



Markovian tricks for non-Markovian tree: contour processes, extinction and scaling limits

Bertrand Cloez, Benoît Henry

► To cite this version:

Bertrand Cloez, Benoît Henry. Markovian tricks for non-Markovian tree: contour processes, extinction and scaling limits. 2018. hal-01678359v1

HAL Id: hal-01678359

<https://hal.science/hal-01678359v1>

Preprint submitted on 9 Jan 2018 (v1), last revised 24 Jan 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Markovian tricks for non-Markovian tree: contour processes

Extinction and Scaling limits

BERTRAND CLOEZ¹, BENOIT HENRY^{2,3}

Abstract

In this work, we study a family of non-Markovian trees modeling populations where individuals live and reproduce independently with possibly time-dependent birth-rate and lifetime distribution. To this end, we use the coding process introduced by Lambert. We show that, in our situation, this process is no longer a Lévy process but remains a Feller process and we give a complete characterization of its generator. This allows us to study the model through well-known Markov processes techniques. On one hand, introducing a scale function for such processes allows us to get necessary and sufficient conditions for extinction or non-extinction and to characterize the law of such trees conditioned on these events. On the other hand, using Lyapounov drift techniques, we get another set of, easily checkable, sufficient criteria for extinction or non-extinction and some tail estimates for the tree length. Finally, we also study scaling limits of such random trees and observe that the Bessel tree appears naturally.

MSC 2000 subject classifications: Primary 60J25, 92D25; secondary 60J80.

Key words and phrases: contour processes; branching processes; time-inhomogeneous splitting tree; Markov processes; scale function; Lyapunov function ; scaling limit.

Contents

1	Introduction	2
2	Description of the model	4
3	The contour process of an inhomogeneous splitting tree	6
3.1	Construction of the contour process	6
3.2	Markov and Feller properties	8
3.3	A jump process point of view for the contour	15

¹MISTEA, INRA, Montpellier SupAgro, Univ. Montpellier

²IECL, Université de Lorraine, Site de Nancy, B.P. 70239, F-54506 Vandœuvre-lès-Nancy Cedex, France

³CNRS, IECL, UMR 7502, Vandœuvre-lès-Nancy, F-54506, France

E-mail: Bertrand.cloez@inra.fr, benoit.henry@univ-lorraine.fr

4	Scale-type functions and hitting times	17
4.1	Definition of the scale function	17
4.2	Tree properties through scale function	20
4.3	Applications to particular cases	22
5	Lyapounov functions: sufficient condition and tail estimates	24
5.1	General condition	24
5.2	Applications to particular cases	29
6	Scaling limits of non-homogeneous splitting trees	29
6.1	Scaling limit of asymptotically critical IST	29
6.2	No scaling limits for naive JCCP	34

1 Introduction

In this paper we study non-Markovian trees modeling time-inhomogeneous populations where individuals live and reproduce independently but where birth-rate b and lifetimes distributions K are time-dependent. In addition, we assume the population starts from a single individual called the *root* or the *ancestor* and that individuals only give birth to a single child at time. In particular, as the time-homogeneous case is known has splitting trees, we simply call such tree *inhomogeneous splitting trees* (IST). The homogeneous case has been introduced in [16] but one of the main step in the study of these trees was done by Lambert [26]. The inhomogeneous case has also been studied recently by Lambert and Popovic [29] and by Lambert and Stadler [30]. In [30], the authors are interested in the so-called *reconstructed phylogenetic tree* of this model (among various others) which summarizes de genealogical relations between the individuals. They question whether this reconstructed tree is a CPP or not. A CPP is a kind of random tree where all the tips lay at a same distance from the root and where the coalescent times of the tips are independent random quantities. In particular, they obtain explicit expression for the branch length distribution which we can recover in our context. We also refer the interested reader to [29, 27].

A powerful techniques to study such model, on which this work relies, is based on coding processes, or contour processes. Contour processes are central objects in the study of trees in probability. They allow to substitute the study of trees with the study of real-valued functions which may be considerably easier in many situations. In particular, one of the most famous result involving contour processes is the proof of the convergence of conditioned Galton-Watson trees to the Brownian tree [2] using the Harris paths (which is a particular type of contour process). This result relies on a Donsker like Theorem. A considerable amount of research involved different types of contour processes of different types of random trees [2, 31, 28, 14].

In this work, we use the so-called Jump Chronological Contour Process (JCCP) which were introduced in [26]. In [26], Lambert studies time-homogeneous splitting trees and introduce the JCCP as a coding process for such tree. It turns out that this stochastic process is a Lévy process. In particular, using the well-developed theory of fluctuations of Lévy processes, this property allowed to understand many properties of the splitting trees and of related models [26, 7, 40, 39]. In more

recent work, Lambert and Uribe [28] demonstrate that JCCPs can be used to code for a far more general family of random trees known as *Totally Ordered Measured trees* (TOM trees). In our case, the JCCP is no longer a Lévy process, and all the arguments used in the homogeneous case fall down.

In our situation, we show that the JCCP loses its space homogeneity. The main consequence is that the theory of fluctuation of Lévy process can't be used in this case preventing to extend the methods of the homogeneous case [5]. However it remains a Feller process and we derive its generator:

$$Lf(x) = f'(x) + b(x) \int_{\mathbb{R}_+} (f(x+y) - f(x)) K(x, dy), \quad x \in \mathbb{R}_+, \quad (1.1)$$

with a full characterization of its domain. Parameters b, K denotes respectively the birth rate and the lifetime distribution; a full construction of the tree and details on these parameters is given in Section 2 below. As a consequence, we can exploit the rich literature on Markov processes [15, 38, 41, 34, 22] to derive properties of the tree.

One of the typical method to study the fluctuation of Markov processes is the scale functions theory. Unfortunately, scale function theories only exist in the case of Lévy or diffusion processes. In this work, we introduce a new type of scale function which is relevant for the contour process and is consistent with the Lévy and diffusion theories. Among other results, this allows us to obtain

Theorem 1.1 (Probability of extinction). *Let $E(v)$ be the probability of extinction of the population starting from an individual of fixed lifetime v . Then, under Assumption 2.1, E satisfies*

$$E(v) = e^{-\int_0^v b(u)du} + e^{-\int_0^v b(u)du} \int_0^v b(s) e^{\int_0^s b(u)du} \int_{[0,+\infty)} E(s+x) K(s, dx) ds, \quad \forall v \geq 0.$$

See Corollary 4.5. This last result allows obtaining necessary and sufficient conditions for the extinction of the population. Another interesting application of our scale functions is the conditioning of trees. Indeed, we are able to obtain the distribution of IST conditioned either to extinction or non-extinction. A surprising aspect is that homogeneous splitting tree conditioned to non-extinction becomes non-homogeneous in time (see Proposition 4.7 below).

In this work, we also obtain more practical criteria of extinction than those who derive from Theorem 1.1 (which are optimal but sometimes abstract). To this end, we use Lyapounov drift methods. In particular, denoting $m(x)$ for the mean value of $K(x, dy)$, we show that “weighted means” of $(m(s)b(s) - 1)$, such as

$$\int_0^\infty \left((m(s)b(s) - 1) e^{-\int_s^\infty b(u)du} \right) ds,$$

can be used to determine if the population extinct almost surely; see Corollary 5.9. This nicely reminds criticality criteria in homogeneous cases as in [26, Proposition 2.2]. Also it enables to derive bounds on the tail of the total length of the tree; see Proposition 5.10. This type of result seems new (even in the homogeneous case) and is in general a difficult question for Markovian tree see for instance [1, 25].

Finally, using convergence methods for Markov processes, we are able to study scaling limits of such trees. We have two main results in this thematic. First, we look at the global shape of

asymptotically critical IST and show that, seen from far away, such tree always looks like Bessel trees. The second result studies sequences $(\mathbb{T}_n)_{n \geq 1}$ of tree with respective parameters K_n and b_n satisfying $b_n \rightarrow \infty$ and $K_n \rightarrow \delta_0$. Our result highlight that the JCCP as introduced in [26] is not enough to study the convergence such sequences of trees, and reveal the need of the more general the theory introduced in [28]. Indeed for such parameters, a process generated by (1.1) can not converge to a diffusive limit; see Theorem 6.7.

In Section 2, we introduce the construction of the inhomogeneous splitting trees. Although this construction is quite standard, the notations of this section will be used in the sequel. Section 3 is devoted to the study of the JCCP of the IST. In particular, Subsection 3.1 is a remainder of [26, 28] on the construction of the JCCP. In Subsection 3.2, we show that the contour process is Markov, has the Feller property and derive its generator. Subsection 3.3 gives another construction of the contour seen as a stand-alone process. In Section 4, we introduce a notion of scale functions for the contour allowing to study many properties of the population. In Section 5, we introduce a notion of criticality for the tree and study it through the contour processes using Lyapounov drift conditions. In the last Section 6, we look at scaling limit of such trees using convergence results on Markov processes.

Finally, as we use different notions of generators in the paper, we write an appendix on these notions in the end of the paper which at least defines properly the terms and notations we use.

2 Description of the model

In this section, we introduce the construction of the time-inhomogeneous population model. Although this construction is classical, we recall it for sake of presentation. This model is described by a branching tree, where individuals live and reproduce independently from each other. The main point is that the birth-rate and the lifetime distribution of the individuals are time-dependent. More precisely, we suppose that the birth-rate is given by some measurable function

$$t \in \mathbb{R}_+ \mapsto b(t) \in \mathbb{R}_+$$

and the lifetime distribution of the individual is given by a probabilistic kernel

$$t \in \mathbb{R}_+ \mapsto K(t, dx) \in \mathcal{M}_1(\overline{\mathbb{R}}_+),$$

where $\mathcal{M}_1(\overline{\mathbb{R}}_+)$ denotes the space of probability measure on the extended positive real line. Such model can be called inhomogeneous splitting tree (IST) in reference to the splitting trees of [26]. Although different from Jagers-Nermann general branching processes, IST are constructed in a similar way. However, contrary to Jagers-Nermann general branching processes, ISTs do not present renewal structures, which is the core point of the study of Jagers-Nermann branching processes.

To give the construction of ISTs, let us introduce some notations. In the following, we denote by \mathcal{U} the so-called Ulam-Harris-Neveu set, that is

$$\mathcal{U} = \bigcup_{n \geq 0} \mathbb{N}^n.$$

This set is meant to label individuals in a genealogically consistent way and to describe the discrete genealogy of the population. For two elements $v = (v_1, \dots, v_n)$ and $w = (w_1, \dots, w_m)$ in \mathcal{U} , let $v.w$ be the concatenation of v and w , that is $v.w = (v_1, \dots, v_n, w_1, \dots, w_m)$ of \mathcal{U} . In addition, for any positive integer k , we denote $v^k = (v_1, \dots, v_k)$ with the convention that $v^k = v$ if $k \geq n$.

As for standard splitting trees, ISTs are random chronological trees. A chronological tree \mathbb{T} is defined as a subset of $\mathcal{U} \times \mathbb{R}_+$ such that $(\sigma, t) \in \mathbb{T}$ if and only if the σ -labeled individual is alive at time t . Hence, every individual $v \in \mathcal{U}$ in the tree is formalized by a subset of the form $\{v\} \times (a, b]$ where a is its birthdate and b is its deathdate. From this remark, it is clear that a subset of $\mathcal{U} \times \mathbb{R}_+$ must satisfy some properties to be admissible as a chronological tree. We do not recall these properties which are well known and refer the interested reader to the original paper of Lambert [26] for more details. In the sequel, $P_{\mathcal{U}}$ (resp. $P_{\mathbb{R}_+}$) denotes the canonical projection of $\mathcal{U} \times \mathbb{R}_+$ to \mathcal{U} (resp. \mathbb{R}_+). In particular, for a chronological tree \mathbb{T} , $P_{\mathcal{U}}(\mathbb{T})$ gives the discrete genealogy of \mathbb{T} .

We now give the construction of ISTs. Let $(\mathcal{N}_u)_{u \in \mathcal{U}}$ be an i.i.d. family of Poisson random measures with common intensity $b(t) dt$, where dt refers to the Lebesgue measure on \mathbb{R}_+ . We recursively define an increasing (for the inclusion order) of sequence trees. Let

$$\begin{cases} T_1 = \{\xi_\emptyset\} \times (0, \xi_\emptyset], \\ B_\emptyset = 0 \end{cases}$$

where ξ_\emptyset is some random variable whose distribution is given later. Now defines recursively,

$$T_n = \bigcup_{v \in P_{\mathcal{G}}(\mathbb{T}_{n-1})} \bigcup_{i \geq 1}^{\mathcal{N}_v(B_v, B_v + \xi_v]} \{v.i\} \times (B_{v.i}, B_{v.i} + \xi_{(v,i)}],$$

with

$$\begin{aligned} B_{v.1} &= \inf\{t > 0 \mid \mathcal{N}_v(B_v, B_v + t] > 0\}, \quad \text{if } \mathcal{N}_v(B_v, B_v + \xi_v] > 0, \\ B_{v.i} &= \inf\{t > B_{(v,i-1)} \mid \mathcal{N}_v(B_v, B_v + t] > i\}, \quad 2 \leq i \leq \mathcal{N}_v(0, \xi_\emptyset], \end{aligned}$$

and $\xi_{v.i}$ is a random variable with conditional distribution with respect to $B_{(v,i)}$ given by $K(B(v, i), dx)$, assuming, in addition, that $\xi_{v.i}$ independent from any other random quantities except from \mathcal{N}_v . Finally, the IST is given by

$$\mathbb{T} = \bigcup_{n \geq 1} T_n.$$

Let us highlight that, for any $\sigma \in \mathbb{P}_{\mathcal{U}}(\mathbb{T})$, B_σ refers to the birthdate of individual σ . Similarly, for any $\sigma \in \mathbb{P}_{\mathcal{U}}(\mathbb{T})$, let us set

$$A_\sigma = B_\sigma + \xi_\sigma,$$

the death-date of individual σ .

This procedure define a probability measure on the space of chronological trees. In the sequel, we consider tree starting from an ancestor whose lifetime is a fixed real number. For this reason, we denote by \mathbb{P}_x the probability distribution on the space of chronological trees satisfying $\mathbb{P}_x(\xi_\emptyset = x) = 1$.

Before ending this section, let us make our general assumptions in order to have a well behaved contour process.

Assumption 2.1 (Regularity assumptions).

1. $x \mapsto b(x)$ is a measurable and locally bounded function.
2. $x \mapsto K(x, dy)$ is weakly continuous in $\mathcal{M}_1(\overline{\mathbb{R}}_+^*)$.

3 The contour process of an inhomogeneous splitting tree

3.1 Construction of the contour process

In this section, we briefly recall some facts on the construction of the JCCP which is due to Lambert in [26]. Recently, Lambert endowed this construction in the more general framework of TOM trees (totally ordered measured trees) which we briefly recall here because it is used in Section 6. A TOM tree $(\mathcal{T}, \leq, \lambda)$ is a real tree (see e.g. [31] for the definition) equipped with a total order \leq and a σ -finite measure λ satisfying some properties (see [28]). Given the order relation \leq , one can define the left of an element x of \mathcal{T} by

$$L_x = \{y \in \mathcal{T} \mid y \leq x\},$$

and the inverse exploration process by

$$\begin{aligned} \varphi^{-1} : \mathcal{T} &\rightarrow \mathbb{R}_+, \\ x &\mapsto \lambda(\{y \mid y \leq x\}). \end{aligned}$$

As explain in [28], the exploration process φ is the unique càd-làg extension of the generalized inverse of φ^{-1} . This construction is used in the next section in order to characterize the law of this process when the underlying chronological tree is an IST.

In our particular case, we use the formalism introduced in [26] which makes the endowed the trees with a total order and measured structure. Let \mathbb{T} be some chronological tree. In [26], Lambert introduce on \mathbb{T} a measure λ and a total order relation \leq which can be summarized as follows: for two elements (σ, t) and (δ, s) , we say that $(\sigma, t) \leq (\delta, s)$ if and only if (σ, t) and (δ, s) satisfy one of the two following condition (see also Figure 1):

$$\begin{cases} \delta \preceq \sigma \text{ and } P_{\mathbb{R}}((\delta, s) \wedge (\sigma, t)) \geq s & (C1) \\ \text{or} \\ \exists n \in \mathbb{N}, \sigma^n \preceq \delta \text{ and } t > B_{\sigma^n}, & (C2) \end{cases}$$

where \preceq is the classical order relation on \mathcal{U} . Informally speaking, the measure λ can be thought as the length of the segments in the tree. We refer the reader to [26] for more details. In the sequel, we set

$$\mathcal{L}(\mathbb{T}) := \lambda(\mathbb{T}) \text{ and } \mathcal{H}(\mathbb{T}) := \sup P_{\mathbb{R}}(\mathbb{T}), \quad (3.1)$$

which are respectively the total length and the height of the tree \mathbb{T} . In particular, $\mathcal{H}(\mathbb{T}) < \infty$ means that the population gets extinct in finite time. This event is denoted

$$\text{Ext} := \{\mathcal{H}(\mathbb{T}) < \infty\}. \quad (3.2)$$

Let us now recall an important result concerning finite chronological tree.

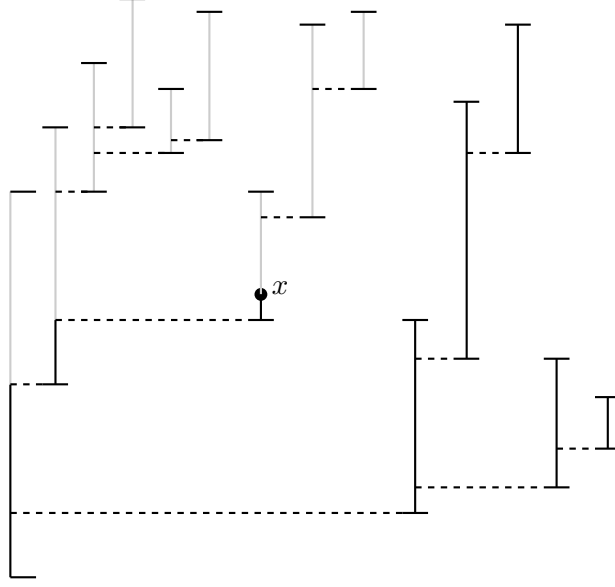


Figure 1: In gray, the set $\{y \in \mathbb{T} \mid y \leq x\}$

Proposition 3.1 (Lambert 2010). *Let \mathbb{T} be a chronological tree such that $\mathcal{L}(\mathbb{T}) < \infty$. Then the function $(\mathcal{E}_s(\mathbb{T}), s \in [0, \mathcal{L}(\mathbb{T})])$ defined by*

$$\mathcal{E}_s(\mathbb{T}) = \inf \{x \in \mathbb{T} \mid \lambda(\{y \in \mathbb{T} \mid y \leq x\}) \geq s\}, \quad s \in [0, \mathcal{L}(\mathbb{T})].$$

is an increasing bijection from $[0, \mathcal{L}(\mathbb{T})]$ to \mathbb{T} . This process is called the exploration process of \mathbb{T} (see Figure 2).

An important consequence of the construction of the contour process which is used in the sequel is that, for any point (δ, t) of t ,

$$\{\mathcal{E}_s(\mathbb{T}) \mid s \leq \mathcal{E}_{(\delta, t)}^{-1}(\mathbb{T})\} = \{x \in \mathbb{T} \mid x \leq (\delta, t)\}. \quad (3.3)$$

In other words, the part of the tree which is explored by the exploration process up to the exploration of (δ, t) is exactly the left of (δ, t) .

From $(\mathcal{E}_s(\mathbb{T}), s \in [0, \mathcal{L}(\mathbb{T})])$ the contour process is now defined by

$$\mathcal{C}_s(\mathbb{T}) = P_{\mathbb{R}_+}(\mathcal{E}_s(\mathbb{T}))\mathbb{1}_{s \leq \mathcal{L}(\mathbb{T})}, \quad s \in \mathbb{R}_+.$$

This means that the contour process is the height in the tree of the exploration process at a given time until the exploration process hits the point $(\emptyset, 0)$ at time $\mathcal{L}(\mathbb{T})$. After time $\mathcal{L}(\mathbb{T})$, the contour remains equal to 0.

Remark 3.2 (Truncated tree). *In the following, we consider trees \mathbb{T} with infinite total length (i.e. $\mathcal{L}(\mathbb{T}) = \infty$). For such tree, one cannot define properly the contour process. To avoid this technicality, we consider truncated trees above some threshold associated to tree \mathbb{T} . More precisely, for some time t , the truncated tree \mathbb{T}^t of \mathbb{T} above level t is defined by*

$$(\sigma, s) \in \mathbb{T}^t \Leftrightarrow ((\sigma, s) \in \mathbb{T} \text{ and } s \leq t).$$

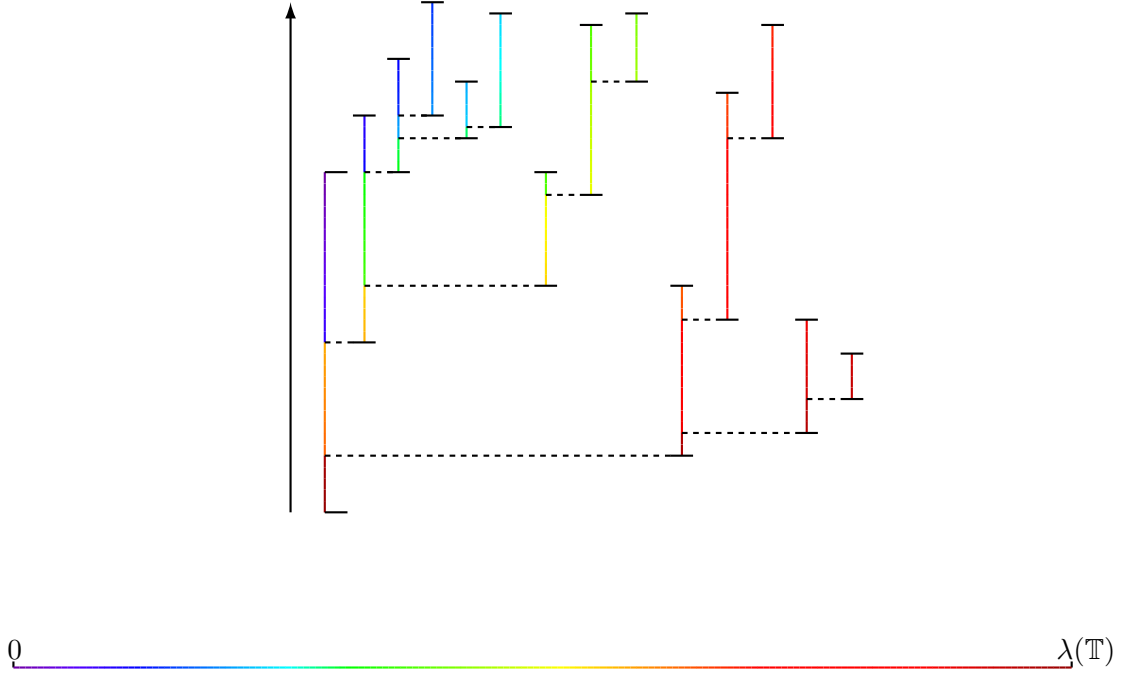


Figure 2: A graphical representation of the exploration process. The one-to-one correspondence between $[0, \lambda(\mathbb{T})]$ and \mathbb{T} is represented by corresponding colors.

An important remark is that the time-inhomogeneity of tree become a space in-homogeneity for the contour process (see Figure 3). Hence, the process remains time-homogeneous but is not a Lévy process, as in the case of splitting trees, due to space-inhomogeneity.

3.2 Markov and Feller properties

The purpose of this section is to study the law of the contour process of an IST. So let \mathbb{T} be an IST with birth-rate b and birth-kernel K . To study the contour process of \mathbb{T} , we need to introduce a modification of the tree \mathbb{T} called contracted tree. This modification is useful in order to understand dependencies in the tree. Let (σ, s) be an element of \mathbb{T} , the contracted tree $\mathbb{T}_{(\sigma, s)}$ of \mathbb{T} at point (σ, t) is defined by (see also Figure 4)

$$\forall (\delta, t) \in \mathbb{T}, \{(\delta, t) \in \mathbb{T}_{(\sigma, s)} \Leftrightarrow \exists n \in \mathbb{N}, (\sigma^n \cdot \delta, t) \in \mathbb{T} \text{ and } B_{\sigma^n \cdot \delta^1} \leq B_{\sigma^{n+1}} \wedge s\}.$$

Our first step is to show that the contour process is a Feller Markov process. To this end, we need the following proposition which use the contracted tree to understand the dependencies in the tree seen from the exploration process.

Proposition 3.3. *Let \mathbb{T} be an IST with birth-kernel K and birth-rate b . Let X be a random variable with value in $\mathcal{U} \times \mathbb{R}_+$ almost surely in \mathbb{T} , and $(\mathcal{G}_s, s \in \mathbb{R}_+)$ be the natural filtration associated to $(\mathcal{C}_s(\mathbb{T}), s \in \mathbb{R}_+)$. Then, \mathbb{T}_X is independent of $\mathcal{G}_{\mathcal{E}_X^{-1}(\mathbb{T})}^{-1}$ conditionally on $P_{\mathbb{R}}(X)$.*

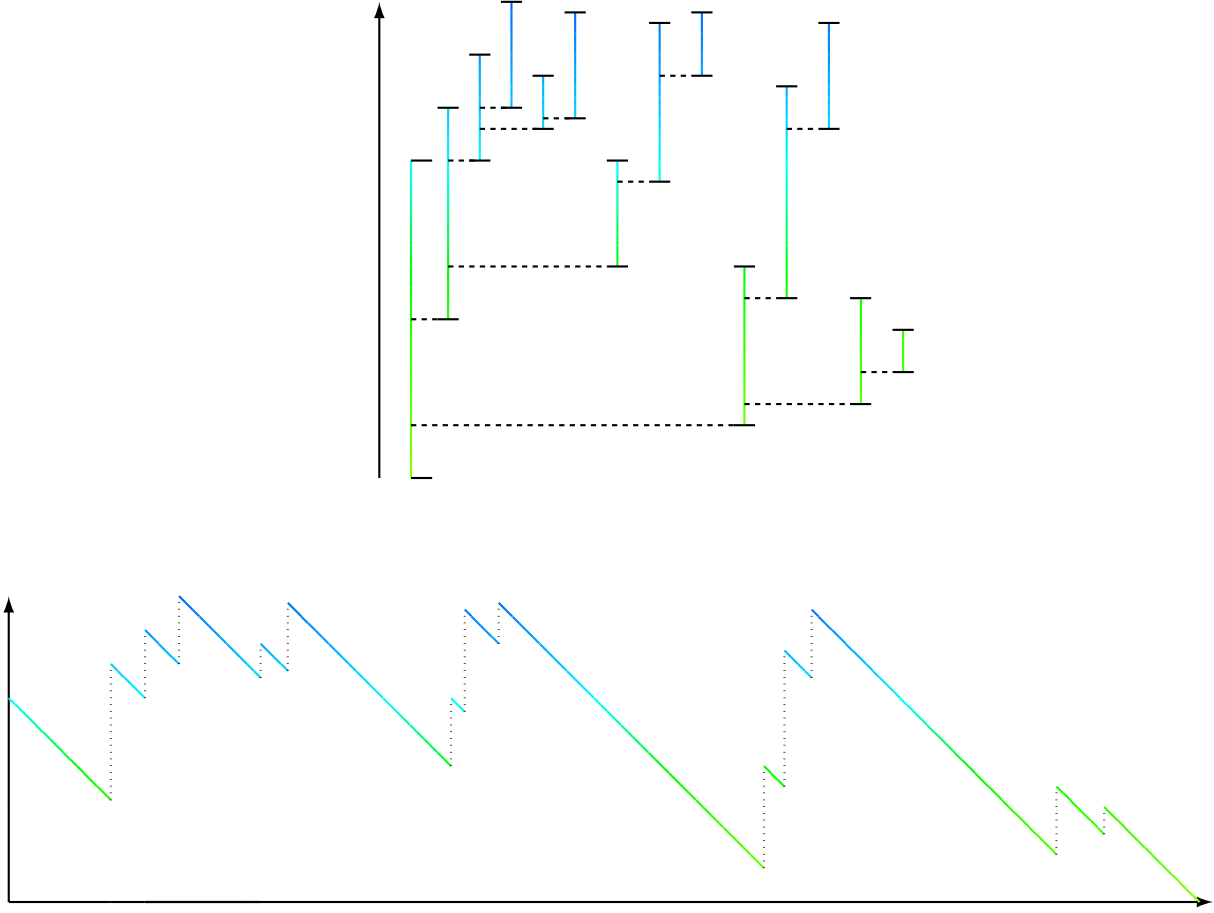


Figure 3: IST with varying birth-rate. The color on both contour and tree denotes the value of the birth-rate.

Proof. In order to lighten notations, let us denote, for any $x \in \mathbb{T}$,

$$\xi_x = \xi_{P_U(x)} \text{ and } \mathcal{N}_x = \mathcal{N}_{P_U(x)},$$

where we recall that ξ_σ and \mathcal{N}_σ stands respectively for the lifetime and the birth point process of individual $\sigma \in \mathbb{T}$, as defined in Section 2.

Now, let Z be a random variable measurable with respect to

$$\sigma \left\{ \xi_{\mathcal{E}_s(\mathbb{T})}, \mathcal{N}_{\mathcal{E}_s(\mathbb{T})} \left(P_{\mathbb{R}_+}(\{x \in \mathbb{T} \mid x \leq X\}) \cap \cdot \right) \mid s \leq \mathcal{E}_X^{-1}(\mathbb{T}) \right\} \supset \mathcal{G}_{\mathcal{E}_X^{-1}(\mathbb{T})},$$

where the inclusion follows from (3.3). Hence, there exists a measurable function ψ such that

$$Z = \psi \left(\left(\xi_{\mathcal{E}_s(\mathbb{T})} \right)_{s \leq \mathcal{E}_X^{-1}(\mathbb{T})}, \left(\mathcal{N}_{\mathcal{E}_s(\mathbb{T})} \left(P_{\mathbb{R}_+}(\{x \in \mathbb{T} \mid x \leq X\}) \cap \cdot \right) \right)_{s \leq \mathcal{E}_X^{-1}(\mathbb{T})} \right).$$

By construction, if $y \in \mathbb{T} \setminus \{x \in \mathbb{T}, x \leq X\}$ and $P_U(y) \neq P_U(X)$, then ξ_y and \mathcal{N}_y are independents of Z . Consequently, for any y in $\mathbb{T}_X \setminus \{\emptyset\} \times \mathbb{R}_+$, V_y and \mathcal{N}_y are independents of Z . So, it only

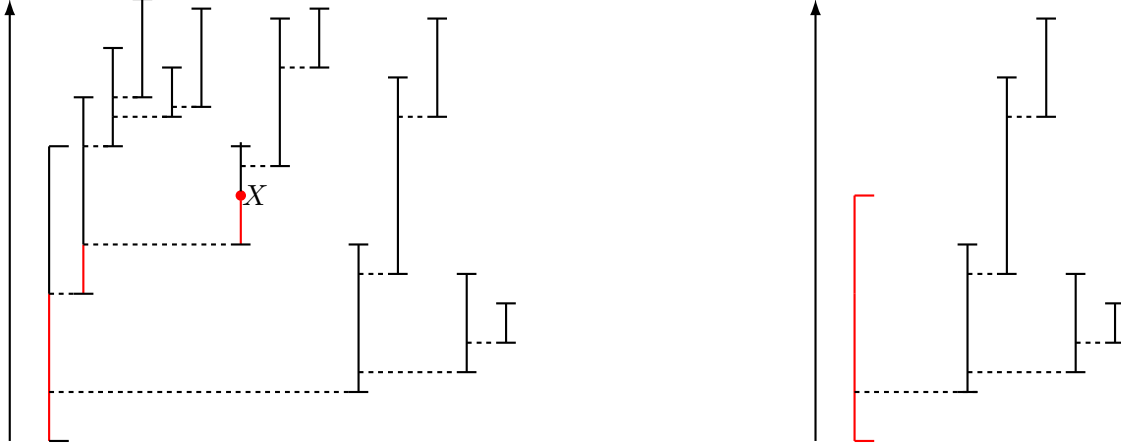


Figure 4: The original tree \mathbb{T} (left) and the contracted tree \mathbb{T}_X (right).

remains to understand the dependencies between Z and the root of \mathbb{T}_X . Now, denote by \mathcal{P} the point process of births associated to the root of \mathbb{T}_X . By construction, for any measurable set A ,

$$\mathcal{P}(A) = \sum_{v \preceq P_{\mathcal{U}}(X)} \mathcal{N}_v(A \cap \{s \in \mathbb{R}_+ \mid (v, t) \preceq X\}),$$

where $(v, t) \preceq X$ if and only if $v \preceq P_{\mathcal{U}}(X)$ and $X \leq (v, t)$. In particular, the set $\{x \in \mathbb{T} \mid x \preceq X\}$ corresponds to the path in the tree which link the bottom of the root to point X (see Figure 4). Now, since $\{x \in \mathbb{T} \mid x \preceq X\}$ and $\{x \in \mathbb{T} \mid x \leq X\}$ are disjoint sets, we get, using the independence properties of Poisson random measure, that \mathcal{P} is independent of Z . Hence, the dependence between Z and \mathbb{T}_X only relies on the lifetime of the root. Finally, since the lifetime of the root of \mathbb{T}_X is given by $P_{\mathbb{R}}(X)$, we can conclude that \mathbb{T}_X is independent of Z conditionally on $P_{\mathbb{R}}(X)$. This ends the proof. \square

We can now show that the contour process of a truncated IST is a Feller process.

Proposition 3.4. *Let \mathbb{T} be an IST with birth-kernel K and birth-rate b satisfying Assumptions 2.1. Let $(\mathbb{P}_x)_{x \in \mathbb{R}_+}$ a family of probability measure such that $\mathbb{P}_x(\xi_0 = x) = 1$. Denote, for any positive real number t , \mathbb{T}^t the truncated tree above t as defined in Remark 3.2. Then, $((\mathcal{C}_s(\mathbb{T}^t), s \in \mathbb{R}_+), \mathbb{P}_x)$ is a Feller process.*

Proof. We begin the proof by showing that $(\mathcal{C}_s(\mathbb{T}^t), s \in \mathbb{R}_+)$ is a Markov process. To this end let us denote by $\mathcal{F} = (\mathcal{F}_s)_{s \in \mathbb{R}_+}$ the natural filtration associated with $(\mathcal{C}_s(\mathbb{T}^t), s \in \mathbb{R}_+)$ and, as above, $\mathcal{G}^t = (\mathcal{G}_s^t)_{s \in \mathbb{R}_+}$ the natural filtration of $(\mathcal{E}_s(\mathbb{T}^t), s \in \mathbb{R}_+)$. Let f be some bounded measurable function. We have, for any positive real numbers s and h ,

$$\mathbb{E} [f(\mathcal{C}_{s+h}(\mathbb{T}^t)) \mid \mathcal{G}_s^t] = \mathbb{E} [f(\mathcal{C}_h(\mathbb{T}_{\mathcal{E}_s(\mathbb{T}^t)}^t)) \mid \mathcal{G}_s^t].$$

However, according to Proposition 3.3, $\mathbb{T}_{\mathcal{E}_s(\mathbb{T}^t)}^t$ is independent of \mathcal{G}_s^t conditionally on $P_{\mathbb{R}}(\mathcal{E}_s(\mathbb{T}^t)) = \mathcal{C}_s(\mathbb{T}^t)$. Consequently,

$$\mathbb{E} [f(\mathcal{C}_h(\mathbb{T}_{\mathcal{E}_s(\mathbb{T}^t)}^t)) \mid \mathcal{G}_s^t] = \mathbb{E} [f(\mathcal{C}_h(\mathbb{T}_{\mathcal{E}_s(\mathbb{T}^t)}^t)) \mid \mathcal{C}_s(\mathbb{T}^t)],$$

which gives the Markov property. As a consequence of the Markov property,

$$f \mapsto \mathbb{E}_x [f(\mathcal{C}_s(\mathbb{T}^t))]$$

defines a semigroup $(P_s, s \in \mathbb{R}_+)$ on the space $\mathcal{B}_b([0, t])$ of bounded measurable function on $[0, t]$. To prove the Feller property, it remains to show that the subspace $\mathcal{C}([0, t])$ of continuous function on $[0, t]$ is invariant under the action of P_t , for any $t \geq 0$.

Let f be an element of $\mathcal{C}([0, t])$ and x, y be two positive real numbers such that $t \geq x > y$. Using Markov property and denoting by T the time of first jump of $(\mathcal{C}_s(\mathbb{T}^t), s \in \mathbb{R}_+)$, we have, with $s > (x - y)$, that

$$\begin{aligned} \mathbb{E}_x [f(\mathcal{C}_s(\mathbb{T}^t))] &= \mathbb{E}_x [f(\mathcal{C}_s(\mathbb{T}^t)) \mathbf{1}_{T > x-y}] + \mathbb{E}_x [f(\mathcal{C}_s(\mathbb{T}^t)) \mathbf{1}_{T \leq x-y}] \\ &= \mathbb{E}_y [f(\mathcal{C}_{s-(x-y)}(\mathbb{T}^t))] \mathbb{P}_x(T > x-y) + \mathbb{E}_x [f(\mathcal{C}_s(\mathbb{T}^t)) \mathbf{1}_{T \leq x-y}]. \end{aligned}$$

Now, by construction

$$\mathbb{P}_x(T > x-y) = \mathbb{P}(\mathcal{N}_\emptyset([x, y]) = 0) = \exp\left(-\int_y^x b(u) du\right).$$

Consequently,

$$\begin{aligned} |\mathbb{E}_x [f(\mathcal{C}_s(\mathbb{T}^t))] - \mathbb{E}_y [f(\mathcal{C}_s(\mathbb{T}^t))]| &\leq \left| P_{s-(x-y)}f(y) \exp\left(-\int_y^x b(u) du\right) - P_sf(y) \right| \\ &\quad + \|f\|_\infty \left(1 - \exp\left(-\int_y^x b(u) du\right)\right) \\ &\leq |P_{s-(x-y)}f(y) - P_sf(y)| + 2 \left(1 - \exp\left(-\int_y^x b(u) du\right)\right) \|f\|_\infty. \end{aligned}$$

Now, using once again the Markov property, we have that

$$\begin{aligned} \mathbb{E}_y [f(\mathcal{C}_{s-(x-y)}(\mathbb{T}^t))] &= \mathbb{E}_y \left[f(\mathcal{C}_s(\mathbb{T}^t) + (x-y)) \exp\left(-\int_{\mathcal{C}_s(\mathbb{T}^t)}^{\mathcal{C}_s(\mathbb{T}^t)+(x-y)} b(u) du\right) \right] \\ &\quad + \mathbb{E}_y \left[f(\mathcal{C}_{s-(x-y)}(\mathbb{T}^t)) \left(1 - \exp\left(-\int_{\mathcal{C}_s(\mathbb{T}^t)}^{\mathcal{C}_s(\mathbb{T}^t)+(x-y)} b(u) du\right)\right) \right], \end{aligned}$$

leading to

$$\begin{aligned} &|P_{s-(x-y)}f(y) - P_sf(y)| \\ &\leq |\mathbb{E}_y [f(\mathcal{C}_s(\mathbb{T}^t)) - f(\mathcal{C}_s(\mathbb{T}^t) + (x-y))]| + 2\mathbb{E} \left[\left(1 - \exp\left(-\int_{\mathcal{C}_s(\mathbb{T}^t)}^{\mathcal{C}_s(\mathbb{T}^t)+(x-y)} b(u) du\right)\right) \right] \|f\|_\infty. \end{aligned}$$

Finally, we obtain, for any $x > y$ and $s > (x - y)$,

$$\begin{aligned} &|P_sf(x) - P_sf(y)| \\ &\leq |\mathbb{E}_y [f(\mathcal{C}_s(\mathbb{T}^t)) - f(\mathcal{C}_s(\mathbb{T}^t) + (x-y))]| + 4(1 - \exp(-(x-y)\|b\|_\infty)) \|f\|_\infty. \quad (3.4) \end{aligned}$$

Now, we have thanks to Equation 3.4 and Lebesgue's dominated convergence theorem that

$$|P_s f(x) - P_s f(y)| \xrightarrow{x \rightarrow y} 0,$$

for any f in $\mathcal{C}([0, t])$ and any $s \geq 0$. Hence, $P_s f$ is continuous on $(0, t]$. So it remains to prove that $P_s f$ is continuous at 0. To this end, let x be a positive real number, we have

$$|\mathbb{E}_x [f(\mathcal{C}_s(\mathbb{T}^t))] - f(0)| \leq |f(0)\mathbb{P}_x(T > x) - f(0)| + \|f\|_\infty(1 - \mathbb{P}_x(T > x)).$$

So according to the above computation, this gives the continuity of f at point 0. Hence, the space $\mathcal{C}([0, t])$ is invariant under the action of P_s for any $s \geq 0$, which gives the Feller property. \square

Since the contour process is a Feller process, it induces a strongly continuous semi-group on the space $\mathcal{C}([0, t])$ of continuous function on $[0, t]$. Our next goal is to obtain the generator of the process as well as its domain. Looking at the behavior of the tree and, thus, of the contour, a natural candidate is given by

$$L^{(t)} f(x) = -f'(x) + b(x) \int_{\mathbb{R}_+} (f((x+y) \wedge t) - f(x)) K(x, dy)$$

with domain

$$D(L^{(t)}) = \left\{ f \in \mathcal{C}^1([0, t]), L^{(t)} f(0) = 0 \right\}.$$

Our strategy relies on two steps: the first step is to show that

$$\frac{1}{t} (\mathbb{E}_x [f(\mathcal{C}_t(\mathbb{T}))] - f(x)) \xrightarrow{t \rightarrow 0} L^{(t)} f,$$

in $\mathcal{C}([0, t])$. But this is not enough to ensure that $(L^{(t)}, D(L^{(t)}))$ is the generator of the contour. Hence, our second step is to show that $(L^{(t)}, D(L^{(t)}))$ generate a strongly continuous semi-group on $\mathcal{C}([0, t])$. As consequence, using the maximality property of such operator (see [21, Lemma 17.12]), we obtain that $(L^{(t)}, D(L^{(t)}))$ is indeed the generator of the contour. We begin by showing that $(L^{(t)}, D(L^{(t)}))$ generate a strongly continuous semi-group.

Proposition 3.5. *Let*

$$L^{(t)} f(x) = -f'(x) + b(x) \int_{\mathbb{R}_+} (f((x+y) \wedge t) - f(x)) K(x, dy)$$

on $\mathcal{C}([0, t])$ defined on the domain

$$D(L^{(t)}) = \left\{ f \in \mathcal{C}^1([0, t]), L^{(t)} f(0) = 0 \right\}.$$

Then $(L^{(t)}, D(L^{(t)}))$ generate a strongly continuous semi-group on $C_0(\mathbb{R}_+)$.

Proof. In the following, the supremum norm $\|\cdot\|_\infty$ refers to the one of $\mathcal{C}([0, t])$. Let us first consider the operator

$$Tf = -f' - 2\|b\|_\infty f,$$

defined on the domain

$$D(T) = \left\{ f \in \mathcal{C}^1([0, t]), L^{(t)} f(0) = 0 \right\}.$$

It is easily seen that $L^{(t)}$ is a perturbation of T by a bounded operator. Hence, deducing that $L^{(t)}$ generate a strongly continuous semi-group is easy if T does. Hence, we begin to show, using Lumer-Philips Theorem (see [37, Chapter 1, Theorem 4.3]), that T generate a strongly continuous semi-group. First of all, note that, since $L^{(t)}$ is continuous in $\mathcal{C}^1([0, t])$, then its kernel is closed in this space, making T a closed operator.

Let f in $D(T)$ and x^* be a global maximum of $|f|$.

It is, hence, easily seen that the measure $f(x^*)\delta_{x^*}$ belongs to the duality set of f [37, Chapter 1, Section 1.4]. In addition, we have

$$\langle -f' - 2\|b\|_\infty f, f(x^*)\delta_{x^*} \rangle = (-f'(x^*) - 2\|b\|_\infty f(x^*)) f(x^*) = -\frac{1}{2}(f^2)'(x^*) - 2\|b\|_\infty f(x^*)^2.$$

Since x^* is an extremum of f ,

$$\langle -f' - \|b\|_\infty f, f(x^*)\delta_{x^*} \rangle \leq -\|b\|_\infty f(x^*)^2, \quad (3.5)$$

as soon as $x^* \in (0, t)$. On the other hand, x^* is a maximum for f^2 , so (3.5) also holds if $x^* = t$. Finally, if $x^* = 0$, we have

$$\begin{aligned} (-f'(x^*) - 2\|b\|_\infty f(x^*)) f(x^*) &= -\|f\|_\infty b(0) \int_{\mathbb{R}_+} f(y) K(0, dy) - (2\|b\|_\infty - b(0))\|f\|_\infty^2 \\ &\leq (b(0) - \|b\|_\infty)\|f\|_\infty^2. \end{aligned}$$

Finally, in any case, $\langle -f' - \|b\|_\infty f, f(x^*)\delta_{x^*} \rangle \leq 0$ which implies that T is a dissipative operator (see [37, Chapter 1, Section 1.4]). In addition, si $L^{(t)}$ is unbounded, its kernel is dense in $C([0, t])$. To apply Lumer-Philips theorem, it remains to show that the range of $\lambda - T$ is dense in $C([0, t])$. However, taking $g \in C([0, t])$, solving the problem

$$\begin{cases} \lambda f + f' = g \\ L^{(t)} f(0) = 0, \end{cases}$$

is straightforward using elementary ODE solving techniques. As a consequence, T generate a strongly continuous semi-group over $C([0, t])$.

Finally, since

$$f \mapsto b(x) \int_{\mathbb{R}_+} (f((x+y) \wedge t) - f(x)) K(x, dy)$$

is a bounded operator over $C([0, t])$, $L^{(t)}$ is obtain from T through the perturbation by a bounded operator. As a consequence, $L^{(t)}$ is a closed operator with the same domain as T and generating a strongly continuous semi-group (see [37, Chapter 3, Theorem 1.1]). \square

We can finally show that $(L^{(t)}, D(L^{(t)}))$ is the generator of the truncated contour process.

Theorem 3.6. Let \mathbb{T} with birth-kernel K and birth-rate b . Let $(\mathbb{P}_x)_{x \in \mathbb{R}_+}$ a family of probability measure such that $\mathbb{P}_x(\xi_\emptyset = x) = 1$. Denote, for any positive real number t , \mathbb{T}^t the truncated tree above t . Then, $((\mathcal{C}_s(\mathbb{T}^t), s \in \mathbb{R}_+), \mathbb{P}_x)$ is a Feller process, with generator

$$L^{(t)} f(x) = -f'(x) + b(x) \int_{\mathbb{R}_+} (f((x+y) \wedge t) - f(x)) K(x, dy),$$

on the domain

$$D(L^{(t)}) = \left\{ f \in \mathcal{C}^1([0, t]), L^{(t)} f(0) = 0 \right\}.$$

Proof. We can now derive the generator of this process. Assume now that f belongs to the set $D(L^{(t)})$ defined above. In order to lighten notation, $\mathcal{E}_s(\mathbb{T}^t)$ is now simply denoted \mathcal{E}_s .

$$\mathbb{E}_x [f(\mathcal{C}_h)] = \mathbb{E}_x [f(\mathcal{C}_h) \mathbf{1}_{\Delta \mathcal{C}([0, h])=0}] + \mathbb{E}_x [f(\mathcal{C}_h) \mathbf{1}_{\Delta \mathcal{C}([0, h])=1}] + \mathbb{E}_x [f(\mathcal{C}_h) \mathbf{1}_{\Delta \mathcal{C}([0, h])>1}],$$

where $\Delta \mathcal{C}([0, h])$ denotes the number of jump of \mathcal{C} over the time interval $[0, h]$. By construction of the exploration process $\Delta \mathcal{C}([0, h]) = 1$ if and only if $\mathcal{N}_\emptyset([x-h, x]) > 0$ and $\mathcal{N}_{N_\emptyset}([x+\xi_{N_\emptyset}-, x]) = 0$. Hence,

$$\mathbb{E}_x [f(\mathcal{C}_h) \mathbf{1}_{\Delta \mathcal{C}([0, h])=1}] = \mathbb{E}_x \left[f(x + V_{N_\emptyset}) \mathbf{1}_{\mathcal{N}_\emptyset([x-h, x])>0} \mathbf{1}_{\mathcal{N}_{N_\emptyset}([x+\xi_{N_\emptyset}-h, x+\xi_{N_\emptyset}-T_1])=0} \right].$$

By construction, ξ_{N_\emptyset} , $\mathcal{N}_{N_\emptyset}$ and \mathcal{N}_\emptyset are independent. It follows that

$$\begin{aligned} \mathbb{E}_x [f(\mathcal{C}_h) \mathbf{1}_{\Delta \mathcal{C}([0, h])=1}] &= \mathbb{E}_x \left[f(x + \xi_{N_\emptyset}) \mathbf{1}_{\mathcal{N}_\emptyset([x-h, x])>0} \mathbf{1}_{\mathcal{N}_{N_\emptyset}([x+\xi_{N_\emptyset}-h, x+\xi_{N_\emptyset}-T_1])=0} \right] \\ &= \mathbb{E}_x \left[\int_{\mathbb{R}_+} f(x+y) \mathbf{1}_{T \leq h} \mathbf{1}_{\mathcal{N}_{N_\emptyset}([x+y-h, x+y-T_1])=0} K(x-T, dy) \right] \\ &= \int_{[0, h]} \int_{\mathbb{R}_+} f(x+y) e^{-\int_{x+y-h}^{x+y-z} b(u) du} K(x-z, dy) b(x-z) e^{-\int_{x-z}^x b(u) du} dz. \end{aligned}$$

Set

$$I = \int_{[0, h]} b(x-z) e^{-\int_{x-z}^x b(u) du} dz \int_{\mathbb{R}_+} f(x+y) K(x, dy).$$

Now, it is easily seen that

$$\begin{aligned} &\left| \int_{[0, h]} \int_{\mathbb{R}_+} f(x+y) e^{-\int_{x+y-h}^{x+y-z} b(u) du} K(x-z, dy) b(x-z) e^{-\int_{x-z}^x b(u) du} dz - I \right| \\ &\leq 3 \left| \int_{[0, h]} \int_{\mathbb{R}_+} f(x+y) (K(x-z, dy) - K(x, dy)) b(x-z) e^{-\int_{x-z}^x b(u) du} dz \right| \\ &\leq 3h \|b\|_{L^\infty([0, t])} \sup_{z \in [0, h]} |\langle K(x, \cdot) - K(x-z, \cdot), f \rangle|, \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality bracket between $C([0, t])$ and its dual. But since $x \mapsto K(x, dy)$ is weakly continuous, we get that

$$h \|b\|_{L^\infty([0, t])} \sup_{z \in [0, h]} |\langle K(x, \cdot) - K(x - z, \cdot), f \rangle| = o(h).$$

Moreover,

$$\frac{1}{h} I \xrightarrow{h \rightarrow 0} b(x) \int_{\mathbb{R}_+} f(x + y) K(x, dy),$$

we obtain the same limit for $h^{-1} \mathbb{E}_x [f(\mathcal{C}_h) \mathbf{1}_{\Delta \mathcal{C}([0, h])=1}]$. Using similar argument, one can also get that

$$\mathbb{E}_x [f(\mathcal{C}_h) \mathbf{1}_{\Delta \mathcal{C}([0, h])=0}] = f(x) e^{\int_{x-h}^x b(u) du},$$

and

$$\mathbb{E}_x [f(\mathcal{C}_h) \mathbf{1}_{\Delta \mathcal{C}([0, h])>1}] \leq \|f\|_\infty \mathbb{P}(\mathcal{P}([0, h]) > 1),$$

where \mathcal{P} is a Poisson random measure on $[0, t]$ with constant intensity given by $\|b\|_{L^\infty([0, t])}$, which implies that

$$\mathbb{E}_x [f(\mathcal{C}_h) \mathbf{1}_{\Delta \mathcal{C}([0, h])>1}] = o(h),$$

uniformly in x . Hence, we get that

$$\begin{aligned} \left| \frac{1}{h} (\mathbb{E}_x [f(\mathcal{C}_h)] - f(x)) - L^{(t)} f(x) \right| &\leq \left| \frac{1}{h} (f(x - h) - f(x)) e^{-\int_{x-h}^x b(u) du} + f'(x) \right| \\ &\quad + \left| b(x) - \frac{1}{h} \left(1 - e^{-\int_{x-h}^x b(u) du} \right) \right| \|f\|_\infty + o(1). \end{aligned}$$

Using the uniform continuity of f' and b on $[0, t]$ finally gives the desired result. \square

3.3 A jump process point of view for the contour

Let us fix $T \in \mathbb{R}_+ \cup \{+\infty\}$. In this section, we define a Markov process $(X_t^{(T)})_{t \geq 0}$, which is distributed as $(\mathcal{C}_s(\mathbb{T}^T))_{s \geq 0}$ but whose the construction is separated from the tree. The case $T = +\infty$ corresponds to the non-truncated tree \mathbb{T} and we omit the symbol (T) , i.e. $(X_t^{(\infty)})_{t \geq 0} = (X_t)_{t \geq 0}$. In addition, for being relevant with the tree notation, we also denote by \mathbb{P}_x the probability measure on $\mathbb{D}(\mathbb{R}_+)$ satisfying $\mathbb{P}_x(X_0^T = x) = 1$, for $0 < T \leq \infty$.

This process has $[0, T]$ (or \mathbb{R}_+ when $T = +\infty$) as state space and evolves deterministically and linearly (with slope -1) between some positive jumps. These jumps arise at the non-homogeneous rate b and the sizes are given by the kernel K_T satisfying for any function f and $x \in [0, T]$,

$$\int_{\mathbb{R}_+} f(y) K_T(x, dy) = \int_{\mathbb{R}_+} f(y \wedge T) K(x, dy).$$

The process is also supposed to be absorbed at 0. Let us give further explanations on its dynamics. Starting from some $x > 0$, $(X^{(T)})_{t \geq 0}$ experiences its first jump at some random time $\tau \in \mathbb{R}_+ \cup \{+\infty\}$ such that

$$\forall t \geq 0, \quad \mathbb{P}(\tau > t) = \exp \left(- \int_0^{t \wedge x} b(x - s) ds \right),$$

and $\mathbb{P}(\tau = \infty) = \exp\left(-\int_0^x b(u)du\right)$, corresponding to the absorption of the process at 0. For $x \in [0, \tau)$, $X_s^T = (x - s)\mathbf{1}_{s \leq x}$, and conditionally on the event $\{\tau < \infty\} = \{\tau < x\}$, the jump size $(X_\tau^{(T)} - X_{\tau-}^{(T)})$ is distributed according to $K_T(X_{\tau-}^{(T)}, \cdot) = K_T(x - \tau, \cdot)$. After this jump time, the process $(X_t^{(T)})_{t \geq 0}$ recursively follows these preceding steps.

If $T < +\infty$ then, since b is locally bounded (see Assumptions 2.1) and jumps are bounded, it is easy to see that the process is well defined. Namely, there is no explosion, that is the sequence of jump times does not converge to some finite value, namely if $(T_k)_{k \geq 0}$ is the sequence of jump times of $(X_t)_{t \geq 0}$, we have $\lim_{k \rightarrow \infty} T_k = +\infty$ (even if it can be a finite sequence whose the last term is $+\infty$). This may not be the case when $T = +\infty$. Note that this definition of explosion is consistent with the definition of [36].

Due to the exponential-like distribution of τ and its iterated construction, $(X_t^{(T)})_{t \geq 0}$ is a strong Markov process. It is neither a diffusion process (because it is not continuous) nor a Lévy process (because it is space-inhomogeneous). This type of process belongs to the class of Piecewise Deterministic Markov Process (PDMP). Introduced in [9], this class of processes has recently motivated a considerable amount of research in various context (see for instance [3, 8, 33, 11] for surveys). Among many results, simulation algorithms can be found in [10, 11, 32].

The present section is devoted to the study of this particular process. In particular, we can derive its generator as above and, even, characterize its extended generator.

Theorem 3.7 (Strong generator). *Under Assumptions 2.1, if $T < +\infty$, then $(X_t^T)_{t \geq 0}$ is a Feller process with full generator*

$$\begin{aligned} L^{(T)}f(x) &= -f'(x) + b(x) \int_{\mathbb{R}_+} (f((x+y) \wedge T) - f(x)) K(x, dy) \\ &= -f'(x) + b(x) \int_{\mathbb{R}_+} (f(x+y) - f(x)) K_T(x, dy), \end{aligned}$$

on $C([0, T])$ defined on the domain

$$D(L^{(T)}) = \left\{ f \in C^1([0, T]), L^{(T)}f(0) = 0 \right\}.$$

Proof. The Feller property comes from [10, Theorem (27.6) p 27]. The remaining of the proof is analogous to the one of Theorem 3.6. \square

An easy consequence is the following (see for instance [15, Proposition 1.2.9]).

Corollary 3.8 (The PDMP and the contour process). *Processes $(\mathcal{C}_t(\mathbb{T}^{(T)}))_{t \geq 0}$ and $(X_t^{(T)})_{t \geq 0}$ have the same law.*

In addition, we can also obtain the extended generator of the process.

Theorem 3.9 (Extended generator: case $T < +\infty$). *Under Assumptions 2.1, if $T < +\infty$, then the domain $\widehat{D}(L^{(T)})$ of the extended generator of $(X_t^{(T)})_{t \geq 0}$ on $\mathcal{C}([0, T])$ consists of the absolutely continuous functions f . Moreover for such function f , we have*

$$L^{(T)}f(x) = -f'(x) + b(x) \int_{\mathbb{R}_+} (f(x+y) - f(x)) K_T(x, dy),$$

where f' has to be understood in the sense of Radon-Nikodym.

Proof. This is a direct consequence of [10, Theorem (26.14) p.69]. \square

Theorem 3.10 (Extended generator: case $T = +\infty$). *Under Assumptions 2.1, if $T = +\infty$. Moreover, if we assume that the process is non-explosive. Then the domain $\widehat{D}(L)$ of the extended generator of $(X_t)_{t \geq 0}$ on $\mathcal{C}([0, +\infty))$ consists of the absolutely continuous functions f such that*

$$\forall n \geq 1, \quad \mathbb{E} \left[\left| \sum_{k \geq 1, T_k \leq \tau_n} f(X_{T_k}) - f(X_{T_k-}) \right| \right] < \infty,$$

for at least one sequence of stopping time τ_n converging to infinity. Moreover for such function, we have

$$Lf(x) = -f'(x) + b(x) \int_{\mathbb{R}_+} (f(x+y) - f(x)) K(x, dy).$$

Proof. This is a direct consequence of [10, Theorem (26.14) p.69]. \square

4 Scale-type functions and hitting times

4.1 Definition of the scale function

Here, we will define and study a *scale* type function for the contour process. Before introducing it, let us recall briefly the definition of the scale function in the context of diffusion processes and Lévy processes. The interested reader may also take a look at [38] for the diffusion case and [20] for the Lévy case. For a diffusion processes $(Z_t)_{t \geq 0}$ on \mathbb{R} , the (or actually a) scale function $s : \mathbb{R} \rightarrow \mathbb{R}$ verifies, for every $a < b$,

$$\forall x \in [a, b], \quad \mathbb{P}_x(\tau_a > \tau_b) = \mathbb{P}_x(\tau_{(-\infty, a]} > \tau_{[b, +\infty)}) = \frac{s(x) - s(a)}{s(b) - s(a)}, \quad (4.1)$$

where τ_A is the hitting time of the set or the point A for the process $(Z_t)_{t \geq 0}$ (see for instance [38, Chapter VII, Definition (3.3)]). This function is known to be continuous, strictly increasing, unique up to an affine function (see [38, Chapter VII, Proposition (3.2)]). Moreover, from [22, Equation (5.42) 339], we have

$$s : x \mapsto \int_c^x \exp \left(-2 \int_c^y \frac{b(z)}{\sigma^2(z)} dz \right) dy,$$

for any constant $c \in \mathbb{R}$, where b, σ are respectively the drift and the diffusion coefficients of the diffusion process. Moreover, in general, $(s(Z_{t \wedge T_{[a, b]^c}}))_{t \geq 0}$ is a martingale and $\mathcal{L}s = 0$, where \mathcal{L} is the generator of $(Z_t)_{t \geq 0}$ (see [22, Equation (5.43)]). For a Lévy process $(L_t)_{t \geq 0}$, the definition of the scale function, generally denoted by W , is closely related to the previous one but slightly differ. Indeed, from [5, Theorem 8 p.194], we have, for every $a \leq b$,

$$\forall x \in [a, b], \quad \mathbb{P}_x(\tau_{(-\infty, a]} > \tau_{[b, +\infty)}) = \frac{W(x-a)}{W(b-a)},$$

where τ_A is now the hitting time associated to the Lévy process $(L_t)_{t \geq 0}$. In particular, in the case of Lévy process with no positive jumps, the map $x \mapsto W(x)$ is increasing and verifies

$$\int_0^\infty e^{-\lambda x} W(x) dx = \frac{1}{\psi(\lambda)},$$

for large λ and ψ is the Laplace exponent of (L_t) ; see [5, p.188]. Again, $(W(L_t))$ is a martingale and $\mathcal{L}W = 0$, where \mathcal{L} is the generator of $(L_t)_{t \geq 0}$ (this is a rewriting of the last expression, up to a Laplace transform). The map W is fundamental in the study of homogeneous splitting tree and related models [6, 7, 19, 26, 39, 40]. In our case, it appears for instance in the parameter of the geometric distribution associated to the number of individuals (see [26]).

Both these functions are particular cases of harmonic functions of Markov processes (see [24, Chapter 6]) that are solutions to the Dirichlet problem. More precisely, for an open set U and a function $h \in C(\partial U)$, the associated Dirichlet problem for a Markov process $(X_t)_{t \geq 0}$, with generator L corresponds to the existence of a function u of

$$\begin{cases} Lu(x) = 0, & \forall x \in U \\ u(x) = h(x), & \forall x \in \partial U. \end{cases}$$

Under irreducibility assumption $x \mapsto u(x) = \mathbb{E}_x[h(X_\tau)]$ is a solution of this equation (see [24, Theorem 6.2.3]), where τ equals $\tau_{\partial U}$, the hitting time of the boundary of U . In particular, regularity of u is unknown and the expression Lu is taken in a weak sense. For instance, it is not clear that for a diffusion processes that $Lu = au'' + bu'$. Moreover, scale functions do not depend on the boundary a, b of the interval. Indeed, here u depends on U and h , and S, W should depends on $U = [a, b]$ and $h = \mathbf{1}_{[b, +\infty)}$.

Before introducing our definition of a scale function, let us point out some differences between diffusion processes, (spectrally negative) Lévy processes and our PDMP. Firstly, in contrast with both the others processes, the diffusion process has infinite quadratic variation and then $\lim_{x \rightarrow a} \mathbb{P}_x(\tau_a > \tau_b) = 0$ and $\lim_{x \rightarrow b} \mathbb{P}_x(\tau_a > \tau_b) = 1$. Also, the Lévy and the diffusion processes are stochastically monotonous (see [24, Section 5.9]), although our PDMP is not in general. These properties entail, for these two processes, the monotonicity of $x \mapsto \mathbb{P}_x(\tau_{(-\infty, a]} > \tau_{[b, +\infty)})$ and the continuity property of the associated scale function.

We can now give our definition of the scale function for the process $(X_t)_{t \geq 0}$.

Theorem 4.1 (Definition of the scale function). *Under Assumptions 2.1, for every $T \in \mathbb{R}_+$, there exists a bounded function S_T on \mathbb{R}_+ such that $S_T(t) = 0$, for all $t \geq T$, S_T is absolutely continuous on $[0, T]^1$ and*

$$\forall t \in [0, T), \quad S_T(t) = e^{-\int_0^t b(u) du} + e^{-\int_0^t b(u) du} \int_0^t b(s) e^{\int_0^s b(u) du} \int_{[0, T-s)} S_T(s+v) K(s, dv) ds.$$

In particular, S_T belongs to the domain of the extended generator of $(X_{t \wedge \tau_{(-\infty, a]} \wedge \tau_{[b, +\infty)}})_{t \geq 0}$ and $LS_T(t) = 0$ for almost all $t \in [0, T]$. In addition, for all $s \leq t \leq T$, $(S_T(X_{t \wedge \tau_{(-\infty, s]} \wedge \tau_{[T, +\infty)}}))$ is a

¹Usually absolutely continuity is defined on compact intervals, here it means that S_T is absolutely continuous over all compact subintervals of $[0, T]$.

Martingale and

$$\mathbb{P}_t(\tau_{[0,s]} > \tau_{[T,+\infty)}) = \frac{S_T(t) - S_T(s)}{S_T(T) - S_T(s)} = \frac{S_T(s) - S_T(t)}{S_T(s)}. \quad (4.2)$$

Proof. Set $\mathcal{C} = \{H \in C[0, T] \mid H(0) = 1, \|H\|_\infty \leq 1, H \geq 0\}$. This is a non-empty convex closed bounded subset in the uniformly convex Banach space of continuous functions $C([0, +\infty))$. Let $\mathcal{A} : \mathcal{C} \rightarrow \mathcal{C}([0, +\infty))$, defined, for every $x \in [0, b]$ and $S \in \mathcal{C}$ by

$$\mathcal{A}S(t) = e^{-\int_0^t b(u)du} + e^{-\int_0^t b(u)du} \int_0^t b(s) e^{\int_0^s b(u)du} \int_{[0, T-s)} S_T(s+v) K(s, dv) ds$$

We have $\mathcal{A}S(0) = 1$ and $\|\mathcal{A}S\|_\infty \leq 1$, thus $\mathcal{A}S \in \mathcal{C}$. The map \mathcal{A} is also non expansive, namely for all $S_1, S_2 \in \mathcal{C}$,

$$\|\mathcal{A}S_1 - \mathcal{A}S_2\|_\infty \leq \|S_1 - S_2\|_\infty.$$

Then, by the Browder fixed point theorem \mathcal{A} admits a fixed point S_\star in \mathcal{C} . Set $S_T(t) = S_\star(t) \mathbf{1}_{t \in [0, T]}$. It is an absolutely continuous function on $[0, T)$ and then belongs to the extended generator of $(X_{t \wedge \tau_{(-\infty, s]} \wedge \tau_{[T, +\infty)}})$; see [10, Theorem (26.14) p.69] or Theorem 3.9. In particular, by the Itô-Dynkin formula, see [10, Section 26, page 66], the process $(S_T(X_{t \wedge \tau_{(-\infty, s]} \wedge \tau_{[T, +\infty)}}))$ is a martingale. Since hitting times are finite and this martingale is bounded, then using the stopping time theorem and the dominated convergence theorem, we find the expected formula (4.2). \square

For Lévy processes, we can set $W_T : t \mapsto S_T(T-t) \mathbf{1}_{t < T}$ and using the invariance by translation property of these processes, we recover that W_T does actually not depend on T .

We also naturally recover the natural property that S_T is constant (namely $\mathbb{P}_t(\tau_{[0,s]} > \tau_{[T, +\infty)}) = 0$) on $[0, T)$ if and only if $K(t, [0, T)) = 1$ for all $t \in [0, T]$.

Corollary 4.2 (Integro-differential equation for S_T). *If K is Feller, b and $t \mapsto K(t, [0, T-t))$ are continuous then $S_T \in C^1([0, T) \cup (T, +\infty))$ and*

$$\forall t \in [0, T), \quad -S'_T(t) + b(t) \left(\int_{[0, T-t)} S_T(t+s) K(t, ds) - S_T(t) \right) = 0. \quad (4.3)$$

Note that if $K(t, ds) = \varphi(t, s) ds$, with a continuous density φ , then K is Feller and $t \mapsto K(t, [0, T))$ is continuous.

Comparing with the application for the tree, we recover the formula established in [Lambert, Stadler 2013, Proposition 4] and [27, Proposition 3.2.6]. We even improve this result, because we establish the regularity of the function S_T and precise the sense of the derivative.

Corollary 4.3 (Monotony properties of the scale function). *For every $T \geq t \geq 0$, we have*

$$S_T(t) = \mathbb{P}_t(\tau_0 < \tau_{[T, +\infty)}).$$

In particular $t \mapsto S_T(t)$ is decreasing and $T \mapsto S_T(t)$ is increasing.

4.2 Tree properties through scale function

We can now express certain properties of the tree through this new defined scale function. We set Ξ_t the number of individuals at time t .

Proposition 4.4 (Number of individuals at a fixed time). *For every $t_0 \leq t$, we have*

$$\mathbb{P}_{t_0}(\Xi_t = 0) = \mathbb{P}_{t_0}(\mathcal{H}(\mathbb{T}) \leq t) = S_t(t_0). \quad (4.4)$$

Moreover, conditionally on $\{\Xi_t \neq 0\}$, Ξ_t is geometrically distributed; that is

$$\mathbb{P}_{t_0}(\Xi_t = k \mid \Xi_t \neq 0) = S_t(t-)(1 - S_t(t-))^k \quad (4.5)$$

This result is a direct adaptation of [26, Proposition 5.6]. The proof is written for sake of completeness.

Proof. The height of the tree is lower than t if and only if $(X_t)_t$ hits 0 before (t, ∞) . Theorem 4.1 then entails Equation 4.4. Now, as Ξ_t correspond of the number of times that $(X_t)_{t \geq 0}$ hits t . By the Markov property, it is then a sequence of i.i.d. excursions of $(X_t)_{t \geq 0}$ from t on $(0, t]$, stopped at the first one that exits it from the bottom. Hence, from Theorem 4.1, we obtain (4.5). \square

Corollary 4.5 (Extinction probability). *For every $t_0 \geq 0$, we have*

$$\mathbb{P}_{t_0}(\text{Ext}) = \mathbb{P}_{t_0}(\exists t \geq 0, \Xi_t = 0) = \mathbb{P}_{t_0}(\mathcal{L}(\mathbb{T}) < \infty) = \lim_{t \rightarrow \infty} S_t(t_0).$$

In particular, $S : t_0 \mapsto \lim_{t \rightarrow \infty} S_t(t_0)$ exists, is decreasing and it is solution to the functional equation

$$\forall t \geq 0, S(t) = e^{-\int_0^t b(u)du} + e^{-\int_0^t b(u)du} \int_0^t b(s) e^{\int_0^s b(u)du} \int_{[0, +\infty)} S(s+v) K(s, dv) ds,$$

Proof. Function S is well defined and decreasing thanks to Corollary 4.3 and the functional equation holds because of the Beppo Levi Theorem. \square

Again under the regularity assumptions of Corollary 4.2, S is solution to an integro-differential equation similar to (4.3).

In the extinction case, as in the homogeneous setting ([26, Proposition 5.8], it is possible to establish Yaglom-type limiting result; that means convergence to a quasi-stationary type distribution.

Corollary 4.6 (Quasi-limiting behavior of (Ξ_t)). *We have*

- *If $\lim_{t \rightarrow \infty} S_t(t-) = q \in (0, 1)$ then*

$$\lim_{t \rightarrow \infty} \mathbb{P}_{t_0}(\Xi_t = k \mid \Xi_t \neq 0) = q(1 - q)^k \quad (4.6)$$

- *If $\lim_{t \rightarrow \infty} S_t(t-) = 0$ then*

$$\forall x > 0, \quad \lim_{t \rightarrow \infty} \mathbb{P}_{t_0}(\Xi_t S_t(t-) > x \mid \Xi_t \neq 0) = e^{-x}. \quad (4.7)$$

In case of Lévy process, it is direct that $t \mapsto S_t(t-)$ is decreasing and then converges although it is not generally the case. Thus, the previous corollary does not give an exhaustive description of possible behaviors.

However thanks to Corollary 4.3, we have, for all $s \geq 0$,

$$\lim_{t \rightarrow \infty} S_t(t-) \leq \lim_{t \rightarrow \infty} S_t(s) = S(s) = \mathbb{P}_s(\text{Ext}),$$

and then $\lim_{t \rightarrow \infty} S_t(t-) \leq \lim_{t \rightarrow \infty} S(t)$.

Finally Corollary 4.6 (Equation (4.7)) gives also the deterministic growth under non-extinction. Indeed, as

$$\mathbb{P}_{t_0}(\text{Ext}) = \lim_{t \rightarrow \infty} \mathbb{P}_{t_0}(\Xi_t = 0) = \lim_{t \rightarrow \infty} S_t(t_0),$$

we have

$$\lim_{t \rightarrow \infty} \mathbb{P}_{t_0}(\Xi_t S_t(t-) > x \mid \Xi_t \neq 0) = \lim_{t \rightarrow \infty} \mathbb{P}_{t_0}(\Xi_t S_t(t-) > x \mid \text{Ext}^c).$$

Hereafter, we will say that the tree is supercritical when $\mathbb{P}_x(\mathcal{L}(\mathbb{T}) = +\infty) > 0$. Several details are given in the next section. In the supercritical case, one can go further that this result by looking the tree conditioning on being finite or infinite. Indeed, From Corollary 4.5 (see also the proof of Proposition 4.7 below), S is an harmonic function, which is non-trivial, under the supercritical assumption, and one can then use a Doob h -transform (a Cameron-Martin-Girsanov type change of measure, see [12] or [41, Page 83]) to define a new (Markov) process which is distributed as the process conditioned on hitting 0 (or never hit it).

Before going further, let us introduce some notations. For any positive measurable function f and g , we set

$$Kf(x) = \int_{\mathbb{R}_+} f(x+y) K(x, dy), \quad K^g(f) := K(gf)/K(g).$$

We can now state our result on conditioned ISTs.

Proposition 4.7 (Conditioned tree). *Assume that \mathbb{T} is supercritical. We have*

- *The law of \mathbb{T} conditionally on $\text{Ext} = \{\mathcal{L}(\mathbb{T}) < +\infty\}$ is the same as an IST with birth rate bKS/S and death kernel K^S .*
- *The law of \mathbb{T} conditionally on $\text{Ext}^c = \{\mathcal{L}(\mathbb{T}) = +\infty\}$ is the same as a IST with birth rate $b(1-KS)/(1-S)$ and death kernel $K^{(1-S)}$.*

Proof. Let $(P_t)_{t \geq 0}$ be the Markov semi-group of $(X_t)_{t \geq 0}$. Using the Markov property, we easily see that for every $t \geq 0$ that

$$P_t S(t) = \mathbb{E}_t[\mathbb{P}_{X_t}(\exists s \geq 0, X_s = 0)] = \mathbb{P}_t(\exists s \geq 0, X_s = 0) = S(t).$$

This shows that the function S is an harmonic function. Let $(P_t^S)_{t \geq 0}$ defined, for every continuous and bounded functions $f : \mathbb{R}_+^* \rightarrow \mathbb{R}$ vanishing at 0 (and that we extend in 0 by $f(0) = 0$), by

$$P_t^S f = \frac{P_t(Sf)}{P_t S} = \frac{P_t(Sf)}{S}$$

is a Markov semi-group on $\mathbb{R}_+^* = (0, +\infty)$. We have, for all f such that $f/S \in D(L)$, $\partial_t P_t^S f = P_t^S \mathcal{L}^S f$ where,

$$\mathcal{L}^S f(t) = \frac{\mathcal{L}(fS)(t)}{S(t)} = -f'(t) + \frac{b(t)KS(t)}{S(t)} \left(\frac{K(Sf)(t)}{KS(t)} - f(t) \right).$$

This semi-group then correspond to the semi-group of the PDMP which decreases linearly (with rate 1) and jumps at rate bKS/S with kernel K^S .

Using again the Markov property, we have, for every continuous and bounded functions $f : \mathbb{R}_+^* \rightarrow \mathbb{R}$ and $t > 0$,

$$P_t^S f(t) = \frac{\mathbb{E}_t [f(X_t) \mathbb{P}_{X_t}(\exists s \geq 0, X_s = 0)]}{\mathbb{P}_t(\exists s \geq 0, X_s = 0)} = \frac{\mathbb{E}_t [f(X_t) \mathbf{1}_{\{\exists s \geq t, X_s = 0\}}]}{\mathbb{P}_t(\exists s \geq 0, X_s = 0)} = \mathbb{E}_t [f(X_t) \mid \exists s \geq 0, X_s = 0].$$

In other words, the law of X_t conditioned to the extinction is described by P_t . By successive conditioning, it gives that the law of $(X_t)_{t \geq 0}$ conditionally on extinction is the same as the Markov process with semi-group $(P_t)_{t \geq 0}$.

The proof is similar when conditioning on non-extinction with $h = 1 - S$ as harmonic function. We then set (P_t^h) and \mathcal{L}^h the operators $(P_t(\cdot)/h)_{t \geq 0}$ and $\mathcal{L}(\cdot)/h$. The only difference is that h is not bounded around 0 and we need to prove that the associated Markov process does not explode. But as $\mathcal{L}^h(1/h) = 0$, [36, Theorem 2.1] ends the proof. \square

Remark 4.8. Using S_T as harmonic function permits to have the law of tree conditioned on $\mathcal{H}(\mathbb{T}) < T$ or $\mathcal{H}(\mathbb{T}) \geq T$.

4.3 Applications to particular cases

4.3.1 Deterministic life duration

Here we assume that $K(t, du) = \delta_1$, meaning that individuals live a fixed time equals to 1. The map $t \mapsto K(t, [0, T-t]) = \mathbf{1}_{\{t+1 \leq T\}}$ is not continuous. We have

$$\begin{aligned} S_T(t) &= e^{-\int_0^t R(u)du} + e^{-\int_0^t R(u)du} \int_0^t R(s) e^{\int_0^s R(u)du} S_T(s+1) \mathbf{1}_{s+1 < T} ds \\ &= e^{-\int_0^t R(u)du} + e^{-\int_0^t R(u)du} \int_0^{t \wedge T-1} R(s) e^{\int_0^s R(u)du} S_T(s+1) ds. \end{aligned}$$

In particular, for $t \in [T-1, T)$, $S_T(t) = e^{-\int_0^t R(u)du} C$, for some $C > 0$ and then we recover the property $\mathbb{P}_t(T_{[0, T-1]} < T_{[T, +\infty)}) = e^{-\int_{(T-1)}^t R(u)du} = S_T(t)/S_T(T-1)$ that can be obtain by regarding if the process has a jump before hitting $T-1$.

4.3.2 Fixed death moment

Let us consider fixed death times; for instance at times 1 or 2. For $x \in [0, 1]$,

$$K(t, du) = \frac{1}{2} \delta_{1-t}(du) + \frac{1}{2} \delta_{2-t}(du)$$

and for $t \in (1, 2]$, $K(t, du) = \delta_{2-t}(du)$. Then for $t \leq 1$,

$$\begin{aligned} S_T(x) &= e^{-\int_0^t R(u)du} + e^{-\int_0^t R(u)du} \int_0^t R(s) e^{-\int_0^s R(u)du} \frac{1}{2} (S(1)\mathbf{1}_{T \geq 1} + S(2)\mathbf{1}_{T \geq 2}) ds \\ &= e^{-\int_0^t R(u)du} + \left(1 - e^{-\int_0^t R(u)du}\right) C_T, \end{aligned}$$

for some $C_T > 0$. For $t \in [1, 2)$

$$\begin{aligned} S_T(x) &= e^{-\int_0^t R(u)du} + e^{-\int_0^t R(u)du} \left(\int_0^1 R(s) e^{-\int_0^s R(u)du} C_T ds + \int_1^t R(s) e^{-\int_0^s R(u)du} S(2)\mathbf{1}_{T \geq 2} ds \right) \\ &= e^{-\int_0^t R(u)du} + \left(1 - e^{-\int_0^1 R(u)du}\right) C_T + \left(e^{-\int_0^1 R(u)du} - e^{-\int_0^t R(u)du}\right) c_T, \end{aligned}$$

for some $c_T > 0$. If $S(2) = 0$ then $c_T = 0$, $C_T = 2e^{-\int_0^1 R(u)du} / \left(1 + e^{-\int_0^1 R(u)du}\right)$, and for all $t \in [0, 2)$,

$$S_T(x) = e^{-\int_0^t R(u)du} + 2e^{-\int_0^1 R(u)du} \frac{1 - e^{-\int_0^{t \wedge 1} R(u)du}}{1 + e^{-\int_0^1 R(u)du}}.$$

In particular S is not differentiable at $t = 1$.

4.3.3 Inhomogeneous-time Markovian tree

Let $K(t, du) = d(t+u)e^{-\int_t^{t+u} d(v)dv} du$ and the equation of Corollary 4.2 reads

$$\begin{aligned} S'_T(x) &= b(t) \left(\int_0^{T-t} S_T(t+s) d(t+s) e^{-\int_t^{t+s} d(u)du} ds - S_T(t) \right) \\ &= b(t) \left(\int_t^T S_T(s) d(s) e^{-\int_t^s d(u)du} ds - S_T(t) \right). \end{aligned} \tag{4.8}$$

In particular, if b, d are C^1 then

$$\begin{aligned} S''_T(t) &= b'(t) \left(\int_t^T S_T(s) d(s) e^{-\int_t^s d(u)du} ds - S_T(t) \right) \\ &\quad - b(t) S_T(t) d(t) + b(t) \int_t^T S_T(s) d(s) d(t) e^{-\int_t^s d(u)du} ds - b(t) S'_T(t). \end{aligned}$$

Using Equation (4.8), this can be simplified to $S''_T(t) = \frac{b'(t)}{b(t)} S'_T(t) + d(t) S'_T(t) - b(t) S'_T(t)$ with also the following boundary condition: $S_T(0) = 1$, $S'_T(T) = -b(T) S(T)$. This simple equation can be solved. Indeed

$$S'_T(t) = -b(T) S_T(T) e^{-\int_t^T \frac{b'(u)}{b(u)} + d(u) - b(u) du},$$

and

$$S_T(t) = 1 - b(T) S(T) \int_0^t e^{-\int_s^T \frac{b'(u)}{b(u)} + d(u) - b(u) du} ds.$$

Using $t = T$, we finally have

$$\begin{aligned} S_T(t) &= 1 - \frac{b(T) \int_0^t e^{-\int_s^T \frac{b'(u)}{b(u)} + d(u) - b(u) du} ds}{1 + b(T) \int_0^T e^{-\int_s^T \frac{b'(u)}{b(u)} + d(u) - b(u) du} ds} = \frac{1 + b(T) \int_t^T e^{-\int_s^T \frac{b'(u)}{b(u)} + d(u) - b(u) du} ds}{1 + b(T) \int_0^T e^{-\int_s^T \frac{b'(u)}{b(u)} + d(u) - b(u) du} ds} \\ &= \frac{1 + \int_t^T b(s) e^{-\int_s^T d(u) - b(u) du} ds}{1 + \int_0^T b(s) e^{-\int_s^T d(u) - b(u) du} ds}. \end{aligned}$$

In particular, we recover [27, Exercise 3.2.8]. When $b \neq d$ are constant, we have

$$S_T(t) = \frac{d - be^{(b-d)(T-t)}}{d - be^{(b-d)T}},$$

so we recover the classical function [7],

$$W(t) = \frac{d - be^{(b-d)t}}{d - b}.$$

4.3.4 Homogeneous tree

Let us consider the case of a standard homogeneous splitting tree. In such situation, b is constant and the kernel $K(dy)$ is not time-dependent. Moreover, we have $S(t) = e^{-\alpha t}$ for non-negative constant α .

In such situation the semi-group of the process conditioned to Ext provided by Proposition 4.7 is given by

$$L^S f(x) = -f'(y) + b \int_{\mathbb{R}_+} e^{-\alpha y} K(dy) \left(\frac{\int_{\mathbb{R}_+} e^{-\alpha y} f(x+y) K(dy)}{\int e^{-\alpha y} K(dy)} - f(x) \right).$$

In this case, we recover, the results of [26].

A more interesting situation is the case of splitting tree conditioned to non-extinction. In this case we get

$$L^h f(x) = -f'(y) + b \frac{\int_{\mathbb{R}_+} (1 - e^{-\eta(x+y)}) K(dy)}{(1 - e^{-\eta x})} \left(\frac{\int_{\mathbb{R}_+} (1 - e^{-\eta(x+y)}) f(x+y) K(dy)}{\int (1 - e^{-\eta(x+y)}) K(dy)} - f(x) \right),$$

which shows that the conditioned tree is inhomogeneous, but tends to recover its homogeneity in the long time limit.

5 Lyapounov functions: sufficient condition and tail estimates

5.1 General condition

The aim of this section is to generalise [26, Proposition 2.2 p. 12] which links the probability to extinction (finiteness of the tree) to the drift of the contour process. However, in contrast with the case where the contour is a Lévy process (see for instance [5, Corollary 2 p. 190]), there is no simple formulation on the drift describing the long time behaviour of a general PDMP. Still, we have

Theorem 5.1 (Extinction and explosion: the merciless dichotomy). *Let τ_0 be the hitting time of 0. Then, almost surely, either τ_0 is finite or $\lim_{t \rightarrow \infty} X_t = +\infty$. Equivalently either the height $\mathcal{H}(\mathbb{T})$ of the tree is infinite (and then the length $\mathcal{L}(\mathbb{T})$ is also infinite) or $\mathcal{L}(\mathbb{T})$ is finite.*

Proof. From Theorem 3.6, it is enough to prove that either the hitting time τ_0 of 0 is finite or $\lim_{t \rightarrow \infty} X_t = +\infty$. But as the process $(X_t)_{t \geq 0}$ is a δ_0 -irreducible T -process, this is a direct consequence of [35, Theorem 3.2] (δ_0 -Harris recurrence is equivalent to absorption in finite time). \square

Remark 5.2. *Remark that Hypothesis 2.1 implies the local boundedness of b . This implies that there can only be a finite number of individuals in the truncated tree $\mathbb{T}^{(T)}$, which implies $\mathcal{L}(\mathbb{T}^{(T)}) < \infty$. In particular, $\mathcal{L}(\mathbb{T}) < \infty$ if and only if $\mathcal{H}(\mathbb{T}) < \infty$. Let us insist on the fact that this relies on the local boundedness of b .*

From the Theorem 5.1 and from the classical definition for Galton-Watson processes or homogeneous splitting tree, we can define the typical behaviours of the population.

1. The supercritical case: if one of the following equivalent assertions holds

- For all $x > 0$, $\mathbb{P}_x(\lim_{t \rightarrow \infty} X_t = +\infty) = 1 - \mathbb{P}_x(\tau_0 < +\infty) > 0$,
- $\mathbb{P}_x(\mathcal{L}(\mathbb{T}) = +\infty) > 0$,
- $\mathbb{P}_x(\mathcal{H}(\mathbb{T}) = +\infty) > 0$.

2. The critical case: if one of the following equivalent assertions holds

- $\tau_0 < +\infty$ almost surely but $\mathbb{E}_x[\tau_0] = +\infty$, for all $x > 0$.
- $\mathbb{P}_x(\mathcal{L}(\mathbb{T}) = +\infty) > 0$ but $\mathbb{E}_x[\mathcal{L}(\mathbb{T})] < +\infty$, for all $x > 0$.

3. The subcritical case: if one of the following equivalent assertions holds

- $\tau_0 < +\infty$ almost surely and there exists $\theta > 0$ $\mathbb{E}_x[e^{\theta\tau_0}] < +\infty$, for all $x > 0$.
- $\mathbb{E}_x[\mathcal{L}(\mathbb{T})] < +\infty$, for all $x > 0$.

This division also comes from the different recurrence notions for the contour. Note that the situation is not as homogeneous-time Galton-Watson for whose we have

$$\mathbb{E}[\tau_0] < \infty \Rightarrow \exists \theta > 0, \mathbb{E}_t[e^{\theta\tau_0}] < +\infty.$$

Let us also highlight that, as in the Galton-Watson case, a non-supercritical population almost surely extinct.

As pointed out in the beginning of this section, it is not easy to have simple condition to ensure in which case we are. Let us end this section by giving sufficient condition using drift-type conditions. These conditions are based on classical Lyapunov functions. Even they may not be optimal, their proof gives a method to show how exploit the contour to answer some questions on the size of the tree.

First of all, let us set, for all $x \geq 0$,

$$\begin{cases} m(x) = \int_{\mathbb{R}_+} y K(x, dy), \\ m_p(x) = \int_{\mathbb{R}_+} |y|^p K(x, dy), \quad p \geq 1. \end{cases} \quad (5.1)$$

We can, now, introduce the following necessary condition for supercriticality.

Assumption 5.3 (No age barrier). $\mathbb{P}_x(\exists t > 0, X_t = y) > 0, \quad \forall x, y > 0$.

Remark 5.4. *This last assumption can fail in some pathological situations. For instance, if new individual lifetime became shorter and shorter or if b is null on large time interval (or a combination of these two phenomena). It is difficult to write a necessary and sufficient condition on b and K for this assumption. A sufficient one is*

$$\begin{cases} b(x) > 0, & \forall x \geq 0, \\ K(x, [y, \infty)) > 0, & \forall x, y \geq 0. \end{cases}$$

We can now state the drift condition.

Proposition 5.5 ((Continuous) Drift conditions for extinction/survival).

1. *If there exists a positive $V \in \widehat{D}(L)$ such that $LV \leq 0$ outside a compact set and*

$$V(x) \xrightarrow{x \rightarrow \infty} +\infty,$$

then $\tau_0 < +\infty$ almost surely, which means that \mathbb{T} is not supercritical.

2. *If Assumption 5.3 holds and there exists a positive $V \in \widehat{D}(L)$ and a compact $K \supset \{0\}$ such that*

$$\begin{cases} LV(x) \leq 0, & \forall x \in \mathbb{R}_+ \setminus K, \\ V(x) < \inf_{y \in K} V(y), & \forall x \in K, \end{cases}$$

then \mathbb{T} is supercritical.

Proof. For the first point set V satisfies (CD1) of [36] and then by the result of [36, Section 3], the process does not drift to infinity, and then goes to 0 according to Theorem 5.1. For the second point, it is classic that if V is a map such that $LV \leq 0$ then

$$\mathbb{P}_x(\tau_0 < +\infty) \leq \frac{V(x)}{V(0)}, \quad \forall x \in \mathbb{R}_+. \quad (5.2)$$

Indeed, as $(V(X_{t \wedge \tau_0}))_{t \geq 0}$ is a super-martingale (from Theorem 3.9), it is a consequence of the stopping time theorem and Fatou Lemma. Then for large x , $\mathbb{P}_x(\tau_0 < +\infty) < 1$. Using Assumption 5.3 ends the proof. \square

Corollary 5.6 (Sufficient asymptotic conditions for extinction/survival).

1. *If $\limsup_{x \rightarrow \infty} b(x)m(x) < 1$, then \mathbb{T} is not supercritical.*

2. If $\liminf_{x \rightarrow \infty} b(x)m(x) > 1$ and $\sup_{x \geq 1} \frac{b(x)m_2(x)}{x} < +\infty$, then \mathbb{T} is supercritical.

Proof. For the first point use $V : x \mapsto x$ in Proposition 5.5. For the second point, set $V : x \mapsto Mx^{-\alpha}\mathbf{1}_{x \geq M} + \mathbf{1}_{x < M}$, for some $M, \alpha > 0$ fixed hereafter. A rapid calculation using Jensen inequality shows that, for $x > M$,

$$LV(x) \leq \frac{\alpha}{x^{\alpha+1}} \left[(1 - b(x)m(x)) + \frac{1}{\alpha + 1} b(x) \frac{1}{x} m_2(x) \right].$$

Then for M and α large enough, we have $LV \leq 0$. \square

Remark 5.7 (Assumption). *The assumption $\sup_{x \geq 1} \frac{b(x)m_2(x)}{x} < +\infty$ is a technical but no confining assumption. Moreover, it is not clear if it is only a technical one (that can be removed) or if it has a real sense. Indeed, let us recall that fluctuation of the environment has an effect of the probability of extinction in classical Galton-Watson chain (see [17, Section 2.9.2 p.49] for instance).*

Anyway, it can be weaken in such a way: if V is decreasing with increasing derivative function V' then it is enough that $\sup_{x \in K^c} b(x)m_2(x)V''(x) < +\infty$ for some compact set K .

Proposition 5.8 ((Discrete) Drift conditions for extinction/survival). *Let P be the transition kernel defined, for every positive function f , by $Pf(0) = f(0)$ and for all $x > 0$ by*

$$Pf(x) = \int_0^x \int_{\mathbb{R}_+} f(y+z) K(y, dz) b(y) e^{-\int_y^x b(s) ds} dy + f(0) e^{-\int_0^x b(s) ds}.$$

1. *If there exists a positive function V such that $PV - V \leq 0$ outside a compact set and*

$$V(x) \xrightarrow{x \rightarrow \infty} +\infty,$$

then \mathbb{T} is not supercritical.

2. *If Assumption 5.3 holds and there exists a non-constant decreasing and positive V such that $PV - V \leq 0$ outside a compact set K including 0 then the tree is supercritical.*

Proof. Let us consider $(\hat{X}_n)_{n \geq 0}$ be the (post-jump) embedded (or skeleton) chain associated to $(X_t)_{t \geq 0}$; namely it is the Markov chain defined by $\hat{X}_n = X_{T_n}$, where T_n is the n th jump time of $(X_t)_{t \geq 0}$. Its transition kernel is P and similarly to Proposition 5.5, the statement is a consequence of the classical result [34, Theorem 8.4.3 p. 191] and [34, Proposition 8.4.1 p.189] (in discrete times now). \square

Corollary 5.9 (Sufficient integral condition for extinction). *If*

$$\limsup_{x \rightarrow \infty} \int_0^x \left((m(s)b(s) - 1) e^{-\int_s^x b(u) du} \right) ds < 0,$$

then \mathbb{T} is not supercritical.

Proof. Again function $V : x \mapsto x$ satisfies

$$Pf(x) = \int_0^x (s + m(s)) b(s) e^{-\int_s^x b(u) du} ds.$$

An integration by parts ends the proof. \square

In case of finite tree, it is possible to give some bound on the tail of the length of the tree $\mathcal{L}(\mathbb{T})$ using again drift conditions.

Proposition 5.10 (Tail estimate of the tree length). *Assume $\limsup_{x \rightarrow \infty} b(x)m(x) < 1$.*

1. *If there exists $a > 0$, such that for any large enough x , $\int_{\mathbb{R}_+} e^{ay} K(x, dy) < +\infty$ then*

$$\mathbb{P}_x(\mathcal{L}(\mathbb{T}) \geq t) \leq C(1 + e^{\theta x})e^{-\lambda t}, \quad \forall t \geq 0, \quad (5.3)$$

and some constants $C, \lambda, \theta > 0$.

2. *If there exists $p \in \mathbb{N}, p \geq 2$, such that for any x large enough, $m_p(x) < +\infty$ and*

$$\lim_{x \rightarrow \infty} x^{k-p+1} m_{p-k}(x) = 0,$$

then

$$\mathbb{P}_x(\mathcal{L}(\mathbb{T}) \geq t) \leq C(t^{-1/p} x^p + e^{\theta x} t^{-1/p+1/p^2}), \quad \forall t \geq 0, \quad (5.4)$$

and some constant $C > 0$.

Proof. First, from Theorem 3.6, we have

$$\mathbb{P}(\mathcal{L}(\mathbb{T}) \geq 2t) = \mathbb{P}(\tau_0 > t) = \|\mathcal{L}(X_t) - \delta_0\|_{\text{TV}},$$

where $\|\cdot\|_{\text{TV}}$ is the classical total variation distance. Hence, it is enough to work with classical results on the convergence of Markov processes to equilibrium. In particular (5.3) is a consequence of [36, Theorem 6.1] and (5.4) is a consequence of [13, Theorem 3.10, (3.6)]. For the first point, let us check [36, (CD3)] with $V_\theta(x) : t \mapsto e^{\theta x}$. In the first hand, we have

$$LV_\theta(x) = V(x) \left(-\theta + b(x) \int_{\mathbb{R}_+} (e^{\theta y} - 1) K(x, dy) \right) = \lambda_\theta V(x), \quad \forall x \geq 0.$$

On the second hand $\lambda_0 = 0$, and $\partial_\theta \lambda_\theta(x) =_{|\theta=0} b(x)m(x) - 1$ which is negative for large x . Then there exists a small θ such that [36, (CD3)] and [36, Theorem 6.1] hold. For the second point now, the map $V_p : x \mapsto x^p$ satisfies

$$\mathcal{L}V_p(x) = (b(x)m(x) - 1)pV_p^{1-1/p}(x) + \sum_{k=0}^{p-2} \binom{p}{k} x^{k-p+1} m_{p-k}(x)$$

and then the drift condition of [13, Theorem 3.11, (3.10)] (see also [18, Theorem 4.1 point 3]) holds. This ends the proof. \square

5.2 Applications to particular cases

5.2.1 Periodic environment

We will assume that $b = \beta$ is constant and m is periodic (with period T). Set

$$\psi : s \mapsto \beta m(s) - 1, \quad \Psi : t \mapsto e^{-t} \int_0^t \psi(s) e^{\beta s} ds.$$

From Corollary 5.9, the process goes to extinction if $\limsup_{t \rightarrow \infty} \Psi(t) < 0$. Using the periodicity assumption, we have, for every $n \geq 0$,

$$\begin{aligned} \Psi(nT) &= e^{-\beta t - \beta(n-1)T} \Psi(T) + \Psi(t) = e^{-\beta t} \Psi(T) \sum_{k=0}^{n-1} e^{-\beta kT} + \Psi(t) \\ &= \Psi(T) \frac{e^{-\beta t} (1 - e^{-\beta nT})}{(1 - e^{-\beta kT})} + \Psi(t). \end{aligned}$$

In particular, if we set

$$\Phi : t \mapsto \Psi(T) \frac{e^{-\beta t}}{(1 - e^{-\beta kT})} + \Psi(t) = e^{-\beta t} \left(\frac{e^{-\beta T}}{(1 - e^{-\beta kT})} \int_0^T \psi(s) e^{\beta s} ds + \int_0^t \psi(s) e^{\beta s} ds \right),$$

then Φ is periodic and $\lim_{t \rightarrow \infty} |\Phi(t) - \Psi(t)| = 0$. To verify Corollary 5.9 (1), it is then enough to have $\sup_{t \in [0, T]} \Phi(t) < 0$. As an instance, when $\psi : t \mapsto \cos(t) + c$ (c fixed hereafter), we have

$$\forall t \in [0, 2\pi], \quad \Phi(t) = \frac{1}{1 + \beta^2} \left[-\beta e^{-\beta t} (1 + e^{-2\pi\beta}) + \beta \cos(t) + \sin(t) \right]$$

For $\beta = 1$, $\sup_{t \in [0, T]} \Phi(t) \approx 0.5072555 + c$. Then the process goes to extinction for $c < 0.5073$. Note that for such c , we can have $\limsup_{t \rightarrow \infty} \psi(t) = 1 + c > 0$ (this is then a better result than those of Corollary 5.6 (1))

6 Scaling limits of non-homogeneous splitting trees

In this section, we are interested in scaling limits of ISTs. The first subsection is concerned by the overall shape of asymptotically critical IST. In the second part, we consider the asymptotic behavior of ISTs as the birth-rate goes to ∞ whereas the birth-kernel $K(x, dy)$ tends to have zero means. In such situation, we consider two sequences $(b_n)_{n \geq 1}$ and $(K_n)_{n \geq 1}$ converging point-wise receptively to ∞ and δ_0 . We show that one cannot obtain scaling limit for such trees without endowing it in the larger space of TOM trees. This is quite a frustrating result since, intuitively speaking, the limiting tree (from a topological point of view) should not depend on a measured structure.

6.1 Scaling limit of asymptotically critical IST

In this section we look up to the global shape of asymptotically critical IST. More precisely, we suppose that b and K satisfy

$$b(x) \int_{\mathbb{R}_+} y K(x, dy) \xrightarrow{x \rightarrow \infty} 1.$$

To this end, we consider the time contraction \mathbb{T}_n of \mathbb{T} on the scale $(a_n)_{n \geq 1}$ by

$$(\delta, t) \in \mathbb{T}_n \Leftrightarrow (\delta, a_n t) \in \mathbb{T}(x),$$

where $a_n \rightarrow \infty$ as n goes to infinity. Our goal is to show that the sequence a tree converges to some tree in the sense of the Gromov-Hausdorff-Prokhorov topology using the contour processes. Since the triplet $(\mathbb{T}(x), d, \lambda)$ is a TOM tree, we can define a sequence of TOM trees by setting

$$\begin{cases} d_n((\sigma, s), (\delta, t)) = \frac{d((\sigma, c_n s), (\delta, c_n t))}{a_n}, & \forall (\sigma, a_n s), (\delta, a_n t) \in \mathbb{T}(a_n x) \\ \lambda_n(A) = \frac{a_n}{c_n} \lambda(A), \end{cases}$$

From this construction, it is easily seen that the contour process $(\mathcal{C}(\mathbb{T}_n), s \in \mathbb{R}_+)$ of \mathbb{T}_n has generator given by

$$L_n f(x) = -\frac{a_n}{c_n} f'(x) + a_n b(c_n x) \int_{\mathbb{R}_+} f\left(x + \frac{y}{c_n}\right) - f(x) K(c_n x, dy).$$

In particular, if we assume that f has third derivative we have, using Taylor expansion,

$$\begin{aligned} L_n f(x) = & \frac{a_n}{c_n} (b(c_n x) m(c_n x) - 1) f'(x) + \frac{1}{2} f''(x) \frac{a_n b(c_n x) m_2(c_n x)}{c_n^2} \\ & + a_n b(c_n x) \int_{\mathbb{R}_+} \int_x^{x+y/c_n} \frac{f^{(3)}(s)}{2} \left(s - x - \frac{y}{c_n}\right)^2 ds K(c_n x, dy). \end{aligned} \quad (6.1)$$

Hence, to obtain a non-degenerate limit in n , one should require that

$$\begin{cases} \frac{a_n}{c_n} (b(c_n x) m(c_n x) - 1) & \xrightarrow{n \rightarrow \infty} g_d(x), \\ \frac{a_n}{c_n^2} b(c_n x) m_2(c_n x) & \xrightarrow{n \rightarrow \infty} g_v(x), \end{cases}$$

for some functions g_d and g_v . In particular, this implies that the functions $x \mapsto (b(x)m(x) - 1)$ and $x \mapsto b(x)m_2(x)$ are regularly varying. As a consequence, there exists two slowly varying functions S_d and S_v , and two real numbers β and γ , such that

$$\begin{cases} (b(x)m(x) - 1) = x^\beta S_d(x), \\ b(x)m_2(x) = x^\gamma S_v(x). \end{cases}$$

In addition, the only case where the limit can hold is $\beta = \gamma - 1$. This leads to the following assumptions on b and K .

Assumption 6.1. 1. $(b(x)m(x) - 1) = \frac{c}{x} S_d(x)$ for some constant c and such that S_d is bounded and satisfies $S_d(x) \xrightarrow{x \rightarrow \infty} 1$.

2. $b(x)m_2(x) = S_v(x)$ such that S_v is bounded and satisfies $S_v(x) \xrightarrow{x \rightarrow \infty} 1$.

3. $x \mapsto b(x)m_3(x)$ is bounded.

Remark 6.2. *There are many examples of parameters which satisfy these hypotheses. Among the simpler ones, one could, for instance, think of*

$$\begin{cases} b(x) = 1 + \frac{c}{1+x}, \text{ for some constant } c, \\ K(x, dy) = \delta_1(dy). \end{cases}$$

In addition, in order to lighten notations, we also assume that $c_n^2 = a_n = n$. The proof relies on the two following lemmas. The idea is to show the convergence of the sequence $(L_n)_{n \geq 1}$ of generators using a wise choice of core for the expected limit of this sequence. The good choice appears in the following lemma where we show that this set is dense.

Lemma 6.3. *The subspace*

$$\mathcal{D} = \left\{ f \in C_c^\infty(\mathbb{R}_+) \mid f^{(n)}(0) = 0, \quad \forall n \geq 1 \right\},$$

is dense in $C_0(\mathbb{R}_+)$.

Proof. Let f be a function in $C_0(\mathbb{R}_+)$. Since $C_c^\infty(\mathbb{R}_+)$ is dense in $C_0(\mathbb{R}_+)$, there exists some smooth compactly supported function φ such that

$$\|f - \varphi\|_\infty < \frac{\varepsilon}{2},$$

for some fixed positive real number ε . In addition, since φ is compactly supported, it is also uniformly continuous which implies the existence of some positive real number η_ε such that

$$\omega(\varphi, \eta_\varepsilon) < \frac{\varepsilon}{2}, \tag{6.2}$$

where ω is the modulus of continuity defined by

$$\omega(\varphi, \delta) = \sup_{\substack{x, y \in \mathbb{R}_+ \\ \text{s.t. } |x-y| < \delta}} |\varphi(x) - \varphi(y)|, \quad \forall \delta > 0. \tag{6.3}$$

Now, let us define for any positive integer k ,

$$\begin{aligned} g_k : \mathbb{R}_+ &\rightarrow \mathbb{R}_+, \\ x &\mapsto \begin{cases} x \exp\left(-\frac{1}{kx^2}\right) & \text{if } x > 0, \\ 0 & \text{if } x = 0. \end{cases} \end{aligned}$$

It is easy to check that, for any k , g_k is $C^\infty(\mathbb{R}_+)$ and satisfies $g_k^{(n)}(0) = 0$ for all integer $n \geq 0$. In addition, $g_k(x) \rightarrow \infty$ as $x \rightarrow \infty$ which implies that the function $\varphi \circ g_k$ is in $C_c^\infty(\mathbb{R}_+)$. Finally, using for instance Faà di Bruno's identity, it is also easy to check that $(\varphi \circ g_k)^{(n)}(0) = 0$ for all integer $n \geq 1$. So we get that

$$\varphi \circ g_k \in \mathcal{D}, \quad \forall k \geq 1.$$

Now, using that

$$(1 - \exp(-x^2)) \leq x, \quad \forall x \geq 0,$$

we easily get,

$$\sup_{x \in \mathbb{R}_+} |g_k(x) - x| \leq \frac{1}{\sqrt{k}}.$$

Hence, we have, thanks to (6.2) and (6.3), for all $k > \eta_\varepsilon^{-2}$,

$$\sup_{x \in \mathbb{R}_+} |\varphi(g_k(x)) - \varphi(x)| \leq \omega \left(\varphi, \sup_{x \in \mathbb{R}_+} |g_k(x) - x| \right) < \frac{\varepsilon}{2}.$$

Finally, picking $k > \eta_\varepsilon^{-2}$, we get

$$\|f - \varphi \circ g_k\|_\infty \leq \|f - \varphi\|_\infty + \|\varphi - \varphi \circ g_k\|_\infty \leq \varepsilon,$$

which ends the proof. \square

The next lemma shows that \mathcal{D} is indeed a core for the expected generator.

Lemma 6.4. *Let $(A^c, D(A^c))$ be the generator of a Bessel process with dimension $(c-1)/2$ absorbed in 0 (when it is possible i.e. $c < -1/2$). In particular*

$$A^c f(x) = \frac{c}{x} f'(x) + \frac{1}{2} f''(x),$$

for all twice differentiable functions f in $C_0(\mathbb{R}_+)$. Then, the space \mathcal{D} defined in Lemma 6.3 is a core for $(A^c, D(A^c))$.

Proof. First of all, it is well known that the generator of the Bessel process in any dimension contains the space of twice differentiable functions on $(0, \infty)$ (see [23]). As a consequence, $D(A^c)$ contains \mathcal{D} . Now, according to Lemma 6.3 and Proposition 17.9 of [21], we just need to show that \mathcal{D} is invariant under the action of A^c , which is clearly the case. Consequently, \mathcal{D} is indeed a core for $(A^c, D(A^c))$. \square

We can now prove the main result.

Theorem 6.5. *The sequence of contour processes $\mathcal{C}(\mathbb{T}_n)_{n \geq 1}$ associated with the sequence of TOM trees $(\mathbb{T}_n)_{n \geq 1}$ converges weakly to a (possibly absorbed) Bessel process with dimension $(c-1)/2$.*

Proof. The proof relies on Theorem 2.2.5 and Theorem 1.6.1 of [15], showing that the convergence of the sequence of generators implies the weak convergence of the associated processes. To this end, let f be an element of \mathcal{D} . Now set for every $n \geq 1$,

$$f_n(x) = f(x) - \int_{\mathbb{R}_+} f\left(\frac{y}{\sqrt{n}}\right) - f(0) K(0, dy).$$

Clearly, for any integer $n \geq 1$, f_n belongs to the domain \mathcal{L}_n of the generator L_n . Moreover, it is easily seen that f_n converges to f in $C_0(\mathbb{R}_+)$. Now, according to (6.1), we have, for $x > 0$,

$$\begin{aligned} |L_n f_n(x) - A^c f(x)| &\leq |L_n f(0)| + |L_n f(x) - A^c f(x)| \\ &\leq \left| \sqrt{n} (b(\sqrt{n}x)m(\sqrt{n}x) - 1) - \frac{c}{x} \right| |f'(x)| + \frac{1}{2} |f''(x)| |b(\sqrt{n}x)m_2(\sqrt{n}x) - 1| \\ &\quad + \left| nb(\sqrt{n}x) \int_{\mathbb{R}_+} \int_x^{x+y/\sqrt{n}} \frac{f^{(3)}(s)}{2} \left(s - x - \frac{y}{\sqrt{n}} \right)^2 ds K(\sqrt{n}x, dy) \right| + |L_n f(0)|. \end{aligned}$$

Now, using Assumptions 6.1, we get

$$\begin{aligned} |L_n f_n(x) - A^q f(x)| &\leq |S_d(\sqrt{n}x) - 1| \left| \frac{c|f'(x)|}{x} + \frac{1}{2} |f''(x)| |S_v(\sqrt{n}x) - 1| \right| \\ &\quad + \left| nb(\sqrt{n}x) \int_{\mathbb{R}_+} \int_x^{x+y/\sqrt{n}} \frac{f^{(3)}(s)}{2} \left(s - x - \frac{y}{\sqrt{n}} \right)^2 ds K(\sqrt{n}x, dy) \right| + |L_n f(0)|. \end{aligned} \quad (6.4)$$

Hence, we have, for all $x > 1$,

$$|L_n f_n(x) - A^q f(x)| \leq |L_n f(0)| + \mathcal{C} \left(c |S_d(nx) - 1| + \frac{1}{2} |S_v(nx) - 1| + \frac{nb(\sqrt{n}x)m_3(\sqrt{n}x)}{2n^{3/2}} \right),$$

where

$$\mathcal{C} = \max \left(\sup_{x \geq 1} \left| \frac{f'(x)}{x} \right|, \|f''\|_\infty, \|f^{(3)}\|_\infty \right).$$

Now, using that $|S_d(nx) - 1|$ and $|S_v(nx) - 1|$ converges to 0, as n goes to infinity, uniformly on any set of the form $[a, \infty)$ (for $a > 0$). We get the uniform convergence on $[1, \infty)$. In addition, $|L_n f_n(0) - A^q f(0)| = 0$, so we only have to study the supremum over $(0, 1]$. To this end we use Equation (6.4). Let us assume that

$$\limsup_{n \rightarrow \infty} \sup_{x \in (0, 1]} |S_d(nx) - 1| \frac{c|f'(x)|}{x} = \alpha.$$

This implies that there exists two sequences $(k_n)_{n \geq 1} \subset \mathbb{N}$ and $(x_n)_{n \geq 1} \subset (0, 1]$ such that

$$|S_d(k_n x_n) - 1| \frac{c|f'(x_n)|}{x_n} \geq \alpha, \quad \forall n \geq 1,$$

where $k_n \geq n$, for all positive integer n . Now, if

$$\liminf_{n \rightarrow \infty} x_n > 0,$$

using the uniform convergence property of $|S_d(k_n x) - 1|$ and the boundedness of $\frac{c|f'(x)|}{x}$, α has to be equal to 0. On the other hand, if

$$\liminf_{n \rightarrow \infty} x_n = 0,$$

extracting a subsequence if needed, we have

$$\alpha \leq \limsup_{n \rightarrow \infty} |S_d(k_n x_n) - 1| \frac{c |f'(x_n)|}{x_n} \leq M \limsup_{n \rightarrow \infty} \frac{|f'(x_n)|}{x_n} = 0,$$

where M is an upper bound for $|S_d(x) - 1|$. Let us also point out that $\frac{|f'(x_n)|}{x_n}$ converges to 0 because $f'(x)$ goes to 0 as x goes to 0 faster than any polynomial.

The study of the other terms of (6.4) follows the same lines and, so, are left to the reader. Consequently, we obtain the convergence of the sequence of generators and thus the result. \square

This last result implies, in virtue of the result of [28], the convergence of the sequence of trees.

Corollary 6.6. *The sequence of TOM trees $(\mathbb{T}_n)_{n \geq 1}$ converges weakly, in the Gromov-Hausdorff-Prokhorov topology, to a TOM tree whose contour process is a Bessel process with dimension $(c-1)/2$.*

6.2 No scaling limits for naive JCCP

In this section, we show that the convergence of a sequence of chronological tree satisfying b goes to infinity as t increase and that of $K(t, dy)$ converges to $\delta_0(dy)$ cannot be done using JCCP without considering a renormalisation of the metric and the measure. This is due to the following result.

Proposition 6.7. *There is no diffusive limits to any sequence of processes in $\mathbb{D}([0, 1])$ with generators of the form*

$$L_n f(x) = -f(x) + b_n(x) \int_{\mathbb{R}_+} f(x+y) - f(x) K_n(x, dy).$$

The proof of this result relies on the two following Lemmas.

Lemma 6.8. *Define for any $f \in \mathbb{D}([0, 1])$, the negative variations of f by*

$$V_-(f) = \sup_{\sigma} \sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| \mathbf{1}_{f(\sigma_i) < f(\sigma_{i-1})},$$

where the supremum is taken on the set of all partition σ of $[0, 1]$ with $\sigma_i < \sigma_{i+1}$ for all integer $i \leq |\sigma|$, where $|\sigma|$ is the cardinality of the partition σ . Then, the set

$$V_1 = \{f \in D([0, 1]) \mid V_-(f) \leq 1\}$$

is closed for the Skorokhod topology.

Proof. Let f_n be a sequence of functions of $\mathbb{D}([0, 1])$ converging to $f \in \mathbb{D}([0, 1])$ for the Skorokhod topology. Hence, there exists a sequence of strictly increasing continuous function λ_n such that

$$\begin{cases} \|\lambda_n - I\|_{\infty} \xrightarrow{n \rightarrow \infty} 0, \\ \|f_n \circ \lambda_n - f\|_{\infty} \xrightarrow{n \rightarrow \infty} 0, \end{cases}$$

where $\|\cdot\|_\infty$ stands for the uniform norm over $f \in \mathbb{D}([0, 1])$ and I is the identity function on $[0, 1]$. Now, let σ be some partition of $[0, 1]$, then

$$\sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| \mathbf{1}_{f(\sigma_i) < f(\sigma_{i-1})} = \lim_{n \rightarrow \infty} \sum_{i=1}^{|\sigma|} |f_n(\lambda_n(\sigma_i)) - f_n(\lambda_n(\sigma_{i-1}))| \mathbf{1}_{f_n(\lambda_n(\sigma_i)) < f_n(\lambda_n(\sigma_{i-1}))}. \quad (6.5)$$

On the other hand, since λ_n is strictly increasing, there exists some partition $\tilde{\sigma}$ of $[0, 1]$ such that $\tilde{\sigma}_i = \lambda_n(\sigma_i)$, for all i . In particular,

$$\sum_{i=1}^{|\sigma|} |f_n(\lambda_n(\sigma_i)) - f_n(\lambda_n(\sigma_{i-1}))| \mathbf{1}_{f_n(\lambda_n(\sigma_i)) < f_n(\lambda_n(\sigma_{i-1}))} \leq \sup_{\sigma} \sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| \mathbf{1}_{f(\sigma_i) < f(\sigma_{i-1})} \leq 1. \quad (6.6)$$

Finally, equations (6.5) and (6.6) leads to

$$V_-(f) = \sup_{\sigma} \sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| \mathbf{1}_{f(\sigma_i) < f(\sigma_{i-1})} \leq \liminf_{n \rightarrow \infty} \sup_{\sigma} \sum_{i=1}^{|\sigma|} |f_n(\sigma_i) - f_n(\sigma_{i-1})| \mathbf{1}_{f_n(\sigma_i) < f_n(\sigma_{i-1})} \leq 1,$$

which ends to proof. \square

Lemma 6.9. Denote

$$V(f, x) = \sup_{\sigma} \sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| \mathbf{1}_{f(\sigma_i) < f(\sigma_{i-1})},$$

where the supremum is taken on the set of all partition σ of $[0, x]$. Let f be a function of $\mathbb{D}([0, 1])$ such that

$$\begin{cases} V(f, x) < \infty, & \forall x \in [0, 1), \\ V_-(f) < \infty, \\ V(f, x) \xrightarrow{x \rightarrow 1} \infty. \end{cases}$$

Then,

$$f(x) \xrightarrow{x \rightarrow 1} \infty.$$

Proof. Let f be such a function. Without loss of generality, we may assume that $f(0) = 0$. Because of $V(f, x) \xrightarrow{x \rightarrow 1} \infty$, we have that, for any $M > 0$, there exists some partition σ of $[0, 1]$ such that

$$\sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| > M,$$

and $\sigma_1 = 0$ and $\sigma_{|\sigma|} = 1$. Now, we have

$$\begin{aligned} f(1) &= \sum_{i=1}^{|\sigma|} f(\sigma_i) - f(\sigma_{i-1}) = \sum_{i=1}^{|\sigma|} (f(\sigma_i) - f(\sigma_{i-1})) \mathbf{1}_{f(\sigma_i) \geq f(\sigma_{i-1})} - \sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| \mathbf{1}_{f(\sigma_i) < f(\sigma_{i-1})} \\ &\geq \sum_{i=1}^{|\sigma|} (f(\sigma_i) - f(\sigma_{i-1})) \mathbf{1}_{f(\sigma_i) \geq f(\sigma_{i-1})} - V_-(f). \end{aligned}$$

Moreover, it is easily seen that

$$\begin{aligned} \sum_{i=1}^{|\sigma|} (f(\sigma_i) - f(\sigma_{i-1})) \mathbf{1}_{f(\sigma_i) \geq f(\sigma_{i-1})} &= \sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| - \sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| \mathbf{1}_{f(\sigma_i) < f(\sigma_{i-1})} \\ &\geq \sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| - V_-(f). \end{aligned}$$

Hence,

$$f(1) = \sum_{i=1}^{|\sigma|} f(\sigma_i) - f(\sigma_{i-1}) \geq \sum_{i=1}^{|\sigma|} |f(\sigma_i) - f(\sigma_{i-1})| - 2V_-(f) \geq M - V_-(f).$$

Finally, since M is arbitrary, we obtain the desired result. \square

In particular, let us highlight that a diffusion has unbounded variation on any compact set. So, such a diffusion as one in the preceding lemma would instantaneously explode. We can now prove our proposition.

Proof. Let $(\mu_n)_{n \geq 1}$ be the sequence of measure on $\mathbb{D}([0, 1])$ corresponding to the laws the sequence $(X^n)_{n \geq 1}$. Assume that $(\mu_n)_{n \geq 1}$ converges weakly to some measure μ . Since, the only negative variation of X^n is provided by the negative drift (with slope 1), we have that

$$V_-(X^n) \leq 1, \quad \text{almost surely.}$$

Hence, $\mu_n(V_1) = 1$ for all $n \geq 1$. Now, since V_1 is closed, the portemanteau theorem entails that

$$1 = \limsup_{n \rightarrow \infty} \mu_n(V_1) \leq \mu(V_1).$$

Hence, we necessarily have $\mu(V_1) = 1$. This ends the proof. \square

Appendix: a quick reminder on Generator

Let us begin by reminding some notions on the generator of Markov processes (which also enable to precise our notation).

Given any Markov process $(Z_t)_{t \geq 0}$ on \mathbb{R}_+ (or any more general state space). One can define its associated semi-group $(P_t)_{t \geq 0}$ by

$$\forall t \geq 0, \forall x \in \mathbb{R}_+, \quad P_t f(x) = \mathbb{E}[f(Z_t) \mid Z_0 = x],$$

for every function f in some space of functions L . This space of functions can be chosen as the space of continuous functions vanishing at infinity $C_0(\mathbb{R}_+)$, the space of bounded and continuous function $C_b(\mathbb{R}_+)$ or the space $L^2(\mathbb{R}_+, \pi)$ of function which are square-integrable function with respect to the invariant measure of the semi-group (when it exists); see for instance [4, 15]. If we assume that $(L, \|\cdot\|_L)$ is a Banach space (for some norm $\|\cdot\|_L$) and $P_t : L \rightarrow L$, for every $t \geq 0$ is strongly continuous; that means that

$$\lim_{t \rightarrow 0} \|P_t f - f\|_L = 0,$$

for every continuous function $f \in L$, then one can define the *infinitesimal generator* or *strong generator* A of the semi-group $(P_t)_{t \geq 0}$ on L by

$$Af = \lim_{t \rightarrow 0} \frac{1}{t} (T_t f - f),$$

for every f in $\mathcal{D}_L(A) \subset L$, the domain of A , which corresponds to the functions $f \in L$ such that the previous limit holds. Recall that when $L = C_0$ then the process $(Z_t)_{t \geq 0}$ is called a Feller process. A powerful property of the generator is given by the Dynkin formula and reads as the process $(M_t)_{t \geq 0}$, given by

$$M_t = f(X_t) - \int_0^t Af(X_s) ds, \quad (6.7)$$

is a martingale for any $f \in L$. see in particular [10, Theorem (14.13) p.31]. In order to generalise this property, one can define the *full generator* of $(Z_t)_{t \geq 0}$ on the domain

$$\overline{D}_L(A) = \left\{ f \in L \mid \exists g_f \in L, \forall t \geq 0 \quad P_t f = f + \int_0^t P_t g_f ds \right\}.$$

For $f \in \overline{D}_L(A)$, we set $Af = g_f$. Again, (M_t) defined by (6.7) is a martingale for all $f \in \overline{D}(A)$; see [15, Proposition 1.7 p.162]. We can again weaken the assumption on f by considering directly the space $\widehat{D}_L(A)$ of functions $f \in L$ such that there exists $g_f \in L$, such that $(M_t)_{t \geq 0}$, defined by

$$\forall t \geq 0 \quad , M_t = f(Z_t) - \int_0^t g_f(X_s) ds,$$

is a local Martingale. Again we set $Af = g_f$, and A with its domain $\widehat{D}_L(A)$ is called the extended generator. Note that the extended and full generator are multi-valued, in the sense that it is possible to find different g_f for the same f . However, these functions will be different only on a negligible set (see [10, page 32-33]). Also, there is no problem of notation because $D_L(A) \subset \overline{D}_L(A) \subset \widehat{D}_L(A)$ and if $L' \subset L$ then $D_{L'}(A) \subset D_L(A)$, $\overline{D}_{L'}(A) \subset \overline{D}_L(A)$ and $\widehat{D}_{L'}(A) \subset \widehat{D}_L(A)$.

References

- [1] Louigi Addario-Berry, Luc Devroye, and Svante Janson. Sub-gaussian tail bounds for the width and height of conditioned galtonwatson trees. *Ann. Probab.*, 41(2):1072–1087, 03 2013.
- [2] David Aldous. The continuum random tree iii. *The Annals of Probability*, pages 248–289, 1993.
- [3] R. Azaïs, J.-B. Bardet, A. Génadot, N. Krell, and P.-A. Zitt. Piecewise deterministic Markov process – recent results. *ESAIM: Proceedings*, 44:276–290, 2014.
- [4] Dominique Bakry. L’hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, Berlin, 1994.
- [5] J. Bertoin. *Lévy Processes*. Cambridge Tracts in Mathematics. Cambridge University Press, 1998.

- [6] Nicolas Champagnat and Benoît Henry. Moments of the frequency spectrum of a splitting tree with neutral poissonian mutations. *Electron. J. Probab.*, 21:34 pp., 2016.
- [7] Nicolas Champagnat, Amaury Lambert, and Mathieu Richard. Birth and death processes with neutral mutations. *Int. J. Stoch. Anal.*, pages Art. ID 569081, 20, 2012.
- [8] B. Cloez, R. Dessalles, A. Genadot, F. Malrieu, A. Marguet, and R. Yvinec. Probabilistic and Piecewise Deterministic models in Biology. *ArXiv e-prints*, June 2017.
- [9] M. H. A. Davis. Piecewise-deterministic markov processes: a general class of nondiffusion stochastic models. *J. Roy. Statist. Soc. Ser. B*, 46(3):353–388, With discussion, 1984.
- [10] M. H. A. Davis. *Markov models and optimization*, volume 49. Monographs on Statistics and Applied Probability, Chapman & Hall, 1993.
- [11] Benoîte de Saporta, François Dufour, and Huilong Zhang. *Numerical methods for simulation and optimization of piecewise deterministic Markov processes*. John Wiley & Sons, 2015.
- [12] Joseph L Doob. Conditional brownian motion and the boundary limits of harmonic functions. *Bull. Soc. Math. France*, 85(4):431–458, 1957.
- [13] Randal Douc, Gersende Fort, and Arnaud Guillin. Subgeometric rates of convergence of f-ergodic strong markov processes. *Stochastic Processes and their Applications*, 119(3):897 – 923, 2009.
- [14] Thomas Duquesne. The coding of compact real trees by real valued functions. *arXiv preprint math/0604106*, 2006.
- [15] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics, 1986.
- [16] J. Geiger and G. Kersting. Depth-first search of random trees, and Poisson point processes. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 111–126. Springer, New York, 1997.
- [17] P. Haccou, P. Jagers, V.A. Vatutin, and International Institute for Applied Systems Analysis. *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge Studies in Adaptive Dynamics. Cambridge University Press, 2005.
- [18] Martin Hairer. Convergence of markov processes. <http://www.hairer.org/notes/Convergence.pdf>, January 2016.
- [19] Henry, Benoît. Central limit theorem for supercritical binary homogeneous crump-mode-jagers processes. *ESAIM: PS*, 21:113–137, 2017.
- [20] F Hubalek and E Kyprianou. Old and new examples of scale functions for spectrally negative lévy processes. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 119–145. Springer, 2011.

- [21] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [22] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, 2014.
- [23] John Kent. Some probabilistic properties of Bessel functions. *Ann. Probab.*, 6(5):760–770, 1978.
- [24] Vassili N Kolokoltsov. *Markov processes, semigroups, and generators*, volume 38. Walter de Gruyter, 2011.
- [25] Igor Kortchemski. Sub-exponential tail bounds for conditioned stable bienaymé–galton–watson trees. *Probability Theory and Related Fields*, 168(1-2):1–40, 2017.
- [26] Amaury Lambert. The contour of splitting trees is a Lévy process. *Ann. Probab.*, 38(1):348–395, 2010.
- [27] Amaury Lambert. Probabilistic models for the (sub)tree(s) of life. *Braz. J. Probab. Stat.*, 31(3):415–475, 08 2017.
- [28] Amaury Lambert and Gerónimo Uribe Bravo. Totally ordered measured trees and splitting trees with infinite variation. *arXiv preprint arXiv:1607.02114*, 2016.
- [29] Amaury Lambert and Lea Popovic. The coalescent point process of branching trees. *Ann. Appl. Probab.*, 23(1):99–144, 02 2013.
- [30] Amaury Lambert and Tanja Stadler. Birth–death models and coalescent point processes: The shape and probability of reconstructed phylogenies. *Theoretical population biology*, 90:113–128, 2013.
- [31] Jean-François Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311, 2005.
- [32] V. Lemaire, M. Thieullen, and N. Thomas. Exact simulation of the jump times of a class of Piecewise Deterministic Markov Processes. *ArXiv e-prints*, February 2016.
- [33] Florent Malrieu. Some simple but challenging Markov processes. *Ann. Fac. Sci. Toulouse Math. (6)*, 24(4):857–883, 2015.
- [34] Sean Meyn, Richard L. Tweedie, and Peter W. Glynn. *Markov Chains and Stochastic Stability*. Cambridge Mathematical Library. Cambridge University Press, 2 edition, 2009.
- [35] Sean P. Meyn and R. L. Tweedie. Stability of markovian processes ii: continuous-time processes and sampled chains. *Advances in Applied Probability*, 25(3):487–517, 1993.
- [36] Sean P. Meyn and R. L. Tweedie. Stability of markovian processes iii: Foster-lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3):518–548, 1993.
- [37] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

- [38] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2013.
- [39] M. Richard. Limit theorems for supercritical age-dependent branching processes with neutral immigration. *Advances in Applied Probability*, 43(1):276300, 2011.
- [40] Mathieu Richard. *Arbres, Processus de branchement non Markoviens et Processus de Lévy*. Theses, Université Pierre et Marie Curie - Paris VI, December 2011.
- [41] LCG Rogers and David Williams. Diffusions, markov processes and martingales: Itô calculus , volume 2 of cambridge mathematical library, 2000.