)$.
- $\mathbb{P}\big(\{(n+1)-$th customer opens up a new table$\}\big) \propto (\theta + k\alpha)$.

This model doesn't give any idea about the location of the tables, hence it is termed as unordered CRP(α, θ). We will end our discussion by stating an ordered CRP(α, θ).

Definition 2.2 ([RW22], oCRP(α, θ))**.** We will follow the same insertion rule of the customers as described in the above model. The following transitions and rates are equipped with this model with the following bounds $0 \leq \alpha \leq 1$ and $\theta \geq 0$:

- $\mathbb{P}\big(\{$a new table opens up to the right of all tables$\}\big) \propto \theta$.
- $\mathbb{P}\big(\{$a new table open up to the left of table $i\}\big) \propto \alpha$.

Following this constructions, some more results related to exchangeable partitions and exchangeable compositions can be found in [RW22].

2.1.2 Continuous-time Chinese Restaurants $\big[$oCRP$(\alpha, \theta)\big]$

This is a class of continuous time Markov chain in the state space $\mathcal{C} := \big\{(n_1, n_2, \cdots, n_k) : k \geq 0,\ n_j \in \mathbb{N} \text{ for all } 1 \leq j \leq k\big\}$. The sequence $n_1, n_2, \cdots, n_k$ denotes the number of customers at an ordered list of k tables. The following transitions and rates are equipped with this model:

- $\mathbb{P}\big(\{\text{a new customer joins the } i\text{−th table}\}\big) \propto n_i - \alpha$ and the transition for this event leads to state $(n_1, \cdots, n_{i-1}, n_i + 1, \cdots, n_k)$ for $1 \leq i \leq k$.
- $\mathbb{P}\big(\{\text{a new customer inserts a new table to the left of the } i\text{−th table}\}\big) \propto \alpha$ and the transition for this event leads to state $(n_1, \cdots, n_{i-1}, 1, n_i, n_{i+1}, \cdots, n_k)$ for $1 \leq i \leq k$.
- $\mathbb{P}\big(\{\text{a new customer inserts a new table in the right-most position}\}\big) \propto \theta$ and the transition for this event leads to state $(n_1, \cdots, n_k, 1)$.
- For $1 \leq i \leq k$, $\mathbb{P}\big(\{\text{one of the customers at table } i \text{ leaves}\}\big) \propto n_i$ and the transition for this event leads to either $(n_1, \cdots, n_{i-1}, n_i - 1, n_{i+1}, \cdots, n_k)$ if $n_i \geq 2$ or $(n_1, \cdots, n_{i-1}, n_{i+1}, \cdots, n_k)$ if $n_i = 1$.

These transition rates give rise to a $\mathcal{C}-$valued continuous-time Markov chain if $0 \leq \alpha \leq 1$ and $\theta \geq 0$, which we call a Poissonized up-down ordered Chinese Restaurant Process with parameters (α, θ), or just up-down oCRP(α, θ). The last part of this thesis shows the construction in the introduction is equivalent to Poissonized up-down oCRP$(\alpha, 0)$.

2.2 Lévy processes

In this section, we introduce basic concepts of Lévy processes, and we will restrict ourselves to discussing Lévy processes on $\mathbb{R}$. The following definitions and the literature can be found in [Ber96].

Definition 2.3 (Infinitely divisible distributions)**.** Let μ be a probability measure on $\mathbb{R}$ and its characteristic function is given by $\varphi_\mu(\theta) = \int_{\mathbb{R}} e^{i\theta x}\mu(dx)$ for all $\theta \in \mathbb{R}$. We say μ is infinitely divisible if for all $n \in \mathbb{N}$ there exists a probability measure $\varphi_{\mu_n}(\theta)$ such that $\big(\varphi_{\mu_n}(\theta)\big)^n = \varphi_\mu(\theta)$ for all $\theta \in \mathbb{R}$. An **alternate definition** for infinitely divisible distribution, an $\mathbb{R}-$valued random variable Z is said to be infinitely divisible, if for all $n \geq 1$, there exist $\mathbb{R}-$valued IID random variables $Y_{1,n}, Y_{2,n}, \cdots, Y_{n,n}$ such that $Z \overset{d}{=} Y_{1,n} + Y_{2,n} + \cdots + Y_{n,n}$.

Example 2.4 (Poisson distribution is a infinitely divisible distribution)**.** Let

$X \sim \text{Poisson}(\lambda)$ where $\lambda > 0$. Then its characteristic function,

$$\begin{aligned}
\varphi(\theta) &= \mathbb{E}\big[e^{i\theta X}\big] \\
&= \sum_{k\in\mathbb{N}_0} e^{i\theta k}\mathbb{P}\big(\{X = k\}\big) \\
&= \sum_{k\in\mathbb{N}_0} e^{i\theta k}\frac{e^{-\lambda}\lambda^k}{k!} \\
&= e^{-\lambda}\sum_{k\in\mathbb{N}_0}\frac{(e^{i\theta}\lambda)^k}{k!} \\
&= e^{-\lambda}e^{e^{i\theta}\lambda} \\
&= \{e^{\frac{\lambda}{n}(e^{i\theta}-1)}\}^n \\
&= (\varphi_n(\theta))^n,
\end{aligned}$$

where $\varphi_n(\theta) = e^{\frac{\lambda}{n}(e^{i\theta}-1)}$, which is the characteristic function of Poisson$(\frac{\lambda}{n})$.

Definition 2.5 ([Ber96], Lévy process)**.** Consider the probability space $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ with $\mathbb{P}\big(\{\zeta = \infty\}\big) = 1$. We say $X = \big\{X_t\big\}_{t\geq 0}$ is a Lévy process for $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ if $\mathbb{P}\big(\{X_0 = 0\}\big) = 1$ and for all $s, t \geq 0$, $X_{t+s} - X_t \stackrel{d}{=} X_s$ and is independent of the process $\big(X_u\big)_{0\leq u\leq t}$.

In other words, a Lévy process is a stochastic process $X = \{X_t : t \geq 0\}$ that has the following properties:
I): $X_0 = 0$ almost surely.
II): For any $0 \leq t_1 < t_2 < \cdots t_n < \infty$, $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \cdots, X_{t_n} - X_{t_{n-1}-1}$ are mutually independent,
III): For any $s < t$, $X_t - X_s \stackrel{d}{=} X_{t-s}$,
IV): For any $\epsilon > 0$ and $t \geq 0$ it holds that $\lim_{h\to 0}\mathbb{P}\left(|X_{t+h} - X_t| > \epsilon\right) = 0$.

Remark 2.6. A Lévy process has right-continuous sample paths for $\theta \in \mathbb{R}$ the functions $t \mapsto \varphi_{X_t}(\theta)$ are right-continous, where $\varphi_{X_t}(\theta)$ denotes the characteristic function of X_t.

Example 2.7 (An easy construction of a Lévy process)**.** Consider a one dimensional Brownian motion $\big(B_t\big)_{t\geq 0}$ which is independent of a Poisson process

$\left(N_t\right)_{t\geq 0}$ of rate λ. Define $X(t) = B(t) + N(t)$. Note that $X(t)$ is right continuous with left limits. Then,

$$\begin{aligned} X(t+s) - X(t) &= (B(t+s) + N(t+s)) - (B(t) + N(t)) \\ &= \underbrace{(B(t+s) - B(t))}_{\text{independent of } (B(u))_{0\leq u\leq t} \text{ and } (N(u))_{u\geq 0}} + \underbrace{(N(t+s) - N(t))}_{\text{independent of } (N(u))_{0\leq u\leq t} \text{ and } (B(u))_{u\geq 0}} \\ &\stackrel{d}{=} B(s) + N(s) \end{aligned}$$

Thus $\left(X_t\right)_{t\geq 0}$ is a Lévy process.

Remark 2.8. More generally, sum of any finite number of independent Lévy processes is also a Lévy process. Note that since Lévy processes have stationary independent increments, they can be thought of as analogues of random walks in continuous time.

Lemma 2.9. *Let X be a Lévy process. For $t > 0$, X_t is an infinitely divisible distribution.*

Proof. Recall that, $X_0 = 0$ almost surely. Observe that for $t > 0$,

$$\begin{aligned} X_t &= \sum_{k=1}^{n} \left\{X_{kt/n} - X_{(k-1)t/n}\right\} \\ &= \left\{X_{t/n} - X_0\right\} + \left\{X_{2t/n} - X_{t/n}\right\} + \cdots + \left\{X_{nt/n} - X_{(n-1)t/n}\right\}. \end{aligned}$$

Now by property of independence of increments (I) and stationary increments (II) in Definition 1.5, we have for all $n \geq 1$, the random variables $\left\{X_{t/n} - 0\right\}$, $\left\{X_{2t/n} - X_{t/n}\right\}, \cdots, \left\{X_{nt/n} - X_{(n-1)t/n}\right\}$ are independent and identically distributed. Hence, by using the alternate definition of infinitely divisible distributions, X_t is infinitely divisible for all $t > 0$.

□

Remark 2.10. One common example of a Lévy process is a Poisson process. The next lemma shows that if $\left(X_t\right)_{t\geq 0}$ is a counting Lévy process, then $\left(X_t\right)_{t\geq 0}$ is a Poisson process. Before that we would to like state a result on the memorylessness property of continuous random variables.

Lemma 2.11 (Memorylessness)**.** *Let X be a continuous random variable. Then for all $t, s \geq 0$ it satisfies $\mathbb{P}\big(\{X > t+s\}\big) = \mathbb{P}\big(\{X > t\}\big)\mathbb{P}\big(\{X > s\}\big)$ if and only if X has Exponential distribution.*

Proof. Let X be a continuous random variable which satisfies the memorylessness property. Then,

$$\log \mathbb{P}\big(\{X > t+s\}\big) = \log \mathbb{P}\big(\{X > t\}\big)\mathbb{P}\big(\{X > s\}\big)$$

Consider $g(t) = \log \mathbb{P}\big(\{X > t\}\big)$, then g satisfies $g(t+s) = g(t) + g(s)$ for all $t, s > 0$. Thus g is a additive function on $\mathbb{R}^+ \Rightarrow g$ is $\mathbb{Q}-$linear $\to g(q) = cq$ for all $q \in \mathbb{Q}$ and $c > 0$. Note that g is right-continuous. Consider $r \in \mathbb{Q}^c$. Then there exists a sequence $(q_n)_{n\geq 1} \subseteq \mathbb{Q}$ such that $q_n > r$ and $q_n \to r$. By right continuity of g, we have $g(q_n) \to g(r)$. Again, $g(q_n) \to cr$. Thus $g(r) = cr$, which shows $g(t) = ct$ for all $t \in \mathbb{R}^+ \Rightarrow \log \mathbb{P}\big(\{X > t\}\big) = ct \Rightarrow \mathbb{P}\big(\{X > t\}\big) = e^{ct}$. This shows that X follows an Exponential distribution.

Conversely, note that an Exponential distribution satisfies the memorylessness property. □

Theorem 2.12. *There is a bijection between the class of Lévy processes and the class of infinitely divisible distributions.*

Lemma 2.13 ([Ber96], Lévy process follows strong Markov property)**.** *For every finite stopping time T, $\big(X_s\big)_{s\leq T}$ is independent of $\big(\hat{X}_s = X_{s+T} - X_T\big)_{s\geq 0}$, and the latter has the same distribution as $\big(X_u\big)_{u\geq 0}$.*

Lemma 2.14. *Let $\big(X_t\big)_{t\geq 0}$ be a Lévy process which is also a counting process. Then $\big(X_t\big)_{t\geq 0}$ is a Poisson process.*

Proof. Consider D_1 to be the first time of jump of $(X_t)_{t\geq 0}$. Then $\mathbb{P}(\{D_1 > t+s\}|\{D_1 > t\}) = \mathbb{P}(\{X_{t+s} = 0\}|\{X_t = 0\}) = \mathbb{P}(\{X_{t+s} - X_t = 0\}|\{X_t = 0\})$. Now by independence and stationary of increments,

$$\begin{aligned}\mathbb{P}(\{X_{t+s} - X_t = 0\}|\{X_t = 0\}) &= \mathbb{P}(\{X_s = 0\}) \\ &= \mathbb{P}(\{D_1 > s\})\end{aligned}$$

Thus D_1 has an Exponential distribution, say with rate α. Next we show that $(X_{D_1+t} - X_{D_1})_{t\geq 0}$ is independent of D_1 and has the same distribution as that of $\big(X_t\big)_{t\geq 0}$. Consider the filtration $\mathcal{F}_t = \big\{X(t) > 1\big\}$ for all $t \geq 0$. Note that $\big\{D_1 \leq t\big\} \in \big\{X(t) > 1\big\}$. Thus D_1 is a stopping time. Again note that D_1 is a function of $\big(X_s\big)_{s\leq D_1}$. Thus by strong Markov property, $(X_{D_1+t} - X_{D_1})_{t\geq 0}$ is independent of D_1 and has the same distribution as that of $\big(X_t\big)_{t\geq 0}$. This gives D_2 has Exponential distribution with rate α. In a similar way, one can show that $(X_{D_2+t} - X_{D_2})_{t\geq 0}$ is independent of D_2 and has the same distribution as that of $\big(X_t\big)_{t\geq 0}$, hence inductively $\big(D_i\big)_{i\geq 1}$ is IID Exp(α). □

Theorem 2.15 (Lévy–Itô decomposition)**.** *Every Lévy process can always be represented as an independent sum of upto a countably infinite number of other Lévy processes out of which at most one will be a linear Brownian motion and the remaining processes will be compound Poisson processes with drift.*

2.3 Chronological tree

In order to understand more about chronological trees, we begin our discussion from discrete to continuous trees. The following literature can be found explicitly in [Gal05] and [Lam10].

Definition 2.16. A (discrete) rooted plane tree $\mathcal{T}$ is a finite subset of $\mathcal{U}$ that satisfies the following properties:

- $\emptyset \in \mathcal{T}$;
- For all $u \in \mathcal{T} \setminus \{\emptyset\}$, we have $\pi(u) \in \mathcal{T}$;
- For all $u \in \mathcal{T}$, there exists an integer $k_u(\mathcal{T}) \geq 0$ such that for every $j \in \mathbb{N}$, $uj \in \mathcal{T}$ if and only if $1 \leq j \leq k_u(\mathcal{T})$.

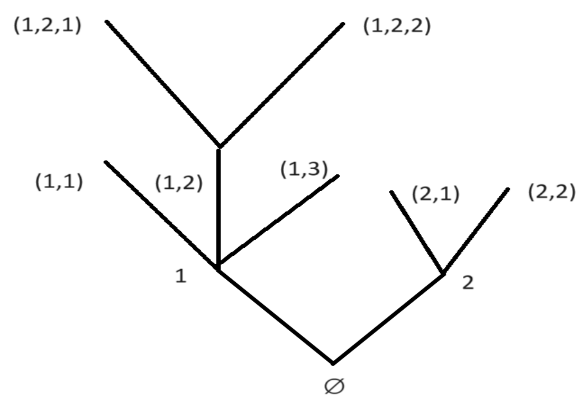

Figure 2.1: Tree $\mathcal{T}$

We present some discussions about the discrete function mainly used to code a discrete tree and its relation with Lukasiewicz paths as mentioned in [Gal05], in Appendix A. There are yet other ways to code a discrete tree, one of them being the contour function, which we shall state while discussing the JCCP later in Chapter 2.

We now state the definition of the Galton-Watson Branching process and the construction of Galton-Watson trees.

Definition 2.17 (Galton-Watson Branching process)**.** A stochastic process $\{X_n\}_{n\geq 0}$ which follows the recurrence relation $X_0 = 1$ and $X_{n+1} = \sum_{j=1}^{X_n} Y_j^n$, where $\{Y_j^n : n, j \in \mathbb{N}\}$ is a set of IID $\mathbb{N}-$valued random variables. Then

$\{X_n\}_{n\geq 0}$ is a Galton-Watson process. In contrast, X_n can be thought of as the number of descendants of the n^{th} generation, while Y_j^n can be thought of as the number of children of the j^{th} of these descendants. The probability of final extinction is given by

$$\lim_{n\to\infty} \mathbb{P}(X_n = 0).$$

Proposition 2.18 ([Gal05], Proposition 1.3)**.** *Let μ be a probability measure on $\mathbb{Z}^+$ such that $\sum_{k=1}^{\infty} k\mu(k) \leq 1$. Let $(k_u : u \in \mathcal{U})$ be a collection of IID random variables with distribution μ and $\mathcal{U}$ being the label set. Define*

$$\theta = \{u = (u^1, u^2, \cdots, u^n) \in \mathcal{U} : u^j \leq k_{(u^1,u^2,\cdots,u^{j-1})}\ \forall 1 \leq j \leq n\} \tag{2.3.1}$$

Then θ is a.s. a finite tree. Again, if $X_n = |\{u \in \theta : |u| = n\}|$, then $\{X_n\}_{n\geq 0}$ is a Galton-Watson process with probability μ.

Definition 2.19. Tree θ constructed in Proposition 2.18 is called a Galton-Watson tree.

Next, we state an informal definition of the Crump-Mode-Jagers (CMJ) branching process. For a formal definition, one can refer to [Kom16].

Definition 2.20 (CMJ)**.** In a CMJ model, each individual exists for an random duration and reproduces randomly throughout their lifespan. Their offspring undergo independent evolution, following the same process. Additionally, every individual possesses a characteristic that can be observed over time. The characteristic of an individual i at a certain age s could serve as an indicator of whether the individual lives beyond that age s, or it could represent the individual's fitness at age s.

Before we move on to the definition of chronological tree, we discuss the notion of a $\mathbb{R}-$tree. A chronological tree is a particular type of a $\mathbb{R}-$tree that can be viewed as the set of edges of a discrete rooted plane tree where each edge length denotes a lifespan. More generally, each element in the discrete structure is characterized by a starting point, denoted by α, and an ending point, denoted by ω, where $0 < \alpha < \omega$. Additionally, each element may have offspring whose birth times fall within the range (α, ω) with the possibility of these times being zero.

Definition 2.21 ([Gal05], $\mathbb{R}-$tree)**.** A compact metric space $(\mathcal{T}, d)$ is a real tree if the following properties hold:

- For all $a, b \in \mathcal{T}$ there exist an unique isometry $f_{a,b} : [0, d(a, b)] \to \mathcal{T}$ such that $f_{a,b}(0) = a$ and $f_{a,b}(d(a, b)) = b$.

- For all $a, b \in \mathcal{T}$ if g is a continuous injective map from $[0,1]$ into $\mathcal{T}$ such that $g(0) = a$ and $g(1) = b$, then $g([0,1]) = f_{a,b}([0, d(a,b)])$.

Definition 2.22 ([Lam10], Chronological tree)**.** Let $\mathbb{U} = \mathcal{U} \times [0, \infty)$ and $\rho := (\emptyset, 0)$. Let $\mathbb{T} \subseteq \mathbb{U}$ and consider the projection

$$\mathcal{T} := p(\mathbb{T}) = \{u \in \mathcal{U} : \ \exists\, \sigma \geq 0 \text{ such that } (u, \sigma) \in \mathbb{T}\} \tag{2.3.2}$$

Then $\mathbb{T}$ is a chronological tree if the following holds:

- The root $\rho \in \mathbb{T}$;
- $\mathcal{T}$ is a (discrete) rooted plane tree;
- For all $u \in \mathcal{T}$, there exist $0 \leq \alpha(u) < \omega(u) \leq \infty$ such that $(u, \sigma) \in \mathbb{T}$ iff $\sigma \in (\alpha(u), \omega(u)]$ except for $u = \emptyset, \sigma = 0$;
- For all $u \in \mathcal{T}$ and $j \in \mathbb{N}$ such that $uj \in \mathcal{T}$, then $\alpha(uj) \in (\alpha(u), \omega(u))$;
- For all $u \in \mathcal{T}$ and $i, j \in \mathbb{N}$ such that $ui, uj \in \mathcal{T}$, then for $i < j$ we have $\alpha(ui) < \alpha(uj)$.

We call $\alpha(u)$ the birth time of u and $\omega(u)$ its death time. The lifespan is denoted by

$$\zeta(u) := \omega(u) - \alpha(u).$$

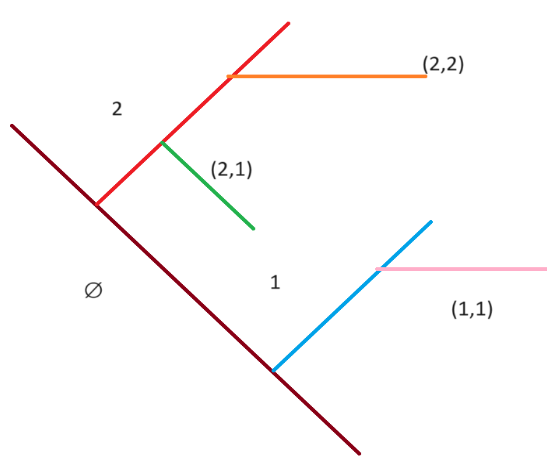

Figure 2.2: Chronological tree $\mathbb{T}$

2.4 Genealogical and metric structure of chronological trees

A detailed study can be found in [Lam10].

Let $\prec_{\mathcal{U},a}$ denote the ordering by ancestry in a discrete rooted plane tree. In Figure 2.1, $\emptyset \preceq_{\mathcal{U},a} 1$ and $\emptyset \preceq_{\mathcal{U},a} (2,1)$ but $1 \npreceq_{\mathcal{U},a} (2,1)$.

Definition 2.23 (Reverse-chronological depth-first search order, $<_{\mathcal{U},r}$)**.** We order $\mathcal{U}$ by reverse-chronological depth-first search order defined as follows,

- If w is a prefix of w', e.g. $w = (2,1)$ and $w' = (2,1,1,3,2)$ then $w <_{\mathcal{U},r} w'$ (ancestors before descendants).
- If neither $w = (n_1, n_2, \cdots, n_k)$ nor $w' = (n_1^{'}, \cdots, n_{k'}^{'})$ is a prefix of the other, then there is some $i = \min\left\{j > 0 : n_j \neq n_j^{'}\right\}$. Then either $w <_{\mathcal{U},r} w'$ iff $n_i > n_i'$ or $w' <_{\mathcal{U},r} w$ iff $n_i < n_i^{'}$ (lineage of the last born before lineage of the first born).

Definition 2.24 (Most recent common ancestor)**.** . We define the most recent common ancestor in a discrete tree $\mathcal{T}$. Let $u = (u_1, u_2, \cdots, u_m)$ and $v = (v_1, v_2, \cdots, v_n) \in \mathcal{T}$ then

$$u \wedge v = \begin{cases} (u_1, u_2, \cdots, u_k) & \text{if } \exists\ k = \min\{j : u_{j+1} \neq v_{j+1}\} \\ u & u \text{ is an ancestor of } v \\ v & v \text{ is an ancestor of } u \\ \emptyset & \text{otherwise} \end{cases}$$

Similarly we define the most recent common ancestor in a chronological tree $\mathbb{T}$. Let $x = (u, \sigma), y = (v, \tau) \in \mathbb{T}$. Let $N = \min\{j : u_{j+1} \neq v_{j+1}\}$ then $u \wedge v = (u_1, u_2, \cdots, u_N) = (v_1, v_2, \cdots, v_N)$. Define

$$x \wedge y = (u \wedge v, \alpha\left((u \wedge v) \min\left(u_{N+1}, v_{N+1}\right)\right)) \tag{2.4.1}$$

Definition 2.25 (Partial order by ancestry in a chronological tree, $\prec_{\mathbb{T},a}$)**.** Let $x = (u, \sigma), y = (v, \tau) \in \mathbb{T}$, similar to discrete trees we denote $x \prec_{\mathbb{T},a} y$ meaning x is an ancestor of y if either:

- $u = v$, and $\sigma \leq \tau$ or,
- $u \prec_{\mathcal{U},a} v$, and $\sigma \leq \alpha(uj)$, where j is the unique integer such that $uj \prec_{\mathcal{U},a} v$.

Definition 2.26 (Segments in chronological trees)**.** . We define $[\rho, x]$ to denote the set of ancestors of x i.e., $[\rho, x] = \left\{y \in \mathbb{T} : y \prec_{\mathbb{T},a} x\right\}$. Observe that $x \wedge y$

is the unique point in $\mathbb{T}$ such that

$$[\rho, x] \cap [\rho, y] = [\rho, x \wedge y] \tag{2.4.2}$$

Clearly, $x \wedge y$ is the point of maximum height in $\mathbb{T}$ such that $x \wedge y \prec_{\mathbb{T},a} x$ and $x \wedge y \prec_{\mathbb{T},a} y$. The segment $[x, y]$ is defined as

$$[x, y] := [\rho, x] \cup [\rho, y] \setminus [\rho, x \wedge y) \tag{2.4.3}$$

The natural distance d on $\mathbb{T}$ can be defined in the context of a second projection p_2 onto the second coordinate,

$$d(x, y) := p_2(x) + p_2(y) - 2p_2(x \wedge y) \tag{2.4.4}$$

The map p_2 can also be viewed as distance to the root.

Definition 2.27 (Degree in a chronological tree $\mathbb{T}$)**.** In the context of chronological tree set $\mathbb{T}$, the degree of a point x is defined as the number of connected components of $\mathbb{T} \setminus \{x\}$. This count can be 1, 2, or 3. Excluding the root ρ, points with a degree of 1 are termed as death points or leaves, and they are characterized by having coordinates $(u, \omega(u))$. Points with degree 2 are termed as simple points. Points with degree 3 are labeled as birth points, and they possess coordinates $(u, \alpha(uj))$ for some integer j where $1 \leq j \leq K_u$.

Definition 2.28 (Total order in a chronological tree, $<_{\mathbb{T},l}$)**.** Let $x, y \in \mathbb{T}$. For any $z \in \mathbb{T}$ denote by $\theta(z)$ the descendant of z, i.e. $\theta(z) = \{w \in \mathbb{T} : z \prec_{\mathbb{T},a} w\}$. The descendants of x can be split into left and right descendants namely, $\theta_l(x)$ and $\theta_r(x)$. Their definitions depends on whether x is a branching point or not. If x is not a branching point, then $\theta_l(x) = \theta(x)$ and $\theta_r(x) = \emptyset$ and if $x = (u, \sigma)$ is a branching point $\sigma = \alpha(uj)$ for some $j \leq K_u$ and

$$\theta_l(x) = \bigcup_{\epsilon>0} \theta(u, \sigma + \epsilon) \qquad \text{and} \qquad \theta_r(x) = \{x\} \cup \bigcup_{\epsilon>0} \theta(uj, \sigma + \epsilon) \tag{2.4.5}$$

Assume that $x \wedge y \notin \{x, y\}$. Then either $y \in \theta_r(x \wedge y)$ and $x \in \theta_l(x \wedge y)$ or $x \in \theta_r(x \wedge y)$ and $y \in \theta_l(x \wedge y)$. Define

$$\begin{aligned} x <_{\mathbb{T},l} y &\iff [y \prec_{\mathbb{T},a} x \text{ or } x \in \theta_l(x \wedge y)] \\ &\iff [y \prec_{\mathbb{T},a} x \text{ or } y \in \theta_r(x \wedge y)] \end{aligned}$$

Thus the relation $<_{\mathbb{T},l}$ defined on $\mathbb{T}$ is a total order whereas $\prec_{\mathbb{T},a}$ only defines a partial order.

Note that if $\mathcal{T} \subseteq \mathcal{U}$ is the discrete tree underlying a chronological tree $\mathbb{T}$ then for $u, v \in \mathcal{T}$, $u \prec_{\mathcal{U},r} v$ is the reverse chonological depth first search order iff $(u, \omega(u)) <_{\mathbb{T},l} (v, \omega(v))$ in the linear order.

Note that the Borel $\sigma-$field of a chronological tree $\mathbb{T}$ can be defined as the $\sigma-$field generated by segments. Denote the measure $\lambda(\mathbb{S})$ of a Borel subset $\mathbb{S}$ of $\mathbb{T}$. The total length of the tree

$$\lambda(\mathbb{T}) = \sum_{u \in \mathcal{T}} \zeta(u) \leq \infty \tag{2.4.6}$$

2.5 Jumping chronological contour process (JCCP)

Theorem 2.29 ([Lam10]). *Let $\mathbb{T}$ be a chronological tree and $l := \lambda(\mathbb{T}) < \infty$. Consider the measure space $([0,l], \mathcal{B}([0,l]), \mu)$ equipped with the Lebesgue measure μ. For all $x \in \overline{\mathbb{T}}$ define $\mathbb{S}(x) = \{y \in \mathbb{T} : y <_{\mathbb{T},l} x\}$. Consider the mapping $\varphi : \overline{\mathbb{T}} \to [0,l]$ given by $\varphi(x) := \lambda(\mathbb{S}(x))$ for all $x \in \overline{\mathbb{T}}$. Then φ is the unique order-preserving and measure-preserving bijection from $\overline{\mathbb{T}}$ onto $[0,l]$.*

Remark 2.30. Recall that $\mathcal{T}$ is countable and the set of leaves of $\mathbb{T}$ are in bijection with $\mathcal{T}$. $\mathbb{S}(x) \setminus \{x\}$ is the union of segments of the form $[z,y]$ where $z = y \wedge x$, $y \in \mathbb{T}$ and $y <_{\mathbb{T},l} x$, which shows $\mathbb{S}(x) \setminus \{x\}$ is a countable union of Borel sets, thus $\mathbb{S}(x) \in \mathcal{B}(\mathbb{T})$.

Lemma 2.31. *For any $x, y \in \mathbb{T}$ such that $x \prec_{\mathbb{T},a} y$, we have $\lambda([x,y]) \leq \varphi(x) - \varphi(y)$.*

Proof. Note that $\mathbb{S}(y) \cap [x,y] \subseteq \mathbb{S}(x)$, thus by monotonicity of measure we have $\lambda\left(\mathbb{S}(y) \cup [x,y]\right) \leq \lambda\big(\mathbb{S}(x)\big)$. Thus,

$$\lambda\big(\mathbb{S}(y)\big) + \lambda\left([x,y]\right) - \lambda\big(\mathbb{S}(y) \cap [x,y]\big) \leq \lambda\big(\mathbb{S}(x)\big)$$

Now $\lambda\big(\mathbb{S}(y) \cap [x,y]\big) = \lambda(\{y\}) = 0$. This gives $\lambda([x,y]) \leq \varphi(x) - \varphi(y)$. ☐

Definition 2.32. The exploration process is defined as $\Bigg(\varphi^{-1}(t) : t \in [0,l]\Bigg)$. The jumping chronological contour process (JCCP) is defined as

$$X_t := p_2 \circ \varphi^{-1}(t),\ t \in [0,l],$$

where p_2 again denotes projection onto the second coordinate.

Theorem 2.33 ([Lam10], Theorem 3.3). *The JCCP $\Bigg(X_t : t \in [0,l]\Bigg)$ has a càdlàg path and the height of each jump is the lifespan of one individual. Also,*

$$X_t = -t + \sum_{\varphi(v,\omega(v)) \leq t} \zeta_v \textit{ for all } 0 \leq t \leq l$$

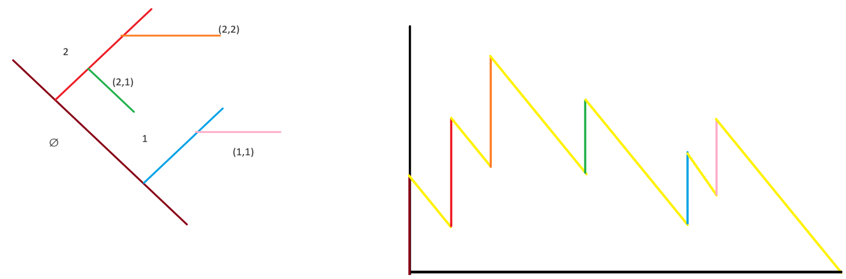

Figure 2.3: JCCP of the chronological tree $\mathbb{T}$ in figure 2.2

2.5.1 Splitting tree

The following is the discrete case of Lambert's definition in [Lam10].

Given a probability measure μ on $(0, \infty)$ and $\lambda > 0$. Let $(\zeta_u)_{u \in \mathcal{U} \setminus \{\emptyset\}}$ be IID μ and let $\zeta_\emptyset$ be a positive random variable independent of this sequence. Let $((N_u(t), t \geq 0))_{u \in \mathcal{U}}$ be IID Poisson processes of rate λ, independent of $(\zeta_u)_{u \in \mathcal{U}}$. We define

$$\mathcal{T}_n := \bigcup_{v \in \mathcal{T}_{n-1}} \{vj : 1 \leq j \leq N_u(\zeta_v)\}$$

$$\mathcal{T} := \bigcup_{n=0}^{\infty} \mathcal{T}_n$$

Fix $(\alpha(\emptyset), \omega(\emptyset)) := (0, \zeta_\emptyset)$. For $u = vj \in \mathcal{T}_n$, $n \geq 1$, define

$$\alpha(u) := \alpha(v) + \inf\{t \geq 0 : N_v(t) \geq j\}$$

$$\omega(u) := \alpha(u) + \zeta_u$$

Let $\mathbb{T} = \bigcup_{u \in \mathcal{T}} \{u\} \times (\alpha(u), \omega(u)]$. This is a random chronological tree, in the sense of definition 2.22.

Definition 2.34. [Splitting tree] A chronological tree with the probability distribution of $\mathbb{T}$ constructed above is called a splitting tree with birth rate λ and lifetime distribution μ.

Theorem 2.35 ([Lam10], Theorem 4.3)**.** *The JCCP of a splitting tree is a Lévy process.*

Chapter 3

A dependent study of $\mathrm{Exp}(\lambda)$

As discussed in chapter 1, we present the findings that were explored while understanding JCCPs. We begin with a key proposition which states that if we consider the sum of M IID copies of $\mathrm{Exp}(\lambda)$, where $M \sim \mathrm{Geom}(p)$ and M is independent of the copies, then the conditional sum follows $\mathrm{Exp}(\lambda p)$.

Proposition 3.1. *Suppose $Q_1, Q_2, \cdots$ follows IID $\mathrm{Exp}(\lambda)$ and M follow $\mathrm{Geom}(p)$ which is independent of $(Q_i)_{i=1}^{\infty}$, then $\sum_{i=1}^{M} Q_i \sim \mathrm{Exp}(\lambda p)$, with the convention that support of M starts with 1.*

Proof. Take $S = \sum_{i=1}^{M} Q_i$. Recall that the sum of n IID copies of $\mathrm{Exp}(\lambda)$ gives $\mathrm{Gamma}(n, \lambda)$. Then S has probability density function given by,

$$\begin{aligned}
f_S(s) &= \sum_{m \in \mathbb{N}} f_{S|M}(s|m)\mathbb{P}_M(m) \\
&= \sum_{m \in \mathbb{N}} \left(\frac{\lambda^m}{(m-1)!} s^{m-1} e^{-\lambda s} \right) (1-p)^{m-1} p \\
&= \lambda p e^{-\lambda s} \sum_{m \in \mathbb{N}} \frac{\lambda^{m-1}}{(m-1)!} [s(1-p)]^{m-1} \\
&= \lambda p e^{(-\lambda s + \lambda s - \lambda s p)} \\
&= \lambda p e^{-\lambda p s}.
\end{aligned}$$

This is the PDF of the claimed $\mathrm{Exp}(\lambda p)$ distribution. □

The next two lemmas are well-known and we will use them in Chapter 4.

Lemma 3.2 (Competing exponential clocks)**.** *If $X \sim \mathrm{Exp}(\lambda_1)$, $Y \sim \mathrm{Exp}(\lambda_2)$ and are independent of each other. Then,*

- $Z = \min\{X, Y\} \sim \mathrm{Exp}(\lambda_1 + \lambda_2)$,
- $\mathbb{P}(\{X < Y\}) \frac{\lambda_1}{\lambda_1 + \lambda_2}$, *and*

- *Z is independent of $\{X < Y\}$.*

Proof.

$$\begin{aligned}\mathbb{P}\big(\{Z \le k\}\big) &= \mathbb{P}\big(\{\min\{X, Y\} \le k\}\big)\\ &= 1 - \mathbb{P}\big(\{\min\{X, Y\} > k\}\big)\\ &= 1 - \mathbb{P}\big(\{X > k\} \cap \{Y > k\}\big)\\ &= 1 - \mathbb{P}\big(\{X > k\}\big)\mathbb{P}\big(\{Y > k\}\big) \quad \text{(since } X \text{ and } Y \text{ are independent)}\\ &= 1 - e^{-(\lambda_1+\lambda_2)k}.\end{aligned}$$

This shows, $Z \sim \text{Exp}(\lambda_1 + \lambda_2)$.

$$\begin{aligned}\mathbb{P}\big(\{Z \le k\}\big|\{X < Y\}\big) &= 1 - \mathbb{P}\big(\{Z > k\}\big|\{X < Y\}\big)\\ &= 1 - \frac{\mathbb{P}\big(\{Z > k\} \cap \{X < Y\}\big)}{\mathbb{P}\big(\{X < Y\}\big)}.\end{aligned}$$

Now,

$$\begin{aligned}\mathbb{P}\big(\{X < Y\}\big) &= \lambda_1\lambda_2 \int_0^\infty \int_x^\infty e^{-\lambda_1 x} e^{-\lambda_2 y} dy dx\\ &= \lambda_1\lambda_2 \bigg(\int_0^\infty e^{-\lambda_1 x}\Big[-(1/\lambda_2)e^{-\lambda_2 y}\Big]_x^\infty\bigg)dx\\ &= \lambda_1 \int_0^\infty e^{-\lambda_1 x} e^{-\lambda_2 x} dx\\ &= \frac{\lambda_1}{\lambda_1 + \lambda_2}.\end{aligned}$$

Next,

$$\begin{aligned}\mathbb{P}\big(\{Z > k\} \cap \{X < Y\}\big) &= \int_0^\infty \mathbb{P}\big(\{k < X < y\}\big)\lambda_2 e^{-\lambda_2 y} dy\\ &= e^{-\lambda_1 k}\int_k^\infty \lambda_2 e^{-\lambda_2 y} dy - \lambda_2 \int_k^\infty e^{-(\lambda_1+\lambda_2)y} dy\\ &= e^{-(\lambda_1+\lambda_2)k} - \frac{\lambda_2}{\lambda_1+\lambda_2} e^{(\lambda_1+\lambda_2)k}\\ &= e^{-(\lambda_1+\lambda_2)k}\frac{\lambda_1}{\lambda_1+\lambda_2}.\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{P}\big(\{Z \leq k\}\big|\{X < Y\}\big) &= 1 - \mathbb{P}\big(\{Z > k\}\big|\{X < Y\}\big)\\ &= 1 - e^{-(\lambda_1+\lambda_2)k}\\ &= \mathbb{P}\big(\{Z \leq k\}\big)\end{aligned}$$

Hence, Z is independent of $\{X < Y\}$. □

Based on Lemma 3.2, we provide a similar result on Geometric random variables.

Lemma 3.3 (Competing geometric clocks)**.** *Let $(X_n)_{n\geq 1}$ be an IID sequence with $\mathbb{P}(\{X_n = a\}) = p_a$, $\mathbb{P}(\{X_n = b\}) = p_b$ and $\mathbb{P}(\{X_n = c\}) = p_c$ with $p_a + p_b + p_c = 1$. Let $N_a = \inf\{n : X_n = a\} \sim \mathrm{Geom}(p_a)$ and $N_b = \inf\{n : X_n = b\} \sim \mathrm{Geom}(p_b)$. Let $N_{ab} = \min\{N_a, N_b\}$. Then,*

- $N_{ab} \sim \mathrm{Geom}(p_a + p_b)$,
- *N_{ab} is independent of $\{N_a < N_b\}$, and*
- $\mathbb{P}(\{N_a < N_b\}) = \frac{p_a}{p_a+p_b}$.

Lemma 3.4 (Strong Markov property of IID sequences)**.** *Suppose $\big(X_i\big)_{i\geq 1}$ and $\big(Y_i\big)_{i\geq 1}$ are IID sequences from the same distribution as each other and the two sequences are independent of each other. Let $\mathcal{G}$ be a $\sigma-$algebra independent of both sequences. Suppose J is a $\mathbb{N}-$valued random variable with $\{J = n\} \in \sigma\big(X_1, X_2 \cdots, X_n, \mathcal{G}\big)$ for all $n \in \mathbb{N}$. Define,*

$$W_i := \begin{cases} X_i & \text{if } i \leq J \\ Y_i & \text{otherwise} \end{cases}$$

Then $\big(W_i\big)_{i\geq 1} \overset{d}{=} \big(X_i\big)_{i\geq 1}$.

Proposition 3.5. *Suppose $(X_i)_{i\geq 1}$ follows IID $\mathrm{Exp}(a)$, $(Y_i)_{i\geq 1}$ follows IID $\mathrm{Exp}(b)$ and $(Z_i)_{i\geq 1}$ follows IID $\mathrm{Exp}(a)$, all are independent of each other. Then, $G = \min\{i \geq 1 : X_i < Y_i\} \sim \mathrm{Geom}\big(\frac{a}{a+b}\big)$. Define,*

$$\hat{X}_i := \begin{cases} X_i & \text{if } i \leq G \\ Z_i & \text{otherwise} \end{cases}$$

Then $\big(\hat{X}_i\big)_{i\geq 1}$ is IID $\mathrm{Exp}(\alpha)$.

Proof. Note that this proposition is just a special case of the Lemma 3.4. □

Lemma 3.6. *Consider a time T that is independent of the Poisson process $(N(t), t \geq 0)$ of rate λ. Let, $S_0 = 0 < S_1 < S_2 < \cdots$ denote the arrival times of the Poisson process. Let*

$$Y := \begin{cases} S_1 & \text{if } S_1 \geq T \\ T - S_{N(T)} & \text{if } S_1 < T \end{cases}$$

Then, $Y \sim \mathrm{Exp}(\lambda)$. (Note that $S_{N(T)}$ denotes the last arrival time before T)

Proof. For $t \in [0, T]$, let

$$N_1(t) := N(T) - \lim_{h \to 0} N((T - t) - h)$$

(An example illustrating N_1 : If $T = 10$, and N has arrivals at 3, 8 and 9, then N_1 has arrivals at times 1, 2 and 7) and for $t \geq 0$, let

$$N_2(t) := N(T + t) - N(T).$$

By the time-reversibility of the Poisson process $(N(t))_{t \geq 0}$, $(N_1(t))_{t \leq T}$ is also a Poisson process stopped at an independent random time T.

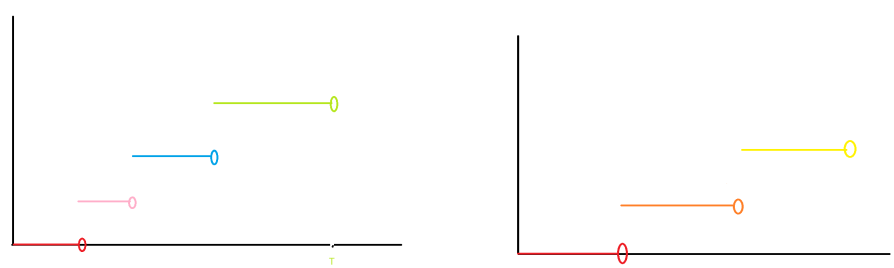

Figure 3.1: Poisson processes N_1 and N_2

Note that $N_1(T) = N(T) - N(T - T) = N(T)$. By the Lévy process property of $(N(t))_{t \geq 0}$, $(N_1(t))_{t \leq T}$ is independent of $(N_2(t))_{t \geq 0}$, and the latter is also a Poisson process of rate λ. Therefore by concatenating the increments of these two parts,

$$\hat{N}(t) := \begin{cases} N_1(t) & \text{if } 0 \leq t \leq T \\ N_2(t) + N_1(T) & \text{if } t > T \end{cases}$$

is also a Poisson process of rate λ. Now observe that Y is the first arrival time in $(\hat{N}(t))_{t \geq 0}$, thus $Y \sim \mathrm{Exp}(\lambda)$.

□

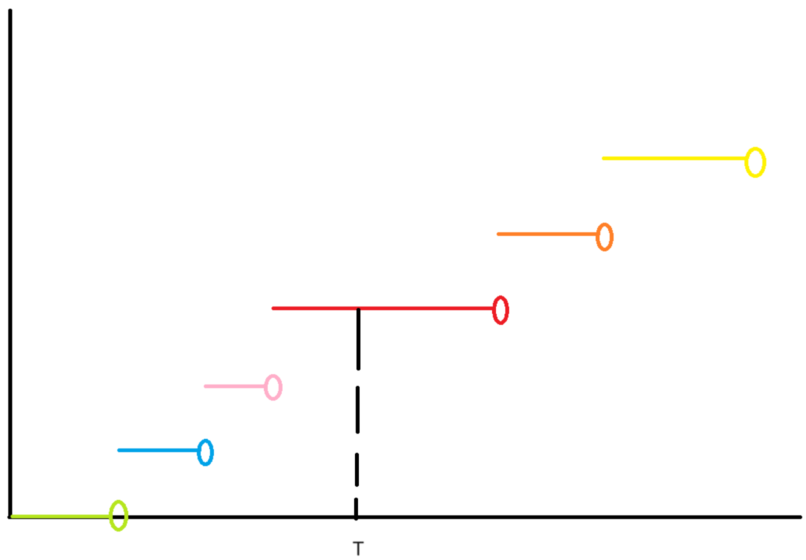

Figure 3.2: Concatenation of the increments of N_1 and N_2

Lemma 3.7. *Suppose $Z : \Omega \to S_1$ and $T : \Omega \to S_2$ are independent random variables and $f : S_1 \times S_2 \to S_3$ is a measurable function. Let $\mathcal{L}_T$ denote the law of T on S_2 and let $\mu_t = \mathcal{L}_{f(Z,t)}$ denote the law of $f(Z,t)$ on S_3 for all $t \in S_2$. If there is some law μ on S_3 such that $\mu_t = \mu$ for $\mathcal{L}_T-$a.e. $t \in S_2$ then $f(Z,T) \sim \mu$.*

Proof. Refer to [Kal97] Lemma 8.6. □

Lemma 3.8. *Let $(T_j)_{j\geq 1}$ be a positive deterministic sequence with $\sum_{j=1}^{\infty} T_j = \infty$, and let $(X_j)_{j\geq 1}$ be IID $\text{Exp}(\alpha)$. Define $J = \min\{j \in \mathbb{N} : X_j \leq T_j\}$. Let $Y = \sum_{j=1}^{J-1} T_j + X_J$, and $Y = 0$ if $J = \infty$. Then Y follows $\text{Exp}(\alpha)$.*

Proof. Let $S_j = \sum_{i=1}^{j} T_i$ and $N_m = \min\{k : S_k \geq m\}$ for all $m \in \mathbb{R}^+$. First let's compute the distribution of J. Note that,

$$\{J = k\} = \{X_j'\text{s are greater than } T_j \text{ for all } 1 \leq j \leq k-1\} \bigcap \{X_k \leq T_k\}.$$

Now since $(X_j)_{j\geq 1}$ follows IID $\text{Exp}(\alpha)$,

$$\begin{aligned}\mathbb{P}(\{J = k\}) &= \mathbb{P}(\{X_1 > T_1\} \bigcap \{X_2 > T_2\} \bigcap \cdots \bigcap \{X_{k-1} > T_{k-1}\} \bigcap \{X_k \leq T_k\}) \\ &= e^{-\alpha S_{k-1}}(1 - e^{-\alpha T_k})\end{aligned}$$

Observe that $\{Y \leq m\} = \bigsqcup_{j=1}^{N_m} (\{S_{j-1} + X_j \leq m\} \bigcap \{J = j\})$, thus

$$\begin{aligned}\mathbb{P}\big(\{Y \le m\}\big) &= \sum_{j=1}^{N_m} \mathbb{P}\big(\{S_{j-1} + X_j \le m \big| \{J = j\}\big)\mathbb{P}\big(\{J = j\}\big) \\ &= \sum_{j=1}^{N_m - 1} \mathbb{P}\big(\{J = j\}\mathbb{P}\big(\{X_j + S_{j-1} \le m\}\big|\{J = j\}\big) \\ &\qquad + \mathbb{P}\big(\{J = N_m\}\big)\mathbb{P}\big(\{X_{N_m} \le m - S_{N_m - 1}\}\big|\{X_{N_m} \le T_{N_m}\}\big).\end{aligned}$$

Now $\forall 1 \le j \le N_m - 1$, we have $m - S_{j-1} > T_j \iff m > S_{j-1} + T_j = S_j$, thus we have an immediate inclusion $\{X_j \le T_j\} \subseteq \{X_j \le m - S_{j-1}\}$, hence

$$\mathbb{P}\big(\{S_{j-1} + X_j \le m\}\big|\{J = j\}\big) = \mathbb{P}\big(\{X_j \le m - S_{j-1}\}\big|\bigcap_{i=1}^{j}\{X_i > T_i\}\bigcup\{X_j \le T_j\}\big).$$

Next,

$$\begin{aligned}&\mathbb{P}\big(\{X_{N_m} \le m - S_{N_m - 1}\}\big|\{J = N_m\}\big)\mathbb{P}\big(\{J = N_m\}\big) \\ &= \frac{\mathbb{P}\big(\{X_{N_m} \le m - S_{N_m - 1}\}\big)\mathbb{P}\big(\{J = N_m\}\big)}{\mathbb{P}\big(\{X_{N_m} \le T_{N_m}\}\big)} \\ &= \frac{\{1 - e^{-\alpha(m - S_{N_m - 1})}\}\big(1 - e^{-\alpha T_{N_m}}\big)\{e^{-\alpha(T_1 + T_2 + \cdots + T_{N_m - 1})}\}}{\big(1 - e^{-\alpha T_{N_m}}\big)} \\ &= e^{-\alpha S_{N_m - 1}} - e^{-\alpha m}.\end{aligned}$$

Finally,

$$\begin{aligned}&\mathbb{P}\big(\{Y \le m\}\big) \\ &= \mathbb{P}\big(\{J = 1\}\big) + \sum_{j=2}^{N_m - 1} \mathbb{P}\big(\{J = j\}\big) + \mathbb{P}\big(\{X_{N_m} \le m - S_{N_m - 1}\}\big|\{J = N_m\}\big)\mathbb{P}\big(\{J = N_m\}\big) \\ &= \big(1 - e^{-\alpha T_1}\big) + \sum_{j=2}^{N_m - 1} \big[e^{-\alpha(T_1 + T_2 + \cdots + T_{j-1})} - e^{-\alpha(T_1 + T_2 + \cdots + T_j)}\big] + e^{-\alpha S_{N_m - 1}} - e^{-\alpha m} \\ &= 1 - e^{-\alpha T_1} + e^{-\alpha T_1} - e^{-\alpha(T_1 + T_2 + \cdots + T_{N_m - 1})} + e^{-\alpha S_{N_m - 1}} - e^{-\alpha m} \\ &= 1 - e^{-\alpha m}.\end{aligned}$$

Thus Y follows $\mathrm{Exp}(\alpha)$. □

Lemma 3.9. *Let $\big(T_j\big)_{j \ge 1}$ be a sequence of positive random variables with $\mathbb{P}\big(\{\sum_{j=1}^{\infty} T_j = \infty\}\big) = 1$. Let $\big(X_j\big)_{j \ge 1}$ be a sequence of IID* $\mathrm{Exp}(\alpha)$ *which is independent of $\big(T_j\big)_{j \ge 1}$. Define $Y = \big(\sum_{j=1}^{J-1} T_j\big) + X_J$ or $Y = 0$ if $J = \infty$.*

Then $Y \sim \mathrm{Exp}(\alpha)$. *Also,* $\mathbb{P}\big(\{J = \infty\}\big) = 0$.

Proof. Let $X \sim \mathrm{Exp}(\alpha)$. Define $k : (0,\infty)^{\mathbb{N}} \times \mathcal{B}(\mathbb{R}) \to [0,1]$ given by,

$$k\big((t_j)_{j=1}^{\infty}, A\big) := \begin{cases} \mathbb{P}\big(\{X \in A\}\big); & \sum t_j = \infty \\ 0; & \text{otherwise} \end{cases}$$

where $\big(t_j\big)_{j\geq 1}$ is a positive sequence. Note that k is the regular conditional distribution for Y given $\big(T_j\big)_{j\geq 1}$ by Lemma 3.8. In other words, given $\big(T_j\big)_{j\geq 1}$ is conditionally $\mathrm{Exp}(\alpha)$, thus Y is overall $\mathrm{Exp}(\alpha)$ distributed. □

Throughout the remainder of the chapter, $(J(t))_{t\geq 0}$ denotes a Poisson process of rate α and $0 = S_0 < S_1 < S_2 < \cdots$ its sequence of arrival times with $D_j = S_j - S_{j-1}$ for all $j \geq 1$.

Lemma 3.10. *Consider a Poisson process* $(J_1(t))_{t\geq 0}$ *of rate* α. *Let* τ *be the stopping time of* $(J_1(t))_{t\geq 0}$. *Let* $K = J_1(\tau)$. *Let* $(Y_j)_{j\geq 1}$ *be IID* $\mathrm{Exp}(\alpha)$ *and be independent of the Poisson process. Let*

$$\hat{D}_j = \begin{cases} D_j & \text{if } 1 \leq j \leq K \\ \tau - S_K + Y_1 & \text{if } j = K+1 \\ Y_{j-K} & \text{if } j \geq K+2 \end{cases}$$

Then $(\hat{D}_j)_{j\geq 1}$ *is IID* $\mathrm{Exp}(\alpha)$.

Proof. Let $(J_2(t)_{t\geq 0})$ be another Poisson process of rate α whose inter-arrival times are Y_j. Define

$$\hat{J}(t) := \begin{cases} J_1(t) & \text{if } 0 \leq t \leq \tau \\ J_2(t) + J_1(\tau) & \text{if } t > \tau \end{cases}$$

Clearly $(J_2(t))_{t\geq 0}$ is independent of $(J_1(t))_{t\leq \tau}$ and has the same distribution as that of $(J_1(t))_{t\geq 0}$. Thus by the Lévy process property, concatenating the increments, we have that $(\hat{J}(t))_{t\geq 0}$ is also a Poisson process of rate α. Observe that $\hat{D}_j'$s are the inter-arrival times of $(\hat{J}(t))_{t\geq 0}$.

Thus $(\hat{D}_j)_{j\geq 1}$ follows IID $\mathrm{Exp}(\alpha)$. □

Proposition 3.11. *Let* $(J(t))_{t\geq 0}$ *be a Poisson process of rate* α, *then for any* $t > 0$, $S_{J(t)+1} - t \sim \mathrm{Exp}(\alpha)$.

Proof. Let $t > 0$. Write $\mathcal{L} = S_{J(t)+1} - t$. Note that for any given $t > 0$, there exists $k \in \mathbb{N}$ such that $S_k \leq t < S_{k+1}$, and we shall condition the event

$\{\mathcal{L} > l\}$ by $\{J(t) = k\}$.
By definition $D_{k+1} \sim \text{Exp}(\alpha)$ which is independent of S_k. Again,

$$S_k = \sum_{j=1}^{k} D_j,$$

where $(D_i)_{i\geq 1}$ is IID Exp(α). Thus, $S_k \sim \text{Gamma}(k, \alpha)$. Given $\{J(t) = k\}$, the first arrival time after t is the first arrival after the arrival at S_k, which gives $\{\mathcal{L} > l\} = \{D_{k+1} > l + t - S_k\}$, thus,

$$\mathbb{P}\big(\{\mathcal{L} > l\}\big|\{J(t) = k\}\big) = \mathbb{P}\big(\{D_{k+1} > l + t - S_k\}\big|\{D_{k+1} > t - S_k\} \bigcap \{S_k < t\}\big).$$

Let $A = \{D_{k+1} > l + t - S_k\}, B = \{D_{k+1} > t - S_k\}$ and $C = \{S_k < t\}$. Note that $A \subseteq B$, thus $A \cap B = A$.

$$\begin{aligned}\mathbb{P}\big(\{D_{k+1} > l + t - S_k\}\big|\{D_{k+1} > t - S_k\} \bigcap \{S_k < t\}\big) &= \mathbb{P}\big(A|B \cap C\big) \\ &= \frac{\mathbb{P}\big(A \cap C\big)}{\mathbb{P}\big(B \cap C\big)}\end{aligned}$$

Since S_k is independent of D_{k+1}, we have

$$\begin{aligned}\mathbb{P}\big(A \cap C\big) &= \int_0^t \left(\int_{l+t-x}^{\infty} f_{D_{k+1},S_k}(y,x)dy \right) dx \\ &= \int_0^t \left(\int_{l+t-x}^{\infty} f_{D_{k+1}}(y) f_{S_k}(x) dy \right) dx \\ &= \int_0^t \left(\int_{l+t-x}^{\infty} \alpha e^{-\alpha y} \frac{\alpha^k x^{k-1} e^{-\alpha x}}{(k-1)!} dy \right) dx \\ &= \int_0^t \left(\frac{\alpha^k x^{k-1} e^{-\alpha x}}{(k-1)!} \int_{l+t-x}^{\infty} \alpha e^{-\alpha y} dy \right) dx \\ &= \int_0^t \left(\frac{\alpha^k x^{k-1} e^{-\alpha x}}{(k-1)!} \big[-e^{-\alpha y}\big]_{l+t-x}^{\infty} \right) dx \\ &= \int_0^t \left(\frac{\alpha^k x^{k-1} e^{-\alpha x}}{(k-1)!} e^{-\alpha(l+t-x)} \right) dx\end{aligned}$$

$$
\begin{aligned}
&= \int_0^t \left(\frac{\alpha^k x^{k-1} e^{-\alpha(l+t)}}{(k-1)!} \right) dx \\
&= \frac{\alpha^k e^{-\alpha(l+t)}}{(k-1)!} \int_0^t x^{k-1} dx \\
&= \frac{\alpha^k e^{-\alpha(l+t)}}{(k-1)!} \Big[\frac{x^k}{k} \Big]_0^t \\
&= \frac{(\alpha t)^t}{k!} e^{-\alpha(l+t)},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{P}\big(B \cap C\big) &= \mathbb{P}\big(\big\{J(t) = k\big\}\big) \quad [J(t) \sim \text{Poisson}(\alpha t)] \\
&= \frac{(\alpha t)^k}{k!} e^{-\alpha t}.
\end{aligned}
$$

Thus

$$
\mathbb{P}\big(\big\{D_{k+1} > l + t - S_k\big\} \big| \big\{D_{k+1} > t - S_k\big\} \bigcap \big\{S_k < t\big\}\big) = e^{-\alpha l}
$$

which is independent of k. This shows that $S_{J(t)+1} - t \sim \text{Exp}(\alpha)$.

□

Lemma 3.12. *Let $(J(t))_{t\geq 0}$ be a Poisson process of rate α and T be an independent time w.r.t the process. Then $S_{J(T)+1} - T \sim \text{Exp}(\alpha)$.*

Proof. The proof of this lemma follows in a similar way to that of Lemma 3.9. □

Example 3.13. Illustration of a fact that conditioning by an event messes up the whole model
Let $(N_t)_{t\geq 0}$ be a Poisson process of rate λ, $N(0) = 0$ and $(Y_j)_{j\geq 1}$ be the inter-arrival times. Define $Z(t) = 1 - mt + N(t)$ and $\tau = \inf\{t \geq 0 : Z(t) = 0\}$. For $1 \leq j \leq N(\tau)$, $D_j = Y_j$ and $j > N(\tau)$, $D_j = \infty$.
Note that $\mathbb{P}(\{D_1 = \infty\}) = \mathbb{P}(\{N(\tau) < 1\}) = \mathbb{P}(\{N(\tau) = 0\})$. By definition of

τ, we have $Z(\tau) = 0$.

$$\begin{aligned}
\mathbb{P}\left(\{D_1 = \infty\}\right) &= \mathbb{P}\left(\{N(\tau) = 0\}\right) \\
&= \mathbb{P}\left(\{Z(\tau) - 1 + m\tau = 0\}\right) \\
&= \mathbb{P}\left(\left\{\tau = \frac{1}{m}\right\}\right) \\
&= \mathbb{P}\left(\left\{Z(\frac{1}{m}) = 0\right\}\right) \\
&= \mathbb{P}\left(\left\{N(\frac{1}{m}) = 0\right\}\right) \\
&= e^{\frac{-\lambda}{m}}
\end{aligned}$$

Now for $u \leq \frac{1}{m}$,

$$\begin{aligned}
\mathbb{P}\left(\{D_1 \leq u\} \,\middle|\, \{ D_1 \neq \infty\}\right) &= \frac{\mathbb{P}\big(\{D_1 \leq u\}, \{D_1 \neq \infty\}\big)}{\mathbb{P}(\{D_1 \neq \infty\})} \\
&= \frac{\mathbb{P}\big(\{D_1 \leq u\}, \{D_1 \neq \infty\}\big)}{1 - \mathbb{P}\big(\{D_1 = \infty\}\big)} \\
&= \frac{\mathbb{P}\big(\{D_1 \leq u\}\big)}{1 - \mathbb{P}\big(\{D_1 = \infty\}\big)} \\
&= \frac{1 - e^{-\lambda u}}{1 - e^{\frac{-\lambda}{m}}}
\end{aligned}$$

This shows that given $\{D_1 \neq \infty\}$, D_1 no longer follows $\text{Exp}(\lambda)$.

Chapter 4

The Main Result

The main aim of this chapter is to show that the skewer process of the contour obtained from the splitting tree, which in turn is derived from our construction, follows oCRP$(\alpha, 0)$. This will provide oCRP$(\alpha, 0)$ with a branching process representation.

4.1 Branching Process construction

We now repeat the two stage construction from the introduction for the readers convenience. This setup deals with countably many herds, which branched off of from a starting herd and the transitions are birth, death and branch events. This setup introduces a parameter $\alpha \in (0, 1)$, which is also eventually absorbed.

We construct a single continuous time Markov chain $(Z_t, K_t)_{t \geq 0}$ on $\mathbb{N}^2$ with

$$Z_0 = n \in \mathbb{N},\ K_0 = 0.$$

- We will view Z_t as the number of bison at time t and $K(t)$ as the number of herds that branched off of upto time t. We will construct a branching process $(Z_t)_{t \geq 0}$ that is eventually absorbed at 0.
- Let $(T_k)_{k \geq 1}$ denote the sequence of random times when Z makes a transition. Given $(Z_t)_{t \in [0,\ T_n]}$, $\left(T^{k+1} - T^k\right)$ is conditionally $\operatorname{Exp}(2Z_{T^k})$. In this notation of Proposition 1.2, $\lambda(n, j) = 2n$.
- K and Z stay constant on each interval $[T_{k+1},\ T_k)$, then at T_{k+1} any of the following transitions occurs.

  Birth: $Z_{T_{k+1}} = Z_{T_k} + 1$ and $K_{T_{k+1}} = K_{T_k}$ with probability $\dfrac{1}{2}\left(1 - \dfrac{\alpha}{Z_{T_k}}\right)$.

Death: $Z_{T_{k+1}} = Z_{T_k} - 1$ and $K_{T_{k+1}} = K_{T_k}$ with probability $\dfrac{1}{2}$

Branch: $Z_{T_{k+1}} = Z_{T_k}$ and $K_{T_{k+1}} = K_{T_k} + 1$ with probability $\left(\dfrac{\alpha}{2Z_{T_k}}\right)$.

We now construct our full process with multiple herds. Let $\left(\left(\hat{Z}_t^{(w)},\ \hat{K}_t^{(w)}\right)_{t\geq 0},\ w \in \mathcal{U}\right)$ be independent copies of the previously constructed chain with

$$\hat{Z}_0^{(\emptyset)} = n \in \mathbb{N},\ \hat{Z}_0^{(w)} = 1 \text{ for } w \neq \emptyset \text{ and } \hat{K}_0^{(w)} = 0 \text{ for all } w.$$

We will use these chains as ingredients in our construction. Define the absorption time of each herd w by

$$\zeta_w := \inf\left\{t \geq 0 :\ \hat{Z}_t^{(w)} = 0\right\}. \tag{4.1.1}$$

Recursively we define $\mathcal{T}_0 := \{\emptyset\}$ and

$$\mathcal{T}_n := \bigcup_{w \in \mathcal{T}_{n-1}} \left\{(wi) \in \mathbb{N}^{n-1} \times \mathbb{N} :\ 1 \leq i \leq \hat{K}_{\zeta_w}^{(w)}\right\}. \tag{4.1.2}$$

Define

$$\mathcal{T} := \bigcup_{n=0}^{\infty} \mathcal{T}_n. \tag{4.1.3}$$

Note $\mathcal{T}$ denotes the set of all herds w that arise in the process.
Define $\beta(\emptyset)$ to be the starting time of herd $\emptyset$. We will denote by $\beta(w)$ the time when herd w first appears in our process branching off of its parent herd $\pi(w)$. For $wi \in \mathcal{T}_n$ (so $w \in \mathcal{T}_{n-1}$), recursively we define

$$\beta(wi) := \beta(w) + \inf\left\{t > 0 :\ \hat{K}_t^{(w)} \geq i\right\}. \tag{4.1.4}$$

Adopt the convention $\beta(w) = \infty$ for $w \in \mathcal{U} \setminus \mathcal{T}$. Define

$$\left(Z^{(w)}(t),\ K^{(w)}(t)\right) := \begin{cases} (0,0) & \text{if } t < \beta(w) \\ \left(\hat{Z}^{(w)}\left(t-\beta(w)\right),\ \hat{K}^{(w)}\left(t-\beta(w)\right)\right) & \text{if } t \geq \beta(w) \end{cases} \tag{4.1.5}$$

Our interest is in the process

$$\left(\left(Z_t^{(w)},\ K_t^{(w)}\right)_{t\geq 0},\ w \in \mathcal{T}\right).$$

Let

$$W(t) := \left\{ w \in \mathcal{U} : \ Z_t^{(w)} > 0 \right\}. \tag{4.1.6}$$

Define $\tilde{Z}_t = \sum_{w \in W(t)} Z_t^{(w)}$.

4.2 Alternative construction

Given $\left\{ Z_{T_i} = m \right\}$, let $D^+_{i+1} \sim \text{Exp}(m - \alpha)$, $D^-_{i+1} \sim \text{Exp}(m)$ and $D^{br}_{i+1} \sim \text{Exp}(\alpha)$. Define $D^{no}_{i+1} = \min \left\{ D^+_{i+1}, D^-_{i+1} \right\}$, $D_{i+1} = \min \left\{ D^+_{i+1}, D^-_{i+1}, D^{br}_{i+1} \right\}$ and $T_{i+1} = T_i + D_{i+1}$. The transitions are given by,

$$\begin{aligned} &\text{Death if } D^-_{i+1} = D_{i+1} \\ &\text{Birth if } D^+_{i+1} = D_{i+1} \\ &\text{Branch if } D^{br}_{i+1} = D_{i+1} \end{aligned}$$

This is an alternative construction of $(Z_t, K_t)_{t \geq 0}$ for a single herd. Extension to multiple herds is same as before.

4.3 Rogers & Winkel's theorem

Recall the definition of Q_α Markov chains from Chapter 1.

Theorem 4.1 ([RW22]). *Fix $\alpha \in (0,1)$ and $n_0 \geq 1$. Let Z^i be independent Q_α- Markov Chains with $Z^0(0) = n_0$ and $Z^i(0) = 1$ for $i \geq 1$. Let $\zeta_i = \inf\{ s \geq 0 : Z^i_s = 0 \}$ be the absorption times. Let $\left(J_t\right)_{t \geq 0}$ be an independent Poisson process of rate α and*

$$X_t = -t + \sum_{i=0}^{J_t} \zeta_i, t \geq 0 \text{ and } T = \inf \left\{ t > 0 : X_t = 0 \right\}.$$

Let $(T_i)_{i \geq 1}$ denote jump times of X. Consider the process

$$\left(X, Z\right) := \left((X_t)_{0 \leq t \leq T}, \left(Z^i\right)_{0 \leq i \leq J_T} \right)$$

with Z^i as the mark of the i-th jump of X. Denote the set of indices of jumps that cross level y by

$$I(y) := \left\{ i : \ X(T_i -) \in \left(y - \zeta^i, y\right] \right\}.$$

Define the Skewer of $\left(Z_y^i\right)_{y\geq 0}$ *and* $(X_t)_{t\geq 0}$ *by*

$$\text{Skewer}\left(X, \left(\left(Z_y^i,\ y\geq 0\right),\ i\geq 0\right)\right)(y) := \left(Z^i\left(y - X(T_i-)\right)\right)_{i\in I(y)}.$$

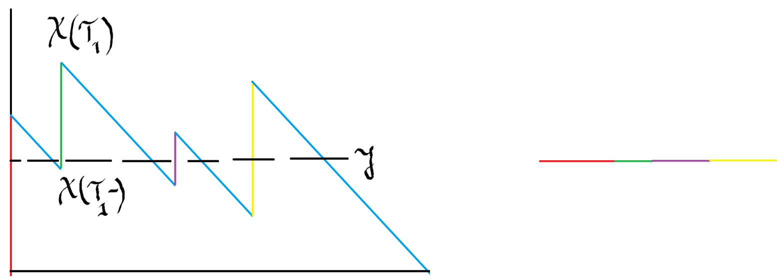

Figure 4.1: JCCP and its corresponding skewer process

Then the Skewer process of $\left(X, Z\right)$ *is up-down* $\text{oCRP}(\alpha, 0)$.

Finally, we start with the main work of the thesis, showing that the branching process construction is equivalent to the process described in the Theorem 4.1.

4.4 Steps for the equivalence

The following three propositions prove the main theorem:

Proposition 4.2. $\left(Z_t\right)_{t\geq 0}$ *is a* Q_α *Markov chain.*

Proposition 4.3. *Let* $\phi : \{1, 2, \cdots, \#\mathcal{T}\} \to \mathcal{T}$ *be the reverse chronological depth first search order preserving bijection. Define the splitting tree [refer to subsection* 2.5.1*] of bison herds by*

$$\mathbb{T} := \bigcup_{w\in\mathcal{T}} \{w\} \times (\beta(w),\ \beta(w) + \zeta_w].$$

There exists a Poisson process $(J_t)_{t\geq 0}$ *of rate* α *such that the JCCP of the splitting tree in the sense of definition* 2.29 *is given by*

$$X_t = -t + \sum_{i=0}^{J_t} \zeta_{\phi(i)},\ t\geq 0 \text{ and } T = \inf\{t\geq 0 : X_t = 0\} \tag{4.4.1}$$

and $\left(\zeta_{\phi(i)}\right)_{i\geq 1}$ *is independent of* $(J_t)_{t\geq 0}$*, where* $\zeta_{\phi(i)}$ *denotes the absorption time of herd* $\phi(i)$*.*

Proposition 4.4. $\left(\left(Z^{\varphi_t(i)}(t),\ 1 \le i \le W(t)\right),\ t \ge 0\right)$ *is up-down* oCRP$(\alpha, 0)$.

4.5 Proof of Proposition 4.2

Define $T_n :=$ the time when the n^{th} transition occurs for all $n \ge 1$. Then

$$\left(T_{n+1} - T_n \middle| (Z_t)_{t\in[0,\ T_n]}\right) \sim \text{Exp}\left(2Z_{T_n}\right).$$

First we prove that the time gaps between two non-branch events follow Exp$(2m - \alpha)$ distribution. For a fixed w, let

$$J_0 = 0,\ J_{n+1} := \inf\left\{j > J_n : Z_{T_j} \ne Z_{T_{J_n}}\right\} \text{ for all } n \ge 0.$$

i.e., if $J_{n+1} = k$, then the $(n+1)^{th}$ change in population of herd w occurs at time T_k. Define $X_m := (T_{m+1} - T_m)$ for all $m \ge 0$.

We show that given $\left\{Z_{T_{J_n}} = m\right\}$, the conditional distribution of $\sum_{j=J_n+1}^{J_{n+1}} X_j = \sum_{i=1}^{J_{n+1}-J_n} X_{J_n+i}$ follows Exp$(2m - \alpha)$ for all $n \ge 0$.

Lemma 4.5. *Suppose $\mathcal{L}_1$ and $\mathcal{L}_2$ are probability distributions on $\mathbb{R}$. Suppose $\left(W_i\right)_{i\ge1}$ and $\left(Y_i\right)_{i\ge1}$ are both IID $\mathcal{L}_1$, $\left(X_i\right)_{i\ge1}$ and $\left(Z_i\right)_{i\ge1}$ are both IID $\mathcal{L}_2$, and all four are jointly independent. Let T be a $\mathbb{N}-$valued stopping time for both sequences of pairs $\left((W_i, Y_i)\right)_{i\ge1}$ and $\left((X_i, Z_i)\right)_{i\ge1}$. Define*

$$\hat{W}_i := \begin{cases} W_i & \text{if } i \le T \\ Y_i & \text{if } i > T \end{cases}$$

and

$$\hat{X}_i := \begin{cases} X_i & \text{if } i \le T \\ Z_i & \text{if } i > T \end{cases}$$

Then $\left((\hat{W}_i, \hat{X}_i)\right)_{i\ge1}$ has the same distribution as that of $\left((Y_i, Z_i)\right)_{i\ge1}$.

Proof. By Lemma 3.4, $\left(\hat{W}_i\right)_{i\ge1}$ follows IID $\mathcal{L}_1$ and $\left(\hat{X}_i\right)_{i\ge1}$ follows IID $\mathcal{L}_2$. The sequence $(W_i, X_i)_{i\ge1}$ is IID with $\left(W_1, X_1\right) \sim \mathcal{L}_1 \otimes \mathcal{L}_2$. Likewise $\left(Y_i, Z_i\right)_{i\ge1}$ has the same IID distribution. By the strong Markov property of IID sequences, given $\left(W_i, X_i\right)_{1\le i\le T}$ the process $\left((W_{T+i}, X_{T+i})\right)_{i\ge1}$ is conditionally IID $\mathcal{L}_1 \otimes \mathcal{L}_2$. Thus $\left(W_i, X_i\right)_{1\le i\le T}$ is independent of $\left((W_{T+i}, X_{T+i})\right)_{i\ge1}$. Likewise $\left((Y_{T+i}, Z_{T+i})\right)_{i\ge1}$ is conditionally $\mathcal{L}_1 \otimes \mathcal{L}_2$, which is independent of $\left(W_i, X_i\right)_{1\le i\le T}$. This completes the proof. □

Proposition 4.6. *Given $\{Z_{T_{J_n}} = m\}$, define*

$$\hat{D}_i^{no} := \begin{cases} D_{J_n+i}^{no} & \text{if } i \leq G_{n+1} \\ Y_i^{no} & \text{if } i > G_{n+1} \end{cases}$$

and

$$\hat{D}_i^{br} := \begin{cases} D_{J_n+i}^{br} & \text{if } i \leq G_{n+1} \\ Y_i^{br} & \text{if } i > G_{n+1} \end{cases}$$

where $\left(Y_i^{no}\right)_{i=1}^{\infty}$ follows IID $\mathrm{Exp}(2m-\alpha)$ independent of $\left(Y_i^{br}\right)_{i=1}^{\infty}$ which follows IID $\mathrm{Exp}(\alpha)$, and both are independent of everything in our construction and $G_{n+1} = J_{n+1} - J_n$. Then $\left(\left(\hat{D}_i^{no}, \hat{D}_i^{br}\right)\right)_{i\geq 1}$ has the same distribution as that of $\left(\left(Y_i^{no}, Y_i^{br}\right)\right)_{i\geq 1}$.

Proof. Note that this is a special case of Lemma 4.5. □

Proposition 4.7. *Let $(Y_i)_{i=1}^{\infty}$ be IID $\mathrm{Exp}(2m)$ independent of everything in the construction. Define*

$$W_i := \begin{cases} X_{J_n+i} & \text{if } i \leq G_{n+1} \\ Y_i & \textit{otherwise} \end{cases}$$

where $G_{n+1} = J_{n+1} - J_n$. Then G_{n+1} is conditionally independent of $(W_i)_{i=1}^{\infty}$ given $\{Z_{T_{J_n}} = m\}$ and $(W_i)_{i\geq 1}$ is IID $\mathrm{Exp}(\alpha)$.

Proof. Recall that $X_i = \min\{D_i^{br}, D_i^{no}\} = T_{i+1} - T_i$. Now, observe that given $\{Z_{T_{J_n}} = m\}$, $G_{n+1} = \left(J_{n+1} - J_n\right)$ equals the number of branch events between two non-branch events which follows $\mathrm{Geom}\left(1 - \frac{\alpha}{2m}\right)$, since we are waiting for a non-branch event, the probability of which is $\left(1 - \frac{\alpha}{2m}\right)$ at each Poissonian event.
Consider as defined in Proposition 4.6,

$$\hat{D}_i^{no} := \begin{cases} D_{J_n+i}^{no} & \text{if } i \leq G_{n+1} \\ Y_i^{no} & \text{if } i > G_{n+1} \end{cases}$$

and

$$\hat{D}_i^{br} := \begin{cases} D_{J_n+i}^{br} & \text{if } i \leq G_{n+1} \\ Y_i^{br} & \text{if } i > G_{n+1} \end{cases}$$

where $\left(Y_i^{no}\right)_{i\geq 1}$ follows IID $\mathrm{Exp}(2m-\alpha)$ independent of $\left(Y_i^{br}\right)_{i\geq 1}$ which follows IID $\mathrm{Exp}(\alpha)$, and both are independent of everything in our construction.
Then $W_i := \min\{\hat{D}_i^{no}, \hat{D}_i^{br}\}$ is as in the statement of Proposition 4.7. Let $Y_i = \min\{Y_i^{no}, Y_i^{br}\}$ for all $i \geq 1$. Then $\left(Y_i\right)_{i\geq 1}$ follows IID $\mathrm{Exp}(2m)$ and is

independent of everything except $\left(Y_i^{no}, Y_i^{br}\right)_{i\geq 1}$ and

$$W_i = \begin{cases} \min\{D^{no}_{J_n+i}, D^{br}_{J_n+i}\} = X_{J_n+i} & \text{if } i \leq G_{n+1} \\ \min\{Y_i^{no}, Y_i^{br}\} = Y_i & \text{if } i > G_{n+1} \end{cases}$$

Let $\sum_1 = \sigma\big(\{\hat{D}_i^{no}, \hat{D}_i^{br}\}, i \geq 1\big)$ and $\sum_2 = \sigma\big(\min\{\hat{D}_i^{no}, \hat{D}_i^{br}\}, i \geq 1\big)$. Note that by Proposition 3.5 given $\{Z_{T_{J_n}} = m\}$, $\big(\hat{D}_i^{no}\big)_{i\geq 1}$ follows IID Exp$(2m-\alpha)$ and is independent of $\big(\hat{D}_i^{br}\big)_{i\geq 1}$ which follows IID Exp(α) (note that the independence follows by Lemma 4.5). Thus $\big(W_i\big)_{i\geq 1}$ conditionally follows Exp$(2m)$. Moreover $G_{n+1} = \min\{i \geq 1 : \hat{D}_i^{no} < \hat{D}_i^{br}\}$ which follows from $G_{n+1} = \min\{i \geq 1 : D^{no}_{J_n+i} < D^{br}_{J_n+i}\}$. Hence given $\{Z_{T_{J_n}} = m\}$, G_{n+1} is conditionally independent of $\big(W_i\big)_{i\geq 1}$ by Lemma 3.2. □

Proposition 4.8. *Fix $n \in \mathbb{N}_0$.*

$$\Big((X_j)_{j\in[J_n+1, J_{n+1}]} \,\big|\, \{Z_{T_{J_n}} = m\}, J_n, J_{n+1}\Big) \sim (\text{Exp}(2m))^{J_{n+1}-J_n}$$

Proof. This is an easy consequence of Proposition 4.7. □

We know that $J_{n+1} - J_n$ has conditional distribution Geom$(\frac{2m-\alpha}{2m})$ given $\{Z_{T_{J_n}} = m\}$. Combining Proposition 3.1 with $\lambda = 2m$ and $p = (\frac{2m-\alpha}{2m})$ and Proposition 4.7 we get, $\sum_{i=1}^{J_{n+1}-J_n} X_{J_n+i}$ follows Exp$(2m-\alpha)$ for all n. In general, $\Big(\sum_{i=1}^{J_{n+1}-J_n} X_{J_n+i}\Big)_{n\geq 1}$ is an IID sequence.

Note that in a similar way, we can show that time gaps between two branch events follow Exp(α) which can be showed in a similar way as in the above claim. A heuristic idea behind it is that the number of non-branch events between two branch events follow Geom$(\frac{\alpha}{2m})$, since we are waiting for a branch event, the probability of which is $(\frac{\alpha}{2m})$ at each Poissonian event.

Given $\{Z_{T_{J_n}} = m\}$, let

$$N_a = \min\left\{j > 0 : \; Z_{T_{J_n+j}} = Z_{T_{J_n+j-1}} + 1\right\}$$

and

$$N_b = \min\left\{j > 0 : \; Z_{T_{J_n+j}} = Z_{T_{J_n+j-1}} - 1\right\}$$

Let $N_{ab} = \min\{N_a, N_b\}$. By our construction we have $N_{ab} = J_{n+1} - J_n$ and $\{N_a < N_b\} = \Big\{Z_{T_{J_{n+1}}} = Z_{T_{J_n}} + 1\Big\}$ thus by Lemma 3.3 (competing geometric clocks) we get,

$$\begin{aligned}\mathbb{P}\left(\left\{Z_{T_{J_{n+1}}} = Z_{T_{J_n}} + 1\right\}\right) &= \frac{\frac{m-\alpha}{2m}}{\frac{m-\alpha}{2m} + \frac{1}{2}} \\ &= \frac{m-\alpha}{2m-\alpha}\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}\left(\left\{Z_{T_{J_{n+1}}} = Z_{T_{J_n}} - 1\right\}\right) &= \frac{\frac{1}{2}}{\frac{m-\alpha}{2m} + \frac{1}{2}} \\ &= \frac{m}{2m-\alpha}\end{aligned}$$

Upon rescaling the above probabilities by multiplying with $(2m - \alpha)$ since $\text{Exp}(2m - \alpha)$ are the holding times in context of Proposition 4.6 and refer back to Proposition 1.2, we get the respective rates to be $(m - \alpha)$ and m. Thus $(Z_t)_{t\geq 0}$ is a Q_α Markov chain.

4.6 Proof of Proposition 4.3

Recall Definition 2.32 of the JCCP. By Theorem 2.33, the JCCP is given by $X_t = -t + \sum_{\varphi(v,\omega(v))\leq t} \zeta_v$ for all $0 \leq t \leq l$. The JCCP obtained from the splitting tree of the bison herd population can be written as

$$X_t = -t + \sum_{i=1}^{J_t} \zeta_{\phi(i)} \text{ and } J_t = \#\left\{v \in \mathcal{T} : \ \varphi(v,\omega(v)) \leq t\right\}$$

where ϕ is the reverse chronological depth first search order preserving surjection. By Theorem 2.35, $\left(X_t\right)_{t\geq 0}$ is a Lévy process. Thus by Lévy–Itô decomposition, for $0 \leq t \leq \sum_{i\geq 1} \zeta_{\phi(i)}$ we have

$$X_t = \underbrace{-t}_{\text{drift component}} + \underbrace{\sum_{i=0}^{J_t} \zeta_{\phi(i)}}_{\text{compound Poisson component}}$$

This gives that $\left(J_t\right)_{t\geq 0}$ is a Poisson process and is independent of $\left(\zeta_{\phi(i)}\right)_{i\geq 1}$. Let $U_1 := \{t \geq 0 : \ J_t \geq 1\}$ be the first arrival time of $(J_t)_{t\geq 0}$. We will show that $U_1 \sim \text{Exp}(\alpha)$.

I would like to thank my advisor Dr. Noah Forman for providing the proofs of Lemma 4.9 and Proposition 4.10.

Lemma 4.9. *Let $(\hat{Z}_t, t \geq 0)$ be a Q_α Markov chain and let $(\hat{K}(t), t \geq 0)$ be an independent Poisson process with rate α. Let $\hat{\zeta}$ denote the absorption time of $\hat{Z}$. Then*

$$\left(\left(\hat{Z}(t), \hat{K}(t \wedge \hat{\zeta})\right),\ t \geq 0\right) \stackrel{d}{=} ((Z_t, K_t), t \geq 0), \tag{4.6.1}$$

where the latter is the process constructed at the start of this chapter.

Proof. Both are continuous-time Markov processes, so it suffices to show that they have the same intensity matrix. Fix $s \geq 0$. Let

$$T_Z = \inf\left\{t > s \colon \hat{Z}(t) \neq \hat{Z}(s)\right\}, \quad T_K = \inf\left\{t > s \colon \hat{K}(t) \neq \hat{K}(s)\right\}.$$

Let

$$E = \left\{\left(\hat{Z}(s), \hat{K}(s)\right) = (n, j)\right\}$$

for some fixed $n \geq 1$, $j \geq 0$. Note that, by construction and the Poisson property, T_K is jointly independent of $\left(\hat{Z}(t), t \geq 0\right)$ and the event E, with $T_K \sim \mathrm{Exp}(\alpha)$. On the other hand, given E, the variable T_Z has conditional distribution $\mathrm{Exp}(2n - \alpha)$, by definition of Q_α Markov chains. Thus, given E, the variable $T = T_Z \wedge T_K$ is conditionally $\mathrm{Exp}(2n)$, by the principle of competing exponentials, Lemma 3.2, and

$$\mathbb{P}\left(\left\{\hat{K}(T) = \hat{K}(s) + 1\right\} \mid E\right) = \mathbb{P}\left(\{T_K < T_Z\} \mid E\right) = \frac{\alpha}{2n}.$$

On the other hand,

$$\begin{aligned}\mathbb{P}\left(\left\{\hat{Z}(T) = \hat{Z}(s) + 1\right\} \mid E\right) &= \mathbb{P}\left(\{T_Z < T_E\} \mid E\right) \mathbb{P}\left(\left\{\hat{Z}(T) = \hat{Z}(s) + 1\right\} \mid E \cap \{T_Z < T_E\}\right) \\ &= \frac{2n - \alpha}{2n} \times \frac{n - \alpha}{2n - \alpha} = \frac{n - \alpha}{2n},\end{aligned}$$

and likewise,

$$\mathbb{P}\left(\left\{\hat{Z}(T) = \hat{Z}(s) + 1\right\} \mid E\right) = \frac{2n - \alpha}{2n} \times \frac{n}{2n - \alpha} = \frac{1}{2}.$$

Moreover, by the Markov properties of $\hat{Z}$ and $\hat{K}$, the same holds if we condition on the histories of these processes up to time s.

Finally, if we condition on

$$E' = \left\{\left(\hat{Z}(s), \hat{K}(s)\right) = (0, j)\right\}$$

instead, then $\left(\hat{Z}(t), \hat{K}(t \wedge \hat{\zeta})\right)$ remains constant on all $t \geq s$. Thus, we conclude that $\left(\left(\hat{Z}(t), \hat{K}(t \wedge \hat{\zeta})\right),\ t \geq 0\right)$ has the same intensity matrix as $((Z_t, K_t),\ t \geq 0)$, as desired. $\square$

Now let $(J_t)_{t\geq 0}$ be the Poisson process discussed previously. Let U_1 denote the first arrival time in this process and $T = \inf\{t\colon X(t) = 0\}$ as discussed previously.

Proposition 4.10. $U_1 \sim \text{Exp}(\alpha)$.

Proof. Let
$$\tilde{U} = \inf\{t > 0\colon J_{T+t} > J_T\},$$
so $\tilde{U} \sim \text{Exp}(\alpha)$ and this is independent of $(X(t),\ t \in [0, T])$.

By Lemma 4.9 (on a sufficiently large probability space) we may assume the existence of a Poisson process $\big(\hat{K}(t),\ t \geq 0\big)$ of rate α, independent of $\big(Z_t^{(\emptyset)},\ t \geq 0\big)$ such that $K_t^{(\emptyset)} = \hat{K}(t \wedge \zeta_\emptyset)$ for all t. In particular, $\hat{K}$ is independent of $\zeta_\emptyset$. Let $(S_j)_{j\geq 1}$ denote the arrival times for $\hat{K}$. Note that

$$\{U_1 > T\} = \{\mathcal{T} = \{\emptyset\}\} = \left\{K^{(\emptyset)}_{\zeta_\emptyset} = 0\right\} = \{S_1 > \zeta_\emptyset\}. \tag{4.6.2}$$

By definition of the JCCP X,

$$U_1 = \begin{cases} \zeta_\emptyset - S_{\hat{K}(\zeta_\emptyset)} & \text{if } U_1 < T \\ T + \tilde{U} & \text{otherwise.} \end{cases}$$

However in the event $U_1 > T$ we get $T = \zeta_\emptyset$. Thus,

$$\begin{aligned} U_1 &= \begin{cases} \zeta_\emptyset - S_{\hat{K}(\zeta_\emptyset)} & \text{if } S_1 < \zeta_\emptyset \\ \zeta_\emptyset + \tilde{U} & \text{otherwise} \end{cases} \\ &\overset{d}{=} \begin{cases} \zeta_\emptyset - S_{\hat{K}(\zeta_\emptyset)} & \text{if } S_1 < \zeta_\emptyset \\ S_1 & \text{otherwise,} \end{cases} \end{aligned}$$

with the last line following from the memorylessness of S_1 given $S_1 > \zeta_\emptyset$.

Finally, we conclude by Lemma 3.6 that $U_1 \sim \text{Exp}(\alpha)$, as desired. $\square$

4.7 Proof of Proposition 4.4

Recall that $\phi : \{1, 2, \cdots, \#\mathcal{T}\} \to \mathcal{T}$ is the reverse chronological depth first search order preserving bijection. By Proposition 4.2, $\left(Z_t^{\phi(i)}\right)_{t\geq 0}$ is a Q_α Markov chain for all $i \geq 1$. By Proposition 4.3, there exists a Poisson process

$(J_t)_{t\geq 0}$ of rate α such that JCCP of $\mathbb{T}$ is given by

$$X_t = -t + \sum_{i=0}^{J_t} \zeta_{\phi(i)}$$

where $\left(\zeta_{\phi(i)}\right)_{i\geq 1}$ is independent of $(J_t)_{t\geq 0}$ and there exists T such that $T = \inf\{t \geq 0 : X_t = 0\}$. Recall φ_t from Theorem 1.4. We want to show that

$$\begin{aligned}&\left(\left(Z^{\varphi_t(i)}(t),\ 1 \leq i \leq W(t)\right),\ t \geq 0\right)\\ &\quad\stackrel{d}{=} \left(\left(Z^i(y - X(T_i-)) :\ i \text{ satisfies } X(T_i-) \leq y < X(T_i)\right),\ y \geq 0\right).\end{aligned}$$

Observe that

$$\varphi_t(i) = \varphi\left(\min\left\{j :\ \#\left(W(t) \cap \varphi(\{1, 2, \cdots, j\})\right) = i\right\}\right).$$

or equivalently if $\varphi(i) \in W(t)$ then

$$\varphi(i) = \varphi_t\left(\#\left(W(t) \cap \varphi(\{1, 2, \cdots, j\})\right)\right).$$

Thus

$$\begin{aligned}&\left(\left(Z^{\varphi_t(i)}(t),\ 1 \leq i \leq W(t)\right),\ t \geq 0\right)\\ &\quad= \left(\left(\hat{Z}^{\varphi(i)}(t - \beta(\varphi(i))) :\ 1 \leq i \leq \#\mathcal{T}, t \in [\beta(\varphi(i)), \beta(\varphi(i)) + \zeta_{\varphi(i)})\right)\ t \geq 0\right).\end{aligned}$$

By our construction

$$\begin{aligned}&\left(\left(\hat{Z}^{\varphi(i)}(t - \beta(\varphi(i))) :\ 1 \leq i \leq \#\mathcal{T}, t \in [\beta(\varphi(i)), \beta(\varphi(i)) + \zeta_{\varphi(i)})\right)\ t \geq 0\right)\\ &\quad\stackrel{d}{=} \left(\left(Z^i(y - X(T_i-)) :\ i \text{ satisfies } X(T_i-) \leq y < X(T_i)\right),\ y \geq 0\right).\end{aligned}$$

Now by Theorem 4.1,

$$\left(\left(\hat{Z}^{\varphi(i)}(t - \beta(\varphi(i))) :\ 1 \leq i \leq \#\mathcal{T}, t \in [\beta(\varphi(i)), \beta(\varphi(i)) + \zeta_{\varphi(i)})\right)\ t \geq 0\right)$$

is up-down oCRP$(\alpha, 0)$. Thus

$$\left(\left(Z^{\varphi_t(i)}(t),\ 1 \leq i \leq W(t)\right),\ t \geq 0\right)$$

is up-down oCRP$(\alpha, 0)$. Hence proved.

Appendix A

Additional properties of Lévy processes

Lemma A.1. *Let* $\{X_t\}_{t\geq 0}$ *be a Lévy process. Let* φ_{X_t} *be the characteristic function of* X_t *. Then for any* $s, t > 0$, $\varphi_{X_{s+t}}(u) = \varphi_{X_s}(u)\varphi_{X_t}(u)$.

Proof. Let's begin with the characteristic function of X_{s+t}.

$$\begin{aligned}
\varphi_{X_{s+t}}(u) &= \mathbb{E}[e^{iuX_{s+t}}] \\
&= \mathbb{E}\Big[e^{iu(X_{s+t}-X_t+X_t)}\Big] \\
&= \mathbb{E}\Big[e^{iu(X_{s+t}-X_t)}e^{iuX_t}\Big] \\
&= \mathbb{E}\Big[e^{iu(X_{s+t}-X_t)}\Big]\mathbb{E}\Big[e^{iuX_t}\Big] \quad \Bigg[\text{by property (II) in Definition 2.5}\Bigg] \\
&= \mathbb{E}\Big[e^{iuX_s}\Big]\mathbb{E}\Big[e^{iuX_t}\Big] \quad \Bigg[\text{by property (I) in Definition 2.5}\Bigg] \\
&= \varphi_{X_s}(u)\varphi_{X_t}(u)
\end{aligned}$$

□

Lemma A.2. *Let* $\{X_t\}_{t\geq 0}$ *be a Lévy process. Then there exists a continuous, complex-valued function* $\psi(\theta)$, *where* $\theta \in \mathbb{R}$ *such that for all* $t \geq 0$ *and* $\theta \in \mathbb{R}$,

$$\varphi_{X_t}(\theta) = e^{t\psi(\theta)}$$

Proof. By Lemma $A.1$ $\varphi_{X_{t+s}}(u) = \varphi_{X_t}(u)\varphi_{X_s}(u)$ for all $t, s \geq 0$. Again X has right-continuous sample paths, thus for $u \in \mathbb{R}$ the function $t \mapsto \varphi_{X_t}(u)$ is right-continuous. $\varphi_{X_0}(u) = \mathbb{E}\Big[e^{iuX_0}\Big] = 1$. Since φ_{X_t} is right-continuous, thus $\lim_{t\to 0^+}\varphi_{X_t}(u) = 1$. Consider $f(t) = \varphi_{X_t}(u)$ for all $t > 0$. Thus f is a function on $\mathbb{R}^+$ and $f(t+s) = f(t)f(s)$ for all $t, s \in \mathbb{R}^+$. It's easy to note that for any rational $\frac{p_n}{q_n}$ where $p_n, q_n \in \mathbb{N}$, we have $f(\frac{p_n}{q_n}) = f(1)^{\frac{p_n}{q_n}}$. Now consider $r \in \mathbb{Q}^+$ and $r > 0$, then there exists a sequence $(q_n)_{n\geq 1} \subseteq \mathbb{Q}$ such that $q_n > r$ and $q_n \to r$. By right-continuity of f, we have $f(q_n) \to f(r)$. Again, $f(q_n) \to f(1)^r$. Thus $f(r) = f(1)^r$. Clearly $f(t) = f(1)^t$ for all $t > 0$. Thus for $u \in \mathbb{R}$, $\varphi_{X_t}(u) = e^{\log \varphi_{X_1}(u) t}$. Take $\psi(u) = \log \varphi_{X_t}(u)$ which is a continuous complex valued function. □

Theorem A.3 ([Ber96], Lévy-Khintchine representation)**.** *Consider an infinitely divisible distribution μ on $\mathbb{R}$. Then there exists $a, b \in \mathbb{R}$ and a measure ν on $\mathbb{R}$ satisfying*

$$\nu(\{0\}) = 0 \text{ and } \int_{\mathbb{R}} (1 \wedge x^2)\nu(dx) < \infty$$

such that the characteristic function of μ is given by

$$\varphi_\mu(\theta) = \exp\left\{ia\theta - \frac{1}{2}b^2\theta^2 + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - i\theta x \mathbb{1}_{|x|<1}\right)\nu(dx)\right\}$$

for all $\theta \in \mathbb{R}$.

Example A.4. (**Compound Poisson process**)
Let $N = \big(N(t)\big)_{t\geq 0}$ be a Poisson process of rate λ and let $(\xi_i)_{i\geq 1}$ be a sequence of IID random variables which is independent of $\big(N(t)\big)_{t\geq 0}$. Consider the compound Poisson process given by $X_t = \sum_{i=1}^{N(t)} \xi_i$ for all $t \geq 0$.

Note that for $t > s \geq 0$, we can write $X_t = X_s + \sum_{N(s)+1}^{N(t)} \xi_i$. Now $\big(N(t)\big)_{t\geq 0}$ has stationary independent increments and is also mutually independent with $(\xi_i)_{i\geq 1}$, thus X_t can be expressed as a sum of X_s and independent copy of X_{t-s}. Fix $\epsilon > 0$.

$$\begin{aligned}
\lim_{h\to 0} \mathbb{P}\left(\left\{\left|X_{t+h} - X_t\right| > \epsilon\right\}\right) &= \lim_{h\to 0} \mathbb{P}\left(\left\{\left|\sum_{i=1}^{N(t+h)} \xi_i - \sum_{i=1}^{N(t)} \xi_i\right| > \epsilon\right\}\right) \\
&= \lim_{h\to 0} \mathbb{P}\left(\left\{\left|\sum_{N(t)+1}^{N(t+s)} \xi_i\right| > \epsilon\right\}\right) \\
&= 0 \quad (\text{since } N \text{ is right continuous, thus } N(t) + 1 > \lim_{h\to 0} N(t+h)).
\end{aligned}$$

This shows $(X_t)_{t\geq 0}$ is a Lévy process. Let F be the distribution function of the IID variables.

$$
\begin{aligned}
\varphi_{X_t}(\theta) &= \mathbb{E}\left[e^{i\theta \sum_{i=1}^{N(t)} \xi_i}\right] \\
&= \sum_{n\geq 0} \mathbb{E}\left[e^{i\theta \sum_{i=1}^{n} \xi_i}\right] e^{\lambda} \frac{\lambda^n}{n!} \qquad \left[\text{by law of total expectation}\right] \\
&= \sum_{n\geq 0} \left(\int_{\mathbb{R}} e^{i\theta x} F(dx) \right)^n e^{-\lambda} \frac{\lambda^n}{n!} \\
&= e^{-\lambda} e^{\lambda \left(\int_{\mathbb{R}} e^{i\theta x} F(dx) \right)} \\
&= e^{-\lambda \int_{\mathbb{R}} \left(1 - e^{i\theta x}\right) F(dx)} \qquad \left[\text{since } \int_{\mathbb{R}} F(dx) = 1\right]
\end{aligned}
$$

Thus the Lévy Khintchine formula for a compound Poisson process takes the form $\psi(\theta) = \lambda \int_{\mathbb{R}} \left(1 - e^{i\theta x}\right) F(dx)$.

Appendix B

Additional path representations of trees

Recall $\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$. Let $\mathcal{T} \subset \mathcal{U}$ denote a discrete rooted plane tree.

Definition B.1 (Height function)**.** Let the elements of $\mathcal{T}$ be denoted as $\emptyset, u_1, u_2, \cdots, u_{\#(\mathcal{T})-1}$ in lexicographical order. Define, $h_{\mathcal{T}}(n) = |u_n|$ for all $0 \leq n \leq \#(\mathcal{T})$. $h_{\mathcal{T}}$. Thus $h_{\mathcal{T}}$ is the sequence of the generations of the individuals of $\mathcal{T}$, when these individuals are listed in the lexicographical order.

Proposition B.2 ([Gal05], Proposition 1.1)**.** *Let* $\boldsymbol{A}$ *denote the set of all discrete trees and* $\mathcal{S}$ *denote the set of all finite sequences of non-negative integers* $a_1, a_2, \cdots, a_p$ *such that* $a_1 + a_2 + \cdots + a_i \geq i$ *for all* $i \in \{1, 2, \cdots, p-1\}$ *and* $a_1 + a_2 + \cdots + a_p = p - 1$. *Then* $\Phi : \mathcal{T} \to (k_{u_0}(\mathcal{T}), k_{u_1}(\mathcal{T}), \cdots, k_{u_{\#(t)-1}}(\mathcal{T}))$ *is a bijective map from* $\boldsymbol{A}$ *onto* $\mathcal{S}$.

Proof. Refer [Gal05]. □

Definition B.3 (Lukasiewicz path)**.** Let $\mathcal{T} \in \mathbf{A}$ and $p = \#(\mathcal{T})$. Consider the sequence given by $x_n = \sum_{i=1}^{n}(m_i - 1)$ for all $0 \leq n \leq p$, where $\Phi(t) = (m_1, m_2, \cdots, m_p)$ and $(x_n)_{0 \leq n \leq p}$ which satisfies the following properties:

- $x_0 = 0$ and $x_p = -1$;
- $x_n \geq 0$ for every $0 \leq n \leq p - 1$;
- $x_i - x_{i-1} \geq -1$ for every $1 \leq i \leq p$.

Then $(x_n)_{0 \leq n \leq p}$ is termed as a Lukasiewicz path.

Now we are in a position to state the relation between height function and Lukasiewicz path.

Proposition B.4 ([Gal05], Proposition 1.2)**.** *The height function $h_{\mathcal{T}}$ of a tree $\mathcal{T}$ is related to the Lukasiewicz path of $\mathcal{T}$ by the formula,*

$$h_{\mathcal{T}} = \#\{j \in \{0, 1, \cdots, n-1\} : x_j = \inf_{j \le l \le n} x_l\}$$

for every $n \in \{0, 1, \cdots, \#(\mathcal{T}) - 1\}$.

Proof. Refer [Gal05]. □

In order to state next theorem, we need to define the notion of Skorokhod space. We begin with the definition of Càdlàg functions, more generally Càdlàg paths. Some literature about Càdlàg paths can be found in [CPS23].

Definition B.5 (Càdlàg functions)**.** Consider a metric space (X, d), and let $A \subseteq \mathbb{R}$. A function $f : A \to X$ is termed as Càdlàg function if f is right-continuous with left limits i.e.

- for all $x \in X$, the left hand limit $f(t-) = \lim_{s \to t^-} f(s)$ exists;
- for all $x \in X$, the right limit $f(t+) = \lim_{s \to t^+} f(s)$ exists and equals $f(t)$.

Definition B.6 (Skorokhod space)**.** The set of all Càdlàg functions from A to X is denoted by $\mathbb{D}(A, X)$ and is termed as the Skorokhod space. For a general construction of the Skorokhod metric space refer [Bil68].

Theorem B.7 ([Gal05], Theorem 1.8)**.** *Let $\theta_1, \theta_2, \cdots$ be a sequence of Galton-Watson trees with probability measure μ and finite variance σ^2 and let $(H_n : n \ge 0)$ be the associated height process. Then $(\frac{1}{\sqrt{p}} H_{[pt]}, t \ge 0)$ converges weakly to $(\frac{2}{\sigma}\gamma_t, t \ge 0)$ on the Skorokhod space $\mathbb{D}(\mathbb{R}^+, \mathbb{R}^+)$, where γ is a reflected Brownian motion.*

Bibliography

[Ber96] J. Bertoin. *Lévy Processes*. Cambridge Tracts in Mathematics. Cambridge University Press, 1996. ISBN: 9780521562430. URL: https://books.google.ca/books?id=A2CMMQEACAAJ.

[Bil68] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Mathematical Statistics. Wiley, 1968. ISBN: 9780471072423. URL: https://books.google.ca/books?id=O9oQAQAAIAAJ.

[CPS23] Christa Cuchiero, Francesca Primavera, and Sara Svaluto-Ferro. *Universal approximation theorems for continuous functions of càdlàg paths and Lévy-type signature models*. 2023. arXiv: 2208.02293 [math.PR].

[For+18] Noah Forman, Soumik Pal, Douglas Rizzolo, and Matthias Winkel. "UNIFORM CONTROL OF LOCAL TIMES OF SPECTRALLY POSITIVE STABLE PROCESSES". In: *The Annals of Applied Probability* 28.4 (2018), pp. 2592–2634. ISSN: 10505164, 21688737. URL: https://www.jstor.org/stable/26542463 (visited on 02/11/2024).

[For+20a] Noah Forman, Soumik Pal, Douglas Rizzolo, and Matthias Winkel. "Diffusions on a space of interval partitions: construction from marked Lévy processes". In: *Electronic Journal of Probability* 25.none (2020), pp. 1–46. DOI: 10.1214/20-EJP521. URL: https://doi.org/10.1214/20-EJP521.

[For+20b] Noah Forman, Soumik Pal, Douglas Rizzolo, and Matthias Winkel. "Diffusions on a space of interval partitions: Poisson–Dirichlet stationary distributions, to appear in Ann". In: *Probab.(2020+), preprint available as* (2020).

[Gal05] Jean-François Le Gall. "Random trees and applications". In: *Probability Surveys* 2.none (2005), pp. 245–311. DOI: 10.1214/154957805100000140. URL: https://doi.org/10.1214/154957805100000140.

[Kal97] Olav Kallenberg. *Foundations of modern probability*. 3rd ed. Springer, 1997.

[Kom16] Julia Komjathy. *Explosive Crump-Mode-Jagers branching processes.* 2016. arXiv: 1602.01657 [math.PR].

[Lam10] Amaury Lambert. "The contour of splitting trees is a Lévy process". In: *The Annals of Probability* 38.1 (2010), pp. 348–395. DOI: 10.1214/09-AOP485. URL: https://doi.org/10.1214/09-AOP485.

[Pit06] Jim Pitman. *Combinatorial stochastic processes: Ecole d'eté de probabilités de saint-flour xxxii-2002*. Springer, 2006.

[Pit] Jim Pitman. "Combinatorial stochastic processes, volume 1875 of". In: *Lecture Notes in Mathematics* (), pp. 7–24.

[PW09] Jim Pitman and Matthias Winkel. "Regenerative tree growth: Binary self-similar continuum random trees and Poisson–Dirichlet compositions". In: *The Annals of Probability* 37.5 (2009), pp. 1999–2041. DOI: 10.1214/08-AOP445. URL: https://doi.org/10.1214/08-AOP445.

[RW22] Dane Rogers and Matthias Winkel. "A Ray–Knight representation of up-down Chinese restaurants". In: *Bernoulli* 28.1 (2022), pp. 689–712. DOI: 10.3150/21-BEJ1364. URL: https://doi.org/10.3150/21-BEJ1364.