UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y
DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

ON CONDITIONED RANDOM WALKS, MULTITYPE RANDOM
FORESTS AND EXCHANGEABLE INCREMENT PROCESSES

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA:
OSVALDO ANGTUNCIO HERNÁNDEZ

DIRECTOR DE LA TESIS

GERÓNIMO URIBE BRAVO
INSTITUTO DE MATEMÁTICAS

MIEMBROS DEL COMITÉ TUTOR

FERNANDO BALTAZAR LARIOS
FACULTAD DE CIENCIAS, DEPARTAMENTO DE MATEMÁTICAS

RAMSÉS MENA CHÁVEZ
INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y EN SISTEMAS

CIUDAD DE MÉXICO, JULIO DEL 2019

Contents

INTEGRATES DEL JURADO

Titulares:

Dra. María Emilia Caballero Acosta
Dr. Loïc Chaumont
Dra. María Clara Fittipaldi
Dra. Sandra Palau Calderón
Dr. Gerónimo Francisco Uribe Bravo

AGRADECIMIENTOS

El primer y mayor agradecimimento va dedicado a mi familia. Ellos me procuraron una muy buena vida y educación, dándome lo suficiente para que pudiera lograr mis objetivos. A mi mamá María Teresa Hernández Jerónimo porque eres una persona increíble y me llenaste de amor. Mi papá Gilberto Angtuncio Cid que estás disponible cuando lo necesito. Mi tía Amelia Hernández Gerónimo que has sido mi segunda madre y un gran apoyo en mi vida. Gracias también hermano Gilberto Angtuncio Hernández, que aunque seas una molestia, aún así te quiero. También te agradezco Blanca Anai García Mendoza por ser una muy buena persona e increíble madre para mis sobrinas favoritas Lian y Valeria Angtuncio García, que le dan una alegría indescriptible a mi vida. Tía y abuela Elvira, y tía Cristina, gracias por procurarme constantemente y estar pendientes de mi.

El siguiente agradecimiento va dedicado a mi tutor Gerónimo Uribe Bravo. Todos saben la gran admiración que te tengo, así como el profundo respeto por el matemático que eres y por tu dedicación; aspiro a alcanzar tu nivel.

También agradezco a la Dra. María Emilia Caballero Acosta, te debo mucho, por tus consejos, constante apoyo, interés, así como por tus enseñanzas matemáticas. Gracias a mi comité tutorial Ramsés Mena Chávez y Fernando Baltazar Larios, que me ayudaron cuando lo necesité, resolvieron mis dudas y me aconsejaban siempre con muy buena intención. Al Consejo Nacional de Ciencia y Tecnología CoNaCyT, por la beca doctoral con número de CVU 419082; así como el proyecto del Fondo Institucional FOINS como estudiante de proyecto de investigación FC-2016-1946 y el proyecto UNAM-DGAPA-PAPIIT con número IN115217. A la Dra. María Emilia Caballero Acosta y al CoNaCyT por brindarme el estímulo económico como ayudante de investigador nacional nivel III con expediente de ayudante 15682.

Al jurado de esta tesis Dra. María Emilia Caballero Acosta, Dr. Loïc Chaumont, Dra. María Clara Fittipaldi y Dra. Sandra Palau Calderón, así como los suplentes Dr. Adrián González Casanova, Dr. Ramsés Mena Chávez, todos son personas con quienes colaborar es un objetivo en mi vida, y a quienes les he aprendido mucho.

Un agradecimiento especial para mi mejor amigo Alejandro Betanzos Gómez y a Mayra Elizabeth Muñiz Lozano. Contigo soy tan afín, me divierto tanto y siempre disfruto nuestras conversaciones. Muchas gracias por tus consejos y por tu patrocinio financiero incondicional. A los amigos del cubículo Miriam, Mónica y César, por tantas pizzas y risas juntos. Miriam, muchísimas gracias por tu compañia y ayuda este tiempo que me son muy valiosas. Los amigos de la maestría Uriel Díaz, Juan Manuel Sánchez, Jorge González, Oscar Peralta, Alan Rivapalacio, Michelle Anzarut. Gracias a mis amigos de juegos Jorge Olivera Cabrera y Marco Andrés Vázquez Hernández, espero nunca se pierda la tradición. Mis buenos amigos de la juventud Antonio Mata, Osvaldo Sánchez y a los demás amigos del Tae, por nuestras tantas aventuras. A mi nana Josefina por todo el amor que me brindó. A Salvador García Pacheco porque fuiste un guía en muchos aspectos de mi vida.

Finalmente te agradezco profundamente Tania Paola Berrueco Ávila, eres inconmensurablemente importante para mí y representas gran parte de lo que soy.

Chapter 1

INTRODUCTION

This thesis focuses on 4 topics:

1. Construction of a d-dimensional random walk $X = (X^1, \dots, X^d)$ conditioned to have its components ordered forever, that is, conditioned on $A = \{X_j^1 \leq \cdots \leq X_j^d \text{ for all } j \geq 0\}$. Our hypotheses are minimal, in particular, the components can be dependent, have drift, and no moment condition is imposed. This is done conditioning X to stay ordered up to an geometric time, and seeing this as a Doob h-transform. Then let the parameter of the geometric tend to infinity. The proof also uses a generalization of the ladder height times of a random walk, allowing us to recover several expressions of the (sub)-harmonic function h in the unidimensional case, that is, a random walk conditioned to stay positive [Tan89, Ber93, CD05].

2. Convergence of the rescaled profile process (giving the number of individuals in each generation) of uniform trees with a given degree sequence (TGDS). By mixing, this model is a Bienamé-Galton-Watson (GW) tree conditioned on its size. Hence, we generalize the results of Drmota and Gittenberger [DG97] and Kersting [Ker11], about the convergence of the profile of GW trees conditioned on its number of individuals (cGW). Our result relies on an encoding of the TGDS through the Vervaat transformation (introduced in [Ver79]) of a discrete time exchangeable increments (EI) process. Then, a time-change equation in the discrete case, relating the profile with such encoding, is proved to hold in the limit. To prove that the profile converges, we need to prove that such continuous time-change equation converges either to the zero function or the unique positive solution. This is done introducing a novel path transformation, implying that in the limit, the tree is not thin near the root. We also give a direct and simple proof of the convergence of the profile for a cGW tree with offspring distribution in the domain of attraction of a stable law. This is based on a stochastic bound of the cGW tree with Kesten's tree, that is, a tree conditioned on non-extinction. We prove that Kesten's tree is not thin at the base, hence neither the cGW tree.

3. Construction and simulation of uniform multitype random forests with a given degree sequence (MFGDS), and of multitype Galton-Watson forests conditioned by its number of individuals by type (cMGW). Under an independence assumption, mixing this model results in a cMGW. To obtain MFGDS, we define a multitype degree sequence, construct a multidimensional discrete time EI process, and generalize the Vervaat transform (the latter using the results in [CL16]). The joint law of the number of individuals by types in a cMGW is also obtained, generalizing the Otter-Dwass formula [Ott49, Dwa69], which gives the law of the size of a GW tree. We obtain enumerations

of forests with prescribed roots and individuals by types, having a *combinatorial structure*, namely, plane, labeled and binary multitype forests. We also give an algorithm to simulate cMGW using MFGDS, generalizing Devroye's algorithm [Dev12] for the unitype case.

4. For X a continuous time EI process of infinite variation and $t \in [0,1)$ fixed, we prove that the superior and inferior limits of $(X_{t+h} - X_t)/h$ are $+\infty$ and $-\infty$, respectively, as $h \downarrow 0$. This extends Rogozin's result for Lévy processes [Rog68]. Our main tool is a change of measure of X, which reduces to the Esscher transform in the Lévy process setting. The applications of this result are the following:

   (a) Both half-lines $(0,\infty)$ and $(-\infty,0)$ are visited immediately (upward and downward regularity).

   (b) X is continuous at its infimum iff its upward and downward regular, generalizing Millar's zero-one law [Mil77] for Lévy processes.

   (c) Weak convergence of X conditioned to have its infimum above $-\varepsilon$ to the Vervaat transform of X, as ε goes to zero. This generalizes results of [DIM77, UB14, CUB15].

   (d) Construction of the convex minorant of X, generalizing [PUB12] for Lévy processes. As a consequence of this result, we also obtain for EI processes the formula $\mathbb{E}(\inf_{0\le s\le 1} X_s) = \int_0^1 \frac{\mathbb{E}(X_l \wedge 0)}{l}\, dl$ (in the discrete case, this is Kac's formula [Kac54]).

Each topic is presented as a chapter. Chapter 2 has been accepted for publication in the Volume of the XIII Symposium on Probability and Stochastic Processes. We present first a brief discussion of each chapter, along with our main results.

1.1 MULTIDIMENSIONAL RANDOM WALKS CONDITIONED TO STAY ORDERED VIA GENERALIZED LADDER HEIGHT FUNCTIONS

1.1.1 Preliminaries and motivation

In Chapter 2 we obtain a new construction of a d-dimensional random walk $X = (X^1,\dots,X^d)$, started at $x = (x_1,\dots,x_d)$ with $x_1 < \cdots < x_d$, conditioned on the event

$$A = \{X_j^1 < \cdots < X_j^d \text{ for all } j \ge 0\}.$$

Such processes have an important relation with random matrix theory, Dyson's Brownian motion being the classical example [Dys62] (a process with the same law as d independent Brownian motions conditioned to stay ordered forever). Another well-studied model is the case $d = 2$, which is equivalent to condition a process to stay non-negative, see [Tan89, Ber93, CD05]. Non-intersecting paths are also used to model some physical phenomena [Fis84]. It also has connections with Young diagrams and domino tiling [Kön05].

It is well-known that whenever the components of X are independent and $\mathbb{E}(X_1) = 0$, then $\mathbb{P}(A) = 0$. Hence, one has to give a meaning to a *random walk conditioned to have ordered components forever*. We achieve this by conditioning X to stay ordered up to a geometric time G_c with parameter $1 - e^{-c}$ and let $c \downarrow 0$ (see Proposition 2.2 and Theorem 2.1). Under certain hypotheses, similar models of ordered random walks have been constructed: in [EK08] when X has i.i.d. components, satisfies a moment condition and a local limit theorem; in [DW15] for random walks in cones, assuming zero drift, zero covariance between the components of X, and a moment condition on X_1; in [Dur14b] when the drift is not zero, under a Cramér condition; in [Ign18] for substochastic transient random walks taking values on a countable semi-group with transition probabilities satisfying a certain inequality. It is important to remark that our model has minimal assumptions, namely, on some cases, we assume the existence of positive $\varepsilon_1,\dots,\varepsilon_{d-1}$ such that

$$\mathbb{P}\Big(X_1^2 - X_1^1 \geq \varepsilon_1,\dots,X_1^d - X_1^{d-1} \geq \varepsilon_{d-1}\Big) > 0, \tag{1.1}$$

needed only to prove the finiteness of the (sub)harmonic function h (see Subsection 2.3.2). In particular, neither a moment condition nor a countable space has to be assumed, as in the cited papers. We prove that our construction gives a Markovian chain (sub-Markovian) whenever the first time two components are not ordered has infinite (finite) expectation (see Lemma 2.6). The proof of our theorem is based on a generalization of the *ladder height process* of a random walk, giving us several generalizations of formulas for the unidimensional case (see Section 2.4).

1.1.2 Statement of the results

For ease of notation, our results are stated for $d=3$ and for a random walk having state space $\mathbb{R}^d$. Let $X=(X^1,X^2,X^3)$ be a 3-dimensional random walk on $\mathbb{R}^3\cup\{\dagger\}$, starting at $X_0=0$, having lifetime $\zeta=\sup\{n : X_n\neq\dagger\}$. Its increments are denoted by $W=(W^1,W^2,W^3)$, and W_1 has law $\mathbb{P}$. We denote by $Y=(Y^1,Y^2)=(X^2-X^1,X^3-X^2)$ the size of the gap between components. The law of X killed at time $n\in\mathbb{N}$, that is, on the event $\zeta=n$, will be denoted by $\mathbb{P}^n$. When killing X at an independent geometric law $N\in\{0,1,\dots,\}$ with parameter $1-e^{-c}$, its law will be $\mathbb{P}^c$. The σ-algebra considered will be $\mathscr{F}_n=\sigma(X_1,\dots,X_n)$.

The notation that we use is component-wise, hence $\min\{Y_i, i\in\mathscr{I}\}=(\min\{Y_i^1, i\in\mathscr{I}\},\min\{Y_i^2, i\in\mathscr{I}\})$, for any index set $\mathscr{I}\subset\mathbb{Z}_+$. We define for $[n]=\{1,2,\dots,n\}$ and $[n]_0=\{0,1,\dots,n\}$, the processes $\underline{Y}_n=\min\{Y_i, i\in[n]_0\}$, $\underline{\underline{Y}}_n=\min\{Y_i, i\in[n]\}$ and $Y_n\vee(y_1,y_2)=(Y_n^1\vee y_1, Y_n^2\vee y_2)$, where $y_1,y_2\in\mathbb{R}$. We also put $(x_1,\dots,x_d)<(y_1,\dots,y_d)$ whenever component-wise the strict inequality is satisfied.

For $\mathbb{W}=\{(x_1,\dots,x_d)\in\mathbb{R}^d : x_1<\cdots<x_d\}$, a *positive regular function* or *harmonic function*, with respect to the transition kernel of X to $\mathbb{W}$, is a function $h:\mathbb{W}$ such that

$$\mathbb{E}_x\left(h(X_1);\tau>1\right)=h(x) \qquad x\in\mathbb{W},$$

where

$$\tau:=\min\{n : X_n\notin\mathbb{W}\}.$$

A function h is *subharmonic* (*superharmonic*) if $\mathbb{E}_x\left(h(X_1);\tau>1\right)\leq h(x)$ ($\mathbb{E}_x\left(h(X_1);\tau>1\right)\geq h(x)$) for every $x\in\mathbb{W}$.

To avoid trivial cases, we assume that Y has components taking positive and negative values with positive probability. Besides that, the construction works with no further hypothesis if either some component of Y drifts to $-\infty$, or every component of Y drifts to $+\infty$. When such conditions are not satisfied, we need Hypothesis (1.1), needed only for the finiteness of the (sub)harmonic function h. Our main result is the following.

Theorem 1.1. *Let X be a d-dimensional random walk. Let N be a geometric time with parameter $1-e^{-c}$, independent of X. Assume that*

$$h^{\uparrow}(x) := 1+\mathbb{E}\left(\sum_{n=1}^{J_1-1} \mathbf{1}_{\overline{Y}_{n-1}-Y_n<y}\right) < \infty \qquad x=(x_1,\dots,x_d)\in\mathbb{W},$$

with $y=(x_2-x_1,\dots,x_d-x_{d-1})$ and $J_1=\inf\{n>0 : \overline{Y}^k_{n-1} < Y^k_n, k\in[d-1]\}$. Then, for every $x\in\mathbb{W}$, every finite $\mathscr{F}_n$-stopping time T and $\Lambda\in\mathscr{F}_T$

$$\lim_{c\to 0^+} \mathbb{P}_x\left(\Lambda, T\le N | X(i)\in\mathbb{W}, i\in[N]\right) = \mathbb{P}^{\uparrow}_x(\Lambda, T<\zeta) := \frac{1}{h^{\uparrow}(x)}\mathbb{E}^Q_x\left(h^{\uparrow}(X_T)\mathbf{1}_{\Lambda,T<\zeta}\right),$$

where $\mathbb{E}^Q_x$ is the expectation under the law of X killed at the first exit time of $\mathbb{W}$. The limit law is a Markov chain with transition probabilities

$$p^{\uparrow}(w,dz) = \mathbf{1}_{z\in\mathbb{W}}\frac{h^{\uparrow}(z)}{h^{\uparrow}(w)}p(w,dz) \qquad w\in\mathbb{W}. \tag{1.2}$$

Moreover, it is a probability measure if $\mathbb{E}(\tau)=\infty$, or a subprobability measure if $\mathbb{E}(\tau)<\infty$.

Indeed, this theorem comprises several of our results: in Lemma 2.5 we prove the existence of the limit of the resulting (sub)harmonic function when conditioning to be ordered up to a geometric time, say $h^{\uparrow}=\lim_{c\downarrow 0} h^{\uparrow}_c$; in Theorem 2.1 we prove the limit is a Markov chain and has as (sub)harmonic function $h^{\uparrow}$; in Lemma 2.6 we characterize when $h^{\uparrow}$ is harmonic or subharmonic; and the limits of the probability laws conditioning up to a geometric time, is given in Lemma 2.8.

It has to be noted that when $d=1$, the above formula for $h^{\uparrow}$ coincides with Bertoin's function h, see [Ber93, p. 22].

We also give simple conditions to ensure $h^{\uparrow}$ is finite.

Lemma 1.1. *Assume that either*

1. *some component of Y drifts to $-\infty$,*
2. *every component of Y drifts to $+\infty$ and $\mathbb{P}(\tau=\infty)>0$,*
3. *there exists $\varepsilon=(\varepsilon_1,\dots,\varepsilon_{d-1})\in\mathbb{R}^{d-1}_+$ such that*

$$\mathbb{P}(Y_1>\varepsilon)>0.$$

Then

$$h^{\uparrow}(x)<\infty \qquad \forall x\in\mathbb{W}.$$

The following is an application of Theorem 1.1.

Example 1.1. *For $d\ge 1$, consider a multidimensional random walk with partial sums $X_n=(an, X^1_n,\dots,X^d_n, bn)$, where $a<b$ and $a,b\in\mathbb{R}$. Assume that $\mathbb{P}\left(X^1_1-a>\varepsilon_1, X^2_1-X^1_1>\varepsilon_2,\dots,b-X^d_1>\varepsilon_{d+1}\right)>0$ for some $\varepsilon_i>0$, $i\in\{1,\dots,d+1\}$, and $\mathbb{P}\left(X^1_1-a<0, X^2_1-X^1_1<0,\dots,b-X^d_1<0\right)>0$. Then, the construction of Theorem 1.1 provides us with a d-dimensional random walk conditioned to have ordered components and staying inside the set $\{x\in\mathbb{R} : at<x<bt, \forall\, t\in\mathbb{R}_+\}$.*

Under some hypothesis concerning the drift of Y, we can reexpress the (sub)harmonic function h as follows (see Lemma 2.9).

Lemma 1.2. *Fix $x \in \mathbb{W}$. When some component of Y drifts to $-\infty$, we have*

$$h^{\uparrow}(x) = \frac{\mathbb{E}_x(\tau)}{\mathbb{E}(\tau)}.$$

When every component drifts to $+\infty$ and $\mathbb{P}(\tau = \infty)$ is positive, then

$$h^{\uparrow}(x) = \frac{\mathbb{P}_x(\tau = \infty)}{\mathbb{P}(\tau = \infty)}.$$

Another reexpression of $h^{\uparrow}$ is the following. For $k = 1, 2$ denote by $(\beta_i^k, i \in \mathbb{N})$ the strict descending ladder times of Y^k, so $\beta_0^k = 0$ and $\beta_i^k = \min\{n : Y^k_{\beta^k_{i-1}+n} < Y^k_{\beta^k_{i-1}}\}$. Let $\{\beta_1, \dots\}$ be the ordered union of the ladder times $\{\beta_i^1, \beta_j^2, i, j \geq 1\}$. Set $\beta_0 = 1$. Then we have the following (see Proposition 2.4).

Proposition 1.1. *Fix $x \in \mathbb{W}$. Then*

$$h^{\uparrow}(x) = 1 + \sum_{n=1}^{\infty} \mathbb{P}\left(-\underline{\underline{Y}}_{\beta_n} < y\right).$$

We remark that the above formulas also extend the ones known in the unidimensional case, see [BD94, p. 2155].

1.2 ON THE PROFILE OF TREES WITH A GIVEN DEGREE SEQUENCE

1.2.1 Preliminaries and motivation

For a tree, we mean a rooted plane tree, that is, a connected graph with no cycles having a distinguished vertex, together with a natural identification of each vertex by a finite sequence of non-negative integers (indicating its location on the tree). The formal definition of rooted plane trees is stated in Section 3.1, see also Figure 3.1 fon an example of such labeling. A branching tree τ with offspring distribution ν, is a random tree, starting with a common ancestor and where each individual, independently of the others in the same generation, has offspring according to ν. This is also called Bienaymé-Galton-Watson (BGW) tree, or simply GW tree.

Branching theory has its origin in the second half of the XIXth century, arising from a demographic question about the probability of extinction of surnames in noble families. Galton formulated the question as follows:

"A large nation, of whom we will only concern ourselves with adult males, N in number, and who each bear separate surnames colonize a district. Their law of population is such that, in each generation, a_0 percent of the adult males have no male children who reach adult life; a_1 have one such male child; a_2 have two; and so on up to a_5 who have five. Find (1) what proportion of their surnames will have become extinct after r generations; and (2) how many instances there will be of the surname being held by m persons."

Bienaymé [Bie45] was the one to ensure (without mathematical justification), that the probability of extinction of τ is one if the mean of the reproduction law is smaller than one. It was until 1875, that Galton and Watson [WG75] used a method involving generating functions. Unfortunately, their conclusion was incorrect; it was until 1930 that Steffensen gave the correct solution [Ste30].

Undoubtedly, branching processes serve as the most simple model for genealogical evolutions and the dynamic of populations of individuals living and giving birth independently of the others. It has applications in different areas such as demography, genealogy, genetics, cell kinetics, epidemics, computer communications networks, statistical physics, data storage algorithms, ecology, image processing, etc. [Jag75, AJ97, Pak03, HJV07, All11, ABD18].

Branching processes are also an important tool in pure mathematics. At the end of the past century, there was a special interest in the asymptotic behavior of certain properties of random trees. Properties such as the individual with maximum offspring, the number of individuals with a fixed degree, or the profile of the tree. Among such large random trees studied, are random trees conditioned to be large. The book of Drmota [Drm09] reviews several of such models. But in the 90's, Aldous [Ald91a] studied the convergence of the whole random tree to a new random object, the so-called *Continuum Random Tree* (CRT). This is a universal limit in the sense that the normalized contour function of every GW conditioned to have n vertices, with critical offspring distribution having finite variance converges as $n \to \infty$, to the normalized Brownian excursion, see [Ald91b, Ald93]. The formalism was further developed by [LGLJ98, DLG02, EPW06], where the theory of $\mathbb{R}$-trees (seeing such trees as abstract metric spaces) was used to prove the convergence for more general offspring laws, such as laws belonging to the domain of attraction of a stable law. This theory resulted in the introduction of Lévy trees.

The previous convergence to the CRT, links in a very nice way the branching processes theory with combinatorics. Depending on the offspring distribution, one can obtain asymptotic features on certain classes of *combinatorial trees* with a given number of vertices. For example, the law of a GW tree with geometric offspring distribution conditioned to have n individuals, is the uniform law on the set of all rooted plane trees with n vertices. Similarly, with a Poisson offspring distribution the conditioned GW tree has the uniform law on the set of all rooted Cayley (labeled) trees with n vertices. As a last example, a Bernoulli-type distribution gives rise to the uniform law over binary trees with n vertices. See [Pit98] for references. Since we know the convergence of the random conditioned trees to the CRT, we can obtain several features of *typical* trees on such classes: height of the tree, number of vertices in the tree between given generations, number of vertices in a generation that have descendants at a given generation, etc. As a remark, not all the asymptotic properties of the tree can be obtained in this way. In particular, the convergence of the whole tree does not imply the convergence of its profile (number of individuals at each generation). Indeed, Broutin and Marckert [BM14b] proved the convergence of a certain class of conditioned random trees, the so-called uniform trees with a given degree sequence, and we proved that there is also convergence of the profile. Actually, we proved this result in a much more general setting than in [BM14b].

Uniform trees with a given degree sequence are constructed as follows: let $\mathbf{S} = (N_i, i \geq 0)$ be a sequence of non-negative integers such that $s := \sum_i N_i = 1 + \sum_i iN_i$, and take a uniform tree from the set of trees with degree sequence $\mathbf{S}$. Let $\mathbb{P}_{\mathbf{S}}$ be the distribution that samples uniformly at random a tree from the set of trees with degree sequence $\mathbf{S}$. Such trees are a particular case of uniform random graphs with a given degree sequence, which are important models for real-world networks (see [BA99, CDS11]). Indeed, it has been shown that many of the latter have special features such as a degree sequence having power law tails. This feature is not present in the most well-known model of random graphs, namely, the Erdős-Rényi random graph [ER60].

To understand the construction of random trees and scaling limits of its characteristics, it is easier to code them with random paths and analyze the convergence of the latter.

1.2.2 Forests and excursions

A rooted plane forest F is a finite sequence of rooted plane trees, say $(T_1,\dots,T_k)$. For any individual u in a tree T, denote by $c(u)$ its number of children. We code and order the individuals of a plane forest T in two different ways.

Depth-first order: Order the vertices the tree T_1 according to the *lexicographical order* (e.g., $\varnothing < 1 < 21 < 22$), and assign label i to the ith vertex in T_1. Similarly, assign label $|T_1|+\cdots|T_{j-1}|+i$ to the ith vertex in T_j, with $2\leq j\leq k$.

Breadth-first order: To define this, assign label 1 to the root of T_1. Suppose the first generation (offspring of the root) of u_1 has size z_1. Order the first generation in lexicographical order, and assign label i to the ith vertex, for $i\in\{2,\dots,1+z_1\}$. Do the same for each consecutive generation, and after that, for each consecutive tree.

For any $n\in\mathbb{N}$ let $[n]_0:=\{0,1,\dots,n\}$. We say that the path $y:[n]_0\mapsto\mathbb{Z}$ is a *downward skip-free chain*, if $y_{k+1}-y_k\in\mathbb{Z}_+\cup\{-1\}$. There exists two known bijections between plane forests and downward skip-free chains.

Lemma 1.3 (Lemma 6.3 [Pit06] or Proposition 1.1 of [LG05])**.** *Given a plane forest F with n vertices and k trees, let $u_1,\dots,u_n$ be the vertices of F labeled in the breadth-first (or depth-first) order. Then the coding*

$$T\mapsto(c(u_1),\dots,c(u_n))$$

sets up a bijection between the set of plane forests with k trees and n individuals, and sequences of non-negative integers $(x_1,\dots,x_n)$ such that the downward skip-free chain, starting at zero and with steps x_i-1, first reaches $-k$ at time n.

The *depth-first walk* (DFW) of a tree will be the walk started at one, and with ith increment $c(u_i)-1$. The *breadth-first walk* (BFW) of a tree will be the walk started at one, and with ith increment $c(u_i)-1$. These codings are also called *excursions*. See Figures 3.1 and 3.2 for examples. For a minor technical reason, when $k\geq 2$, we start the DFW and BFW of the forest at zero.

1.2.3 Uniform trees with a given degree sequence and discrete time EI processes

In the paper [BM14b], the authors give an algorithm to construct a tree with law $\mathbb{P}_{\mathbf{S}}$ from a degree sequence $\mathbf{S}$, and which is easier to simulate. To introduce it, we need to define a discrete time EI process.

Definition 1.1. *Let $s\in\mathbb{N}$. A discrete time process $(W^b(j),j\in[s]_0)$ with increments $\Delta W^b(i)=W^b(i)-W^b(i-1)$ has exchangeable increments (EI) if for every deterministic permutation σ on $[s]$*

$$\left(\Delta W^b(1),\dots,\Delta W^b(s)\right)\overset{d}{=}\left(\Delta W^b(\sigma_1),\dots,\Delta W^b(\sigma_s)\right).$$

To transform W^b to an excursion we use a *cyclic permutation*. For $s\in\mathbb{N}$, consider any application $y:[s]_0\mapsto\mathbb{Z}$ with $y(0)=0$. The s-cyclical permutations of y are the s applications $\theta_{s,q}(y)$, for $q\in[s]_0$ given by

$$\theta_{s,q}(y)=\begin{cases}y(j+q)-y(q) & j\leq s-q\\ y(j+q-s)+y(s)-y(q) & s-q\leq j\leq s.\end{cases}$$

This transformation can be described as cutting the original path at time q obtaining two paths and exchanging them, making the new path to start at zero. The next path transformation, introduced by Vervaat in [Ver79], is used to code random trees from discrete time EI processes.

Definition 1.2. *The* discrete Vervaat transform *of* $y : [s]_0 \mapsto \mathbb{Z}$*, denoted by* $V(y)$*, is the* i^**th cyclic shift of* y*, where* $i^* = \min\{i \in [s] : y(i) = \min_{j\in[s]} y(j)\}$ *is the first time* y *reaches its minimum.*

See Figure 3.3 for an example in the continuous setting.

Now we construct a tree with law $\mathbb{P}_{\mathbf{S}}$. First, obtain the child sequence $\mathbf{c}(\mathbf{S}) := \mathbf{c} = (c_1, \ldots, c_s)$, a vector with N_0 zeros, N_1 ones, and so on. Then, obtain the tree as follows:

1. Let $\pi = (\pi(1), \ldots, \pi(s))$ be a uniform random permutation on $[s] := \{1, \ldots, s\}$.
2. Define $W^b(j) = \sum_1^j (c \circ \pi(j) - 1)$ for $j \in [s]$, with $W^b(0) = 1$.
3. Construct $W := V(W^b)$.

In [BM14b], the authors indicate that W codes the depth-first walk of a tree with law $\mathbb{P}_{\mathbf{S}}$.

1.2.4 Profile and Lamperti transform

For some degree sequence $\mathbf{S_n}$, consider a tree τ_n with law $\mathbb{P}_{\mathbf{S_n}}$. The number of individuals of τ_n will be denoted by s_n. For any $j \geq 0$, let $C_{s_n}(j)$ be the total number of vertices up to generation j of τ_n. The process C_{s_n} is called the *cumulative profile* or *cumulative population profile*. If $Z_{s_n}(j)$ denotes the total number of vertices *at* generation j, then we call Z_{s_n} the *population profile* or simply the *profile* of τ_n.

Label the tree using the breadth-first order. Then, the number of individuals up to generation $j+1$ is the number of individuals up to generation j, plus all the children from the individuals in the jth generation, that is

$$C_{s_n}(j+1) = C_{s_n}(j) + c(C_{s_n}(j-1)+1) + \cdots + c(C_{s_n}(j)).$$

This leads us to

$$C_{s_n}(j+1) = 1 + \sum_{k=1}^{C_{s_n}(j)} c(k).$$

Note that if $W_{s_n}(j) = \sum_{k=1}^{j}(c(j)-1)$, then

$$Z_{s_n}(j+1) = C_{s_n}(j+1) - C_{s_n}(j) = 1 + \sum_{k=1}^{C_{s_n}(j)} (c(k)-1) = W_{s_n} \circ C_{s_n}(j). \tag{1.3}$$

First we rescale W_{s_n}, call this X^n, so that X^n converges to some process X. Then we show that a.s., any subsequential limit of $(C_{s_n}, n \in \mathbb{N})$ converges to a process C which, if not the zero function, is positive on $(0, \infty)$, and a solution of

$$Z(t) = X \circ C(t), \tag{1.4}$$

for $t \geq 0$, where $C(t) = \int_0^t Z(s)ds$. We call the above process Z the *Lamperti transform* of X, and Z_{s_n} the *discrete Lamperti transform* of W_{s_n} (see [Lam67a]). To describe the limit of the profile in Equation (1.4) we need to describe the limit X of the rescaled BFW's.

1.2.5 Continuous time EI processes

As in the unidimensional case, it turns out that the limit X can be described as the continuous Vervaat transform of a continuous time process X^b with exchangeable increments (EI process, or also known as interchangeable increment process).

A continuous time càdlàg $\mathbb{R}$-valued stochastic process X has *exchangeable increments* if for every n, the increments $(X_{i/n} - X_{(i-1)/n}, i \in [n])$ are exchangeable, in the sense of Definition 1.1. From Kallenberg's representation [Kal73], any EI on $[0,1]$ has the form

$$X^b(t) = \alpha t + \sigma b(t) + \sum_1^\infty \beta_j \left(\mathbf{1}\left(U_j \leq t\right) - t\right) \qquad t \in [0,1], \tag{1.5}$$

where b is a Brownian bridge on $[0,1]$, $(U_j, j \geq 1)$ are independent of b and i.i.d. with uniform law on $[0,1]$, and with independent three parameters $\alpha \in \mathbb{R}$, $\sigma \geq 0$, and $(\beta_j, j \geq 1)$ satisfying $\beta_1 \geq \beta_2 \geq \cdots > 0$ and $\sum \beta_j^2 < \infty$. We say that the process X^b has canonical parameters (α, σ, β). In our case, the EI process will have deterministic canonical parameters $(0, \sigma, \beta)$.

The "excursion-type" process X, is obtained from X^b as follows. Denote by $\{t\}$ the fractional part of a real number t. Let $D[0,1]$ be the space of real-valued càdlàg functions w with domain $[0,1]$ starting at zero (with an analogous definition for $D[0,\infty)$). Let $D' \subset D[0,1]$ be the subset of functions w such that $w(0) = w(1) = 0$, and w hits its infimum in a unique time and continuously. Define for every $u \in [0,1]$ the transformation $\theta_u : D' \mapsto D'$ by

$$\theta_u(w)(t) = w(\{t+u\}) - w(u).$$

Definition 1.3. *The* continuous Vervaat transform *of a càdlàg function $w \in D'$ is defined as $V(w) = \theta_\rho(w)$, where ρ is the unique time w hits its infimum.*

Figure 3.3 shows an example.

1.2.6 Statement of the results

Our main theorem is the following.

Theorem 1.2. *Consider a sequence $(\mathbf{s_n}, n \geq 1)$ of degree sequences $\mathbf{s_n} = (N_i^n, i \geq 0)$, satisfying*

1. $s_n \to \infty$.
2. *There exists a sequence of positive numbers $(b_{s_n}, n \geq 1)$ going to infinity, and $M \in \mathbb{N} \cup \{+\infty\}$, such that*
$$\left(\frac{1}{b_{s_n}^2} \sum (j-1)^2 N_j^n, \frac{\tilde{c}(1)}{b_{s_n}}, \frac{\tilde{c}(2)}{b_{s_n}}, \dots\right) \to \left(\sigma^2 + \sum_1^M \beta_j^2, \beta_1, \beta_2, \dots\right)$$
for some $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_M > 0$, and $\beta_{M+j} = 0$ for every j, and such that $\sum_1^M \beta_j^2 < \infty$ and $\sigma^2 \in [0,\infty)$, and where $(\tilde{c}(i), i \geq 1)$ is the child sequence in decreasing order.
3. *Either $\sigma > 0$ or $\sum \beta_j = \infty$.*
4. $s_n / b_{s_n} \to \infty$.

Define the rescaled processes

$$X^n = \left(\frac{1}{b_{s_n}} W_{s_n} \left(\lfloor s_n t \rfloor \right), t \in [0,1] \right),$$

$$C^n = \left(\frac{1}{s_n} C_{s_n} \left(\left\lfloor \frac{s_n}{b_{s_n}} t \right\rfloor \right), t \geq 0 \right) \quad \textit{and} \quad Z^n = \left(\frac{1}{b_{s_n}} Z_{s_n} \left(\left\lfloor \frac{s_n}{b_{s_n}} t \right\rfloor \right), t \geq 0 \right).$$

Then, we have the convergence

$$X^n \xrightarrow{d} X,$$

under the Skorohod topology $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$*. The limit* X *is the Vervaat transform of an EI process with parameters* $(0, \sigma, \beta)$*. Furthermore, if*

$$\int_{1/2}^{1} \frac{1}{X_s} ds < \infty \qquad a.s., \tag{1.6}$$

then, we have the joint convergence

$$(X^n, C^n, Z^n) \xrightarrow{d} (X, C, Z) \tag{1.7}$$

under the product Skorohod topology $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)^3$*. The limit* C *has an inverse* I *given by*

$$I(t) = \int_0^t \frac{1}{X(s)} ds, \quad t \in [0,1],$$

and is the unique solution to

$$C(t) = \int_0^t X \circ C(s)\, ds, \tag{1.8}$$

which is strictly increasing on $[0, I(1)]$ *if it is not the constant function. Finally* $Z = X \circ C$*.*

As a corollary we obtain the convergence of the profile of TGDS with a *finite variance condition*.

Corollary 1.1. *Consider a sequence* $(\mathbf{s_n}, n \geq 1)$ *of degree sequences, such that*

1. $s_n \to \infty$.
2. *For*

$$\sigma_n^2 := \sum_{j \geq 1} \frac{N^n(j)}{s_n - 1} j^2 - 1,$$

   we have $\sigma_n^2 \to \sigma^2$ *for some* $\sigma^2 \in (0, \infty)$.

3. *The maximum degree satisfies*

$$\Delta_n := \max\{j : N^n(j) > 0\} = o(s_n^{1/2}).$$

Define the rescaled processes X^n*,* C^n *and* Z^n *as in the previous theorem, with* $b_{s_n} = s_n^{1/2}$*. Then, we have the joint convergence*

$$(X^n, C^n, Z^n) \xrightarrow{d} (\sigma \mathbf{e}, C, Z) \tag{1.9}$$

in the space $\mathscr{C}([0,1], \mathbb{R}_+^3)$*, with* C *and* Z *defined as in the previous theorem, but driven by* $\sigma \mathbf{e}$*.*

We prove our main theorem in several steps. The convergence of the rescaled BFW's is given in Proposition 3.6. The correct rescaling of the profile and cumulative profile is given in 3.3.1. The convergence of subsequential limits of the cummulative profile is given in Subsection 3.3.2, and of the profile in Subsection 3.3.3.

To prove this theorem, the technical difficulty is to show that any subsequential limit of the cumulative Lamperti transform in (1.3), if converges to the non-zero function, then converges to the *unique solution of* (1.4) *which is positive on* $(0,\infty)$. When it converges to the non-zero function, this is equivalent to prove that such subsequential limits are not zero on $(0,\Lambda)$, for some $\Lambda \in (0,\infty]$. This can be interpreted as trees with law $\mathbb{P}_{\mathbf{s_n}}$ not having a thin base. This is proved in Section 3.4. Kersting gave another interpretation for those trees having few individuals at the base (see Lemma 1.4 below).

In the literature, there are related results on convergence of trees with law $\mathbb{P}_{\mathbf{s_n}}$ (when rescaled in an appropriate way). The paper [BM14b] proves the convergence of such trees to the CRT, when the degree sequence satisfies some *finite variance condition*. Even that the whole tree converges, from such results it cannot be proved the convergence of the profile. This is true since the profile is not a continuous functional in the Skorohod space. Therefore, we expand the analysis on convergence for such models, as well as validate that the profile of such trees satisfying the finite variance condition, converges to the Lamperti transform of the Brownian excursion. Another very related model are the **p**-trees, whose continuum versions are related also by EI processes with some restrictions of the parameters. The paper [AMP04] proves the convergence of their profile. Here, we prove that the profile of random trees converges in a more general setting than such papers. Recently, the result of [BM14b] was extended to forests in [Lei17], under a similar finite variance assumption. It is left as an **open problem** to prove the convergence of the profile of uniform random forests with a given degree sequence.

1.2.7 Application to CGW trees

We prove that the law of a GW tree conditioned by its size, is a mixture of laws of TGDS (Lemma 3.10). Hence, as a particular case of Theorem 1.2, we can prove a conjecture due to Aldous [Ald91b], about the convergence of the profile for CGW trees with finite variance offspring. Indeed, our result also applies to the case where the offspring distribution is in the domain of attraction of a stable law (in DA). Let us give some definitions.

Galton-Watson Trees

We briefly describe the definition of a GW tree, as well as its law on the set of all rooted plane trees. The formal statements can be found in [Nev86, LG05], and we follow the definitions of [AP98].

Let $\mathbb{T}^{(\infty)}$ the set of (possibly infinite) plane trees. For any $k \in \mathbb{N}$, let $\mathbb{T}^{(k)}$ be the set of plane trees with height (latest generation having individuals) at most k. Consider the restriction map $r_k : \mathbb{T}^{(\infty)} \mapsto \mathbb{T}^{(k)}$, where $r_k t$ is the subtree of $t \in \mathbb{T}^{(\infty)}$, formed by all the vertices up to generation k. A tree $t \in \mathbb{T}^{(\infty)}$ is identified by the sequence $(r_k t, k \geq 0)$.

A *random family tree* τ is a random element of $\mathbb{T}^{(\infty)}$, specified by the sequence $(r_k \tau, k \geq 0)$, where each $r_k \tau$ is a random variable taking values on $\mathbb{T}^{(k)}$, and $r_k \tau = r_k(r_{k+1}\tau)$ for every k.

Let μ be a distribution on $\mathbb{Z}_+$. A GW tree τ with offspring distribution μ, satisfies

$$\mathbb{P}(r_k \tau = t) = \prod_{v \in r_{k-1} t} \mu(c(v)) \qquad \forall t \in \mathbb{T}^{(k)}, k \geq 1,$$

where the product is taken over all vertices v of t up to generation $k-1$, and $c(v)$ is the number of children of v.

Convergence of the profile of CGW trees with offspring distribution in the domain of attraction of a stable law

Let μ be a *critical* and *aperiodic* distribution on $\mathbb{Z}_+$, that is, has mean one and the greatest common divisor of all n such that $\mu_n > 0$ is one. We exclude the case $\mu(1) = 1$. We are interested in the following distributions.

Definition 1.4. *Consider a sequence of i.i.d. random variables* $(\xi_n, n \in \mathbb{N})$. *If for some positive numbers* $b_n = n^{1/\alpha} L(n)$, *with* L *a positive function satisfying* $\lim_{x\to\infty} L(tx)/L(x) = 1$ *for every positive* t *and* $\alpha \in (1,2]$, *we have*

$$\frac{\sum_1^n \xi_k - n}{b_n} \to S_\alpha,$$

for some non-degenerate limit, then we say μ belongs to the domain of attraction of a stable law. *For simplicity, we simply say* $\mu \in DA(\alpha)$.

A well-known fact, is that the convergence of a rescaled sum of i.i.d. variables implies that the rescaling $(b_n, n \in \mathbb{N})$ should be of the form mentioned above (see [BGT89]).

In the next proposition we work with a CGW(n) tree whose rescaled BFW, cumulative profile and profile are denoted by X^n, C^n and Z^n respectively, and are defined as in Theorem 1.2, but replacing s_n with n and b_{s_n} with b_n.

Proposition 1.2. *Assume* μ *is a critical, aperiodic law in* $DA(\alpha)$, *for some* $\alpha \in (1,2)$. *Then, for a CGW*(n) *tree with offspring distribution* μ *we have the joint convergence*

$$(X^n, C^n, Z^n) \xrightarrow{d} (X, C, Z)$$

under the product Skorohod topology $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)^3$, *with* C *and* Z *defined as in Theorem 1.2, but driven by the stable excursion* X.

We prove this in Proposition 3.10. The proof of this result was given by Drmota and Gittenberger [DG97] in the finite variance case, and by Kersting [Ker11] in the stable case. Note that our main result, Theorem 1.2, is more general since rather than trees constructed from independent individuals, we allow dependencies between them.

1.2.8 Another proof for the convergence of the profile of CGW in DA

As stated before, the main difficulty to prove Theorem 1.2 is to prove that all subsequential limits of the cumulative profile are not thin at the base. Kersting deals with this issue in the case of CGW trees.

Lemma 1.4 (Lemma 9 of [Ker11]). *Let* C^n *be the rescaled cumulative Lamperti transform of a GW tree* τ *under* $\mathbb{P}(\cdot \,|\, |\tau| = n)$, *having offspring distribution* μ. *If* μ *is critical, aperiodic and in* $DA(\alpha)$ *for* $\alpha \in (1,2]$, *then for every* $\lambda > 0$

$$\lim_{\varepsilon \downarrow 0} \overline{\lim_{n}} \, \mathbb{P}(C^n(\lambda) \le \varepsilon \,|\, |\tau| = n) = 0.$$

We give a different and simple proof of this Lemma in Section 3.6. The idea is based on a comparison between the first generations of a CGW tree and the *Q-process*. Since the latter is not zero on the limit, neither the former.

The Q-process is a Markov chain on $\mathbb{N}$ having transition probabilities

$$\mathbb{P}^{\uparrow}\left(\bar{Z}_n = j | \bar{Z}_0 = i\right) = \frac{j}{i} m^{-n} \mathbb{P}\left(\bar{Z}_n = j | \bar{Z}_0 = i\right) \qquad i, j \geq 1,$$

where $\bar{Z}$ is a GW with offspring distribution μ and mean $m \in (0, \infty)$. Define the *size-biased distribution* $\tilde{\mu}(x) := x\mu(x)$.

We define the *infinite size-biased tree*, as the tree with profile the Q-process. Its construction is:

1. The root of the tree is marked, and has offspring distribution $\tilde{\mu}$.
2. At generation $n \geq 1$, there is only one marked individual, which also has offspring law $\tilde{\mu}$.
3. At generation $n+1$, chose one children uniformly at random from the children of the marked individual, and mark it. Every other individual has offspring distribution according to μ.

The mentioned comparison to prove Lemma 1.4 is

$$\overline{\lim_n} \mathbb{P}(C^n(\lambda) \leq \varepsilon | |\tau| = n) \leq c_\varepsilon \overline{\lim_n} \mathbb{P}^{\uparrow}\left(C^n(\lambda) \leq \varepsilon\right), \tag{1.10}$$

for fixed $\varepsilon > 0$, where c_ε is a positive constant bounded from above as ε goes to zero (see Theorem 3.4 and page 96).

It will be proved that the Q-process has the same distribution as a *GW process with immigration*, GWI for short (see Equation 3.35). This is a model like the GW but at each generation arrives i.i.d. immigrants following some offspring law, independent of the actual population. More explicitly, the Q process has the same law as the profile of the GWI

$$Z_n(m+1) = 1 + W \circ C_n(m) + \tilde{W}(m+1) \qquad m \geq 0,$$

where W is a skip-free random walk with increments having law $(\mu(j+1), j \geq -1)$, and $\tilde{W}$ is an independent random walk having increments with law $\tilde{\mu}$. Finally, using the results about the convergence of GWI in [CPGUB13], we can prove that the rescaled Q-process converges to a positive function on $(0, \infty)$ (see Lemma 3.14). It follows that the right-hand side of Equation (1.10) converges to zero as $\varepsilon \downarrow 0$.

1.3 ON MULTITYPE RANDOM FORESTS WITH A GIVEN DEGREE SEQUENCE, THE TOTAL POPULATION OF BRANCHING FORESTS AND ENUMERATIONS OF MULTITYPE FORESTS

In Chapter 3 we prove the convergence of the profile of uniform trees with a given degree sequence, as explained in the previous section. The natural generalization of this model is the profile of uniform multitype forests with a given degree sequence (MFGDS). Analogously as in the unitype case, MFGDS are a more general model than conditioned multitype Galton-Watson (MGW) forests, since under an independence assumption, by mixing the former we obtain a MGW conditioned with the number of individuals by types. Because a construction of both random forests has not been developed, we first define them, generalizing several well-known results of random forests to the multidimensional case. In future work, we will analyze the convergence of the profile of MFGDS.

1.3.1 Preliminaries and motivation

In Chapter 4 we construct two types of constrained multitype random forests:

1. Uniform multitype forest from the set of forests with a given degree sequence.

2. Multitype Galton-Watson forests conditioned with the number of individuals by types.

Multitype random forests serve to model the genealogical evolution of random populations, whenever there are different types of individuals coexisting. Those forests have plenty applications, like in biology, demography, genetics, medicine, epidemics, and language theory (see [Har63, San71, Jag75, GP75, CKB$^+$, AJ97, All11]). Multitype random forests also give rise to nice probability theory, for example in the field of continuum random forests. When the progeny distribution has a finite variance, Miermont [Mie08] proved its convergence to Aldou's CRT (see also [BO18]). This convergence unveils natural questions, as well as poses open problems regarding the convergence of *multitype Lévy forests* and generalizations of the Ray-Knight theorems, proved in the unitype case in [LGLJ98, DLG02].

Conditioned MGW forests also provides us with a lot of theoretical applications. There are several ways to condition such a forest leading to a generalization of the so-called Kesten's infinite tree [Nak78], as well as the Q-process. Pénisson [P1́6] proved that critical MGW forests conditioned on a special proportion of its total progeny converges locally to a MGW forest under some moment condition. Minimal assumptions of this result were found in [ADG18]. Another conditionings (differently to non-extinction in the near future) were given in [P1́6], like the process reaching a positive threshold or a non-absorbing state. Conditioning on a lineal combination of the population sizes by types was also proved in [Ste18], to converge locally to the multitype Kesten's tree, giving also applications on random maps.

Another way to condition a forest, which is also an active field of research, is to fix its degree sequence. Several authors study conditioning a (unitype) forest to have a fixed degree sequence, as well as random graphs with a given degree sequence [MR95, MR98, VL05, CDS11, AB12, HM12, BM14b, Jos14, Lei17, Mar19]. As explained previously in Subsection 1.2.1, random graphs with a given degree sequence are important models for real-world networks. They can be used to model specific features such as graphs having degree sequence with power law tails, a property not present in the Erdős-Rényi random graph. To our knowledge, there is no construction of multitype random forests with a given degree sequence, so we present one, as well as a simulation algorithm (see Section 4.3 and Algorithm 7). Just as in the unitype case, this is based on discrete time EI processes, and a generalization of the Vervaat transformation [Ver79]. An important tool to generalize the latter is the result in [CL16], giving us the number of cyclical permutations of the EI processes leading to paths coding multitype forests.

Using also the result in [CL16], we generalize the Otter-Dwass formula [Ott49, Dwa69], obtaining the law of the total population by types in a MGW, under certain conditions. This is done in Section 4.5. The unitype case of Otter and Dwass says that, if $\#\tau_k$ is the size of a GW forest with k trees, having offspring distribution μ, then

$$\mathbb{P}(\#\tau_k = n) = \frac{k}{n}\mathbb{P}(X_n = n-k),$$

where X is a random walk with step law μ. In Subsection 1.2.1 we pointed out the relation between GW forests with a given size and certain combinatorial classes (plane, labeled and binary forests with a given size). We also exploit such connection for the same classes in the multitype case, using our generalization of the Otter-Dwass formula (see Subsection 4.5.1).

Finally, we extend Devroye's algorithm about the simulation of CGW trees in Subsection 4.6.5. Indeed, Devroye's algorithm generates a CGW tree from a tree uniformly chosen from the set of trees with a given degree sequence, this last one obtained from the offspring distribution. So, we use both of our constructions to simulate a MGW forest conditioned by its total population by types.

1.3.2 Coding of multitype forests and the Multivariate Cyclic Lemma

Before stating our results, we need some definitions. Recall form Section 1.2.2 the coding of forests of size n with k trees with downward skip-free chains, starting at zero and first reaching $-k$ at time n. We briefly recall such coding for the multitype case, introduced in [CL16].

Define $[n] = \{1,\dots,n\}$ and $[n]_0 = \{0,1,\dots,n\}$ for $n \in \mathbb{N}$. For a forest F, let $c_F : v(F) \mapsto [d]$ be an application from the set of vertices of F to $[d]$, such that the children of each vertex are ordered by color, that is, if $u_i, u_{i+1} \dots, u_{i+j} \in v(F)$ have the same parent, then $c_F(u_i) \le c_F(u_{i+1}) \le \cdots \le c_F(u_{i+j})$. The couple (F, c_F) is a *d-multitype forest*. A *subtree of type i* of (F, c_F), denoted by $T^{(i)}$, is a maximal connected subgraph of (F, c_F) whose all vertices are of type i. Subtrees of type i are ranked according to the order of their roots, and with this ordering, we define the *subforest of type i* of (F, c_F) as $F^{(i)} = \{T_1^{(i)},\dots,T_k^{(i)},\dots\}$ For $u \in v(F)$, denote by $p_i(u)$ the number of children of type i of u. Let $n_i \ge 0$ be the number of vertices in the subforest $F^{(i)}$ of (F, c_F). The coding of the forest is the d-dimensional chain $x^{(i)} = (x^{i,1},\dots,x^{i,d}) \in \mathbb{Z}^d$ with length $n_i \in \mathbb{N}$, defined for $0 \le n \le n_i - 1$ by

$$x^{i,j}_{n+1} - x^{i,j}_{n} = p_j(u^{(i)}_{n+1}) - \mathbf{1}_{i=j} \qquad i, j \in [d]. \tag{1.11}$$

We set $x_0^{(i)} = 0$. The set $(u_n^{(i)}; n \ge 1)$ is the labeling of the subforest $F^{(i)}$ in its own breadth-first order.

To generalize the Vervaat transform, we define the type of cyclical permutations that we apply to our chains. For $n \in \mathbb{N}$, consider any application $y : [n]_0 \mapsto \mathbb{Z}^d$ with $y(0) = 0$. The n-cyclical permutations of y are the n applications $\theta_{n,q}(y)$, for $q \in [n-1]_0$ given by

$$\theta_{q,n}(y) = \begin{cases} y(j+q) - y(q) & j \le n-q \\ y(j+q-n) + y(n) - y(q) & n-q \le j \le n. \end{cases}$$

We say that the path $y : \mathbb{N} \mapsto \mathbb{Z}$ is a *downward skip-free chain*, if $y_{k+1} - y_k \in \mathbb{Z}_+ \cup \{-1\}$. The possible paths that a coding of multitype forest can take values are the following.

Definition 1.5. *Fix any $\mathbf{n} = (n_1,\dots,n_d) \in \mathbb{Z}_+^d$, and define S_d as the set of $[\mathbb{Z}^d]^d$-valued sequences $x = (x^{(1)},\dots,x^{(d)})$ such that for all $i \in [d]$, $x^{(i)} = (x^{i,1},\dots,x^{i,d})$ is a $\mathbb{Z}^d$-valued sequence starting at zero of length n_i, and where $x^{i,j} = (x_k^{i,j}, k \in [n_i]_0)$ is non-decreasing when $i \ne j$, and a downward skip-free chain when $i = j$.*

The $\mathbf{n}$-cyclical permutations of $x \in S_d$ are given by

$$\theta_{\mathbf{q},\mathbf{n}}(x) := (\theta_{q_1,n_1}(x^{(1)}),\dots,\theta_{q_d,n_d}(x^{(d)})) \qquad \forall\, \mathbf{q} = (q_1,\dots,q_d) \text{ such that } 0 \le \mathbf{q} \le \mathbf{n} - 1_d,$$

with $1_d = (1,\dots,1)$ of length d. Each sequence $\theta_{\mathbf{q},\mathbf{n}}(x)$ will be called a cyclical permutation of x.

For $\mathbf{m}, \mathbf{n} \in \mathbb{Z}_+^d$, write $\mathbf{m} < \mathbf{n}$ if $\mathbf{m} \le \mathbf{n}$ (the inequality understood component-wise) and if there exists i such that $m_i < n_i$. Sequences $x \in S_d$ will be denoted by $x = (x_k^{i,j}, k \in [n_i]_0, i, j \in [d])$, and the vector

$\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}_+^d$, is called the length of x. Fix any such x of length $\mathbf{n}$, and $\mathbf{r} = (r_1, \ldots, r_d) \in \mathbb{Z}_+^d$ with $\sum r_i > 0$. We say that the system $(\mathbf{r}, x)$ admits a solution if there exists $\mathbf{m} \leq \mathbf{n}$ such that

$$r_j + \sum_{i=1}^{d} x^{i,j}(m_i) = 0 \qquad \forall\, j \in [d]. \tag{1.12}$$

If there is no smaller solution $\mathbf{m} < \mathbf{n}$ for the system $(\mathbf{r}, \theta_{\mathbf{q},\mathbf{n}}(x))$, then we call $\theta_{\mathbf{q},\mathbf{n}}(x)$ a *good cyclical permutation*. It is proved in [CL16] that only such good cyclical permutations code multitype forests, and the next lemma tells us how many there are.

Lemma 1.5 (Multivariate Cyclic Lemma [CL16])**.** *Let $x \in S_d$ with $x^{i,i}(n_i) \neq 0$ for every $i \in [d]$. Consider the system $(\mathbf{r}, x)$ with solution $\mathbf{n}$ as above. Then, the number of good cyclical permutations of x is $det((-x^{i,j}(n_i))_{i,j\in[d]})$.*

Since in most of the cases, we fix the number of roots or number of individuals of each type, we need the following definition.

Definition 1.6 (Root-type and individuals-type)**.** *We say a multitype plane forest with $d \in \mathbb{N}$ types has* root-type $\mathbf{r} = (r_1, \ldots, r_d) \in \mathbb{N}^d$*, if it has r_i roots of type i for $i \in [d]$, with $\mathbf{r} > 0$. Also, it has* individuals-type $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{N}^d$ *if it has n_i individuals of type i, for $i \in [d]$.*

1.3.3 Multitype Galton-Watson Forests

A *multitype Galton-Watson* forest in d-types, is a branching forest, where each individual has a type $i \in [d]$, and has children independently of the other individuals of its generation or below, according to a law ν_i on $\mathbb{Z}_+^d$. The *progeny distribution* of the forest is $\nu = (\nu_1, \ldots, \nu_d)$. The formal definition is the following.

Definition 1.7. *A* multitype Galton-Watson process *is a Markov chain $Z = \left((Z_n^{(1)}, \ldots, Z_n^{(d)}); n \geq 0\right)$ on $\mathbb{Z}_+^d$, with transition function*

$$\mathbb{P}(Z_{n+1} = (k_1, \ldots, k_d)\,|Z_n = (r_1, \ldots, r_d)) = \nu_1^{*r_1} * \cdots * \nu_d^{*r_d}(k_1, \ldots, k_d),$$

where ν is the progeny distribution, and ν_i^{*j} is the jth iteration of the convolution product of ν_i by itself, with $\nu_i^{*0} = \delta_0$.

For $\mathbf{r} \in \mathbb{Z}_+^d$, the probability measure $\mathbb{P}_\mathbf{r}$ is the law $\mathbb{P}(\cdot|Z_0 = \mathbf{r})$. As in Theorem 1.2 in [CL16], we consider MGW trees satisfying the following. For $i, j \in [d]$, let $m_{i,j} = \sum_{z\in\mathbb{Z}_+^d} z_j \nu_i(z)$ be the mean number of children type j given by an individual type i, and set $M = (m_{i,j})_{i,j}$ as the mean matrix of the MGW tree. Whenever M is irreducible, by the Perron-Frobenius Theorem (see [AN04, Chapter V.2]), it has a unique eigenvalue which is simple, positive and with maximal modulus. We say in such case that the MGW tree is *irreducible*. If the unique eigenvalue equals one (is less than one), then we say the tree is *critical (subcritical)*. The tree is *non-degenerate* if individuals have exactly one offspring with probability different from one.

1.3.4 Statement of the results

A *multitype degree sequence* $\mathbf{S} = (\mathbf{S}_{i,j}, i, j \in [d])$ is a sequence of sequences of non-negative integers $\mathbf{S}_{i,j} = (N_{i,j}(k); k \in [m_{i,j}]_0)$, where $m_{i,j} \in \mathbb{N}$, satisfying:

1. $n_i = \sum_k N_{i,j}(k)$ for every $i \in [d]$,
2. $n_j = r_j + \sum_k kN_{1,j}(k) + \cdots + \sum_k kN_{d,j}(k)$, for every $j \in [d]$,
3. $det(-k_{i,j}) > 0$ with $k_{i,j} := \sum kN_{i,j}(k) - n_i \mathbf{1}_{i=j}$ and $k_{i,i} < 0$ for every $i \in [d]$.

The value $N_{i,j}(k)$ represents the number of individuals of type i with k children of type j, so n_i represents the total number of type i individuals. Clearly, the total number of vertices is $s := n_1 + \cdots + n_d = \sum_k N_{1,j}(k) + \cdots + \sum_k N_{d,j}(k)$ for $j \in [d]$. The last condition is required to obtain the degree sequence of a forest (see page 109). An example for $d = 2$ is given in Table 1.1.

$\mathbf{S}_{1,1} = (N_{1,1}(0), \ldots, N_{1,1}(m_{1,1}))$	$\mathbf{S}_{1,2} = (N_{1,2}(0), \ldots, N_{1,2}(m_{1,2}))$	$n_1 = \sum_k N_{1,j}(k)$
$\mathbf{S}_{2,1} = (N_{2,1}(0), \ldots, N_{2,1}(m_{2,1}))$	$\mathbf{S}_{2,2} = (N_{2,2}(0), \ldots, N_{2,2}(m_{2,2}))$	$n_2 = \sum_k N_{2,j}(k)$
$n_1 = r_1 + \sum_k kN_{1,1}(k) + kN_{2,1}(k)$	$n_2 = r_2 + \sum_k kN_{1,2}(k) + kN_{2,2}(k)$	$n_1 + n_2 = s$

Table 1.1: Relations on a 2-type degree sequence.

The canonical *child sequence* $\mathbf{c} = (\mathbf{c}_{i,j}, i, j \in [d])$ is constructed from the degree sequence as follows: let $\mathbf{c}_{i,j}$ be a sequence whose first $N_{i,j}(0)$ entries are zeros, the next $N_{i,j}(1)$ entries are ones, and so on. Let $\sigma_{i,j}$ be any permutation on $[n_i]$, and construct $w^b = \{w^b_{i,j}; i, j \in [d]\}$, where

$$w^b_{i,j}(k) = \sum_{l=1}^{k} \left(\mathbf{c}_{i,j} \circ \sigma_{i,j}(l) - \mathbf{1}_{i=j} \right), \quad k \in [n_i].$$

Remark 1.1. *Note that $k_{i,j} = w^b_{i,j}(n_i)$ does not depend on the permutation, so it is deterministic. Also, note that the system of equations $(\mathbf{r}, w^b)$ admits $\mathbf{n}$ as a solution, since by definition*

$$r_j + \sum_{i=1}^{d} w^b_{i,j}(n_i) = r_j - n_j + \sum_{i=1}^{d} \sum kN_{i,j}(k) = 0 \qquad \forall\, j \in [d].$$

From the Multivariate Cyclic Lemma 1.5, we know that $det(k_{i,j})$ is the number of good cyclical permutations of w^b. From such set we define a Vervaat-type transformation of w^b. Such transformation is given by choosing uniformly at random a good-cyclical permutation from all the good-cyclical permutations. After that, the algorithm is similar to the unidimensional case.

Definition 1.8 (Multidimensional Vervaat Transform)**.** *For any w^b as constructed above and any $u \in [\det(k_{i,j})]$, define $V(w^b, u)$ as follows: enumerate the $\det(k_{i,j})$ good cyclical permutations of w^b, using the lexicographic order on the set of $\mathbf{q}$ such that $\theta_{\mathbf{q},\mathbf{n}}(w^b)$ codes a forest; then, let $V(w^b, u)$ be the u-th good cyclical permutation.*

Define $\mathbb{F}_{\mathbf{S},\mathbf{r}}$ as the set of multitype plane forests with degree sequence $\mathbf{S}$, having root-type $\mathbf{r}$ and individuals-type $\mathbf{n}$. Our construction of MFGDS is the following (the proof is given on page 109).

Theorem 1.3 (Uniform multitype forest with a given degree sequence)**.** *Fix the degree sequence* $\mathbf{S}$ *of a multitype forest having root-type* $\mathbf{r}$ *and individuals-type* $\mathbf{n}$*. Let* $\mathbf{W}$ *be the BFW coding a forest (see* (1.11)*) taken uniformly at random from* $\mathbb{F}_{\mathbf{S},\mathbf{r}}$*. Let* $\pi = (\pi_{i,j}, i, j \in [d])$ *be independent random permutations, where* $\pi_{i,j}$ *takes values on* $[n_i]$*, and let* U *be an independent uniform variable on* $[\det(k_{i,j})]$*. Define the processes* $\mathbf{W}^b = (W^b_{i,j}, i, j \in [d])$ *as*

$$W^b_{i,j}(k) = \sum_{l=1}^{k} (\mathbf{c}_{i,j} \circ \pi_{i,j}(l) - \mathbf{1}_{i=j}), \quad k \in [n_i],$$

where $\mathbf{c} = \big(c_{i,j}, i, j \in [d]\big)$ *is the child sequence of* $\mathbf{S}$*. Then*

$$V(\mathbf{W}^b, U) \stackrel{d}{=} \mathbf{W}.$$

From the proof, we obtain the number of multitype forests with a given degree sequence $\mathbf{S}$:

$$|\mathbb{F}_{\mathbf{S},\mathbf{r}}| = \frac{\det(-k_{i,j})}{\prod n_i} \prod\prod \binom{n_i}{\mathbf{S}_{i,j}}.$$

MGW forests conditioned by types

Before turning to the joint law of the number of individuals of type $i \in [d]$, of a MGW forest, we prove that the latter model is a mixture of MFGDS in Section 4.4. This justifies the importance of the latter model.

Let $S^{i,j}$ be a random walk with increments having law the jth marginal of ν_i. Our hypotheses are the following:

H1 For every $i \in [d]$, the law ν_i has independent components, with

$$\nu_i^{*n_i}(k_1, \dots, k_d) = \prod_j \mathbb{P}\big(S^{i,j}_{n_i} = k_j\big) \qquad k_1, \dots, k_d \in \mathbb{N}_0.$$

H2 For every $i, j \in [d]$, with $i \neq j$

$$\mathbb{E}\left(S^{i,j}_{n_i}; \sum_{l \in [d]} S^{l,j}_{n_l} = n_j - r_j\right) = \frac{n_i(n_j - r_j)}{n} \mathbb{P}\left(\sum_{l \in [d]} S^{l,j}_{n_l} = n_j - r_j\right).$$

Using those hypotheses, we obtain the following result (see page 112), which is a generalization of the Otter-Dwass formula.

Theorem 1.4. *Consider an irreducible, non-degenerate and (sub)critical MGW forest, and let* $n_i > 0$ *for every* i*. Suppose that* **H1** *and* **H2** *and are also satisfied. If* O_i *is the number of type* i *individuals, then*

$$\mathbb{P}_{\mathbf{r}}(O_i = n_i, i \in [d]) = \frac{r}{n} \prod_{i=1}^{d} \mathbb{P}\left(\sum_{l \in [d]} S^{l,i}_{n_l} = n_i - r_i\right),$$

where $r = r_1 + \dots + r_d$ *and* $n = n_1 + \dots + n_d$*, and* $r_i < n_i$*.*

Remark 1.2. *The assumption $n_i > 0$ for every i makes the proof easier, but we think this hypothesis can be dropped as in [CL16].*

Remark 1.3. *After the proof of the theorem, we obtain the case when $n_i = r_i$ for some i's. Since Theorem 1.4 has a different formula on such case, the law of $\sum_{i\in[d]} O_i$ (computed on Corolary 4.1) does not have a nice expression.*

For the next results denote by $\mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}}$, $\mathbb{F}^{labeled}_{\mathbf{r},\mathbf{n}}$ and $\mathbb{F}^{binary}_{\mathbf{r},\mathbf{n}}$, the set of d-type plane, labeled and binary forests having root-type $\mathbf{r}$ and individuals-type $\mathbf{n}$, with $r_i < n_i$ for every i and $r > 0$. Our labeled multitype forests have labels on $[n]$, that is, for $F \in \mathbb{F}^{labeled}_{\mathbf{r},\mathbf{n}}$, each individual v has a unique label $i \in [n]$ and a type $c_F(v) \in [d]$; also, F has fixed *root set* $[r]$, that is, the r_1 type 1 roots have labels on $\{1,\dots,r_1\}$, the r_2 type 2 roots have labels on $\{r_1+1,\dots,r_1+r_2\}$, and so on. Using Theorem 1.4, we give in Subsection 4.5.1 three examples of distributions were the law of a MGW forest conditioned by the number of individuals of each type can be computed. This generalizes the constructions given in [Pit98].

Proposition 1.3. *For fixed $p \in (0,1)$, let $\mathscr{G}_{\mathbf{r},p}$ be a d-type GW forest with root-type $\mathbf{r}$, having geometric offspring distribution with parameter p independently for each type, that is, $\nu_i(k_1,\dots,k_d) = \prod_i p(1-p)^{k_i}$ for $k_i \geq 0$. Let $\#_i\mathscr{G}_{\mathbf{r},p}$ be the number of type i individuals in $\mathscr{G}_{\mathbf{r},p}$. Then*

$$\mathbb{P}(\mathscr{G}_{\mathbf{r},p} = F | \#_i\mathscr{G}_{\mathbf{r},p} = n_i, i \in [d]) = \frac{1}{\frac{r}{n}\prod_{i\in[d]}\binom{n+n_i-r_i-1}{n_i-r_i}} \qquad \forall F \in \mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}},$$

thus, such conditioned forest is uniform on $\mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}}$.

Proposition 1.4. *For $\mu \in \mathbb{R}^+$, let $\mathscr{P}_{\mathbf{r},\mu}$ be a d-type GW forest with root-type $\mathbf{r}$, having Poisson offspring distribution of parameter μ independently for each type, that is, $\nu_i(k_1,\dots,k_d) = \prod_i e^{-\mu}\mu^{k_i}/k_i!$ for $k_i \geq 0$. Let $\#_i\mathscr{P}_{\mathbf{r},p}$ be the number of type i individuals in $\mathscr{P}_{\mathbf{r},p}$. If $\mathscr{P}^*_{\mathbf{r},\mathbf{n}}$ is $\mathscr{P}_{\mathbf{r},\mathbf{n}}$ relabeled by d uniform random permutations, one for each type, then*

$$\mathbb{P}\big(\mathscr{P}^*_{\mathbf{r},p} = F | \#_i\mathscr{P}_{\mathbf{r},p} = n_i, i \in [d]\big) = \frac{1}{\frac{r}{n}n^{n-r}} \qquad \forall F \in \mathbb{F}^{labeled}_{\mathbf{r},\mathbf{n}},$$

thus, such conditioned forest is uniform on $\mathbb{F}^{labeled}_{\mathbf{r},\mathbf{n}}$.

Proposition 1.5. *For $0 < p < 1$, let $\mathscr{B}_{\mathbf{r},p}$ be a d-type GW forest with root-type $\mathbf{r}$, having Bernoulli offspring distribution with parameter p, for each vertex independently of the type, that is, $\nu_i(k_1,\dots,k_d) = \prod p^{k_i}(1-p)^{1-k_i}$ with $k_i \in \{0,1\}$. Assume $n_i - r_i$ is an even number for every $i \in [d]$. Since any vertex v has zero or two children with probability p or $1-p$ respectively, then $\nu_i(c_1(v),\dots,c_d(v)) = \prod p^{c_i(v)/2}(1-p)^{1-c_i(v)/2}$. Let $\#_i\mathscr{B}_{\mathbf{r},p}$ be the number of type i individuals in $\mathscr{B}_{\mathbf{r},p}$. Then*

$$\mathbb{P}(\mathscr{B}_{\mathbf{r},p} = F | \#_i\mathscr{B}_{\mathbf{r},p} = n_i, i \in [d]) = \frac{1}{\frac{r}{n}\prod\binom{n}{(n_i-r_i)/2}} \qquad \forall F \in \mathbb{F}^{binary}_{\mathbf{r},\mathbf{n}},$$

thus, such conditioned forest is uniform on $\mathbb{F}^{binary}_{\mathbf{r},\mathbf{n}}$.

As a simple consequence of our results, we obtain the following enumerations.

Lemma 1.6. *The number of d-type plane, labeled, and binary forest, with root-type* $\mathbf{r}$ *and individuals-type* $\mathbf{n}$ *is given respectively by*

$$\frac{r}{n}\prod_{i\in[d]}\binom{n+n_i-r_i-1}{n_i-r_i}, \quad \frac{r}{n}n^{n-r} \quad \text{and} \quad \frac{r}{n}\prod\binom{n}{(n_i-r_i)/2}.$$

Finally, we give an algorithm to simulate MGW processes conditioned by its types. This is done using the following proposition and the well-known Accept-Reject method (see Algorithm 8).

Proposition 1.6. *Let W be the breadth-first walk of a MCGW($n_1,\ldots,n_d$) forest satisfying the Hypotheses of Theorem 1.4, having offspring distribution ν, and root-type $\mathbf{r}$, with $0<r_i<n_i$ for every i. Generate independent multinomial vectors $\mathbf{S}_{i,j}=(N_{i,j}(0),N_{i,j}(1),\ldots)$ with parameters $(n_i,\nu_{i,j}(0),\nu_{i,j}(1),\ldots)$, and stop the first time that $r_j+\sum_i\sum_k kN_{i,j}(k)=n_j$ for every j. Denote by $\mathbf{S}$ the multitype degree sequence obtained, and let $V(\mathbf{W}^b,U)$ be the breadth-first walk generated by Theorem 1.3 using the degree sequence $\mathbf{S}$. Then,*

$$\mathbb{P}\Big(V(\mathbf{W}^b,U)=w\Big)=\frac{1}{\frac{n}{r}\frac{det(k_{i,j})}{\prod n_i}}\mathbb{P}_{\mathbf{r}}\left(\mathscr{F}=F|\#_j\mathscr{F}=n_j,\forall\, j\right),$$

for every multitype forest F with root-type $\mathbf{r}$ and individuals-type $\mathbf{n}$, coded by w and with $k_{i,j}=\sum kn_{i,j}(k)-n_i\mathbf{1}_{i=j}$.

1.4 DINI DERIVATIVES FOR EXCHANGEABLE INCREMENT PROCESSES AND APPLICATIONS

1.4.1 Preliminaries and motivation

In Theorem 1.2 we stated our result about the convergence of the profile of TGDS. One of our hypotheses for the convergence, is that $\sigma>0$ or $\sum\beta_j=\infty$. As will be seen in Proposition 3.6, this hypothesis implies that the rescaled BFW's converge to the Vervaat transform of a continuous time EI process. Indeed, the facts needed are that the EI process has a unique infimum and that it is continuous there. In our particular case, that is, when the EI process has deterministic canonical parameters, the jumps are positive and $\sigma>0$ or $\sum\beta_j=\infty$, the proof of Lemma 6 of [Ber01] tells us that the process has a unique infimum where it is continuous.

Now, we review the known results of the previous facts, for *Lévy processes*. Recall that a Lévy process X is a càdlàg stochastic process, starting at zero and having stationary and independent increments. By the celebrated Lévy-Khintchine formula, its characteristic function is of the form

$$\mathbb{E}\Big(e^{i\theta X_t}\Big)=e^{-t\psi(\theta)} \qquad t\geq 0,\ \theta\in\mathbb{R}_+,$$

where

$$\psi(\theta)=-i\alpha\theta+\frac{1}{2}\sigma^2\theta^2+\int_{\mathbb{R}}\Big(1-e^{i\theta x}+i\theta x\mathbf{1}_{|x|\leq 1}\Big)\Pi(dx),$$

for $\alpha\in\mathbb{R},\ \sigma\geq 0$ and Π is a measure satisfying

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

We say that a process starting at zero is *upward regular* (*downward regular*), if it hits $(0,\infty)$ immediately (respectively $(-\infty,0)$). It is *regular* if it is both upward and downward regular. For Lévy processes the whole picture is clear:

1. If it is of infinite variation, then it is regular (Rogozin [Rog68]).

2. If it hits is infimum in a unique place, then it is continuous there iff it is regular (this is Millar's zero-one law [Mil77]).

This was the motivation for Chapter 5: obtain a proof for the convergence of the profile independent of Bertoin's result [Ber01], and prove the above two points for the most general EI processes. It turns out, that we were able to prove both for general EI processes: any infinite variation EI process is regular; and any EI process hitting its infimum only once, is continuous at its infimum iff it is regular. As a matter of fact, we were able to prove more.

1.4.2 Statement of the results

Our main result of Chapter 5 is the following.

Theorem 1.5. *Let X be an EI process of infinite variation. Then, for any fixed t almost surely,*

$$\limsup_{h\to 0\pm}\frac{X_{t+h}-X_t}{h}=\infty \quad \textit{and} \quad \liminf_{h\to 0\pm}\frac{X_{t+h}-X_t}{h}=-\infty \tag{1.13}$$

both from the left and from the right.

This theorem says that any typical path of such EI processes cannot be enclosed by two lines starting at the origin. Note that X is of infinite variation iff $\sigma>0$ or $\sum|\beta_j|=\infty$.

The idea of the proof is to use a change of measure, depending on a parameter $\theta\in\mathbb{R}$. In the particular case of Lévy processes, this change of measure is the Esscher transform, see Section 5.2. For general EI processes, this change of measure adds a drift (which depends on θ) and takes away some of the original jumps, thus X takes the form $\alpha^\theta\,\mathrm{Id}+Y^\theta$ under the new measure (this is proved in Proposition 5.3). Then, we prove that α^θ is not bounded on θ, under the infinite variation assumption, hence neither the Dini derivative (in page 137 we give a very simple idea of the proof).

Theorem 1.5 was proved by Rogozin [Rog68] in the case of Lévy processes, approximating the latter with compound Poisson processes and using an integro-differential equation (which is proved in [Wat71]). Other proofs for Lévy processes based in fluctuation theory are given in [Sat99, Ch. 9§47] and [Vig02, Theorem 5.3.2]. In the particular case that X is an EI process with non-negative jumps, the limsup was proved infinite by [Ber02], using martingale arguments and a result from [Fri72]. Kallenberg uses couplings with Lévy processes to obtain rates of growth for EI processes as in Theorem 1.5, but of the form $\limsup_{h\to 0}X_h/f(h)$ where f satisfies certain conditions. Proposition 3.5 of [CUB15] proves Theorem 1.5 when $\sigma>0$ or under an additional hypothesis on β; such proof uses the law of the iterated logarithm for Brownian motion or Lévy couplings, respectively. We remark that from those results our theorem is not obtainable.

Regularity of half-lines for a Lévy process has many applications: it helps in obtaining perfectness of the zero set and in constructing a continuous (Markovian) local time (Theorem 6.6 of [Kyp14]); it implies uniqueness for solutions of time-change equations used to construct multitype branching processes

(Lemma 6 in [CPGUB17]); regularity of $(-\infty,0)$ has been used when pricing perpetual American put options, as a condition for smooth pasting (see the discussion on Section 1.4.4 in [KL05]).

Now, we state the applications of our theorem. The first is an easy consequence.

Corollary 1.2. *Let X be an infinite variation EI process. Then X is regular.*

An EI process is called *extremal* if its canonical parameters are deterministic. We use Knight's result [Kni96], about the necessary and sufficient conditions for X to admit a unique minimum:

UM either $\sigma \neq 0$, or $\sum_i \mathbf{1}_{\beta_i \neq 0} = \infty$, or $\sum_i \mathbf{1}_{\beta_i \neq 0} < \infty$ and $\sum_i \beta_i \neq \alpha$.

The next result is an extension of Millar's zero-one law for Lévy processes at its infimum (items a and b of [Mil77, Thm. 3.1]); its proof is given in page 149. It characterizes the behavior of X at its infimum with its regularity.

Theorem 1.6. *Let X be an extremal EI process satisfying UM. Let $\underline{X}_1 = \inf_{s\in[0,1]} X_s$ and let ρ be the unique element of $\{t \in [0,1] : X_t \wedge X_{t-} = \underline{X}_1\}$. Then $X_\rho > \underline{X}_1$ if and only if X is irregular upward and $X_{\rho-} > \underline{X}_1$ if and only if X is irregular downward. In particular, X is continuous at ρ if and only if X is both upward and downward regular and this holds on the set where X has paths of infinite variation.*

Next, we obtain the weak limit of an EI process X ending at zero, conditioned to remain above $-\varepsilon$, as $\varepsilon \to 0$. The limit has the same law as the Vervaat transform $V(X)$ of X, which was explained in Definition 1.3.

Theorem 1.7. *Let X be a regular EI process with $\alpha = 0$. Consider $\varepsilon > 0$ and let X^ε have the law of X conditionally on $\underline{X}_1 > -\varepsilon$. Then $X^\varepsilon \xrightarrow{d} V(X)$ as $\varepsilon \to 0$.*

For Brownian bridges from 0 to 0, this was proved in [DIM77]. Our proof follows directly from [CUB15], and is given in Subsection 5.4.2. Indeed, the missing part to obtain our result in such paper, was only the zero-one law at the minimum of EI processes, which we already proved in Theorem 1.6.

Our final application, is to generalize the results in [PUB12] about the convex minorant of a Lévy process, but for EI processes. Our hypothesis is:

NPL $\sigma > 0$ or $\sum_i \mathbf{1}_{\beta_i \neq 0} = \infty$.

This is equivalent to work with EI processes not having piecewise linear trajectories. We introduce several definitions.

Definition 1.9. *The* convex minorant *of a càdlàg function $f : [0,1] \to \mathbb{R}$ is the greatest convex function c that is bounded above by f. The* excursion set *is the open set*

$$\mathscr{O} = \{t \in [0,1] : f(t) > c(t)\}.$$

Its maximal components, intervals of the form (g,d), are termed excursion intervals *and they have an associated* length $d-g$, increment $f(d)-f(g)$, slope $(f(d)-f(g))/(d-g)$ *and* excursion $e(t) = f(g+t) - c(g+t)$ *defined for $t \in [0, d-g]$.*

Since the supremum of a set of convex functions bounded from above is a convex function, let C be the convex minorant of X, where the latter satisfies NPL. We prove that its excursion set is open and of Lebesgue measure 1. Thus, an independent i.i.d. sequence $(U_i, i \in \mathbb{N})$ of uniform variables on $[0,1]$ falls a.s. inside the excursion set. Let (g_1,d_1), (g_2,d_2),... be the distinct excursion intervals successively discovered by $(U_i, i \in \mathbb{N})$, and $(e_i, i \in \mathbb{N})$ the corresponding excursions.

Now, consider another independent sequence $(V_i, i \in \mathbb{N})$ of i.i.d. uniform variables on $[0,1]$. The partition on $[0,1]$ given by the *stick-breaking process* $(L_i, i \in \mathbb{N})$ constructed from $(V_i, i \in \mathbb{N})$, is defined by

$$S_0 = 0, \quad S_{i+1} = S_i + L_i \quad \text{and} \quad L_{i+1} = (1 - S_i)V_{i+1}.$$

Definition 1.10. *Consider an EI process Y starting at zero on the interval $[0,t]$ for $t \in (0,1]$, satisfying UM. We define the* Knight bridge *of Y as $\tilde{K}_s = Y_s - (s/t)Y_t$. If ρ is the location of its infimum, define the* Knight transform *of Y as*

$$K_s = \tilde{K}_{(\rho+s) \bmod t} - \tilde{K}_\rho \wedge \tilde{K}_{\rho-} \text{ for } s \in [0,t].$$

For any i, let K^i be the Knight transform of $X - X_{S_{i-1}}$ on $[0, L_i]$. The next description of the convex minorant of an EI process, generalizes the result in [PUB12]. As before, the proof is a simple consequence of such paper, which lacked only the right hypotheses for the regularity of an EI process (see Subsection 5.4.3).

Theorem 1.8. *Assume that the EI process X satisfies NPL. Then, its excursion set $\mathscr{O}$ is open and of Lebesgue measure 1, also the following equality in law holds:*

$$\left(d_i - g_i, X_{d_i} - X_{g_i}, e^i\right)_{i \geq 1} \stackrel{d}{=} \left(L_i, X_{S_i} - X_{S_{i-1}}, K^i\right)_{i \geq 1}.$$

Using this theorem, we can obtain the following formula for NPL EI processes:

$$\mathbb{E}\left(\inf_{0 \leq s \leq 1} X_s\right) = \int_0^1 \frac{\mathbb{E}(X_l \wedge 0)}{l}\, dl,$$

also known as Kac's formula in the discrete case (see [Kac54]).

Chapter 2

MULTIDIMENSIONAL RANDOM WALKS CONDITIONED TO STAY ORDERED VIA GENERALIZED LADDER HEIGHT FUNCTIONS

In this chapter, we define a d-dimensional random walk conditioned to have ordered components forever, for any $d \in \mathbb{N}$. This is done defining the Doob h-transform of the random walk and the associated (sub)harmonic functions. The construction is equivalent as the limit as $c \downarrow 0$ of the random walk conditioned to have ordered components up to a geometric time of parameter $1 - e^{-c}$. Several reexpresions of the (sub)harmonic function are given, reducing to well-known formulas for the one dimensional case.

2.1 Introduction

Let $X^1, \ldots, X^d$ be independent, simple, symmetric random walks on $\mathbb{Z}$. Let $\mathbb{P}_{\mathbf{i}}$ be the probability measure of $X = (X^1, \ldots, X^d)$ starting at $\mathbf{i} = (i_1, \ldots, i_d)$. In [KOR02] and [EK08], the authors define X *conditioned to have ordered components*. The interest on such processes is by their relation with random matrix theory, for example, with Dyson's Brownian motion [Dys62], which can be interpreted as d Brownian motions conditioned to stay ordered at all times. As another important connection, conditioning a 2-dimensional random walk to have ordered components, is equivalent to condition a random walk to stay non-negative, a theory with a long history (see [Tan89, Kee92, Ber93, Cha94, BD94, Hir01, Tan04, CD05, CC08, Par08, GLP16]). The conditioning event can be written as

$$A = \{X_j^1 < \cdots < X_j^d \text{ for all } j \geq 0\}.$$

Denoting by $\mathbb{W} = \{x \in \mathbb{R}^d : x_1 < \cdots < x_d\}$ the *Weyl chamber*, the conditioning event A can be rewritten as $\{X_j \in \mathbb{W}, \forall j \geq 0\}$. Note that, even in the case $i_1 < \cdots < i_d$, we have

$$\mathbb{P}_{\mathbf{i}}(A) = 0,$$

since $X^2 - X^1$ is an oscillating random walk. Therefore, a rigorous definition of the law of X conditioned on A should be given. This is done in [KOR02] and [EK08] (as a particular case of their results), introducing the event

$$A_n = \{X_j^1 < \cdots < X_j^d \text{ for all } j \in [n]\} \qquad n \in \mathbb{N},$$

with $[n] = \{1, \ldots, n\}$, and proving that for every $k \in \mathbb{N}$, the limit as $n \to \infty$

$$\mathbb{P}_{\mathbf{i}}\left(X_0 = \mathbf{i}_0, \ldots, X_k = \mathbf{i}_k | A_n\right)$$

exists and is a probability measure. In fact, Karlin-McGregor's formula (cf. [KM59]) gives us an expression for $\mathbb{P}_{\mathbf{i}}(A_n)$ and implies

$$\lim_n \mathbb{P}_{\mathbf{i}}\left(X_0 = \mathbf{i}_0, \ldots, X_k = \mathbf{i}_k | A_n\right) = \mathbb{E}_{\mathbf{i}}\left(\frac{\Delta(\mathbf{i}_k)}{\Delta(\mathbf{i}_0)} \mathbf{1}_{X_0 = \mathbf{i}_0, \ldots, X_k = \mathbf{i}_k}\right),$$

where, for $x = (x_1, \ldots, x_d) \in \mathbb{W}$

$$\Delta(x) = \prod_{1 \le i < j \le d} (x_j - x_i) = \det\left(\left(x_j^{i-1}\right), i, j \in [d]\right). \tag{2.1}$$

is the Vandermonde's determinant. Hence, the conditioning is made using Doob's *h-transform*. Similar transformations, also called *h-transforms* or *h-process*, appear in [Saw97]. We must emphasize that most of the papers constructing ordered random walks are based on finding the limit as n goes to infinity of $\mathbb{P}_{\mathbf{i}}(A_n)/\mathbb{P}(A_n)$. Such limit will be the associated *h-function* of the process.

The objective is to generalize known constructions to d-dimensional random walks conditioned to have ordered components. In particular, the components of X could be dependent or have different distributions. Some models in the literature are [EK08, Dur14b, DW15, GR16, Ign18]. A general construction (when the drift is zero) is given in [DW15], for random walks in cones. The assumptions on the step distribution are that each of its components has mean zero, variance one, and zero covariance between components; also, a moment assumption is made on the step distribution. The case of non-zero drift was solved in [Dur14b], using a Cramér condition. Also, the recent paper [Ign18] constructs ordered Markov chains without moment conditions, but the state space must be countable.

Our result has minimal assumptions. Define $Y = (X^2 - X^1, X^3 - X^2, \ldots, X^d - X^{d-1})$. In order to avoid trivial cases, we assume Y has components taking negative and positive values with positive probability; besides that, the construction works with no further hypotheses when some component Y^k drifts to $-\infty$, or every component Y^k drifts to $+\infty$ (see Lemma 2.9). In the remaining cases, the assumption on Y is the existence of positive $\varepsilon_1, \ldots, \varepsilon_{d-1}$ such that

$$\mathbb{P}\left(X_1^2 - X_1^1 \ge \varepsilon_1, \ldots, X_1^d - X_1^{d-1} \ge \varepsilon_{d-1}\right) > 0.$$

This condition is used only to prove the finiteness of the (sub)harmonic function h.

Our method is to analyze a random walk conditioned to have ordered components up to an independent geometric time N of parameter $1 - e^{-c}$, and take the limit as $c \to 0$. The main tool is to construct a *ladder height function* for the random walk, which is based on a generalization of the ladder times in the unidimensional case. These ideas are adapted from the unidimensional case given for random walks in [Ber93], and for Lévy processes in [CD05, Don07].

2.1.1 Statement of the results

For ease of notation, our results are stated for $d = 3$ and for a random walk having state space $\mathbb{R}^d$. Let $X = (X^1, X^2, X^3)$ be a 3-dimensional random walk on $\mathbb{R}^3 \cup \{\dagger\}$, starting at $X_0 = 0$, having lifetime $\zeta = \sup\{n : X_n \neq \dagger\}$. Its increments are denoted by $W = (W^1, W^2, W^3)$, and W_1 has law $\mathbb{P}$. We denote by $Y = (Y^1, Y^2) = (X^2 - X^1, X^3 - X^2)$ the size of the gap between components, and $y = (x_2 - x_1, x_3 - x_2)$ for $x \in W$. The law of X killed at time $n \in \mathbb{N}$, that is, on the event $\zeta = n$, will be denoted by $\mathbb{P}^n$. When killing X at an independent geometric law $N \in \{0, 1, \ldots, \}$ with parameter $1 - e^{-c}$, its law will be $\mathbb{P}^c = \sum_0^\infty e^{-cn}(1 - e^{-c})\mathbb{P}^n$. The σ-algebra considered will be $\mathscr{F}_n = \sigma(X_1, \ldots, X_n)$.

The notation that we use is component-wise, hence $\min\{Y_i, i \in \mathscr{I}\} = (\min\{Y_i^1, i \in \mathscr{I}\}, \min\{Y_i^2, i \in \mathscr{I}\})$, for any index set $\mathscr{I} \subset \mathbb{Z}_+$. We define for $[n] = \{1, 2, \ldots, n\}$ and $[n]_0 = \{0, 1, \ldots, n\}$, the processes $\underline{Y}_n = \min\{Y_i, i \in [n]_0\}$, $\underline{\underline{Y}}_n = \min\{Y_i, i \in [n]\}$, $\overline{Y}_n = \max\{Y_i, i \in [n]_0\}$, and $Y_n \vee (y_1, y_2) = (Y_n^1 \vee y_1, Y_n^2 \vee y_2)$, where $y_1, y_2 \in \mathbb{R}$. We also put $(x_1, \ldots, x_d) < (y_1, \ldots, y_d)$ whenever component-wise the strict inequality is satisfied.

For $\mathbb{W} = \{(x_1, \ldots, x_d) \in \mathbb{R}^d : x_1 < \cdots < x_d\}$, a *positive regular function* or *harmonic function*, with respect to the transition kernel of X to $\mathbb{W}$, is a function $h : \mathbb{W} \mapsto \mathbb{R}_+$ such that

$$\mathbb{E}_x\left(h(X_1); \tau > 1\right) = h(x) \qquad x \in \mathbb{W},$$

where

$$\tau := \min\{n : X_n \notin \mathbb{W}\}.$$

A function h is *subharmonic* (*superharmonic*) if $\mathbb{E}_x\left(h(X_1); \tau > 1\right) \leq h(x)$ ($\mathbb{E}_x\left(h(X_1); \tau > 1\right) \geq h(x)$) for every $x \in \mathbb{W}$. The resulting (sub)harmonic function associated with a Doob h-transform will also be called h-function.

To avoid trivial cases, we assume that Y has components taking positive and negative values with positive probability. Besides that, the construction works with no further hypothesis if either some component of Y drifts to $-\infty$, or every component of Y drifts to $+\infty$. When such conditions are not satisfied, we need Hypothesis (1.1), needed only for the finiteness of the (sub)harmonic function h. Our main result is the following, which justifies our construction can be interpreted as a random walk X conditioned to stay ordered forever.

Theorem 2.1. *Let N be a geometric time with parameter $1 - e^{-c}$, independent of X. Assume that*

$$h^{\uparrow}(x) := 1 + \mathbb{E}\left(\sum_{n=1}^{J_1 - 1} \mathbf{1}_{\overline{Y}_{n-1} - Y_n < y}\right) < \infty \qquad x = (x_1, \ldots, x_d) \in \mathbb{W},$$

with $y = (x_2 - x_1, \ldots, x_d - x_{d-1})$ and $J_1 = \inf\{n > 0 : \overline{Y}_{n-1}^k < Y_n^k, k \in [d-1]\}$. Then, for every $x \in \mathbb{W}$, every finite $\mathscr{F}_n$-stopping time T and $\Lambda \in \mathscr{F}_T$

$$\lim_{c \to 0^+} \mathbb{P}_x\left(\Lambda, T \leq N | X(i) \in \mathbb{W}, i \in [N]\right) = \mathbb{P}_x^{\uparrow}(\Lambda, T < \zeta) := \frac{1}{h^{\uparrow}(x)} \mathbb{E}_x^Q\left(h^{\uparrow}(X_T)\mathbf{1}_{\Lambda, T < \zeta}\right),$$

where $\mathbb{E}_x^Q$ is the expectation under the law of X killed at the first exit time of the Weyl chamber. The limit law is a Markov chain with transition probabilities

$$p^{\uparrow}(w, dz) = \mathbf{1}_{z \in \mathbb{W}} \frac{h^{\uparrow}(z)}{h^{\uparrow}(w)} p(w, dz) \qquad w \in \mathbb{W}. \tag{2.2}$$

Moreover, it is a probability measure if $\mathbb{E}(\tau) = \infty$, or a subprobability measure if $\mathbb{E}(\tau) < \infty$.

We also give simple conditions to ensure $h^{\uparrow}$ is finite.

Lemma 2.1. *Assume that either*

1. *some component of Y drifts to $-\infty$,*
2. *every component of Y drifts to $+\infty$ and $\mathbb{P}(\tau = \infty) > 0$,*
3. *there exists $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_{d-1}) \in \mathbb{R}_+^{d-1}$ such that*

$$\mathbb{P}(Y_1 > \varepsilon) > 0.$$

Then

$$h^{\uparrow}(x) < \infty \qquad \forall x \in \mathbb{W}.$$

The following is an application of Theorem 2.1.

Example 2.1. *For $d \geq 1$, consider a multidimensional random walk with partial sums $X_n = (an, X_n^1, \ldots, X_n^d, bn)$, where $a < b$ and $a, b \in \mathbb{R}$. Assume that $\mathbb{P}\big(X_1^1 - a > \varepsilon_1, X_1^2 - X_1^1 > \varepsilon_2, \ldots, b - X_1^d > \varepsilon_{d+1}\big) > 0$ for some $\varepsilon_i > 0$, $i \in \{1, \ldots, d+1\}$, and $\mathbb{P}\big(X_1^1 - a < 0, X_1^2 - X_1^1 < 0, \ldots, b - X_1^d < 0\big) > 0$. Then, the construction of Theorem 2.1 provides us with a d-dimensional random walk conditioned to have ordered components and staying inside the set $\{x \in \mathbb{R} : at < x < bt, \forall\, t \in \mathbb{R}_+\}$.*

Depending on the drift of its components, we reexpress our function $h^{\uparrow}$.

Lemma 2.2. *Let $x \in \mathbb{W}$. If some component of Y drifts to $-\infty$, the $h^{\uparrow}$-transform is given by*

$$h^{\uparrow}(x) = \frac{\mathbb{E}_x(\tau)}{\mathbb{E}(\tau)}.$$

If every component drifts to $+\infty$ and $\mathbb{P}(\tau = \infty) > 0$, then

$$h^{\uparrow}(x) = \frac{\mathbb{P}_x(\tau = \infty)}{\mathbb{P}(\tau = \infty)}.$$

We also express $h^{\uparrow}$ as a renovation function. For $k \in [d-1]$, denote by $(\beta_i^k, i \in \mathbb{N})$ the strict descending ladder times of Y^k, that is $\beta_0^k = 0$ and for $i \in \mathbb{N}$ the time β_i^k is the smallest index n such that $Y^k(\beta_{i-1}^k + n) < Y^k(\beta_{i-1}^k)$. Let $\{\beta_0, \beta_1, \beta_2, \ldots\}$ be the ordered union of all such ladder times, with $\beta_0 = 1$. Denoting by $g_n = \beta_n$ and $d_n = \beta_{n+1}$ for $n \geq 0$, the set $\{g_n, g_n + 1 \ldots, d_n - 1\}$ is the nth interval where $\underline{\underline{Y}}$ remains constant.

Proposition 2.1. *Let $x \in \mathbb{W}$. The h-transform can be expressed as*

$$h^{\uparrow}(x) = 1 + \sum_{n=1}^{\infty} \mathbb{P}\Big(-\underline{\underline{Y}}_{\beta_n} < y\Big).$$

We conjecture that our h-function is subharmonic when X has i.i.d. components taking values in $\mathbb{R}$, satisfies the hypotheses of [EK08, DW10, DW15], and $d > 2$. The reason is that on such papers, it is computed the tail of the distribution of $\tau := \inf\{n : X \notin \mathbb{W}\}$, which we prove characterizes the harmonicity of our h. In fact, we prove in Lemma 2.6 that h is harmonic (subharmonic) iff the random time τ has infinite (finite) mean. In [DW10] it is proved that $\mathbb{P}_x(\tau > n)$ is of the order $n^{-d(d-1)/4}$ (see Subsection

2.5), implying $\mathbb{E}_x(\tau) < \infty$ when $d \geq 3$ and $x \in \mathbb{W}$. Even that $\mathbb{E}_x(\tau) \leq \mathbb{E}(\tau)$, it is expected that $\mathbb{E}(\tau) < \infty$. This represents a difference with respect to [EK08], since, regardless of the dimension, their h-function is harmonic. We give some conditions on the drift of the components of Y, to obtain harmonic h-function.

Such difference has been observed recently in [Ign18]. As an application of its results, the author on such paper characterizes the h-functions of centered irreducible random walks taking values on a countable set, with slowly varying hitting probabilities and other minor assumptions. The author proved that when $\mathbb{E}(\tau) = \infty$, any harmonic function for the process is proportional to $V(\cdot) = \mathbb{E}_\cdot(\mathscr{T})$, where $\mathscr{T}$ is the exit time of some *ladder height process*, and also that $\lim \mathbb{P}_x(\tau > n)/\mathbb{P}(\tau > n) = V(x)$, with $x \in \mathbb{W}$. On the case $\mathbb{E}(\tau) < \infty$, the author proved that $V(x) \leq \underline{\lim}\, \mathbb{P}_x(\tau > n)/\mathbb{P}(\tau > n)$. We remark that such notion of ladder height process is not as intuitive as ours, and the results in [Ign18] are only given when the process takes values on a countable set, while our results works for $\mathbb{R}^d$.

It is important to address that even in the unidimensional case, the harmonicity of the h-function highly depends on $\mathbb{E}(\tau)$ and even in the way we choose the approximating events. An example appears in [BD94] for random walks not drifting to $+\infty$. They obtain limits of some random walks, under two different conditionings to stay positive. It is proved that for oscillating random walks both limits are the same. But when the drift is negative, depending on the upper tail of the step distribution it can happen: both limits are the same and the h-function is subharmonic; both limits are different with harmonic h-function. Results in the same spirit for Lévy processes, are given in [BD94]. Also, the h-function of the Brownian motion with negative drift conditioned to stay positive is harmonic or subharmonic, depending on the approximation: it is proved in [MSM94] that is harmonic conditioning with $\{\tau > t\}$ and letting $t \to \infty$; while in [CD05] is subharmonic when conditioning with $\{\tau > E/c\}$, an exponential random variable with mean 1, and letting $c \to 0$.

In fact, the set of non-negative harmonic functions for random walks in $\mathbb{R}^d$, having non-zero drift, and killed when leaving general cones with zero as a vertex, was proved to be uncountable in [Dur14a].

We adapt the idea of using geometric times to condition a process to stay positive, given for random walks in [Ber93], and for Lévy processes in [CD05, Don07]. Thus, we analyze a random walk conditioned to have ordered components up to an independent geometric time N of parameter $1 - e^{-c}$, and take the limit as $c \to 0$.

This chapter is organized as follows. In Subsection 2.1.2 we review similar models of particles conditioned to stay in some subspace and applications. Subsection 2.1.3 is devoted to provide an updated literature and explicit results about ordered random walks. We include the first constructions of ordered Brownian motions, several references to ordered processes, as well as the description of the most general result of ordered random walks in cones, given in [DW15]. Our construction is given in Section 2.2, where first we condition the random walk to stay ordered up to an independent geometric time, and prove this is a Markov chain. There, we obtain a function $h_c^\uparrow$ which plays the role of an h-function of the random walk, but up to the geometric time. In Subsections 2.2.1 and 2.2.2 we reexpress $h_c^\uparrow$ using a partition $\mathbb{N}$ on intervals, making the random walk to be like *excursions* on each interval. This allows us to obtain in Lemma 2.5 the limit $h^\uparrow$ of $h_c^\uparrow$ as $c \downarrow 0$, and in Theorem 2.1 the limit of the conditioned random walk, as a Markov chain using a change of measure with $h^\uparrow$. We characterize in Section 2.3 when $h^\uparrow$ is harmonic or subharmonic, give a condition to ensure its finiteness, and prove that the law of the random walk using the h-function $h^\uparrow$ is the same as the limit of the random walk law conditioned to stay ordered up to a geometric time. In Section 2.4 we obtain several reexpresions of $h^\uparrow$. Finally, in Section 2.5 we review known results for the order of $\mathbb{P}_x(\tau > n)$, thus, allowing us to know some cases where $\mathbb{E}_x(\tau)$ is finite.

2.1.2 Related models

There exists several models which apply the theory of ordered stochastic processes. For random walks, names like *non-colliding, non-intersecting* or *vicious walkers* are also employed. The model of *vicious random walks* was investigated in [Bai00]. Johansson used non-intersecting trajectories for the analysis of the *corner-growth* model in [Joh00], and for the *Artic circle* model in [Joh02]. There exists applications to *series of queues in tandem* given in [O'C03, KOR02].

When the processes are Brownian motions, many results have been found. A complex random matrix, having eigenvalues process a vector of Brownian motions ordered up to time s, is obtained in [KT03b]. In [O'C12], it is related a transformation of the partition function of a Brownian directed polymer model with dimension 1+1 in a random environment, with the distribution of the largest eigenvalue of the random matrix GUE. The latter has the same distribution as the first component of a multidimensional Brownian motion conditioned to stay ordered. Also, physical applications are given in [KT04] and [IK11] (Makoto Katori has done extensive research on non-colliding processes, mainly for Brownian motion). Brownian bridges conditioned to non-collision where obtained in [BS07], using approximating random walk bridges. The paper also analyzes the connections with the random matrix central limit theorem. The case of squared Bessel processes is obtained in [KO01]. Wolfgang König in Section 4 of [Kön05], surveys several results on non-colliding particle systems.

Some ordered infinite particle systems has been obtained before, for example in [KNT04] and [Bai00]. The former gives an N Brownian particle system conditioned to non-collision up to time s, and let N and s tend to infinity. The latter uses a system of point in $\mathbb{Z}_+^2$, with each point in $2\mathbb{N}$ and at each step they move upwards one unity, and move left or right one unity as long as the place is empty.

Results about non-colliding systems with a wall are given in [Gra99, KT04, TW07, BFP+09, LW17]. In [BFP+09], it is related the law of the maximum of Dyson's Brownian motion with d particles, with models of non-intersecting Brownian particles which also: never hits a wall at the origin if d is par or the wall is reflecting if d odd. The case of ordered random walks conditioned to stay on the Weyl chambers of type C or D is given in [KS10], where

$$\mathbb{W}^C = \{x = (x_1, \ldots, x_k) \in \mathbb{R}^k : 0 < x_1 < \cdots < x_k)\},$$

and

$$\mathbb{W}^D = \{x = (x_1, \ldots, x_k) \in \mathbb{R}^k : |x_1| < x_2 < \cdots < x_k)\}.$$

Another related models are processes conditioned to stay in a subset of the state space. For example, Lévy processes conditioned to stay in an interval, studied in [Lam00], non-colliding systems on the unit circle as in [LW16], or more generally, Markov processes conditioned to never leave a subspace, as seen in [PR04].

2.1.3 Known results

First we review the results obtained in the unidimensional case, that is, conditioning a random walk to stay positive. The paper [Ber93], gives the h-transform of a random walk X conditioned to stay positive forever, under the unique hypothesis that the walk takes positive and negative values with positive probability (see Theorem 2.3 of such paper). The explicit formula for the h-function in such case is

$$h^{\uparrow}(x) = 1 + \mathbb{E}\left(\sum_{1}^{\alpha_1^+ - 1} \mathbf{1}_{-x < X_k}\right), \tag{2.3}$$

where $x \geq 0$ and α_1^+ is the index of the first visit to $(0,\infty)$.

Such function, can be reexpresed in several ways. We review some formulas in this case (see [BD94]). The first hitting times, respectively, in $(-\infty,0)$ and in $[n,\infty)$ are denoted by $\tau = \min\{k \geq 1 : X_k < 0\}$ and $\sigma_n = \min\{k \geq 1 : X_k \geq n\}$. Let $(H,T) = ((H_k,T_k), k \geq 0)$ be the strict ascending ladder point process of the reflected random walk $-X$. That is, we have $T_0 = 0$ and

$$H_k = -X_{T_k} \quad \text{and} \quad T_{k+1} = \min\{j > T_k : -X_j > H_k\}.$$

The convention is $H_k = \infty$ if $T_k = \infty$. The renewal function associated with H_1 is

$$V(x) = \sum_{k=0}^{\infty} \mathbb{P}(H_k \leq x), \qquad x \geq 0.$$

This is a non-decreasing right-continuous function. But the duality lemma gives us

$$V(x) = \mathbb{E}\left(\sum_{j=0}^{\sigma_0 - 1} \mathbf{1}_{-x \leq X_j}\right) = h^{\uparrow}(x).$$

In the case S drifts to $+\infty$

$$V(x) = \frac{\mathbb{P}_x(\tau = \infty)}{\mathbb{P}(\tau = \infty)}.$$

When S drifts to $-\infty$, we have $\mathbb{E}_x(\tau) < \infty$ for every $x \geq 0$ and

$$V(x) = \frac{\mathbb{E}_x(\tau)}{\mathbb{E}(\tau)}.$$

The main result of this chapter is to generalize formula (2.3) for the multidimensional case. We also give conditions, to obtain the remaining formulas in the multidimensional case.

In the multidimensional case, Dyson linked non-colliding particles with the theory of random matrices. Dyson [Dys62], introduced the theory of processes conditioned to non-collision. For $i < j$ and $i,j \in [d]$, let $B_{i,j}^1$ and $B_{i,j}^2$ be i.i.d. real standard Brownian motions starting at zero, and define the Hermitian random matrix $M = (M_{i,j})_{i,j=1}^d$ with $M_{i,j} = B_{i,j}^1 + iB_{i,j}^2$. Then, the process $(B + B^T)/\sqrt{2}$ is a *Gaussian Orthogonal Ensemble process* (or GOE process). The author identifies a distributional relation between the eigenvalues process of the matrix hermitian Brownian motion, and Brownian particle systems of dimension one, such that the repulsive forces of two particles are proportional to the inverse of their distance. Such process, say $Y = (Y_1,\ldots,Y_d)$, known as Dyson's Brownian motion, can be described by the stochastic differential equations

$$dY_i(t) = dB_i(t) + \frac{\beta}{2} \sum_{1 \leq j \leq k, j \neq i} \frac{1}{Y_i(t) - Y_j(t)} dt \qquad t \geq 0, \;\; i \in [d], \tag{2.4}$$

with $\beta = 1,2,4$ for the GOE, GUE and GSE respectively[1], and $(B_i, i \in [d])$ independent standard Brownian motions of dimension one.

Those *repulsive forces* in Dyson's Brownian motion, do not allow collisions between particles. That is, the process in (2.4) for $\beta = 2$, satisfies that $Y_0 \in \mathbb{W}$ implies $Y_t \in \mathbb{W}$ for every t with probability one. It has also been proved that the process can be started at the origin, see [OY02].

To prove that such a process can be viewed as Brownian motions conditioned never to collide, the two most used methods are:

[1] Gaussian unitary, ortogonal and symplectic ensembles

1. using Doob's h-transforms with harmonic functions,
2. finding the limit law of the process conditioned to non-collision up to the (possibly random) time T, and let $T \to \infty$.

Let us recall both methods. To obtain harmonic functions, the paper [Ras14] lists several methods. Grabiner proved in [Gra99], that the density of the transition probabilities of (2.4) with $\beta = 2$ is given by

$$p(s,x;t,y) = \frac{\Delta(y)}{\Delta(x)} f(t-s,y|x),$$

where Δ is the Vandermonde's determinant, and $f(t-s,y|x)$ is the transition density of the absorbing Brownian motion in $\mathbb{W}$, from state x at time s to state y at time t. Such f is found using Karlin-McGregor's formula [KM59]. As Δ is a positive harmonic function on $\mathbb{W}$, the process Y can be seen as an h-transform in Doob's sense, implying that the eigenvalues process of GUE has the same distribution that the h-transform of an absorbing Brownian motion in the Weyl's chamber.

Conditioning to non-collision up to a deterministic time is given in [KT03a]. They consider *vicious walker* models, which is a system of non-intersecting random walks. Proving a functional central limit theorem, they obtain Brownian motions $X = (X_1, \dots, X_d)$ which do not intersect on the time interval $(0,T]$, and prove X converges to Y as $T \to \infty$.

2.2 The random walk conditioned to be ordered up to a geometric time as an h-transform

Recall the notation at the beginning of Section 2.1.1. Consider $x = (x_1,x_2,x_3) \in \mathbb{W}$ and let $y = (y_1,y_2) = (x_2-x_1,x_3-x_2)$. Recall the definition of $\tau = \inf\{n : X_n \notin \mathbb{W}\}$, the first exit time from the Weyl chamber. For any $n \in \mathbb{N}$ and $A = A_0 \times A_1 \times \cdots \times A_n \in \mathscr{B}(\mathbb{R})^{n+1}$, we find the limit as $c \to 0+$ of

$$\mathbb{P}^c\left(\bigcap_0^n \{X_i + x_i \in A_i\} \middle| \tau > N\right) = \mathbb{P}^c\left(\bigcap_0^n \{X_i + x_i \in A_i\} \middle| X_j^1 + x_1 < X_j^2 + x_2 < X_j^3 + x_3, j \in [N]\right).$$

First we prove this is a Markov chain.

Proposition 2.2. *Under $\mathbb{P}^c$ and for any $x = (x_1,x_2,x_3) \in \mathbb{W}$, the chain $X + x$ conditioned to be ordered up to time N is a Markov chain with transition probabilities*

$$\mathbb{P}_c^{\uparrow}(w,dy) = \mathbf{1}_{y \in \mathbb{W}} \frac{h_c^{\uparrow}(y)}{h_c^{\uparrow}(w)} e^{-c} p(w,dy),$$

with $w = (w_1,w_2,w_3) \in \mathbb{W}$, $y = (y_1,y_2,y_3)$, $p(w,dy) = \mathbb{P}(W_1 + w \in dy)$ and

$$h_c^{\uparrow}(w) = \frac{\mathbb{P}^c\left(X_i + w \in \mathbb{W}, i \in [\zeta]\right)}{\mathbb{P}^c\left(X_i \in \mathbb{W}, i \in [\zeta]\right)}.$$

Proof. We compute the n-step transition probabilities

$$\mathbb{P}_x^c \left(X_i \in dw_i, i \in [n]_0 \mid \tau > N\right),$$

where $w_i \in \mathbb{R}^3$ for $i \in [n]_0 := \{0,\dots,n\}$ and $w_0 = x \in \mathbb{W}$. Then, the numerator of the n-step transition probability is given by

$$\mathbb{P}^c \left(X_i + x \in dw_i, i \in [n]_0, X_i + x \in \mathbb{W}, i \in [N]\right).$$

On such set we have $N \geq n$, and for $i \in [n]_0$ we have $X_i + x = w_i \in \mathbb{W}$, while for $n+i \in \{n, n+1, \dots, N\}$

$$\{X_n + x \in dw_n, X_{n+i} + x \in \mathbb{W}\} = \{X_n + x \in dw_n, X_{n+i} - X_n + w_n \in \mathbb{W}\}.$$

Summing over the values of N, and using independent and stationary increments of X, the numerator is equal to

$$\begin{aligned}
&\mathbf{1}_{\cap_1^n \{w_i \in \mathbb{W}\}} \mathbb{P}\left(X_i + x \in dw_i, i \in [n]_0\right) \\
&\times \sum_{k \geq n} \mathbb{P}\left(X_{n+i} - X_n + w_n \in \mathbb{W}, n+i \in [k]\right) \mathbb{P}(N = k) \\
&= \mathbf{1}_{\cap_1^n \{w_i \in \mathbb{W}\}} \mathbb{P}\left(X_i + x \in dw_i, i \in [n]_0\right) e^{-cn} \\
&\times \mathbb{P}^c \left(X_i + w_n \in \mathbb{W}, i \in [N]\right),
\end{aligned}$$

by the lack of memory property of N. Therefore, the n-step transition probability is given by

$$= \mathbf{1}_{\cap_1^n \{w_i \in \mathbb{W}\}} \mathbb{P}\left(X_i + x \in dw_i, i \in [n]_0\right) e^{-cn} \times \frac{\mathbb{P}^c \left(w_n + X_i \in \mathbb{W}, i \in [N]\right)}{\mathbb{P}^c \left(x + X_i \in \mathbb{W}, i \in [N]\right)}.$$

Denote by $X^\uparrow$ the random walk X conditioned to stay ordered up to time N. Considering $w_i \in \mathbb{W}$ for $i \in [n-1]_0$, we obtain, using that X is a random walk

$$\begin{aligned}
&\mathbb{P}^c \left(X^\uparrow(n) \in dw_n | X^\uparrow(0) = w_0, X^\uparrow(1) \in dw_1, \dots, X^\uparrow(n-1) \in dw_{n-1}\right) \\
&= \mathbf{1}_{\{w_n \in \mathbb{W}\}} \mathbb{P}\left(X_n + x \in dw_n | X_{n-1} + x \in dw_{n-1}\right) e^{-c} \frac{h_c^\uparrow(w_n)}{h_c^\uparrow(w_{n-1})} \\
&= \mathbf{1}_{\{w_n \in \mathbb{W}\}} \mathbb{P}\left(X_1 \in dw_n | X_0 \in dw_{n-1}\right) e^{-c} \frac{h_c^\uparrow(w_n)}{h_c^\uparrow(w_{n-1})},
\end{aligned}$$

which is the one-step transition probability, and depends only on w_{n-1} and w_n. $\square$

Now we analyze the function $h_c^\uparrow$.

2.2.1 Reexpression of the h-function of the ordered RW up to a geometric time

A priori, $h_c^\uparrow$ is the division of two probabilities converging to zero. We reexpress $h_c^\uparrow$ to prove it converges. Working with the numerator of $h_c^\uparrow(x)$, first sum over all possible values of N

$$\begin{aligned}
&\mathbb{P}^c \left(X_i + x \in \mathbb{W}, i \in [\zeta]\right) \\
&= \left(1 - e^{-c}\right) \left(1 + \sum_1^\infty e^{-cn} \mathbb{P}^n \left(X_i + x \in \mathbb{W}, i \in [n]\right)\right).
\end{aligned}$$

Recall that $Y=(X^2-X^1,X^3-X^2)$ and $y=(x_2-x_1,x_3-x_2)$. It follows that

$$\mathbb{P}^c\left(X_i+x\in\mathbb{W},i\in[\zeta]\right)/(1-e^{-c})-1=\mathbb{E}\left(\sum_1 e^{-cn}\mathbf{1}_{-y<\underline{\underline{Y}}_n}\right). \tag{2.5}$$

For any $n\in\mathbb{N}$, it is known that X and the time-reversed process X^* has the same distribution, with

$$X_i^*=X_n-X_{n-i} \qquad \text{for } 0\le i\le n.$$

This chain has components $X^*=(X^{1,*},X^{2,*},X^{3,*})$, and similarly for Y^*. Then

$$\left\{-y<\underline{\underline{Y}}_n\right\}\overset{d}{=}\left\{\max\{-Y_j^{1,*},j\in[n]\}<x_2-x_1,\max\{-Y_j^{2,*},j\in[n]\}<x_3-x_2\right\}.$$

Define $\overline{Y}_n^k=\max\{Y_j^k,0\le j\le n\}$ for $k=1,2$ and $\overline{Y}_n=(\overline{Y}_n^1,\overline{Y}_n^2)$. Add and subtract the term $Y_n^{k,*}$ and use $Y_i^{k,*}-Y_n^{k,*}=Y_n^k-Y_{n-i}^k-Y_n^{k,*}=-Y_{n-i}^k$ for $k=1,2$, so

$$\left\{-y<\underline{\underline{Y}}_n\right\}\overset{d}{=}\left\{\overline{Y}_{n-1}-Y_n<y\right\}. \tag{2.6}$$

This implies that $h_c^\uparrow$ can be reexpresed as

$$h_c^\uparrow(x)=\frac{1+\sum_1^\infty e^{-cn}\mathbb{P}\left(\overline{Y}_{n-1}-Y_n<y\right)}{1+\sum_1^\infty e^{-cn}\mathbb{P}\left(\overline{Y}_{n-1}-Y_n<0\right)} \tag{2.7}$$

2.2.2 Partitioning $\mathbb{N}$ via the times of a multidimensional ladder height function to obtain the limit of the approximated h-function

In this subsection, we partition $\mathbb{N}$ at some particular times $\{J_i,i\in\mathbb{N}\}$. Those are the times in common among the ascending ladder times of Y^1 and Y^2, that is, if $(\alpha_j^k,j\ge 0)$ are the strict ascending ladder times of Y^k, then J_i is the ith time such that $\alpha_j^1=\alpha_l^2$ for some $j,l\in\mathbb{N}$. We prove that the subpaths $\{Y_{J_i+n},0\le n<J_{i+1}-J_i\}_i$ are i.i.d., and at the times J_i, every component of the walk Y is at least as big as the current cumulative maximum. In this sense, the reader should think on those subpaths as *excursions* of Y.

Let $J_0=0$ and for $i\in\mathbb{N}$, define

$$J_{i+1}=\min\left\{n>J_i:\overline{Y}_{n-1}^k<Y_n^k,k=1,2\right\},$$

the first time after J_i, such that both walks reach the current maximum at the same time.

Remark 2.1. *Note that $\overline{Y}_{J_i}=Y_{J_i}$, since both processes are at the same maximum. Also, since we assumed $\mathbb{P}(Y_1>0)>0$, then $\mathbb{P}(J_1=1)=\mathbb{P}\left(\overline{Y}_0<Y_1\right)=\mathbb{P}(0<Y_1)$ has positive probability.*

We prove $(J_i,i\ge 0)$ are stopping times. Let $(\mathscr{F}_n,n\in\mathbb{N})$ be the natural filtration of X. For any $m\in\mathbb{N}$, the event $\{J_1=m\}$ is equal to

$$\left\{\overline{Y}_{j-1}^k\ge Y_j^k,1\le j\le m-1 \text{ for } k=1 \text{ or } k=2\right\}\cap\left\{\overline{Y}_{m-1}^k<Y_m^k,k=1,2\right\},$$

which is in $\mathscr{F}_m$. Assuming J_i is a stopping time, the event $\{J_{i+1}=m\}$ is equal to

$$\bigcup_{l=i}^{m-1}\left(\{J_i=l\}\cap\left\{\overline{Y}^k_{j-1}\geq Y^k_j, l+1\leq j\leq m-1 \text{ for } k=1 \text{ or } k=2\right\}\cap\left\{\overline{Y}^k_{m-1}<Y^k_m, k=1,2\right\}\right),$$

which also belongs to $\mathscr{F}_m$.

We prove the independence and obtain the distribution between such times.

Lemma 2.3. *For every $i\in\mathbb{N}$, the walk $\{\overline{Y}_{J_i+n-1}-Y_{J_i+n}, n\geq 1\}$ is independent of $\mathscr{F}_{J_i}$ and has the same distribution as $\{\overline{Y}_{n-1}-Y_n, n\geq 1\}$.*

Proof. Let $T<\infty$ be a stopping time. For $n\geq 2$, decompose $\overline{Y}_{T+n-1}$ as the maximum up to time T and the maximum between times $\{T+1,\ldots,T+n-1\}$. Hence

$$\overline{Y}_{T+n-1}-Y_{T+n}=\left(\overline{Y}_T-Y_T\right)\vee\max\{Y_{T+l}-Y_T, l\in[n-1]\}-(Y_{T+n}-Y_T),$$

and for $n=1$

$$\overline{Y}_T-Y_{T+1}=\left(\overline{Y}_T-Y_T\right)\vee(0,0)-(Y_{T+1}-Y_T).$$

We substitute $T=J_i$ for $i\in\mathbb{N}$ and recall $\overline{Y}_{J_i}=Y_{J_i}$. For $n\in\mathbb{N}$ and $A_m\in\mathbb{R}^2$ with $m\in[n]$, the events

$$\begin{aligned}&\bigcap_{m=1}^{n}\{\overline{Y}_{J_i+m-1}-Y_{J_i+m}\in A_m\}\\ &=\bigcap_{m=1}^{n}\{(0,0)\vee\max\{Y_{J_i+l}-Y_{J_i}, l\in[m-1]\}-(Y_{J_i+m}-Y_{J_i})\in A_m\}\end{aligned}$$

are independent of $\mathscr{F}_{J_i}$ under $\{J_i<\infty\}$, by the strong Markov property. They also have the same distribution as

$$\bigcap_{m=1}^{n}\{(0,0)\vee\max\{Y_l, l\in[m-1]\}-Y_m\in A_m\}=\bigcap_{m=1}^{n}\left\{\overline{Y}_{m-1}-Y_m\in A_m\right\},$$

recalling that $\overline{Y}_0=Y_0=(0,0)$ under $\mathbb{P}$. □

The following result is crucial to partition the sums in (2.7).

Lemma 2.4. *The times $\{J_{i+1}-J_i, i\in\mathbb{N}\}$ are i.i.d. and $J_{i+1}-J_i=J_1\circ\theta_{J_i}$, where θ is the translation operator.*

Proof. For $i\in\mathbb{N}$ we have

$$\begin{aligned}J_{i+1}-J_i&=\min\{n>0:\overline{Y}_{J_i+n-1}<Y_{J_i+n}\}\\ &=\min\{n>0:\max\{Y_{J_i+m}-Y_{J_i}; 0\leq m\leq n-1\}-(Y_{J_i+n}-Y_{J_i})<0\}\\ &=J_1\circ\theta_{J_i}.\end{aligned}$$

Then $J_{i+1}-J_i$ is independent of $\mathscr{F}_{J_i}$ and has the same law as J_1, by Lemma 2.3. □

Lemma 2.5. *The h-function $h_c^{\uparrow}$ converges as $c \downarrow 0$ to*

$$h^{\uparrow}(x) = 1 + \mathbb{E}\left(\sum_{n=1}^{J_1-1} \mathbf{1}_{\overline{Y}_{n-1}-Y_n<y}\right),$$

recalling that $y = (x_2 - x_1, x_3 - x_2)$.

Proof. Recall Equation (2.7). Partition $\mathbb{N}$ at times $(J_i, i \in \mathbb{N})$

$$\begin{aligned}&\mathbb{E}\left(\sum_1 e^{-cn}\mathbf{1}_{\overline{Y}_{n-1}-Y_n<y}\right)\\ &= \mathbb{E}\left(\sum_0 e^{-cJ_i}\mathbf{1}_{J_i<\infty}\sum_{n=1}^{J_{i+1}-J_i} e^{-cn}\mathbf{1}_{\overline{Y}_{n-1+J_i}-Y_{n+J_i}<y}\right).\end{aligned}$$

Conditioning with $\mathscr{F}_{J_i}$ and summing over the values taken by $J_{i+1} - J_i$, the previous equation is equal to

$$\mathbb{E}\left\{\sum_0 e^{-cJ_i}\mathbf{1}_{J_i<\infty}\sum_{m\geq 1}\sum_{n=1}^{m} e^{-cn}\mathbb{P}\left(J_{i+1}-J_i = m, \overline{Y}_{n-1+J_i} - Y_{n+J_i} < y \,|\, \mathscr{F}_{J_i}\right)\right\}.$$

Using lemmas 2.3 and 2.4, we obtain

$$\begin{aligned}&\mathbb{P}^c\left(X_i + x \in \mathbb{W}, i \in [\zeta]\right)/(1-e^{-c})\\ &= 1 + \mathbb{E}\left(\sum_0^{\infty} e^{-cJ_i}\mathbf{1}_{J_i<\infty}\right)\mathbb{E}\left(\sum_{n=1}^{J_1} e^{-cn}\mathbf{1}_{\overline{Y}_{n-1}-Y_n<y}\right).\end{aligned}$$

Since $e^{-cJ_i}\mathbf{1}_{J_i=\infty} = 0$, we can ignore the indicator $\mathbf{1}_{J_i<\infty}$. When $x = 0$, the only term that remains in the second expectation above is e^{-cJ_1}, since $\overline{Y}^1_{n-1} \geq Y^1_n$ or $\overline{Y}^2_{n-1} \geq Y^2_n$ for $n < J_1$. It follows that

$$\mathbb{P}^c\left(X_i \in \mathbb{W}, i \in [\zeta]\right)/(1-e^{-c}) = 1 + \mathbb{E}\left(\sum_0^{\infty} e^{-cJ_i}\right)\mathbb{E}\left(e^{-cJ_1}\right).$$

Dividing both terms, and using

$$\mathbb{E}\left(\sum_{n=1}^{J_1} e^{-cn}\mathbf{1}_{\overline{Y}_{n-1}-Y_n<y}\right) - \mathbb{E}\left(e^{-cJ_1}\right) = \mathbb{E}\left(\sum_{n=1}^{J_1-1} e^{-cn}\mathbf{1}_{\overline{Y}_{n-1}-Y_n<y}\right),$$

we have

$$h_c^{\uparrow}(x) = 1 + \mathbb{E}\left(\sum_{n=1}^{J_1-1} e^{-cn}\mathbf{1}_{\overline{Y}_{n-1}-Y_n<y}\right)\frac{\mathbb{E}\left(\sum_0^{\infty} e^{-cJ_i}\right)}{1+\mathbb{E}(\sum_0^{\infty} e^{-cJ_i})\,\mathbb{E}(e^{-cJ_1})}.$$

Since $J_i = \sum_0^{i-1}(J_{k+1} - J_k)$ is a sum of i.i.d. random variables, then

$$\mathbb{E}\left(\sum_0^{\infty} e^{-cJ_i}\right) = \left(1 - \mathbb{E}\left(e^{-cJ_1}\right)\right)^{-1},$$

implying

$$h_c^{\uparrow}(x) = 1 + \mathbb{E}\left(\sum_{n=1}^{J_1-1} e^{-cn}\mathbf{1}_{\overline{Y}_{n-1}-Y_n<y}\right).$$

The result follows from the monotone convergence theorem. □

The latter result implies Theorem 2.1. In the next section, we prove that $h^{\uparrow}$ is (sub)harmonic, and give a simple condition that ensures it is finite.

2.3 Properties of the limiting h-function and the interpretation of the walk as conditioned to stay ordered forever

2.3.1 The harmonicity of the h-function depends on the first exit time to $\mathbb{W}$

We know that the first exit time from the Weyl chamber is given by

$$\tau = \min\{n > 0 : Y_n^1 \wedge Y_n^2 \leq 0\}.$$

By Lemma 2.5, we rewrite $h^{\uparrow}$ as

$$h^{\uparrow}(x) = \lim_{c \to 0^+} \frac{\mathbb{P}_x^c\left(\underline{\underline{Y}}_N > 0\right)}{\mathbb{P}^c\left(\underline{\underline{Y}}_N > 0\right)} = \lim_{c \to 0^+} \frac{\mathbb{P}_x(\tau > N)}{\mathbb{P}(\tau > N)}.$$

Let Q_x be the law of X killed at the first exit of the Weyl chamber, that is, for $n \in \mathbb{N}$ and $\Lambda \in \mathscr{F}_n$

$$Q_x(\Lambda, n < \zeta) = \mathbb{P}_x(\Lambda, n < \tau).$$

Expectations under Q_x will be denoted by $\mathbb{E}_x^Q$. The next lemma gives us conditions to know if $h^{\uparrow}$ is harmonic or subharmonic. It is based on Lemma 1 of [CD05].

Lemma 2.6. *Let $x \in \mathbb{W}$. If $\mathbb{E}(\tau) < \infty$, then $h^{\uparrow}$ is subharmonic and*

$$\mathbb{E}_x^Q\left(h^{\uparrow}(X_n)\mathbf{1}_{n<\zeta}\right) < h^{\uparrow}(x).$$

If $\mathbb{E}(\tau) = \infty$, then $h^{\uparrow}$ is harmonic and

$$\mathbb{E}_x^Q\left(h^{\uparrow}(X_n)\mathbf{1}_{n<\zeta}\right) = h^{\uparrow}(x).$$

Proof. Since we proved in Lemma 2.5 that the convergence of $h_c^{\uparrow}$ to $h^{\uparrow}$ is monotone, then

$$\mathbb{E}_x^Q\left(h^{\uparrow}(X_n)\mathbf{1}_{n<\zeta}\right) = \lim_{c \to 0^+} \mathbb{E}_x\left(\frac{\mathbb{P}_{X_n}(\tau > N)}{\mathbb{P}(\tau > N)}\mathbf{1}_{n<\tau}\right). \tag{2.8}$$

Using the Markov property

$$\begin{aligned}
\mathbb{P}_x(\tau > n+N) &= \mathbb{E}_x\left(\mathbf{1}_{Y_k^2 \wedge Y_k^1 > 0, k \in [n+N]}\right) \\
&= \mathbb{E}_x\left(\mathbf{1}_{Y_k^2 \wedge Y_k^1 > 0, k \in [n]}\mathbf{1}_{Y_k^2 \wedge Y_k^1 > 0, k \in [N]} \circ \theta_n\right) \\
&= \mathbb{E}_x\left(\mathbf{1}_{\tau > n}\mathbb{P}_{X_n}(\tau > N)\right),
\end{aligned}$$

which is the numerator in the right-hand side of Equation (2.8). Summing over all the values of N

$$\begin{aligned}\mathbb{P}_x(\tau > n+N) &= \sum_k \mathbb{P}_x(\tau > n+k, N=k)\\ &= e^{cn}\sum_{k\geq n}\mathbb{P}_x(\tau > k)(1-e^{-c})e^{-ck}.\end{aligned}$$

Starting the sum from $k=0$, we obtain

$$\mathbb{P}_x(\tau > n+N) = e^{cn}\left\{\mathbb{P}_x(\tau > N) - \sum_0^{n-1}\mathbb{P}_x(\tau > k)\,\mathbb{P}(N=k)\right\}.$$

Thus, the right-hand side of Equation (2.8) is equal to

$$\begin{aligned}&\lim_{c\to 0^+} e^{cn}\left\{\frac{\mathbb{P}_x(\tau > N)}{\mathbb{P}(\tau > N)} - \sum_0^{n-1}\frac{\mathbb{P}_x(\tau > k)\,e^{-ck}}{\sum\mathbb{P}(\tau > m)\,e^{-cm}}\right\}\\ &= h^{\uparrow}(x) - \frac{1}{\mathbb{E}(\tau)}\sum_0^{n-1}\mathbb{P}_x(\tau > k),\end{aligned}$$

which proves the lemma, since $\mathbb{P}_x(\tau > 0) = \mathbb{P}(x+X_0\in\mathbb{W}) = 1$. □

2.3.2 Finiteness of the h-function

To prove $h^{\uparrow}(x) < \infty$ for every $x\in\mathbb{W}$, we use the remark of Lemma 1 in [Tan89]. In this subsection, the inequality $x > z$ for $x,z\in\mathbb{R}^3$ means there is strict inequality component-wise.

Lemma 2.7. *Assume there exists $\varepsilon = (\varepsilon_1,\varepsilon_2)\in\mathbb{R}_+$ such that*

$$\mathbb{P}\big((X_1^2 - X_1^1, X_1^3 - X_1^2) > \varepsilon\big) > 0.$$

Then

$$h^{\uparrow}(x) < \infty \qquad \forall x\in\mathbb{W}.$$

Proof. Note that Lemma 2.6 was independent of the finiteness of $h^{\uparrow}$. Hence, from such lemma and $x\in\mathbb{W}$ we have

$$h^{\uparrow}(x) \geq \int\mathbb{P}(x+X_1\in dz, 1<\tau)\,h^{\uparrow}(z) = \int_{z\in\mathbb{W}}\mathbb{P}(x+X_1\in dz)\,h^{\uparrow}(z).$$

Define $g(x) = (x_2-x_1, x_3-x_2)$ for $x\in\mathbb{R}^3_+$. For simplicity, instead of $g(x)$ we write x^-. So, for instance $X^- = (X^2-X^1, X^3-X^2)$ and $x^- := (x_2-x_1, x_3-x_2)$. Then, we have

$$\mathbb{P}(x+X_1\in dz) = \mathbb{P}\big(x_1+X_1^1\in dz_1, x^-+X_1^-\in dz^-\big).$$

Note from Lemma (2.5) that $h^{\uparrow}(x)$ depends on x only trough x^-. Define $h^- : \mathbb{R}^2_+\cup\{(0,0)\}\mapsto\mathbb{R}_+$ as $h^-(x^-) := h^{\uparrow}(x)$, so

$$h^-(x^-) = 1 + \mathbb{E}\left(\sum_{n=1}^{J_1-1}\mathbf{1}_{\overline{Y}_{n-1}-Y_n<x^-}\right).$$

It follows that

$$h^-(x^-) \geq \int_{z\in\mathbb{W}} \mathbb{P}\big(x_1 + X_1^1 \in dz_1, x^- + X_1^- \in dz^-\big)\, h^-(z^-). \tag{2.9}$$

Also, note that $h^-(0) = 1$, since $\mathbf{1}_{\overline{X}_0^- - X_1^- \leq 0} = 1$ implies $J_1 = 1$.

Assume that $h^-(z^-) = \infty$ for every $z^- > \varepsilon$. Fix any $x_1 \in \mathbb{R}$, and use $x = (x_1, x_1, x_1)$ in (2.9) to obtain

$$1 = h^-(0) = h^\uparrow(x) \geq \int_{z\in\mathbb{W}\cap\{z\in\mathbb{R}^3 : z^- > \varepsilon\}} \mathbb{P}\big(x_1 + X_1^1 \in dz_1, x^- + X_1^- \in dz^-\big)\, h^-(z^-).$$

Hence, it should be the case that

$$0 = \mathbb{P}\big(x_1 + X_1^1 \in \mathbb{R}, x^- + X_1^- > \varepsilon\big) = \mathbb{P}\big(X_1^- > \varepsilon\big),$$

contradicting the hypothesis. Therefore, there exists $z_{(1)} = (x_1, x_1 + z_{(1),1}, x_1 + z_{(1),1} + z_{(1),2}) \in \mathbb{W}$ such that

$$z_{(1)}^- > \varepsilon \qquad \text{and} \qquad h^\uparrow(z_{(1)}) = h^-(z_{(1)}^-) < \infty.$$

Now, assume $h^-(z^-) = \infty$ for every $z^- > \varepsilon + z_{(1)}^-$. Use $x = z_{(1)}^-$ in (2.9) to obtain

$$\infty > h^-(z_{(1)}^-) \geq \int_{z\in\mathbb{W}\cap\{z\in\mathbb{R}^3 : z^- > \varepsilon + z_{(1)}^-\}} \mathbb{P}\Big(x_1 + X_1^1 \in dz_1, z_{(1)}^- + X_1^- \in dz^-\Big)\, h^-(z^-).$$

Then, it should happen that

$$0 = \mathbb{P}\Big(x_1 + X_1^1 \in \mathbb{R}, z_{(1)}^- + X_1^- > \varepsilon + z_{(1)}^-\Big) = \mathbb{P}\big(X_1^- > \varepsilon\big),$$

again contradicting the hypothesis. Hence, there exists $z_{(2)} = (x_1, x_1 + z_{(2),1}, x_1 + z_{(2),1} + z_{(2),2}) \in \mathbb{W}$ such that

$$z_{(2)}^- > \varepsilon + z_{(1)}^- \qquad \text{and} \qquad h^\uparrow(z_{(2)}) = h^-(z_{(2)}^-) < \infty.$$

Continuing in this way, there is some subsequence $(z_{(n)}, n \in \mathbb{N})$, with $z_{(n)} = (x_1, x_1 + z_{(n),1}, x_1 + z_{(n),1} + z_{(n),2}) \in \mathbb{W}$ satisfying

$$z_{(n)}^- > \varepsilon + z_{(n-1)}^- \qquad \text{and} \qquad h^\uparrow(z_{(n)}) = h^-(z_{(n)}^-) < \infty,$$

for every n.

Fix any $x = (x_1, x_2, x_3) \in \mathbb{W}$. We prove that $h^\uparrow(x) < \infty$. Note that in the previous analysis, x_1 was arbitrary. Let $n \in \mathbb{N}$ such that

$$z_{(n),1}^- \wedge z_{(n),2}^- > (n\varepsilon_1) \wedge (n\varepsilon_2) > (x_3 - x_2) \vee (x_2 - x_1).$$

It follows that

$$\begin{aligned} h^\uparrow(x) &= 1 + \mathbb{E}\left(\sum_{n=1}^{J_1 - 1} \mathbf{1}_{\overline{X}_{n-1}^- - X_n^- < x^-}\right) \\ &\leq 1 + \mathbb{E}\left(\sum_{n=1}^{J_1 - 1} \mathbf{1}_{\overline{X}_{n-1}^- - X_n^- < z_{(n)}^-}\right), \end{aligned}$$

which is finite by construction. $\square$

2.3.3 Ordered random walks as the limit law of random walks conditioned to stay ordered up to a geometric time

Let $(q_n, n \geq 1)$ be the transition probabilities of (X, Q). From Theorem 2.1, denote by $(p_n^{\uparrow}, n \geq 1)$ the transition probabilities of X conditioned to stay ordered

$$p_n^{\uparrow}(w, dz) = \frac{h^{\uparrow}(z)}{h^{\uparrow}(w)} q_n(w, dz) \qquad w \in \mathbb{W},\ n \in \mathbb{N}.$$

The law of the Markov process with transition probabilities $(p_n^{\uparrow}, n \geq 1)$ and starting from $x \in \mathbb{W}$ is denoted by $\mathbb{P}_x^{\uparrow}$. Hence, for $n \in \mathbb{N}$ and $\Lambda \in \mathscr{F}_n$

$$\mathbb{P}_x^{\uparrow}(\Lambda, n < \zeta) = \frac{1}{h^{\uparrow}(x)} \mathbb{E}_x^{Q}\left(h^{\uparrow}(X_n) \mathbf{1}_{\Lambda, n<\zeta}\right). \tag{2.10}$$

Its lifetime is $\mathbb{P}_x^{\uparrow}$-finite if $h^{\uparrow}$ is subharmonic and $\mathbb{P}_x^{\uparrow}$-infinite if it is harmonic. Let us prove $(X, \mathbb{P}_x^{\uparrow})$ is the limit as $c \to 0+$ of $(X, \mathbb{P}_x)$ conditioned to have ordered components up to a geometric time.

Lemma 2.8. *Let N be geometric time with parameter $1 - e^{-c}$, independent of $(X, \mathbb{P})$. Then, for every $x \in \mathbb{W}$, every finite $\mathscr{F}_n$-stopping time T and $\Lambda \in \mathscr{F}_T$*

$$\lim_{c \to 0^+} \mathbb{P}_x\left(\Lambda, T \leq N | X(i) \in \mathbb{W}, i \in [N]\right) = \mathbb{P}_x^{\uparrow}(\Lambda, T < \zeta).$$

Proof. First we use a deterministic time $T \in \mathbb{N}$. Note that $\{T < \tau\} = \{X(i) \in \mathbb{W}, i \in [T]\}$. We work with $\mathbb{P}_x\left(\Lambda, T \leq N, X(i) \in \mathbb{W}, i \in [N]\right)$. Separating in the first T values of X and summing over all the values of N

$$\begin{aligned}
&\mathbb{P}_x\left(\Lambda, T \leq N, X(i) \in \mathbb{W}, i \in [N]\right) \\
&= \sum_{n \geq T} \mathbb{P}_x\left(\Lambda, T \leq n, X(i) \in \mathbb{W}, i \in [T], X(T+i) \in \mathbb{W}, i \in [n-T]\right) \mathbb{P}(N = n),
\end{aligned}$$

starting the sum at zero and using the Markov property at $\mathscr{F}_T$

$$\begin{aligned}
&\mathbb{P}_x\left(\Lambda, T \leq N, X(i) \in \mathbb{W}, i \in [N]\right) \\
&= e^{-cT} \sum_{n \geq 0} \mathbb{P}_x\left(\Lambda, T < \tau, \mathbb{P}_x\left(X(T+i) \in \mathbb{W}, i \in [n] | \mathscr{F}_T\right)\right) \mathbb{P}(N = n) \\
&= e^{-cT} \mathbb{P}_x\left(\Lambda, T < \tau, \mathbb{P}_{X_T}\left(N < \tau\right)\right) \\
&= \mathbb{P}_x\left(\Lambda, T < \tau, t \leq N, \mathbb{P}_{X_T}\left(N < \tau\right)\right).
\end{aligned}$$

Now, consider $c_0 > 0$ and any $c \in (0, c_0)$. Recall from Lemma 2.5 that $h_c^{\uparrow}$ increases to $h^{\uparrow}$, hence

$$\begin{aligned}
\mathbf{1}_{\Lambda, T<\tau, T \leq N} \frac{\mathbb{P}_{X_T}\left(\tau > N\right)}{\mathbb{P}_x\left(\tau > N\right)} &= \mathbf{1}_{\Lambda, T<\tau, T \leq N} \frac{h_c^{\uparrow}(X_T)}{h_c^{\uparrow}(x)} \\
&\leq \mathbf{1}_{\Lambda, T<\tau} \frac{h^{\uparrow}(X_T)}{h_{c_0}^{\uparrow}(x)}.
\end{aligned}$$

Taking expectations on both sides and using Lemma 2.6

$$\mathbb{E}_x\left(\mathbf{1}_{\Lambda,T<\tau,T\leq N}\frac{\mathbb{P}_{X_T}(\tau>N)}{\mathbb{P}_x(\tau>N)}\right)\leq\frac{h^{\uparrow}(x)}{h^{\uparrow}_{c_0}(x)},$$

and the right-hand side is finite by Lemma 2.7. Hence, by Lebesgue's dominated convergence theorem

$$\lim_{c\to 0^+}\mathbb{P}_x(\Lambda,T\leq N|\tau>N)=\lim_{c\to 0^+}\mathbb{E}_x\left(\mathbf{1}_{\Lambda,T<\tau,T\leq N}\frac{\mathbb{P}_{X_T}(\tau>N)}{\mathbb{P}_x(\tau>N)}\right)=\mathbb{P}^{\uparrow}_x(\Lambda,T<\zeta).$$

Let us prove the same convergence for any finite stopping time T. Summing over all the values of T, the equality

$$\mathbb{P}_x(\Lambda,T\leq N<\tau)=\mathbb{P}_x(\Lambda,T<\tau,t\leq N,\mathbb{P}_{X_T}(N<\tau))$$

and Equation (2.10) holds for T. We need to prove Lemma 2.6 holds true for any stopping time $T<\infty$ a.s. Summing over all the values of T, in the subharmonic case

$$\begin{aligned}\mathbb{E}^{Q}_x\left(h^{\uparrow}(X_T)\mathbf{1}_{T<\zeta}\right)&=\sum_n\mathbb{E}_x\left(h^{\uparrow}(X_n)\mathbf{1}_{n<\tau,T=n}\right)\\&\leq h^{\uparrow}(x)\sum_n\mathbb{P}^{\uparrow}_x(T=n,n<\zeta)\\&=h^{\uparrow}(x)\mathbb{P}^{\uparrow}_x(T<\infty,T<\zeta),\end{aligned}$$

which is smaller than $h^{\uparrow}(x)$. In the harmonic case, the inequality above is an equality, so it remains to prove that $\mathbb{P}^{\uparrow}_x(T<\infty,T<\zeta)=1$. But this is clear since

$$\mathbb{P}^{\uparrow}_x(T=\infty,T<\zeta)=\lim_n\frac{1}{h^{\uparrow}(x)}\mathbb{E}_x\left(h^{\uparrow}(X_T)\mathbf{1}_{T>n}\right)=0,$$

by monotone convergence. □

In the next section, we obtain several reexpresions of the h-function.

2.4 Reexpressions of the h-function

2.4.1 Reexpresions using the minimum of the descending ladder times of the components

Changing the measure to start at zero, we have

$$\mathbb{E}\left(\sum_1^{\infty}e^{-cn}\mathbf{1}_{-y<\underline{Y}_n}\right)=\mathbb{E}_x\left(\sum_1^{\infty}e^{-cn}\mathbf{1}_{0<\underline{Y}_n}\right).$$

For $k=1,2$, denote by $(\beta^k_i,i\in\mathbb{N})$ the strict descending ladder times of Y^k, that is $\beta^k_0=0$ and for $i\in\mathbb{N}$ the time β^k_i is the smallest index n such that $Y^k(\beta^k_{i-1}+n)<Y^k(\beta^k_{i-1})$. The above sum stops when one

component Y^k becomes negative, that is, at $\beta_1^1 \wedge \beta_1^2$. Then

$$\begin{aligned} h_c^{\uparrow}(x) &= \frac{1+\mathbb{E}_x\left(\sum_1^\infty e^{-cn}\mathbf{1}_{0<\underline{Y}_n}\right)}{1+\mathbb{E}\left(\sum_1^\infty e^{-cn}\mathbf{1}_{0<\underline{Y}_n}\right)} \\ &= \frac{1+\mathbb{E}_x\left(\sum_1^{\beta_1^1\wedge\beta_1^2-1} e^{-cn}\right)}{1+\mathbb{E}\left(\sum_1^{\beta_1^1\wedge\beta_1^2-1} e^{-cn}\right)}. \end{aligned}$$

This equality allows us to prove the next proposition.

Lemma 2.9. *If some component of Y drifts to $-\infty$, the h-function $h^{\uparrow}$ is given by*

$$h^{\uparrow}(x) = \frac{\mathbb{E}_x\left(\beta_1^1\wedge\beta_1^2\right)}{\mathbb{E}\left(\beta_1^1\wedge\beta_1^2\right)} = \frac{\mathbb{E}_x(\tau)}{\mathbb{E}(\tau)}.$$

If every component drifts to $+\infty$ and $\mathbb{P}\left(\beta_1^1\wedge\beta_1^2=\infty\right)>0$, then

$$h^{\uparrow}(x) = \frac{\mathbb{P}_x\left(\beta_1^1\wedge\beta_1^2=\infty\right)}{\mathbb{P}\left(\beta_1^1\wedge\beta_1^2=\infty\right)} = \frac{\mathbb{P}_x(\tau=\infty)}{\mathbb{P}(\tau=\infty)}.$$

Proof. If Y^k drifts to $-\infty$ for some $k=1,2$, then $\mathbb{E}\left(\beta_1^1\wedge\beta_1^2\right)\le\mathbb{E}\left(\beta_1^k\right)<\infty$ by Proposition 9.3, page 167 of [Kal02]. Therefore, by the monotone convergence theorem

$$h_c^{\uparrow}(x) \to \frac{1+\mathbb{E}_x\left(\beta_1^1\wedge\beta_1^2-1\right)}{1+\mathbb{E}\left(\beta_1^1\wedge\beta_1^2-1\right)}.$$

If every component drifts to $+\infty$, then $Y^1\wedge Y^2$ has a finite minimum with positive probability. By hypothesis $\mathbb{P}\left(\beta_1^1\wedge\beta_1^2=\infty\right)>0$, so

$$\begin{aligned} h_c^{\uparrow}(x) &= \frac{\mathbb{E}_x\left(\sum_0^{\beta_1^1\wedge\beta_1^2-1} e^{-cn}\right)}{\mathbb{E}\left(\sum_0^{\beta_1^1\wedge\beta_1^2-1} e^{-cn}\right)} \\ &= \frac{\mathbb{E}_x\left(\left(1-e^{-c(\beta_1^1\wedge\beta_1^2)}\right)\mathbf{1}_{\beta_1^1\wedge\beta_1^2<\infty}+\mathbf{1}_{\beta_1^1\wedge\beta_1^2=\infty}\right)}{\mathbb{E}\left(\left(1-e^{-c(\beta_1^1\wedge\beta_1^2)}\right)\mathbf{1}_{\beta_1^1\wedge\beta_1^2<\infty}+\mathbf{1}_{\beta_1^1\wedge\beta_1^2=\infty}\right)} \\ &\to \frac{\mathbb{P}_x\left(\beta_1^1\wedge\beta_1^2=\infty\right)}{\mathbb{P}\left(\beta_1^1\wedge\beta_1^2=\infty\right)}. \end{aligned}$$

□

2.4.2 Reexpresions using the union of the descending ladder times

Let $\{\beta_1,\beta_2,\ldots\}$ be the ordered union of the positive strict descending ladder times of Y^1 and Y^2, that is, the ordered union of $\{\beta_i^1,\beta_j^2,i,j\ge1\}$. Define $\beta_0=1$. Denoting by $g_n=\beta_n$ and $d_n=\beta_{n+1}$ for $n\ge0$, the set $\{g_n, g_n+1, \ldots, d_n-1\}$ is the nth interval where $\underline{\underline{Y}}$ remains constant. Partitioning $\mathbb{N}$ on such intervals, from Equation (2.5)

$$\begin{aligned}
&1+\mathbb{E}\left(\sum_1^\infty e^{-cn}\mathbf{1}_{-y<\underline{\underline{Y}}_n}\right)\\
&=\mathbb{E}\left(\sum_{n\geq 0}\mathbf{1}_{g_n<\infty}\sum_{k=g_n}^{d_n-1} e^{-c(k-g_n)}e^{-cg_n}\mathbf{1}_{-y<\underline{\underline{Y}}_k}\right)\\
&=\mathbb{E}\left(\sum_{n\geq 0} e^{-cg_n}\mathbf{1}_{g_n<\infty,-y<\underline{\underline{Y}}_{g_n}}\sum_{k=0}^{d_n-g_n-1} e^{-ck}\right).
\end{aligned}$$

The above equation for $x=0$ is

$$\mathbb{E}\left(\sum_{n\geq 0} e^{-cn}\mathbf{1}_{0<\underline{\underline{Y}}_n}\right)=\mathbb{E}\left(\sum_{k=0}^{d_0-1} e^{-ck}\right).$$

Note that $d_0=\beta_1^1\wedge\beta_1^2$. Also, note that $-y\leq 0\leq \underline{\underline{Y}}_{d_0-1}$, therefore

$$\begin{aligned}
h_c^{\uparrow}(x)&=\frac{\mathbb{E}\left(\sum_0^{d_0-1} e^{-ck}\right)+\sum_{n\geq 1}\mathbb{E}\left(e^{-cg_n}\mathbf{1}_{g_n<\infty,-y<\underline{\underline{Y}}_{g_n}}\sum_{k=0}^{d_n-g_n-1} e^{-ck}\right)}{\mathbb{E}\left(\sum_0^{d_0-1} e^{-ck}\right)}\\
&=1+\left(\mathbb{E}\left(\sum_0^{d_0-1} e^{-ck}\right)\right)^{-1}\sum_{n\geq 1}\mathbb{E}\left(e^{-cg_n}\mathbf{1}_{g_n<\infty,-y<\underline{\underline{Y}}_{g_n}}\sum_{k=0}^{d_n-g_n-1} e^{-ck}\right).
\end{aligned}$$

As before, depending on the asymptotic behavior of the components, we can obtain a limit.

Proposition 2.3. *If some component of Y drifts to $-\infty$, then*

$$\mathbb{E}(d_n-g_n)<\infty \qquad \forall n,$$

and the h-function is

$$h^{\uparrow}(x)=1+\sum_{n\geq 1}\frac{\mathbb{E}\left(d_n-g_n; g_n<\infty, -y<\underline{\underline{Y}}_{g_n}\right)}{\mathbb{E}(d_0)}.$$

If every component drifts to $+\infty$ and $\mathbb{P}(d_0=\infty)>0$, then

$$h^{\uparrow}(x)=1+\sum_{n\geq 1}\frac{\mathbb{P}\left(g_n<\infty, -y<\underline{\underline{Y}}_{g_n}, d_n-g_n=\infty\right)}{\mathbb{P}(d_0=\infty)}.$$

Proof. As before, if the component k drifts to $-\infty$, the first case follows by monotone convergence theorem and $d_n-g_n\leq\beta_j^k-\beta_{j-1}^k$ for some j. The second case follows by

$$\sum_0^{d_n-g_n-1} e^{-ck}=\frac{1-e^{-c(d_n-g_n)}}{1-e^{-c}}\mathbf{1}_{d_n<\infty}+\frac{1}{1-e^{-c}}\mathbf{1}_{d_n=\infty},$$

and using monotone convergence theorem. $\square$

2.4.3 Reexpresion as a renovation function

Recall from the previous section that $(\beta_n, n \geq 0)$ is the ordered union of the strict descending ladder times of Y^1 and Y^2. We have the following result.

Proposition 2.4. *The h-function $h\uparrow$ can be expressed as*

$$h^{\uparrow}(x) = 1 + \sum_{n=1}^{\infty} \mathbb{P}\Big(-\underline{\underline{Y}}_{\beta_n} < y\Big).$$

Proof. First we express $h^{\uparrow}$ as an infinite sum, using Tonelli's theorem and Theorem 2.1, we have

$$h^{\uparrow}(x) - 1 = \mathbb{E}\left(\sum_{n=1}^{J_1-1} \mathbf{1}_{\overline{Y}_{n-1} - Y_n < y}\right) = \sum_{n=1}^{\infty} \mathbb{E}\Big(\mathbf{1}_{\overline{Y}_{n-1} - Y_n < y} \mathbf{1}_{J_1 > n}\Big). \tag{2.11}$$

The event $\{J_1 > n\}$ means that for every $j \in [n]$, there is some k, such that the running maximum at time $j-1$ of Y^k is at least Y_j^k. This is written as

$$\{J_1 > n\} = \bigcap_{j=1}^{n} \bigcup_{k} \Big\{\max\Big\{Y_l^k; 0 \leq l \leq j-1\Big\} - Y_j^k \geq 0\Big\}.$$

Recall the equality in distribution between Y and Y^*, which is Y reversed in time. Also, recall the equality in distribution of $-\underline{\underline{Y}}$ and $\overline{Y}_{\cdot-1} - Y_{\cdot}$ of Equation (2.6). Hence, we have

$$\begin{aligned}
\{J_1 > n\} &\stackrel{d}{=} \bigcap_{j=1}^{n} \bigcup_{k} \Big\{\max\Big\{Y_l^{k,*}; 0 \leq l \leq j-1\Big\} - Y_j^{k,*} > 0\Big\} \\
&= \bigcap_{j=1}^{n} \bigcup_{k} \Big\{\max\Big\{-Y_{n-l}^k; 0 \leq l \leq j-1\Big\} + Y_{n-j}^k > 0\Big\} \\
&= \bigcap_{j=1}^{n} \bigcup_{k} \Big\{\min\Big\{Y_{n-l}^k; 0 \leq l \leq j-1\Big\} < Y_{n-j}^k\Big\}.
\end{aligned}$$

In a similar way, we can prove that

$$\Big\{\overline{Y}_{n-1} - Y_n \leq y, J_1 > n\Big\} \stackrel{d}{=} \Big\{\underline{\underline{Y}}_n^k \geq -y_k, \forall k\Big\} \bigcap \bigcap_{j=1}^{n} \bigcup_{k} \Big\{\min\Big\{Y_{n-l}^k; 0 \leq l \leq j-1\Big\} < Y_{n-j}^k\Big\}.$$

Now we prove the last term means n is a strict descending ladder time of some Y^k. In fact, reordering the index set, the last term is equal to

$$\begin{aligned}
&\bigcap_{j=1}^{n} \bigcup_{k} \Big\{\min\Big\{Y_l^k; n-j+1 \leq l \leq n\Big\} < Y_{n-j}^k\Big\} \\
&= \bigcap_{j=1}^{n} \bigcup_{k} \Big\{\min\Big\{Y_l^k; j \leq l \leq n\Big\} < Y_{j-1}^k\Big\}.
\end{aligned}$$

The right-hand side means the future minimum of Y^k up to time n is always smaller than the current value of Y^k, for some k. Thus, the time n is a strict descending ladder time of some Y^k, and

$$h^{\uparrow}(x) = 1 + \sum_{n=1}^{\infty} \mathbb{P}\left(\underline{Y}_{\beta_n} > -y\right). \qquad \square$$

The next section is devoted to obtain conditions for the finiteness of $\mathbb{E}(\tau)$.

2.5 Known results about the expectation of the first exit time of $\mathbb{W}$, to ensure the h-function is harmonic

In Theorem 1 of [DW10], the tail of the distribution of τ is computed. Explicitly, let $X = (X^1, \ldots, X^d)$ be a random walk with i.i.d. components on $\mathbb{R}$. Under the assumptions that the step distribution has mean zero and the α moment is finite for $\alpha = d-1$ if $d > 3$, and $\alpha > 2$ if $d = 3$, they prove

$$\lim_n n^{d(d-1)/4} \mathbb{P}_x(\tau > n) = KV(x),$$

where K is an explicit constant and V is given by

$$V(x) = \Delta(x) - \mathbb{E}_x(\Delta(X(\tau))) \qquad x \in \mathbb{W} \cap S^d,$$

with $S \subset \mathbb{R}$ the state space of the random walks, and Δ defined in (2.1). This implies that for $x \in \mathbb{W} \cap S^d$

$$\mathbb{E}_x(\tau) < \infty \qquad \text{whenever } d \geq 3.$$

This suggests that $\mathbb{E}(\tau) < \infty$ in this case.

In the paper [Dur14a] the author obtains the asymptotic behavior of τ_x, for random walks with non-zero drift killed when leaving general cones on $\mathbb{R}^d$. Under some assumptions, in particular, the step distribution having all moments and a drift pointing out of the cone, it is proved the existence of a function U such that

$$\mathbb{P}(\tau_x > n) \sim \rho c^n n^{-p-d/2} U(x).$$

The value $p \geq 1$ is the order of some homogeneous function, and $c \in [0,1]$. This suggests $\mathbb{E}(\tau_x) \leq \rho U(x) \sum c^n < \infty$ whenever $c \in (0,1)$.

As mentioned in Section 2.1.3, in the paper [DW15] the authors obtain

$$\mathbb{P}(\tau_x > n) \sim cV(x) n^{-p/2} \qquad n \to \infty,$$

for random walks in a cone, with components having zero mean, variance one, covariance zero, and some finite moment. In that case, the value p is

$$p = \sqrt{\lambda_1 + (d/2-1)^2} - (d/2-1) > 0.$$

Thus, the expectation of τ_x is infinite iff

$$\begin{aligned} 1 \geq p/2 &\iff (d/2+1)^2 \geq \lambda_1 + (d/2-1)^2 \\ &\iff 2d \geq \lambda_1. \end{aligned}$$

The paper [GR16], computes the asymptotic exit time probability for random walks in cones, under some general conditions. The first is that the support of the probability measure of $X(1)$ is not included in any linear hyperplane. The second is that, if L is the Laplace transform of the random walk having x^* as a minimum, then L is finite on an open neighborhood of x^*, and that this value belongs to the dual cone. Under such hypotheses, they prove that

$$\lim_{n\to\infty} \mathbb{P}_x(\tau > n)^{1/n} = L(x^*),$$

for all $x \in K_\delta := K + \delta v$, for some $\delta \geq 0$ and some fixed v in K^o. The authors note that in general, there is no explicit link between the drift m of the walk (if exists), x^* and $L(x^*)$. The only exception is when $m \in K$. In such case, $L(x^*) = 1$ iff $x^* = 0$. Furthermore, when the drift m exists, then $m \in K$ iff $x^* = 0$. Hence, if we want that $\mathbb{E}(\tau_x) = \infty$, we should restrict to the case $L(x^*) = 1$.

Chapter 3

ON THE PROFILE OF TREES WITH A GIVEN DEGREE SEQUENCE

Let $\mathbf{s_n} = (N_i^n, i \geq 0)$ be a sequence of non-negative integers satisfying $s_n := \sum_i N_i^n = 1 + \sum_i iN_i^n$. Such an $\mathbf{s_n}$ is called a *degree sequence*. Let $\mathbb{P}_{\mathbf{s_n}}$ be the uniform distribution on all rooted plane trees with given degree sequence $\mathbf{s_n}$. We give conditions for the convergence of the profile (sequence of generation sizes), giving a more general formulation and a proof of a conjecture due to Aldous [Ald91b], concerning the convergence of the profile for conditioned Galton-Watson trees. Our formulation contains and extends results in this direction obtained previously by Drmota and Gittenberger [DG97] and Kersting [Ker11]. A technical result is needed to ensure trees with law $\mathbb{P}_{\mathbf{s_n}}$ have enough individuals in the first generations, and this is handled through a novel path transformation of exchangeable increment processes.

3.1 Introduction and statement of the results

Trees are an important concept in both pure and applied mathematics, appearing (for example) in biology to represent genealogies, in computer science as a fundamental data structure, as well as an important example of a combinatorial class or species (cf. [Knu98], [Knu06], [Joy81], [FS09], [Drm09]). Random trees, on the other hand, have been shown to be useful in analyzing the asymptotic behavior of certain families of deterministic trees. For example, certain Galton-Watson (GW) trees conditioned to have size n (denoted CGW(n)), have been shown to be uniform in classes of trees of size n, like plane trees, binary plane trees or Cayley trees (cf. [Pit98]).

We will analyze the class of trees with a given degree sequence. The reason is that for several important real-world networks that have been analyzed, their degree sequence might have a certain feature (see for example [BA99, CDS11]), and the simplest way to build an associated model of random trees is through the uniform distribution on trees whose degree sequence has the observed feature. One way to understand the shape of rooted trees is through their profile, which counts the quantity of elements in the successive generations. In this chapter, we will focus on the profile of uniform trees with a given degree sequence. The introductions in [FHN06, GK12] summarize certain applications and references on the profile of random trees.

Trees with a given degree sequence are more general than CGW(n) trees, since the latter laws can be obtained as *mixtures* of the former (see Section 3.5). A conjecture due to Aldous [Ald91b] for CGW(n) having a finite variance offspring distribution, states that the rescaled profile converges in distribution to a multiple of the total local time process of the normalized Brownian excursion (NBE). Aldous's conjecture was proved in [DG97] as a complex application of analytic combinatorics.

The latter work was generalized in [Ker11] to the case where the offspring distribution is in the domain of attraction of a stable law. We show a much more general version of Aldous's conjecture in the setting of trees with a given degree sequence, satisfying a *finite variance condition*. Also, in Section 3.5, we obtain as a particular case of our results another proof of Kersting's theorem [Ker11], and consequently of Aldous's conjecture.

Let us turn to the formal statements of our results. We define rooted plane trees following [Nev86]. Let $\mathbb{Z}_+ = \{1,2,\dots\}$ be the set of positive integers, and define $\mathscr{U} = \bigcup_{n=0}^{\infty} \mathbb{Z}_+^n$ as the set of all labels, using the convention $\mathbb{Z}_+^0 = \{\varnothing\}$. An element of $\mathscr{U}$ is a sequence $u = u_1 \cdots u_n$ of positive integers, where $|u| = n$ represents the *generation* of u. If $u = u_1 \cdots u_i$ and $v = v_1 \cdots v_j$ belong to $\mathscr{U}$, write $uv = u_1 \cdots u_i v_1 \cdots v_j$ for the concatenation of u and v. By convention $u\varnothing = \varnothing u = u$. For any $n \in \mathbb{N}$, let $[n] = \{1,\dots,n\}$ with $[0] = \varnothing$.

Definition 3.1. *A* rooted plane tree T *is a finite subset of* $\mathscr{U}$ *such that:*

1. $\varnothing \in T$,
2. *if* $v \in T$ *and* $v = uj$ *for some* $j \in \mathbb{Z}_+$, *then* $u \in T$,
3. *for every* $u \in T$, *there exists a number* $c(u) \in \mathbb{N}$, *such that* $uj \in T$ *iff* $j \in [c(u)]$.

In the previous definition, the value $c(u)$ represents the number of children of u in T. The size of a tree T (the number of individuals) will be denoted by $|T|$. In the following, by a tree we mean a rooted plane tree. See Figure 3.1 for a graphical representation.

Order the vertices of the tree according to the *lexicographical order* (e.g., $\varnothing < 1 < 21 < 22$), and assign label i to the ith vertex, for $i \in [|T|]$. This ordering is also called *depth-first order*. The *depth-first walk* (DFW) of the tree will be the walk with ith increment $c(i)-1$, started at one. See Figure 3.1 for an example.

Another labeling of the tree, is the *breadth-first order*. To define this, assign label 1 to the root. Suppose the first generation (offspring of the root) has size z_1. Order the first generation in lexicographical order, and assign label i to the ith vertex, for $i \in \{2,\dots,1+z_1\}$. Do the same for each consecutive generation. The *breadth-first walk* (BFW) of the tree will be the walk with ith increment $c(i)-1$, started at one. See Figure 3.2 for an example.

Now we introduce the trees analyzed in this paper. For every $n \in \mathbb{N}$, let $\mathbf{s_n} = (N_i^n, i \geq 0)$ be a *degree sequence*, that is, a sequence of non-negative integers satisfying $s_n := |\mathbf{s_n}| = \sum N_i^n = 1 + \sum i N_i^n$. The values N_i^n represent the number of vertices with i children of some rooted plane tree T_n. Let $\mathbb{P}_{\mathbf{s_n}}$ be the distribution which samples uniformly at random from the set of all *trees with the given degree sequence* $\mathbf{s_n}$. Consider a tree τ_n with law $\mathbb{P}_{\mathbf{s_n}}$. The BFW (DFW) of τ_n will be denoted by W_{s_n}.

Denote by $C_{s_n}(j)$ the total number of vertices of τ_n up to generation j. This is called the *cumulative profile* or *cumulative population*. Denote by $Z_{s_n}(j)$ the number of vertices in generation j, called in this paper the *profile*. Other names given in the literature for the profile are the *population profile*, *horizontal profile* or *height profile*. Note that Z_{s_n} can be recursively obtained as follows: $Z_{s_n}(0) = 1$ and

$$Z_{s_n}(j+1) = W_{s_n} \circ C_{s_n}(j), \tag{3.1}$$

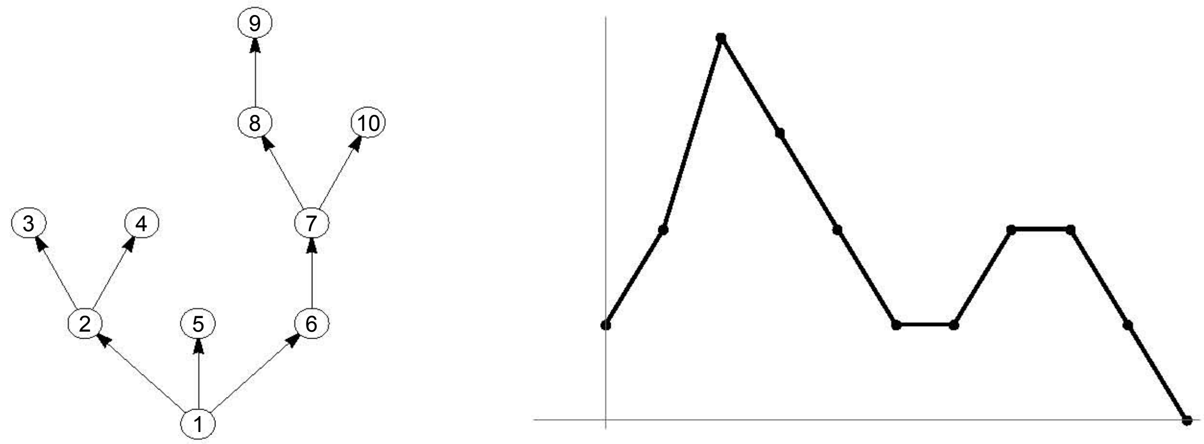

Figure 3.1: Tree labeled in lexicographical order (depth-first order) and the corresponding DFW. The explicit order of the vertices in lexicographical order is $\varnothing < 1 < 11 < 12 < 2 < 3 < 31 < 311 < 3111 < 312$.

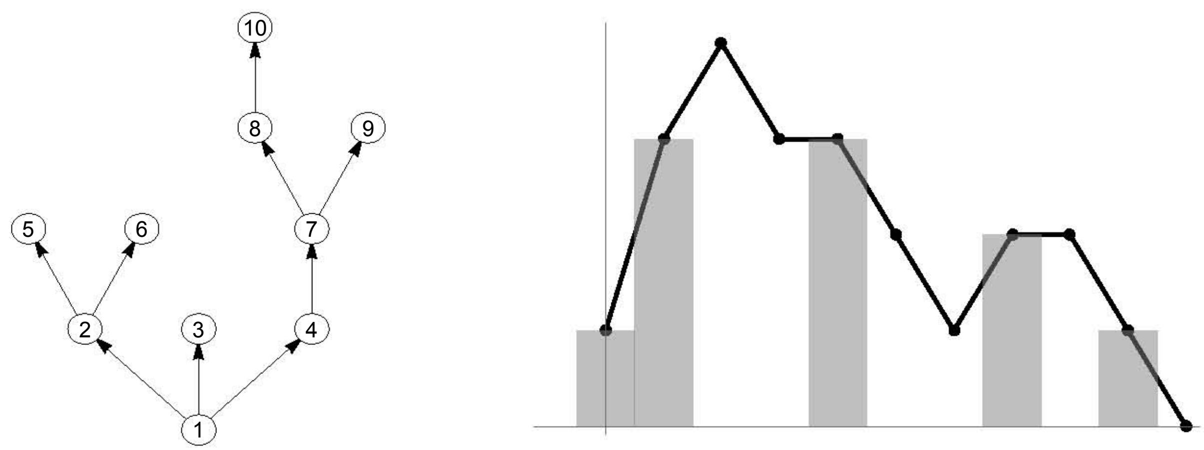

Figure 3.2: Tree labeled in breadth-first order and the relation between the profile and the BFW. The profile of the tree is $(1,3,3,2,1)$, which is the BFW evaluated at times $0,1,4,7,9$.

where W_{s_n} is the BFW of τ_n and $C_{s_n}(j) = \sum_0^j Z_{s_n}(i)$ (cf. Chapter 9 of [EK86], and the introduction in [CPGUB13]). By analogy with the continuous-time case (cf. [Lam67a]), Z_{s_n} is called the *discrete Lamperti transform* of W_{s_n}. Figure 3.2 shows an example of a BFW and its Lamperti transform.

To study the scaling limits of the profiles, the general idea is: rescale Equation (3.1) and prove the limit of the profile is a particular solution to

$$Z(t) = X \circ C(t) \quad t \geq 0, \tag{3.2}$$

where X is the limit of the rescaled BFWs and $C(t) = \int_0^t Z(s)\,ds$.

In our case, the limit X is related to a continuous time process X^b with exchangeable increments (EI process, or also known as interchangeable increments process) of the form

$$X^b(t) = \sigma b(t) + \sum_1^\infty \beta_j \left(\mathbf{1}\left(U_j \leq t\right) - t\right) \qquad t \in [0,1], \tag{3.3}$$

where b is a Brownian bridge on $[0,1]$, $(U_j, j \geq 1)$ are independent of b and i.i.d. with uniform law on $[0,1]$, and with constants $\sigma \in \mathbb{R}^+$, $\beta_1 \geq \beta_2 \geq \cdots > 0$ with $\sum \beta_j^2 < \infty$. From Kallenberg's representation [Kal73], the process X^b has canonical parameters $(0, \sigma, \beta)$.

The "excursion-type" process X, is obtained from X^b using the *Vervaat* transformation, defined as follows. Denote by $\{t\}$ the fractional part of a real number t. Let $D[0,1]$ be the space of real-valued càdlàg functions w with domain $[0,1]$ starting at zero (with an analogous definition for $D[0,\infty)$). Let $D' \subset D[0,1]$ be the subset of functions w such that $w(0) = w(1) = 0$, and w hits its infimum in a unique time and continuously. Define for every $u \in [0,1]$ the transformation $\theta_u : D' \mapsto D'$ by

$$\theta_u(w)(t) = w(\{t+u\}) - w(u).$$

This transformation can be described as cutting the path of w at u obtaining two paths (to the left and to the right of u), and interchanging them.

Definition 3.2. *The* Vervaat transform *of a càdlàg function $w \in D'$ is defined as $V(w) = \theta_\rho(w)$, where ρ is the unique time w hits its infimum.*

Figure 3.3 shows an example.

For a given degree sequence $\mathbf{s_n}$, let $(c(j), j \in [s_n])$ be the associated *child sequence*, obtained by writing N_0^n zeros, N_1^n ones, etc. Note that $N_j^n = |\{i : c(i) = j\}|$. Our main theorem is the following.

Theorem 3.1. *Assume the sequence $(\mathbf{s_n}, n \geq 1)$ of degree sequences satisfies*

1. $s_n \to \infty$.

2. *There exists a sequence of positive numbers $(b_{s_n}, n \geq 1)$ going to infinity, and $M \in \mathbb{N} \cup \{+\infty\}$, such that*

$$\left(\frac{1}{b_{s_n}^2}\sum (j-1)^2 N_j^n, \frac{\tilde{c}(1)}{b_{s_n}}, \frac{\tilde{c}(2)}{b_{s_n}}, \dots\right) \to \left(\sigma^2 + \sum_1^M \beta_j^2, \beta_1, \beta_2, \dots\right)$$

for some $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_M > 0$, and $\beta_{M+j} = 0$ for every j, and such that $\sum_1^M \beta_j^2 < \infty$ and $\sigma^2 \in [0,\infty)$, and where $(\tilde{c}(i), i \geq 1)$ is the child sequence in decreasing order.

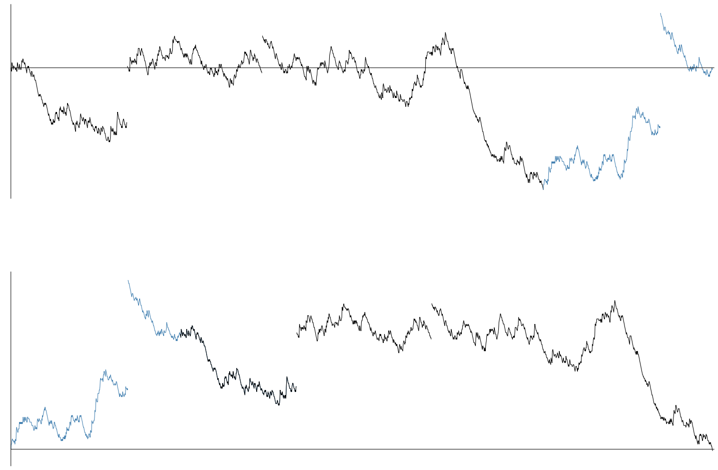

Figure 3.3: Continuous EI process (top) and its Vervaat transform (bottom). Note that if the EI process attains its infimum in a unique time, the transformed path is positive.

3. *Either* $\sigma > 0$ *or* $\sum \beta_j = \infty$.
4. $s_n / b_{s_n} \to \infty$.

Define the rescaled processes

$$X^n = \left(\frac{1}{b_{s_n}} W_{s_n} \left(\lfloor s_n t \rfloor \right), t \in [0,1] \right),$$

$$C^n = \left(\frac{1}{s_n} C_{s_n} \left(\left\lfloor \frac{s_n}{b_{s_n}} t \right\rfloor \right), t \geq 0 \right) \quad \text{and} \quad Z^n = \left(\frac{1}{b_{s_n}} Z_{s_n} \left(\left\lfloor \frac{s_n}{b_{s_n}} t \right\rfloor \right), t \geq 0 \right).$$

Then, we have the convergence

$$X^n \xrightarrow{d} X,$$

under the Skorohod topology $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$*. The limit* X *is the Vervaat transform of an EI process with parameters* $(0, \sigma, \beta)$*. Furthermore, if*

$$\int_{1/2}^{1} \frac{1}{X_s} ds < \infty \qquad a.s., \tag{3.4}$$

then, we have the joint convergence

$$(X^n, C^n, Z^n) \xrightarrow{d} (X, C, Z) \tag{3.5}$$

under the product Skorohod topology $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)^3$*. The limit* C *has an inverse* I *given by*

$$I(t) = \int_0^t \frac{1}{X(s)} ds, \quad t \in [0,1],$$

and is the unique solution to

$$C(t) = \int_0^t X \circ C(s)\, ds, \tag{3.6}$$

which is strictly increasing on $[0, I(1)]$ *if it is not the zero function. Finally,* $Z = X \circ C$.

In Figures 3.4 and 3.6 we show some simulations of trees with a given degree sequence. Figures 3.5 and 3.7 are their respective cumulative and population profiles. We also show in Figures 3.9 and 3.8 some trees depicted generation by generation, showing the population profiles.

Let us discuss the hypotheses of the theorem. Hypothesis 1 just means that the sizes of the trees we are considering go to infinity. Hypothesis 2 gives the asymptotic quantity of children of the individuals with most children (also called the hubs) and bounds the contribution of other members of the population. Hypothesis 3 guarantees that the breadth-first walks converge to a function with infinite variation so that I (and therefore C) can be non-trivial. Hypothesis 4 implies that a bigger and bigger proportion of the population is not a direct descendant of the hubs. Finally, the Hypothesis in Equation (3.4) ensures we can find a giant subtree at the top of the tree, which is used to prove C is positive. At this moment we believe this hypothesis can be removed, if one can obtain a giant subtree with an arbitrary small height.

One application of our theorem is for TGDS such that their rescaled BFW converges to the normalized Brownian excursion $\mathbf{e}$ (intuitively, a Brownian motion conditioned to be positive and hit zero for the first time at $t = 1$). For such trees, the spacial rescaling turns out to be $b_{s_n} = s_n^{1/2}$. Since $\mathbf{e}$ is continuous, we have to ensure that the maximum number of children an individual has is $o(s_n^{1/2})$.

Corollary 3.1. *Consider a sequence* $(\mathbf{s_n}, n \geq 1)$ *of degree sequences, such that*

1. $s_n \to \infty$.
2. *For*
$$\sigma_n^2 := \sum_{j \geq 1} \frac{N^n(j)}{s_n - 1} j^2 - 1,$$
we have $\sigma_n^2 \to \sigma^2 \in (0, \infty)$.
3. *The maximum degree satisfies*
$$\Delta_n := \max\{j : N^n(j) > 0\} = o(s_n^{1/2}).$$

Define the rescaled processes X^n, C^n *and* Z^n *as in Theorem 3.1, with* $b_{s_n} = s_n^{1/2}$. *Then, we have the joint convergence*

$$(X^n, C^n, Z^n) \xrightarrow{d} (\sigma \mathbf{e}, C, Z) \tag{3.7}$$

in the space $\mathscr{C}([0,1], \mathbb{R}_+^3)$, *where* C *is described as in Equation (3.6) but driven by* $\sigma \mathbf{e}$ *and* Z *is the derivative of* C.

The proof of this corollary is simple. The convergence of the rescaled BFWs is given in [BM14b, Lemma 7]. Several ways to prove (3.4) (about the integrability of $1/\mathbf{e}$ on $[0,1]$) are given in [Jan06, Remark 5.2]. Note that the previous corollary is also true if the rescaled BFW's converge to the Vervaat transform of an EI with parameters $(0, \sigma, \beta^m)$, where $\beta^m = \sum_1^m \delta_{\beta_j \in dx}$ is a finite sum of positive jumps. This is true because the integral in (3.4) is not affected by a finite number of jumps in X.

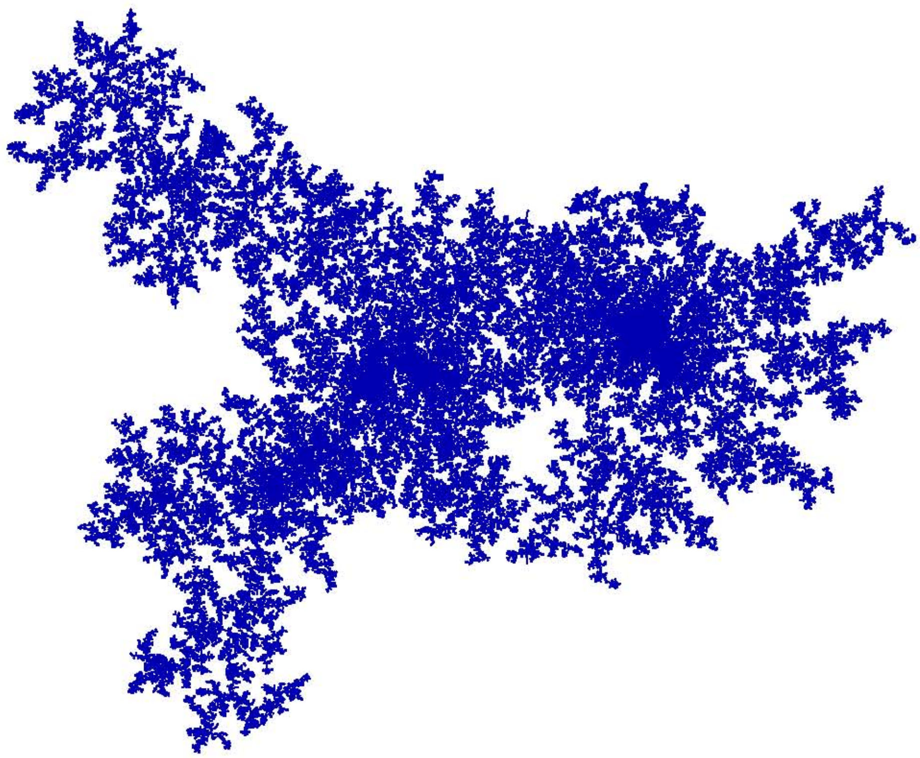

Figure 3.4: Uniformly sampled tree with a given degree sequence, approximating an EI process with zero drift, $\sigma = 2$ and $(\beta_i, i \geq 1) = (1/i, i \geq 1)$. The degree sequence is close to a geometric distribution, in the sense that $N_i \approx s_n(1/2)^{i+1}$, but also some individuals have a lot of descendants of the order $1/i$.

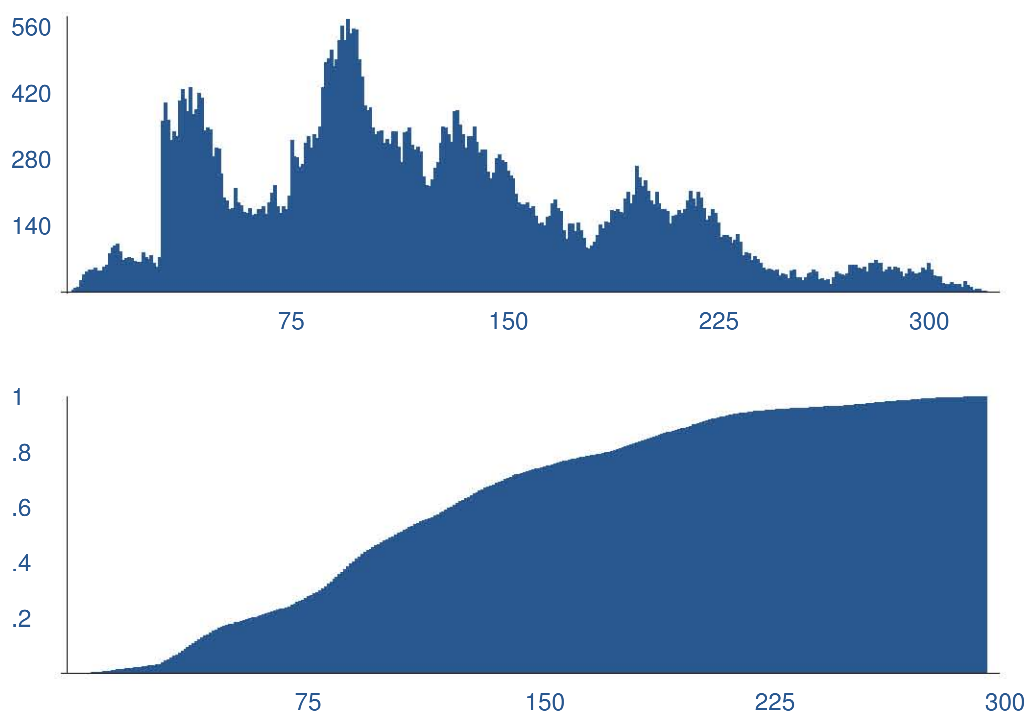

Figure 3.5: Profile and cumulative profile generated by the tree of Figure 3.4.

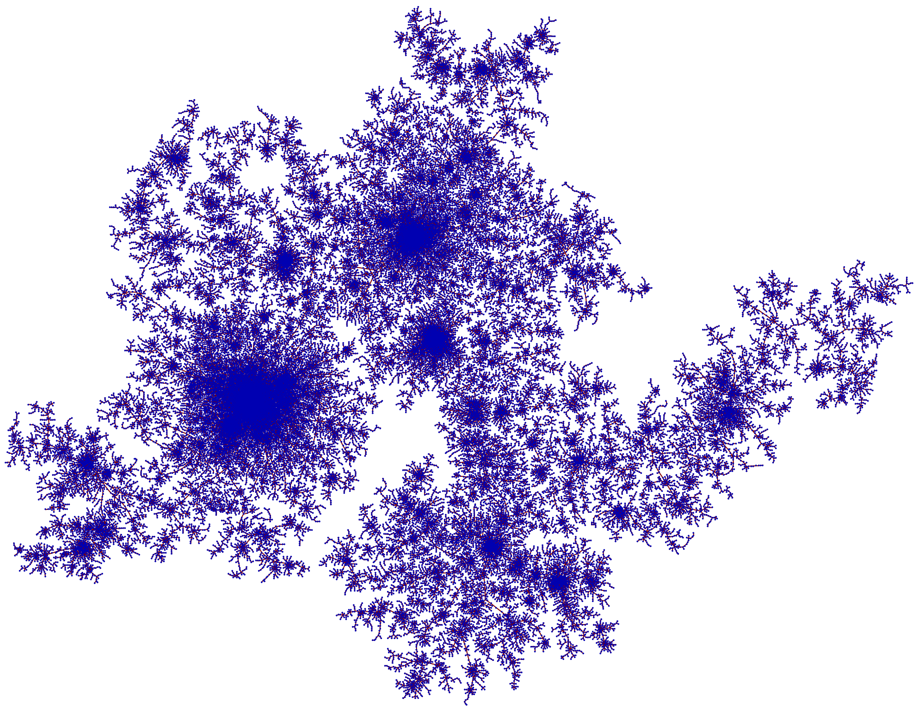

Figure 3.6: Uniformly sampled tree with a given degree sequence, approximating an EI process with zero drift, $\sigma = 0$ and $(\beta_i, i \geq 1) = (1/i, i \geq 1)$. The degree sequence is close to a Pareto distribution, in the sense that $N_i \approx s_n(1/i)^{\alpha+1}$.

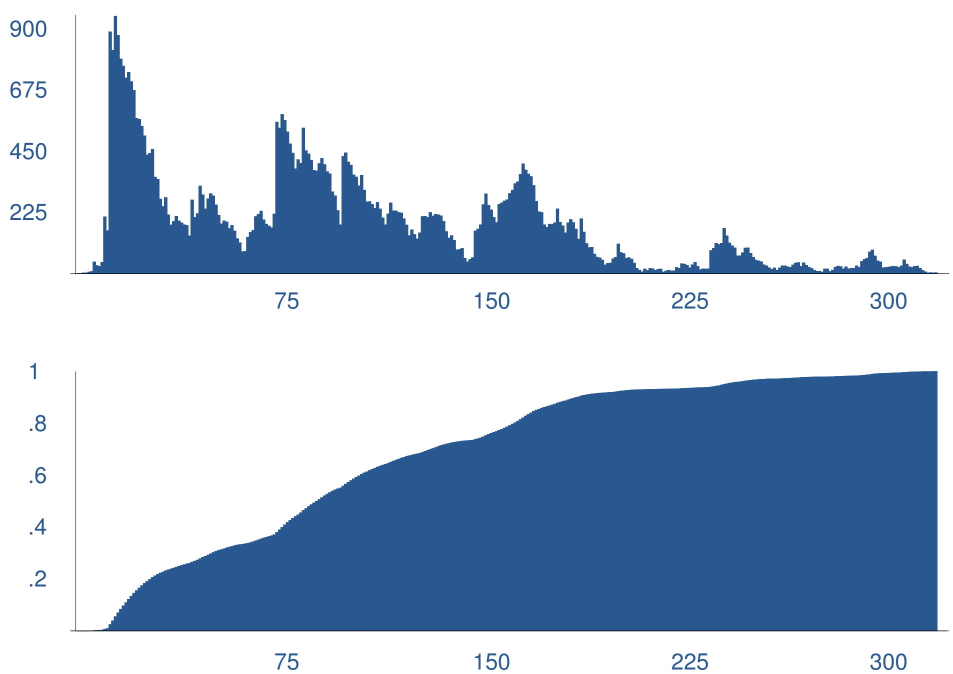

Figure 3.7: Profile and cumulative profile generated by the tree of Figure 3.6.

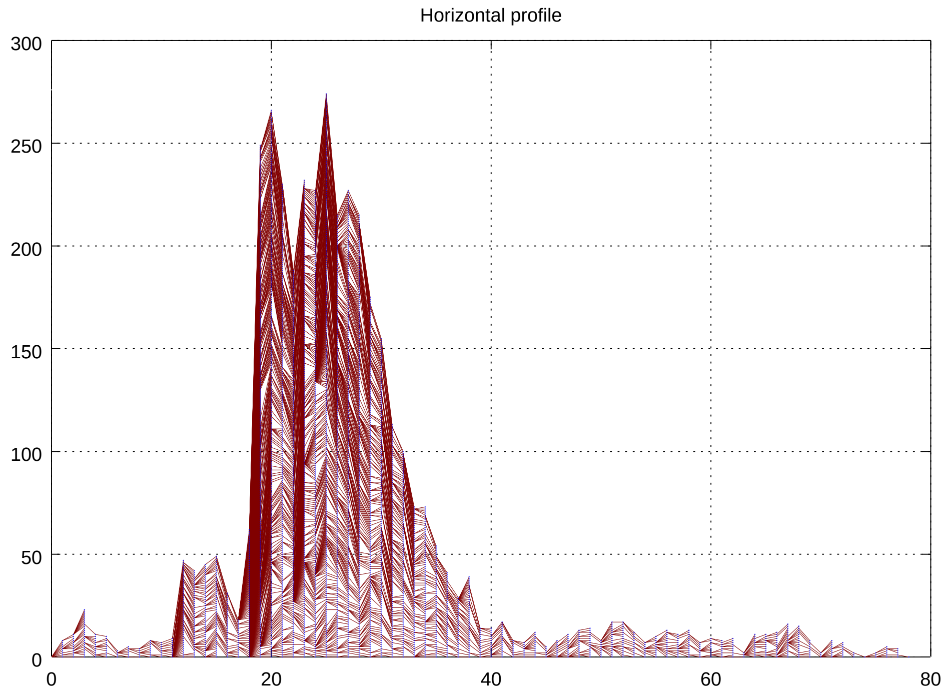

Figure 3.8: Uniformly sampled tree with a given degree sequence depicted generation by generation from left to right. The degree sequence is of the form of Figure 3.4.

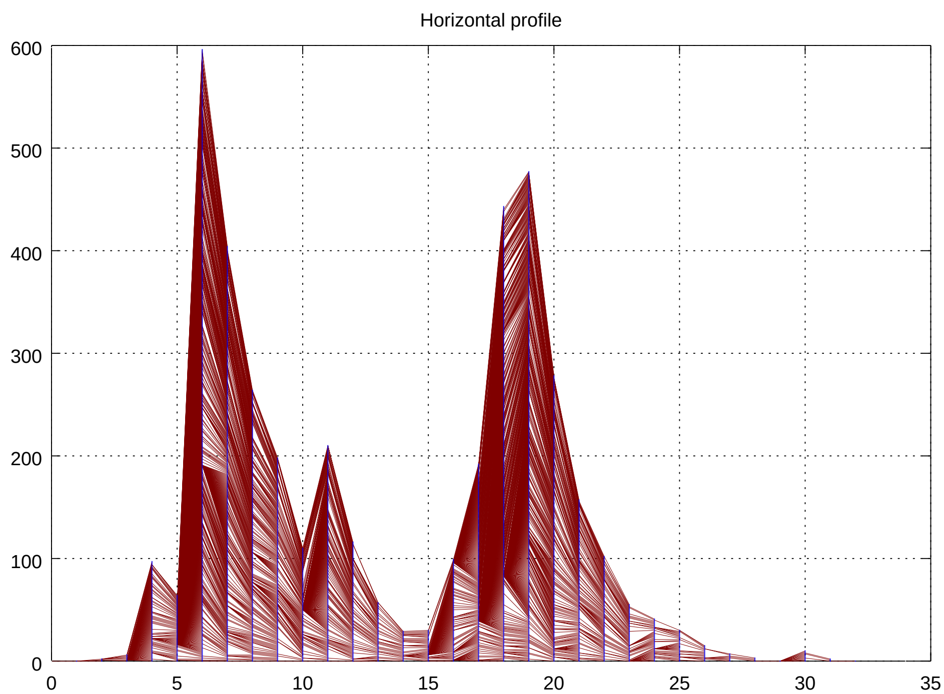

Figure 3.9: Uniformly sampled tree with a given degree sequence depicted generation by generation from left to right. The degree sequence is of the form of Figure 3.6.

Another application of our main theorem is for the convergence of the rescaled profile for CGW trees having offspring distribution of finite variance. Recall that a distribution $\mu = (\mu_n, n \geq 0)$ is called *critical* if $\sum n\mu_n = 1$, and *aperiodic* if the greatest common divisor of all n with $\mu_n > 0$ is one.

The hypotheses on μ will be the following:

$\mathbf{H}_2$ The distribution μ is critical, aperiodic and has finite variance.

Proposition 3.1. *Consider a CGW(n) tree with offspring distribution μ, satisfying hypothesis H_2. Denote by X^n, C^n and Z^n its rescaled breadth-first walk, cumulative profile and profile as in Theorem 3.1, but with s_n replaced with n and b_{s_n} replaced with $\sqrt{n}$. Then, we have the joint convergence*

$$(X^n, C^n, Z^n) \xrightarrow{d} (\sigma \mathbf{e}, C, Z)\,. \tag{3.8}$$

under the space $\mathscr{C}([0,1], \mathbb{R}_+^3)$, where C is described as in Equation (3.6) but for $\sigma\mathbf{e}$ and Z is the derivative of C.

This gives another proof of Aldous's conjecture [Ald91b], first proved in [DG97]. Actually, our method is robust enough to handle the case where μ does not have a finite variance, and belongs to the *domain of attraction of a stable law*. Kersting proved such case in [Ker11], and we recover its result as an application of our main theorem (see Section 3.5).

Let us discuss about the proof of Theorem 3.1.

Definition 3.3. *The process C defined in Theorem 3.1 is called the* cumulative Lamperti transform *of X, and its right-hand derivative (satisfying Equation (3.2)) is called the* Lamperti transform *of X.*

To explain our technique, we note that the cumulative Lamperti transform can be used to describe the possible limits in (3.5).

Proposition 3.2. *Any subsequential limit of the rescaled cumulative profile $(C^n, n \in \mathbb{N})$ satisfies (3.6), and is of the form C^Λ for some random $\Lambda \in [0,\infty]$, where*

$$C^\Lambda(t) = \begin{cases} 0 & \text{if } t \in [0,\Lambda] \\ C(t-\Lambda) & \text{if } t \in [\Lambda, \Lambda + I(1)] \\ 1 & \text{if } t \in [\Lambda + I(1), \infty), \end{cases}$$

with the convention $C^\infty \equiv 0$.

Theorem 3.1 will follow from Proposition 3.2 by showing that Λ can only be zero under Hypothesis (3.4). Indeed, we prove in Section 3.4 that

$$\mathbb{P}\left(\Lambda \in (0,\infty], \int_{1/2}^1 ds/X(s) < \infty\right) = 0.$$

If $\Lambda \in (0,\infty)$, the equality $C^\Lambda(\Lambda) = 0$ can be interpreted as an *asymptotic thickness of the base* of our random trees. To give another interpretation of this thickness, by Skorohod's representation Theorem, let $(n_k)_k$ be a subsequence such that $C^{n_k} \to C^\Lambda$ a.s. Then, for any $\lambda \in \mathbb{R}_+$ we have

$$\mathbb{P}\left(C^\Lambda(\lambda) = 0\right) = \lim_{\varepsilon \downarrow 0} \mathbb{P}\left(C^\Lambda(\lambda) < \varepsilon\right) = \lim_{\varepsilon \downarrow 0} \mathbb{P}\left(\underline{\lim}_k \{C^{n_k}(\lambda) < \varepsilon\}\right) \leq \lim_{\varepsilon \downarrow 0} \overline{\lim_k}\, \mathbb{P}(C^{n_k}(\lambda) < \varepsilon)\,.$$

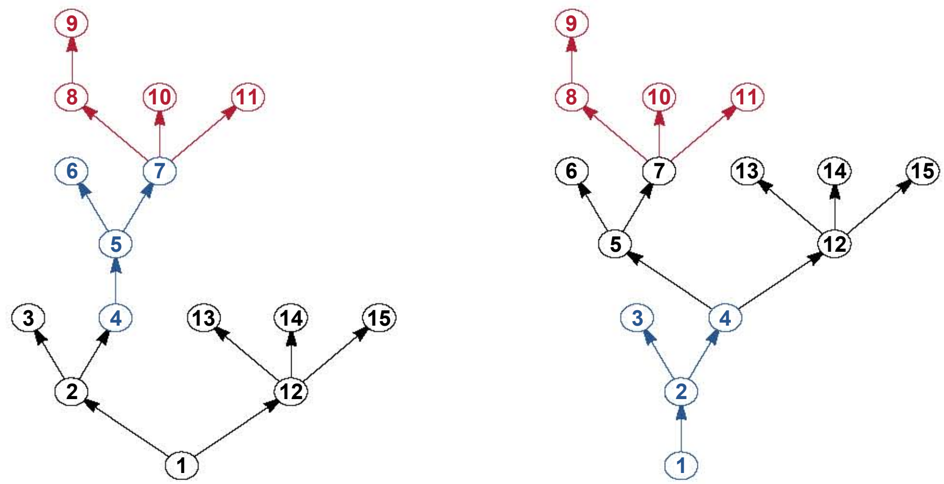

Figure 3.10: On the left, a tree labeled in depth-first order. We choose the vertex $u = 7$ and its ancestor $n = 4$. We cut the tree at vertex n and u, and interchange the subtrees with the original root and with root n. On the right, we show the transformed tree.

The double limit on the right-hand side, is what Lemma 9 of Kersting [Ker11] proves as non-thickness of the base.

A novel path transformation for discrete EI process, called the *213 transformation*, is introduced to show that, under the conditions of Theorem 3.1, our random tree sequence does not have an asymptotically thin base.

Definition 3.4 (213 Transformation)**.** *Let T be a plane tree with $|T|$ individuals. Let u be a vertex in T with ancestor $n \in \{2,\dots,|T|\}$. Cut T at n and u, obtaining three subtrees keeping their original labels: $T(1,n)$ having the original root, $T(n,u)$ with root n, and $T(u)$ with root u (if u is a leaf, then $T(u)$ is empty). Construct a new tree $\Psi_{n,u}(T)$ by grafting $T(u)$ at the leaf n of $T(1,n)$ (that is, paste the root of $T(u)$ at n), and this resulting subtree is grafted at the leaf u of $T(n,u)$. If w is the depth-first walk of T, denote by $\Psi_{n,u}(w)$ the depth-first walk of $\Psi_{n,u}(T)$.*

We obtain in this way a new tree with the same degree distribution. An example is shown in Figure 3.10.

In order to apply the 213 transformation and preserve the law $\mathbb{P}_{\mathbf{s_n}}$, we need to choose two vertices N and U in an specific way. Given a natural number $h \geq 1$, define the condition

$$U \text{ has height greater than } h. \tag{3.9}$$

Proposition 3.3. *Let W_{s_n} be the DFW of a tree with law $\mathbb{P}_{\mathbf{s_n}}$. Consider an independent uniform r.v. U on $\{2,\dots,s_n\}$, and a natural number $h \geq 1$. When (3.9) is satisfied for the Uth vertex of the tree generated by W_{s_n}, let N be the only ancestor of U at distance h, and define $\tilde{W}_{s_n} := \Psi_{N,U}(W_{s_n})$ as the 213 transformation of W_{s_n} at N and U. If (3.9) is not satisfied, set $\tilde{W}_{s_n} = W_{s_n}$. Then, we have*

$$\left(\tilde{W}_{s_n}, U\right) \stackrel{d}{=} \left(W_{s_n}, U\right).$$

Consider an arbitrary $\lambda \in \mathbb{R}$ and a sequence of trees τ_n with law $\mathbb{P}_{\mathbf{s_n}}$. The idea to prove the main theorem is: first we prove that on the event $\{\Lambda \in (\lambda,\infty), I(1)-I(1/2)<\infty\}$, there exists a *giant subtree* at the top of τ_n and of height smaller than $\lambda_1 < \lambda/2$; this subtree has many individuals, almost all of them *before* height λ_1 from its root (up to scaling factors). Apply the 213 transformation when the uniform variable U falls in the subtree, with N being its ancestor at distance λ_1. The transformed subtree will be fat *before* height $2\lambda_1$, thin *between* heights $\{2\lambda_1,\dots,\lambda\}$ and fat *after* height λ, which has probability going to zero. In Figure 3.11 we give an example of such transformed tree.

3.1.1 Plan of the paper

In Section 3.2 we focus on the BFW of uniform trees with a given degree sequence. The 213 transformation and the asymptotic non-thickness of the base of trees are analyzed in Section 3.4. We use those results, however, in Section 3.3 to prove our main theorem. Section 3.5 is devoted to the convergence of the rescaled (cumulative) profile for CGW(n) trees with offspring distribution in the domain of attraction of a stable law.

3.2 Convergence of the BFW to the Vervaat transform of an EI process

Consider a plane tree T, and let $c(u)$ be the number of children of the uth vertex in T. The degree sequence of T is obtained from the sequence $(c(u), u\in[s_n])$ as follows:

$$n_i(T)=n_i=|\{u\in V(T):c(u)=i\}|.$$

Recall that $\mathbb{T}_{\mathbf{s}}$ is the set of plane rooted trees with a given degree sequence $\mathbf{s}=(n_i,i\geq 0)$ and that $\mathbb{P}_{\mathbf{s}}$ is the uniform distribution on $\mathbb{T}_{\mathbf{s}}$. If $u_1,\dots,u_p$ are the vertices of T labeled increasingly in the breadth-first (or depth-first) order, define the breadth-first (resp depth-first) walk (abridged BFW or DFW) of T by x, where $x_0=1$ and $x_n=1+c(u_1)+\cdots c(u_n)-n$. Its increments will be denoted $k_1,\dots,k_p$. It is known that $T\mapsto(k_1,\dots,k_p)$ is a bijection between the the set of trees with p individuals and the set of sequences

$$E=\left\{\mathbf{K}=(k_1,\dots,k_p):k_i\in\mathbb{N}\cup\{-1\},\sum_1^{|\mathbf{s}|}k_i=0,\sum_1^{j}k_i>0\ \text{ if } j<|\mathbf{s}|\right\},$$

see Lemma 6.3 of [Pit06] or Proposition 1.1 of [LG05]. For $i\geq 0$, let $\tilde{n}_i=\tilde{n}_i(\mathbf{K})=|\{j\in[|\mathbf{s}|]:k_j=i-1\}|$ be the number of increments of the excursion which are equal to $i-1$, and call $(\tilde{n}_i,i\in\mathbb{N})$ the *degree sequence* of the excursion $\mathbf{K}$.

The BFW (DFW) of a tree τ with law $\mathbb{P}_{\mathbf{s}}$ is related with a discrete time EI process.

Definition 3.5. *A discrete time process $(W^b(j),0\leq j\leq|\mathbf{s}|)$ with increments $\Delta W^b(i)=W^b(i)-W^b(i-1)$ has exchangeable increments (EI) if for every permutation σ on $[|\mathbf{s}|]$*

$$\left(\Delta W^b(1),\dots,\Delta W^b(|\mathbf{s}|)\right)\overset{d}{=}\left(\Delta W^b(\sigma_1),\dots,\Delta W^b(\sigma_{|\mathbf{s}|})\right).$$

Write $\theta_i(W^b) = W^{b,(i)}$ for the *cyclic shift* of W^b at i, that is, the sequence of length $|\mathbf{s}|$ whose jth increment is $\Delta W^b(i+j)$ with $i+j$ interpreted mod $|\mathbf{s}|$. A path transformation, introduced by Vervaat in [Ver79], is used to code discrete random trees from EI processes. The *discrete Vervaat transform* of W^b, denoted by $V(W^b)$, is the i^*th cyclic shift of W^b, where $i^* = \min\{i \in [\mathbf{s}] : W^b(i) = \min_{j\in[\mathbf{s}]} W(j)\}$ is the first minimum of W^b.

The next proposition is an easy consequence of the definitions (cf. the proof of Lemma 7 in [BM14b]).

Proposition 3.4. *Let W be the BFW (DFW) of a tree with law $\mathbb{P}_{\mathbf{s}}$. and U an independent and uniform random variable on $[|\mathbf{s}|]$. Then, the cyclic shift of W at U is a discrete time EI process.*

Since the BFW and the DFW of a tree have the same increments, but in different order, we deduce the following.

Corollary 3.2. *The breadth-first walk and the depth-first walk of a tree with law $\mathbb{P}_{\mathbf{s}}$, have the same distribution.*

Consider a sequence of degree sequences $(\mathbf{s_n})_{n\geq 1} = (N_i^n, i \geq 0)_{n\geq 1}$ and define $s_n = |\mathbf{s_n}|$. Using a uniform permutation π_n on $[s_n]$ and a child sequence $(c(j), j \in [s_n])$, construct the process

$$W_{s_n}^b(\lfloor s_n t\rfloor) = 1 + \sum_{j\leq \lfloor s_n t\rfloor} (c(\pi_n(j)) - 1), \ \ t \in [0,1],$$

where b stands for *bridge*. It has exchangeable increments on $\{k/s_n : 0 \leq k \leq s_n\}$.

Now, we analyze the convergence of $W_{s_n}^b$ to an EI process. We introduce the notation for convergence in the Skorohod topology, obtained from Chapter 3 of [Bil99]. Throughout this chapter, by an increasing function we mean a strictly increasing function. For càdlàg functions $f^n, f \in D[0,1]$ (the space of càdlàg functions with domain $[0,1]$), the functions f^n converge to f in the Skorohod topology, denoted by $f^n \to f$, if there exists increasing bijections $\alpha^n : [0,1] \mapsto [0,1]$ such that

$$||\alpha^n - \mathrm{Id}|| \to 0 \quad \text{and} \quad ||f^n - f\circ\alpha^n|| \to 0,$$

where $||f|| = \sup_{t\in[0,1]} |f_t|$ is the uniform norm, and Id the identity function.

Proposition 3.5. *Suppose the sequence of degree sequences $(\mathbf{s_n}, n \geq 1)$ satisfies Hypotheses (1) and (2) of Theorem 3.1. Then we have*

$$\left(\frac{1}{b_{s_n}} W_{s_n}^b(\lfloor s_n t\rfloor), t \in [0,1]\right) \xrightarrow{d} \left(X_t^b, t \in [0,1]\right)$$

in the Skorohod topology, where X^b an EI process on $[0,1]$, having parameters $(0,\sigma,\beta)$.

Proof. Define $\xi_j = (c(\pi_n(j)) - 1)/b_{s_n}$ and $\Delta_{s_n} = \max\{i : N_i^n > 0\}$. By Theorem 2.2 of [Kal73], we know that

$$W_{s_n}^b(\lfloor s_n \cdot\rfloor)/b_{s_n} \xrightarrow{d} X^b \text{ if and only if } \left(\sum \xi_j, \sum \xi_j^2, \tilde{\xi}_1, \tilde{\xi}_2, \ldots\right) \xrightarrow{d} (\alpha, \sigma^2 + \sum \beta_j^2, \beta_1, \beta_2, \ldots)$$

where $(\tilde{\xi}_j, j \in [s_n])$ are the $(\xi_j, j \in [s_n])$ in decreasing order.

The convergence to the drift α is due to the fact that $W^b_{s_n}(s_n) = 0$. For the sum of the squares, using the definition of a child sequence

$$\sum \xi_j^2 = \sum (c(j)-1)^2/b_{s_n}^2 = \sum_0^{\Delta_{s_n}} \frac{(j-1)^2 N_j^n}{b_{s_n}^2} \to \sigma^2 + \sum_1^M \beta_j^2.$$

Finally, by hypothesis $\tilde{\xi}_j \to \beta_j$ for every $j \in [M]$. □

Remark 3.1. *It should be obvious that we can modify hypotheses (1) and (2) of Theorem 3.1 by $s_n \to \infty$, $\Delta_{s_n} = o(b_{s_n})$ and $\sum (j-1)^2 N_j^n / b_{s_n}^2 \to \sigma^2 \in (0,\infty)$. In this case, we obtain in the limit the Brownian bridge on $[0,1]$, as in [BM14b].*

From Proposition 3.4 we deduce that the discrete Vervaat transform of $W^b_{s_n}$ is the BFW of a tree with law $\mathbb{P}_{\mathbf{s_n}}$. Denote this BFW by $W_{s_n} = V(W^b_{s_n})$.

Proposition 3.6. *Under the hypothesis of Proposition 3.5, assume also that either $\sigma > 0$ or $\sum \beta_j = \infty$. Then*

$$X^n = \left(\frac{1}{b_{s_n}} W_{s_n}(\lfloor s_n t \rfloor), t \in [0,1]\right) \xrightarrow{d} V(X^b) = X, \tag{3.10}$$

the Vervaat transform of X^b.

Proof. By the proof of Lemma 6 of [Ber01], the process X^b hits its infimum in a unique time and continuously. Hence, by Lemma 3 of [Ber01] (or Lemma 14 of [Ker11]) the Vervaat transform X^n of the rescaled bridges $W^b_{s_n}$ converges to X, the Vervaat transform of X^b. □

3.3 Convergence of the profile to the Lamperti transform

The objective of this section is to prove our main theorem on convergence of profiles stated as Theorem 3.1 by using Theorem 3.2 from Section 3.4. First, we study the effect of scaling on the discrete Lamperti transformation in Subsection 3.3.1. Then, we analyze the subsequential limits of the cummulative profile in Subsection 3.3.2 and of the profile in Subsection 3.3.3. Finally, the proof of Theorem 3.1 is obtained in Subsection 3.3.4.

3.3.1 Rescaling the functional relation of the BFW and the cumulative profile

Recall that Z_{s_n}, C_{s_n} and W_{s_n} are the profile, the cumulative profile and the BFW of a tree with law $\mathbb{P}_{s_n}$, respectively. Define the scaling operators S_a^b acting on the Skorohod space by $S_a^b f(t) = f(at)/b$. From Equation (3.1) the discrete Lamperti transform satisfies for any $j \geq 0$

$$Z_{s_n}(j+1) = W_{s_n} \circ C_{s_n}(j). \tag{3.11}$$

Extend C_{s_n} to $\mathbb{R}_+$ by linear interpolation, and consider the sequence

$$a_n = \frac{s_n}{b_{s_n}} \qquad n \in \mathbb{N},$$

which is the temporal scaling of the cumulative profile. This is true because, by a simple change of variables

$$
\begin{aligned}
S^{s_n}_{a_n} C_{s_n}(t) = \frac{1}{s_n} C_{s_n}(a_n t) &= \frac{1}{s_n} + \sum_{j=0}^{\lfloor a_n t \rfloor} \frac{1}{s_n} W_{s_n} \circ C_{s_n}(j) + (a_n t - \lfloor a_n t \rfloor) \frac{1}{s_n} W_{s_n} \circ C_{s_n}(\lfloor a_n t \rfloor) \\
&= \frac{1}{s_n} + \int_0^{a_n t} \frac{1}{s_n} W_{s_n} \circ C_{s_n}(\lfloor u \rfloor)\, du \\
&= \frac{1}{s_n} + \int_0^{t} \frac{a_n}{s_n} W_{s_n} \circ s_n \frac{1}{s_n} C_{s_n}(\lfloor a_n u \rfloor)\, du \\
&= \frac{1}{s_n} + \int_0^{t} S^{b_n}_{s_n} W_{s_n} \circ S^{s_n}_{a_n} C_{s_n}(\lfloor a_n u \rfloor / a_n)\, du.
\end{aligned}
$$

We conclude that the scaling of the cumulative profile satisfies an equation similar to (3.6). This will be the basis of our convergence analysis.

Define $C^n(0) = 0$, and for $i \geq 0$ and $t \in [t_i, t_{i+1})$, with $t_i = i/a_n$, write

$$
C^n(t) := \int_0^t X^n \circ C^n(\lfloor a_n u \rfloor / a_n) du = C^n(t_i) + (t - t_i) X^n \circ C^n(t_i), \tag{3.12}
$$

with $X^n = S^{b_n}_{s_n} W_{s_n}$. In the notation of [CPGUB13], the above equation is $\mathrm{IVP}_{1/a_n}(X^n, 0)$, which has a unique solution (see [CPGUB13, page 1594]). Thus, we have

$$
C^n(1/a_n) = \frac{1}{a_n} X^n(0) = \frac{1}{s_n} = S^{s_n}_{a_n} C_{s_n}(0/a_n),
$$

and for $i \geq 1$ and $t \in [t_i, t_{i+1})$, by induction

$$
\begin{aligned}
C^n(t) &= C^n(t_i) + (t - t_i) X^n \circ C^n(t_i) \\
&= S^{s_n}_{a_n} C_{s_n}(t_{i-1}) + (t - t_i) X^n \circ S^{s_n}_{a_n} C_{s_n}(t_{i-1}) \\
&= S^{s_n}_{a_n} C_{s_n}(t_{i-1}) + (t - 1/a_n - t_{i-1}) X^n \circ S^{s_n}_{a_n} C_{s_n}(t_{i-1}) \\
&= S^{s_n}_{a_n} C_{s_n}(t - 1/a_n),
\end{aligned}
$$

and for $t \in [0, t_1)$ note that $C^n(t) = t X^n(0)$. Therefore, to prove the convergence of the rescaled cumulative profile, it is enough to prove the convergence of C^n. The rescaled profile is defined by $Z^n = D_+ C^n$. First we give conditions to ensure Hypothesis 2 of Theorem 3.1, that is $a_n \to \infty$.

Lemma 3.1. *Assume either one of the following is true*

$$
\sum \beta_j = \infty \qquad \text{or} \qquad \frac{\sum_{i<j} c^n_i c^n_j}{\sum (c^n_i)^2} \to \infty.
$$

Then $a_n \to \infty$.

Proof. Order the child sequence in decreasing order as $c^n_{(1)} \geq c^n_{(2)} \geq \cdots$ and fix any $k \in \mathbb{N}$. The Hypothesis 2 of Theorem 3.1 implies that for n big enough

$$
\frac{c^n_{(i)}}{b_{s_n}} \geq \frac{1}{2} \beta_i \quad \text{for every } i \in [k].
$$

Since $s_n = 1 + \sum c^n_{(i)}$, this implies

$$\varliminf_n \frac{s_n}{b_{s_n}} \geq \varliminf_n \frac{1+\sum_1^k c^n_{(i)}}{b_{s_n}} \geq \frac{1}{2}\sum_1^k \beta_j.$$

This being true for any k, the conclusion follows when $\sum \beta_j = \infty$. Now assume the other hypothesis of the lemma is true. By Hypothesis 2 of Theorem 3.1, we have

$$\sum \frac{(c^n_i - 1)^2}{b^2_{s_n}} = \sum \frac{(c^n_i)^2 - s_n + 2}{b^2_{s_n}} \to \sigma^2 + \sum \beta_j^2.$$

This implies

$$b_{s_n} \sim \left(\sum (c^n_i)^2 - s_n\right)^{1/2}.$$

Therefore

$$\frac{s_n^2}{b^2_{s_n}} \sim \frac{(\sum c^n_i)^2}{\sum (c^n_i)^2 - s_n}.$$

Therefore, a_n goes to infinity whenever

$$\frac{s_n + \sum_{i<j} c^n_i c^n_j}{\sum (c^n_i)^2} \to \infty. \qquad \square$$

3.3.2 All subsequential limits of the cumulative profile satisfy the IVP

We introduce the functions f which play the role of the sample paths of X.

Definition 3.6. *An* admissible breadth-first function *is a non-negative function f, which is càdlàg without negative jumps, starting at a non-negative value and with absorption time $\zeta = \zeta(f) = \inf\{t > 0 : f(t) = 0\} \in (0, \infty)$.*

For admissible breadth-first functions f, the initial value problem IVP(f) is defined as

$$D_+ c = f \circ c \text{ and } c(0) = 0.$$

For $\lambda \in [0, \infty]$, define the function i^λ by

$$i^\lambda(t) = \lambda + \int_0^t \frac{dr}{f(r)} \qquad t \in [0, \zeta],$$

and the value

$$\lambda_{0+} := i^\lambda(0+) = \inf\{t > 0 : c^\lambda(t) > 0\}$$

where c^λ is the right-continuous inverse of i^λ, with $c^\lambda \equiv 0$ if either $\lambda = \infty$ or $\int_0^t dr/f(r) = \infty$ for some $t \in (0, \zeta)$. By convention we write $i := i^0$ and $c := c^0$. It is known the inverse c of i satisfies IVP(f). This is a solution which is identically zero if $i(t) = \infty$, or immediately becomes positive when $i(t) < \infty$, for some $t \in (0, \zeta)$. The function c is the cumulative Lamperti transform of f, and plays an important role in the next proposition.

For an admissible breadth-first function f define the functional inequality

$$\int_s^t f_- \circ c(r)\,dr \le c(t) - c(s) \le \int_s^t f \circ c(r)\,dr \qquad s \le t. \tag{3.13}$$

Note that any solution to this inequality is continuous and non-decreasing on $[0,\infty)$, because f_- is a non-negative function. Also, any solution to IVP(f) satisfies (3.13), so we have always a solution (as noted after Equation (5) in [CPGUB13]). In fact, all the solutions are characterized as follows.

Proposition 3.7. *For $\lambda \in [0,\infty)$ define the function*

$$c^\lambda(t) = \begin{cases} 0 & \text{if } t \in [0,\lambda] \\ c(t-\lambda) & \text{if } t \in [\lambda, \lambda + i(\zeta)) \\ \zeta & \text{if } t \in [\lambda + i(\zeta), \infty). \end{cases}$$

Define also $c^\lambda \equiv 0$ if $\lambda = \infty$ or $\int_0^t dr/f(r) = \infty$ for some $t \in (0,\zeta]$. Then, all the solutions of (3.13) are of the form c^λ. In particular, when $i(t) = \infty$ for $t \in (0,\zeta)$, the unique solution is the zero function.

Proof. Assume there exists a non-zero solution d to (3.13) and define

$$\lambda = \inf\{t > 0 : d(t) > 0\}.$$

Then, for $t > 0$

$$\int_\lambda^{\lambda+t} f_- \circ d(r)\,dr \le d(\lambda+t) - d(\lambda) \le \int_\lambda^{\lambda+t} f \circ d(r)\,dr.$$

The function $d_\lambda(\cdot) = d(\lambda + \cdot)$ satisfies (3.13) and is positive on the interval $(0,\varepsilon)$, for some $\varepsilon > 0$. We now show that $d_\lambda \equiv c$, which proves the proposition. To ease notation, we write d instead of d_λ.

Let $\tau(d) = \inf\{t > 0 : d(t) = \zeta\}$. Then d is constant on $(\tau(d),\infty)$ because f is absorbed at time ζ. By definition, d is positive on the interval $(0,\varepsilon')$ for some $\varepsilon' > 0$. Consider $\varepsilon = \tau(d) \wedge \varepsilon'$. The left-hand side of (3.13) is positive on $(0,\varepsilon)$ and therefore d is strictly increasing on this interval. So, we have equalities in (3.13). Hence, d is a solution to IVP(f) which is positive on $(0,\infty)$.

Let $\hat{i}$ be the inverse of d on $[0,\zeta)$. On this interval, $\hat{i}$ is increasing, continuous, and with values on $[0,\infty)$. Let $0 < r < \zeta$. From the definition of the IVP, the right-hand derivative of d evaluated at $\hat{i}(r)$ is equal to $f \circ d(\hat{i}(r)) = f(r) > 0$. Hence, by the formula for the derivative of the inverse function and the fundamental theorem of calculus (for Riemann integrable derivatives)

$$\infty > \hat{i}(t) - \hat{i}(s) = \int_s^t D_+\hat{i}(r)\,dr = \int_s^t \frac{dr}{f(r)} = i(t) - i(s) \qquad 0 < s < t < \zeta.$$

Because d is continuous at 0, then $\hat{i}(s) \to 0$ as $s \downarrow 0$. Therefore, we obtain

$$\int_0^t \frac{dr}{f(r)} = \hat{i}(t) < \infty \qquad \forall t \in [0,\zeta). \tag{3.14}$$

It remains to prove that $\tau(d) = i(\zeta)$. But this is clear from (3.14), because $\lim_{t\uparrow\zeta} \hat{i} = \lim_{t\uparrow\zeta} i$. This implies that $d \equiv c$.

Finally, when $i(t) = \infty$ for $t \in (0,\zeta)$, the unique solution is the zero function because, a positive solution would imply by (3.14), that $i(t) < \infty$ for $t \in [0,\zeta)$. $\square$

Remark 3.2. *Note that c^{λ} satisfies IVP(f) for every λ.*

As in the previous section, we denote by $\alpha^n : [0,1] \mapsto [0,1]$ the increasing bijections required for the convergence on $[0,1]$ of f^n to f. For the next result, we introduce the definition of convergence in the space $D(\mathbb{R}_+,\mathbb{R})$ of càdlàg functions with domain $\mathbb{R}_+$ taking values on $\mathbb{R}$, under the Skorohod topology. This is obtained from Theorem 16.2 Chapter 3 of [Bil99] (cf. Section 2 of [Whi80]). Consider functions $g^n, g \in D(\mathbb{R}_+,\mathbb{R})$. For $0 < v < \infty$, denote by $||g||_v = \sup_{r\leq v} |g(r)|$ the uniform norm of g on $[0,v]$. More generally, for $0 \leq u \leq v < \infty$ we write $||g||_{[u,v]} = \sup_{u\leq r\leq v} |g(r)|$. The functions g^n converge to g in the Skorohod topology, denoted by $g^n \to g$, if there exists increasing bijections $\beta^n : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that for every continuity point $v < \infty$ of g, we have

$$||\beta^n - \mathrm{Id}\,||_v \to 0 \quad \text{and} \quad ||g^n - g\circ\beta^n||_v \to 0. \tag{3.15}$$

The functions $(\beta^n, n\in\mathbb{N})$ can be chosen as the identity when g is continuous. Note that the space $D(\mathbb{R}_+,\mathbb{R})$ is not complete under the metric we are defining. Nevertheless, Theorem 12.1 of [Bil99] says this metric is equivalent to some d^0 such that $(D(\mathbb{R}_+,\mathbb{R}), d^0)$ is a complete metric space.

Define, for a random variable $\Lambda \in [0,\infty]$,

$$I^{\Lambda}(t) = \Lambda + \int_0^t \frac{ds}{X(s)} \quad t\in[0,1] \tag{3.16}$$

with the conventions $I^{\Lambda} \equiv \infty$ on $\{\Lambda = \infty\} \cup \{\int_0^t ds/X(s) = \infty\}$ for some $t \in (0,1]$, and $I := I^0$, with X the Vervaat transform of an EI process on $[0,1]$. Define C^{Λ} as the *right-continuous inverse* of I^{Λ}, and $\Lambda_{0+} = \inf\{t > 0 : C^{\Lambda}(t) > 0\}$. Thus, C^{Λ} is defined as c^{λ} from Proposition 3.7, but using the cumulative Lamperti transform C of X. For the next result, we use Proposition 3.7 and a simplified version of the proof given in Theorem 3 of [CPGUB13] (also, compare with Theorem 1.5, Chapter 6 of [EK86]).

Proposition 3.8. *Under the assumptions of Theorem 3.1, let $(X^n, n\geq 1)$ be the rescaled breadth-first walks of trees $(\tau_n, n \geq 1)$ with law $\mathbb{P}_{\mathbf{s_n}}$, as in (3.10). Let X be the limit of $(X^n, n\in\mathbb{N})$. Consider the rescaled cumulative profile C^n of τ_n. Then, $(C^n, n\geq 1)$ is sequentially compact for convergence in distribution uniformly on compact sets, and every subsequential limit of $((X^n,C^n), n \geq 1)$ is of the form (X,C) where C satisfies $C_t = \int_0^t X\circ C_s\, ds$.*

Proof. We prove tightness of $(C^n, n\in\mathbb{N})$, which together with the tightness of $(X^n, n\in\mathbb{N})$, implies tightness of $((X^n,C^n), n\in\mathbb{N})$. Recall that $0 \leq C^n \leq 1$, so the sequence $(C^n, n\in\mathbb{N})$ is uniformly bounded. Note that

$$0 \leq D_+C^n(s) = X^n \circ C^n(\lfloor a_n s\rfloor/a_n) \leq \|X^n\|$$

and so the modulus of continuity

$$\omega_n(\delta) = \sup\{|C^n(t) - C^n(s)| : |t-s| \leq \delta\}$$

of C^n satisfies $\omega_n(\delta) \leq \|X^n\|\delta$. Therefore

$$\mathbb{P}(\omega_n(\delta) > \varepsilon) \leq \mathbb{P}(\|X^n\| > \varepsilon/\delta)\,.$$

Using Theorem 13.2 in [Bil99], the right-hand side can be made as small as we want uniformly in n if δ is small enough. Hence $((X^n,C^n), n\in\mathbb{N})$ is tight.

Suppose that (X,D) is a subsequential limit of $((X^{n_l},C^{n_l}),l\in\mathbb{N})$. By Skorohod's theorem, we assume the convergence takes place almost surely. Suppose we have proved that for any $T>0$

$$X_-\circ D\le\liminf_l X^{n_l}\circ C^{n_l}\le\limsup_l X^{n_l}\circ C^{n_l}\le X\circ D\quad \text{on } [0,T]. \tag{3.17}$$

Then, using Fatou's lemma, for $0\le s<t\le T$

$$\int_s^t X_-\circ D(r)\,dr\le D(t)-D(s)\le\int_s^t X\circ D(r)\,dr. \tag{3.18}$$

It remains to prove (3.17). We start with the equality

$$X^{n_l}\circ C^{n_l}=(X^{n_l}\circ C^{n_l}-X\circ\alpha^{n_l}\circ C^{n_l})+X\circ\alpha^{n_l}\circ C^{n_l}.$$

The difference in parenthesis is bounded above by $||X^{n_l}-X\circ\alpha^{n_l}||$, which goes to zero. The convergence of $\alpha^{n_l}\circ C^{n_l}\to D$ follows by adding the terms $\pm C^{n_l}$:

$$\sup_{u\le v}|\alpha^{n_l}\circ C^{n_l}(\lfloor a_nu\rfloor/a_n)-D(u)|\le||\alpha^{n_l}-\mathrm{Id}||+\sup_{u\le v}|C^{n_l}(\lfloor a_nu\rfloor/a_n)-D(u)|\to 0. \tag{3.19}$$

Then, because X is càdlàg and has only positive jumps

$$X_-\circ D\le\liminf_l X\circ\alpha^{n_l}\circ C^{n_l}\le\limsup_l X\circ\alpha^{n_l}\circ C^{n_l}\le X\circ D. \qquad \square$$

3.3.3 Obtaining all subsequential limits of the profile

The next step is to obtain all subsequential limits of $X^n\circ C^n$. We need some definitions.

Let c be a non-negative, non-decreasing continuous function $c:\mathbb{R}_+\mapsto[0,1]$. Denote by b and i its left and right-continuous inverses, respectively.

Definition 3.7. *An* admissible cumulative profile function *is a non-negative continuous function* $c:[0,\infty)\mapsto[0,1]$*, such that*

1. *is zero on the interval* $[0,i(0)]$*,*
2. *is strictly positive on* $[i(0),b(1)]$*,*
3. *stays at one on* $[b(1),\infty)$*.*

Note that in the definition, we allow $i(0)=\infty$ or $b(1)=\infty$. If we consider $\lambda:=i(0)$, we can also denote $c=c^\lambda$ to express the dependence on λ, and analogously $i^\lambda:=i$ for its right-continuous inverse.

Lemma 3.2. *Consider a sequence of non-negative càdlàg functions* $(f^n,n\in\mathbb{N})$*, starting at a non-negative value and with absorption time* $\zeta(f^n)=1$*. Let* f *be an admissible breadth-first function with absorption time* $\zeta(f)=1$*. Also, consider admissible cumulative profile functions* $(c^n,n\in\mathbb{N})$ *and* c^λ*, where* $c^n\not\equiv 0$*. Suppose that* $f^n\to f$ *in the Skorohod topology and* $c^n\to c^\lambda$ *uniformly on compact sets. Then*

$$f^n\circ c^n\overset{a.s.}{\to}f\circ c^\lambda$$

in the Skorohod topology.

Remark 3.3. *Lemma 3.2 could be proved as in Theorem 3 of [CPGUB13], using Theorem 1.2 of [Wu08]. Indeed the conditions of the latter hold since f is continuous at 0 and 1, which are the only possible discontinuities of i^λ. For completeness, we include the following proof, adapted from the proof of Lemma 3 of [Ker11].*

Proof. First suppose that $\lambda < \infty$. Let $0 < u < v < r$ be continuity points of $f \circ c^\lambda$. Using Lemma 2.2 of [Whi80], convergence of $f^n \circ c^n$ on $[0,r]$ follows from its convergence on each of the subintervals $[0,u]$, $[u,v]$ and $[v,r]$.

The simplest is the middle interval whenever $0 < i^\lambda(\varepsilon) = u < v = i^\lambda(1-\varepsilon)$ for some $\varepsilon \in (0,1/2)$ where ε and $1-\varepsilon$ are continuity points of f. Indeed, the definition of admissible cumulative profile function implies that for n large enough, c^n and c^λ are strictly increasing on $[u,v]$. If $\|f^n \circ \alpha^n - f\|_{[\varepsilon,1-\varepsilon]} \to 0$ and $(\alpha^n, n \in \mathbb{N})$ are increasing homeomorphisms on $[\varepsilon, 1-\varepsilon]$ converging uniformly to the identity, we can define $\beta^n = i^\lambda \circ \alpha^n \circ c^n$ to obtain

$$\|i^\lambda \circ \alpha^n \circ c^n - \mathrm{Id}\|_{[u,v]} \to 0$$

because $i^\lambda \circ \alpha^n \circ c^n \to i^\lambda \circ c^\lambda = \mathrm{Id}$ on $[u,v]$. Also,

$$\|f^n \circ c^n - f \circ c^\lambda \circ \beta^n\|_{[u,v]} = \|f^n \circ c^n - f \circ \alpha^n \circ c^n\|_{[u,v]} = \|f^n - f \circ \alpha^n\|_{[c^n(u), c^n(v)]}$$

The right-hand side goes to zero because $c^n(u) \to \varepsilon$, $c^n(v) \to 1-\varepsilon$, and both limits are continuity points of f.

Now, we choose v and r such that $||f^n \circ c^n - f \circ c^\lambda||_{[v,r]}$ is as small as we want. Let $\tau^\downarrow(\varepsilon) = \inf\{t : ||f||_{[t,1]} < \varepsilon\}$. For $\varepsilon \in (0,1/2)$, set $\tilde{v} = \tau^\downarrow(\varepsilon)$, $v = i^\lambda(\tau^\downarrow(\varepsilon))$ and choose any $r > v$. Since $\|f\|_{[\tilde{v},1]} \le \varepsilon$ then $\|f^n\|_{[\tilde{v},1]} \le 2\varepsilon$ for n large enough. Then

$$\|f \circ c^\lambda\|_{[v,r]} = \|f\|_{[c^\lambda(v), c^\lambda(r)]} \le \|f\|_{[\tilde{v},1]} \le \varepsilon \quad \text{and} \quad \|f^n \circ c^n\|_{[v,r]} \le \|f^n\|_{[c^n(v),1]}.$$

Also, since the interval $[c^n(v),1]$ converges to $[\tilde{v},1]$ and f is continuous at $\tilde{v}$ (by the lack of negative jumps), then $\|f^n\|_{[c^n(v),1]} \le 3\varepsilon$ for large enough n. Hence $||f^n \circ c^n - f \circ c^\lambda||_{[v,r]} \le 4\varepsilon$ for large enough n.

A similar argument proves that $||f^n \circ c^n - f \circ c^\lambda||_u$ can be made smaller than ε for large enough n by choosing u adequately. The latter works also when $\lambda = \infty$ (in which case $f \circ c^\infty \equiv 0$), but for any $u > 0$. □

3.3.4 All subsequential limits of the (cumulative) profile converge to the (cumulative) Lamperti transform

Using Theorem 3.2 of the next section, we now prove that the rescaled cumulative profile converges to the cumulative Lamperti transform of X. Actually, we prove the joint convergence under the product topology of the rescaled BFW, the rescaled cumulative profile and the rescaled profile.

For the next lemma, denote by $Z = X \circ C$ the Lamperti transform of X.

Lemma 3.3. *Assume the hypotheses of Theorem 3.1. Furthermore, assume $\int_{1/2}^1 1/X(s)ds < \infty$ a.s. Let Z^n be the rescaled profile of a tree with law $\mathbb{P}_{\mathbf{s_n}}$. Then, for any bounded continuous function F on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)^3$ we have the joint convergence*

$$\mathbb{E}(F(X^n, C^n, Z^n)) \to \mathbb{E}(F(X, C, Z)).$$

Proof. In Theorem 3.2 (below) we prove that

$$\mathbb{P}\left(\Lambda \in (0,\infty], \int_{1/2}^{1} ds/X(s) < \infty\right) = 0.$$

Since $\int_{1/2}^{1} ds/X(s) < \infty$ a.s., then $\mathbb{P}(\Lambda \in (0,\infty]) = 0$. Recall from Proposition 3.8 that every subsequence has a further subsequence such that $Y^{n_l} := (X^{n_l}, C^{n_l}) \to Y^{\Lambda} := (X, C^{\Lambda})$, and set $Y = (X,C)$, where C is the cumulative Lamperti transform of X. For a continuity set A of the distribution function of Y^{Λ}, we have

$$\begin{aligned} \lim \mathbb{P}(Y^{n_l} \in A) &= \mathbb{P}\left(Y^{\Lambda} \in A\right) \\ &= \mathbb{P}\left(Y^{\Lambda} \in A, \Lambda = 0\right) \\ &= \mathbb{P}(Y \in A, \Lambda = 0) \\ &= \mathbb{P}(Y \in A). \end{aligned}$$

Since the limit does not depend on the subsequence and by Portmonteau's theorem, we have $Y^n \xrightarrow{d} Y$.

By Skorohod's theorem, we can assume such convergence takes place almost surely. Note that, C and C^n are admissible cumulative profile functions for every n. Hence, from Lemma 3.2 we have

$$X^n \circ C^n \xrightarrow{a.s.} X \circ C = Z.$$

This implies $(X^n, C^n, Z^n) \to (X, C, Z)$ a.s. on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)^3$, and therefore convergence in distribution follows. □

From the previous lemma we have

$$(Z(t), t \in [0, I(1)]) \stackrel{d}{=} \left(X \circ I^{(-1)}(t), t \in [0, I(1)]\right),$$

which has been established in a similar context in Equation (3) of [AMP04]. The equality is a generalization of Jeulin's identity, given in Chapter 4 of [JY85].

It remains to show that having a positive limit of the form C^{Λ} for $\Lambda \in (0,\infty]$ is the result of trees with an asymptotically thin base.

3.4 Asymptotic thickness of the base

Through this section, we assume the hypotheses of Theorem 3.1. Recall from Proposition 3.8 that $(C^n, n \in \mathbb{N})$ is sequentially compact. The objective of this section is to prove the following result.

Theorem 3.2. *Assume the hypotheses of Theorem 3.1 are satisfied. Let C^{Λ} be any subsequential limit of $(C^n, n \in \mathbb{N})$, and $I^{\Lambda}(\cdot) = \Lambda + \int_0^{\cdot} ds/X(s)$ its right-continuous inverse. Then*

$$\mathbb{P}\left(\Lambda \in (0,\infty], \int_{1/2}^{1} ds/X(s) < \infty\right) = 0. \tag{3.20}$$

The reader can imagine the event on the left-hand side, as the limit of trees having a base of the form of a cord with size approximately Λ, after that, a giant subtree starts to grow, attached to such cord, there can be other cords with (possibly) giant subtrees growing after height Λ (see the tree on the left in Figure 3.11).

To prove Theorem 3.2, we use a path transformation for discrete time EI processes which is very easily visualized on the tree they code. Recall the 213 transformation was introduced in Definition 3.4 for plane trees, and an example is given in Figure 3.10. Consider a tree τ_n with law $\mathbb{P}_{\mathbf{s_n}}$, an independent uniform variable $U_n \in \{2,\dots,s_n\}$, and $h \in \mathbb{N}$. We define a new tree $\tilde{\tau}_n$ as follows: if the height of U_n is greater than h, then $\tilde{\tau}_n$ is the 213 transformation of τ_n cutting at U_n and N_n, where N_n is the ancestor of U_n at distance h; if the height of U_n is at most h, then $\tilde{\tau}_n = \tau_n$. We will use this transformation for individuals *near the top of the tree*. To give a formal definition of what *near the top* means, we need some lemmas regarding the convergence of the heights of the first and last individuals.

We work on the space where there is a.s. convergence $(X^{n_l}, C^{n_l}) \to (X, C^{\Lambda})$, for some deterministic subsequence $(n_l, l \geq 1)$. Write only $(X^l, C^l) \to (X, C^{\Lambda})$ to avoid cumbersome notation. We introduce the hitting times of (ε, ∞) by C^l.

Definition 3.8. *For $\varepsilon \in (0,1)$ and $l \in \mathbb{N}$, define the first height that the rescaled cumulative profile has more than ε individuals as*
$$\Lambda_{l,\varepsilon} := \inf\{t > 0 : C^l(t) > \varepsilon\}.$$

Recall from (3.16), that $\inf\{t > 0 : C^{\Lambda}(t) > \varepsilon\} = I^{\Lambda}(\varepsilon) = \Lambda + I(\varepsilon)$ for any $\varepsilon \in (0,1)$. We prove that the height given in the above definition, converges to the height where the limit C^{Λ} first accumulates ε.

Lemma 3.4. *For every $\varepsilon \in (0,1)$, we have $\Lambda_{l,\varepsilon} \overset{a.s.}{\to} \Lambda + I(\varepsilon)$ as $l \to \infty$.*

The proof of this Lemma is given in Subsection 3.4.2. By monotonicity and Borel-Cantelli, we can consider limits in which ε of the previous lemma also depends on l. This is stated in the next lemma. To state it, we introduce notation for typographical convenience. For some subsequences $(n_{l^1_k}, k \in \mathbb{N})$ and $(n_{l^2_k}, k \in \mathbb{N})$ we write them as $(n(l^1(k)), k \in \mathbb{N})$ and $(n(l^2(k)), k \in \mathbb{N})$. Also, we will only write $\Lambda_{k,\varepsilon_k} := \Lambda_{n(l^1(k)),\varepsilon_{n(l^2(k))}}$ and $\Lambda_{k,1-\varepsilon_k} := \Lambda_{n(l^1(k)),1-\varepsilon_{n(l^2(k))}}$ when referring to the hitting times evaluated on the previous subsequences. Define $\Lambda_{0+} = \inf\{t > 0 : C^{\Lambda}(t) > 0\}$, and note that $\Lambda_{0+} = \Lambda$ whenever $I(1/2) < \infty$, and $\Lambda_{0+} = \infty$ otherwise. We prove that jointly, the height of the first individuals in the tree converges to the height were C^{Λ} starts to grow, and the height of the last individuals in the tree (from the height of half the tree) converges to the corresponding heights of C^{Λ}.

Lemma 3.5. *Consider any sequence $\varepsilon_{n_l} \downarrow 0$. Then, there exists deterministic subsequences $(n(l^1(k)), k \in \mathbb{N})$ and $(n(l^2(k)), k \in \mathbb{N})$ such that*
$$(\Lambda_{k,\varepsilon_k}, \Lambda_{k,1-\varepsilon_k} - \Lambda_{k,1/2}) \overset{a.s.}{\to} (\Lambda_{0+}, I(1) - I(1/2)) \qquad \textit{as } k \to \infty,$$
where in the right-hand side, we interpret $I(1) - I(1/2)$ as $\int_{1/2}^{1} ds/X(s)$.

Indeed, from the proof of this lemma, we can prove that the previous convergence takes place together with X^n and C^n on the product space $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)^2$. Also, we remark that the choice of 1/2 is arbitrary, and the lemma works for any $a \in (0,1)$.

Remark 3.4. *In the following, we will use the temporal rescaling $a_{s_k} = s_k/b_{s_k}$, which comes from the Lamperti transform in Subsection 3.3.1. For ease of notation, we sometimes put the superscript (k) to refer of such rescaling, for example, for $\lambda > 0$, we write $\lambda^{(k)} := \lfloor \lambda a_{s_k} \rfloor$, and also $\Lambda^{(k)}_{k,\varepsilon_k} = \lfloor \Lambda_{k,\varepsilon_k} a_{s_k} \rfloor$.*

To prove Theorem 3.2, we fix any $\lambda > 0$ and prove

$$\mathbb{P}(\Lambda \in (\lambda,\infty), I(1) - I(1/2) < \infty) = 0. \tag{3.21}$$

We use this case to deduce $\mathbb{P}(\Lambda = \infty, I(1) - I(1/2) < \infty) = 0$. The proof of Theorem 3.2 is done in the following way. It is proved in Section 3.4.1 that $(\tilde{\tau}_k, U) \stackrel{d}{=} (\tau_k, U)$. Let $\delta_n \downarrow 0$ and $\lambda_1 \in (0, \lambda/2)$. In Subsection 3.4.2 we prove that for n big enough and under the event of Equation (3.21), *near the top of the tree*, that is, at height $\Lambda^{(k)}_{k,1-\varepsilon_k} - \lambda^{(k)}_1$, there is one vertex v_k having at least $\lfloor \delta_n s_k \rfloor$ descendants, for every k big enough. This will be referred to as a *giant subtree*. It is a subtree with a lot of individuals, almost all of them between its first $\lambda^{(k)}_1$ generations. In Section 3.4.3, we use the 213 transformation to obtain a tree which is fat near the root, thin in the middle and fat after that. The ideas are:

1. Intersect with the set $\Lambda_{k,1-\varepsilon_k} - \lambda_1 > \Lambda_{k,1/2}$.
2. Consider two heights $h(v_k, 1/2) < h(v_k, 1/4)$ where the giant subtree is fat: *up to* $h(v_k, 1/2)$ there is at least half of the size of the giant subtree, *between* $h(v_k, 1/2)$ and $h(v_k, 1/4)$ there is at least quarter of the size of the giant subtree.
3. The probability that U_k (the uniform variable used for the 213 transformation) be in the giant subtree, *between* heights $h(v_k, 1/2)$ and $h(v_k, 1/4)$, is at least $\delta_n/4$.
4. Apply the 213 transformation cutting at U_k and N_k, this one having distance $\lambda^{(k)}_1$ from U_k. Note that N_k is an ancestor of v_k.
5. On such events, the transformed tree $\tilde{\tau}_k$ has
   (a) at least $\delta_n s_k/2$ individuals up to height $(2\lambda_1)^{(k)}$,
   (b) less than $2\varepsilon_k s_k$ individuals between heights $\{(2\lambda_1)^{(k)}, \ldots, \lambda^{(k)}\}$,
   (c) approximately at least $s_k/2$ individuals after $\lambda^{(k)}$.
6. Such tree has the same distribution as the original, but its cumulative Lamperti transform is constant between two intervals where it increases. This has probability zero.

In Figure 3.11 we show a representation of point (5).

3.4.1 The 213 transformation

The aforementioned transformation of a tree is easy to formalize using the associated depth-first walk. Nevertheless, it will be applied to the breadth-first walk.

Let $s_n \in \mathbb{N}$ and $x_i \in \mathbb{N} \cup \{-1\}$ for $i \in [s_n]$ such that $\sum_1^{s_n} x_i = -1$. Consider the discrete time excursion w with partial sums $w_j = \sum_1^j x_i$, starting at zero and non-negative up to time $w_{s_n} = -1$. The tree T_n generated by w is labeled in depth-first order. We now define the *213 transformation* of T_n, denoted $\Psi_{n,u}(T_n)$. Let

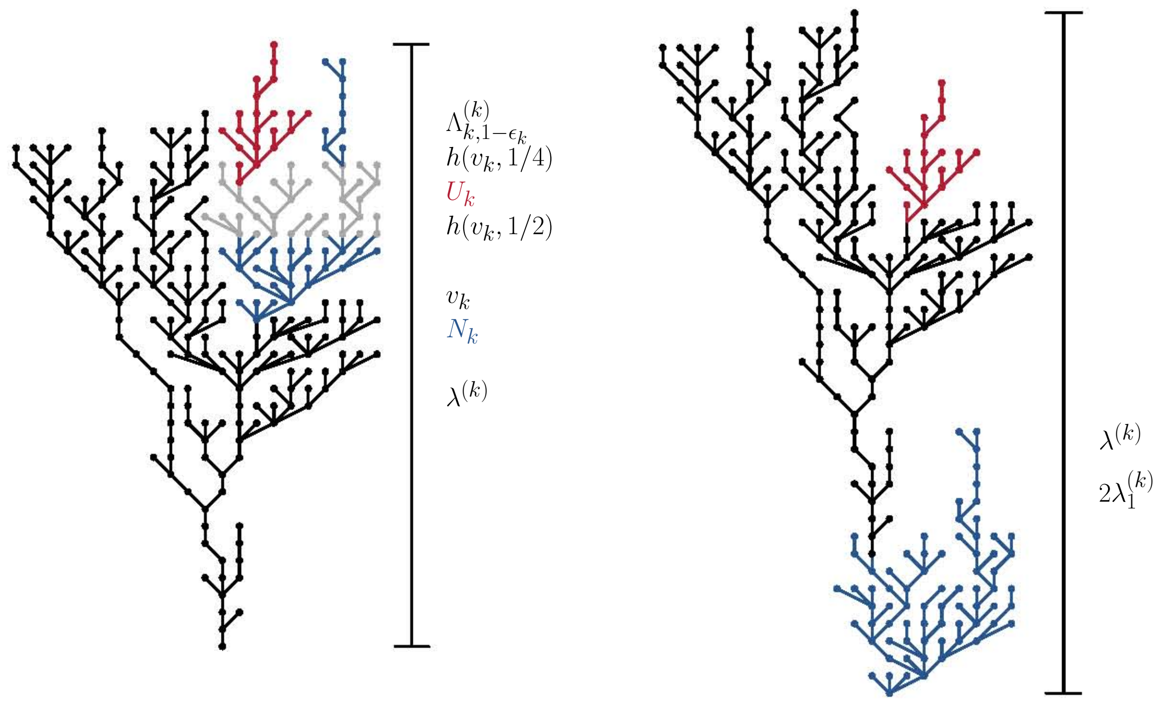

Figure 3.11: On the left, the original tree. The giant subtree has root v_k, and the distance between v_k and any of its descendants at height $\Lambda^{(k)}_{k,1-\varepsilon_k}$ is bounded by $\lambda^{(k)}_1$. The uniform variable U_k is in such subtree, between heights $h(v_k,1/2)$ and $h(v_k,1/4)$ (gray). The ancestor N_k of U_k at distance $\lambda^{(k)}_1$, is an ancestor of v_k. After applying the 213 transformation (figure on the right), the new tree is fat before height $(2\lambda_1)^{(k)}$, thin between heights $\{(2\lambda_1)^{(k)},\dots,\lambda^{(k)}\}$, and again fat after $\lambda^{(k)}$.

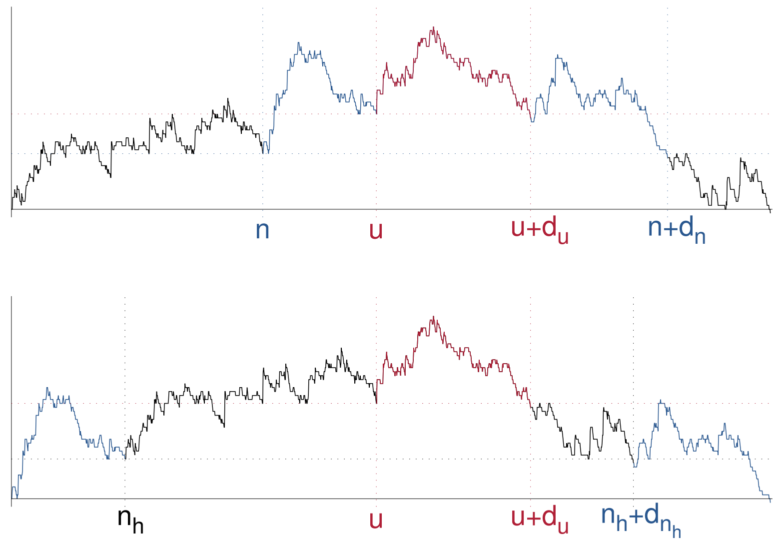

Figure 3.12: Recall the notation of the subtrees obtained after cutting at n and u the tree T_n, as in Definition 3.4. In the top excursion, the subtree $T_n(1,n)$ is represented in black, $T_n(n,u)$ in blue and $T_n(u)$ in red. The bottom excursion is T_n after applying the 213 transformation. The root of the grafted subtree $T_n(1,n)$ will have label $n_h = u-n+1$ in $\tilde{T}_n$.

$\Delta w_j := w_j - w_{j-1}$ be the jth increment of w (number of children of j). Choose a vertex $n \in \{2,\ldots,s_n\}$ in the tree which is neither the root nor a leaf, and let

$$d_n = \inf\{j \in \{1,\ldots,s_n-n+1\} : w_{n-1+j} - w_{n-1} = -1\}$$

be the length of the excursion starting at n (this is equivalent to $|T_n(n)|+1$). Notice that a vertex n is a leaf iff $d_n = 1$. Consider any vertex $u \in \{n+1,\ldots,n-1+d_n\}$, implying n is its ancestor.

Denoting by $\Psi_{n,u}(w)$ the depth-first walk of the transformed tree $\Psi_{n,u}(T_n)$, we decree:

$$\Delta\Psi_{n,u}(w)_j = \begin{cases} \Delta w_{n-1+j} & \text{if } 1 \le j \le u-n \\ \Delta w_{j-(u-n)} & \text{if } u-n+1 \le j \le u-1 \\ \Delta w_{u+j-u} & \text{if } u \le j \le u-1+d_u \\ \Delta w_{n+d_n+j-(u+d_u)} & \text{if } u+d_u \le j \le s_n+u+d_u-n-d_n \\ \Delta w_{u+d_u+j-(s_n+u+d_u-n-d_n+1)} & \text{if } s_n+u+d_u-n-d_n+1 \le j \le s_n. \end{cases}$$

Figure 3.12 shows the transformation of the depth-first walk.

Using a uniform law on the vertex u and choosing n in an adequate way, we prove the invariance of the law $\mathbb{P}_{\mathbf{s_n}}$ under a 213-type transformation.

Define the space of all excursions coding a tree with given degree sequence $\mathbf{s_n}$ by

$$\mathscr{E}_{\mathbf{s_n}} = \{w : w \text{ is an excursion with degree sequence } \mathbf{s_n}\}.$$

Given an excursion $w \in \mathscr{E}_{\mathbf{s_n}}$, the natural numbers $u \in [s_n] \setminus \{1\}$ and $h \ge 1$, we construct $\tilde{w}$ as follows. In the tree generated by w, if $h(u)$ denotes the height of the vertex u (its distance to the root), when

$$h(u) > h, \tag{3.22}$$

let n be the ancestor of u at distance h, and define $\tilde{w} = \Psi_{n,u}(w)$. If condition (3.22) is not satisfied, define $\tilde{w} = w$.

Lemma 3.6. *For fixed $h \geq 1$ and $u \in \{2, \ldots, s_n\}$, the transformation $\Phi_{h,u} : \mathscr{E}_{\mathbf{s_n}} \to \mathscr{E}_{\mathbf{s_n}}$ sending w to $\tilde{w}$ is bijective.*

Proof. It is easy to prove that u satisfies (3.22) for w iff u satisfies (3.22) for $\tilde{w}$. Being a transformation between finite sets, it suffices to prove it is onto. Consider $\tilde{w} \in \mathscr{E}_{\mathbf{s_n}}$. If u does not satisfy (3.22) for $\tilde{w}$, define $w := \tilde{w}$. If u satisfies (3.22) for $\tilde{w}$, choose the vertex $n_h = u - n + 1$ at height h which is its ancestor. Define in this case $w = \Psi_{u-n+1,u}(\tilde{w})$. Hence, the explicit bijection is

$$\tilde{w} = \Psi_{n,u}(\Psi_{u-n+1,u}(\tilde{w})). \qquad \square$$

We now prove the aforementioned equality in distribution.

Proposition 3.9. *Let W_{s_n} be the depth-first walk of a tree with law $\mathbb{P}_{\mathbf{s_n}}$. Consider an independent uniform r.v. U on $\{2, \ldots, s_n\}$ and a natural number $h \geq 1$. When (3.22) is satisfied for the Uth vertex in the tree generated by W_{s_n}, let N be the only ancestor of U at distance h, and define $\tilde{W}_{s_n} := \Psi_{N,U}(W_{s_n})$ the 213 transformation of W_{s_n} at N and U. If (3.22) is not satisfied, set $\tilde{W}_{s_n} := W_{s_n}$. Then, we have*

$$\left(\tilde{W}_{s_n}, U\right) \stackrel{d}{=} (W_{s_n}, U).$$

Proof. Consider any excursion $w \in \mathscr{E}_{\mathbf{s_n}}$ and $u \in \{2, \ldots, s_n\}$. Using the bijection $\Phi_{h,u}$ of Lemma 3.6 and the independence between the tree and U

$$\mathbb{P}\left(\tilde{W}_{s_n} = w, U = u\right) = \mathbb{P}\left(W_{s_n} = \Phi_{h,u}^{-1}(w), U = u\right) = \mathbb{P}(W_{s_n} = w, U = u),$$

using that W_{s_n} is the excursion of the tree τ_n with a uniform law. $\square$

We will use this proposition with trees labeled in BFO. In that case, let (W_{s_n}, U) be the BFW of a tree $\tau_{\mathbf{s_n}}$ with law $\mathbb{P}_{\mathbf{s_n}}$ and U a uniform r.v. on $\{2, \ldots, s_n\}$. We define $\tilde{W}_{s_n}$ exactly as before, that is, if $h(U) > h$ we cut both at U and its ancestor at distance h, and leave unchanged the excursion if $h(U) \leq h$. Since there is a bijection between $(\tau_{\mathbf{s_n}}, U)$ and the same tree but labeled in DFO, together with the new label of U, then also in this case $\left(\tilde{W}_{s_n}, U\right) \stackrel{d}{=} (W_{s_n}, U)$.

3.4.2 Existence of a giant subtree

In the following, assume that we are working on a space where the rescaled BFW's and cumulative Lamperti transform converge a.s.

Proof of Lemma 3.4. Recall that $\Lambda_{0+} = \inf\{t > 0 : C^\Lambda(t) > 0\}$. First assume $\omega \in \{\Lambda_{0+} = \infty\}$. Using such ω implicitly, then $C^l \to 0$ on $[0, M]$ for any $M > 0$. Hence, we have $C^l(M) \leq \varepsilon$ and $\Lambda_{l,\varepsilon} \geq M$ for every l big enough. It follows that $\Lambda_{l,\varepsilon} \to \infty$.

Now, assume

$$\omega \in \{\Lambda_{0+} < \infty\} = \{C^\Lambda(\infty) > 0\} = \{I^\Lambda(a) = \Lambda + I(a) < \infty, \forall a \in (0,1)\}. \tag{3.23}$$

Using such ω implicitly, for $\varepsilon \in (0,1)$, consider δ such that $0 < \varepsilon - \delta < \varepsilon + \delta < 1$. Then, for every $u_0 > \Lambda + I(\varepsilon + \delta)$ exists L_{u_0} such that

$$C^{\Lambda}(u) - \delta < C^l(u) < C^{\Lambda}(u) + \delta \quad \forall\, l \geq L_{u_0},\ u \in [0, u_0].$$

Substituting u in this inequality with the values $\Lambda + I(\varepsilon - \delta)$ and $\Lambda + I(\varepsilon + \delta)$, we obtain

$$C^l(\Lambda + I(\varepsilon - \delta)) < \varepsilon < C^l(\Lambda + I(\varepsilon + \delta)).$$

Therefore, for every $l \geq L_{u_0}$

$$\Lambda + I(\varepsilon - \delta) < \Lambda_{l,\varepsilon} \leq \Lambda + I(\varepsilon + \delta).$$

Letting $l \to \infty$ and then $\delta \downarrow 0$ we obtain $\Lambda_{l,\varepsilon} \to \Lambda + I(\varepsilon)$. □

To prove the convergence of the heights for the first and last individuals, we need some definitions. Fix any points $0 < a < b \leq 1$ of X, and note X is continuous on both points since it only jumps at uniform times. For $\Lambda^l = \lfloor \Lambda_{l,a} a_{s_l} \rfloor / a_{s_l}$, define the processes

$$X_a^l = (X^l(C^l(\Lambda^l) + v), v \in [0, b - C^l(\Lambda^l)]) \qquad \text{and} \qquad X_a = (X(a+v), v \in [0, b-a]).$$

Since $X_a(0) = X(a) > 0$ and X_a is absorbed at zero or positive, by Proposition 1 of [CPGUB13], the equation $D_+ C_a = X_a \circ C_a$ has a unique solution C_a, with inverse I_a. It is easy to see that

$$I_a(v) = \int_0^v \frac{ds}{X_a(s)} = I(a+v) - I(a) \qquad v \in [0, b-a].$$

Consider also the solution C_a^l to the equation

$$C_a^l(t) = \int_0^t X_a^l \circ C_a^l(\lfloor u a_{s_l} \rfloor / a_{s_l}) du,$$

which is unique by its definition as a recursion (see Equation (3.12)). We prove that

$$C_a^l(t) = C^l(\Lambda^l + t) - C^l(\Lambda^l), \qquad t \in [0, \Lambda_{l,b} - \Lambda^l] \tag{3.24}$$

is such solution. Using Equation (3.12), and substituting the values of X_a^l and C_a^l, we have for $t_i \leq t < t_{i+1}$

$$\begin{aligned}
C_a^l(t) &= \int_0^t X_a^l \circ C_a^l(\lfloor u a_{s_l} \rfloor / a_{s_l}) du \\
&= C_a^l(t_i) + (t - t_i) X_a^l \circ C_a^l(t_i) \\
&= C^l(\Lambda^l + t_i) - C^l(\Lambda^l) + (t - t_i) X^l \circ C^l(\Lambda^l + t_i) \\
&= C^l(\Lambda^l + t_i) + (\Lambda^l + t - t_i - \Lambda^l) X^l \circ C^l(\Lambda^l + t_i) - C^l(\Lambda^l) \\
&= C^l(\Lambda^l + t) - C^l(\Lambda^l),
\end{aligned}$$

using the fact that there is a j such that for $t_j := \Lambda^l + t_i$, then $t_j \leq \Lambda^l + t < t_{j+1}$. From these definitions, we can prove the following lemma.

Lemma 3.7. *For any $0 < a < b < 1$, we have $\Lambda_{l,b} - \Lambda_{l,a} \to \int_a^b ds/X(s)$ a.s. as $l \to \infty$.*

Remark 3.5. *Recall that* $I(b)-I(a):=\int_a^b ds/X(s)$. *Thus, even if* $\Lambda_{l,a}\to\infty$ *because* $\Lambda=\infty$, *the integral* $\Lambda_{l,b}-\Lambda_{l,a}$ *converges to a finite quantity in this case.*

Proof. By definition

$$a=C^l(\Lambda_{l,a})=C^l(\Lambda^l)+(\Lambda_{l,a}-\Lambda^l)X^l\circ C^l(\Lambda^l),$$

and note that $\Lambda_{l,a}-\Lambda^l\le 1/a_{s_l}\to 0$ and $X^l\circ C^l(\Lambda^l)\le 2||X||<\infty$ for l large enough. This implies $C^l(\Lambda^l)\to a$, and therefore $X^l_a\to X_a>0$ a.s. by Lemma 2.2 of [Whi80] and Proposition 6.5.a of [EK86]. Then, by Theorem 3 of [CPGUB13] we have $C^l_a\to C_a$ uniformly on $[0,I(b)-I(a)]$. This convergence implies the convergence of its inverses, analogously as in the proof of Lemma 3.4 (cf. Theorem 7.2 in [Whi80]); since I_a is continuous, the convergence takes place uniformly. If I^l_a is the inverse of C^l_a, then

$$\begin{aligned}\Lambda_{l,b}-\Lambda_{l,a}&=\Lambda_{l,b}-\Lambda^l+\Lambda^l-\Lambda_{l,a}\\&=\inf\{t:C^l(\Lambda^l+t)-C^l(\Lambda^l)>b-C^l(\Lambda^l)\}+\Lambda^l-\Lambda_{l,a}\\&=I^l_a(b-C^l(\Lambda^l))+\Lambda^l-\Lambda_{l,a}\\&\to I_a(b-a),\end{aligned}$$

which equals $I(b)-I(a)$ by definition. □

Proof of Lemma 3.5. Recalling Equation (3.23), we have $\Lambda_{0+}+I(\varepsilon)\to\Lambda_{0+}$ as $\varepsilon\downarrow 0$, since either both sides are finite or infinite. Thus

$$(\Lambda_{0+}+I(\varepsilon_{n_l}),I(1-\varepsilon_{n_l})-I(1/2))\overset{a.s.}{\to}(\Lambda_{0+},I(1)-I(1/2))\qquad\text{as } l\to\infty.$$

Hence, this convergence is also in probability. Let $(\delta_k;k\in\mathbb{N})$ be any summable sequence of positive reals decreasing to zero. Fixing $k\in\mathbb{N}$, and using a distance d that generates the usual topology on $[0,\infty]^2$, we know that for δ_k exists $l^2_k\in\mathbb{N}$ such that

$$\mathbb{P}(d\left((\Lambda_{0+}+I(\varepsilon_{n_l}),I(1-\varepsilon_{n_l})-I(1/2)),(\Lambda_{0+},I(1)-I(1/2))\right)>\delta_k/2)<\delta_k/2\quad\forall\, l\ge l^2_k.$$

Without loss of generality, assume $l^2_k>l^2_{k-1}$. By Lemmas 3.4 and 3.7, we have

$$\left(\Lambda_{l,\varepsilon_{n(l^2_k)}},\Lambda_{l,1-\varepsilon_{n(l^2_k)}}-\Lambda_{l,1/2}\right)\overset{a.s.}{\to}\left(\Lambda_{0+}+I(\varepsilon_{n(l^2_k)}),I(1-\varepsilon_{n(l^2_k)})-I(1/2)\right)\qquad\text{as } l\to\infty.$$

Again, this implies convergence in probability. Thus, for the same $\delta_k>0$ exists $l^1_k\in\mathbb{N}$ such that

$$\mathbb{P}\left(d\left((\Lambda_{l,\varepsilon_{n(l^2_k)}},\Lambda_{l,1-\varepsilon_{n(l^2_k)}}-\Lambda_{l,1/2}),(\Lambda_{0+}+I(\varepsilon_{n(l^2_k)}),I(1-\varepsilon_{n(l^2_k)})-I(1/2))\right)>\delta_k/2\right)<\delta_k/2\quad\forall\, l\ge l^1_k.$$

We assume $l^1_k>l^1_{k-1}$. Joining the previous inequalities and using the triangle inequality

$$\mathbb{P}\left(d\left((\Lambda_{k,\varepsilon_k},\Lambda_{k,1-\varepsilon_k}-\Lambda_{k,1/2}),(\Lambda_{0+},I(1)-I(1/2))\right)>\delta_k\right)<\delta_k.$$

We conclude the proof using Borel-Cantelli lemma. □

Remark 3.6. *In the following, we work with the subsequences given in Lemma 3.5.*

Sometimes, we will use the notation $I^k(\cdot):=\Lambda_{k,\cdot}$ for the inverse of C^k.

3.4.3 Trees are not asymptotically thin at the base

In this section we prove our main result, Theorem 3.2. In the proof, we will apply the 213-type transformation to a tree τ_k using a uniform variable U_k, and its ancestor N_k at distance $\lambda_1^{(k)}$ (see Proposition 3.9 and its discussion); the value λ_1 will be specified later. Define $\tilde{X}^k$, $\tilde{C}^k$ and $\tilde{\Lambda}_{k,\varepsilon_k}$ as X^k, C^k and $\Lambda_{k,\varepsilon_k}$ where defined, but for the transformed tree $\tilde{\tau}_k$. Because $X^k \overset{d}{=} \tilde{X}^k$, then $C^k \overset{d}{=} \tilde{C}^k$ and $\Lambda_{k,\varepsilon_k} \overset{d}{=} \tilde{\Lambda}_{k,\varepsilon_k}$.

To prove the result, we first restrict us to $\Lambda \in (\lambda,\infty)$ for some fixed $\lambda \in (0,\infty)$, and compute

$$\mathbb{P}(\Lambda \in (\lambda,\infty), I(1)-I(1/2)<\infty) = \lim_{\lambda_2'\uparrow\infty} \mathbb{P}\big(\Lambda \in (\lambda,\infty), I(1)-I(1/2)<\lambda_2'\big)$$

Fix $\lambda_2'>0$ and define $A_{\lambda_2'}$ as the set inside the last probability. We need a lower bound on the distance between $\Lambda_{k,1/2}$ and $\Lambda_{k,1-\varepsilon_k}$, hence for any fixed $\lambda_2 \in (0,\lambda_2')$ we split

$$\mathbb{P}\Big(A_{\lambda_2'}\Big) = \mathbb{P}\Big(A_{\lambda_2'} \cap \{I(1)-I(1/2) \in (\lambda_2,\lambda_2')\}\Big) + \mathbb{P}\Big(A_{\lambda_2'} \cap \{I(1)-I(1/2) \le \lambda_2\}\Big) \tag{A}$$

The first term on the right-hand side will be denoted by A1. We will prove that A1 equals zero for every (fixed) $\lambda_2 \in (0,\lambda_2')$. This would imply

$$\mathbb{P}\Big(A_{\lambda_2'}\Big) \le \varlimsup_{\lambda_2\downarrow 0} \mathbb{P}(I(1)-I(1/2)\le\lambda_2) \le \mathbb{P}(I(1)-I(1/2)=0), \tag{A'}$$

which has probability zero since $\int_{1/2}^1 ds/X(s)>0$ a.s.

To prove A1 has zero probability, we decompose it using the set where there is a giant subtree at the top of the tree. For a plane tree T_k with s_k vertices, let $u\in[s_k]$ be any vertex on T_k. We denote by $T_k(u)$ the *subtree generated by* u in the tree T_k (a tree with root u and all its descendants). When $v\in[0,1]$ is such that $\lfloor vs_k\rfloor \in [s_k]$, we also refer to v as a vertex in the tree (and write $T_k(v)$ instead of $T_k(\lfloor vs_k\rfloor)$). Let $\delta_n\downarrow 0$. For any $n,k\in\mathbb{N}$ and $\lambda_1\in(0,\lambda\wedge\lambda_2)$, consider:

$$A(n,k)=\{\text{there exists a vertex } v_k \text{ at height } \Lambda_{k,1-\varepsilon_k}^{(k)}-\lambda_1^{(k)} \text{ such that} |\tau_k(v_k)|>\delta_n s_k\}.$$

Recall that Lemma 3.5 implies $\Lambda_{k,\varepsilon_k}\to\Lambda_{0+}\ge\Lambda$. Then, intersecting with $A(n,k)$ and $A(n,k)^c$, we bound A1 as follows

$$\begin{aligned}
&\mathbb{P}\big(\Lambda\in(\lambda,\infty), I(1)-I(1/2)\in(\lambda_2,\lambda_2')\big)\\
&\le \mathbb{P}\Big(\varliminf_k \{\Lambda_{k,\varepsilon_k}>\lambda, \Lambda_{k,1-\varepsilon_k}-\Lambda_{k,1/2}\in(\lambda_2,\lambda_2')\}\Big)\\
&\le \varlimsup_n \varlimsup_k \mathbb{P}\big(\{\Lambda_{k,\varepsilon_k}>\lambda, \Lambda_{k,1-\varepsilon_k}-\Lambda_{k,1/2}\in(\lambda_2,\lambda_2')\}\cap A(n,k)\big)\\
&+\varlimsup_n \varlimsup_k \mathbb{P}\big(\{\Lambda_{k,1-\varepsilon_k}-\Lambda_{k,1/2}\in(\lambda_2,\lambda_2')\}\cap A(n,k)^c\big).
\end{aligned} \tag{B}$$

The first and second term on the right-side will be denoted by B1 and B2, respectively.

For arbitrary $\varepsilon\in(0,1/2)$, we split B2 as

$$\begin{aligned}
&\varlimsup_n \varlimsup_k \mathbb{P}\big(\{\Lambda_{k,1-\varepsilon_k}-\Lambda_{k,1/2}\in(\lambda_2,\lambda_2')\}\cap A(n,k)^c\big)\\
&\le \varlimsup_n \varlimsup_k \mathbb{P}\Big(\Big\{\Lambda_{k,1-\varepsilon_k}-\Lambda_{k,1/2}\in(\lambda_2,\lambda_2'), C^k(\Lambda_{k,1-\varepsilon_k}-\lambda_1)<1-\varepsilon\Big\}\cap A(n,k)^c\Big)\\
&+\varlimsup_k \mathbb{P}\Big(\Lambda_{k,1-\varepsilon_k}-\Lambda_{k,1/2}\in(\lambda_2,\lambda_2'), C^k(\Lambda_{k,1-\varepsilon_k}-\lambda_1)\ge 1-\varepsilon\Big),
\end{aligned} \tag{C}$$

and denote by C1 the first term on the right. Similarly as in A, we shall prove that C1 is zero for fixed ε. Thus

$$\begin{aligned}&\overline{\lim_n}\,\overline{\lim_k}\,\mathbb{P}\left(\left\{\Lambda_{k,1-\varepsilon_k}-\Lambda_{k,1/2}\in(\lambda_2,\lambda_2')\right\}\cap A(n,k)^c\right)\\&\leq\mathbb{P}\left(\overline{\lim_{\varepsilon\downarrow 0}}\,\overline{\lim_k}\left\{\Lambda_{k,1-\varepsilon_k}-\Lambda_{k,1/2}\in(\lambda_2,\lambda_2'),C^k(\Lambda_{k,1-\varepsilon_k}-\lambda_1)\geq 1-\varepsilon\right\}\right).\end{aligned}\tag{D}$$

Denote the right-hand side by D1.

Informally, the reason to decompose in such events is to obtain a giant subtree growing in the last λ_1 generations (which is B1), and prove that $\Lambda>0$ cannot happen using the 213 transformation by moving such giant subtree to the base of the tree. For the existence of the giant subtree, we have to prove there are a lot of individuals in the last λ_1 generations, that is, to prove D1 equals zero; then use this to prove that when there are a lot of individuals in the last λ_1 generations, one vertex at such height has a lot of descendants, which is C1. The steps of the proof are the following. First we prove D1 is zero, then prove the same for C1. Both imply B2 is zero. Then we prove B1 is zero. Hence, we have that A1 is zero, which also implies, by A', that $\mathbb{P}(\Lambda\in(\lambda,\infty),I(1)-I(1/2)<\infty)=0$.

D1 equals zero: the last λ_1 generations have a lot of individuals

Let $a\in(0,1/2)$ and consider $0<\lambda_1<\lambda_2$. Here, we prove that even when $\Lambda=\infty$ (that is, in the limit there are zero individuals at any fixed height), if we erase the first $\lfloor as_k\rfloor$ individuals, the limit of the cumulative profile will be positive, using Lemma 3.5. This means that above height $\Lambda_{k,a}$, there are a lot of individuals. Actually, we prove that

$$D:=\overline{\lim_{\varepsilon\downarrow 0}}\,\overline{\lim_k}\left\{\Lambda_{k,1-\varepsilon_k}-\Lambda_{k,1/2}\in(\lambda_2,\infty),C^k(\Lambda_{k,1-\varepsilon_k}-\lambda_1)\geq 1-\varepsilon\right\}=\emptyset.$$

On the event D, over a subsequence, at least $1-\varepsilon$ individuals are below height λ_1 from the top of the tree. To simplify notation, we denote with the same k such a subsequence. Assume that on D, there exists $a_{\lambda_1}\in(0,1-a)$, such that

$$\lim_k\left(C^k(\Lambda_{k,1-\varepsilon_k})-C^k(\Lambda_{k,1-\varepsilon_k}-\lambda_1)\right)=1-a-a_{\lambda_1}>0.\tag{3.25}$$

This would imply that on D, adding to $C^k(\Lambda_{k,1-\varepsilon_k}-1/a_{s_k})$ the terms $\pm C^k(\Lambda_{k,1-\varepsilon_k}-\lambda_1)$, the condition $C^k(\Lambda_{k,1-\varepsilon_k}-\lambda_1)\geq 1-\varepsilon$ implies

$$1-\varepsilon_k\geq 1-\varepsilon+C^k(\Lambda_{k,1-\varepsilon_k}-1/a_{s_k})-C^k(\Lambda_{k,1-\varepsilon_k}-\lambda_1).$$

Hence, taking the limit first with k and then with ε, by (3.25) we obtain

$$D\subset\overline{\lim_{\varepsilon\downarrow 0}}\left\{I(1)-I(1/2)\in[\lambda_2,\infty),1\geq 1-\varepsilon+1-a-a_{\lambda_1}\right\}=\emptyset.$$

Hence, it remains to prove (3.25) to obtain D1 equals zero. Note that any fixed $a\in(0,1/2)$ is a continuity point of X (since X only jumps at uniform times). Consider the processes X_a^k and X_a on $[0,1-a]$, defined on page 78. Consider also the associated solution C_a and its inverse I_a of IVP(X_a). By Equation (3.24), the expression inside parenthesis in (3.25) can be written as

$$C^k(I^k(1-\varepsilon_k))-C^k(I^k(1-\varepsilon_k)-\lambda_1)=C_a^k(I^k(1-\varepsilon_k)-\Lambda^k)-C_a^k(I^k(1-\varepsilon_k)-\Lambda^k-\lambda_1),$$

were we recall that $\Lambda^k = \lfloor \Lambda_{l,a} a_{s_l} \rfloor / a_{s_l}$. Note that the above expression is well-defined since $a \in (0,1/2)$ and

$$I^k(1-\varepsilon_k) - \Lambda^k \geq I^k(1-\varepsilon_k) - I^k(1/2) \geq \lambda_2 > \lambda_1. \tag{3.26}$$

From Lemma 3.5, we obtain $I^k(1-\varepsilon_k) - \Lambda^k \to I(1) - I(a) = I_a(1-a)$, which is finite on the event of interest. Also, in the proof of Lemma 3.7 we obtained the convergence $C_a^k \to C_a$ uniformly on compact sets. This implies

$$\lim \Big(C^k(I^k(1-\varepsilon_k)) - C^k(I^k(1-\varepsilon_k) - \lambda_1) \Big) = C_a(I_a(1-a)) - C_a(I_a(1-a) - \lambda_1). \tag{3.27}$$

Let $a_{\lambda_1} \in (0, 1-a)$ be such that

$$\lambda_1 = I_a(1-a) - I_a(a_{\lambda_1}).$$

Such value exists because the function $g(x) = I_a(1-a) - I_a(x)$ is continuous on $[0, 1-a]$, and satisfies $g(1-a) = 0$ and $g(0) > \lambda_1$ by (3.26). Hence, the right-hand side of (3.27) equals $C_a(I_a(1-a)) - C_a(I_a(a_{\lambda_1})) = 1 - a - a_{\lambda_1} > 0$.

C1 equals zero: some vertex at height λ_1 from the top of the tree has a lot of descendants

Denote the event in C1 as $C(n,k)$. On $C(n,k)$, every individual at height $\Lambda^{(k)}_{k,1-\varepsilon_k} - \lambda_1^{(k)}$ has less than $\delta_n s_k$ descendants (by the definition of $A(n,k)^c$). Also on such set, the number of individuals at height greater than $\Lambda^{(k)}_{k,1-\varepsilon_k} - \lambda_1^{(k)}$ are greater than εs_k, since

$$1 = C^k(\Lambda_{k,1-\varepsilon_k} - \lambda_1) + C^k(\infty) - C^k(\Lambda_{k,1-\varepsilon_k} - \lambda_1) < 1 - \varepsilon + C^k(\infty) - C^k(\Lambda_{k,1-\varepsilon_k} - \lambda_1).$$

Since the number of individuals at height greater than $\Lambda^{(k)}_{k,1-\varepsilon_k} - \lambda_1^{(k)}$ is equal to the number of descendants of the vertices at that height, which is also the elements of the subtrees growing at that height minus its roots, then

$$\begin{aligned}
\varepsilon &< \sum_{j: h(v_j) = \Lambda_{k,1-\varepsilon_k} - \lambda_1} \frac{|\tau_k(v_j)| - 1}{s_k} \\
&\leq Z^k(\Lambda_{k,1-\varepsilon_k} - \lambda_1)(\delta_n - 1/s_k) \\
&\leq ||X^k|| \delta_n \\
&\leq 2||X|| \delta_n,
\end{aligned}$$

for k big enough, where the last rows are true since $X^k \to X$ and $Z^k = X^k \circ C^k$ (see the definition of Z^k in page 66). Thus, for ε fixed, we can bound C1 with

$$\mathbb{P}\Big(\overline{\lim_n} \{ \varepsilon < 2||X|| \delta_n \} \Big) = 0.$$

B1 equals zero: moving a giant subtree to the base of the tree

We prove that, for fixed $\lambda_1 \in (0, \lambda/2)$ small enough we have

$$\overline{\lim_n}\,\overline{\lim_k}\,\mathbb{P}\big(\Lambda_{k,\varepsilon_k} > \lambda, \Lambda_{k,1-\varepsilon_k} - \Lambda_{k,1/2} \in (\lambda_2, \lambda_2'),\\ \text{exists a vertex } v_k \text{ at height } \Lambda^{(k)}_{k,1-\varepsilon_k} - \lambda^{(k)}_1 \text{ with } |\tau_k(v_k)| > \delta_n s_k\big) = 0.$$

Denote the set inside the probability as $B(n,k)$. Now we formalize the steps given in page 74, using the 213 transformation to move a giant subtree to the base of the tree. First we need some definitions.

Definition 3.9. *Let T be a tree and v a non-leaf vertex in T. For any $h_1, h_2 \in \mathbb{N} \cup \{\infty\}$ with $h(v) \le h_1 \le h_2$, let*

$$|T(v)|(h_1, h_2) = \{\text{number of individuals in } T(v) \text{ having height } h \in \{h_1, \dots, h_2\} \text{ in the tree } T\}.$$

We also define $|T(v)|(h_1) := |T(v)|(h(v), h_1)$, the number of individuals in $T(v)$ up to height h_1 in the tree T.

Lemma 3.8. *Let $2\lambda_1 < \lambda_2 < \lambda_2'$ and $2\lambda_1 < \lambda$. Then for every n big enough*

$$\overline{\lim_k}\,\mathbb{P}(B(n,k)) = 0.$$

Proof. Define the first height where the subtree $\tau_k(v_k)$ has at least half of its size:

$$h(v_k, 1/2) = \inf\{h > h(v_k) : |\tau_k(v_k)|(h) \ge |\tau_k(v_k)|/2\},$$

and the first height h after $h(v_k, 1/2)$, where the subtree $\tau_k(v_k)$ accumulates a quarter of its size between $\{h(v_k, 1/2) + 1, \dots, h\}$:

$$h(v_k, 1/4) = \inf\{h > h(v_k, 1/2) : |\tau_k(v_k)|(h(v_k, 1/2) + 1, h) \ge |\tau_k(v_k)|/4\}.$$

Note that $h(v_k, 1/4) < \Lambda^{(k)}_{k,1-\varepsilon_k}$ with high probability. This is true because, letting $z_k(v_k, 1/2) := |\tau_k(v_k)|(h(v_k, 1/2), h(v_k, 1/2))$ be the number of individuals in $\tau_k(v_k)$ at height $h(v_k, 1/2)$, by definition

$$\begin{aligned}
&|\tau_k(v_k)|(h(v_k, 1/4) - 1)\\
&= z_k(v_k, 1/2) + |\tau_k(v_k)|(h(v_k, 1/2) - 1) + |\tau_k(v_k)|(h(v_k, 1/2) + 1, h(v_k, 1/4) - 1)\\
&< 3|\tau_k(v_k)|/4 + z_k(v_k, 1/2).
\end{aligned}$$

By definition, above height $\Lambda^{(k)}_{k,1-\varepsilon_k}$ there are at most $s_k\varepsilon_k$ individuals. Hence, on the event $B(n,k) \cap \{h(v_k, 1/4) \ge \Lambda^{(k)}_{k,1-\varepsilon_k}\}$, all the individuals in $\tau_k(v_k)$ having height at least $h(v_k, 1/4)$ are at most $s_k\varepsilon_k$, thus

$$s_k\varepsilon_k \ge |\tau_k(v_k)|(h(v_k, 1/4), \infty) \ge |\tau_k(v_k)|/4 - z_k(v_k, 1/2) \ge \delta_n s_k/4 - Z_{s_k}(h(v_k, 1/2)).$$

Therefore, on the event of interest

$$\delta_n/4 - \varepsilon_k < (b_{s_k}/s_k)||Z^k|| \le (b_{s_k}/s_k)||X^k|| \le (b_{s_k}/s_k)2||X||,$$

for k big enough since $X^k \to X$. The above has probability going to zero as $k \to \infty$ for fixed n, using Hypothesis (4) of Theorem 3.1. Therefore, we can assume $h(v_k, 1/4) < \Lambda^{(k)}_{k,1-\varepsilon_k}$. In a similar way, we can prove that $h(v_k, 1/2) < h(v_k, 1/4)$.

Consider a uniform and independent random variable U_k. An important remark is that U_k is independent of τ_k, since the subsequence $(n(l^1(k)), k \in \mathbb{N})$ obtained in Lemma 3.5 was deterministic. The probability for U_k to be in $\tau_k(v_k)$, between heights $\{h(v_k, 1/2)+1, \ldots, h(v_k, 1/4)\}$, is at least $\delta_n/4$, by definition of $h(v_k, 1/4)$. It follows that

$$\begin{aligned}
&\frac{\delta_n}{4} \varlimsup_k \mathbb{P}(B(n,k)) \\
&\leq \varlimsup_k \mathbb{P}\Big(B(n,k) \cap \{U_k \in \tau_k(v_k), h(U_k) \in \{h(v_k,1/2)+1, \ldots, h(v_k,1/4)\}, h(v_k,1/4) < \Lambda^{(k)}_{k,1-\varepsilon_k}\}\Big).
\end{aligned}$$

Denote the event on the right-hand side by $B_1(n,k)$. Note that the ancestor N_k at distance $\lambda_1^{(k)}$ of U_k, is also an ancestor of v_k. Indeed, we have

$$h(N_k) = h(U_k) - \lambda_1^{(k)} \leq h(v_k, 1/4) - \lambda_1^{(k)} \leq \Lambda^{(k)}_{k,1-\varepsilon_k} - \lambda_1^{(k)} = h(v_k).$$

By construction, on the set $B_1(n,k)$ the transformed tree $\tilde{\tau}_k$ satisfies

1. There are at least $\delta_n s_k/2$ individuals up to height $(2\lambda_1)^{(k)}$. This is true, because

$$h(v_k,1/2) - h(N_k) \leq \Lambda^{(k)}_{k,1-\varepsilon_k} - h(U_k) + \lambda_1^{(k)} \leq \Lambda^{(k)}_{k,1-\varepsilon_k} - h(v_k) + \lambda_1^{(k)} = (2\lambda_1)^{(k)}$$

   and

$$|\tau_k(N_k)|(h(v_k,1/2)) \geq |\tau_k(v_k)|(h(v_k,1/2)) \geq \delta_n s_k/2.$$

2. Between heights $\{(2\lambda_1)^{(k)}, \ldots, \lambda^{(k)}\}$ there are at most $2\varepsilon_k s_k$ individuals. This is true, since on one hand

$$|\tau_k(N_k)|(\Lambda^{(k)}_{k,1-\varepsilon_k}, \infty) \leq |\tau_k|(\Lambda^{(k)}_{k,1-\varepsilon_k}, \infty) \leq \varepsilon_k s_k \quad \text{and} \quad |\tau_k|(\lambda^{(k)}) \leq |\tau_k|(\Lambda^{(k)}_{k,\varepsilon_k} - 1) \leq \varepsilon_k s_k,$$

   and, on the other hand

$$\Lambda^{(k)}_{k,1-\varepsilon_k} - h(N_k) \leq h(U_k) + \lambda_1^{(k)} - h(N_k) = (2\lambda_1)^{(k)}.$$

3. After height $\lambda^{(k)}$ there are at least $s_k/2 - \varepsilon_k s_k$ individuals. This holds, since

$$h(N_k)/a_{s_k} \geq \Lambda_{k,1-\varepsilon_k} - 2\lambda_1 \geq \Lambda_{k,1/2} + \lambda_2 - 2\lambda_1 \geq \Lambda_{k,1/2},$$

   and $\lambda_1^{(k)} < \Lambda^{(k)}_{k,\varepsilon_k} \leq \Lambda^{(k)}_{k,1-\varepsilon_k}$ implies

$$C^k(h(N_k)/a_{s_k}) - C^k(\lambda) \geq 1/2 - \varepsilon_k.$$

These three properties are illustrated in Figure 3.11. Using the equality in distribution $\tau_k \stackrel{d}{=} \tilde{\tau}_k$ of Proposition 3.9, we bound

$$\begin{aligned}\overline{\lim_k}\, \mathbb{P}(B_1(n,k)) &\leq \overline{\lim_k}\, \mathbb{P}\Big(\tilde{C}^k(2\lambda_1) \geq \delta_n/2, \tilde{C}^k(\lambda) - \tilde{C}^k(2\lambda_1) \leq 2\varepsilon_k, 1 - \tilde{C}^k(\lambda) \geq 1/2 - \varepsilon_k\Big) \\ &= \overline{\lim_k}\, \mathbb{P}\Big(C^k(2\lambda_1) \geq \delta_n/2, C^k(\lambda) - C^k(2\lambda_1) \leq 2\varepsilon_k, 1 - C^k(\lambda) \geq 1/2 - \varepsilon_k\Big) \\ &\leq \mathbb{P}\Big(C^\Lambda(2\lambda_1) \geq \delta_n/2, C^\Lambda(\lambda) - C^\Lambda(2\lambda_1) = 0, 1 - C^\Lambda(\lambda) \geq 1/2\Big),\end{aligned}$$

which is zero, since C^Λ cannot be constant inside an interval where it is strictly increasing. This concludes the proof. □

Proof of Theorem 3.2

On the set $\omega \in \{\Lambda = \infty, I(1) - I(1/2) < \infty\}$, since $C^k \to C^\Lambda \equiv 0$, we have $\Lambda_{k,\varepsilon_k} \to \infty$ using Lemma 3.5. It follows that

$$\begin{aligned}\mathbb{P}(\Lambda = \infty, I(1) - I(1/2) < \infty) &= \mathbb{P}\left(\bigcap_\lambda \bigcup_{\lambda_2'} \left\{\underline{\lim_k}\, \{\Lambda_{k,\varepsilon_k} > \lambda, \Lambda_{k,1-\varepsilon_k} - \Lambda_{k,1/2} < \lambda_2'\}\right\}\right) \\ &\leq \lim_{\lambda\uparrow\infty} \lim_{\lambda_2'\uparrow\infty} \underline{\lim_k}\, \mathbb{P}\big(\Lambda_{k,\varepsilon_k} > \lambda, \Lambda_{k,1-\varepsilon_k} - \Lambda_{k,1/2} < \lambda_2'\big),\end{aligned}$$

which can be proved to be zero by the previous case. Since $\lambda > 0$ was arbitrary, the claim follows.

3.5 Application to CGW(n) trees

In this section we obtain the joint convergence of the rescaled BFW, cumulative profile and profile of CGW(n) trees with offspring distribution in the domain of attraction of an stable law under the product topology. This proves Aldous's conjecture [Ald91b], and recovers the results of Drmota and Gittenberger [DG97], and Kersting [Ker11]. The key point is that CGW(n) trees are a finite mixture of uniform trees with a given degree sequence. This allows us to apply our previous results to CGW(n) trees. First we give some definitions.

Let $\mathbb{T}^{(\infty)}$ be the set of (possibly infinite) plane trees. For any $k \in \mathbb{N}$, let $\mathbb{T}^{(k)}$ be the set of plane trees with height at most k. Consider the restriction map $r_k : \mathbb{T}^{(\infty)} \mapsto \mathbb{T}^{(k)}$, where $r_k t$ is the subtree of $t \in \mathbb{T}^{(\infty)}$, formed by all the vertices up to generation k. A tree $t \in \mathbb{T}^{(\infty)}$ is identified by the sequence $(r_k t, k \geq 0)$.

A *random family tree* τ is a random element of $\mathbb{T}^{(\infty)}$, specified by the sequence $(r_k\tau, k \geq 0)$, where each $r_k\tau$ is a random variable taking values on $\mathbb{T}^{(k)}$, and $r_k\tau = r_k(r_{k+1}\tau)$ for every k.

For a GW tree τ with offspring distribution μ, we have

$$\mathbb{P}(r_k\tau = t) = \prod_{v \in r_{k-1}t} \mu(c(v)) \qquad \forall t \in \mathbb{T}^{(k)}, k \geq 1, \tag{3.28}$$

where the product is taken over all vertices v of t up to generation $k-1$, and $c(v)$ is the number of children of v.

A distribution $\mu = (\mu_n, n \geq 0)$ is called *critical* if $\sum n\mu_n = 1$, and *aperiodic* if the greatest common divisor of all n with $\mu_n > 0$ is one. Let $(\xi_n, n \in \mathbb{N})$ be i.i.d. with law μ. If there exists a sequence of positive numbers $(b_n, n \in \mathbb{N})$ such that

$$\frac{\xi_1 + \cdots \xi_n - n}{b_n} \xrightarrow{d} S_\alpha,$$

where the limit S_α is non-degenerate, we say that μ belongs to the *domain of attraction of a stable law*, DA for short.

It is well-known that if μ is in DA, then there exists an $\alpha \in (1,2]$ such that $b_n = n^{1/\alpha}L(n)$ where $L : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is a *slowly varying function*, that is, $\lim_{x\to\infty} L(tx)/L(x) = 1$ for all $t > 0$. Alternatively, we write DA(α) to indicate the dependence on the parameter.

The hypotheses on μ will be the following:

$\mathbf{H}_\alpha$ The distribution μ is critical, aperiodic and belongs to DA(α) for $\alpha \in (1,2]$.

Throughout this section, the distribution $\mu = (\mu_n, n \geq 0)$ satisfies H_α. Let τ be a GW tree with offspring distribution μ, and denote by $\mathbb{P}_\mu$ its probability distribution. Consider the empirical degree sequence $\hat{n}(\tau) = (\hat{n}_i(\tau), i \geq 0)$ of τ, that is

$$\hat{n}_i(\tau) = \sum_{j=1}^{|\tau|} 1\{c(j) = i\}, \tag{3.29}$$

with $c(j)$ the number of children of individual j. We simply write $\hat{n}$ when the tree τ is obvious from the context. Define the normalized empirical degree sequence $\hat{\mu} = (\hat{\mu}_i, i \geq 0)$ of τ as $\hat{\mu}_i = \hat{n}_i/|\tau|$. The law of the CGW($n$) is denoted by $\mathbb{P}^n_\mu(\cdot) = \mathbb{P}_\mu(\cdot \,|\, |\tau| = n)$, considering the n for which this has sense.

By an *α-stable Lévy process* we mean a stable Lévy process Y of index $\alpha \in (1,2]$, without negative jumps, standardized so we have

$$\mathbb{E}(\exp(-qY(t))) = \exp(-tq^\alpha) \qquad q > 0.$$

Chaumont's path construction (see [Cha97]) of the *normalized α-stable bridge X^b* and the *normalized α-stable excursion X*, states that

$$X^b \stackrel{d}{=} \left(c^{-1/\alpha} Y(ct), t \in [0,1]\right),$$

where $c = \sup\{t \in [0,1] : Y(t) = 0\}$, and

$$X = \left((d-g)^{-1/\alpha}\left(Y_{g+(d-g)s} - Y_g\right), s \in [0,1]\right),$$

where $g = \sup\{s \leq 1 : Y_s = \inf_{u\in[0,s]} Y_u\}$ and $d = \inf\{s > 1 : Y_s = \inf_{u\in[0,s]} Y_u\}$.

Similarly as in Corollary 3.2, it can be proved that the DFW and the BFW of a CGW(n) tree τ_n, have the same distribution. Hence, the following well-known result can be stated in terms of BFWs. A proof can be found in [Duq03].

Lemma 3.9. *Let $W(n) = \sum_1^n \xi_i - n$ and $W(0) = 0$, where $(\xi_i, i \geq 1)$ are i.i.d. with distribution μ satisfying H_α. Then, the rescaled BFW of a CGW(n) tree with offspring distribution μ, converges to the normalized α-stable excursion, that is, under $\mathbb{P}^n_\mu$*

$$X^n = \left(\frac{1}{b_n} W(\lfloor nt \rfloor), t \in [0,1] \,\middle|\, \inf\{j : W(j) = -1\} = n\right) \xrightarrow{d} X. \tag{3.30}$$

Define the set of all possible degree sequences taken by τ when $|\tau| = n$, as

$$DS(n) = \{\mathbf{s} = (N_i, i \geq 0) : n = \sum N_i = 1 + \sum iN_i, N_i^n \geq 0\}.$$

The next lemma relates a GW tree with a uniform tree with a given degree sequence.

Lemma 3.10. *For every $\mathbf{s} \in DS(n)$, the distribution $\mathbb{P}_{\mathbf{s}}$ is the same as the distribution of a GW tree with offspring distribution μ conditioned to have degree sequence $\mathbf{s}$. Therefore, the law of a CGW(n) tree is a finite mixture of the laws $(\mathbb{P}_{\mathbf{s}}, \mathbf{s} \in DS(n))$.*

Proof. The first assertion can be verified directly from the definition (3.28) as follows. Let T_1 and T_2 be two trees on $\mathbb{T}_{\mathbf{s}}$, the set of trees with degree sequence $\mathbf{s}$. Then

$$\mathbb{P}_\mu\left(\tau = T_1, \hat{n}_i(\tau) = N_i, \forall i\right) = \prod_{v \in T_1} \mu_{c(v)} = \prod_{i \geq 0} \mu_i^{N_i} = \mathbb{P}_\mu\left(\tau = T_2, \hat{n}_i(\tau) = N_i, \forall i\right).$$

Hence, the probability that a tree conditioned to have degree sequence $\mathbf{s}$, is the same for T^1 and T^2, so

$$\mathbb{P}_\mu\left(\tau = T | \hat{n}_i(\tau) = N_i, \forall i\right) = \frac{1}{\#\mathbb{T}_{\mathbf{s}}}.$$

To prove the second assertion, for every $\mathbf{s} \in DS(n)$ consider the probabilities

$$\lambda_\mu^{\mathbf{s}} = \mathbb{P}_\mu\left(\hat{n} = \mathbf{s} | \, |\tau| = n\right),$$

and notice that $\sum_{\mathbf{s} \in DS(n)} \lambda_\mu^{\mathbf{s}} = 1$. Let τ be a GW tree with offspring distribution μ, then

$$\mathbb{P}_\mu^n(\tau \in \cdot) = \sum_{\mathbf{s} \in DS(n)} \frac{\mathbb{P}_\mu\left(\hat{n} = \mathbf{s}, |\tau| = n\right)}{\mathbb{P}_\mu\left(|\tau| = n\right)} \mathbb{P}_\mu\left(\tau \in \cdot | \, \hat{n} = \mathbf{s}\right) = \sum_{\mathbf{s} \in DS(n)} \lambda_\mu^{\mathbf{s}} \mathbb{P}_{\mathbf{s}}(\cdot).$$

This proves the lemma. □

We define the rescaled cumulative profile and rescaled profile with offspring distribution μ in DA(α), exactly .

Define the rescaled BFW of τ_n, a CGW(n) tree, as $X^n = S_n^{b_n} \tilde{W}_n$, where $\tilde{W}_n$ is the BFW of τ_n started at one. Using the results of Subsection 3.3.1, the temporal rescaling of the profile will be $a_n = n/b_n$. The rescaled cumulative profile is defined as $C^n(t) = C^n(t_i) + (t - t_i)X^n \circ C^n(t_i)$ for $t_i = i/a_n$ and $t_i \leq t < t_{i+1}$. Thus, for $t \geq 0$

$$Z^n(t) = D_+ C^n(t) \quad \text{and} \quad C^n(t) = \begin{cases} S_{a_n}^n C_n(t - 1/a_n) & t > 1/a_n \\ tX^n(0) & t \leq 1/a_n. \end{cases} \tag{3.31}$$

To prove the convergence of such processes, we need to prove that $\int_{1/2}^1 1/X_s ds$ is finite, a condition imposed in Theorem 3.1. Let $H(\tau_n)$ be the height of τ_n, that is, the last generation having individuals. Theorem 4 of [Kor15a] proves that, for every $\delta \in (0, \alpha)$, there exists positive constants C_1 and C_2 such that for every $u \geq 0$ and every $n \geq 1$, we have

$$\mathbb{P}\left(H(\tau_n) \geq \frac{n}{b_n} u\right) \leq c_1 \exp\left(-c_2 u^\delta\right).$$

This also tells us that the correct scaling is $H^n := b_n H(\tau_n)/n$. The next lemma relies on Kortchemski's bound.

Lemma 3.11. *Let X be the normalized excursion of the α-stable process withouth negative jumps. Then*

$$\int_0^1 \frac{1}{X_s} ds < \infty \qquad a.s.$$

Proof. In Proposition 3.8 we proved that $(C^n, n \in \mathbb{N})$ is sequentially compact for convergence in distribution uniformly on compact sets, and that every subsequential limit is of the form C^Λ (a result which only depended on the convergence of the rescaled BFW's). Denote by $(n_l, l \in \mathbb{N})$ the subsequence such that $C^{n_l} \to C^\Lambda$. Such subsequential limit, has inverse

$$I^\Lambda(t) = \Lambda + \int_0^t \frac{ds}{X(s)} \qquad t \in [0,1].$$

As in Lemma 3.5, for $\varepsilon_n \downarrow 0$, we find deterministic subsequences $(n(l^1(k)), k \in \mathbb{N})$ and $(n(l^2(k)), k \in \mathbb{N})$, such that $\Lambda_{k,1-\varepsilon_k} := \Lambda_{n(l^1(k)),1-\varepsilon_{n(l^2(k))}}$ satisfies

$$\Lambda_{k,1-\varepsilon_k} \overset{a.s.}{\to} \Lambda_{0+} + I(1),$$

with $\Lambda_{0+} = \inf\{t > 0 : C^\Lambda(t) > 0\}$. By the definition in (3.31), note that

$$\Lambda_{k,1-\varepsilon_k} < \sup\{t : C^k(t - 1/a_k) = 1\} = (H(\tau_k) + 1/a_k)/a_k = H^k + 1/a_k^2,$$

for $H^k = H^{n(l^1(k))}$ and $a_k = a_{n(l^2(k))}$. Thus, using Fatou's Lemma, we obtain for every continuity point $u > 0$ of the distribution function of $\Lambda_{0+} + I(1)$

$$\mathbb{P}(\Lambda_{0+} + I(1) > u) \leq \varliminf_k \mathbb{P}\big(\Lambda_{k,1-\varepsilon_k} > u\big) \leq C_1 \exp\left(-C_2(u - 1/a_k^2)^\delta\right).$$

Taking limits we conclude

$$\Lambda_{0+} + \int_0^1 \frac{1}{X_s} ds < \infty \quad a.s. \qquad \square$$

We are ready to prove that jointly, the rescaled BFW, profile, and cumulative profile converge for CGW trees with Pareto offspring distribution.

Proposition 3.10. *Suppose assumption H_α is satisfied for μ. Let τ_n be a CGW(n) tree with law $\mathbb{P}_\mu^n$. Let X^n, C^n and Z^n be the rescaled BFW, cumulative Lamperti transform and Lamperti transform of τ_n respectively, as in* (3.30) *and* (3.31)*. Then, we have the joint convergence*

$$(X^n, C^n, Z^n) \overset{d}{\to} (X, C, Z)$$

under the product Skorohod topology $D(\mathbb{R}_+, \mathbb{R}_+)^3$, where X is a normalized α-stable excursion, C and Z are the cumulative Lamperti transform and Lamperti transform of X.

Proof. Note that Hypotheses (1), (3), (4) and 3.4 of Theorem 3.1 are satisfied in this case. Also, from Lemma 3.9 we have $X^n \to X$ in distribution. This convergence of the BFW's was the result in Section 3.2, and replaces Hypothesis (2) of Theorem 3.1. Note that this is enough to obtain Proposition 3.8, because the latter only depended on the convergence of the rescaled BFWs. Now we prove the equality in distribution of Proposition 3.9 but for CGW(n) trees. Let U^n be a uniform variable on $\{2, \ldots, n\}$ independent of τ, a

GW tree with offspring law μ. Let W be the BFW of τ and $\hat{n}(W)$ its degree sequence. Define $\tilde{W}$ as in Proposition 3.9. As seen from the proof of Lemma 3.10, the probability that τ equals a fixed tree, does not change for trees with the same degree sequence. This, the proof of Proposition 3.9 and the independence with U^n, imply

$$\mathbb{P}_\mu(W = w, U = U) = \mathbb{P}_\mu(\tilde{W} = w, U = U), \tag{3.32}$$

for excursions w having degree sequence $\mathbf{s} \in DS(n)$. To prove the same under the law $\mathbb{P}^n_\mu$ we sum over all the possible degree sequences. First note that the law of (W, U) is uniform on $\mathbb{T}_\mathbf{s} \times [n]$ under the law $\mathbb{P}_\mu(\tau \in \cdot | \hat{n}(\tau) = \mathbf{s})$, for $\mathbf{s} \in DS(n)$. We also have that

$$\lambda_\mu^{\mathbf{s_n}} := \mathbb{P}_\mu\left(\hat{n}(\tau) = \mathbf{s} \middle|\, |\tau| = n\right) = \mathbb{P}_\mu\left(\hat{n}(\tilde{\tau}) = \mathbf{s} \middle|\, |\tilde{\tau}| = n\right) =: \tilde{\lambda}_\mu^{\mathbf{s_n}},$$

using (3.32) and that both τ and $\tilde{\tau}$ have the same degree sequence and size. Joining the previous equalities, we have

$$\begin{aligned}\mathbb{P}^n_\mu(W = w, U^n = u) &= \sum_{\mathbf{s} \in DS(n)} \lambda^\mathbf{s}_\mu \mathbb{P}_\mu\left(W = w, U^n = u \,\middle|\, \hat{n}(W) = \mathbf{s}\right) \\ &= \sum_{\mathbf{s} \in DS(n)} \tilde{\lambda}^\mathbf{s}_\mu \mathbb{P}_\mu\left(\tilde{W} = w, U^n = u \,\middle|\, \hat{n}(\tilde{W}) = \mathbf{s}\right) \\ &= \mathbb{P}^n_\mu(\tilde{W} = w, U^n = u).\end{aligned}$$

Now, the results of Section 3.4 are applied straightforward in this case, because they rely on the previous equality in distribution, in Proposition 3.8 and on trajectorial arguments. Therefore, we also obtain $\mathbb{P}\left(\Lambda \in (0,\infty], \int_0^1 ds/X(s) < \infty\right) = 0$. □

Remark 3.7. *Proposition 3.10 also works for a more general model than GW trees, the so called* simply generated trees *(see [Jan12], on such paper a* degree sequence *is what we called* child sequence*). Indeed, given a fixed* weight sequence $w = (w_k, k \geq 0)$ *of non-negative real numbers with* $w_0 > 0$ *and* $w_k > 0$ *for some* $k \geq 2$*, we define the* weight *of a tree* T *in the set* $\mathbb{T}$ *of all finite plane trees as*

$$w(T) = \prod_{v \in T} w_{c(v)} = \prod_{i \geq 0} w_i^{n_i(T)}, \tag{3.33}$$

where $(n_i(T), i \in \mathbb{N})$ *is the degree sequence of* T*, and the first product is over the set of vertices* v *in* T*. Such trees with weights are called* simply generated trees*. For every* $n \in \mathbb{N}$*, let* τ_n *be a random tree chosen from the set* $\mathbb{T}_n$ *of plane trees with* n *vertices proportional to its weight*

$$\mathbb{P}(\tau_n = T) = \frac{w(T)}{\sum_{T \in \mathbb{T}_n} w(T)} \quad T \in \mathbb{T}_n,$$

when $\sum_{T \in \mathbb{T}_n} w(T) > 0$*.*

From formula (3.33), every tree $T \in \mathbb{T}_n$ with the same degree sequence $\mathbf{s_n}$ has the same distribution. Then, simply generated trees conditioned on having the degree sequence $\mathbf{s_n}$ are uniform on the set of all trees in $\mathbb{T}_n$ with degree sequence $\mathbf{s_n}$, and a similar result as in Proposition 3.10 could be obtained. Nevertheless, it has been proved in [Kor15b], that for the particular case of subcritical GW trees with offspring distribution in DA(α), it is not possible to scale the tree to obtain convergence in the Gromov-Hausdorff topology. As a consequence of Theorem 5 of [Kor15b], Kortchemski proved that there is no non-trivial scaling limit for the contour function of these trees. It is left as an **open problem** to obtain the convergence of the profile for simply generated trees, other than critical CGW trees.

3.6 On Kersting's condition for the convergence of the profile of a CGW tree with Pareto offspring distribution

The objective of this section is to give another proof for the asymptotic thickness of the base of CGW trees with Pareto offspring distribution. Although we proved this result in Section 3.5 as an application of Theorem 3.1, this proof is simpler, as it relies on simple facts about GW trees and α-stable distributions. Note that proving this result, together with the results in Section 3.3, imply the convergence of the profile of such CGW trees, thus giving another proof of Kersting's result [Ker11].

Recall the hypotheses for an offspring distribution $\mu = (\mu(x), x \geq 0)$:

$\mathbf{H}_\alpha$ The distribution μ is critical, aperiodic and belongs to DA(α) for $\alpha \in (1,2)$.

In [Ker11], Kersting gives a condition to prove that CGW(n) trees are not thin at the base.

Lemma 3.12 (Lemma 9 of [Ker11]). *Let C^n be the rescaled cumulative Lamperti transform of a CGW(n) tree τ, having offspring distribution μ. If μ satisfies Hypotheses $\mathbf{H}_\alpha$, then for every $\lambda > 0$*

$$\lim_{\varepsilon \downarrow 0} \overline{\lim_{n}} \mathbb{P}(C^n(\lambda) \leq \varepsilon \,|\, |\tau| = n) = 0.$$

In this section we use the results of Subsection 3.3.3 and [CPGUB13], to give a different proof than in [Ker11].

This is done via a comparison between the first generations of the CGW(n) tree and the *Q-process*. This process is defined in Subsection 3.6.1. Then, in Subsection 3.6.2 we prove convergence of the rescaled Q-process to a *continuous-state branching process with immigration*, and define this limit. Finally, in Subsection 3.6.3 we define the tree having population profile the Q-process, and compare it with the CGW(n) tree, to prove Lemma 9 of [Ker11].

Throughout this section, some of the results we use are stated for (sub)critical offspring distributions, so we say μ satisfies Hypotheses $\mathbf{H}^{(s)}$ if it is (sub)critical, aperiodic and is in $DA(\alpha)$ for $\alpha \in (0,1)$.

3.6.1 Galton-Watson processes conditioned on non-extinction.

We briefly review some well-known results of the Q-process and GW processes with immigration (see [Lam07]). Let $(\bar{Z}(n), n \geq 0)$ be a Galton-Watson process with offspring distribution μ under the law $\mathbb{P}$, having positive and finite mean m. Label the tree in BFO, and let $\xi(j)$ be the number of children of individual $j \in \mathbb{N}$, thus $\xi(j)$ has distribution μ. From Section 1.14 of [AN04], the conditional probabilities $\mathbb{P}(\cdot | \inf\{n : \bar{Z}(n) = 0\} > k)$ converge as k goes to infinity to a probability measure $\mathbb{P}^{\uparrow}$, in the sense of finite-dimensional distributions. This defines a Markov chain with values in $\{1, 2, \ldots\}$ called the *Q-process*. Its transition probabilities are given by

$$\mathbb{P}^{\uparrow}(\bar{Z}_n = j | \bar{Z}_0 = i) = \frac{j}{i} m^{-n} \mathbb{P}(\bar{Z}_n = j | \bar{Z}_0 = i) \qquad i, j \geq 1.$$

This process is related to a *GW process with immigration* (GWI process, for short), which is defined as follows.

Consider independent random walks $(W(j), j \geq 1)$ and $(\tilde{W}(j), j \geq 1)$, having steps $(\xi(j) - 1, j \geq 1)$ and $(\tilde{\xi}(j), j \geq 1)$, where $\xi(1)$ has distribution $(\mu(j); j \geq 0)$ and $\tilde{\xi}(1)$ has distribution $\tilde{\mu} = (\tilde{\mu}(x), x \geq 0)$.

Let $k \geq 1$ be the initial population. Similarly as in Equation (3.1), the population profile and cumulative population profile, are respectively

$$Z(n+1) = k + W \circ C(n) + \tilde{W}(n+1) \quad \text{and} \quad C(n+1) = C(n) + Z(n+1),$$

with $Z(0) = C(0) = k$ and $n \geq 0$. The value $\tilde{W}(m)$ represents the quantity of immigrants arriving at generations less than or equal to m, not counting the initial k members of the population as immigrants.

The process Z just defined is a GWI, having offspring distribution μ and immigration $\tilde{\mu}$. It can be proved that for $s \in (0,1)$ and $i \geq 0$

$$\mathbb{E}_i\left(s^{Z(1)}\right) = f(s)^i g(s),$$

where f and g are the probability generating functions of μ and $\tilde{\mu}$, respectively, and $\mathbb{E}_i$ is the law of the process starting with i individuals. Denote this process as GWI(f,g).

To obtain the relation between Q-processes and GWI processes, set $\tilde{\mu}$ as the *size-biased distribution* of μ:

$$\tilde{\mu}(x) = x\mu(x) \qquad x \geq 1. \tag{3.34}$$

Then, it is true that

$$\mathbb{E}_i^{\uparrow}\left(s^{\bar{Z}(1)-1}\right) = f(s)^{i-1} f'(s)/m. \tag{3.35}$$

The latter implies that $(Z(n)-1, n \geq 0)$ under the law $\mathbb{P}_i^{\uparrow}$ is a GWI$(f, f'/m)$.

The interpretation of the Q-process as a GWI gives us a recursive construction of the *infinite size-biased tree*:

1. Start with a marked individual (the root), which has offspring distribution $\tilde{\mu}$.

2. At generation $n \geq 1$, there is only one marked individual having offspring distribution $\tilde{\mu}$.

3. Choose one children of such marked individual uniformly at random and mark it. That individual also has offspring distribution $\tilde{\mu}$. All the other descendants in that generation have offspring distribution μ.

After removing the marked individuals, this construction provides us with a tree having population profile a GWI, with offspring distribution μ and immigration $\tilde{\mu}$. In Subsection 3.6.3 we give the formal definition of such tree.

3.6.2 Convergence of the rescaled Q-process.

Consider the triangular array $(\xi^n(j)/a_n, j \in [b_n])$ for $n \in \mathbb{N}$, of independent (in the individual series) random variables, where $a_n = n^{1/\alpha} L_a(n)$, for $L_a \in SV$, a slowly varying function. Define the sum

$$S(n) = (\xi^n(1) + \cdots + \xi^n(b_n) - n)/a_n.$$

In the case $\alpha \in (0,1)$ and the $(\xi^n(j)/a_n, j \geq 1, n \geq 1)$ are i.i.d., by IX.8, page 315 of [Fel71], the centering constants $-n/a_n$ are irrelevant for the convergence of $(\xi^n(1) + \cdots + \xi^n(b_n))/a_n$. Assume that for fixed $\varepsilon > 0$

$$\lim_n \max_{j \in [b_n]} \mathbb{P}(|\xi^n(j)/a_n| > \varepsilon) = 0,$$

which says that the terms in the sum are uniformly infinitesimal. Denote by F_n and F_j^n the distribution functions of $S(n)$ and $\xi^n(j)/a_n$, respectively. Define

$$\sigma_n^\varepsilon = \sum_1^{b_n} Var\left(\frac{\xi^n(j)}{a_n}\mathbf{1}_{\xi^n(j)/a_n<\varepsilon}\right) \qquad \varepsilon > 0.$$

The following result is a form of Corollary 15.16 in [Kal02].

Theorem 3.3. *Consider in $\mathbb{R}$ an i.i.d. array $(\xi_{nj}, j \le b_n)_n$ and let X be a Lévy process with characteristics $(0,0,\pi)$. Then $\sum_j \xi_{nj} \xrightarrow{d} X$ iff for any $h > 0$ with $\pi\{|x| = h\} = 0$, we have*

1. $b_n \mathscr{L}(\xi_{n1}) \to \pi$,
2. $b_n \mathbb{E}\left(\xi_{n1}^2; |\xi_{n1}| \le h\right) \to \int_{|x|\le h} x^2 \pi(dx)$
3. $b_n \mathbb{E}(\xi_{n1}; |\xi_{n1}| \le h) \to \int_{h<|x|\le 1} x\pi(dx)$,

where $\mathscr{L}(\xi_{n1})$ is the law of ξ_{n1}.

Let the critical offspring distribution μ be of the form

$$\mu(n) = c_\alpha \frac{L(n)}{n^{\alpha+1}}, \tag{3.36}$$

where $\alpha \in (1,2)$, $L \in SV$ and $c_\alpha > 0$ is a constant depending on α. By Theorem 5.10 we have that $n\mu(n)/\mu((n,\infty)) \to \alpha$, hence the tail of μ is of regular variation with index α. Without loss of generality, write $\mu((n,\infty)) = c'_\alpha n^{-\alpha} L(n)$, for some constant c'_α and using the same slowly varying function L. By the same theorem, the tail of the size-biased distribution satisfies $\lim n^2 \mu(n)/\tilde{\nu}((n,\infty)) = \alpha - 1$, so write $\tilde{\nu}((n,\infty)) = c''_\alpha n^{1-\alpha} L(n)$, for some constant c''_α. The latter implies that, if $\tilde{W}$ is a random walk with step distribution $\tilde{\mu}$, then $\tilde{W}(n)/d_n$ converges to an $(\alpha-1)$-stable variable, say $Y(\alpha-1)$, where $d_n \in RV(1/(1-\alpha))$. In order to obtain convergence of the rescaled Q-process, we need the following lemma, whose proof is postponed to the Appendix.

Lemma 3.13. *Let $\tilde{W}$ be a random walk having increments $(\tilde{\xi}(j), j \ge 1)$ with distribution $\tilde{\mu}$. Suppose μ satisfies Hypotheses $\mathbf{H}^{(s)}$ and is defined in (3.36). Then*

$$\frac{\tilde{W}(\lfloor n/a_n \rfloor)}{a_n} \to Y(\alpha - 1).$$

Let $d_n = \lfloor n/a_n \rfloor$. If we set

$$X^n(t) = \frac{W(nt)}{a_n} \quad \text{and} \quad Y^n(t) = \frac{\tilde{W}(d_n t)}{a_n} \qquad t \ge 0,$$

then we have $X^n \to X$ and $Y^n \to Y$ under the Skorohod topology, where X is an α-stable process without negative jumps, and Y is an independent $(\alpha-1)$-stable subordinator.

In analogy with Subsection 3.3.1, we define the appropriate rescalings for the Lamperti and cumulative Lamperti transform of the Q-process.

Let $k_n \geq 1$ such that $k_n/a_n \to x \geq 0$. The profile and cumulative profile of a GWI$(f, f'/m)$ process with k_n initial individuals are

$$Z_n(m+1) = k_n + W \circ C_n(m) + \tilde{W}(m+1) \quad \text{and} \quad C_n(m+1) = C_n(m) + Z_n(m+1),$$

with $Z_n(0) = C_n(0) = k_n$ and $m \geq 0$. We work with the linear interpolation of C_n and its right-hand derivative, and for simplification, denote them by C_n and Z_n respectively.

Consider the processes

$$S_{d_n}^{a_n} Z_n(t) = \frac{1}{a_n} Z_n(d_n t) \quad \text{and} \quad S_{d_n}^{n} C_n(t) = \frac{1}{n} C_n(d_n t) \quad t \geq 0.$$

Then, define the rescaled cumulative Lamperti and Lamperti transform as

$$C^n(t) = \frac{k_n}{a_n} t + \int_0^t X^n \circ C^n(s) ds + \int_0^t Y^n(s) ds,$$

and

$$Z^n(t) = \frac{k_n}{a_n} + X^n \circ C^n(t) + Y^n(t).$$

The latter has the continuum analogous

$$Z(t) = x + X \circ C(t) + Y(t), \quad C(t) = \int_0^t Z(s) ds, \tag{3.37}$$

where Z is a continuous-state branching process with immigration. We now briefly recall simple aspects of *continuous-state branching processes* (CB processes) and *continuous-state branching processes with immigration* (CBI processes). Those models were introduced in [Jr58, Lam67b, Sil68, KW71], see also [Lam02, CPGUB13] for other detailed descriptions. A $[0, \infty]$-valued strong Markov process Z is a CB process if its paths are càdlàg and the sum of two independent copies started at x and at y has the same law as the process started at $x + y$. For a CB process, the states 0 and ∞ are absorbing. Its Laplace exponent is given by

$$\mathbb{E}_x(\exp(-\theta Z_t)) = \exp(-x u_t(\theta)) \qquad \forall x > 0, t, \theta \geq 0,$$

where u_t satisfies

$$\frac{\partial u_t}{\partial t}(\theta) + \psi(u_t(\theta)) = 0, \tag{3.38}$$

with initial condition $u_0(\theta) = \theta$, and for $\lambda \geq 0$

$$\psi(\lambda) = -q - d\lambda + \sigma^2 \lambda^2/2 + \int_{\mathbb{R}_+} \left(e^{-\lambda x} - 1 + \lambda x \mathbf{1}_{x<1} \right) \Pi(dx),$$

where $q, \sigma \geq 0$, $d \in \mathbb{R}$, and Π is a measure supported on $\mathbb{R}_+$ satisfying $\int_{\mathbb{R}_+} (1 \wedge x^2) \Pi(dx) < \infty$. The function ψ is called the *branching mechanism* of the CB process.

A $[0, \infty]$-valued strong Markov process Z is a CBI process with branching mechanism ψ and *immigration mechanism* ϕ (CBI(ψ, ϕ) for short) if the following holds: its paths are càdlàg and

$$\mathbb{E}_x(\exp(-\theta Z_t)) = \exp\left(-x u_t(\theta) - \int_0^t \phi(u_{t-s}(\theta)) ds \right) \qquad \forall x, t > 0, \theta \geq 0,$$

where u_t is the unique solution to (3.38), and

$$\phi(\theta) = \delta\theta + \int_{\mathbb{R}_+} (1 - e^{-\theta x})\Upsilon(dx) \qquad \theta \geq 0,$$

with $\delta \geq 0$, and Υ a measure concentrated on $\mathbb{R}_+$ such that $\int_{\mathbb{R}_+} (1 \wedge x)\Upsilon(dx) < \infty$.

By Proposition 2 and Theorem 2 of [CPGUB13], Equation (3.37) has a unique solution Z, which is a CBI(ϕ, ϕ') starting at x. Using this, we can prove the convergence of the Q-process to such CBI.

Lemma 3.14. *Assume $k_n/a_n \to x \geq 0$ and that the offspring distribution μ satisfies Hypotheses $\mathbf{H}_\alpha$. Then, the rescaled Lamperti (resp. cumulative Lamperti) transform of the Q-process starting with k_n individuals, converges to Z (resp. C) defined by (3.37), under the Skorohod topology (resp. uniformly on compact sets). The process Z is a CBI(ϕ, ϕ') with $\phi(\lambda) = \lambda^\alpha$ for $\lambda \geq 0$.*

Proof. By Skorohod's representation theorem, we work on a space where $X^n \to X$ and $Y^n \to Y$ a.s. Note that Y is a strictly increasing Lévy process. Using Proposition 4 and Theorem 3 of [CPGUB13], we have that

$$Z^n \to Z \qquad \text{and} \qquad C^n \to C,$$

where the first convergence holds under the Skorohod topology and the second uniformly on compact sets. $\square$

3.6.3 The CGW(n) tree and the size-biased tree compared up to the height they first have εn individuals.

In this subsection we compare a CGW(n) tree with the size-biased GW tree. This is used to prove that a small probability for the size-biased tree of having a thin base, implies a small probability for the CGW(n) tree.

Size-biased GW trees.

We state the construction of the *infinite size-biased GW tree* $\tau^{(\infty)}$, from a (sub)critical GW tree τ. Recall the definition of the restriction map, given in (3.28).

Proposition 3.11 ([Kes86])**.** *Let Z_n be the number of individuals in the nth generation of a GW tree τ having (sub)critical offspring distribution μ, with $\mu(0) < 1$. Then*

$$\mathscr{L}(\tau | Z_n > 0) \to \mathscr{L}(\tau^{(\infty)})$$

as $n \to \infty$, where the law on the right-hand side is specified by

$$\mathbb{P}\left(r_k\tau^{(\infty)} = t\right) = \frac{z_k(t)}{m^k}\mathbb{P}(r_k\tau = t) \qquad \forall t \in \mathbb{T}^{(k)}, k \geq 0, \tag{3.39}$$

and $z_k(t)$ denotes the number of individuals in generation k of t.

Fix any $n \in \mathbb{N}$ and $\varepsilon \in (0,1)$. Define $\Lambda^n_\varepsilon := \inf\{k : C_n(k) > \varepsilon n\}$, and denote by $\mathbb{T}_{n,\varepsilon}$ the set of trees t such that $C_t(h(t)-1) \leq \varepsilon n < C_t(h(t))$, where $h(t)$ is the height of the tree, and C_t is the cumulative profile of t. In the next theorem, for a GW tree τ conditioned to have size n, we consider it up to height Λ^n_ε. This restricted tree takes values on $\mathbb{T}_{n,\varepsilon}$.

Theorem 3.4. *Let τ be a GW tree with offspring distribution satisfying Hypotheses $\mathbf{H}_\alpha$. Then, there exists a constant $c_\varepsilon > 0$ (depending only on α and ε) such that*

$$\overline{\lim_n} \sup_{B\subset\mathbb{T}_{n,\varepsilon}} \mathbb{P}\big(r_{\Lambda^n_\varepsilon}\tau \in B | s(\tau)=n\big) \le c_\varepsilon \overline{\lim_n} \sup_{B\subset\mathbb{T}_{n,\varepsilon}} \mathbb{P}\Big(r_{\Lambda^n_\varepsilon}\tau^{(\infty)} \in B\Big), \tag{3.40}$$

for fixed $\varepsilon \in (0,1)$. Also, $\overline{\lim}_{\varepsilon\downarrow 0} c_\varepsilon < c$ for some finite c.

Proof. Let $\varepsilon' \in (0,\varepsilon)$. First we prove inequality (3.40) restricted to trees in $\mathbb{T}_{n,\varepsilon}$, such that $(\varepsilon-\varepsilon')n \le C_t(h(t)-1)$. On the event $\{r_{\Lambda^n_\varepsilon}\tau = t\}$, we have $\Lambda^n_\varepsilon = h(t)$, and so $C_n(\Lambda^n_\varepsilon - 1) = C_t(h(t)-1)$. Let $z_{h(t)}$ be the number of individuals in generation $h(t)$ of t. By the Markov property, there are $z_{h(t)}$ GW trees growing from generation Λ^n_ε (denote this GW forest by $\mathscr{F}_{z_{h(t)}}$), and this is independent of $r_{\Lambda^n_\varepsilon}\tau$. Then, using Otter-Dwass formula and Equation (3.39),

$$\begin{aligned}
&\mathbb{P}\big(r_{\Lambda^n_\varepsilon}\tau = t, s(\tau)=n\big)\\
&= \mathbb{P}\big(r_{\Lambda^n_\varepsilon}\tau = t\big)\,\mathbb{P}\Big(s(\mathscr{F}_{z_{h(t)}}) = n - C_t(h(t)-1)\Big)\\
&= z_{h(t)}\mathbb{P}\big(r_{\Lambda^n_\varepsilon}\tau = t\big)\,\frac{1}{n-C_t(h(t)-1)}\mathbb{P}\big(W(n-C_t(h(t)-1)) = -z_{h(t)}\big)\\
&= \mathbb{P}\Big(r_{\Lambda^n_\varepsilon}\tau^{(\infty)} = t\Big)\,\frac{1}{n-C_t(h(t)-1)}\mathbb{P}\big(W(n-C_t(h(t)-1)) = -z_{h(t)}\big).
\end{aligned}$$

It should be noted that Equation (3.39) can be applied since $\Lambda^n_\varepsilon = h(t)$ is deterministic. Dividing by $\mathbb{P}(s(\tau)=n)$ on both sides

$$\mathbb{P}\big(r_{\Lambda^n_\varepsilon}\tau = t | s(\tau)=n\big) = \mathbb{P}\Big(r_{\Lambda^n_\varepsilon}\tau^{(\infty)} = t\Big)\frac{n}{n-C_t(h(t)-1)}\frac{\mathbb{P}\big(W(n-C_t(h(t)-1)) = -z_{h(t)}\big)}{\mathbb{P}(W(n)=-1)}. \tag{3.41}$$

We use the local limit theorem 5.8 on the right-hand side. From I.4 page 21 of [Zol86], it is known that g is uniformly bounded and continuous on $\mathbb{R}$. Then, there exists constants $c_1, c_2 > 0$ and $N\in\mathbb{N}$ such that for every $n \ge N$

$$\mathbb{P}\Big(W(n-C_t(h(t)-1)) = -z_{h(t)}\Big) \le \frac{c_1}{a_{n-C_t(h(t)-1)}} \qquad \text{and} \qquad \mathbb{P}(W(n)=-1) \ge \frac{c_2}{a_n}.$$

Since by assumption $(\varepsilon-\varepsilon')n \le C_t(h(t)-1) < \varepsilon n$, then

$$\frac{n}{n-C_t(h(t)-1)}\frac{\mathbb{P}\big(W(n-C_t(h(t)-1)) = -z_{h(t)}\big)}{\mathbb{P}(W(n)=-1)} \le \frac{c_1}{c_2}\left(\frac{1}{1-\varepsilon}\right)^{1+1/\alpha}\frac{L(n)}{L(n-C_t(h(t)-1))}.$$

Using the Potter bounds (see [BGT89, Theorem 1.5.6]), there exists $N\in\mathbb{N}$ depending only on ε,ε' such that for every $n\ge N$

$$\frac{L(n)}{L(n-C_t(h(t)-1))} \le 2\frac{n}{n-C_t(h(t)-1)} \le 2\frac{n}{n-\varepsilon n}.$$

This, together with (3.41) implies (3.40) but for B restricted to trees satisfying $(\varepsilon-\varepsilon')n \le C_t(h(t)-1)$. To get rid of this, note that by definition $C_n(\Lambda^n_\varepsilon) \ge \varepsilon n$, therefore

$$A_n := \{C_n(\Lambda^n_\varepsilon - 1) < (\varepsilon-\varepsilon')n\} \subset \big\{Z_n(\Lambda^n_\varepsilon) = C_n(\Lambda^n_\varepsilon) - C_n(\Lambda^n_\varepsilon - 1) \ge \varepsilon' n\big\}. \tag{3.42}$$

In particular, for the CGW(n) we have $W \circ C_n(\Lambda^n_\varepsilon - 1) \geq \varepsilon' n$. The latter implies that on A_n

$$\frac{a_n}{n} \sup_{0 \leq t \leq \varepsilon - \varepsilon'} X^n(t) = \sup_{0 \leq t \leq \varepsilon - \varepsilon'} \frac{1}{n} W(nt) \geq \varepsilon'.$$

Without loss of generality, we can assume ε' is a point of continuity for the distribution of $\lim_n \sup_{t \leq \varepsilon - \varepsilon'} X^n(t) = \sup_{t \leq \varepsilon - \varepsilon'} X(t)$. Hence, the probability of A_n goes to zero as $n \to \infty$.

Letting $\mathbb{T}_{n,\varepsilon,\varepsilon'}$ be the subset of trees t in $\mathbb{T}_{n,\varepsilon}$ with $(\varepsilon - \varepsilon')n \leq C_t(h(t) - 1)$, then

$$\begin{aligned}
&\overline{\lim_n} \sup_{B \subset \mathbb{T}_{n,\varepsilon}} \mathbb{P}\big(r_{\Lambda_{n,\varepsilon}} \tau \in B | s(\tau) = n\big) \\
&\leq \overline{\lim_n} \sup_{B \subset \mathbb{T}_{n,\varepsilon,\varepsilon'}} \mathbb{P}\big(r_{\Lambda_{n,\varepsilon}} \tau \in B | s(\tau) = n\big) + \overline{\lim_n} \mathbb{P}\big(C_n(\Lambda_{n,\varepsilon} - 1) < (\varepsilon - \varepsilon')n | s(\tau) = n\big) \\
&\leq \overline{\lim_n} \sup_{B \subset \mathbb{T}_{n,\varepsilon,\varepsilon'}} \mathbb{P}\big(r_{\Lambda_{n,\varepsilon}} \tau \in B | s(\tau) = n\big).
\end{aligned}$$

Since the right-hand side is bounded from above by the left hand side, the previous inequalities are equalities. Using the same reasoning, it is easy to prove the above equality for the Q-process. Indeed, on the set A_n in Equation (3.42), but defined for the cumulative profile of the Q-process we have

$$\frac{a_n}{n} \sup_t Z^n(t) \geq \frac{Z_n(\Lambda_{n,\varepsilon})}{n} \geq \varepsilon'.$$

This probability goes to zero as $n \to \infty$ by the convergence of the Lamperti transform of Lemma 3.14. It follows that

$$\overline{\lim_n} \sup_{B \subset \mathbb{T}_{n,\varepsilon}} \mathbb{P}\Big(r_{\Lambda_{n,\varepsilon}} \tau^{(\infty)} \in B\Big) = \overline{\lim_n} \sup_{B \subset \mathbb{T}_{n,\varepsilon,\varepsilon'}} \mathbb{P}\Big(r_{\Lambda_{n,\varepsilon}} \tau^{(\infty)} \in B\Big),$$

This proves (3.40). □

Now, we are ready to prove Kersting's condition, by bounding the event that the cumulative profile of the CGW(n) tree is small, with the same event but for the Q-process. The latter converges by Lemma 3.14, to the event that the cumulative profile of a CBI is small, which happens with small probability.

Proof of Lemma 3.12. Recall that Λ^n_ε is the first height that the tree has more than εn descendants. Define $\lambda_n = \lfloor \lambda n / a_n \rfloor$, and consider as before $\varepsilon' \in (0, \varepsilon)$. Denoting by $\mathbb{T}_{n,\varepsilon,\varepsilon',\lambda}$ the set of trees t in $\mathbb{T}_{n,\varepsilon,\varepsilon'}$ such that $C_t(\lambda_n) \leq \varepsilon n$, then for every n big enough

$$\begin{aligned}
&\mathbb{P}\big(C^n(\lambda) \leq \varepsilon, (\varepsilon - \varepsilon')n \leq C_n(\Lambda^n_\varepsilon - 1) | s(\tau) = n\big) \\
&= \sum_{t \in \mathbb{T}_{n,\varepsilon,\varepsilon',\lambda}} \mathbb{P}\big(r_{\Lambda^n_\varepsilon} \tau = t | s(\tau) = n\big) \\
&\leq c_\varepsilon \sum_{t \in \mathbb{T}_{n,\varepsilon,\varepsilon',\lambda}} \mathbb{P}\Big(r_{\Lambda^n_\varepsilon} \tau^{(\infty)} = t\Big) \\
&= c_\varepsilon \mathbb{P}^\uparrow \left(C^n(\lambda) \leq \varepsilon, (\varepsilon - \varepsilon')n \leq C_n(\Lambda^n_\varepsilon - 1)\right),
\end{aligned}$$

using the first part of the proof of Theorem 3.4. Since the CGW(n) process on the event $\{C^n(\lambda) \leq \varepsilon, C_n(\Lambda^n_\varepsilon - 1) < (\varepsilon - \varepsilon')n\}$ converges to zero as in the proof of Theorem 3.4, it follows that

$$\overline{\lim_n} \mathbb{P}(C^n(\lambda) \leq \varepsilon | s(\tau) = n) \leq c_\varepsilon \overline{\lim_n} \mathbb{P}^\uparrow \left(C^n(\lambda) \leq \varepsilon\right) \leq c_\varepsilon \mathbb{P}^\uparrow \left(C(\lambda) \leq \varepsilon\right),$$

by Portmanteau's theorem and Lemma 3.14. The right-hand side goes to zero as $\varepsilon \downarrow 0$ because $C(\lambda) > 0$ $\mathbb{P}^{\uparrow}$-a.s. and c_ε is bounded from above. $\square$

Chapter 4

ON MULTITYPE RANDOM FORESTS WITH A GIVEN DEGREE SEQUENCE, THE TOTAL POPULATION OF BRANCHING FORESTS AND ENUMERATIONS OF MULTITYPE FORESTS

In this chapter we introduce the model of multitype random forests chosen uniformly at random from the set of multitype forest with a given degree sequence (MFGDS). The unitype case was studied in [BM14b]. By mixing our model, one obtains multitype Galton-Watson (MGW) forests conditioned by their degree sequence. The construction of MFGDS is done using the results in [CL16], and a novel path transformation on multidimensional discrete exchangeable increments processes, which is a generalization of the Vervaat transformation [Ver79]. We also obtain the joint law of the number of individuals by types in a MGW forest, thus, generalizing the Otter-Dwass formula (which is shown to hold in the unitype case in [Ott49, Dwa69]). This allows us to obtain enumerations of multitype forests with a combinatorial structure (plane, labeled and binary forest), having a prescribed number of roots and individuals by types. Finally, under certain hypotheses, we give an algorithm to simulate MGW forests conditioned by the number of individuals of each type (cMGW), generalizing Devroye's algorithm [Dev12] for the unitype case.

4.1 Introduction

Bienaymé-Galton-Watson forests (GW forests) are a simplified model for the genealogy of populations, where individuals have the same reproduction law. A natural generalization of such model are the Multitype Galton-Watson forest (MGW forests), used when several types of individuals coexist, leading to different reproduction rates. Such multitype forest are used in a variety of areas, for example, having applications in biology, demography, genetics, medicine, epidemics, and language theory (see [Har63, San71, Jag75, GP75, CKB$^+$, AJ97, All11]). MGW forest have several applications also for pure mathematics. Miermont [Mie08] has proved that under certain conditions, MGW forests converge also to the CRT. A very nice improvement of such result, would be to generalize the results of [LGLJ98, DLG02] to obtain the covergence of MGW forests to *multitype Lévy forests*, as well as generalizations of the Ray-Knight theorems.

Conditioned random forests also provides us with several applications. In the unitype setting, Le Gall [LG05] proved the convergence of GW trees conditioned to have a fixed number of vertices, towards Aldou's Continuum Random Tree (CRT), see [Ald91a]. In [Ald91b], GW trees conditioned to have size n with Poisson offspring distribution (equivalently, uniform labeled trees on $[n] := \{1,\dots,n\}$) where used to study the component sizes in the Erdős-Rényi random graph [ER60] (see also page 379 of [ABBG12]). Another way to condition a forest is by its degree sequence [BM14b, Lei17], that is, the number of individuals having a fixed number of offspring. With no doubt, one application of this model is to study invariance principles for random graphs with a prescribed degree sequence (see the discussion in [BM14b]). The interest in such random graphs lies in its matching with observations in large real-world networks (having features not present in the Erdős-Rényi random graph). For example, with this model, one can obtain forests having degree sequence with power law tails.

Another reason to study MFGDS, is that they are a more general model than conditioned MGW forests. Indeed, under an independence assumption on the progeny distribution, a MGW forest conditioned by its types can be written as a linear combination of the laws of MFGDS (see Section 4.4). Thus, some results of the latter model can be recovered for the former.

Most of the papers on multitype conditioned forests are interested on proving the convergence towards a limiting object, such as the multitype generalization of Kesten's infinite forest or a continuum random forest ([Nak78, Mie08]). Some authors studying conditioned multitype forests are [DJ08, PÍ11, PÍ16, CL16, Wan14, ADG18, Ste18]. But simulating such conditioned forests is not trivial, since the independence assumption is generally lost, or the conditioning event is too complicated. Some papers giving explicit algorithms for generating MGW trees are [PV91, AS95, Şte98]. We emphasize that neither constructions of MGW forests conditioned by the total number of individuals by types, nor uniform multitype forests with a given degree sequence, are available in the literature. In this chapter we define both forests and provide easy algorithms for their simulation.

We state the known results in the unidimensional case, and how we generalize them. Consider a unidimensional degree sequence $\mathbf{S}$, that is, a sequence of integers $\mathbf{S} = (N_i, i \geq 0)$ such that $s := 1 + \sum iN_i = \sum N_i$. From such sequence we can obtain the child sequence $\mathbf{c}(\mathbf{S}) := \mathbf{c} = (c_1,\dots,c_s)$, a vector with N_0 zeros, N_1 ones, and so on. In the paper [BM14b], the authors give an algorithm to construct, from a discrete exchangeable increment (EI) process, a uniform tree from the set of trees with given degree sequence $\mathbf{S}$ as follows: define W^b a walk with increments $(c \circ \pi(j) - 1, j \in [s])$, where π is a uniform random permutation on $[s]$, and let W be the walk with increments $(c \circ \pi(i^* + j) - 1, j \in [s])$, where $i^* + j$ is considered modulo s and i^* is the first time W^b reaches its minimum value (that is, apply the Vervaat transformation [Ver79]). From such excursion W we can recover the desired tree. This algorithm was extended to unitype forests in [Lei17], the distinction is that one has to carefully chose the cyclical permutations that lead bridges to excursions.

We extend the previous construction to multitype forest, uniformly chosen from the set of multitype forest with a given degree sequence. In order to do this, we generalize the above algorithm: define a multitype degree sequence, construct *dxd* exchangeable increments (EI) processes, and apply to them a generalized Vervaat transform. We use the results in [CL16] to know how many cyclical permutations lead to paths coding a multitype forest.

Also, using the results in [CL16], we obtain the law of the total population by types of a MGW forest, under certain conditions. The unitype case is known as the Otter-Dwass formula [Ott49, Dwa69]. This formula says that the total number of individuals in a GW forest τ_k with k trees, say $\#\tau_k$, having offspring distribution ν is given by

$$\mathbb{P}(\#\tau_k = n) = \frac{k}{n}\mathbb{P}(X_n = n-k),$$

where X is a random walk with step law ν.

It turns out that, using the law of $\#\tau_k$, it has been obtained the total number of plane, labeled and binary forests having k trees and n vertices, see [Pit98]. This chapter generalizes elementary connections between the combinatorial and probabilistic results about enumerations of forests and lattice paths given in [Pit98]. An example of such connection in the unitype case is the following. Let $\mathscr{F}^{plane}_{k,n}$ be a uniformly distributed forest from the set of plane forests having k trees and n individuals; let $\mathscr{G}_{k,p}$ be a GW forest with k trees and Geometric(p) offspring distribution, with $p \in (0,1)$. Then we have

$$\mathscr{F}^{plane}_{k,n} \stackrel{d}{=} \left(\mathscr{G}_{k,p} \,|\, \#\mathscr{G}_{k,p} = n\right) \qquad \text{and} \qquad \mathbb{P}\left(\mathscr{G}_{k,p} = F \,|\, \#\mathscr{G}_{k,p} = n\right) = \frac{1}{\frac{k}{n}\binom{2n-k-1}{n-k}},$$

for every plane forest F with k trees and n individuals. Similar equalities in distribution are available for the Poisson and the Bernoulli distribution. We generalize the above formulas, obtaining the number of multitype plane, labeled, and binary forests having an specified number of roots and individuals of each type.

Finally, we give an algorithm to simulate MGW forests conditioned to have n_i individuals of type i, for $i \in [d]$, using our constructions. Indeed, an algorithm of Devroye [Dev12] simulates a GW tree conditioned to have size n, using a uniform tree with a given degree sequence; thus, we use both of our constructions to generalize such algorithm. Devroye's algorithm is: generate a multinomial vector $\mathbf{S} = (N_0, N_1, \ldots)$ with parameters $(n; \nu_0, \nu_1, \ldots)$, repeat until $1 + \sum iN_i = n$ and apply the algorithm to generate a uniform tree from the set of trees with degree sequence $\mathbf{S}$. Our algorithm is analogous: generate $d \times d$ multinomial distributions with laws $(n_i; \nu_{i,j}(0), \nu_{i,j}(1), \ldots)$ until they form a multitype degree sequence, and apply the algorithm to generate a uniform multitype forest with such given degree sequence.

4.1.1 Preliminaries

Coding of unititype and multitype forests

A rooted plane tree T is a connected graph with no cycles having a distinguished vertex, together with a natural identification of each vertex by a finite sequence of non-negative integers (denoting its location on the tree). The root of T will be denoted by $r(T)$, or simply r. A rooted plane forest is a directed planar graph whose connected components are rooted plane trees, those are ordered according to its roots. We will only consider finite rooted plane forests in the following.

We consider forests where each tree is labeled according to the *breadth-first order* (BFO), that is, from the initial individual to the top, traverse each tree generation by generation from left to right. We define the vector with ith component, the number of individuals having i children, for any $i \geq 0$.

Definition 4.1. *Let T be a tree. The* degree sequence $\mathbf{S} = (N_0, N_1, \ldots)$ *of T is a vector with*

$$N_i := N_i(T) = |\{u \in T : c(u) = i\}|,$$

where $c(u)$ is the number of children of individual u.

As an example, the tree in Figure 3.1 has degree sequence $(5,2,2,1)$.

Let $(u_i)_i$ be the individuals in BFO of a plane forest. It is well known that the walk with increments $(c(u_i)-1, i \geq 1)$ codes the branching forest, that is, determines its structure completely (see [Pit06, Lemma 6.2]). This is called the *breadth-first walk* (BFW) of the forest. Now, we briefly recall the analogous coding in the multitype case, following [CL16].

Define $[n] = \{1,\dots,n\}$ and $[n]_0 = \{0,1,\dots,n\}$ for $n \in \mathbb{N}$. For a forest F, let $c_F : v(F) \mapsto [d]$ be an application from the set of vertices of F to $[d]$, such that the children of each vertex are ordered by color, that is, if $u_i, u_{i+1}\dots, u_{i+j} \in v(F)$ have the same parent, then $c_F(u_i) \leq c_F(u_{i+1}) \leq \cdots \leq c_F(u_{i+j})$. The couple (F, c_F) is a *d-multitype forest*. A *subtree of type i* of (F,c_F), denoted by $T^{(i)}$, is a maximal connected subgraph of (F,c_F) whose all vertices are of type i. Subtrees of type i are ranked according to the order of their roots, and with this ordering, we define the *subforest of type i* of (F,c_F) as $F^{(i)} = \{T_1^{(i)},\dots,T_k^{(i)},\dots\}$ For $u \in v(F)$, denote by $p_i(u)$ the number of children of type i of u. Let $n_i \geq 0$ be the number of vertices in the subforest $F^{(i)}$ of (F,c_F). The coding of the forest is the d-dimensional chain $x^{(i)} = (x^{i,1},\dots,x^{i,d}) \in \mathbb{Z}^d$ with length $n_i \in \mathbb{N}$, defined for $0 \leq n \leq n_i - 1$ by

$$x_{n+1}^{i,j} - x_n^{i,j} = p_j(u_{n+1}^{(i)}) - \mathbf{1}_{i=j} \qquad i,\, j \in [d]. \tag{4.1}$$

We set $x_0^{(i)} = 0$. The set $(u_n^{(i)}; n \geq 1)$ is the labeling of the subforest $F^{(i)}$ in its own breadth-first order.

The cyclical permutations that we use are the following. For $n \in \mathbb{N}$, consider any application $y : [n]_0 \mapsto \mathbb{Z}^d$ with $y(0) = 0$. The n-cyclical permutations of y are the n applications $\theta_{n,q}(y)$, for $q \in [n-1]_0$ given by

$$\theta_{q,n}(y) = \begin{cases} y(j+q) - y(q) & j \leq n-q \\ y(j+q-n) + y(n) - y(q) & n-q \leq j \leq n. \end{cases}$$

We say that the path $y : \mathbb{N} \mapsto \mathbb{Z}$ is a *downward skip-free chain*, if $y_{k+1} - y_k \in \mathbb{Z}_+ \cup \{-1\}$. The possible paths that a coding of multitype forest can take are the following.

Definition 4.2. *Fix any $\mathbf{n} = (n_1,\dots,n_d) \in \mathbb{Z}_+^d$, and define S_d as the set of $[\mathbb{Z}^d]^d$-valued sequences $x = (x^{(1)},\dots,x^{(d)})$ such that for all $i \in [d]$, $x^{(i)} = (x^{i,1},\dots,x^{i,d})$ is a $\mathbb{Z}^d$-valued sequence starting at zero of length n_i, and where $x^{i,j} = (x_k^{i,j}, k \in [n_i]_0)$ is non-decreasing when $i \neq j$, and a downward skip-free chain when $i = j$.*

The $\mathbf{n}$-cyclical permutations of $x \in S_d$ are given by

$$\theta_{\mathbf{q},\mathbf{n}}(x) := (\theta_{q_1,n_1}(x^{(1)}),\dots,\theta_{q_d,n_d}(x^{(d)})) \qquad \forall\, \mathbf{q} = (q_1,\dots,q_d) \text{ such that } 0 \leq \mathbf{q} \leq \mathbf{n} - 1_d,$$

with $1_d = (1,\dots,1)$ of length d. Each sequence $\theta_{\mathbf{q},\mathbf{n}}(x)$ will be called a cyclical permutation of x.

For $\mathbf{m},\mathbf{n} \in \mathbb{Z}_+^d$, write $\mathbf{m} < \mathbf{n}$ if $\mathbf{m} \leq \mathbf{n}$ (the inequality understood component-wise) and if there exists i such that $m_i < n_i$. Sequences $x \in S_d$ will be denoted by $x = (x_k^{i,j}, k \in [n_i]_0, i,j \in [d])$, and the vector $\mathbf{n} = (n_1,\dots,n_d) \in \mathbb{Z}_+^d$, is called the length of x. Fix any such x of length $\mathbf{n}$, and $\mathbf{r} = (r_1,\dots,r_d) \in \mathbb{Z}_+^d$ with $\sum r_i > 0$. We say that the system $(\mathbf{r},x)$ admits a solution if there exists $\mathbf{m} \leq \mathbf{n}$ such that

$$r_j + \sum_{i=1}^{d} x^{i,j}(m_i) = 0 \qquad \forall\, j \in [d]. \tag{4.2}$$

If there is no smaller solution $\mathbf{m} < \mathbf{n}$ for the system $(\mathbf{r},\theta_{\mathbf{q},\mathbf{n}}(x))$, then we call $\theta_{\mathbf{q},\mathbf{n}}(x)$ a *good cyclical permutation*. It is proved in [CL16] that only such good cyclical permutations code multitype forests, and the next lemma tells us how many there are.

Lemma 4.1 (Multivariate Cyclic Lemma [CL16]). *Let $x \in S_d$ with $x^{i,i}(n_i) \neq 0$ for every $i \in [d]$. Consider the system $(\mathbf{r}, x)$ with solution $\mathbf{n}$ as above. Then, the number of good cyclical permutations of x is $det((-x^{i,j}(n_i))_{i,j \in [d]})$.*

Since in most of the cases, we fix the number of roots or number of individuals of each type, we need the following definition.

Definition 4.3 (Root-type and individuals-type). *We say a multitype plane forest with $d \in \mathbb{N}$ types has* root-type $\mathbf{r} = (r_1, \ldots, r_d) \in \mathbb{N}^d$, *if it has r_i roots of type i for $i \in [d]$, with $\mathbf{r} > 0$. Also, it has* individuals-type $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{N}^d$ *if it has n_i individuals of type i, for $i \in [d]$.*

Multitype Galton-Watson forests

Consider a (unitype) *branching forest* with k trees and progeny distribution ν on $\mathbb{Z}_+$, that is, each of the k individuals at generation 0 has offspring according to ν, then each of its children has offspring independently of the others and with the same law. Such forests are also called *GW forests*. A *multitype Galton-Watson* (MGW) forest in d-types, is a branching forest, where each individual has a type $i \in [d]$, and has children independently of the others, according to a law ν_i on $\mathbb{Z}_+^d$. The *progeny distribution* of the forest is $\nu = (\nu_1, \ldots, \nu_d)$. The formal definition is the following.

Definition 4.4. *A* multitype Galton-Watson process *is a Markov chain $Z = \left((Z_n^{(1)}, \ldots, Z_n^{(d)}); n \geq 0\right)$ on $\mathbb{Z}_+^d$, with transition function*

$$\mathbb{P}(Z_{n+1} = (k_1, \ldots, k_d) \,|Z_n = (r_1, \ldots, r_d)) = \nu_1^{*r_1} * \cdots * \nu_d^{*r_d}(k_1, \ldots, k_d),$$

where ν is the progeny distribution, and ν_i^{*j} is the jth iteration of the convolution product of ν_i by itself, with $\nu_i^{*0} = \delta_0$.

For $\mathbf{r} \in \mathbb{Z}_+^d$, the probability measure $\mathbb{P}_\mathbf{r}$ is the law $\mathbb{P}(\cdot | Z_0 = \mathbf{r})$. As in Theorem 1.2 in [CL16], we consider MGW trees satisfying the following. For $i, j \in [d]$, let $m_{i,j} = \sum_{z \in \mathbb{Z}_+^d} z_j \nu_i(z)$ be the mean number of children type j given by an individual type i, and set $M = (m_{i,j})_{i,j}$ as the mean matrix of the MGW tree. Whenever M is irreducible, by the Perron-Frobenius Theorem (see [AN04, Chapter V.2]), it has a unique eigenvalue which is simple, positive and with maximal modulus. We say in such case that the MGW tree is *irreducible*. If the unique eigenvalue equals one (is less than one), then we say the tree is *critical (subcritical)*. The tree is *non-degenerate* if individuals have exactly one offspring with probability different from one.

4.1.2 Statement of the results

Multitype forests with a given degree sequence

To define uniform d-type forests with a given degree sequence, having root-type $\mathbf{r} > 0$, we first define a multitype degree sequence. A *multitype degree sequence* $\mathbf{S} = (\mathbf{S}_{i,j}, i, j \in [d])$ is a sequence of sequences of non-negative integers $\mathbf{S}_{i,j} = (N_{i,j}(k); k \in [m_{i,j}]_0)$, where $m_{i,j} \in \mathbb{N}$, satisfying:

1. $n_i = \sum_k N_{i,j}(k)$ for every $i \in [d]$,

2. $n_j = r_j + \sum_k kN_{1,j}(k) + \cdots + \sum_k kN_{d,j}(k)$, for every $j \in [d]$,

3. $det(-k_{i,j}) > 0$ with $k_{i,j} := \sum kN_{i,j}(k) - n_i \mathbf{1}_{i=j}$ and $k_{i,i} < 0$ for every $i \in [d]$.

The value $N_{i,j}(k)$ represents the number of individuals of type i with k children of type j, so n_i represents the total number of individuals of type i. Thus, the total number of vertices is $s := n_1 + \cdots + n_d = \sum_k N_{1,j}(k) + \cdots + \sum_k N_{d,j}(k)$ for $j \in [d]$. The last condition is imposed to obtain a forest with such degree sequence (see page 109). For simplicity, we will assume that our multitype degree sequences satisfy the third condition and we focus on the first two conditions. Table 4.1 summarizes the case $d = 2$.

$\mathbf{S}_{1,1} = (N_{1,1}(0), \ldots, N_{1,1}(m_{1,1}))$	$\mathbf{S}_{1,2} = (N_{1,2}(0), \ldots, N_{1,2}(m_{1,2}))$	$n_1 = \sum_k N_{1,j}(k)$
$\mathbf{S}_{2,1} = (N_{2,1}(0), \ldots, N_{2,1}(m_{2,1}))$	$\mathbf{S}_{2,2} = (N_{2,2}(0), \ldots, N_{2,2}(m_{2,2}))$	$n_2 = \sum_k N_{2,j}(k)$
$n_1 = r_1 + \sum_k kN_{1,1}(k) + kN_{2,1}(k)$	$n_2 = r_2 + \sum_k kN_{1,2}(k) + kN_{2,2}(k)$	$n_1 + n_2 = s$

Table 4.1: Relations on the degree sequence of a 2-type forest.

As in the unitype case, we construct the canonical child sequence $\mathbf{c} = (\mathbf{c}_{i,j}, i, j \in [d])$ from the degree sequence, that is, let $\mathbf{c}_{i,j}$ be a sequence whose first $N_{i,j}(0)$ entries are zeros, the next $N_{i,j}(1)$ entries are ones, and so on. Let $\sigma_{i,j}$ be any permutation on $[n_i]$, and construct $w^b = \{w^b_{i,j}; i, j \in [d]\}$, where

$$w^b_{i,j}(k) = \sum_{l=1}^{k} \left(\mathbf{c}_{i,j} \circ \sigma_{i,j}(l) - \mathbf{1}_{i=j} \right), \quad k \in [n_i].$$

Remark 4.1. *Note that $k_{i,j} := w^b_{i,j}(n_i)$ does not depend on the permutation, so it is deterministic. Also, note that the system of equations $(\mathbf{r}, w^b)$ admits $\mathbf{n}$ as a solution, since by definition*

$$r_j + \sum_{i=1}^{d} w^b_{i,j}(n_i) = r_j - n_j + \sum_{i=1}^{d} \sum kN_{i,j}(k) = 0 \qquad \forall\, j \in [d].$$

Finally, note that $-k_{j,j} = r_j + \sum_{i \neq j} k_{i,j}$ since by definition, we have

$$-k_{j,j} = n_j - \sum_k kN_{j,j}(k) = r_j + \sum_i \sum_k kN_{i,j}(k) - \sum_k kN_{j,j}(k) = r_j + \sum_{i \neq j} k_{i,j}.$$

Trivially, whenever $r_j > 0$ we have $-k_{j,j} > 0$.

From the Multivariate Cyclic Lemma 4.1, we know that $det(k_{i,j})$ is the number of good cyclical permutations of w^b. From such set we define a Vervaat-type transformation of w^b. Such transformation is given by choosing uniformly at random a good-cyclical permutation from all the good-cyclical permutations. After that, the algorithm is similar to the unidimensional case.

Definition 4.5 (Multidimensional Vervaat Transform)**.** *For any w^b as constructed above and any $u \in [\det(k_{i,j})]$, define $V(w^b, u)$ as follows: enumerate the* $\det(k_{i,j})$ *good cyclical permutations of w^b, using the lexicographic order on the set of $\mathbf{q}$ such that $\theta_{\mathbf{q},\mathbf{n}}(w^b)$ codes a forest; then, let $V(w^b, u)$ be the u-th good cyclical permutation.*

Let $\mathbb{F}_{\mathbf{S},\mathbf{r}}$ be the set of multitype plane forests with degree sequence $\mathbf{S}$, having root-type $\mathbf{r}$ and individuals-type $\mathbf{n}$. Our construction of MFGDS is the following (the proof is given on page 109).

Theorem 4.1 (Uniform multitype forest with a given degree sequence)**.** *Fix the degree sequence* $\mathbf{S}$ *of a multitype forest having root-type* $\mathbf{r}$ *and individuals-type* $\mathbf{n}$*. Let* $\mathbf{W}$ *be the BFW coding a forest (see* (4.1)*) taken uniformly at random from* $\mathbb{F}_{\mathbf{S},\mathbf{r}}$*. Let* $\pi = (\pi_{i,j}, i, j \in [d])$ *be independent random permutations, where* $\pi_{i,j}$ *takes values on* $[n_i]$*, and let* U *be an independent uniform variable on* $[\det(k_{i,j})]$*. Define the processes* $\mathbf{W}^b = (W^b_{i,j}, i, j \in [d])$ *as*

$$W^b_{i,j}(k) = \sum_{l=1}^{k} (\mathbf{c}_{i,j} \circ \pi_{i,j}(l) - \mathbf{1}_{i=j}), \quad k \in [n_i],$$

where $\mathbf{c} = \big(c_{i,j}, i, j \in [d]\big)$ *is the child sequence of* $\mathbf{S}$*. Then*

$$V(\mathbf{W}^b, U) \stackrel{d}{=} \mathbf{W}.$$

From the proof, we obtain $|\mathbb{F}_{\mathbf{S},\mathbf{r}}|$, the number of multitype forests with a given degree sequence $\mathbf{S}$:

$$|\mathbb{F}_{\mathbf{S},\mathbf{r}}| = \frac{\det(-k_{i,j})}{\prod n_i} \prod\prod \binom{n_i}{\mathbf{S}_{i,j}}.$$

MGW forests conditioned by types

Before turning to the joint law of the number of individuals of type $i \in [d]$, of a MGW forest, we prove that the latter model is a mixture of MFGDS in Section 4.4. This justifies the importance of the latter model.

Let $S^{i,j}$ be a random walk with increments having law the jth marginal of ν_i. Our hypotheses are the following:

H1 For every $i \in [d]$, the law ν_i has independent components, with

$$\nu_i^{*n_i}(k_1, \dots, k_d) = \prod_j \mathbb{P}\big(S^{i,j}_{n_i} = k_j\big) \qquad k_1, \dots, k_d \in \mathbb{N}_0.$$

H2 For every $i, j \in [d]$, with $i \neq j$

$$\mathbb{E}\left(S^{i,j}_{n_i}; \sum_{l \in [d]} S^{l,j}_{n_l} = n_j - r_j\right) = \frac{n_i(n_j - r_j)}{n} \mathbb{P}\left(\sum_{l \in [d]} S^{l,j}_{n_l} = n_j - r_j\right).$$

It is important to remark that we do not assume that the components ν_i have the same distribution. Using those hypotheses, we obtain the following result (see page 112), which is a generalization of the Otter-Dwass formula [Ott49, Dwa69].

Theorem 4.2. *Consider an irreducible, non-degenerate and (sub)critical MGW forest, and let $n_i > 0$ for every i. Suppose that H1 and H2 and are also satisfied. If O_i is the number of type i individuals, then*

$$\mathbb{P}_{\mathbf{r}}\left(O_i = n_i, i \in [d]\right) = \frac{r}{n}\prod_{i=1}^{d}\mathbb{P}\left(\sum_{l\in[d]} S_{n_l}^{l,i} = n_i - r_i\right),$$

where $r = r_1 + \cdots + r_d$ and $n = n_1 + \cdots + n_d$, and $r_i < n_i$.

Remark 4.2. *Under the assumption $n_i > 0$ for every i, the proof is simpler, but we think this hypothesis can be dropped as in [CL16].*

Remark 4.3. *On page 114, we obtain the case when $n_i = r_i$ for some i's. Since Theorem 4.2 has a different formula on such case, the law of $\sum_{i\in[d]} O_i$ (computed on Corolary 4.1) does not have a nice expression.*

For the next results denote by $\mathbb{F}_{\mathbf{r},\mathbf{n}}^{plane}$, $\mathbb{F}_{\mathbf{r},\mathbf{n}}^{labeled}$ and $\mathbb{F}_{\mathbf{r},\mathbf{n}}^{binary}$, the set of d-type plane, labeled and binary forests having root-type $\mathbf{r}$ and individuals-type $\mathbf{n}$, where $r_i < n_i$ for every i and $r > 0$. Our labeled multitype forests have labels on $[n]$, that is, for $F \in \mathbb{F}_{\mathbf{r},\mathbf{n}}^{labeled}$, each individual v has a unique label $i \in [n]$ and a type $c_F(v) \in [d]$; also, F has fixed *root set* $[r]$, that is, the r_1 type 1 roots have labels on $\{1,\ldots,r_1\}$, the r_2 type 2 roots have labels on $\{r_1+1,\ldots,r_1+r_2\}$, and so on. Using Theorem 4.2, we give in Subsection 4.5.1 three examples of distributions were the law of a MGW forest conditioned by the number of individuals of each type can be computed explicitly. This generalizes the constructions given in [Pit98].

Proposition 4.1. *For fixed $p \in (0,1)$, let $\mathscr{G}_{\mathbf{r},p}$ be a d-type GW forest with root-type $\mathbf{r}$, having geometric offspring distribution with parameter p independently for each type, that is, $\nu_i(k_1,\ldots,k_d) = \prod_i p(1-p)^{k_i}$ for $k_i \geq 0$. Let $\#_i\mathscr{G}_{\mathbf{r},p}$ be the number of type i individuals in $\mathscr{G}_{\mathbf{r},p}$. Then*

$$\mathbb{P}(\mathscr{G}_{\mathbf{r},p} = F | \#_i\mathscr{G}_{\mathbf{r},p} = n_i, i \in [d]) = \frac{1}{\frac{r}{n}\prod_{i\in[d]}\binom{n+n_i-r_i-1}{n_i-r_i}} \qquad \forall F \in \mathbb{F}_{\mathbf{r},\mathbf{n}}^{plane},$$

thus, such conditioned forest is uniform on $\mathbb{F}_{\mathbf{r},\mathbf{n}}^{plane}$.

Proposition 4.2. *For $\mu \in \mathbb{R}^+$, let $\mathscr{P}_{\mathbf{r},\mu}$ be a d-type GW forest with root-type $\mathbf{r}$, having Poisson offspring distribution of parameter μ independently for each type, that is, $\nu_i(k_1,\ldots,k_d) = \prod_i e^{-\mu}\mu^{k_i}/k_i!$ for $k_i \geq 0$. Let $\#_i\mathscr{P}_{\mathbf{r},p}$ be the number of type i individuals in $\mathscr{P}_{\mathbf{r},p}$. If $\mathscr{P}_{\mathbf{r},\mathbf{n}}^*$ is $\mathscr{P}_{\mathbf{r},\mathbf{n}}$ relabeled by d uniform random permutations, one for each type, then*

$$\mathbb{P}\left(\mathscr{P}_{\mathbf{r},p}^* = F | \#_i\mathscr{P}_{\mathbf{r},p} = n_i, i \in [d]\right) = \frac{1}{\frac{r}{n}n^{n-r}} \qquad \forall F \in \mathbb{F}_{\mathbf{r},\mathbf{n}}^{labeled},$$

thus, such conditioned forest is uniform on $\mathbb{F}_{\mathbf{r},\mathbf{n}}^{labeled}$.

Proposition 4.3. *For $0 < p < 1$, let $\mathscr{B}_{\mathbf{r},p}$ be a d-type GW forest with root-type $\mathbf{r}$, having Bernoulli offspring distribution with parameter p, for each vertex independently of the type, that is, $\nu_i(k_1,\ldots,k_d) = \prod p^{k_i}(1-p)^{1-k_i}$ with $k_i \in \{0,1\}$. Assume that $n_i - r_i$ is an even number for every $i \in [d]$. Since any vertex v has zero or two children with probability p or $1-p$ respectively, then $\nu_i(c_1(v),\ldots,c_d(v)) = \prod p^{c_i(v)/2}(1-p)^{1-c_i(v)/2}$. Let $\#_i\mathscr{B}_{\mathbf{r},p}$ be the number of type i individuals in $\mathscr{B}_{\mathbf{r},p}$. Then*

$$\mathbb{P}(\mathscr{B}_{\mathbf{r},p} = F | \#_i\mathscr{B}_{\mathbf{r},p} = n_i, i \in [d]) = \frac{1}{\frac{r}{n}\prod\binom{n}{(n_i-r_i)/2}} \qquad \forall F \in \mathbb{F}_{\mathbf{r},\mathbf{n}}^{binary},$$

thus, such conditioned forest is uniform on $\mathbb{F}_{\mathbf{r},\mathbf{n}}^{binary}$.

As a simple consequence of our results, we obtain the following enumerations.

Lemma 4.2. *The number of d-type plane, labeled, and binary forest, with root-type* $\mathbf{r}$ *and individuals-type* $\mathbf{n}$ *is given respectively by*

$$\frac{r}{n}\prod_{i\in[d]}\binom{n+n_i-r_i-1}{n_i-r_i}, \qquad \frac{r}{n}n^{n-r} \quad and \quad \frac{r}{n}\prod\binom{n}{(n_i-r_i)/2}.$$

Finally, we give an algorithm to simulate MGW processes conditioned by its types. This is done using the following proposition and an Accept-Reject method (see Algorithm 8).

Proposition 4.4. *Let W be the breadth-first walk of a MCGW($n_1,\dots,n_d$) forest satisfying the Hypotheses of Theorem 4.2, having offspring distribution ν, and root-type* $\mathbf{r}$*, with $0<r_i<n_i$ for every i. Generate independent multinomial vectors $\mathbf{S}_{i,j}=(N_{i,j}(0),N_{i,j}(1),\dots)$ with parameters $(n_i,\nu_{i,j}(0),\nu_{i,j}(1),\dots)$, and stop the first time that $r_j+\sum_i\sum_k kN_{i,j}(k)=n_j$ for every j. Denote by* $\mathbf{S}$ *the multitype degree sequence obtained, and let $V(\mathbf{W}^b,U)$ be the breadth-first walk generated by Theorem 4.1 using the degree sequence* $\mathbf{S}$*. Then,*

$$\mathbb{P}\Big(V(\mathbf{W}^b,U)=w\Big)=\frac{1}{\frac{n}{r}\frac{det(k_{i,j})}{\prod n_i}}\mathbb{P}_{\mathbf{r}}\left(\mathscr{F}=F|\#_j\mathscr{F}=n_j,\forall\, j\right),$$

for every multitype forest F with root-type $\mathbf{r}$ *and individuals-type* $\mathbf{n}$*, coded by w and with $k_{i,j}=\sum kn_{i,j}(k)-n_i\mathbf{1}_{i=j}$.*

The chapter is organized as follows: in Section 4.2 we construct uniform forests with a given degree sequence. This will be used in Section 4.3, to construct MFGDS and prove Theorem 4.1. In Section 4.4 we prove that under an independence assumption, the cMGW forests are mixtures of MFGDS. Section 4.5 is devoted to prove the joint law of the number of individuals by types in a MGW forest, which is Theorem 4.2. In that section we also obtain in Corollary 4.1, the law of the total population in a MGW forest. Examples satisfying the hypotheses of Theorem 4.2 are given in Subsection 4.5.1. The algorithms are given in Section 4.6.

4.2 Construction of unitype random forests with a given degree sequence

A well-known encoding of forests by skip-free random walks is given as follows. Define the set of all bridges finishing at $-m$ at time s, as

$$\mathbb{B}^{s,m}=\left\{(y(1),\dots,y(s))\in\mathbb{Z}^s : y(j)-y(j-1)\geq -1 \text{ for } j\in[s], y(s)=-m\right\}.$$

For $i\in[s]-1$ and $y\in\mathbb{B}^{s,m}$, define $\theta_i(y)$ as the cyclic permutation of y at i, that is

$$(\theta_i(y))(j)=\begin{cases} y(j+i)-y(i) & j+i\leq s\\ y(j+i-s)+y(s)-y(i) & s-i\leq j\leq s.\end{cases}$$

This transformation puts the last $s-i$ increments of y as the first $s-i$ increments of $\theta_i(y)$, and the first i increments of y as the last i increments of $\theta_i(y)$.

For any $u \in [m]-1$ and $y \in \mathbb{B}^{s,m}$, let τ_u be the time that y hits $min(y)+u$ for the first time. The Vervaat-type transformation V that we use is given by

$$V(y,u) = \theta_{\tau_u}(y).$$

Note that this transformation leads the set of bridges, to the set of excursions of size s finishing at $-m$

$$\mathbb{E}^{s,m} = \{(w(1),\dots,w(s)) \in \mathbb{Z}^s : w(j)-w(j-1) \in \{-1,0,1,\dots\} \text{ for } j \in [s], w \text{ first reaches } -m \text{ at time } s\}.$$

Now, let F be a forest with trees $T_1,\dots,T_m$, for $m \in \mathbb{N}$. The degree sequence of F is given by

$$N_i(F) = \sum_1^m N_i(T_j).$$

Note that, any finite sequence of non-negative integers $\mathbf{S} = (N_i, i \geq 0)$, such that for some $m \in [|\mathbf{S}|]$, we have

$$s := \sum N_i = \sum iN_i + m,$$

is the degree sequence of some forest with m trees. In this case we call $\mathbf{S}$ a degree sequence. The size of the forest F will also be denoted by $|F|$.

Fix any degree sequence $\mathbf{S}$ and obtain its child sequence $\mathbf{c} := \mathbf{c}(\mathbf{S})$. As before, we obtain an EI process using uniform permutations of the child sequence $\mathbf{c}$. Let σ be any permutation on $[s]$. Define the *bridge*

$$w^b(j) = \sum_1^j (c\circ\sigma(i)-1) \qquad j \in [s],$$

with $w^b(0) = 0$. Note that $w^b(s) = -m$. The set of paths taken by w^b is

$$\mathbb{B}_{\mathbf{S}} = \{(y(1),\dots,y(s)) \in \mathbb{B}^{s,m} : |j : y(j)-y(j-1) = i-1| = N_i \text{ for every } i \geq 0\}.$$

From the excursions in $\mathbb{E}^{s,m}$, we consider those with fixed number of increments of given size:

$$\mathbb{E}_{\mathbf{S}} = \{(w(1),\dots,w(s)) \in \mathbb{E}^{s,m} : |j : w(j)-w(j-1) = i-1| = N_i, i \geq 0\}.$$

Define $\mathbb{F}_{\mathbf{S}}$ as the set of all forests with degree sequence $\mathbf{S}$. Using Lemma 6.3 of [Pit06], it can be proved that there exists a bijection between $\mathbb{E}_{\mathbf{S}}$ and $\mathbb{F}_{\mathbf{S}}$, and we know that $|\mathbb{F}_{\mathbf{S}}| = \frac{m}{s}\binom{s}{N_i, i\geq 0}$.

It is clear from a picture, that a bridge is sent to an excursion by the Vervaat transformation. Let us prove this is also the case for bridges in $\mathbb{B}_{\mathbf{S}}$, that is, bridges coming from a degree sequence $\mathbf{S}$. The next three lemmas are inspired in [Lei17].

Lemma 4.3. *For any $u \in [m]-1$ and $y \in \mathbb{B}_{\mathbf{S}}$, the path $V(y,u)$ belongs to $\mathbb{E}_{\mathbf{S}}$.*

Proof. By definition of τ_u, the minimum value that can take $y(\tau_u+j)-y(\tau_u)$ for $j+\tau_u \leq s$ is $-u > -m$. These are the first $s-\tau_u$ times of $V(y,u)$. On the remaining times $\{s-\tau_u,\dots,s\}$, the minimum of $V(y,u)$ is attained for the first time at time s. This implies $V(y,u)(j) > -m$ for $j < s$, and hence $V(y,u) \in \mathbb{E}$. Since the Vervaat transformation only permutes the increments, it is clear that if $y \in \mathbb{B}_{\mathbf{S}}$ then $V(y,u) \in \mathbb{E}_{\mathbf{S}}$. □

Lemma 4.4. *Let $w \in \mathbb{E}_{\mathbf{S}}$. Then, the number of different pairs $(y,u) \in \mathbb{B}_{\mathbf{S}} \times ([m]-1)$ such that $V(y,u) = w$ is exactly s.*

Proof. Consider any $i \in [s]-1$ and the cyclical permutation $y' = \theta_i(w)$. It is clear that $\theta_{s-i}(y') = w$. In fact $V(y',u) = w$ for some $u \in [m]-1$. This holds true since the last $s-i$ increments of w are the first $s-i$ increments of y', then y' hits $y'(s-i) = w(s) - w(s-i)$ for the first time at time $s-i$. Hence, the Vervaat transform can be applied at $u = y'(s-i)$, giving us the path w.

Note that the path of w can be partitioned in m subexcursions, each one of the form $(w(j+T_i), j \in [T_{i+1} - T_i])$, with $i \in [m]-1$ and T_i the first hitting time to $-i$. First assume that w can be partitioned in $k_w \in [m]$ *identical* subexcursions, each of length l_w. It follows $k_w l_w = s$. This is equivalent to say that there exists k_w values u such that $V(w,u) = w$. Those values are $w(jl_w) + m \in [m]-1$ for $j \in [k_w]$. In this case, there are only l_w different cyclic permutations $\theta_i(w)$ of w, each one having k_w distinct values of u such that $V(\theta_i(w),u) = w$. This proves that w has exactly s preimages. If w cannot be partitioned in identical subexcursions, for every cyclical permutation $\theta_i(w)$ there is only one u such that $V(\theta_i(w),u) = w$. This concludes the proof. □

Now, we construct a uniform forest on $\mathbb{F}_{\mathbf{S}}$.

Lemma 4.5. *Consider a degree sequence $\mathbf{S}$ of a forest having m trees and s individuals, and let $\mathscr{F}$ be a forest taken uniformly at random from $\mathbb{F}_{\mathbf{S}}$. Let π be a uniform random permutation on $[s]$, U an independent uniform variable on $[m]-1$, and define the bridge*

$$W^b(j) = \sum_{1}^{j} (c \circ \pi(j) - 1) \qquad j \in [s],$$

where $\mathbf{c}$ is the child sequence of $\mathbf{S}$. Then

$$V(W^b, U) \stackrel{d}{=} \mathscr{F}.$$

Proof. Fix any $w \in \mathscr{E}_{\mathbf{S}}$ and any of its cyclical permutations $w^b \in \mathscr{B}_{\mathbf{S}}$. From the $s!$ possible values that π can take, only $\prod N_i!$ give the same bridge w^b. This is true since we can permute the labels of the N_i individuals having i children and obtain the same bridge. Hence

$$\mathbb{P}\Big(W^b = w^b\Big) = \frac{1}{\binom{s}{N_i, i\geq 0}},$$

which does not depend on w^b.

By the previous lemma, there are s distinct pairs (w^b, u) that are mapped to w. Denote such pairs as $(w^b_k, u_k) \in \mathscr{B}_{\mathbf{S}}$, for $k \in [s]$. Then, using the independence of W^b and U, and that the former is uniform, then

$$\mathbb{P}\Big(V(W^b,U) = w\Big) = \sum_{k=1}^{s} \mathbb{P}\Big(W^b = w^b_k, U = u_k\Big) = \frac{s}{m}\mathbb{P}\Big(W^b = w^b\Big) = \frac{1}{|\mathbb{F}_{\mathbf{S}}|}.$$ □

4.3 Construction of multitype random forests with a given degree sequence

Recall Definition 4.5 of the multidimensional Vervaat transform $V(w^b,u)$, using the bridge w^b as constructed from the degree sequence. The set of values of $V(w^b,u)$ will be denoted by $\mathbb{E}_{\mathbf{S},\mathbf{r}}$. Now we are ready to construct a forest taken uniformly at random from $\mathbb{F}_{\mathbf{S},\mathbf{r}}$, the set of plane forests with degree sequence $\mathbf{S}$ having root-type $\mathbf{r}$.

Proof of Theorem 4.1. First we prove that from any multitype degree sequence, we can construct a multitype forest. From Remark 4.1, since we can associate to the degree sequence a system of equations $(\mathbf{r},w^b)$ with solution $\mathbf{n}$. To such system we can associate a multitype forest using the Multivariate Cyclic Lemma 4.1, since any good cyclical permutation codes a forest.

Now, define $\mathbf{S}_{i,j}=(N_{i,j}(k);k\geq 0)$, and write

$$\binom{n_i}{\mathbf{S}_{i,j}}:=\binom{n_i}{N_{i,j}(0),N_{i,j}(1),\ldots}.$$

Fix any bridge $w^b\in\mathbb{B}_{\mathbf{S},\mathbf{r}}$. From the possible $\prod_i (n_i!)^d$ values taken by the random permutations $(\pi_{i,j},i,j\in[d])$, exactly $\prod_j\prod_i\prod_k N_{i,j}(k)!$ form the bridge w^b. This is true since, permuting the labels of the $N_{i,j}(k)$ individuals type i having k children type j, we obtain the same bridge. This proves the assertion since this is true for every i,j,k. Therefore

$$\mathbb{P}\Big(\mathbf{W}^b=w^b\Big)=\frac{1}{\prod\prod\binom{n_i}{\mathbf{S}_{i,j}}}.$$

Now, fix any $w\in\mathbb{E}_{\mathbf{S},\mathbf{r}}$ and $i\in[d]$. We obtain the number of different pairs (w^b,u) that can be mapped to w using the multidimensional Vervaat transform. The point is that such bridges can only be of the form $\theta_{\mathbf{q},\mathbf{n}}(w)$, that is, cyclical permutations of w. Hence, each component $w^{(i)}$ comes from a Vervaat transform $V(\theta_j(w^{(i)}),u)$ for some j,u. By Lemma 4.4, the number of pairs $(\theta_j(w^{(i)}),u)$ that can be mapped to $w^{(i)}$ are exactly n_i. This being true for every i implies there are $\prod n_i$ unique pairs $(\theta_{\mathbf{q},\mathbf{n}}(w),u)$ such that $V((\theta_{\mathbf{q},\mathbf{n}}(w),u))=w$. Denote such pairs as

$$A(w)=\Big\{(w^b_k,u_k)\in\mathbb{B}_{\mathbf{S},\mathbf{r}}\times[\det(k_{i,j})]:V((w^b_k,u_k))=w,k\in\big[\prod n_i\big]\Big\}.$$

This implies

$$\begin{aligned}\mathbb{P}\Big(V(\mathbf{W}^b,U)=w\Big)&=\mathbb{P}\Big((\mathbf{W}^b,U)\in A(w)\Big)\\&=\sum_{k\in[\prod n_i]}\mathbb{P}\Big((\mathbf{W}^b,U)=(w^b_k,u_k)\Big)\\&=\sum_{k\in[\prod n_i]}\frac{1}{\prod\prod\binom{n_i}{\mathbf{S}_{i,j}}}\frac{1}{\det(-k_{i,j})}\\&=\frac{1}{\frac{\det(-k_{i,j})}{\prod n_i}\prod\prod\binom{n_i}{\mathbf{S}_{i,j}}}.\end{aligned}$$

This concludes the proof since the right-hand side is independent of w, so $V(\mathbf{W}^b,U)$ is uniform. $\square$

Remark 4.4. *From this lemma we can conclude that the set of plane forests with degree sequence* $\mathbf{S}$ *having root-type* $\mathbf{r}$ *is*

$$|\mathbb{F}_{\mathbf{S},\mathbf{r}}| = \frac{\det(-k_{i,j})}{\prod n_i} \prod\prod \binom{n_i}{\mathbf{S}_{i,j}}.$$

4.4 Relation between MFGDS and cMGW forests

Before turning to the results about conditioned MGW forests, let us prove that under an independence condition, a MGW conditioned by its degree sequence has the same law as a MFGDS. For any given $\mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}_+^d$ and $\mathbf{r} = (r_1, \dots, r_d) \in \mathbb{Z}_+^d$ with $\sum r_i > 0$, define the set of all degree sequences having n_i individuals of type i, as

$$DS(\mathbf{n},\mathbf{r}) = \Big\{ \mathbf{S} = (N_{i,j}(k), i, j \in [d], k \geq 0) : \\ n_i = \sum_k N_{i,j}(k),\ n_j = r_j + \sum_i \sum_k kN_{i,j}(k), N_{i,j}(k) \geq 0 \text{ for } i, j \in [d] \Big\}.$$

Also, for any given multitype forest F, define its empirical multitype degree sequence $\hat{N}(F) := \hat{N} = (\hat{N}_{i,j}(k), i, j \in [d], k \geq 0)$ as

$$\hat{N}_{i,j}(k) = \sum_{l: v_l \in F^{(i)}} \mathbf{1}_{\mathbf{c}_{i,j}(l)=k},$$

where $\mathbf{c}_{i,j}(l)$ is the number of children type j, that the lth individual of the subforest $F^{(i)}$ of vertices type i has.

Lemma 4.6. *Fix any* $\mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}_+^d$ *and* $\mathbf{r} = (r_1, \dots, r_d) \in \mathbb{Z}_+^d$ *with* $\sum r_i > 0$. *Consider a multitype degree sequence* $\mathbf{S} = (N_{i,j}(k), i, j \in [d], k \geq 0) \in DS(\mathbf{n},\mathbf{r})$. *Consider a MGW forest with progeny distribution* $\nu = (\nu_1, \dots, \nu_d)$ *such that each* ν_i *has independent components. Then, the law of a MGW forest conditioned to have multitype degree sequence* $\mathbf{S}$ *is the same as* $\mathbb{P}_{\mathbf{S}}$*, the law of a MFGDS. Therefore, the law of a cMGW forest is a finite mixture of the laws* $(\mathbb{P}_{\mathbf{S}}, \mathbf{S} \in DS(\mathbf{n}))$.

Proof. Let $\mathscr{F}$ be a MGW tree. The assumption on independence can be written as $\nu_i(\mathbf{z}) = \prod_j \nu_{i,j}(z_j)$ for any $z = (z_1, \dots, z_d) \in \mathbb{Z}_+^d$ and any j. Let $(F_1, c(F_1))$ and $(F_2, c(F_2))$ be two multitype forests having degree sequence $\mathbf{S} \in DS(\mathbf{n},\mathbf{r})$. Then

$$\begin{aligned}
\mathbb{P}_{\mathbf{r}}\left(\mathscr{F} = F_1, \hat{N}(\mathscr{F}) = \mathbf{S}\right) &= \prod_i \prod_{k: v_k \in F_1^{(i)}} \prod_j \nu_{i,j}(c_{i,j}(k)) \\
&= \prod_i \prod_j \prod_k \nu_{i,j}(k)^{N_{i,j}(k)} \\
&= \prod_i \prod_{k: v_k \in F_2^{(i)}} \prod_j \nu_{i,j}(c_{i,j}(k)) \\
&= \mathbb{P}_{\mathbf{r}}\left(\mathscr{F} = F_2, \hat{N}(\mathscr{F}) = \mathbf{S}\right).
\end{aligned}$$

This implies the first assertion. Let $\mathbf{O} = (O_1, \dots, O_d)$ be the vector with the total number of individuals of each type in a MGW forest. To prove the second assertion, we sum over all the values in $DS(\mathbf{n},\mathbf{r})$, obtaining

$$\begin{aligned}
\mathbb{P}_{\mathbf{r}}\left(\mathscr{F}\in\cdot|\mathbf{O}=\mathbf{n}\right) &= \frac{1}{\mathbb{P}_{\mathbf{r}}\left(\mathbf{O}=\mathbf{n}\right)}\sum_{\mathbf{S}\in DS(\mathbf{n},\mathbf{r})}\mathbb{P}_{\mathbf{r}}\left(\mathscr{F}\in\cdot,\hat{N}(\mathscr{F})=\mathbf{S}\right)\\
&= \sum_{\mathbf{S}\in DS(\mathbf{n},\mathbf{r})}\frac{\mathbb{P}_{\mathbf{r}}\left(\hat{N}(\mathscr{F})=\mathbf{S},\mathbf{O}=\mathbf{n}\right)}{\mathbb{P}_{\mathbf{r}}\left(\mathbf{O}=\mathbf{n}\right)}\mathbb{P}_{\mathbf{r}}\left(\mathscr{F}\in\cdot|\hat{N}(\mathscr{F})=\mathbf{S}\right)\\
&= \sum_{\mathbf{S}\in DS(\mathbf{n},\mathbf{r})}\lambda_{\nu}^{\mathbf{S}}\mathbb{P}_{\mathbf{S}}\left(\mathscr{F}\in\cdot\right),
\end{aligned}$$

where

$$\lambda_{\nu}^{\mathbf{S}}=\mathbb{P}_{\mathbf{r}}\left(\hat{N}(\mathscr{F})=\mathbf{S}|\mathbf{O}=\mathbf{n}\right).$$

Note that trivially $\sum_{\mathbf{S}\in DS(\mathbf{n},\mathbf{r})}\lambda_{\nu}^{\mathbf{S}}=1$. □

4.5 Law of the number of individuals by types of a MGW forest

The main result in [CL16] is the following.

Theorem 4.3 (Theorem 1.2, [CL16])**.** *Let Z be a d-type branching process, which is irreducible, non-degenerate and (sub)critical. For $i\in[d]$, let O_i be the total number of individuals of type i, up to the extinction time T, and for $i\neq j$, let $A_{i,j}$ be the total number of individuals of type j whose parent is of type i, up to time T. Then, for all integers r_i, n_i, $k_{i,j}$ such that $r_i\geq 0$ with $r>0$, $k_{i,j}\geq 0$ for $i\neq j$, $-k_{j,j}=r_j+\sum_{i\neq j}k_{i,j}$, and $n_i\geq -k_{i,i}$, we have*

$$\begin{aligned}
&\mathbb{P}_{\mathbf{r}}\left(O_i=n_i,i\in[d],A_{ij}=k_{ij},i,j\in[d],i\neq j\right)\\
&=\frac{det(-k_{ij})}{\bar{n}_1\cdots\bar{n}_d}\prod_1^d\nu_i^{*n_i}\left(k_{i1},\ldots,k_{i(i-1)},n_i+k_{ii},k_{i(i+1)},\ldots,k_{id}\right),
\end{aligned}\tag{4.3}$$

where $\mathbf{r}=(r_1,\ldots,r_d)$, $\bar{n}_i=n_i\vee 1$ and $(-k_{ij})_{i,j\in[d]}$ is the matrix to which we remove row i and column i, for every i such that $n_i=0$.

For simplicity, we use $n_i>0$ for $i\in[d]$ in the following. Let us give a hint on how to derive the law of the population by types for a 2-type GW, having $r_i<n_i$ type i roots, for every i. Recall the hypothesis **H1** about the independence in the components of ν_i of Theorem 4.2. From Theorem 4.3, summing over all the possible values of $(k_{i,j})$ we have

$$\begin{aligned}
&\mathbb{P}_{\mathbf{r}}\left(O_1=n_1,O_2=n_2\right)\\
&=\sum_{i=0}^{n_1-r_1}\sum_{j=0}^{n_2-r_2}\frac{r_1r_2+r_1j+r_2i}{n_1n_2}\mathbb{P}\left((S_{n_1}^{1,1},S_{n_1}^{1,2})=(n_1-r_1-i,j)\right)\mathbb{P}\left((S_{n_2}^{2,1},S_{n_2}^{2,2})=(i,n_2-r_2-j)\right)\\
&=\sum_{i=0}^{n_1-r_1}\sum_{j=0}^{n_2-r_2}\frac{r_1r_2+r_1j+r_2i}{n_1n_2}\mathbb{P}\left(S_{n_1}^{1,1}=n_1-r_1-i\right)\mathbb{P}\left(S_{n_1}^{1,2}=j\right)\mathbb{P}\left(S_{n_2}^{2,1}=i\right)\mathbb{P}\left(S_{n_2}^{2,2}=n_2-r_2-j\right).
\end{aligned}\tag{4.4}$$

We perform each summation in *columns*, obtaining three terms of the form

$$\sum_{i=0}^{n_1-r_1}k_{i_1,1}\mathbb{P}\left(S_{n_1}^{1,1}=n_1-r_1-i\right)\mathbb{P}\left(S_{n_2}^{2,1}=i\right)\sum_{j=0}^{n_2-r_2}k_{i_2,2}\mathbb{P}\left(S_{n_1}^{1,2}=j\right)\mathbb{P}\left(S_{n_2}^{2,2}=n_2-r_2-j\right),$$

where $k_{i_1,1}$ can be r_1 or i, and $k_{i_2,2}$ can be r_2 or j. Hence, in order to perform the summation for any dimension, we need to expand the determinant $det(-k_{i,j})$, perform the summation in columns, and divide in cases: either a constant or a variable multiply the above probabilities. Note that, in the first case the summation is only a convolution. Is precisely the second case why we need Hypotheses **H2**. First, we describe explicitly $det(-k_{ij})$.

Definition 4.6. *An* elementary forest *is a forest of $\mathscr{F}_d$ that contains exactly one vertex of each type. In particular, each elementary forest contains exactly d vertices and is coded by the d couples (i_j, j) for $j \in [d]$, where i_j is the type of the parent of vertex type j. If the vertex of type j is a root, then we set $i_j = 0$. We define the set D of vectors $(i_1, \dots, i_d)$, with $0 \le i_j \le d$ such that (i_j, j), $i \in [d]$ codes an elementary forest.*

Recall Definition 4.2 of S_d of coding sequences of multitype forests.

Definition 4.7. *For any $\mathbf{r} = (r_1 \dots, r_d) \in \mathbb{Z}_+^d$ with $\mathbf{r} > 0$, let $S_d^{\mathbf{r}}$ be the subset of S_d of sequences x whose length belongs to $\mathbb{N}^d$ and corresponds to the smallest solution of the system $(\mathbf{r}, x)$.*

Joining Lemmas 4.4, 4.5, 4.6 and 4.7 in [CL16], we obtain the following easy consequence, which is a precise description of the number of good cyclical permutations of $x \in S_d^{\mathbf{r}}$.

Lemma 4.7. *Let $x \in S_d^{\mathbf{r}}$, where $\mathbf{r} > 0$. Assume that $\mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}_+^d$ is a solution of the system $(\mathbf{r}, x)$. Then, the number of good cyclical permutations of x is*

$$det(-k_{i,j}) = \sum_{(i_1,\dots,i_d)\in D} \prod_{j=1}^{d} k_{i_j,j},$$

where we set $k_{0,j} = r_j$, and $-k_{j,j} = r_j + \sum_{i\neq j} k_{i,j}$, $j \in [d]$, and $k_{i,j} = x^{i,j}(n_i)$.

Now, we prove our theorem.

Proof of Theorem 4.2. By the independence imposed on ν_i, the product in Equation (4.3) can be expressed as follows:

$$\begin{aligned} P(K) &:= \prod_{1}^{d} \nu_i^{*n_i}\left(k_{i1}, \dots, k_{i(i-1)}, n_i + k_{ii}, k_{i(i+1)}, \dots, k_{id}\right) \\ &= \prod_{j=1}^{d} \mathbb{P}\left(S_{n_j}^{j,j} = n_j - r_j - \sum_{i\neq j} k_{i,j}\right) \prod_{i\neq j} \mathbb{P}\left(S_{n_i}^{i,j} = k_{i,j}\right). \end{aligned}$$

We define the index set

$$A(\mathbf{n}, \mathbf{r}) = \{K = (k_{ij}) : k_{ij} \ge 0 \text{ for } i \neq j, 0 \le -k_{jj} \le n_j, -k_{jj} = r_j + \sum_{i\neq j} k_{ij}, \forall j \in [d]\},$$

and use the notation $\sum_{K\in A(\mathbf{n},\mathbf{r})}$ to denote the summation over all k_{ij} with $i, j \in [d]$ and $i \neq j$, such that $K = (k_{ij}) \in A(\mathbf{n}, \mathbf{r})$. Also, fix $j \in [d]$ and define the index set

$$A_j(\mathbf{n}, \mathbf{r}) = \{K^j = (k_{1j}, \dots, k_{dj}) : k_{lj} \ge 0 \text{ for } l \neq j, 0 \le \sum_{l\neq j} k_{lj} \le n_j - r_j\}.$$

Use the notation $\sum_{K^j \in A_j(\mathbf{n},\mathbf{r})}$ to denote the summation over all k_{lj} with $l \in [d]$ and $l \neq j$, such that $K^j = (k_{1j}, \ldots, k_{dj}) \in A_j(\mathbf{n},\mathbf{r})$. Then, we have

$$\mathbb{P}_{\mathbf{r}}(O_i = n_i, i \in [d]) = \frac{1}{\prod n_i} \sum_{K \in A(\mathbf{n},\mathbf{r})} \sum_{(i_1,\ldots,i_d) \in D} \prod_{j=1}^{d} k_{i_j,j} P(K).$$

Denote by m the number of summands in $det(-K)$ for any $K = (-k_{ij}) \in A(\mathbf{n},\mathbf{r})$. This is the number of elementary forests, hence, it does not depend on K. Note that there exists m functions $\sigma_l : [d] \mapsto [d]_0$ with $\sigma_l(j) \neq j$ for every l, j, such that

$$det(-k_{ij}) = \sum_{l=1}^{m} \prod_{j=1}^{d} k_{\sigma_l(j),j}.$$

This means that for $K' = (-k'_{ij}) \in A(\mathbf{n},\mathbf{r})$, we have

$$det(-k'_{ij}) = \sum_{l=1}^{m} \prod_{j=1}^{d} k'_{\sigma_l(j),j}.$$

Now, fix any of the m permutations σ_l. We analyze the summation $\sum_{K \in A(\mathbf{n},\mathbf{r})} \prod_{j=1}^{d} k_{\sigma_l(j),j} P(K)$. The determinant $det(-k_{i,j})$ is the sum of m terms, each one having one and only one term $k_{\sigma_l(j),j}$ of each column j. Hence, we can join such term with the corresponding probabilities (in the jth column) involving the walks $S^{i,j}_{n_i}$ for $i \in [d]$. This means that, for every $j \in [d]$ we can join the terms $K^j \in A_j(\mathbf{n},\mathbf{r})$ as follows

$$\begin{aligned} &\sum_{K \in A(\mathbf{n},\mathbf{r})} \prod_{j=1}^{d} k_{\sigma_l(j),j} P(K) \\ &= \prod_{j=1}^{d} \sum_{K^j \in A_j(\mathbf{n},\mathbf{r})} k_{\sigma_l(j),j} \mathbb{P}\left(S^{j,j}_{n_j} = n_j - r_j - \sum_{l \neq j} k_{l,j} \right) \prod_{l \neq j} \mathbb{P}\left(S^{l,j}_{n_l} = k_{l,j} \right). \end{aligned} \tag{4.5}$$

Note that whenever $\sigma_l(j) \neq 0$ we have

$$\sum_{K^j \in A_j(\mathbf{n},\mathbf{r})} k_{\sigma_l(j),j} \mathbb{P}\left(S^{j,j}_{n_j} = n_j - r_j - \sum_{l \neq j} k_{l,j} \right) \prod_{l \neq j} \mathbb{P}\left(S^{l,j}_{n_l} = k_{l,j} \right) = \mathbb{E}\left(S^{\sigma_l(j),j}_{n_{\sigma_l(j)}} ; \sum_{l \in [d]} S^{l,j}_{n_l} = n_j - r_j \right),$$

which is related with Hypotheses **H2**. Thus, we define

$$\tilde{k}_{\sigma_l(j),j} = \begin{cases} r_j & \sigma_l(j) = 0 \\ n_{\sigma_l(j)}(n_j - r_j)/n & \sigma_l(j) \neq 0. \end{cases}$$

The idea is that in Equation (4.5), when performing the summation with $k_{\sigma_l(j),j} \neq r_j$ we use the Hypotheses **H2**, whereas when $k_{\sigma_l(j),j} = r_j$ we simply use the convolution formula $\sum_{l \in [d]} S^{l,j}_{n_l}$. This implies

$$\sum_{K \in A(\mathbf{n},\mathbf{r})} \prod_{j=1}^{d} k_{\sigma_l(j),j} P(K) = \prod_{j=1}^{d} \tilde{k}_{\sigma_l(j),j} \mathbb{P}\left(\sum_{l \in [d]} S^{l,j}_{n_l} = n_j - r_j \right),$$

therefore

$$\mathbb{P}_{\mathbf{r}}\left(O_i = n_i, i \in [d]\right) \prod n_i = \prod_{j=1}^{d} \mathbb{P}\left(\sum_{l \in [d]} S_{n_l}^{l,j} = n_j - r_j\right) \sum_{(i_1,\ldots,i_d) \in D} \prod_{j=1}^{d} \tilde{k}_{i_j,j}$$

Define the matrix $\bar{K}$ as a $d \times d$ matrix with entries $\bar{k}_{ij} = n\tilde{k}_{ij} = n_i(n_j - r_j)$ for $i \neq j$, and diagonal

$$-\bar{k}_{jj} = nr_j + \sum_{i \neq j} n\tilde{k}_{ij} = nr_j + \sum_{i \neq j} n_i(n_j - r_j) = nr_j + (n_j - r_j)(n - n_j) = n_j(n - n_j + r_j).$$

Then, using Lemma 4.5 of [CL16], which computes a determinant for integer valued matrices $(k_{i,j})$ satisfying our conditions, we have

$$\sum_{(i_1,\ldots,i_d) \in D} \prod_{j=1}^{d} \tilde{k}_{i_j,j} = n^{-d} det(-\bar{K}).$$

To prove that $det(-\bar{K}) = n^d \frac{r}{n} \prod n_i$, factorize in row i the factor n_i, obtaining

$$det(-\bar{K}) = \prod n_i \begin{vmatrix} n - n_1 + r_1 & -(n_2 - r_2) & \cdots & -(n_d - r_d) \\ \vdots & \vdots & \ddots & \vdots \\ -(n_1 - r_1) & -(n_2 - r_2) & \cdots & n - n_d + r_d \end{vmatrix}.$$

Multiply the last row by minus one, and add it to every other row, to obtain

$$det(-\bar{K}) = \prod n_i \begin{vmatrix} n & 0 & \cdots & -n \\ 0 & n & \cdots & -n \\ \vdots & \vdots & \ddots & \vdots \\ -(n_1 - r_1) & -(n_2 - r_2) & \cdots & n - n_d + r_d \end{vmatrix}.$$

Multiply by $(n_i - r_i)/n$ each row $i \in [d-1]$, and add it to the last row

$$det(-\bar{K}) = \prod n_i \begin{vmatrix} n & 0 & \cdots & -n \\ 0 & n & \cdots & -n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n - \sum_1^d (n_i - r_i) \end{vmatrix},$$

and, being a diagonal matrix, it follows that $det(-\bar{K}) = n^{d-1} r \prod n_i = n^d \frac{r}{n} \prod n_i$ as wanted. □

Now we treat the case $r_i = n_i$ for some i. For simplicity assume $r_i < n_i$ for $i \in [d-1]$ and $r_d = n_d$. This implies neither there are children of type d nor type d individuals have children. Thus, from theorem 4.3 we have

$$\begin{aligned} &\mathbb{P}_{\mathbf{r}}\left(O_i = n_i, i \in [d], A_{ij} = k_{ij}, i,j \in [d-1], i \neq j, A_{ld} = 0, A_{dl} = 0, l \neq d\right) \\ &= \frac{det(-k_{ij})}{n_1 \cdots n_d} \nu_d^{*n_d}(0,0,\ldots,0) \prod_1^{d-1} \nu_i^{*n_i}\left(k_{i1},\ldots,k_{i(i-1)}, n_i + k_{ii}, k_{i(i+1)},\ldots,k_{i(d-1)}, 0\right). \end{aligned}$$

Since the matrix $(-k_{ij})$ has zeros in column and row d, except at $-k_{dd} = r_d = n_d$, we have $det(-k_{ij}) = n_d det(-k_{ij})_{i,j\in[d-1]}$. Using the independence of Hypotheses **H1**, we reduce the problem to the joint law of the first $d-1$ components:

$$\begin{aligned}&\mathbb{P}_{\mathbf{r}}\left(O_i = n_i, i\in[d], A_{ij} = k_{ij}, i,j\in[d-1], i\neq j, A_{ld} = 0, A_{dl} = 0, l\neq d\right)\\&= \mathbb{P}_{(r_1,\dots,r_{d-1})}\left(O_i = n_i, i\in[d-1], A_{ij} = k_{ij}, i,j\in[d-1]\right)\prod_j \mathbb{P}\left(S^{d,j}_{n_d} = 0\right)\prod_{j\neq d}\mathbb{P}\left(S^{j,d}_{n_j} = 0\right).\end{aligned}$$

From this and the proof of Theorem 4.2, we obtain

$$\begin{aligned}\mathbb{P}_{\mathbf{r}}\left(O_i = n_i, i\in[d]\right) &= \prod_j \mathbb{P}\left(S^{d,j}_{n_d} = 0\right)\prod_{j\neq d}\mathbb{P}\left(S^{j,d}_{n_j} = 0\right)\frac{\sum_1^{d-1} r_i}{\sum_1^{d-1} n_i}\prod_{i=1}^{d-1}\mathbb{P}\left(\sum_{l\in[d-1]} S^{l,i}_{n_l} = n_i - r_i\right)\\&= \frac{\sum_1^{d-1} r_i}{\sum_1^{d-1} n_i}\prod_{i=1}^{d}\mathbb{P}\left(\sum_{l\in[d]} S^{l,i}_{n_l} = n_i - r_i, S^{d,i}_{n_d} = 0\right)\\&= \mathbb{P}_{(r_1,\dots,r_{d-1})}\left(O_i = n_i, i\in[d-1]\right)\prod_j \mathbb{P}\left(S^{d,j}_{n_d} = 0\right)\prod_{j\neq d}\mathbb{P}\left(S^{j,d}_{n_j} = 0\right).\end{aligned}$$

This shows that the formula of Theorem 4.2 does not work for the case $n_i = r_i$ for some i. By the same reason, the next result, which is the law of the total number of individuals in a MGW forest, has four additional terms.

Corollary 4.1. *Assume the hypotheses of Theorem 4.2 are satisfied, that $(S^{i,j}, i\in[d])$ are identically distributed for every j, and that $d = 2$. Let $O = \sum_{i\in[d]} O_i$, $n\in\mathbb{N}$ and $r < n$. Let $S^{(j)}_n$ have law $\sum_l S^{l,j}_{n_l}$ for any $n_1 + \dots + n_d = n$. Then*

$$\begin{aligned}\mathbb{P}_{\mathbf{r}}(O = n) =& \frac{r}{n}\mathbb{P}\left(S^{(1)}_n + S^{(2)}_n = n - r\right)\\&+ \frac{r_1}{n - r_2}\mathbb{P}\left(S^{(1)}_{n-r_2} = n - r, S^{(1)}_{r_2} = 0\right)\mathbb{P}\left(S^{(2)}_n = 0\right)\\&+ \frac{r_2}{n - r_1}\mathbb{P}\left(S^{(1)}_n = 0\right)\mathbb{P}\left(S^{(2)}_{r_1} = 0, S^{(2)}_{n-r_1} = n - r\right)\\&- \frac{r}{n}\mathbb{P}\left(S^{(1)}_n = n - r\right)\mathbb{P}\left(S^{(2)}_n = 0\right) - \frac{r}{n}\mathbb{P}\left(S^{(1)}_n = 0\right)\mathbb{P}\left(S^{(2)}_n = n - r\right).\end{aligned}$$

Proof. We have to sum over all n_1, n_2 such that $n_1 + n_2 = n$. Note that $n_i \geq r_i$, and also $r_1 \leq n_1 \leq n - r_2$. It follows

$$\begin{aligned}&\mathbb{P}_{\mathbf{r}}(O = n)\\&= \sum_{n_1: r_1 < n_1 < n - r_2}\mathbb{P}_{\mathbf{r}}(O_1 = n_1, O_2 = n - n_1) + \mathbb{P}_{\mathbf{r}}(O_1 = r_1, O_2 = n - r_1) + \mathbb{P}_{\mathbf{r}}(O_1 = n - r_2, O_2 = r_2)\\&= \frac{r}{n}\sum_{n_1: r_1\leq n_1\leq n - r_2}\mathbb{P}\left(S^{1,1}_{n_1} + S^{2,1}_{n-n_1} = n_1 - r_1\right)\mathbb{P}\left(S^{1,2}_{n_1} + S^{2,2}_{n-n_1} = n - n_1 - r_2\right)\\&\quad + \mathbb{P}_{\mathbf{r}}(O_1 = r_1, O_2 = n - r_1) + \mathbb{P}_{\mathbf{r}}(O_1 = n - r_2, O_2 = r_2)\\&- \frac{r}{n}\mathbb{P}(S^{11}_{n-r_2} + S^{21}_{r_2} = n - r)\mathbb{P}(S^{12}_{n-r_2} + S^{22}_{r_2} = 0) - \frac{r}{n}\mathbb{P}(S^{11}_{r_1} + S^{21}_{n-r_1} = 0)\mathbb{P}(S^{12}_{r_1} + S^{22}_{n-r_1} = n - r).\end{aligned}$$

Making the change of variables $n_1 - r_1 = l_1$, using the convolution formula and Theorem 4.2, gives the desired result. □

4.5.1 Laws of some MGW forests conditioned by the number of individuals of each type

We provide three examples where Hypotheses **H2** are satisfied, under the assumptions of Theorem 4.2. For simplicity, we consider $d = 2$, but the proofs also work for any $d \geq 3$. We perform the summation in Equation (4.4) explicitly in the next examples.

Geometric Offspring

Fix $\mathbf{r} = (r_1, r_2) \in \mathbb{N}^2$ and $\mathbf{n} = (n_1, n_2) \in \mathbb{N}_+^2$. Denote by $\mathbb{F}_{\mathbf{r},\mathbf{n}}^{plane}$ the set of two-type plane forests having r_i roots and n_i individuals of type i, for $i \in [d]$.

On the other hand, for any $p \in (0,1)$ let $\mathscr{G}_{r,p}$ be a two-type Galton-Watson forest with r_i roots of type i, having geometric offspring distribution with parameter p independently for each individual, that is, $\nu_i(k_1,k_2) = p(1-p)^{k_1} p(1-p)^{k_2}$. Recall that for any $F \in \mathbb{F}_{\mathbf{r},\mathbf{n}}^{plane}$, we denote by $F^{(i)}$ the subforest of type i of F. Suppose that $F \in \mathbb{F}_{\mathbf{r},\mathbf{n}}^{plane}$ has $k_{1,2}$ type 2 individuals whose parent is of type 1, and $k_{2,1}$ type 1 individuals whose parent is of type 2. Hence, $n_1 - r_1 - k_{2,1}$ is the number of individuals type 1 whose parent is of type 1, and similarly for the type 2 individuals. Denoting by $c_i(v)$ the number of children type i that vertex v has, then

$$\begin{aligned}
\mathbb{P}(\mathscr{G}_{\mathbf{r},p} = F) &= \prod_{v \in f^{(1)}} \nu_1(c_1(v), c_2(v)) \prod_{v \in f^{(2)}} \nu_2(c_1(v), c_2(v)) \\
&= p^{2n}(1-p)^{n_1 - r_1 - k_{21}}(1-p)^{k_{12}}(1-p)^{k_{21}}(1-p)^{n_2 - r_2 - k_{12}} \\
&= p^{2n}(1-p)^{n-r},
\end{aligned}$$

where $n = n_1 + n_2$ y $r = r_1 + r_2$.

Now, we compute the left-hand side of Hypotheses **H2**. Recall that the sum of k independent geometric random variables with parameter p, has a negative binomial distribution $NB_{k,p}$ of parameters k and p. From Equation (4.4), one obtains the sum

$$\begin{aligned}
&\sum_{i=0}^{n_1 - r_1} i \mathbb{P}\left(S_{n_1}^{1,1} = n_1 - r_1 - i\right) \mathbb{P}\left(S_{n_2}^{2,1} = i\right) \\
&= \sum_{i=0}^{n_1 - r_1} i \binom{n_1 + n_1 - r_1 - i - 1}{n_1 - r_1 - i} \binom{n_2 + i - 1}{i} p^{n_1}(1-p)^{n_1 - r_1 - i} p^{n_2}(1-p)^i \\
&= p^n (1-p)^{n_1 - r_1} \sum_{i=1}^{n_1 - r_1} \binom{n_1 + n_1 - r_1 - i - 1}{n_1 - r_1 - i} \binom{n_2 + i - 2}{i - 1} (n_2 + i - 1) \\
&= p^n (1-p)^{n_1 - r_1} \sum_{i=0}^{n_1 - r_1 - 1} (n_2 + i) \binom{n_1 + n_1 - r_1 - 1 - i - 1}{n_1 - r_1 - 1 - i} \binom{n_2 + i - 1}{i},
\end{aligned}$$

making a change of variable in the last step. For any $m, n_1, n_2 \in \mathbb{N}$, we use the equality

$$\sum_{i=0}^{m} \binom{n_1 + m - i - 1}{m - i} \binom{n_2 + i - 1}{i} = \binom{n_1 + n_2 + m - 1}{m},$$

which can be proved comparing the binomial coefficients in the convolution of two negative binomial random variables. Hence, if we define a function $f : [n_1 - r_1]_0 \mapsto \mathbb{R}_+$, as

$$f(k) = \sum_{i=0}^{n_1-r_1-k} i \binom{n_1+n_1-r_1-k-i-1}{n_1-r_1-k-i} \binom{n_2+i-1}{i},$$

we obtain

$$\begin{aligned} f(0) &= n_2 \binom{n+n_1-r_1-2}{n_1-r_1-1} + \sum_{i=0}^{n_1-r_1-1} i \binom{n_1+n_1-r_1-1-i-1}{n_1-r_1-1-i} \binom{n_2+i-1}{i} \\ &= n_2 \binom{n+n_1-r_1-2}{n_1-r_1-1} + f(1) \\ &= n_2 \sum_{j=1}^{n_1-r_1} \binom{n+n_1-r_1-1-j}{n_1-r_1-j} \\ &= n_2 \sum_{j=0}^{n_1-r_1-1} \binom{n-1+j}{j}, \end{aligned}$$

making a change of variable. Now, for any $m, r \in \mathbb{N}$ use the identity

$$\sum_{j=0}^{m} \binom{r+j}{j} = \binom{r+m+1}{m},$$

to deduce that

$$\sum_{i=0}^{n_1-r_1} i\mathbb{P}\big(S^{1,1}_{n_1} = n_1-r_1-i\big)\,\mathbb{P}\big(S^{2,1}_{n_2} = i\big) = n_2 \binom{n+n_1-r_1-1}{n_1-r_1-1} p^n (1-p)^{n_1-r_1}. \tag{4.6}$$

We compare this quantity with the right-hand side of Hypotheses **H2**:

$$\begin{aligned} \frac{n_2}{n}(n_1-r_1)\mathbb{P}(NB_{n,p} = n_1-r_1) &= \frac{n_2}{n}(n_1-r_1)\binom{n+n_1-r_1-1}{n_1-r_1} p^n (1-p)^{n_1-r_1} \\ &= \frac{n_2}{n} \frac{(n+n_1-r_1-1)!}{(n_1-r_1-1)!(n-1)!} p^n (1-p)^{n_1-r_1} \\ &= n_2 \binom{n+n_1-r_1-1}{n_1-r_1-1} p^n (1-p)^{n_1-r_1}, \end{aligned}$$

which is identical to (4.6).

Thus, using Theorem 4.2, denoting by $\#_i\mathscr{G}_{\mathbf{r},p}$ the number of individuals of type i, we obtain

$$\mathbb{P}(\#_1\mathscr{G}_{\mathbf{r},p} = n_1, \#_2\mathscr{G}_{\mathbf{r},p} = n_2) = \frac{r}{n} \binom{n+n_1-r_1-1}{n_1-r_1} \binom{n+n_2-r_2-1}{n_2-r_2} p^{2n} (1-p)^{n-r}.$$

It follows that

$$\mathbb{P}(\mathscr{G}_{\mathbf{r},p} = \mathbf{f} | \#_1\mathscr{G}_{\mathbf{r},p} = n_1, \#_2\mathscr{G}_{\mathbf{r},p} = n_2) = \frac{1}{\frac{r}{n}\binom{n+n_1-r_1-1}{n_1-r_1}\binom{n+n_2-r_2-1}{n_2-r_2}} \qquad \forall\, \mathbf{f} \in \mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}},$$

being uniform on the set of two-type plane forests with r_i roots type i, and n_i vertices of type i. Note that this implies that the denominator on the right-hand side is the number of two-type plane forests with root-type $\mathbf{r}$ and individuals-type $\mathbf{n}$. We also obtain the distributional equality

$$\mathscr{F}^{plane}_{\mathbf{r},\mathbf{n}} \overset{d}{=} (\mathscr{G}_{\mathbf{r},p} | \#_1\mathscr{G}_{\mathbf{r},p} = n_1, \#_2\mathscr{G}_{\mathbf{r},p} = n_2),$$

where $\mathscr{F}^{plane}_{\mathbf{r},\mathbf{n}}$ is uniform on $\mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}}$.

General case

In the general case $d \in \mathbb{N}$, using Theorem 4.2, denoting by $\#_i\mathscr{G}_{\mathbf{r},p}$ the number of individuals of type i of $\mathscr{G}_{\mathbf{r},p}$, we obtain

$$\mathbb{P}(\#_i\mathscr{G}_{\mathbf{r},p} = n_i, \forall\, i \in [d]) = \frac{r}{n} p^{dn}(1-p)^{n-r} \prod_{i\in[d]} \binom{n+n_i-r_i-1}{n_i-r_i}.$$

It follows that

$$\mathbb{P}(\mathscr{G}_{\mathbf{r},p} = F\,|\#_i\mathscr{G}_{\mathbf{r},p} = n_i, \forall\, i \in [d]) = \frac{1}{\frac{r}{n}\prod_{i\in[d]}\binom{n+n_i-r_i-1}{n_i-r_i}} \qquad \forall\, F \in \mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}},$$

being uniform on $\mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}}$. Note that the previous agrees with the unidimensional case, see Formula (35) in [Pit98].

Poisson Offspring

Let $\mathbf{r} = (r_1, r_2) \in \mathbb{N}^2$ with $r = r_1 + r_1 > 0$ and $\mathbf{n} = (n_1, n_2) \in \mathbb{N}^2$. For $\mu \in \mathbb{R}^+$, let $\mathscr{P}_{\mathbf{r},\mu}$ be a two-type GW with r_i roots type i, and Poisson offspring distribution of parameter μ, for every individual, independently from anyone, that is, $\nu_i(k_1,k_2) = e^{-\mu}\mu^{k_1}e^{-\mu}\mu^{k_2}/(k_1!k_2!)$.

Similarly as in the previous example, consider any $F \in \mathbb{F}^{plane}_{r,n}$ having r_i roots type i, $k_{1,2}$ type 2 individuals whose parent is of type 1, and $k_{2,1}$ type 1 individuals whose parent is of type 2. Then

$$\begin{aligned}
&\mathbb{P}\big(\mathscr{P}_{\mathbf{r},\mu} = F\big)\\
&= e^{-2\mu n_1}\mu^{n_1-r_1-k_{21}}\mu^{k_{12}} \prod_{v\in f^{(1)}} \frac{1}{c_1(v)!} \prod_{v\in f^{(1)}} \frac{1}{c_2(v)!}\, e^{-2\mu n_2}\mu^{n_2-r_2-k_{12}}\mu^{k_{21}} \prod_{v\in f^{(2)}} \frac{1}{c_1(v)!} \prod_{v\in f^{(2)}} \frac{1}{c_2(v)!}\\
&= e^{-2\mu n}\mu^{n-r} \prod_{i:v_i\in f} \frac{1}{c_1(v_i)!c_2(v_i)!},
\end{aligned}$$

where the product is over any enumeration of the n vertices in F.

We compute the left-hand side of Hypotheses H2. Recall that the sum of k independent Poisson random variables with parameter μ, has a Poisson distribution with parameter $k\mu$. Then

$$\begin{aligned}
&\sum_{i=0}^{n_1-r_1} i\mathbb{P}\big(S^{11}_{n_1} = n_1 - r_1 - i\big)\,\mathbb{P}\big(S^{21}_{n_2} = i\big)\\
&= e^{-n\mu}\sum_{i=0}^{n_1-r_1} i\frac{(n_1\mu)^{n_1-r_1-i}}{(n_1-r_1-i)!}\frac{(n_2\mu)^i}{i!}\\
&= e^{-n\mu}\frac{(n_1\mu)^{n_1-r_1}}{(n_1-r_1)!}\sum_{i=0}^{n_1-r_1} i\binom{n_1-r_1}{i}\left(\frac{n_2}{n_1}\right)^i.
\end{aligned}$$

To simplify the sum, note that

$$\begin{aligned}
&\sum_{i=0}^{n_1-r_1} i\binom{n_1-r_1}{i}\left(\frac{n_2}{n_1}\right)^i \\
&= (n_1-r_1)\sum_{i=1}^{n_1-r_1}\frac{(n_1-r_1-1)!}{(i-1)!(n_1-r_1-i)!}\left(\frac{n_2}{n_1}\right)^i \\
&= (n_1-r_1)\frac{n_2}{n_1}\sum_{i=0}^{n_1-r_1-1}\frac{(n_1-r_1-1)!}{i!(n_1-r_1-i-1)!}\left(\frac{n_2}{n_1}\right)^i \\
&= (n_1-r_1)\frac{n_2}{n_1}\left(1+\frac{n_2}{n_1}\right)^{n_1-r_1-1} \\
&= \frac{n_2}{n}(n_1-r_1)\left(\frac{n}{n_1}\right)^{n_1-r_1}.
\end{aligned}$$

Hence, it follows that

$$\sum_{i=0}^{n_1-r_1} i\mathbb{P}\big(S^{11}_{n_1}=n_1-r_1-i\big)\,\mathbb{P}\big(S^{21}_{n_2}=i\big) = e^{-n\mu}\frac{\mu^{n_1-r_1}}{(n_1-r_1)!}\frac{n_2}{n}(n_1-r_1)n^{n_1-r_1}.$$

This is the same as the right-hand side of Hypotheses **H2**, since

$$\begin{aligned}
\frac{n_2}{n}(n_1-r_1)\mathbb{P}\big(P_{n\mu}=n_1-r_1\big) &= e^{-n\mu}\frac{n_2}{n}(n_1-r_1)\frac{(n\mu)^{n_1-r_1}}{(n_1-r_1)!} \\
&= e^{-n\mu}\frac{\mu^{n_1-r_1}}{(n_1-r_1)!}\frac{n_2}{n}(n_1-r_1)n^{n_1-r_1}.
\end{aligned}$$

Denoting by $\#_i\mathscr{P}_{\mathbf{r},p}$ the number of individuals type i, using Theorem 4.2 we obtain

$$\mathbb{P}_r\left(\#_1\mathscr{P}_{\mathbf{r},\mu}=n_1, \#_2\mathscr{P}_{\mathbf{r},\mu}=n_2\right) = \frac{e^{-2n\mu}(n\mu)^{n-r}}{(n_1-r_1)!(n_2-r_2)!}\frac{r}{n},$$

and hence

$$\mathbb{P}\big(\mathscr{P}_{\mathbf{r},\mu}=F\,|\#_1\mathscr{P}_{\mathbf{r},\mu}=n_1, \#_2\mathscr{P}_{\mathbf{r},\mu}=n_2\big) = \frac{\binom{n_1-r_1}{c_1(v_1),\dots,c_1(v_n)}\binom{n_2-r_2}{c_2(v_1),\dots,c_2(v_n)}}{\frac{r}{n}n^{n-r}} \qquad \forall\, \mathbf{f}\in\mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}}.$$

General case

In the general case, from Theorem 4.2 we obtain

$$\mathbb{P}_r\left(\#_i\mathscr{P}_{\mathbf{r},\mu}=n_i, \forall i\in[d]\right) = \frac{r}{n}\frac{e^{-dn\mu}(n\mu)^{n-r}}{\prod_i(n_i-r_i)!}.$$

From which we have

$$\mathbb{P}\big(\mathscr{P}_{\mathbf{r},\mu}=F\,|\#_i\mathscr{P}_{\mathbf{r},\mu}=n_i, \forall i\in[d]\big) = \frac{\prod_i\binom{n_i-r_i}{c_i(v_1),\dots,c_i(v_n)}}{\frac{r}{n}n^{n-r}} \qquad \forall\, F\in\mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}}.$$

Note that this agrees with the unidimensional case, as seen in Formula (39) of [Pit98]. Since the right-hand side depends on F, it is not uniform on the set of plane forests. To obtain a uniform forest, we introduce a function as in [Pit98]. Define $\Psi:\mathbb{F}^{labeled}_{\mathbf{r},\mathbf{n}}\mapsto\mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}}$ as follows:

1. Order the trees of the forest, according to the natural order of the labels in the roots of type 1, then order the type 2 roots, and so on.

2. For each vertex v_i of type i, order its $c_1(v_i)$ children of type 1 according to its labels, its $c_2(v_i)$ children of type 2 according to its labels, and so on.

3. Erase the labels.

Now, we find the number of forests in $\mathbb{F}^{labeled}_{\mathbf{r},\mathbf{n}}$ that are sent to a given plane forest F. For each i, there are $(n_i - r_i)!$ ways to label the type i vertices (recall that our rooted labeled forests have root set $[r]$). But the permutation of the childrens of a fixed type of each vertex also lead to the same forest F. That is, if vertex v has $c_i(v)$ children type i, there are $c_i(v)!$ labelings of such children leading to F. This being true for every type and every vertex, we have

$$\#\Psi^{-1}(F) = \frac{\prod_i (n_i - r_i)!}{\prod_{v\in F}\prod_i c_i(v)!}.$$

This is exactly the numerator in the formula obtained above. Thus, we have the following interpretation: let $\mathscr{F}^{labeled}_{\mathbf{r},\mathbf{n}}$ have uniform distribution over the set of all d-type labeled forests, where the roots are in $[r]$, with roots-type $\mathbf{r}$ and individuals-type $\mathbf{n}$, and let $\mathscr{P}^*_{\mathbf{r},\mathbf{n}}$ be $\mathscr{P}_{\mathbf{r},\mathbf{n}}$ relabeled by d uniform random permutations, one for each type, then

$$\mathscr{F}^{labeled}_{\mathbf{r},\mathbf{n}} \stackrel{d}{=} \left(\mathscr{P}^*_{\mathbf{r},p} \,\middle|\, \#_i \mathscr{P}_{\mathbf{r},p} = n_i, \forall\, i \in [d]\right).$$

We note that the previous formulas coincide with the results in [Pit98, Section 7] for the unitype case. But our formulas relate directly enumerations of multitype labeled forests with the unitype enumerations. Recall our labeling on page 105 for forests in $\mathscr{F}^{labeled}_{\mathbf{r},\mathbf{n}}$. The above formulas also imply that the number of multitype labeled forests in $\mathscr{F}^{labeled}_{\mathbf{r},\mathbf{n}}$ with root set $[r]$ coincides with the number of unitype labeled forests on $[n]$ with root set $[r]$, which by Cayley's formula is $(r/n)n^{n-r}$. This comes from the following bijection. Regard each multitype forest $F \in \mathscr{F}^{labeled}_{\mathbf{r},\mathbf{n}}$ as a unitype labeled forest on $[n]$, together with the following labeling: the roots retain their labels and, according with the order on F, the remaining $n_1 - r_1$ type 1 individuals now have the new labels $\{r+1, \ldots, r+n_1-r_1\}$, the remaining $n_2 - r_2$ type 2 individuals have the new labels $\{r+n_1-r_1+1, \ldots, r+n_1-r_1+n_2-r_2\}$, and so on.

Bernoulli Offspring

Let $\mathbf{r} = (r_1, r_2) \in \mathbb{N}^2$ with $r > 0$ and $\mathbf{n} = (n_1, n_2) \in \mathbb{N}^2$. For $0 < p < 1$, let $\mathscr{B}_{\mathbf{r},p}$ be a two-type GW with r_i roots type i, and Bernoulli offspring distribution of parameter p, for each vertex independently of the others, that is, $\nu_i(k_1, k_2) = p^{k_1}(1-p)^{1-k_1} p^{k_2}(1-p)^{1-k_2}$ with $k_1, k_2 \in \{0, 1\}$. Since any vertex v has zero or two children with probability p or $1-p$ respectively, then $\nu_i(c_1(v), c_2(v)) = p^{c_1(v)/2}(1-p)^{1-c_1(v)/2} p^{c_2(v)/2}(1-p)^{1-c_2(v)/2}$.

As before, consider any $F \in \mathbb{F}^{binary}_{\mathbf{r},\mathbf{n}}$ having r_i roots type i, $k_{1,2}$ type 2 individuals whose parent is of type 1, and $k_{2,1}$ type 1 individuals whose parent is of type 2. Note that $k_{1,2}$ and $k_{2,1}$ are even numbers, as well as $n_i - r_i$ for $i = 1, 2$. Hence

$$\begin{aligned}\mathbb{P}(\mathscr{B}_{\mathbf{r},p} = F) &= (p/(1-p))^{(n_1-r_1-k_{21})/2}(1-p)^{n_1}(p/(1-p))^{k_{12}/2}(1-p)^{n_1}\\ &\quad\times (p/(1-p))^{(n_2-r_2-k_{12})/2}(1-p)^{n_2}(p/(1-p))^{k_{21}/2}(1-p)^{n_2}\\ &= p^{(n-r)/2}(1-p)^{2n-(n-r)/2}.\end{aligned}$$

Recall that twice the sum of n independent Bernoulli random variables with parameter p, has distribution two times the Binomial distribution $B_{n,p}$ of parameters n and p. If n is even, we denote the sum over the even numbers up to n as $\{0,2,\dots,n\} = 2[n/2]_0$. The left-hand side of Hypotheses **H2** is

$$\begin{aligned}
&\sum_{i\in 2[(n_1-r_1)/2]_0} i\mathbb{P}\left(S^{1,1}_{n_1} = n_1 - r_1 - i\right)\mathbb{P}\left(S^{2,1}_{n_2} = i\right)\\
&= \sum i\binom{n_2}{i/2} p^{i/2}(1-p)^{n_2-i/2}\binom{n_1}{(n_1-r_1-i)/2} p^{(n_1-r_1-i)/2}(1-p)^{n_1-(n_1-r_1-i)/2}\\
&= 2p^{(n_1-r_1)/2}(1-p)^{n-(n_1-r_1)/2} n_2 \sum_{i\in 2[(n_1-r_1)/2]} \binom{n_2-1}{i/2-1}\binom{n_1}{(n_1-r_1-i)/2}\\
&= 2p^{(n_1-r_1)/2}(1-p)^{n-(n_1-r_1)/2} n_2 \sum_{i\in 2[(n_1-r_1-2)/2]_0} \binom{n_2-1}{i/2}\binom{n_1}{(n_1-r_1-2-i)/2}.
\end{aligned}$$

Note that, since there are n_2 individuals type 2 and at most each can have 2 children, the number $k_{2,1}$ of individuals type 2 having children type 1 is bounded by $2n_2$. Nevertheless, we have $\mathbb{P}\left(S^{2,1}_{n_2} = i\right) = 0$ for $i > 2n_2$, which agrees with the definition of $\binom{n}{k} = 0$ whenever $n < k$ for positive integers. Using Vandermonde's identity, and adding the term $(n_1 - r_1)/2$ in both the numerator and denominator, the above is equal to

$$\begin{aligned}
&2p^{(n_1-r_1)/2}(1-p)^{n-(n_1-r_1)/2} n_2\binom{n-1}{(n_1-r_1-2)/2}\\
&= \frac{n_2}{n}(n_1-r_1)\binom{n}{(n_1-r_1)/2} p^{(n_1-r_1)/2}(1-p)^{n-(n_1-r_1)/2}.
\end{aligned}$$

The right-hand side of Hypotheses **H2** is

$$\frac{n_2}{n}(n_1-r_1)\mathbb{P}(2B_{n,p} = n_1-r_1) = \frac{n_2}{n}(n_1-r_1)\binom{n}{(n_1-r_1)/2} p^{(n_1-r_1)/2}(1-p)^{n-(n_1-r_1)/2}.$$

Therefore, Hypotheses **H2** are satisfied and

$$\mathbb{P}_{\mathbf{r}}\left(O_1 = n_1, O_2 = n_2\right) = \frac{r}{n}\binom{n}{(n_1-r_1)/2}\binom{n}{(n_2-r_2)/2} p^{(n-r)/2}(1-p)^{2n-(n-r)/2}.$$

Denoting by $\#_i\mathscr{B}_{\mathbf{r},p}$ the number of individuals type i, we have

$$\mathbb{P}(\mathscr{B}_{\mathbf{r},p} = F \,|\, \#_1\mathscr{B}_{\mathbf{r},p} = n_1, \#_2\mathscr{B}_{\mathbf{r},p} = n_2) = \frac{1}{\frac{r}{n}\binom{n}{(n_1-r_1)/2}\binom{n}{(n_2-r_2)/2}} \qquad \forall\, F \in \mathbb{F}^{binary}_{\mathbf{r},\mathbf{n}},$$

being uniform on $\mathbb{F}^{binary}_{\mathbf{r},\mathbf{n}}$.

General case

For $d \in \mathbb{N}$, Theorem 4.2 implies

$$\mathbb{P}_r\left(\#_i\mathscr{B}_{\mathbf{r},\mu} = n_i, \forall i \in [d]\right) = p^{(n-r)/2}(1-p)^{dn-(n-r)/2}\frac{r}{n}\prod_i \binom{n}{(n_i-r_i)/2}.$$

This implies

$$\mathbb{P}(\mathscr{B}_{\mathbf{r},p} = F \,|\#_i\mathscr{B}_{\mathbf{r},p} = n_i, \forall\, i \in [d]) = \frac{1}{\frac{r}{n}\prod\binom{n}{(n_i-r_i)/2}} \qquad \forall\, F \in \mathbb{F}^{binary}_{\mathbf{r},\mathbf{n}},$$

being uniform on the set of binary d-type plane forests with root-type $\mathbf{r}$ and individuals-type $\mathbf{n}$. Compare this formula with the number of binary trees, which is related to the Catalan numbers, see Theorem 2.1 in [Drm09].

4.6 Algorithms

In this section we review the algorithms for the simulation of the unitype random forests presented before. Then, those algorithms are generalized to the multidimensional case.

First, we present how to obtain a degree sequence approaching any given distribution. After that, hubs are added to the algorithm to ensure individuals with a lot of children. Being able to obtain degree sequences, we repeat the algorithm to simulate a uniform tree with such given degree sequence. Using this we describe the simulation of CWG(n) trees given in [Dev12]. The construction for forests was given in Lemma 4.5, but we explicitly write the algorithm. As a side note, a new algorithm is proposed to simulate a tree which has offspring distribution almost as a CGW(n). This algorithm works incredibly fast. This is done by fixing $n_1 \in \mathbb{N}$ and positive ε, and constructing a CGW(n) tree, where $|n-n_1|/n_1 \le \varepsilon$.

After that, we present the counterparts in the multidimensional case. First, the construction of uniformly sampled multitype forests with a given degree sequence, which is Theorem 4.1. Finally, we present an algorithm to obtain MGW forests conditioned by its number of individuals for each type, which is a generalization of Devroye's algorithm.

4.6.1 Simulation of a degree sequence whose normalization approaches a given distribution

Fix any critical distribution $\nu = (\nu_i, i \ge 0)$. The objective is to find a degree sequence $\mathbf{S_n} = (N^n_i, i \ge 0)$ with $s_n = |\mathbf{S_n}|$ such that $\forall\ i \ge 0$

$$\frac{N^n_i}{s_n} \to \nu_i. \tag{4.7}$$

It turns out that for a CGW(n) tree T^n having offspring distribution ν and n vertices, the convergence of the normalized empirical degree sequence has been proved in Lemma 11 of [BM14b]. In such lemma the authors assume that the offspring distribution has finite variance. The latter means that if we want to construct degree sequences, we can construct them roughly setting $N^n_i = \lfloor \nu_i s_n \rfloor$.

Let T be a GW with offspring distribution $\nu = (\nu_i; i \ge 0)$, that is in the domain of attraction of an α-stable law, for short DA(α), with parameter $\alpha \in (1,2)$. This means that $\nu([0,\infty)) = j^{-\alpha}L(j)$ where $L : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is a *slowly varying* function, that is $\lim_{x\to\infty} L(tx)/L(x) = 1$ for every $t > 0$. See [BGT89, Chapter 8.3] for more details. Denote by $\mathbb{P}_\nu$ the probability distribution of T, and by $(\hat{n}_i; i \ge 0)$ the empirical degree sequence of T, that is

$$\hat{n}_i = \sum_{j=1}^{|T|} I\left(c_j = i\right), \tag{4.8}$$

with c_j the number of children of individual j. Define the normalized empirical degree sequence $\hat{\nu} = (\hat{\nu}_i; i \geq 0)$, where $\hat{\nu}_i = \hat{n}_i/|T|$. The law of the CGW($n$) is denoted by $\mathbb{P}^n_\nu(\cdot) = \mathbb{P}_\nu(\cdot|\, |T| = n)$, and we only consider n for which this has sense.

We first generalize the result of [BM14b] in the next Lemma, by dropping the finite variance condition. The proof is given in the Appendix, page 155.

Lemma 4.8. *If ν is critic and aperiodic, then under $\mathbb{P}^n_\nu$*

$$\hat{\nu} \overset{(d)}{\to} \nu,$$

Using this result, we can obtain uniform trees with a given degree sequence, *behaving as trees having a Pareto distribution*, which is the content of Algorithm 1.

Algorithm 1 Generate a degree sequence from a given distribution

Input: A distribution ν and a natural c_n.
Output: A degree sequence, which normalized, approaches ν.

```
Let M = inf{i : ⌊c_n ν_i⌋ = 1}.
for i = 1 to M do
    N_i = ⌊c_n ν_i⌋
end for
Set s_n = 1 + ∑_1^M i N_i.
Set N_0 = s_n − ∑_1^M N_i.
Define the degree sequence S_n = (N_0, ..., N_M).
```

Lemma 4.9. *Let $\mathbf{S_n}$ be defined as in Algorithm 1, and consider any sequence $c_n \uparrow \infty$. Assume the distribution ν satisfies the Hypotheses of Lemma 4.8. Then the convergence in (4.7) holds true.*

Proof. We emphasize the dependence of the degree sequence in n writing N_i^n, and also M^n. Since, for any $i \geq 1$ we have $0 \leq iN_i^n/c_n \leq i\nu_i$ for every n, then, by the Weierstrass test

$$\frac{s_n}{c_n} = \frac{1}{c_n} + \sum_1^{M^n} i\frac{\lfloor c_n \nu_i \rfloor}{c_n} = \frac{1}{c_n} + \sum_1^{\infty} i\frac{\lfloor c_n \nu_i \rfloor}{c_n} \to \sum_1^{\infty} i\nu_i,$$

which equals 1, since ν is critic. This easily implies $N_i^n/s_n \to \nu_i$ for every $i \geq 1$, and also

$$\frac{N_0^n}{s_n} = \frac{s_n}{c_n} - \sum_1^{M^n} \frac{\lfloor c_n \nu_i \rfloor}{c_n} \to 1 - \sum_1^{\infty} \nu_i = \nu_0. \qquad \square$$

Next, we add hubs to the degree sequence, that is, individuals with many children. Those individuals will have $I_{\bar{M}-i+1} = \lfloor \beta_i b_{s_n} \rfloor$ children, for $\bar{M}$ fixed positive reals $\beta_1 > \cdots > \beta_{\bar{M}} > 0$ and where b_{s_n} is the scaling for the BFW (see Hypothesis 2 in Theorem 3.1). If necessary, we choose c_n big enough such that $\lfloor \beta_{i+1} b_{s_n} \rfloor < \lfloor \beta_i b_{s_n} \rfloor$ whenever $\beta_{i+1} < \beta_i$. This condition ensures there are no unnecessary repetitions in the child sequence. We also impose $\lfloor \beta_{\bar{M}} b_{s_n} \rfloor > M$, since M is the maximum number of children obtained in Algorithm 1. This is given in Algorithm 2.

Algorithm 2 Generate a degree sequence from a given distribution and having hubs

Input: A distribution ν, a natural c_n, and $\bar{M}$ positive reals $\beta_1 > \cdots > \beta_{\bar{M}}$.
Output: A degree sequence, which normalized, approaches ν, and has individuals with $\lfloor \beta_i b_{s_n} \rfloor$ children, for $i \in [\bar{M}]$.

Obtain a degree sequence $\mathbf{S_n} = (N_0, \ldots, N_M)$ and s_n from Algorithm 1.
Define the degree sequence $\bar{\mathbf{S}}_\mathbf{n} = (\bar{N}_0, \bar{N}_1, \ldots, \bar{N}_{I_{\bar{M}}})$, as
if $i \in \{1, \ldots, M\}$ **then**
    $\bar{N}_i = N_i$
else if $i \in \{I_1, I_2, \ldots, I_{\bar{M}}\}$ **then**
    $\bar{N}_i = \#\{k : I_k = I_i\}$
else if $i = 0$ **then**
    $\bar{N}_0 = N_0 + \sum_{i=1}^{\bar{M}} (I_i - 1)\bar{N}_{I_i}$
else
    $\bar{N}_i = 0$
end if
Set $\bar{s}_n = \sum \bar{N}_i$.

The fact that this algorithm gives us a degree sequence, follows from

$$\sum \bar{N}_i = \bar{N}_0 + \sum_1^M N_i + \sum_1^{\bar{M}} \bar{N}_{I_i} = s_n + \sum_{i=1}^{\bar{M}} (I_i - 1)\bar{N}_{I_i} + \sum_1^{\bar{M}} \bar{N}_{I_i} = s_n + \sum_{i=1}^{\bar{M}} I_i \bar{N}_{I_i},$$

and

$$1 + \sum i\bar{N}_i = 1 + \sum_1^M iN_i + \sum_1^{\bar{M}} I_i \bar{N}_{I_i} = s_n + \sum_1^{\bar{M}} I_i \bar{N}_{I_i}.$$

Note that the ratio of the number of individuals from the two algorithms is given by

$$\frac{\bar{s}_n}{s_n} = \sum_0^M \frac{N_i}{s_n} + \sum_1^{\bar{M}} \frac{\bar{N}_{I_i}}{s_n}.$$

In Lemma 4.9 we proved the first term goes to one, thus, it suffices to assume the second term goes to zero to ensure such new degree sequence also approaches to the given distribution ν.

Lemma 4.10. *Let $\bar{\mathbf{S}}_\mathbf{n}$ be defined as in Algorithm 2, and consider any sequence $c_n \uparrow \infty$. Assume the distribution ν satisfies the Hypotheses of Lemma 4.8. If*

$$\sum_1^{\bar{M}^n} \frac{\bar{N}^n_{I_i^n}}{s_n} \to 0,$$

then, for every $i \geq 0$ the convergence

$$\frac{\bar{N}_i^n}{\bar{s}_n} \to \nu_i$$

holds true.

4.6.2 Constrained simulation of random trees in the unidimensional case

The paper [Dev12] gives an algorithm to simulate unitype GW trees with offspring distribution ν conditioned to have size n. The idea is: simulate a multinomial vector $(N_0,\dots,N_K)$ with parameters $(n,\nu_0,\nu_1,\dots)$ such that

$$n = \sum N_i = 1 + \sum iN_i, \tag{4.9}$$

that is, simulate the degree sequence of a CGW(n). Then, obtain a uniform tree with degree sequence $(N_0,\dots,N_K)$. The resulting tree will have law $\mathbb{P}^n_\nu$.

First, we give an algorithm to simulate uniform trees with a given degree sequence $\mathbf{S} = (N_i; i \geq 0)$ satisfying (4.9). Algorithm 3 is obtained from [BM14b] and was proved in Lemma 4.5.

Algorithm 3 Generate uniformly sampled trees with a given degree sequence

---

Input: A degree sequence $\mathbf{S} = (N_i; i \geq 0)$ with $\sum N_i = 1 + \sum iN_i = s$.
Output: A uniformly sampled tree with the given degree sequence.

Define the vector $\mathbf{c} = (c_1, c_2, \dots, c_s)$, with N_0 zeros, N_1 ones, etc.
Set $\pi = (\pi_1,\dots,\pi_s)$ a uniform random permutation of $[s]$.
For $j \in [s]$ define the walk

$$W^b(j) = \sum_1^j (\mathbf{c} \circ (\pi_i) - 1),$$

satisfying $W^b(0) = 0$ and $W^b(s) = -1$.
Let $i^* = \min\{j \in [s] : W^b(j) = \min_{1 \leq i \leq s} W^b(i)\}$ be the first time the partial sums reaches its minimum.
For $j \in [s]$ define the walk $V(W^b)$ of length s as

$$V(W^b)(j) = \sum_1^j (\mathbf{c} \circ (\pi_{i^*+j}) - 1),$$

with $i^* + j \bmod s$.
Generate the tree with breadth-first walk $V(W^b)$.

Using Algorithm 3, we give and prove Algorithm 4, which was proposed in [Dev12].

Algorithm 4 Generate a GW tree conditioned to have size n

Input: A distribution ν and a natural n.
Output: A tree with law $\mathbb{P}^n_\nu$.

1: Generate a multinomial vector $\mathbf{S} = (N_0, N_1, \dots)$ with parameters $(n, \nu_0, \nu_1, \dots)$.
2: Let K be the last non-zero component of $\mathbf{S}$, that is $N_j = 0$ for $j > K$ and $N_K > 0$.
3: Define $\Xi(\mathbf{S}) = 1 + \sum iN_i$.
4: **if** $\Xi(\mathbf{S}) = n$ **then**
5:     go to step 9
6: **else**
7:     repeat from step 1
8: **end if**
9: Apply Algorithm 1 to the degree sequence $(N_0,\dots,N_K)$.

The following lemma proves Algorithm 4 gives us an CWG(n) tree.

Lemma 4.11. *Let $\mathbf{S}_{(i)}$ be the ith vector obtained by step 1 of Algorithm 4, and let $K = \inf\{i : \Xi(\mathbf{S}_{(i)}) = n\}$. If $\tau_{(K)}$ is the tree obtained in step 9, then $\tau_{(K)}$ has the same law as a CGW(n) tree.*

Proof. Let $\mathbf{S}$ be a vector with the same distribution as $\mathbf{S}_{(1)}$. For any vector $s = (n_0, \dots, n_k)$ with $\sum n_i = n$, by definition we have

$$\mathbb{P}(\mathbf{S} = s) = \binom{n}{n_0, \dots, n_k} \prod \nu_i^{n_i}.$$

Denote by W^b the bridge with increments $\mathbf{c} \circ \pi_i - 1$, where π is a uniform permutation of the child sequence $\mathbf{c} = (c_i)$, the latter obtained from $\mathbf{S}_{(K)}$. Denote by W its Vervaat transform, which codes the tree $\tau_{(K)}$. Then, for any bridge w^b of size n, having Vervaat transform w and degree sequence $s = (n_0, \dots, n_k)$ we have

$$\mathbb{P}(W = w) = n\mathbb{P}\left(W^b = w^b\right) = \frac{n}{\binom{n}{n_0, \dots, n_k}} \mathbb{P}\left(\mathbf{S}_{(K)} = s\right),$$

since there are n bridges mapped to w by the Vervaat transform, and there are $\binom{n}{n_0, \dots, n_k}$ different labelings of such bridges. For the last term, we sum over all possible values of K and use independence between simulations

$$\begin{aligned}
\mathbb{P}\left(\mathbf{S}_{(K)} = s\right) &= \sum_k \mathbb{P}\left(\mathbf{S}_{(k)} = s, \Xi(\mathbf{S}_{(k)}) = n, \Xi(\mathbf{S}_{(j)}) \neq n,\ j < k\right) \\
&= \mathbb{P}(\mathbf{S} = s, \Xi(\mathbf{S}) = n) \sum_{k \geq 1} \mathbb{P}(\Xi(\mathbf{S}) \neq n)^{k-1} \\
&= \mathbb{P}(\mathbf{S} = s, \Xi(\mathbf{S}) = n) \frac{1}{\mathbb{P}(\Xi(\mathbf{S}) = n)}.
\end{aligned}$$

Note that

$$\mathbb{P}(\Xi(\mathbf{S}) = n) = \sum_{\substack{s = (n_0, \dots): \\ \sum n_i = 1 + \sum i n_i = n}} \mathbb{P}(\mathbf{S} = s) = \sum_{\substack{s = (n_0, \dots): \\ \sum n_i = 1 + \sum i n_i = n}} \binom{n}{n_0, \dots, n_k} \prod \nu_i^{n_i},$$

where the sum is over all degree sequences of plane trees having size n. We relate this with ν^{*n}, the n-th convolution of the law ν with itself. Using the formula for the convolution

$$\nu^{*n}(n-1) = \sum_{\substack{(i_1, \dots, i_n): \\ \sum i_k = n-1}} \prod_{k=1}^{n} \nu_{i_k}.$$

Fix any degree sequence $(n_i, i \geq 0)$ with $\sum n_i = n$, and note that the number of vectors $(i_1, \dots, i_n)$ with $\sum i_k = n - 1$ such that

$$\prod_{k=1}^{n} \nu_{i_k} = \prod_{k \geq 0} \nu_k^{n_k},$$

is equal to the number of different bridges w^b of size n, having degree sequence $(n_i, i \geq 0)$. This number is $\binom{n}{n_0, n_1 \dots}$, therefore

$$\nu^{*n}(n-1) = \sum_{\substack{s = (n_0, \dots): \\ \sum n_i = 1 + \sum i n_i = n}} \binom{n}{n_0, \dots, n_k} \prod \nu_i^{n_i}.$$

The latter, together with the Otter-Dwass formula [Ott49, Dwa69] imply

$$\mathbb{P}(\Xi(\mathbf{S}) = n) = \nu^{*n}(n-1) = n\mathbb{P}(|\tau| = n)\,.$$

If w codes the tree T, then

$$\begin{aligned}
\mathbb{P}(\tau_K = T) &= \frac{n}{\binom{n}{n_0,\ldots,n_k}}\mathbb{P}(\mathbf{S} = s, \Xi(\mathbf{S}) = n)\,\frac{1}{n\mathbb{P}(|\tau| = n)} \\
&= \frac{\prod \nu_i^{n_i}}{\mathbb{P}(|\tau| = n)} \\
&= \mathbb{P}(\tau = T \,||\tau| = n)\,,
\end{aligned}$$

proving the assertion. □

Note that one easy way to generate the multinomial vector $(N_0, N_1, \ldots)$ is using the binomials

$$\begin{aligned}
&N_0 \sim BIN(n, \nu_0) \\
&N_1 \sim BIN(n - N_0, \nu_1/(1-\nu_0)) \\
&N_2 \sim BIN(n - N_0 - N_1, \nu_2/(1-\nu_0-\nu_1)) \\
&\vdots
\end{aligned}$$

Using this conditional construction, the vector $(N_0, N_1, \ldots)$ has the desired multinomial distribution.

We can modify Algorithm 4 to make it faster. 12345

Using this two results, we can relax step 4 of Algorithm 4. Fix the number of initial individuals $n_1 \in \mathbb{N}$. We find n close enough to n_1, and generate an *approximated* CGW(n) tree. Let $\varepsilon > 0$ be the error term allowed in the simulations. Algorithm 5 generates an *almost* CGW tree.

Algorithm 5 Generate an approximated GW tree conditioned to have size n

Input: A distribution ν, a natural number n_1 and an error term ε.

Output: A tree with law $\mathbb{P}^n_\nu$, where $|n - n_1|/n_1 < \varepsilon$.

1: Generate a multinomial vector $(N_0, N_1, \ldots)$ with parameters $(n_1, \nu_0, \nu_1, \ldots)$.
2: Let K be the last non-zero component of $(N_0, N_1, \ldots)$, that is $N_j = 0$ for $j > K$.
3: Define $n = 1 + \sum iN_i$.
4: **if** $|n - n_1|/n_1 < \varepsilon$ **then**
5:     go to step 9
6: **else**
7:     repeat from step 1.
8: **end if**
9: Redefine $N'_0 = n - \sum_1^K N_i$.
10: Apply Algorithm 3 to the degree sequence $(N'_0, N_1, \ldots, N_K)$.

Remark 4.5. *Note that the number of individuals n in the tree is random. But the only difference between the two sequences obtained in Algorithm 5 is that we add some few leaves to obtain a desired degree sequence. By this we mean that we generate an approximated CGW tree.*

Note that Algorithm 5 gives us a degree sequence of size n, since

$$N_0' + \sum_1^K N_i = n - \sum_1^K N_i + \sum_1^K N_i = n \quad \text{and} \quad 1 + \sum iN_i = n.$$

From our Lemma 4.8 (and Lemma 11 in [BM14b]), we know that the empirical degree sequence $(\hat{N}_i; i \geq 0)$ of trees with law $\mathbb{P}_\nu^{n_1}$, rescaled by n_1, converges to ν. Let $\varepsilon_1 > 0$ and consider a vector $M^{n_1} = (N_0^{n_1}, N_1^{n_1}, \ldots,)$ with multinomial distribution having parameters n_1 and $\nu_0, \nu_1, \ldots$. By the Glivenko-Cantelli theorem (see Lemma 5.5 in the Appendix), we have the uniform convergence

$$\left| \frac{N_i^{n_1}}{n_1} - \nu_i \right| < \varepsilon_1$$

for every n_1 big enough. It is not difficult to prove that for every $n \geq M$ for some $M \in \mathbb{N}$, and $i \geq 1$

$$\frac{|N_0^n - (N_0^n)'|}{n_1} = \frac{|n - n_1|}{n_1} < \varepsilon \quad \text{and} \quad \frac{\nu_i - \varepsilon_1}{1+\varepsilon} \leq \frac{N_i^n}{n_1(1+\varepsilon)} < \frac{N_i^n}{n} < \frac{N_i^n}{n_1(1-\varepsilon)} \leq \frac{\nu_i + \varepsilon_1}{1-\varepsilon}.$$

From the last inequalities we get

$$\frac{\nu_i}{1+\varepsilon} \leq \liminf \frac{N_i^n}{n} \leq \limsup \frac{N_i^n}{n} \leq \frac{\nu_i}{1-\varepsilon}.$$

4.6.3 Constrained simulation of random forests in the unidimensional case

Now we give a way to simulate uniformly sampled forests with a given degree distribution, this is Algorithm 6, and was proved in Lemma 4.5.

Algorithm 6 Generate uniformly sampled forests with a given degree sequence

Input: A degree sequence $\mathbf{S} = (N_i; i \geq 0)$ with $\sum N_i = m + \sum iN_i = s$.

Output: A uniformly sampled forest with m trees having the given degree sequence.

Define the vector $\mathbf{c} = (c_1, c_2, \ldots, c_s)$, with N_0 zeros, N_1 ones, etc.

Set $\pi = (\pi_1, \ldots, \pi_s)$ a uniform random permutation of $[s]$.

For $j \in [s]$ define the walk

$$W^b(j) = \sum_1^j (\mathbf{c} \circ \pi_i - 1),$$

satisfying $W^b(0) = 0$ and $W^b(s) = -m$.

Let $i^* = \min\{j \in [s] : W^b(j) = \min_{1 \leq i \leq s} W^b(i)\}$ be the first time the partial sums reaches its minimum.

Define an independent uniform variable U on $[m] - 1$, and $\tau_U = \min\{j : W^b(j) = W^b(i^*) + U\}$.

Define the process $V(W^b, U)$ of length s whose jth term is $\mathbf{c} \circ \pi_{\tau_U + j} - 1$ with $\tau_U + j \mod s$.

Generate the forest with breadth-first walk $V(W^b, U)$.

Algorithm 7 Generate uniformly sampled multitype forests with a given degree sequence

Input: A degree sequence $\mathbf{S}_{i,j} = (N_{i,j}(k); k \in [m_{i,j}])$ satisfying $n_i = \sum_k N_{i,j}(k)$ for every j, and $n_j = r_j + \sum_i \sum_k kN_{i,j}(k)$, for every j.

Output: A uniformly sampled multitype tree with the given degree sequence.

1: Generate the vectors $\mathbf{c}_{i,j} = (c_{i,j}(1), c_{i,j}(2), \ldots, c_{i,j}(n_i))$, with $N_{i,j}(0)$ zeros, $N_{i,j}(1)$ ones, etc., ordered in non-decreasing order $c_{i,j}(k) \leq c_{i,j}(k+1)$.

2: Generate $\pi_{i,j} = (\pi_{i,j}(1), \ldots, \pi_{i,j}(n_i))$, a uniform random permutation of $[n_i]$, everything independent.

3: Define $\mathbf{W}^b = (W^b_{i,j}; i,j \in [d])$, where

$$W^b_{i,j}(k) = \sum_{l=1}^{k} (c_{i,j} \circ \pi_{i,j}(l) - I(i=j)),\ k \in [n_i]$$

satisfying $W^b_{i,i}(0) = 0$ and $W^b_{i,i}(n_i) = -k_i$.

4: Generate an independent uniform random variable U on $[det(k_{i,j})]$, where $k_{i,j} := \sum kN_{i,j}(k) - n_i \mathbf{1}_{i=j}$.

5: Construct the multidimensional Vervaat transform $V(W^b, U)$ of W^b.

6: Generate the multitype forest with breadth-first walk $V(W^b, U)$.

4.6.4 Constrained simulation of random forests in the multidimensional case

Now we propose Algorithm 7, using the multidimensional Vervaat transform as defined in page 103. This algorithm is precisely Theorem 4.1.

4.6.5 Simulation of MCGW($n_1, \ldots, n_d$) forests with given type sizes

For fixed $\mathbf{n} = (n_1, \ldots, n_d)$, we consider the simulation of multitype GW forests conditioned to have n_i individuals of type i (MCGW($n_1, \ldots, n_d$) forests for short). Using Devroye's idea of Algorithm 4 we propose Algorithm 8. We denote by $\mathbb{P}_\nu(\cdot|\#_j\mathscr{F} = n_j, \forall j)$ the law of a MCGW($n_1, \ldots, n_d$), and by $\nu_{i,j}$ the jth component of the distribution ν_i.

The following proposition, stated on the introduction as Proposition 4.4, proves that from Algorithm 8 we construct a MCGW($n_1, \ldots, n_d$).

Proposition 4.5. *Let W be the breadth-first walk of a MCGW($n_1, \ldots, n_d$) forest satisfying the Hypotheses of Theorem 4.2, having offspring distribution ν, and root-type $\mathbf{r}$ with $1 \leq r_i < n_i$ for every i. Generate independent multinomial vectors $\mathbf{S}_{i,j} = (N_{i,j}(0), N_{i,j}(1), \ldots)$ with parameters $(n_i, \nu_{i,j}(0), \nu_{i,j}(1), \ldots)$, and stop the first time $K = \inf\{k : \Xi(\mathbf{S}_{i,j}, i \in [d]) = n_j \text{ for every } j\}$. Denote by $\mathbf{S}_{(K)}$ the multitype degree sequence obtained, and let $V(\mathbf{W}^b, U)$ be the breadth-first walk generated by Algorithm 7 using the degree sequence $\mathbf{S}_{(K)}$. Then,*

$$\mathbb{P}\Big(V(\mathbf{W}^b, U) = w\Big) = \frac{1}{\frac{n}{r}\frac{det(k_{i,j})}{\prod n_i}} \mathbb{P}_{\mathbf{r}}\left(\mathscr{F} = F | \#_j\mathscr{F} = n_j, \forall\, j\right),$$

for every multitype forest F with root-type $\mathbf{r}$ and individuals-type $\mathbf{n}$, coded by w and with $k_{i,j} = \sum kn_{i,j}(k) - n_i\mathbf{1}_{i=j}$.

Algorithm 8 Generate a MGW forest $\mathscr{F}$ conditioned to have n_i individuals of type i

Input: A distribution ν, and natural numbers $n_i > r_i \geq 1$, for $i \in [d]$.
Output: A multitype forest with law $\mathbb{P}_\nu^{\mathbf{n}}$.

1: Generate independent multinomial vectors $\mathbf{S}_{i,j} = (N_{i,j}(0), N_{i,j}(1), \dots)$ with parameters $(n_i, \nu_{i,j}(0), \nu_{i,j}(1), \dots)$.
2: Let $K_{i,j}$ be the last non-zero component of $\mathbf{S}_{i,j}$, that is $N_{i,j}(K_{i,j}) > 0$ and $N_{i,j}(j) = 0$ for $j > K_{i,j}$.
3: Define $\Xi_j := \Xi(\mathbf{S}_{i,j}, i \in [d]) = r_j + \sum_i \sum_k k N_{i,j}(k)$ for every j.
4: **if** $\Xi_j = n_j$ for every j **then**
5:    go to step 9
6: **else**
7:    repeat from step 1.
8: **end if**
9: Apply Algorithm 7 to the degree sequence $\big((N_{i,j}(0), \dots, N_{i,j}(K_{i,j})); i, j \in [d]\big)$, obtaining a multitype forest $\mathscr{F}_0$ with breadth-first walk distributed as $V(\mathbf{W}^b, U)$.
10: Define $k_{i,j} := \sum k N_{i,j}(k) - n_i \mathbf{1}_{i=j}$.
11: Generate an independent uniform variable V on $[0,1]$.
12: **if** $V \leq \frac{det(-k_{i,j})}{(d+1)^{d-1} \prod n_i}$ **then**
13:    Accept $\mathscr{F} = \mathscr{F}_0$
14: **else**
15:    repeat from step 1.
16: **end if**

Proof. We follow the same lines as in Lemma 4.11. Fix any d-type forest F, having r_i roots and n_i vertices of type i, and degree sequence $\mathbf{s} = (n_{i,j}, i, j \in [d])$. Using the same notation as in Theorem 4.1, let w^b be a multidimensional bridge in $\mathbb{B}_{\mathbf{S},\mathbf{r}}$, having multidimensional Vervaat transform $w = V(w^b, u)$ for some $u \in [det(k_{i,j})]$, where $k_{i,j} := \sum k n_{i,j}(k) - n_i \mathbf{1}_{i=j}$. Using that W^b has exchangeable increments, that U is independent and uniform, and that there are $\prod n_i$ pairs $(\theta_{\mathbf{q},\mathbf{n}}(w), u)$ that can be mapped to w (as seen on page 109), then

$$\begin{aligned}
\mathbb{P}\Big(V(\mathbf{W}^b, U) = w\Big) &= \prod_i n_i \, \mathbb{P}\Big(\mathbf{W}^b = w^b, U = u\Big) \\
&= \frac{1}{\frac{det(k_{i,j})}{\prod n_i} \prod_i \prod_j \binom{n_i}{n_{i,j}}} \mathbb{P}\big(\mathbf{S}_{(K)} = \mathbf{s}\big) \\
&= \frac{1}{\frac{det(k_{i,j})}{\prod n_i} \prod \prod \binom{n_i}{n_{i,j}}} \frac{\mathbb{P}(\mathbf{S} = \mathbf{s})}{\mathbb{P}\big(\Xi(\mathbf{S}_{i,j}, i \in [d]) = n_j, \ \forall\, j\big)},
\end{aligned}$$

where $\mathbf{S}$ has the same distribution as $\mathbf{S}_{(1)}$. We compute explicitly the last fraction of the above equation. For the term $\mathbb{P}(\mathbf{S} = \mathbf{s})$ we use the definition of the multinomial distribution

$$\mathbb{P}(\mathbf{S} = \mathbf{s}) = \prod_i \prod_j \binom{n_i}{n_{i,j}} \prod_{l \geq 0} \nu_{i,j}^{n_{i,j}(l)}(l).$$

For the denominator we have

$$\begin{aligned}
\mathbb{P}\big(\Xi(\mathbf{S}_{i,j}, i \in [d]) = n_j,\ \forall\ j\big) &= \sum_{\substack{\mathbf{s}=(n_{i,j}):\\ \sum_i \sum_k k n_{i,j}(k) = n_j - r_j, \forall\, j \\ \sum_k n_{i,j}(k) = n_i, \forall\, i}} \mathbb{P}(\mathbf{S}=\mathbf{s}) \\
&= \sum_{\substack{\mathbf{s}=(n_{i,j}):\\ \sum_i \sum_k k n_{i,j}(k) = n_j - r_j, \forall\, j \\ \sum_k n_{i,j}(k) = n_i, \forall\, i}} \prod_i \prod_j \binom{n_i}{n_{i,j}} \prod_{l \geq 0} \nu_{i,j}^{n_{i,j}(l)}(l).
\end{aligned}$$

On the other hand, note that for fixed j, using the formula for the convolution,

$$\begin{aligned}
\mathbb{P}\left(\sum_{k=1}^{d} S_{n_k}^{k,j} = n_j - r_j\right) &= \sum_{\sum_{k=1}^{d} \sum_{l=1}^{n_k} i_{k,l} = n_j - r_j} \prod_{k=1}^{d} \prod_{l=1}^{n_k} \nu_{k,j}(i_{k,l}) \\
&= \sum_{\substack{\sum_i \sum_k k n_{i,j}(k) = n_j - r_j, \\ \sum_k n_{i,j}(k) = n_i, \forall\, i}} \prod_i \binom{n_i}{n_{i,j}} \prod_{l \geq 0} \nu_{i,j}^{n_{i,j}(l)}(l),
\end{aligned}$$

where in the last equality, we used the fact that $\prod_i \binom{n_i}{n_{i,j}}$ is the number of different bridges having the same degree sequence $(n_{1,j}(0), n_{1,j}(1), \ldots), \ldots, (n_{d,j}(0), n_{d,j}(1), \ldots)$. Note that the above sum only depends on the sequences $(n_{i,j}, i \in [d])$. Thus, multiplying for all j we have

$$\mathbb{P}\big(\Xi(\mathbf{S}_{i,j}, i \in [d]) = n_j,\ \forall\ j\big) = \prod_j \mathbb{P}\left(\sum_k S_{n_k}^{k,j} = n_j - r_j\right).$$

Therefore, using Theorem 4.2

$$\begin{aligned}
\mathbb{P}\Big(V(\mathbf{W}^b, U) = w\Big) &= \frac{1}{\frac{det(k_{i,j})}{\prod n_i}} \frac{\prod_i \prod_j \prod_{l \geq 0} \nu_{i,j}^{n_{i,j}(l)}(l)}{\prod_j \mathbb{P}\Big(\sum_k S_{n_k}^{k,j} = n_j - r_j\Big)} \\
&= \frac{1}{\frac{det(k_{i,j})}{\prod n_i}} \frac{\mathbb{P}_{\mathbf{r}}(\mathbf{W} = w)}{\frac{n}{r} \mathbb{P}_{\mathbf{r}}\big(O_j = n_j, \forall\ j\big)} \\
&= \frac{1}{\frac{n}{r} \frac{det(k_{i,j})}{\prod n_i}} \mathbb{P}_{\mathbf{r}}\big(\mathbf{W} = w | O_j = n_j, \forall\ j\big),
\end{aligned}$$

with $n = \sum n_i$ and $r = \sum r_i$. We remark that $\mathbb{P}_{\mathbf{r}}\big(\mathbf{W} = w | O_j = n_j, \forall\ j\big) = \mathbb{P}_{\mathbf{r}}\big(\mathscr{F} = F | \#_j \mathscr{F} = n_j, \forall\ j\big)$ is the law of the MGW forest conditioned by its sizes.

From Algorithm 8, the first 9 steps are used to obtain a forest $\mathscr{F}_0$ with law $V(\mathbf{W}^b, U)$. The remaining steps are a usual Accept-Reject method to obtain a sample from the law of the conditioned MGW forest. For each multitype forest with root-type $\mathbf{r}$ and individuals type $\mathbf{n}$ coded by w, define

$$c_w = \frac{n}{r} \frac{det(k_{i,j}(w))}{\prod n_i} = \frac{\mathbb{P}_{\mathbf{r}}\big(\mathbf{W} = w | O_j = n_j, \forall\ j\big)}{\mathbb{P}\big(V(\mathbf{W}^b, U) = w\big)}.$$

Recall Definition 4.6 and Lemma 4.7. For $i \neq j$, since $k_{i,j} \leq n_j$ because the maximum number of type j descendants that any type i can have is n_j, then

$$\det(-k_{i,j}) \leq \sum_{(i_1,\ldots,i_d)} \prod_{j=1}^{d} n_j = (d+1)^{d-1} \prod_{j=1}^{d} n_j,$$

where the last inequality is true by the following bijection. We define a function between the set of elementary forests on d types and labeled trees on $[d+1]$ vertices having root with label $d+1$. Regard an elementary forest F on d types as a unitype tree on $d+1$ vertices by adding a root with label $d+1$ having children the roots of F, and assigning label i to the type i vertex (cf. the paragraph before Lemma 4.5 in [CL16], the remark after Proposition 7 in [BM14a]). This implies that the number of elementary forests on d types is $(d+1)^{d-1}$ by Cayley's formula.

The previous paragraph gives us the bound $c_w \leq \frac{n}{r}(d+1)^{d-1} =: c$. Thus the Accept-Reject method (see [Law13, Section 8.2.4]) applies whenever the uniform V satisfies

$$V \leq \frac{\mathbb{P}_{\mathbf{r}}\left(\mathbf{W} = w | O_j = n_j, \forall\ j\right)}{c\mathbb{P}\big(V(\mathbf{W}^b, U) = w\big)} = \frac{c_w}{c} = \frac{det(-k_{i,j})}{(d+1)^{d-1}\prod n_i} \leq 1.$$

This concludes the proof. $\square$

Chapter 5

DINI DERIVATIVES FOR EXCHANGEABLE INCREMENT PROCESSES AND APPLICATIONS

Let X be an exchangeable increment (EI) process whose sample paths are of infinite variation. We prove that, for any fixed t almost surely,

$$\limsup_{h\to 0\pm} (X_{t+h} - X_t)/h = \infty \quad \text{and} \quad \liminf_{h\to 0\pm} (X_{t+h} - X_t)/h = -\infty.$$

This extends a celebrated result of Rogozin for Lévy processes obtained in [Rog68], and completes the known picture for finite-variation EI processes. Applications are numerous. For example, we deduce that both half-lines $(-\infty, 0)$ and $(0, \infty)$ are visited immediately for infinite variation EI processes (called upward and downward regularity). We also generalize the zero-one law of Millar for Lévy processes by showing continuity of X when it reaches its minimum in the infinite variation EI case (cf. [Mil77]); an analogous result for all EI processes links right and left continuity at the minimum with upward and downward regularity. We also consider results of Durrett, Iglehart and Miller on the weak convergence of conditioned Brownian bridges to the normalized Brownian excursion considered in [DIM77] and broadened to a subclass of Lévy processes and EI processes in [UB14, CUB15]. We prove it here for all infinite variation EI processes. We furthermore obtain a description of the convex minorant known for Lévy processes found in [PUB12] and extend it to non-piecewise linear EI processes. Our main tool to study the Dini derivatives is a change of measure for EI processes which extends the Esscher transform for Lévy processes.

5.1 Statement of the results

Undoubtedly, Lévy Processes are one of the most studied classes of stochastic processes. A less known class which contains them is that of Exchangeable Increment (EI) processes considered in general by Kallenberg in [Kal73].

Definition 5.1. *A continuous time càdlàg $\mathbb{R}$-valued stochastic process $X = (X_t, t \in [0,1])$ has* exchangeable increments *if for every $n \geq 1$, the random variables*

$$X_{1/n}, X_{2/n} - X_{1/n}, \ldots, X_1 - X_{(n-1)/n}$$

are exchangeable.

Clearly, all Lévy Processes are EI since iid random variables are exchangeable. Therefore, one can inherit results for Lévy processes from their counterparts for EI processes, as we illustrate in this chapter. However, conditioning a Lévy process X by its final value (to obtain the so called Lévy bridges as in [CUB11] and [UB14]) or considering $(X_t - tX_1, t \leq 1)$ also yield non-Lévy processes, so that our results can be applied more broadly.

Also, the analysis of EI processes is sometimes aided by simple combinatorial considerations. Indeed, for random walks, the combinatorial considerations of [Spi56] lead to a more thorough understanding of the Fluctuation Theory (study of extremes) of random walks and Lévy processes, and in particular of the celebrated arcsine law for symmetric random walks and Lévy processes; it also reobtains the following formula of [Kac54]

$$\mathbb{E}\left(\max_{0\leq k\leq n} X_{k/n}\right) = \sum_{k=1}^{n} \frac{1}{k}\mathbb{E}\left(X_{k/n}^{+}\right). \tag{5.1}$$

More recently, [AP11] introduced a bijection on permutations which ultimately lead to a description of the convex minorant of a (discrete time) EI process and reinterprets the fluctuation theory of random walks. The Kac-Spitzer identity just displayed is interpreted as the equality in law

$$\max_{0\leq k\leq n} X_{k/n} \stackrel{d}{=} \sum_{i=1}^{K_n} \left[X_{S_i/n} - X_{S_{i-1}/n}\right]^{+} \tag{5.2}$$

where $0 = S_0 < S_1 < \cdots < S_{K_n} = n$ is the partition obtained from a uniform stick breaking process on $\{1, \ldots n\}$ independent of X. The link with the typical fluctuation theory (of random walks and Lévy processes) comes from considering a random n independent of X and geometrically distributed. The partition is then seen to arise from a Poisson point process and the right hand side becomes a compound Poisson distribution in the random walk or Lévy process case; cf. Theorem 4 in [AP11]. The description of the convex minorant for discrete time EI processes is used here to prove an analogous theorem for continuous time EI processes. The multidimensional case is much less studied, but the combinatorial lemma of [BNB63], from which one obtains the expected characteristics of the convex hull of (2D) random walks (like perimeter length or area) has been extended in various directions (and dimensions!) including [RFW17, KVZ17a, KVZ17b, VZ18]. Still in the realm of fluctuation theory, [Ber93] constructs (one-dimensional) random walks conditioned to stay positive through a bijection on permutations; this result is used here to study continuity of an EI process when it reaches its minimum. Away from random walks, (discrete time) EI processes of a particular type are associated to trees (with a given degree distribution) in [BM14b] and combinatorial considerations give information on this probabilistic model.

Kallenberg obtained in [Kal73] the following representation of EI processes X: there exist random variables α, $\beta = (\beta_i, i \in \mathbb{N})$, and $\sigma \geq 0$ which are independent of an iid sequence of uniform random variables $(U_i, i \geq 1)$, and of a Brownian bridge b, such that

$$X_t = \alpha t + \sigma b_t + \sum_{i\geq 1} \beta_i \left[\mathbf{1}_{U_i \leq t} - t\right].$$

When α, β and σ are deterministic, the EI process X is termed *extremal*. All EI processes are therefore mixtures of extremal EI processes and we say that X has canonical parameters (α, σ, β).

Remark 5.1. *Our results are stated for extremal processes. They can be generalized by conditioning on the parameters, on the set where these satisfy the given hypotheses.*

The sample paths of an extremal EI process X are of infinite variation if and only if

Infinite variation either $\sigma > 0$ or $\sum |\beta_i| = \infty$.

Our first result is the following:

Theorem 5.1. *Let X be an extremal EI process of infinite variation. Then, for any fixed t almost surely,*

$$\limsup_{h\to 0} \frac{X_{t+h} - X_t}{h} = \infty \quad \textit{and} \quad \liminf_{h\to 0} \frac{X_{t+h} - X_t}{h} = -\infty \tag{5.3}$$

both from the left and from the right.

Reversibility for EI processes (the fact that $(X_1 - X_{(1-t)-}, t \leq 1)$ has the same law as X) implies that it is enough to handle the above theorem for the right-hand derivatives. By exchangeability, it is enough to consider $t = 0$. We define

$$\overline{D}(X) = \limsup_{h\to 0+} \frac{X_h}{h} \quad \text{and} \quad \underline{D}(X) = \liminf_{h\to 0+} \frac{X_h}{h}.$$

In contrast, for finite variation EI processes X, which satisfy $\sigma = 0$ and $\sum |\beta_i| < \infty$, we can write them as $X_t = \tilde{\alpha} t + \sum_i \beta_i \mathbf{1}_{U_i \leq t}$ where $\tilde{\alpha} = \alpha - \sum_i \beta_i$. Finite variation EI processes therefore are characterized by the parameters $(\tilde{\alpha}, \beta)$. It is well known that $\overline{D}(X) = \underline{D}(X) = \tilde{\alpha}$ almost surely (cf. [Kal05, Cor. 3.30]) in this case.

Theorem 5.1 was proved for Lévy processes in [Rog68] by using an integro-differential equation initially found by Cramér and later recognized and analyzed as a resolvent equation by Watanabe in [Wat71]. Additional proofs, based on the fluctuation theory for Lévy processes, may be found in [Sat99, Ch. 9§47] and [Vig02]. Bertoin proved the limsup statement in the spectrally positive EI case (when $\beta_i \geq 0$ for all i) in [Ber02], based on the results of Fristedt from [Fri72]. Kallenberg takes results further by considering upper envelopes of EI processes in [Kal05] by clever couplings with Lévy processes. These results are nevertheless insufficient to obtain Theorem 5.1. A particular case of the above result is found in [CUB15, Prop. 3.5] under an additional hypothesis on β. Additionally, the same proposition proves Theorem 5.1 whenever $\sigma \neq 0$ using the law of the iterated logarithm for Brownian motion. Hence, we could assume that $\sigma = 0$ in our proofs, but the method is robust enough to handle it. Actually, in the Lévy process setting, our method can also handle general Lévy processes and gives and independent proof of Rogozin's result. This is done in Section 5.2, while Theorem 5.1 is proved in Section 5.3.

Our next application is to show that the zero-one laws for Lévy processes of Millar are actually valid (and therefore a consequence) of the following result (cf. items a and b of [Mil77, Thm. 3.1]) which links behavior when reaching the minimum with behavior at time zero.

Definition 5.2. *An EI process X is said to be* upward regular *if* $\inf\{t \in [0,1] : X_t > 0\} = 0$ *almost surely. The process X is* downward regular *if $-X$ upward regular.*

We consider Knight's result [Kni96] in the extreme setting, about the necessary and sufficient conditions for X to admit a unique minimum:

UM either $\sigma \neq 0$ or $\sum_i \mathbf{1}_{\beta_i \neq 0} = \infty$ or $\sum_i \mathbf{1}_{\beta_i \neq 0} < \infty$ and $\sum_i \beta_i \neq \alpha$.

Theorem 5.2. *Let X be an extremal EI process satisfying UM. Let $\underline{X}_1 = \inf_{s\in[0,1]} X_s$ and let ρ be the unique element of $\{t \in [0,1] : X_t \wedge X_{t-} = \underline{X}_1\}$. Then $X_\rho > \underline{X}_1$ if and only if X is irregular upward and $X_{\rho-} > \underline{X}_1$ if and only if X is irregular downward. In particular, X is continuous at ρ if and only if X is both upward and downward regular and this holds on the set where X has paths of infinite variation.*

Millar actually proves the above result in the Lévy process setting at more general random times and refers to this as the pure behavior of Lévy processes, while noting that it is rather exceptional in the class of Markov processes. Millar also remarks that it is this zero-one law which implies that the conditional law of $X_{\rho+\cdot}$ given $X_{\cdot\wedge\rho}$ depends only on $\underline{X}_1$ and X_ρ. The extension to general random times follows quite easily with Millar's arguments from the stated result above. When X is a Lévy process of finite-variation, necessary and sufficient conditions for regularity have been found by Bertoin in [Ber97] in terms of the Lévy measure. We believe that a similar characterization should be available for EI processes in terms of β. This is left as an **open problem**.

Regularity of half-lines for a Lévy process has many other applications: it helps in obtaining perfectness of the zero set and in constructing a continuous (Markovian) local time (Theorem 6.6 of [Kyp14]); it implies uniqueness for solutions of time-change equations used to construct multitype branching processes (Lemma 6 in [CPGUB17]); regularity of $(-\infty,0)$ has been used when pricing perpetual American put options, as a condition for smooth pasting (see the discussion on Section 1.4.4 in [KL05]).

Our second application concerns the weak limit of an EI process X ending at zero, conditioned on remaining above $-\varepsilon$, as $\varepsilon \to 0$. The limiting process is called the Vervaat transform of X, and is defined as

$$V(X) = X_{\cdot+\rho \mod 1}$$

Theorem 5.3. *Let X be an EI process with $\alpha = 0$ which is both upward and downward regular. Consider $\varepsilon > 0$ and let X^ε have the law of X conditionally on $\underline{X}_1 > -\varepsilon$. Then $X^\varepsilon \xrightarrow{d} V(X)$ as $\varepsilon \to 0$.*

Note that the above theorem always applies to infinite variation EI processes thanks to Theorem 5.1. The above theorem was proved when X is a Brownian bridge from 0 to 0 of length 1 by Durrett, Iglehart and Miller in [DIM77]. The form given above is taken from [CUB15] and is more general, but is actually a simple consequence of the results in that paper. What was lacking in the latter reference is the zero-one law at the minimum of our Theorem 5.2 and, in particular, the fact that all infinite variation EI processes reach their minimum continuously.

Our next application is to extend the description of the convex minorant of a Lévy process of [PUB12] to the EI setting. In the latter reference, it is noted that this description gives another interpretation of a fundamental fact of the fluctuation theory for Lévy processes, namely, the Pecherskii-Rogozin identity of [PR69]. We will consider EI processes which do not have piecewise linear trajectories. By considering the extremal case, this happens if and only if

NPL $\sigma > 0$ or $\sum_i \mathbf{1}_{\beta_i \neq 0} = \infty$.

We will call these processes of the NPL type. The setting of infinite variation EI processes of Theorem 5.1 simplifies the proof.

Definition 5.3. *The* convex minorant *of a càdlàg function $f : [0,1] \to \mathbb{R}$ is the greatest convex function c that is bounded above by f. The* excursion set *is the open set*

$$\mathscr{O} = \{t \in [0,1] : f(t) > c(t)\}.$$

Its maximal components, intervals of the form (g,d), are termed excursion intervals *and they have an associated* length $d-g$, increment $f(d)-f(g)$, slope $(f(d)-f(g))/(d-g)$ *and* excursion $e(t) = f(g+t)-c(g+t)$ *defined for* $t \in [0,d-g]$.

Recall that an upper bounded family of convex functions has a convex supremum, which explains why the convex minorant exists. Let C be the convex minorant of an EI process X of the NPL type. As stated in the next result, its excursion set

$$\mathscr{O} = \{t \in [0,1] : X_t > C_t\}$$

is open and of Lebesgue measure 1. We will consider the following precise ordering of the excursion intervals. Let $(U_i, i \geq 1)$ be an iid sequence of uniform random variables on $[0,1]$ and let (g_1,d_1), (g_2,d_2), ... be the sequence of distinct excursion intervals which are successively discovered by the sequence (U_i). With them, we can define the sequence of lengths, slopes and excursions $\left(e^i\right)$. For another independent sequence $(V_i, i \geq 1)$ of iid uniform random variables on $[0,1]$. We will also consider the partition induced by a stick-breaking scheme based on (V_i): define

$$S_0 = 0, \quad S_{i+1} = S_i + L_i \quad \text{and} \quad L_{i+1} = (1-S_i)V_{i+1}.$$

Then (L_i) is the uniform stick-breaking process and S is the partition of $[0,1]$ induced by its cumulative sums. Note that this is a very sparse partition of $[0,1]$ which we can use to analyze X by considering: $X_{S_i} - X_{S_{i-1}}$ and the sequence of Knight bridges where K^i is the Knight transform of $X - X_{S_{i-1}}$ on $[0, L_i]$. The *Knight transform* of an EI process Y starting at zero on an interval $[0,t]$ and satisfying UM is obtained by first defining the Knight bridge $K_s = Y_s - sY_t/t$, letting ρ be the location of its (unique) minimum to finally define

$$s \mapsto K_{(\rho+s) \bmod t} - K_\rho \wedge K_{\rho-} \ \text{ for } s \in [0,t].$$

Theorem 5.4. *Assume that the EI process X satisfies NPL. Then, its excursion set $\mathscr{O}$ is open and of Lebesgue measure* 1. *Furthermore, the following equality in law holds:*

$$\left(d_i - g_i, X_{d_i} - X_{g_i}, e^i\right)_{i\geq 1} \stackrel{d}{=} \left(L_i, X_{S_i} - X_{S_{i-1}}, K^i\right)_{i\geq 1}.$$

Recent papers have used the above description of the convex minorant (in the Lévy process case) to develop an exact simulation method for the maximum of a stable process (found in [GMU18b]) and an approximate simulation method (albeit very efficient, cf. [GMU18a]) for the maximum of Lévy processes whose one-dimensional distributions can be sampled exactly. This is particularly relevant to Monte Carlo methods for ruin probabilities with finite and deterministic horizon. In [CM15], the classical Cramer-Lundberg ruin process is generalized to an exchangeable increment process on $[0,\infty)$ to relax the independence between claim sizes; these are mixtures of Lévy processes. In contrast to the classical setting, when working under the classical net profit condition, the ruin probability might not converge to zero as the initial capital goes to infinity and a new net profit condition is needed. In the finite-horizon case, we would be dealing with an EI process of the type considered here; Theorem 5.4 would give us access to the ruin probabilities.

We end this section with a few comments on the organization of the chapter. Our main result is Theorem 5.1; all others have a simple proof from Theorem 5.1 and more specialized results from the literature. A brief outline of the proof of Theorem 5.1, which explains the organization of the chapter, is as follows.

Step 1: Assume $\alpha, \sigma = 0$ and $\beta_j \geq 0$ for every j (if $\beta_j \leq 0$, apply this step to $-X$).

- $\overline{D}(X) = \infty$**.** This follows from results of [Fri72].
- $\underline{D}(X) = -\infty$**.** We use an exponential change of measure (which reduces to the well known Esscher transform if X is a Lévy process), with parameters $\theta \in \mathbb{R}$ and $T \in (0,1)$, to deduce that $\underline{D}(X) = \alpha^\theta/T + \underline{D}(X^\theta)$, where X^θ is another EI process, α^θ is a random variable (not independent of X^θ, although for Lévy processes, α^θ is deterministic). A lower bound on the probability that an EI process with positive jumps is non-positive (found in [Sat99] for Lévy processes) implies that $\underline{D}(X^\theta) \leq 0$. It remains to notice that the infinite variation hypothesis gives us $\alpha^\theta \to -\infty$ when $\theta \to -\infty$, thereby proving Theorem 5.1 in this case.

Step 2: Assume $\sigma \neq 0$ or $\sum_j \left|\beta_j\right| = \infty$; also set $\alpha = 0$.

- $\underline{D}(X) = -\infty$**.** We note that the aforementioned lower bound is valid when $\sigma \neq 0$ and that it works along deterministic subsequences, so that $\liminf_{n\to\infty} X_{t_n}/t_n \leq 0$ whenever $t_n \downarrow 0$ and X has only positive jumps and a Brownian component. We then write $X = X^p + X^n$, where X^p and X^n are independent, X^n only has negative jumps and X^p only has positive jumps and contains the Brownian component (if any). If X^p or X^n has finite variation, we use [Kal05, Cor. 3.30] and Step 1. Hence, assume both have infinite variation. We then get a random subsequence $T_k \downarrow 0$ such that $X^p_{T_k}/T_k \to -\infty$ and, using independence, we get $\liminf X^n_{T_k}/T_k \leq 0$. Hence, we obtain $\underline{D}(X) = -\infty$.
- $\overline{D}(X) = \infty$**.** Apply the previous case to $-X$.

The chapter is organized as follows: In Section 5.2 we present a simplified proof following the outline above in the setting of Lévy processes. This is because the exponential change of measure and lower bounds on probabilities discussed above are already known. In Section 5.3, we consider Theorem 5.1 in the case of EI processes. Here, we state and prove the exponential change of measure and lower bounds for probabilities. Finally, Section 5.4 is devoted to the applications of our results, and contains the proofs of Theorems 5.2, 5.3 and 5.4.

5.2 The Lévy process case

We now illustrate the proof of Theorem 5.1 in the case of Lévy processes. This proof is the only one published that does not use fluctuation theory for Lévy processes and can be considered to be simpler. It is based on basic facts on Lévy processes and on the Esscher transform. The reader might consult [Ber96] and [Sat99] for these basic facts, some of which we now recall. In particular, Lévy processes satisfy the Blumenthal 0-1 law and therefore the random variables $\overline{D}(X)$ and $\underline{D}(X)$ are actually constant.

Recall that X can be written as the independent sum of two Lévy processes X^1 and X^2, where X^1 has bounded jumps and X^2 is compound Poisson. Since $\lim_{t\to 0} X^2_t/t$ exists and is finite, we see that it suffices to prove Theorem 5.1 when X has bounded jumps.

Assume then, that the jumps of X are bounded by 1; we can then determine X by its Laplace transform

$$\mathbb{E}\left(e^{\lambda X_t}\right) = e^{t\Psi(\lambda)} \quad \text{where} \quad \Psi(\lambda) = \alpha\lambda + \lambda^2\sigma^2/2 + \int_{[-1,1]} [e^{\lambda x} - 1 - \lambda x]\, \pi(dx)$$

by the Lévy-Khintchine formula. Let X be a Lévy process whose paths have infinite variation; equivalently, we assume that

$$\sigma^2 > 0 \quad \text{or} \quad \int |x|\, \pi(dx) = \infty.$$

The Lévy measure π, which is concentrated on $[-1,1]$, satisfies

$$\int x^2 \pi(dx) < \infty.$$

In other words, the characteristic triplet of X is (α, σ, π).

The following result is a trivial extension of the well-known Esscher change of measure for Lévy processes, as found in [Kyp14]. It will imply that the superior and inferior limits in Theorem 5.1 are not finite. Let $\mathscr{F}_t = \sigma(X_s : s \leq t)$.

Proposition 5.1 (Esscher transform)**.** *Fix $\theta \in \mathbb{R}$. Define the measure $\mathbb{Q}$ by its restriction to $\mathscr{F}_t$ as*

$$\mathbb{Q}|_{\mathscr{F}_t} = e^{\theta X_t - t\Psi(\theta)} \cdot \mathbb{P}|_{\mathscr{F}_t}.$$

Then, under $\mathbb{Q}$, the stochastic process X is a Lévy process whose Laplace exponent is:

$$\Psi^\theta(\lambda) = \Psi(\lambda + \theta) - \Psi(\theta).$$

In particular, the characteristic triplet of X under $\mathbb{Q}$ is $(\alpha_\theta, \sigma, \pi_\theta)$ where:

$$\alpha_\theta = \alpha + \theta\sigma^2 + \int_{[-1,1]} \left[e^{\theta x} - 1\right] x\,\pi(dx) \quad \text{and} \quad \pi_\theta(dx) = \mathbf{1}_{[-1,1]}(x)\, e^{\theta x}\, \pi(dx).$$

Note that $\alpha_\theta \to -\infty$ as $\theta \to -\infty$ when $\int |x|\ \pi(dx) = \infty$ or $\sigma^2 > 0$.

We now specialize to the spectrally positive case and then use a (simple) argument to deduce the general case.

5.2.1 The spectrally positive case

We now focus on the spectrally positive case, which corresponds to the Lévy measure π concentrated on $[0,1]$. When X is spectrally positive, a general result of Fristedt implies that $\limsup_{t\to 0} |X_t|/t = \infty$; cf. the proof of part A of Theorem 1 in [Fri72]. Since X_t/t is a reverse martingale with no negative jumps (when t decreases) which does not converge (because of the preceding phrase), Proposition 7.19 in [Kal02] tells us that $\overline{D}(X) = \infty$.

To prove that $\underline{D}(X) = -\infty$, we use the following result of Sato for spectrally positive Lévy processes.

Lemma 5.1. *If X is a spectrally positive Lévy process with parameters (α, σ, π) with jumps bounded by 1 and $\alpha = \mathbb{E}(X_1) \leq 0$ then $\mathbb{P}(X_t \leq 0) \geq 1/16$ for all $t \geq 0$. Also, for any deterministic sequence $t_n \to 0$,*

$$\liminf_{n\to\infty} \frac{X_{t_n}}{t_n} \leq 0.$$

The first statement is Proposition 46.8 in [Sat99]; the proof is simple and based on inequalities of the Paley-Zygmund type (based on exponentials of X) and on properties of the Laplace exponent. However, it should be noted that the theorem cannot hold for all α (consider $\alpha \to \infty$, so that $\mathbb{P}(X_t \leq 0) \to 0$), and that the proof is valid when $\alpha \leq 0$. The second statement is found in the penultimate paragraph of the proof of Theorem 47.1 in [Sat99], and basically follows from the Borel-Cantelli lemma and the first part of Lemma 5.1; however, in the proof, one uses that $\alpha \leq 0$, so that it remains valid. We reprove the lemma in the EI setting using Lemma 5.3 below.

Proof of Theorem 1 for totally asymmetric Lévy processes. As before, we restrict ourselves to the spectrally positive case with jumps bounded by 1.

Note in particular, that Lemma 5.1 implies that $\underline{D}(X) \leq 0$ when $\alpha = 0$. On the other hand, by absolute continuity, we see that $\underline{D}(X)$ takes the same constant value both under $\mathbb{P}$ and under $\mathbb{Q}$. Hence, we can write

$$\underline{D}(X) = \alpha_\theta + \underline{D}(\tilde{X})$$

where $\tilde{X}$ is a spectrally positive Lévy process of characteristics $(0, \sigma, \pi_\theta)$. By the preceding lemma, $\underline{D}(\tilde{X}) \leq 0$. As we remarked, $\alpha_\theta \to -\infty$ as $\theta \to -\infty$, so that $\underline{D}(X) = -\infty$. We deduce that for all $\alpha \in \mathbb{R}$, $\underline{D}(X + \alpha\,\mathrm{Id}) = -\infty$. □

5.2.2 The general case

Let X be a Lévy process of infinite variation and bounded jumps. It suffices to prove $\underline{D}(X) = -\infty$ for any such process and then apply this to $-X$ to conclude that also $\overline{D}(X) = \infty$. Using the Lévy-Itô decomposition, we write $X = X^{\mathrm{neg}} + X^{\mathrm{pos}}$ as the sum of two independent Lévy processes, where X^{neg} is spectrally negative and X^{pos} is spectrally positive; this can be achieved with $\mathbb{E}\big(X_t^{\mathrm{pos}}\big) = 0$ so that Lemma 5.1 applies. In particular, from Theorem 5.1 for totally asymmetric Lévy processes (proved in Subsection 5.2.1), we have $\underline{D}(X^{\mathrm{neg}}) = -\infty$. Hence, there exists a random sequence $V_n \downarrow 0$ such that $X_{V_n}^{\mathrm{neg}} \leq -nV_n$. Since X^{neg} is independent of X^{pos} and the latter is spectrally positive, Lemma 5.1 implies that $\liminf_{n\to\infty} X_{V_n}^{\mathrm{pos}}/V_n \leq 0$. We then conclude that

$$\liminf_{t\to 0} \frac{X_t}{t} \leq \liminf_{n\to\infty} \frac{X_{V_n}}{V_n} \leq \liminf_{n\to\infty} \frac{X_{V_n}^{\mathrm{pos}}}{V_n} - n = -\infty.$$

5.3 Dini derivatives of EI processes in the totally asymmetric case

In this section, we prove Theorem 5.1; note that it suffices to prove it for extremal EI process and obtain the general case by mixing. For concreteness, we assume that X is extremal and only has positive jumps, so that $\beta_i \geq 0$. We first show that Dini derivatives are constant.

Proposition 5.2. *Let X be an extremal EI process with parameters (α, σ, β). Then*

$$\underline{D}(X) = c_1 \quad \text{and} \quad \overline{D}(X) = c_2$$

for some constants $c_1, c_2 \in [-\infty, \infty]$.

Proof. Fix any $k \in \mathbb{N}$, and define $X_t^{(-k)} = \alpha t + \sigma b_t + \sum_{j=k+1}^{\infty} \beta_j \left[\mathbf{1}_{U_j \leq t} - t\right]$. Hence, for any $t < \min\{U_i : i \in [k]\}$ we have

$$\frac{X_t}{t} = \frac{X_t^{(-k)}}{t} - \sum_{1}^{k} \beta_i,$$

which implies

$$\overline{\lim_{t\downarrow 0}} \frac{X_t}{t} = \overline{\lim_{t\downarrow 0}} \frac{X_t^{(-k)}}{t} - \sum_{1}^{k} \beta_i, \tag{5.4}$$

for any k. Let

$$\mathscr{G} = \bigcap_{\varepsilon \in (0,1)} \sigma(b_s : s \leq \varepsilon).$$

The (local) absolute continuity of the Brownian bridge with respect to Brownian motion, and the Blumenthal zero-one law for the latter imply that $\mathscr{G}$ is trivial. Let $\mathscr{F}_k = \sigma(U_k, U_{k+1}, \ldots)$ and note that $\mathscr{G}$ is independent of (any *sigma*-algebra but in particular) $\mathscr{F}_k$. As noted in the proof of [Ber96, Prop. I§2.4], the argument for Kolmogorov's zero-one law tells us that $\bigcap_k \mathscr{G} \vee \mathscr{F}_k$ is trivial. Since the right-hand side of (5.4) is $\mathscr{G} \vee \mathscr{F}_k$-measurable, we deduce that the left-hand side is $\bigcap_k \mathscr{G} \vee \mathscr{F}_k$-measurable and therefore trivial. A similar argument works for the lower Dini derivative. □

We will proceed as in the case of Lévy processes: we first give a change of measure for EI processes, analogous to the Esscher transformation, which has the effect of transforming the drift and the jumps. As in the Lévy case, $\overline{D}(X) = \infty$ follows from simple results on the literature. We then use martingale arguments to prove that $\underline{D}(X) \leq 0$. Finally, our change of measure will imply that $\underline{D}(X) = -\infty$.

Proposition 5.3 (Change of measure). *Let $X = (X_t, t \in [0,1])$ be an extremal EI process with characteristics (α, σ, β), defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Then*

$$\mathbb{E}\left(e^{\theta X_t}\right) = e^{\alpha\theta t + \theta^2\sigma^2 t(1-t)/2} \prod_{j=1}^{\infty} \left[e^{-\theta\beta_j t}[1-t] + e^{\theta\beta_j(1-t)} t\right] < \infty$$

for all $t \in [0,1]$ and $\theta \in \mathbb{R}$.

Fix $T \in (0,1)$, $\theta \in \mathbb{R}$ and let $\mathscr{F}_t = \sigma(X_s : s \leq t)$. Define $\mathbb{Q}$ on $\mathscr{F}_T$ by

$$\mathbb{Q}(A) = \frac{\mathbb{E}\left(e^{\theta X_T} \mathbf{1}_A\right)}{\mathbb{E}\left(e^{\theta X_T}\right)}.$$

Under $\mathbb{Q}$, the stochastic process $(X_t, t \leq T)$ is an EI process whose (random) characteristics $(\alpha^\theta, \sigma, \beta^\theta)$ have the following law. Let (B_j) be independent Bernoulli random variables with parameter p_j given by

$$p_j = \frac{Te^{\theta b_j}}{Te^{\theta b_j} + (1-T)}. \tag{5.5}$$

Then

$$\alpha^\theta = \alpha T + \theta\sigma + \sum_j \beta_j \left[B_j - T\right] \quad \text{and} \quad \beta_j^\theta = \beta_j B_j,$$

where $\sum_j \beta_j \left[B_j - T\right]$ converges almost surely and in L_1.

If $\alpha = 0$ and either $\sigma > 0$ or $\sum_i \mathbf{1}_{\beta_i \neq 0} = \infty$ then $\mathbb{E}\left(e^{\lambda X_t}\right) \to \infty$ as $\lambda \to \infty$ for every $t \in (0,1)$.

Remark 5.2. *When X is a finite variation EI process, driven by the two parameters $(\tilde{\alpha}, \beta)$ (rather than $(\alpha, 0, \beta)$,) as explained in Section 5.1, then X under $\mathbb{Q}$ is also of finite variation and is driven by the two parameters $(\tilde{\alpha}T, \beta^\theta)$. Hence, $\overline{D}(X^\theta) = \underline{D}(X^\theta) = \tilde{\alpha}$, which does not depend on θ; the interpretation is that, in this case, the change of measure is not adding a drift. If we choose not to reparametrize with $(\tilde{\alpha}, \beta^\theta)$, the finite variation case is characterized by the fact that α^θ is bounded in θ. On the other hand, if X is of infinite variation, we shall see that α^θ stochastically increases from $-\infty$ to ∞.*

Proof. Let us begin with the proof that the moment generating function of X is finite. Use the canonical representation $X_t = \alpha t + \sigma b_t + \sum_i \beta_j \left[\mathbf{1}_{U_j \leq t} - t\right]$. Define

$$\phi(t, \theta, \beta) = e^{-\theta\beta t}[1-t] + e^{\theta\beta(1-t)} t.$$

Note that,

$$\mathbb{E}\left(e^{\theta\left[\beta_j \mathbf{1}_{U_j \leq t} - t\right]}\right) = \phi\left(t, \theta, \beta_j\right).$$

For fixed t and θ, we have

$$\phi(t, \theta, \beta) = 1 + \frac{1}{2}\left(\theta^2 t - \theta^2 t^2\right)\beta^2 + O\left(\beta^3\right)$$

as $\beta \to 0$. Therefore, $\prod_{j=1}^{\infty} \phi\left(t, \theta, \beta_j\right)$ exists, since (β_j) is square summable. By Fatou's lemma, we see that

$$\mathbb{E}\left(e^{\theta X_t}\right) \leq e^{\alpha\theta t + \theta^2\sigma^2 t(1-t)} \lim_{n\to\infty} \prod_{j \leq n} \mathbb{E}\left(e^{\theta\left[\beta_j \mathbf{1}_{U_j \leq t} - t\right]}\right) \leq e^{\alpha\theta t + \theta^2\sigma^2 t(1-t)} \prod_{j=1}^{\infty} \phi\left(t, \theta, \beta_j\right) < \infty.$$

But then, Hölder's inequality implies the log-concavity of $\theta \mapsto \mathbb{E}\left(e^{\theta X_t}\right)$, which is then enough to obtain uniform integrability of the sequence $X^n = X - \sum_{j>n} \beta_j \left[\mathbf{1}_{U_j \leq t} - t\right]$. This implies the stated infinite product formula for the moment generating function of X_t.

Consider now a sequence $V = (V_j)$ of independent uniform $(0,1)$ random variables, independent also of b and the (U_j) and define $B_j = \mathbf{1}_{V_j \leq p_j}$. Note that obviously $\sum \beta_j^2 B_j^2 < \infty$. Regarding α^θ, we use the Kolmogorov three series theorem. Indeed, since the sequence (β_j) is bounded, so is the sequence $\left(\beta_j \left[B_j - T\right]\right)$. On the other hand, we have

$$\mathbb{E}\left(\sum_{j=1}^{n} \beta_j \left[B_j - T\right]\right) = \sum_{j=1}^{n} T(1-T)\beta_j \frac{e^{\theta\beta_j} - 1}{Te^{\theta\beta_j} + (1-T)} = O\left(\sum_{j=1}^{\infty} \beta_j^2\right).$$

Finally, we see that

$$\operatorname{Var}\left(\sum_{j=1}^{n} \beta_j \left[B_j - T\right]\right) = \sum_{j=1}^{n} \beta_j^2 p_j (1 - p_j) \leq \sum_{j=1}^{\infty} \beta_j^2.$$

Consider also the process

$$X_t^\theta = \alpha^\theta t/T + \theta\sigma b_{t/T} + \sum_j \beta_j B_j \left[\mathbf{1}_{U_j \leq t/T} - t/T\right]$$

defined on $[0,T]$. Note that X^θ is an EI process on $[0,T]$ with random characteristics $(\alpha^\theta, \sigma^\theta, \beta^\theta)$; we finish the proof by comparing, through moment generating functions, the finite-dimensional distributions of the increments of X under $\mathbb{Q}$ and of X^θ under $\mathbb{P}$.

First of all, by independence of U and b; since the law of b under $e^{\theta b_T} \cdot \mathbb{P}$ equals that of $b + \sigma\theta\,\mathrm{Id}$ (as can be proved through the Gaussian character of b), it suffices to prove the theorem when $\sigma = 0$. Since α is deterministic, it also suffices to consider $\alpha = 0$.

Let $0 = t_0 \leq t_1 \leq \cdots \leq t_n = T$ and $\lambda_1 \cdots, \lambda_n \in \mathbb{R}$. Using similar arguments as for justifying the exchange of expectation and infinite products in the computation of the generating function, we first see that

$$\begin{aligned}
&\mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda_i\left[X_{t_i} - X_{t_{i-1}}\right]} e^{\theta X_T}\right) \\
&= \prod_j \mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda_i\left[\alpha(t_i - t_{i-1}) + \sigma(b_{t_i} - b_{t_{i-1}}) + \beta_j \mathbf{1}_{t_{i-1} \leq U_j \leq t_i}\right]} e^{\theta\left[\alpha T + \sigma b_T + \beta_j \mathbf{1}_{U_j \leq T}\right]}\right).
\end{aligned}$$

Therefore, the Laplace transform of $(X_{t_1},\dots,X_{t_n})$ is finite under $\mathbb{Q}$ and

$$\mathbb{E}_{\mathbb{Q}}\left(\prod_{i=1}^{n} e^{\lambda_i[X_{t_i}-X_{t_{i-1}}]}\right) = \frac{1}{\mathbb{E}\left(e^{\theta X_T}\right)}\mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda_i[X_{t_i}-X_{t_{i-1}}]}e^{\theta X_T}\right)$$
$$= \frac{1}{\mathbb{E}\left(e^{\theta X_T}\right)}\prod_{j=1}^{\infty}\mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda_i\beta_j[\mathbf{1}_{U_j\in[t_{i-1},t_i]}-(t_i-t_{i-1})]}e^{\theta\beta_j[\mathbf{1}_{U_j\in[0,T]}-T]}\right).$$

By considering the interval of the partition $[t_{i-1},t_i]$ on which U_j falls, and recalling the definition of p_j in (5.5), we get

$$\mathbb{E}_{\mathbb{Q}}\left(\prod_{i=1}^{n} e^{\lambda_i[X_{t_i}-X_{t_{i-1}}]}\right)$$
$$= \frac{1}{\mathbb{E}\left(e^{\theta X_T}\right)}\prod_{j=1}^{\infty} e^{-\theta\beta_j T-\beta_j\sum_{i=1}^{n}\lambda_i(t_i-t_{i-1})}\left[(1-T)+\sum_{i=1}^{n}(t_i-t_{i-1})e^{(\lambda_i+\theta)\beta_j}\right]$$
$$= \prod_{j=1}^{\infty} e^{-\beta_j\sum_{i=1}^{n}\lambda_i(t_i-t_{i-1})}\left[(1-p_j)+\sum_{i=1}^{n}\frac{t_i-t_{i-1}}{T}p_j e^{\lambda_i\beta_j}\right].$$

Recall that α and σ are zero in the definition of α^θ and X^θ. On the other hand, using the definition of X^θ, we can use the distributional assumptions on B and U (first independence, then conditioning on B, and finally considering the interval $[t_{i-1},t_i)$ on which U_j falls) to obtain

$$\prod_j \mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda_i\beta_j[B_j-T]\frac{t_i-t_{i-1}}{T}+\lambda_i\beta_j B_j\left[\mathbf{1}_{TU_j\in(t_{i-1},t_i)}-\frac{t_i-t_{i-1}}{T}\right]}\right)$$
$$= \prod_j e^{-\sum_{i=1}^{n}\lambda_i\beta_j(t_i-t_{i-1})}\mathbb{E}\left(\left[(1-p_j)+p_j e^{\sum_{i=1}^{n}\lambda_i\beta_j\mathbf{1}_{TU_j\in(t_{i-1},t_i)}}\right]\right)$$
$$= \prod_j e^{-\sum_{i=1}^{n}\lambda_i\beta_j(t_i-t_{i-1})}\left[(1-p_j)+p_j\sum_{i=1}^{n} e^{\lambda_i\beta_j}\frac{t_i-t_{i-1}}{T}\right]$$

The preceding equation shows that the increments of the left-hand side have the same law as the corresponding increments of $\beta_j[\mathbf{1}_{U_j\le\cdot}-\cdot]$ under $\mathbb{Q}$, for every j. Thus,

$$\prod_j \mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda_i\beta_j[B_j-T]\frac{t_i-t_{i-1}}{T}+\lambda_i\beta_j B_j\left[\mathbf{1}_{TU_j\in(t_{i-1},t_i)}-\frac{t_i-t_{i-1}}{T}\right]}\right) = \mathbb{E}_{\mathbb{Q}}\left(\prod_{i=1}^{n} e^{\lambda_i[X_{t_i}-X_{t_{i-1}}]}\right) < \infty.$$

Similarly as in the proof of Proposition 5.3, using Fatou's lemma and Hölder's inequality we deduce

$$\mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda_i[X^\theta_{t_i}-X^\theta_{t_{i-1}}]}\right) = \prod_j \mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda_i\beta_j[B_j-T]\frac{t_i-t_{i-1}}{T}+\lambda_i\beta_j B_j\left[\mathbf{1}_{TU_j\in(t_{i-1},t_i)}-\frac{t_i-t_{i-1}}{T}\right]}\right).$$

Hence the finite-dimensional distributions of X under $\mathbb{Q}$, and X^θ under $\mathbb{P}$ are the same.

The last part of the statement follows from the fact that if ξ is any random variable on $\mathbb{R}$ with finite generating function g, then $g(\infty)=\infty$ whenever $\mathbb{P}(\xi>0)>0$. The hypotheses on X are chosen so that

$\mathbb{P}(X_t > 0) > 0$ for all $t \in (0,1)$. Indeed, when $\alpha = 0$, $(X_t/(1-t), t < 1)$ is a martingale; the assumption $\mathbb{P}(X_t > 0) = 0$ implies $\mathbb{E}(X_t^-) = 0$ which then gives $\mathbb{P}(X_t = 0) = 1$ and our hypotheses imply that X_t has no atoms for $t \in (0,1)$ as shown in the proof of Lemma 1.2 in [Kni96]. $\square$

We now consider the behavior of the drift α^θ as a function of θ.

Proposition 5.4. *The mapping $\theta \mapsto \alpha^\theta$ is stochastically non-decreasing. Also, $\alpha^\theta \in L_1$, $\theta \mapsto \mathbb{E}(\alpha^\theta)$ is continuous and strictly increasing and, if X is of infinite variation, $\mathbb{E}(\alpha^\theta) \to \pm\infty$ as $\theta \to \pm\infty$.*

Proof. We have already proved that α^θ is a convergent series (plus the couple of constants α and $\theta\sigma$); it is absolutely divergent in the infinite variation case and otherwise absolutely convergent. Using our explicit construction of the random variables B_j as $\mathbf{1}_{V_j \le p_j}$ and the definition of p_j in (5.5), we note that the $\beta_j[B_j - T]$ are increasing in θ, which implies the same for α^θ.

Recall that α^θ is (modulo a constant) a series of independent random variables taking two values, whose means and variances are summable. Hence $\alpha^\theta \in L_1$ and

$$\mathbb{E}\left(\alpha^\theta\right) = \alpha + \theta\sigma + T(1-T)\sum_j \beta_j \frac{e^{\theta\beta_j} - 1}{Te^{\theta\beta_j} + (1-T)}.$$

The above summands are $O\left(\theta\beta_j^2\right)$, uniformly for θ on compact sets. This implies the continuity of $\theta \mapsto \mathbb{E}(\alpha^\theta)$. But the mapping

$$\theta \mapsto \beta_j \frac{e^{\theta\beta_j} - 1}{Te^{\theta\beta_j} + (1-T)}$$

is strictly increasing, and monotone convergence implies the same for $\theta \mapsto \mathbb{E}(\alpha^\theta)$. Finally, note that the preceding function of θ goes to $\pm\beta_j$ as $\theta \to \pm\infty$. When X is of infinite variation, Fatou's lemma can be applied to the series for $\mathbb{E}(\alpha^\theta)$, as the summands in its definition are either all positive or all negative, and conclude that $\mathbb{E}(\alpha^\theta) \to \pm\infty$ as $\theta \to \pm\infty$. $\square$

We now give a version of Lemma 5.1 for EI processes, as well as a simple lemma which uses it.

Lemma 5.2. *Let X be an extremal EI process with parameters $(0, \sigma, \beta)$ such that $\beta_i \le 0$ for all i. Then $\mathbb{P}(X_t \ge 0) \ge 1/16$ for every $t \le 1/2$.*

Proof. Assume we have proved that

$$\mathbb{E}\left(e^{2\lambda X_t}\right) \le \mathbb{E}\left(e^{\lambda X_t}\right)^4 \tag{5.6}$$

for $\lambda > 0$ and $t \le 1/2$. Then, using the Cauchy-Schwarz inequality we would obtain for $\lambda > 0$

$$\begin{aligned}
\mathbb{P}(X_t \ge 0) &\ge \frac{\left(\mathbb{E}\left(e^{\lambda X_t}\right) - \mathbb{E}\left(e^{\lambda X_t}\mathbf{1}_{X_t \le 0}\right)\right)^2}{\mathbb{E}\left(e^{2\lambda X_t}\right)} \\
&\ge \frac{\left(\mathbb{E}\left(e^{\lambda X_t}\right) - 1\right)^2}{\mathbb{E}\left(e^{\lambda X_t}\right)^4}
\end{aligned}$$

By Proposition 5.3 we can chose λ_t such that $\mathbb{E}\left(e^{\lambda_t X_t}\right) = 2$, which implies

$$\mathbb{P}(X_t \geq 0) \geq \frac{1}{16}.$$

Now, let us prove (5.6). First note that (5.6) is an equality for a Brownian bridge. Hence, by the independence of the latter with the (purely discontinuous) jump part of X, it is enough to assume $\sigma = 0$. Defining

$$\phi_j(\lambda) := \ln \mathbb{E}\left(e^{\lambda \beta_j [\mathbf{1}_{U_j \leq t} - t]}\right) := \ln \psi_j(\lambda),$$

it is enough to prove for every $j \in \mathbb{N}$ that

$$0 \leq 4\phi_j(\lambda) - \phi_j(2\lambda) \qquad \lambda \geq 0. \tag{5.7}$$

Since the right-hand side is zero when $\lambda = 0$, proving it has a non-negative derivative implies Equation (5.7). Taking the derivative with respect to λ, we need to prove that

$$\begin{aligned} 0 \leq & 4\frac{\beta_j(1-t)te^{\lambda\beta_j(1-t)} + \beta_j(-t)(1-t)e^{\lambda\beta_j(-t)}}{\psi_j(\lambda)} \\ & - 2\frac{\beta_j(1-t)te^{2\lambda\beta_j(1-t)} + \beta_j(-t)(1-t)e^{2\lambda\beta_j(-t)}}{\psi_j(2\lambda)} \end{aligned}$$

which is equivalent to

$$0 \geq 2\frac{e^{\lambda\beta_j} - 1}{te^{\lambda\beta_j} + 1 - t} - \frac{e^{2\lambda\beta_j} - 1}{te^{2\lambda\beta_j} + 1 - t}.$$

and further equivalent, since the denominators are positive by convexity of the exponential function, to

$$1 - t + (t-2)e^{\lambda\beta_j} + (1+t)e^{2\lambda\beta_j} - te^{3\lambda\beta_j} \geq 0$$

As before, the left-hand side at $\lambda = 0$ is zero, hence, it suffices to prove its derivative is non-negative. We apply an analogous reasoning by evaluation at $\lambda = 0$, differentiation and division by $\beta_j e^{\lambda\beta_j}$ (which is negative) three times! The sequence of derivatives, taking out the factor $\beta_j e^{\lambda\beta_j}$ are

$$\begin{aligned} & t - 2 + 2(1+t)e^{\lambda\beta_j} - 3te^{2\lambda\beta_j}, \\ & 2(1+t) - 6te^{\lambda\beta_j} \text{ and} \\ & -6te^{\lambda\beta_j}. \end{aligned}$$

The penultimate function is non-negative at $\lambda = 0$ when $t \in [0, 1/2]$. The last row then shows that the penultimate one is non-negative, which we can then bootstrap to show inequality (5.7). □

The choose of λ_t such that $f(\lambda_t) = 2$ in the preceding proof seems arbitrary; the reader can check it gives the best bound obtainable by this method.

Lemma 5.3. *Let $X = (X_t, t \geq 0)$ be a càdlàg process such that, for some sequence $t_n \downarrow 0$, the random variable $\liminf_n X_{t_n}/t_n$ is constant. Assume that, for some $\varepsilon, c > 0$, we have $\mathbb{P}(X_t \leq 0) > c$ for every $t \in [0, \varepsilon]$. Then, $\liminf_n X_{t_n}/t_n \leq 0$.*

Remark 5.3. *Note that the above can be applied when the augmented initial σ-algebra of X, given by $\cap_{\varepsilon>0}\sigma(X_s : s \le \varepsilon)$, is trivial (hence for Feller processes) or in the case of extremal EI processes by mimicking the proof of Proposition 5.2; in this case, we can apply the result to any sequence $t_n \downarrow 0$.*

Also, by mixing,we deduce from the above two lemmas that if X is an EI process with random characteristics (α,σ,β), where $\alpha \le 0$ and $\beta_i \ge 0$ almost surely, then $\liminf_n X_{t_n}/t_n \le 0$ almost surely for any sequence $t_n \downarrow 0$.

Proof. By contrapositive, assume that $\liminf_n X_{t_n}/t_n > 0$. Then:

$$\begin{aligned}\lim_{N\to\infty} \mathbb{P}(X_{t_N} \le 0) &\le \lim_{N\to\infty} \mathbb{P}(X_{t_n} \le 0 \text{ for some } n \ge N)\\ &= \mathbb{P}(X_{t_n} \le 0 \text{ infinitely often})\\ &= \mathbb{P}\left(\liminf_n \frac{X_{t_n}}{t_n} \le 0\right) = 0.\square\end{aligned}$$

We are now ready to prove Theorem 5.1 in the case of totally asymmetric EI processes. The proof is similar to the one for (spectrally positive) Lévy processes given in Section 5.2.

Proof of Theorem 5.1 when $\alpha,\sigma = 0$ and $\beta_i \ge 0$. We first prove that $\overline{D}(X) = \infty$. Define the measure $\beta(dx) = \sum \delta_{\beta_j \in dx}$. Using Fubini's theorem

$$\int_0^1 \beta(t,\infty)dt = \int_0^1 \int_t^\infty \beta(dx)\,dt = \int_0^\infty \int_0^{1\wedge x} dt\,\beta(dx) = \sum \beta_j \wedge 1 = \infty.$$

Hence, we obtain for any $k \ge 1$

$$\int_0^1 \beta(kt,\infty)dt = \int_0^k \beta(t,\infty)du/k = \infty.$$

It follows that, for infinitely many $t \in (0,1)$ we have $|\Delta X_t| > kt$, which in turn implies $|X_t| > kt/2$ or $|X_{t-}| > kt/2$, for such t. Since k was arbitrary, we have $\overline{\lim}_{t\downarrow 0} |X_t|/t = \infty$.

Note that $(X_t/t, t \in (0,1])$ is a backward martingale (here, it is important that we restricted ourselves to the case $\alpha = 0$) without positive jumps. Note that it does not converge, thanks to the preceding paragraph. The process $N = (-X_{-t}/t, t \in [-1,0))$ is therefore a martingale without positive jumps (divergent almost surely). If we define τ_c as the first time N reaches $c \in \mathbb{R}_+$, then $(c - N_{-t\wedge\tau_c}, -1 \le t < 0)$ is a non-negative martingale. By the martingale convergence theorem $N_{\cdot\wedge\tau_c}$ converges a.s. to a finite limit as $t \uparrow 0$. If τ_c were infinite, N itself would therefore converge; hence, τ_c is almost surely finite. Since c was arbitrary, then $\overline{\lim}_{t\uparrow 0} N_{-t} = \infty$, which implies $\overline{\lim}_{t\downarrow 0} X_t/t = \infty$. (The above argument was taken from [Ber01] and [Fri72]). We have proved that $\overline{D}(X) = \infty$ for any extremal EI process with parameters $(\alpha, 0, \beta)$ of infinite variation when $\beta_i \ge 0$ for all i. Taking mixtures, we can let α and β be random, as long as $\sum_i \beta_i = \infty$ almost surely. We use this remark in the following paragraph.

We now prove that $\underline{D}(X) = -\infty$. First we apply a change of measure (through Proposition 5.3) to X; call the resulting measure $\mathbb{Q}^\theta$ to stress the dependence on θ. Write $\alpha^\theta \,\mathrm{Id}/T + Y^\theta$ for the (random parameter) EI process whose law is $\mathbb{Q}^\theta$. Recall from Proposition 5.2 that $\underline{D}(X)$ is a constant. Since $\mathbb{Q}^\theta$ is absolutely continuous with respect to $\mathbb{P}$ then $\underline{D}(X) = \alpha^\theta/T + \underline{D}(Y^\theta)$. Even if Y^θ has random parameters, its jumps are almost surely positive. The remark following Lemmas 5.2 and 5.3 implies that $\underline{D}(Y^\theta) \le 0$ almost surely. Taking expectations we see that

$$\underline{D}(X) = \mathbb{E}(\underline{D}(X)) = \mathbb{E}\left(\alpha^\theta/T\right) + \mathbb{E}\left(\underline{D}\left(Y^\theta\right)\right) \le \mathbb{E}\left(\alpha^\theta/T\right) \to -\infty$$

as $\theta \to -\infty$, since $\mathbb{E}(\alpha^\theta/T) \to -\infty$ by Proposition 5.4. $\square$

Proof of Theorem 5.1. As before, assume that $\alpha = 0$ and focus only on the statement $\underline{D}(X) = -\infty$, since we can at the end apply it to $-X$. Write $X = X^{pos} + X^{neg}$, where X^{pos} and X^{neg} are independent extremal EI processes with parameters $(0,\sigma,\beta^{pos})$ and $(0,0,\beta^{neg})$, where β^{pos} are the positive terms of β and β^{neg} the negative ones. We have proved Theorem 5.1 for X^{pos} and for X^{neg} if they are of infinite variation. If one of them is of finite variation, then the other one must be of infinite variation, and then Theorem 5.1 holds for X. Hence, we can assume that both X^{pos} and X^{neg} are of infinite variation.

But then, there exists a random sequence $T_k \downarrow 0$ such that $X^{neg}_{T_k}/T_k \to -\infty$ thanks to Theorem 5.1 for spectrally negative EI processes (just proved). Since (T_k) is independent of X^{pos}, we can apply Lemmas 5.2 and 5.3 to conclude that $\liminf_k X^{pos}_{T_k}/T_k \le 0$. We conclude that $\underline{D}(X) \le \liminf_k X_{T_k}/T_k = -\infty$. $\square$

5.4 Further applications

We now move on to the applications of Theorem 5.1, which were stated as Theorems 5.2, 5.3 and 5.4. We already mentioned that Theorem 5.3 follows from the same arguments as in [CUB15] once we have Theorem 5.1. Again, it suffices to prove the theorems for extremal EI processes.

5.4.1 An extension of Millar's zero-one law at the minimum

To prove Theorem 5.2, we will use a representation of the post-minimum process associated to an EI process found in [Ber93], which is now recalled.

Let X be an extremal EI process with parameters (α,σ,β); according to [Kal05, Ch. 2], such a process is a semimartingale. Let $\tau = \sup\{t \in [0,1] : \underline{X}_1 = X_t \wedge X_{t-}\}$ be the time of the ultimate infimum. Define the post-infimum process $\underset{\rightarrow}{X}$ as

$$\underset{\rightarrow}{X}(t) = \begin{cases} X_{\tau+t} - \underline{X}_1 & t \le 1-\tau \\ \dagger & t > 1-\tau, \end{cases}$$

(where $\dagger$ is a cemetery state) and the reversed pre-infimum process $\underset{\leftarrow}{X}$ as

$$\underset{\leftarrow}{X}(t) = \begin{cases} X_{(\tau-t)-} - \underline{X}_1 & t \le \tau \\ \dagger & t > \tau. \end{cases}$$

We introduce two processes $X^\uparrow$ and $X^\downarrow$ as in [Ber93]. Since X is a semimartingale, it has a semimartingale local time at zero, denote this by L; this local time is actually zero unless $\sigma > 0$ in which case

$$L_t = \lim_{\varepsilon \downarrow 0} \frac{\sigma^2}{2\varepsilon} \int_0^t \mathbf{1}_{|X_s| \le \varepsilon}\, ds.$$

Consider the time spent at $(0,\infty)$ and $(-\infty,0]$ up to time t of X, that is

$$A_t^+ = \int_0^t \mathbf{1}_{X_s>0} ds \quad \text{and} \quad A_t^- = \int_0^t \mathbf{1}_{X_s \le 0} ds,$$

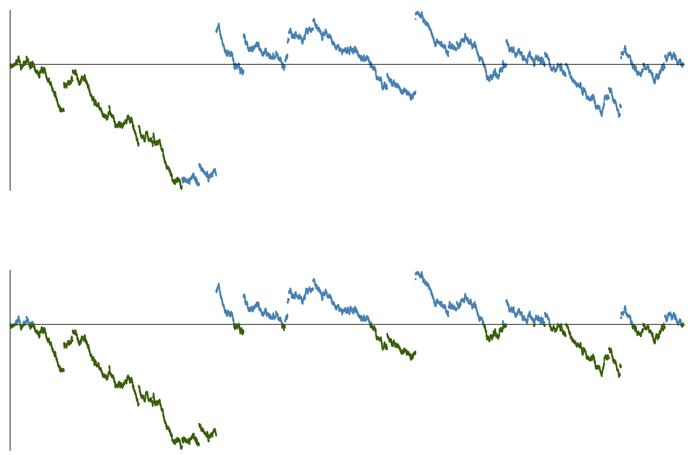

Figure 5.1: The post-infimum and yuxtaposition of the positive excursion processes are depicted in blue. The pre-infimum and yuxtaposition of the negative excursion processes in green.

and consider also their right-continuous inverses $\alpha^{\pm}(t) = \inf\{s : A^{\pm}_s > t\}$. It can be seen by a picture, that using the time change α^+ on X consist on erasing the jumps of X that fall on $(-\infty, 0]$ and closing up the gaps (similarly for α^-). The process of juxtaposition of the excursions in $(0,\infty)$ is given by

$$X^{\uparrow}(t) = \left(X_{\cdot} + \sum_{0<s\leq \cdot} [\mathbf{1}_{X_s\leq 0} X^{+}_{s-} + \mathbf{1}_{X_s>0} X^{-}_{s-}] + L \right) (\alpha^{+}(t)).$$

We remark that an excursion in $(0,\infty)$ includes the possible initial positive jump across 0 and excludes the possible ultimate negative jump across 0. The process of juxtaposition of the excursions in $(-\infty, 0]$ is given by

$$X^{\downarrow}(t) = \left(X_{\cdot} - \sum_{0<s\leq \cdot} [\mathbf{1}_{X_s\leq 0} X^{+}_{s-} + \mathbf{1}_{X_s>0} X^{-}_{s-}] - L \right) (\alpha^{-}(t)).$$

By establishing a bijection for discrete-time EI processes and passing to the limit, Bertoin obtains the following result.

Theorem 5.5 (Theorem 3.1 in [Ber93]). *Let X be an extremal EI process on $[0,1]$ with parameters (α, σ, β). Then, the following equality in law holds*

$$(\underset{\rightarrow}{X}, -\underset{\leftarrow}{X}) \stackrel{d}{=} (X^{\uparrow}, X^{\downarrow}).$$

A graphical representation of such equality in law is given in Figure 5.1.

We first establish the following simple result for EI processes.

Lemma 5.4. *When X satisfies UM, we have $\mathbb{P}\big(X_{U_j}$ or $X_{U_j-} = 0\big) = 0$ for every $j \in \mathbb{N}$.*

Proof. From the proof of Lemma 1.2 in [Kni96] we see that UM implies that $\mathbb{P}(X_t = x) = 0$ for any $t \in (0,1)$ and $x \in \mathbb{R}$. Fix any $j \in \mathbb{N}$ and define $X_t^j = X_t - \beta_j[\mathbf{1}_{U_j \le t} - t]$. Since $X_{U_j} = \beta_j(1-U_j) + X_{U_j}^j$, then

$$\mathbb{P}\big(X_{U_j} = 0\big) = \mathbb{P}\Big(\beta_j(1-U_j) + X_{U_j}^j = 0\Big) = \int_0^1 \mathbb{P}\Big(\beta_j(1-t) + X_t^j = 0\Big)\, dt = 0.$$

The statement for the left limit follows by time-reversibility. □

Proof of Theorem 5.2. Let X be an EI process satisfying UM. The previous lemma tells us that X does not jump into or from 0.

Assume that X is irregular upward. Then, X remains negative up to the time

$$\tau_0^+ = \inf\{t > 0 : X_t > 0\},$$

which is strictly positive. We actually have $\tau_0^+ < 1$ since otherwise X would have to jump at time 1, an event with zero probability. The trajectory of X up to τ_0^+ might comprise several excursions below zero and we will be interested in the first one, which ends at the random time

$$T = \min\{t > 0 : X_t \ge 0\} \le \tau_0^+.$$

Recall the definition of τ as the time of the last minimum. Let us prove that

$$0 < \Delta X_\tau \text{ almost surely.} \tag{5.8}$$

Note that

$$\Delta X_\tau = \Delta \underset{\rightarrow}{X}_0 \overset{d}{=} \Delta X_0^{\uparrow} = \Delta X_{\tau_0^+},$$

where the equality in distribution holds by Theorem 5.5. Since we do not jump into or from 0, then $\Delta X_{\tau_0^+} = 0$ if and only if $X_{\tau_0^+} = 0$ and X is continuous at τ_0^+. Assume that $\Delta X_{\tau_0^+} = 0$ has positive probability. Then the reversed pre-infimum process would hit zero twice (and the process X would hit its infimum twice); this is impossible under UM. Indeed, note that $X = X^{\downarrow}$ on $[0,T]$, and that from the construction of the pre-minimum process in Theorem 5.5, T has the same distribution as S where

$$S = \inf\Big\{t > 0 : \underset{\leftarrow}{X}_t = 0\Big\} = \tau - \sup\{s < \tau : X_{t-} = \underline{X}_1\}.$$

When $\Delta X_{\tau_0^+} = 0$, then $T > 0$ and $X_T = 0$. Hence, with positive probability, we would have that $S < \tau$ and $X_{S-} = \underline{X}_1$, so that the minimum of X is reached at least twice. The contradiction follows from negating (5.8), which proves its validity.

Conversely, assume X jumps from its infimum with positive probability. Then equation (5.8) holds true (though only with positive probability). Since X is continuous at zero, then $\tau_0^+ \in (0,1)$, which implies X is in $(-\infty, 0]$ on $(0, \tau_0^+)$. This means X is irregular upward with positive probability; being irregular upward is a tail event for the uniform random variables defining X, therefore, its probability is zero or one.

Using similar arguments we can prove X is irregular downward if and only if X jumps into its infimum.

Finally, Theorem 5.1 shows that X is both regular upward and downward when it is of infinite variation, so that X reaches its minimum continuously. □

5.4.2 EI processes conditioned to remain positive

The aim of this subsection is to prove Theorem 5.3. As before, let X be an extremal EI process with parameters $(0,\sigma,\beta)$. Assume that X is both upward and downward regular. Since $\alpha = 0$, either X has infinite activity ($\sum_i \mathbf{1}_{\beta_i \neq 0} = \infty$) or a Gaussian component ($\sigma \neq 0$). Otherwise, X would have piecewise linear trajectories with the same slope $\sum_i \beta_i$; but then, X would not be either upward or downward regular. Hence, X satisfies UM; let τ be the unique time X reaches its minimum. Theorem 5.2 tells us that X is continuous at τ. Corollary 3.1 in [CUB15] says that under these hypotheses, the law of X conditioned to remain above $-\varepsilon$ converges weakly to the Vervaat transform of X, given by $X_{\rho+\cdot \bmod 1} - X_\rho$. What was needed in the above cited corollary were conditions that would allow one to apply it and we have identified them in terms of regularity of both half-lines. In the particular case when X is of infinite variation, Theorem 5.1 tells us that X is both upward and downward regular and that therefore, the conclusion of Theorem 5.3 is satisfied.

The reader might wonder why we had to impose $\alpha = 0$. The reference [CUB15] has a description of what could be the limit when $\alpha > 0$ and $\beta = 0$ (that is, for a Brownian bridge from 0 to α). The candidate for a limit is described as a random shift, just as the Vervaat transformation for the case $\alpha = 0$, but it needs a bicontinuous family of (non-zero!) local times in its definition. Defining such a process for an EI process is an open problem; semimartingale local times are only non-zero when $\sigma > 0$, so a different approach is needed. Note that a limit theorem is not provided in [CUB15].

5.4.3 The convex minorant of EI processes

Let X be an extremal EI process with parameters (α,σ,β). To prove Theorem 5.4, we will rely strongly on [PUB12]. First, we establish some basic properties of the convex minorant in analogy with [PUB12, Proposition 1]. They will be fundamental in applying a transformation in Skorohod space, which is continuous on paths satisfying the conclusion.

Proposition 5.5. *Assume that X satisfies NPL and let C be the convex minorant of X. Then*

1. *The open set $\mathscr{O} = \{t \in [0,1] : C_t < X_t \wedge X_{t-}\}$ has Lebesgue measure 1.*
2. *For every connected component (g,d) of $\mathscr{O}$, $\Delta X_g \Delta X_d \geq 0$. If X has infinite variation, $\Delta X_g \Delta X_d = 0$.*
3. *If (g_1,d_1) and (g_2,d_2) are connected components of $\mathscr{O}$, then*

$$\frac{C_{d_1} - C_{g_1}}{d_1 - g_1} \neq \frac{C_{d_2} - C_{g_2}}{d_2 - g_2}.$$

The proof of the above proposition is almost the same as the corresponding one in [PUB12]. We just need to apply different results. For example, the fact that when X has finite variation, $\underline{D}(X) = \overline{D}(X) = \tilde{\alpha}$ (in the parametrization for this case), which is found in [Ber96, Prop. 4, p. 81] for Lévy processes, is now found in [Kal05, Cor. 3.30, p. 161] for EI processes (we have already used this result). Döblin's result that non-piecewise linear Lévy processes have continuous distributions, has a counterpart for EI processes in [Kni96], which also contains the fact that the minimum is reached in a unique place under NPL (which implies UM). One also needs our extensions of Millar's results stated in Theorem 5.2, as well as the fact that

$$\liminf_{t \to 0+} \frac{X_{U_i+t} - X_{U_i}}{t} = -\infty,$$

at any jump time U_i of X. This follows from 5.1 applied to $(X_t - \beta_i[\mathbf{1}_{U_i \leq t} - t])$.

To prove Theorem 5.4, we will use the following path transformation that leaves the laws of EI processes invariant.

Theorem 5.6. *Let X be an extremal EI process of parameters (α, σ, β) satisfying NPL. Define its convex minorant C and the open set of excursion intervals $\mathscr{O}$ as before. Let U be a uniform random variable on $(0,1)$ independent of X and consider the connected component (d,g) of $\mathscr{O}$ that contains U. Define the 3214 transformation X^U of X by means of*

$$X^U_t = \begin{cases} X_{U+t} - X_U, & 0 \leq t < d-U \\ C_d - C_g + X_{g+t-(d+U)} - X_U & d-U \leq t \leq d-g \\ C_d - C_g + X_{t-(d-g)} & d-g \leq t < d \\ X_t & d \leq t \leq 1 \end{cases}.$$

Then, $(U,X) \stackrel{d}{=} (d-g, X^U)$.

Remark that U belongs almost surely to $\mathscr{O}$, since the latter has Lebesgue measure 1 by Proposition 5.5.

The above path transformation can be understood as follows: the random variable U is used to select a face of the convex minorant of X, with endpoints g and d. This divides the trajectory into 4 parts, say $1,2,3$ and 4 which are then rearranged as $3,2,1,4$. Parts 1 and 4 have the same convex minorant as X, with the selected face removed. Parts 3 and 2 are interpreted as an inverse Vervaat transformation; the original trajectory 2 and 3 can be obtained as the Vervaat transform of the Knight bridge of X^U on $[0, d-g]$. One of the consequences is that $d-g$ has the same law as U, which is a remarkable universality result for exchangeable increment processes and is responsible for the stick-breaking process of Theorem 5.4. Indeed, we just need to iterate the path transformation on parts 1 and 4 of the trajectory of X^U. Therefore, Theorem 5.4 follows from Theorem 5.6, whose proof we now sketch, being very similar to the proof for Lévy processes of [PUB12]. It is based on an analogous path transformation for discrete time EI processes stated in [AP11, Theorem 8.1] or [APRUB11, Lemma 7]; the proof of the latter is by means of a bijection between permutations. To pass to the limit, one uses the continuity of the path transformation, on Skorohod space, whenever the trajectory satisfies the basic properties of Proposition 5.5, see Section 6.3 of [PUB12]. Continuity of the path transformation is mucho more simple when X is of infinite variation since then X is continuous at g and d. See Section 6.2 of [PUB12].

We end the chapter with an explanation of the distributional description of the maximum (or minimum, after multiplication by -1) of an EI process, which in discrete time is displayed in Equation (5.2), and how it proves the celebrated formula due to M. Kac, which in discrete time is Equation (5.1). Indeed, note that the infimum $\underline{X}_1$ of X on $[0,1]$ is the sum of the increments of the convex minorant that are negative. Thanks to Theorem 5.4, this gives us the equality in law

$$\underline{X}_1 \stackrel{d}{=} \sum_i \left[X_{S_i} - X_{S_{i-1}}\right]^- .$$

Next, conditioning on the stick-breaking process L, we see that

$$\mathbb{E}(\underline{X}_1) = \mathbb{E}\left(\sum_i \left[X_{S_i} - X_{S_{i-1}}\right]^-\right) = \sum_i \mathbb{E}(f(S_{i-1}, S_i)),$$

where $f(r,s) = \mathbb{E}([X_s - X_r]^-)$ for $r < s$. However, exchangeability implies that $f(r,s) = f(0,s-r)$, so that

$$\mathbb{E}(\underline{X}_1) = \sum_i \mathbb{E}(g(L_i))$$

where $g(l) = f(0,l)$. Finally, recall that the uniform stick-breaking process is invariant under size-biased permutations. Indeed, it is itself a size-biased permutation of a non-decreasing sequence; cf. [Pit95, Corollaries 7 and 8] or the following comment from [Pit06, p. 57]: *[The uniform stick-breaking process] has the same distribution as the size-biased permutation of the jumps of the Dirichlet process[...].* In particular, if conditionally on L, the index I has the law

$$\mathbb{P}(I = i \,|L) = L_i,$$

then $L_1 \stackrel{d}{=} L_I$. Hence, we obtain that for any $h : [0,1] \to \mathbb{R}_-$,

$$\mathbb{E}\left(\sum_i h(L_i)L_i\right) = \mathbb{E}(h(L_I)) = \mathbb{E}(h(L_1)) = \int_0^1 h(s)\, ds.$$

Applying the above result to $h(l) = g(l)/l$ gives

$$\mathbb{E}(\underline{X}_1) = \int_0^1 \frac{\mathbb{E}\big(X_l^-\big)}{l}\, dl.$$

APPENDIX

This appendix contributes with results and proofs of Section 3.6 and Chapter 4, about the convergence of CGW trees with Pareto offspring distribution, and cMGW forests.

Regularly Varying Functions.

We give a brief review of functions of regular and slowly variation.

Definition 5.4 (Slowly varying function)**.** *A positive measurable function* $L : \mathbb{R}_+ \mapsto \mathbb{R}_+$ *is said to be* slowly varying at ∞ *if* $L(x) > 0$ *for x large enough, and for all* $\lambda > 0$

$$\frac{L(\lambda x)}{L(x)} \to 1 \qquad x \to \infty.$$

In this case, we write $L \in SV$

Definition 5.5 (Regularly varying function)**.** *A measurable function* $f > 0$ *such that*

$$\frac{f(\lambda x)}{f(x)} \to \lambda^{\rho} \qquad x \to \infty, \forall \lambda > 0,$$

is called regularly varying *of index* ρ*. In this case, we write* $f \in RV(\rho)$*.*

The Characterization Theorem (see Theorem 1.4.1 of [BGT89]), says that any measurable positive function f is regularly varying iff $f(x) = x^{\rho} L(x)$, where $L \in SV$.

A subset $A \subset \mathbb{Z}$ is said to be lattice if there exists $a \in \mathbb{Z}$ and an integer $b \geq 2$ such that $A \subset a + b\mathbb{Z}$. A measure on $\mathbb{Z}$ is said to be lattice if its support is lattice.

Domains of attraction.

Assume $X_1, X_2, \ldots$ are i.i.d. variables with distribution F, and let G be a distribution not concentrated at one point

Definition 5.6. *The distribution F* belongs to the domain of attraction *of G if there exists constants* $a_n > 0$ *and* b_n *such that the distribution of* $(X_1 + \cdots + X_n)/a_n - b_n$ *tends to G.,*

For $x > 0$ define

$$U(x) = \int_{-x}^{x} y^2 F(dy).$$

An alternative form of the definition is given in Theorem 9.34 of [Bre92].

Theorem 5.7. *A distribution F belongs to the domain of attraction of the normal distribution iff $U \in SV$. It belongs to the domain of attraction of a stable law with exponent $\alpha \in (0,1)$ ($F \in DA(\alpha)$ for short) iff there are constants $M^+, M^- \geq 0$, $M^+ + M^- > 0$ such that as $x \to \infty$*

$$\frac{F(-x)}{1-F(x)} = \frac{M^-}{M^+},$$

and

$$M^+ > 0 \implies \lim \frac{1-F(tx)}{1-F(x)} = t^{-\alpha}$$

$$M^- > 0 \implies \lim \frac{F(-tx)}{F(-x)} = t^{-\alpha}.$$

In our case, the conditions $p_0 > 0$, and $(p_x, x \geq 0)$ being critical and aperiodic imply that $(p_x, x \geq 0)$ is non-lattice. Let g denote the density of X_1, the α-stable process at time one. The following version of the local limit theorem can be found in [Kor13].

Theorem 5.8 (Local Limit Theorem)**.** *Let W be a random walk on $\mathbb{Z}$ started at zero, with jump distribution in the domain of attraction of an α-stable law, with $\alpha \in (1,2]$. Suppose the law of $W(1)$ is critical, non-lattice, and $W(1)$ takes values on $\{-1,0,1,2,\ldots,\}$. Then,*

$$\lim_n \sup_{k\in\mathbb{Z}} |a_n \mathbb{P}(W(n) = k) - g(k/a_n)| = 0,$$

where $a_n = n^{1/\alpha} L(n)$ and L is a slowly varying function.

The following result is given in IX.8, page 312 of [Fel71].

Theorem 5.9. *A distribution F with support on $\mathbb{R}_+$ belongs to some domain of attraction iff there exists $L_\mu \in SV$ such that*

$$\mu(x) := \int_0^x y^2 F(dy) = x^{2-\alpha} L_\mu(x) \qquad x \to \infty, \tag{5.9}$$

with $0 < \alpha \leq 2$.

In fact, from (8.5) page 313 of [Fel71], Equation (5.9) is equivalent to

$$\frac{x^2(1-F(x))}{\mu(x)} \to \frac{2-\alpha}{\alpha}, \tag{5.10}$$

in the sense that the two relations imply each other. The next is Theorem 1.5.11.ii of [BGT89].

Theorem 5.10 (Karamata's Theorem, direct half)**.** *For some $\rho \in \mathbb{R}$, let $f \in RV(\rho)$ be bounded in each compact subset of $\mathbb{R}_+$. Then, for any $\sigma < -(\rho+1)$*

$$\frac{x^{\sigma+1} f(x)}{\int_x^\infty t^\sigma f(t)dt} \to -(\sigma+\rho+1) \quad x \to \infty.$$

Using the previous results, we prove the following lemma of Subsection 3.6.2.

Proof of Lemma 3.13. From Theorem 3.3 applied to the triangular array $(\xi(j)/a_n, j \in [n])$, with $\xi(j)$ having law μ, we have for some constant $c > 0$

$$cx^{-\alpha} = \lim_n n\mathbb{P}(\xi(j) > a_n x) = \lim_n n\mu((a_n x, \infty)) = x^{-\alpha} \lim_n \frac{nL(a_n x)}{a_n^\alpha} = x^{-\alpha} \lim_n \frac{L(a_n x)}{L_a^\alpha(n)},$$

hence the limit on the right-hand side is finite.

Consider the array $(\tilde{\xi}(j)/a_n, j \in [b_n])$. This has terms in the sum that are uniformly infinitesimal, because

$$\lim_n \max_{j\in[b_n]} \mathbb{P}\Big(\tilde{\xi}(j)/a_n > \varepsilon\Big) = \lim_n \tilde{\mu}((a_n\varepsilon, \infty)) = c \lim_n \frac{L(a_n\varepsilon)}{n^{1-1/\alpha} L_a^{\alpha-1}(n)} = c \lim_n \frac{L(a_n\varepsilon)}{L_a^\alpha(n)} \frac{a_n}{n} \to 0.$$

Therefore, for any $x > 0$ we have

$$\lim_n b_n \tilde{\mu}((a_n x, \infty)) = \lim_n \frac{n}{a_n} (a_n x)^{1-\alpha} L(a_n x) = x^{1-\alpha} \lim_n \frac{L(a_n x) L_a^{1-\alpha}(n)}{L_a(n)} = (\alpha - 1) x^{1-\alpha}.$$

On the other hand, from Theorem 5.9, we have $\sigma_x^\varepsilon = x^{2-(\alpha-1)} L_\sigma(x)$, for some $L_\sigma \in SV$. This, together with (5.10) implies the convergence

$$\frac{(a_n\varepsilon)^2 \tilde{\mu}((a_n\varepsilon, \infty))}{Var(\tilde{\xi}(1)\mathbf{1}_{\tilde{\xi}(1)\le a_n\varepsilon})} = \frac{(a_n\varepsilon)^2 (a_n\varepsilon)^{1-\alpha} L(a_n\varepsilon)}{(a_n\varepsilon)^{3-\alpha} L_\sigma(a_n\varepsilon)} = \frac{L(a_n\varepsilon)}{L_\sigma(a_n\varepsilon)} \to \frac{3-\alpha}{\alpha-1}.$$

Therefore

$$\sigma_n^\varepsilon n/a_n = \frac{n}{a_n^3} Var(\tilde{\xi}(1)\mathbf{1}_{\tilde{\xi}(1)<a_n\varepsilon}) = \varepsilon^{3-\alpha} \frac{L_\sigma(a_n\varepsilon)}{L(a_n\varepsilon)} \frac{L(a_n\varepsilon)}{L_a^\alpha(n)} \to \varepsilon^{3-\alpha} c_\alpha''',$$

as $n \to \infty$ for some constant c_α'''. The above quantity goes to zero as $\varepsilon \downarrow 0$. Since the variance is bounded by the second moment, the second condition of the theorem holds. To prove the last condition, note that

$$b_n \mathbb{E}\Big(\tilde{\xi}(1)/a_n; |\tilde{\xi}(1)| \le h a_n\Big) = \frac{n}{a_n^2} (h a_n)^{2-\alpha} L_\mu(h a_n) = \frac{L_\mu(h a_n)}{L_a(n)},$$

which is bounded by a constant. □

Convergence of the normalized empirical degree sequence of CGW trees with offspring distribution in DA

We prove Lemma 4.8, that is, the normalized empirical degree sequence of a CGW tree with offspring distribution in DA(α), for $\alpha \in (1,2)$, converges to the offspring distribution.

Proof of Lemma 4.8. Let us prove that $\hat{\mu}_i \to \mu_i$ in probability for all i. Let W be the DFW associated to the tree T under $\mathbb{P}_\mu^n$ (so, is an excursion of size n). Define the skip-free random walk $\tilde{S}$, with increments $\Delta\tilde{S}_i = \tilde{X}_i$, $\tilde{S}_0 = 0$ and distribution

$$\mathbb{P}\big(\tilde{X}_1 = i\big) = \mu_{i+1} \quad i \ge -1.$$

Then $W \overset{d}{=} \big(\tilde{S} | T_{-1} = n\big)$ under $\mathbb{P}_\mu^n$, where T_{-1} is the first time the walk $\tilde{S}$ hits -1.

Let $K=(K_i;i\geq 0)$, where $K_i=\#\{k:\tilde{X}_k=i-1\}$ is the number of increments having size $i-1$. Using the empirical degree distribution (4.8), for $B\subset\mathbb{N}^\infty$ we compute

$$\begin{aligned}\mathbb{P}^n_\mu((\hat{n}_i;i\geq 0)\in B)&=\mathbb{P}_\mu\left(\#\{k:\Delta W_k=i-1\}\in B_i,i\geq 0|\ |T|=n\right)\\&=\mathbb{P}\left(K\in B|T_{-1}=n\right)\\&=\mathbb{P}\left(K\in B|\tilde{S}_n=-1\right),\end{aligned}$$

where in the last inequality we used the rotation principle (see Chapter 6 of [Pit06] for the Kemperman's formula, and also Proposition 2 of [BK00] and the commentaries before it).

Let $\mathscr{F}_k$ be the σ-algebra generated by the first k increments of $\tilde{S}$. We need to prove that for all $B\in\mathscr{F}_{\lfloor n/2\rfloor}$

$$\mathbb{P}\big(B|\tilde{S}_n=-1\big)\leq c\mathbb{P}(B) \tag{5.11}$$

as inequality (24), in Lemma 11 of [BM14b]. The proof follows the same lines as the mentioned paper, but for the local limit theorems, we use Theorem 1 and Lemma 1 of [Kor13].

As in Lemma 11 of [BM14b], for all $B\in\mathscr{F}_{\lfloor n/2\rfloor}$

$$\mathbb{P}\big(B|\tilde{S}_n=-1\big)=\sum_k\mathbb{P}\big(B,\tilde{S}_{\lfloor n/2\rfloor}=k\big)\frac{\mathbb{P}\big(\tilde{S}_{n-\lfloor n/2\rfloor}=-1-k\big)}{\mathbb{P}\big(\tilde{S}_n=-1\big)}.$$

We bound the numerator and denominator by a constant, and sum over all k. For the numerator, we use Theorem 1 of [Kor13]. For the denominator we use Lemma 1 of the same paper. Thus Equation (5.11) is true. This says that there exists $c_1>0$, such that for n big enough

$$\sup_k\mathbb{P}\big(\tilde{S}_{n-\lfloor n/2\rfloor}=k\big)\leq c_1/b_{n-\lfloor n/2\rfloor}.$$

For the denominator we use Lemma 1 of [Kor13], so for all $\varepsilon\in(0,1)$ there exists $M\in\mathbb{N}$ and $c_2>0$ such that for all $n\geq M$ we have

$$\frac{c_2(1-\varepsilon)}{b_n}\leq\mathbb{P}\big(\tilde{S}_n=-1\big).$$

Then, for n big enough we can find $c>0$ such that

$$\frac{\sup_k\mathbb{P}\big(\tilde{S}_{n-\lfloor n/2\rfloor}=-1-k\big)}{\mathbb{P}\big(\tilde{S}_n=-1\big)}\leq c,$$

and (5.11) is true. The proof that $\hat{\mu}_i\to\mu_i$ in probability is the same as in Lemma 11 of [BM14b]. We work with the random walk $\tilde{S}$ and use equation (5.11) to translate the results for the bridge. For $i\geq 0$, let $K_i^{(1)}=\#\{k\leq\lfloor n/2\rfloor:\tilde{X}_k=i-1\}$ be the number of increments of $(\tilde{X}_1,\ldots,\tilde{X}_{\lfloor n/2\rfloor})$ which are equal to $i-1$. Notice that this random variables are $\mathscr{F}_{\lfloor n/2\rfloor}$ measurable. Then

$$\frac{1}{\lfloor n/2\rfloor}K_i^{(1)}=\frac{1}{\lfloor n/2\rfloor}\sum_{j\geq 1}^{\lfloor n/2\rfloor}\mathbf{1}_{\tilde{X}_j=i-1}\to\mathbb{P}\big(\tilde{X}_1=i-1\big)=\mu_i,$$

by the strong law of large numbers.

If we define

$$K_i^{(2)} = \#\{k \in \{\lfloor n/2 \rfloor + 1, \dots, n\} : \tilde{X}_k = i-1\},$$

and the cyclic permutation σ that exchanges the first $\lfloor n/2 \rfloor$ elements of $[n]$ with the last $n - \lfloor n/2 \rfloor$, then

$$\frac{1}{n - \lfloor n/2 \rfloor} K_i^{(2)} = \frac{1}{n - \lfloor n/2 \rfloor} \sum_{j \geq \lfloor n/2 \rfloor + 1}^{n} \mathbf{1}_{\tilde{X}_j = i-1} \to \mu_i.$$

Let $\varepsilon > 0$, and let

$$K_i^{(1,n)} = \frac{\lfloor n/2 \rfloor}{n} \frac{K_i^{(1)}}{\lfloor n/2 \rfloor}, \quad K_i^{(2,n)} = \frac{n - \lfloor n/2 \rfloor}{n} \frac{K_i^{(2)}}{n - \lfloor n/2 \rfloor}.$$

Notice that $K_i^{(j,n)} \to \mu_i/2$ for $j = 1,2$ in probability under $\mathbb{P}(\cdot | \tilde{S}_n = -1)$. Using triangle inequality, as $K_i/n = K_i^{(1,n)} + K_i^{(2,n)}$, then

$$\begin{aligned}
&\mathbb{P}(|K_i/n - \mu_i| > \varepsilon | X_n = -1) \\
&\leq c\mathbb{P}\Big(|K_i^{(1,n)} - \mu_i/2| > \varepsilon/2\Big) + c\mathbb{P}\Big(|K_i^{(2,n)} - \mu_i/2| > \varepsilon/2\Big) \\
&\to 0.
\end{aligned}$$

□

Let M^n have a multidimensional distribution with parameters n and $\mu_0, \mu_1, \dots$. We know that rescaling any component j by n converges a.s. to μ_j, thanks to the strong law of large numbers. We now prove that this convergence holds also uniformly in j.

Lemma 5.5. *Consider any positive ε and a distribution $(\mu_i, i \geq 0)$. Then, there exists $N \in \mathbb{N}$ such that for any $n \geq N$ and any j, the vector $M^n = (N_0^n, N_1^n, \dots)$ with multinomial distribution having parameters n and $\mu_0, \mu_1, \dots$ satisfies*

$$\left| \frac{N_j^n}{n} - \mu_j \right| \leq \varepsilon.$$

Proof. Define $p_j = \sum_0^j \mu_i$. The multinomial distribution can be obtained from a sequence of i.i.d. uniform variables $(U_i, i \in [n])$ as

$$N_j^n = \#\{i : U_i \in [p_j, p_{j+1})\} = \sum_1^n \mathbf{1}_{i : U_i \in [p_j, p_{j+1})}.$$

Hence we have

$$\frac{\sum_0^j N_i^n}{n} = \frac{\sum_1^n \mathbf{1}_{i : U_i \leq p_{j+1}}}{n}.$$

The Glivenko-Cantelli theorem implies the convergence of the empirical distribution of the U_i's, hence

$$\sup_j \left| \frac{\sum_0^j (N_j^n - n\mu_j)}{n} \right| \leq \varepsilon,$$

for every n big enough. Using this inequality for j and for $j-1$ we obtain

$$-\varepsilon + p_j - \varepsilon - p_{j-1} \leq \frac{N_j^n}{n} \leq \varepsilon + p_j + \varepsilon - p_{j-1},$$

for every n big enough and every j. □

GLOSSARY AND LIST OF SYMBOLS

$\mathbb{N} = \{1, 2, \ldots, \}$

$\mathbb{Z}_+ = \{0, 1, 2, \ldots, \}$

$[n]_0 := \{0, \ldots, n\}$

$[n] := \{1, \ldots, n\}$

$\{0, 2, \ldots, n\} = 2[n/2]_0$

$\{2, \ldots, n\} = 2[n/2]$

$\mathbb{E}(X; A) = \mathbb{E}(X\mathbf{1}_A)$

$\mathbb{W} = \{x \in \mathbb{R}^d : x_1 < \cdots < x_d\}$, the Weyl chamber

$\Delta(\mathbf{x}) = \prod_{1 \le i < j \le d}(x_j - x_i)$, the Vandermonde determinant for $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{W}$

X^*, the time-reversed process of X on $[n]$, that is $X_i^* = X_n - X_{n-i}$ for $0 \le i \le n$

$\max\{Y_i, i \in \mathscr{I}\} = (\max\{Y_i^1, i \in \mathscr{I}\}, \max\{Y_i^2, i \in \mathscr{I}\})$ given any index set $\mathscr{I} \subset \mathbb{Z}_+$

$\overline{Y}_n = \max\{Y_j, 0 \le j \le n\}$

$\underline{Y}_n = \min\{Y_j, 0 \le j \le n\}$

$\overline{\overline{Y}}_n = \max\{Y_j, 1 \le j \le n\}$

$\underline{\underline{Y}}_n = \min\{Y_j, 1 \le j \le n\}$

$Y_n \vee (y_1, y_2) = (Y_n^1 \vee y_1, Y_n^2 \vee y_2)$.

Strict ascending ladder times: for a discrete-time process X those are the times of new maximums $\alpha_0 = 0$, and $\alpha_i = \inf\{n > \alpha_{i-1} : X_n > \overline{X}_{\alpha_{i-1}}\}$. The strict descending ladder times of a process, are the times of new minimums.

$J_{i+1} = \min\left\{n > J_i : \overline{Y}_{n-1}^k < Y_n^k, k = 1, 2\right\}$ for $i \ge 0$ and $J_0 = 0$, are the times in common among the strict ascending ladder times of Y^1 and Y^2.

$(\beta_n, n \ge 0)$ is the ordered union of the strict descending ladder times of Y^1 and Y^2, with $\beta_0 = 1$, that is, the ordered union of $(\beta_j^k, j \in \mathbb{N})$, the strict descending ladder times of Y^k

$\mathbf{e_i}$ vector of zeros, except a one at position i.

$S_a^b f(t) = f(at)/b$ for f a function from $\mathbb{R}_+$ to $\mathbb{R}$, and positive a, b.

$|T|$ size of the tree T.

BFO, BFW are the breadth-first order and breadth-first walk of a tree

DFO, DFW are the depth-first order and depth-first walk of a tree

$T(u)$ is the subtree generated by u, a vertex in the tree T, that is, the subtree with root u and individuals all the descendants of u

$\mathbb{P}_{\mathbf{s}}$ is the law of a uniform tree with given degree sequence $\mathbf{s}$

Z_{s_n}, C_{s_n} and W_{s_n} are the profile, the cumulative profile and the BFW of a tree with law $\mathbb{P}_{s_n}$, respectively. The BFW of such tree satisfies $W_{s_n}(0) = 1$

Z_{s_n} satisfies $Z_{s_n}(j+1) = W_{s_n} \circ C_{s_n}(j)$ for every $j \geq 0$, and is also called the Lamperti transform of W_{s_n}

$C_{s_n}(k) = \sum_0^k Z_{s_n}(j)$ for $k \geq 0$, and is extended to $\mathbb{R}_+$ by linear interpolation

$a_n = s_n/b_{s_n}$ is the spacial scaling of the profile

b_{s_n} is the spacial scaling of the BFW

$X^n = S_{s_n}^{b_n} W_{s_n}$ the rescaled BFW

C^n is defined as follows: $C^n(0) = 0$, and for $i \geq 0$ and $t \in [t_i, t_{i+1})$, with $t_i = i/a_n$, write

$$C^n(t) := \int_0^t X^n \circ C^n(\lfloor a_n u \rfloor / a_n) du = C^n(t_i) + (t - t_i) X^n \circ C^n(t_i).$$

$Z^n = D_+ C^n$ the right-hand derivative of C^n

C is the cumulative Lamperti transform of X, that is, the unique solution to $D_+ C = X \circ C$. This solution is positive on $\mathbb{R}_+$ if not the zero function

$Z = X \circ C$

$\Lambda_{0+} = \inf\{t > 0 : C^\Lambda(t) > 0\}$

$\Lambda_{l,\varepsilon} = \inf\{t > 0 : C^l(t) > \varepsilon\}$

$I(\cdot) = \int_0^\cdot ds/X_s$

$I(b) - I(a) = \int_a^b ds/X_s$ for any $0 < a \leq b \leq 1$

$\inf\{t > 0 : C^\Lambda(t) > \cdot\} = I^\Lambda(\cdot) = \Lambda + I(\cdot)$

$\Lambda_{k,\varepsilon_k} = \Lambda_{n(l^1(k)),\varepsilon_{n(l^2(k))}}$ and $\Lambda_{k,1-\varepsilon_k} := \Lambda_{n(l^1(k)),1-\varepsilon_{n(l^2(k))}}$ are defined in page 73, the deterministic subsequences appearing here are given in Lemma 3.5

$I^k(\cdot) = \Lambda_{k,\cdot}$

$\lambda^{(k)} = \lfloor \lambda a_{s_k} \rfloor$ for any $\lambda \geq 0$

$\tilde{T}$ tree defined in Proposition 3.9. For any tree T coded in DFO by w, $u \in \{2, \ldots, |T|\}$ and $h \in \mathbb{N}$, it is defined as the tree coded by the BFW $\tilde{w}$, using the transformation $\Phi_{h,u}(w)$ (which is defined before Lemma 3.6)

$|T(v)|(h_1, h_2)$ the number of individuals in $T(v)$ having height $h \in \{h_1, \ldots, h_2\}$ in the tree T

$|T(v)|(h_1)$ the number of individuals in $T(v)$ up to height h_1 in the tree T

$h(v_k, 1/2) = \inf\{h > h(v_k) : |\tau_k(v_k)|(h) \geq |\tau_k(v_k)|/2\}$

$h(v_k, 1/4) = \inf\{h > h(v_k, 1/2) : |\tau_k(v_k)|(h(v_k, 1/2) + 1, h) \geq |\tau_k(v_k)|/4\}$

DA(α), a distribution belongs to the domain of attraction of a stable law of index α, see Definition 1.4

Q-process is the profile of the infinite size-biased tree. It is denoted by $\bar{Z}$

Skip-free random chain is a path $x : \mathbb{Z}_+ \mapsto \mathbb{Z}$ with increments on $\{-1, 0, 1, 2, \ldots\}$, it is also called downward skip-free chain

S_d is given in Definition 4.2

$\mathbb{F}^{plane}_{\mathbf{r},\mathbf{n}}$ set of plane d-type forests having $\mathbf{r} = (r_1, \ldots, r_d)$ roots and $\mathbf{n} = (n_1, \ldots, n_d)$ vertices of each type

$\mathbb{F}^{labeled}_{\mathbf{r},\mathbf{n}}$ set of labeled d-type forests having $\mathbf{r} = (r_1, \ldots, r_d)$ roots and $\mathbf{n} = (n_1, \ldots, n_d)$ vertices of each type. Our labeled multitype forests have labels on $[n]$, that is, for $F \in \mathbb{F}^{labeled}_{\mathbf{r},\mathbf{n}}$, each individual v has a unique label $i \in [n]$ and a type $c_F(v) \in [d]$; also, F has fixed *root set* $[r]$, that is, the r_1 type 1 roots have labels on $\{1, \ldots, r_1\}$, the r_2 type 2 roots have labels on $\{r_1 + 1, \ldots, r_1 + r_2\}$, and so on

$\mathbb{F}^{binary}_{\mathbf{r},\mathbf{n}}$ set of plane binary d-type forests having $\mathbf{r} = (r_1, \ldots, r_d)$ roots and $\mathbf{n} = (n_1, \ldots, n_d)$ vertices of each type. In this case vertices have either zero or two children, for each type.

$\mathbb{B}^{s,m}$ set of bridges finishing at $-m$ (skip-free paths) at time s

$\mathbb{E}^{s,m}$ set of excursions finishing at $-m$ at time s

$\mathbf{S} = (N_i, i \geq 0)$ degree sequence, that is a sequence of non-negative numbers such that $\sum N_i = 1 + \sum iN_i$

$\mathbb{B}_\mathbf{S}$ set of bridges finishing at $-m$ and with N_i increments of size $i-1$, for each i

$\mathbb{E}_\mathbf{S}$ set of excursions finishing at $-m$ and with N_i increments of size $i-1$, for each i

$\mathbb{F}_\mathbf{S}$ set of plane forests with degree sequence $\mathbf{S}$

$\mathbb{B}_{\mathbf{S},\mathbf{r}}$ set of $d \times d$ bridges, where $w^b_{i,j}$ finishes at $k_{i,j} = \sum kN_{i,j}(k) - n_i \mathbf{1}_{i=j}$, and with $N_{i,j}(k)$ increments of size $k-1$, for each $i, j \in [d]$ and k

$\mathbb{F}_{\mathbf{S},\mathbf{r}}$ set of multitype plane forests with degree sequence $\mathbf{S}$, root-type $\mathbf{r}$ and individuals-type $\mathbf{n}$

$F^{(i)}$ subtree of type i of the multitype forest F

$\theta_{q,n}(y)$ is the n-cyclical permutation of the vector y with length n at the point q

$\theta_{\mathbf{q},\mathbf{n}}(x)$ is the $\mathbf{n}$-cyclical permutation of x, that is, each component $x^{(i)}$ is cyclically permuted using θ_{q_i,n_i}

$V(\mathbf{w}^b, u)$ Multidimensional Vervaat Transform (see Definition 4.5)

$\mathbb{B}_{\mathbf{S},\mathbf{r}}$ set of paths $w^b_{i,j}$, the latter having increments $c_{i,j} \circ \sigma_{i,j}(l) - \mathbf{1}_{i=j}$, with $l \in [n_i]$, $\mathbf{c}$ the child sequence of $\mathbf{S}$, and for every permutation $\sigma_{i,j}$ of $[n_i]$

$\mathbb{F}_{\mathbf{S},\mathbf{r}}$ set of multitype plane forests with degree sequence $\mathbf{S}$, root-type $\mathbf{r}$ and individuals-type $\mathbf{n}$

$$A(\mathbf{n},\mathbf{r}) = \{K = (k_{ij}) : k_{ij} \geq 0 \text{ for } i \neq j, 0 \leq \sum_{i \neq j} k_{ij} \leq n_j - r_j, -k_{jj} = r_j + \sum_{i \neq j} k_{ij}, \forall j \in [d]\},$$

and the notation $\sum_{K \in A(n,r)}$ means the summation over all k_{ij} with $i, j \in [d]$ and $i \neq j$, such that $K = (k_{ij}) \in A(n,r)$. For fixed $j \in [d]$ define

$$A_j(\mathbf{n},\mathbf{r}) = \{K^j = (k_{1j}, \dots, k_{dj}) : k_{lj} \geq 0 \text{ for } l \neq j, 0 \leq \sum_{l \neq j} k_{lj} \leq n_j - r_j\},$$

and the notation $\sum_{K^j \in A_j(n,r)}$ means the summation over all k_{lj} with $l \in [d]$ and $l \neq j$, such that $K^j = (k_{1j}, \dots, k_{dj}) \in A_j(n,r)$

$n_i = min\{n : x_n^{i,i} = \min_{0 \leq k \leq n_i} x_k^{i,i}\}$

$k_{i,i} = -\min_{0 \leq n \leq n_i} x_n^{i,i} = -x_{n_i}^{i,i}$

Admissible breadth-first pair, a pair of càdlàg functions (f,g) such that f has no negative jumps, g is non-decreasing and $f(0)+g(0) \geq 0$

IVP$_\sigma(f,g)$: Consider an admissible breadth-first pair (f,g) and $\sigma > 0$. For the partition $t_i = i\sigma$ with $i \geq 0$, let c^σ the function defined by $c^\sigma(0) = 0$ and

$$c^\sigma(t) = c^\sigma(t_i) + (t - t_i)[f \circ c^\sigma(t_i) + g(t_i)]^+ \qquad t_i \leq t < t_{+1}.$$

Equivalently, the function c^σ is the unique solution of

$$c^\sigma(t) = \int_0^t [f \circ c^\sigma(\lfloor s/\sigma \rfloor \sigma) + g(\lfloor s/\sigma \rfloor \sigma)]^+ \, ds$$

IVP(f,g): Consider (f,g) an admissible breadth-first pair. The *initial value problem*, denoted by IVP(f,g) is defined as

$$D_+ c = f \circ c + g \quad \text{ and } \quad c(0) = 0.$$

If $g \equiv 0$, we simply write IVP(f), and if $\sigma = 0$ in the previous definition, we write IVP$_0(f,g) =$ IVP(f,g)

Dini derivative: for any function $f : \mathbb{R}_+ \to \mathbb{R}$ we define the upper and lower right-hand Dini derivatives of f at $t \in \mathbb{R}$ by

$$\limsup_{h \to 0+} (f(t+h) - f(t))/h \quad \text{ and } \quad \liminf_{h \to 0+} (f(t+h) - f(t))/h,$$

respectively. If $h \to 0-$ then the above are the left-hand Dini derivatives.

A function f is said to be of bounded variation on an interval $[a,b] \subset \mathbb{R}$ if its total variation is finite, that is

$$\sup_{P \in \mathscr{P}} \sum_0^{n_P - 1} |f(x_{i+1}) - f(x_i)| < \infty,$$

where the supremum is taken over all the partitions $\{x_0, \dots, x_{n_p}\}$ of $[a,b]$ satisfying $x_i \leq x_{i+1}$ for $0 \leq i \leq n_p - 1$

A function f is said to be of infinite variation if its total variation is infinite

Extremal exchangeable increments process: an exchangeable increment process with deterministic characteristics (α, σ, β)

$\rho = \inf\{t : X_t \wedge X_{t-} = \underline{X}_1\}$ for an EI process on $[0,1]$

$V(X) = X_{\cdot+\rho,\text{ mod }1}$ the Vervaat transform of an EI process X with $\alpha = 0$

$\underrightarrow{X}$ post-infimum process

$\underleftarrow{X}$ reversed pre-infimum process

$X^{\uparrow}$ process of juxtaposition of the excursions of X in $(0, \infty)$

$X^{\downarrow}$ process of juxtaposition of the excursions of X in $(-\infty, 0]$

Bibliography

[AB12] Louigi Addario-Berry, *Tail bounds for the height and width of a random tree with a given degree sequence*, Random Structures & Algorithms **41** (2012), no. 2, 253–261. 20

[ABBG12] L. Addario-Berry, N. Broutin, and C. Goldschmidt, *The continuum limit of critical random graphs*, Probab. Theory Related Fields **152** (2012), no. 3-4, 367–406. MR 2892951 99

[ABD18] Romain Abraham, Maïtine Bergounioux, and Pierre Debs, *Automatic choice of the threshold of a grain filter via Galton-Watson trees: application to granite cracks detection*, J. Math. Imaging Vision **60** (2018), no. 1, 50–69. MR 3742847 12

[ADG18] Romain Abraham, Jean-François Delmas, and Hongsong Guo, *Critical multi-type Galton-Watson trees conditioned to be large*, J. Theoret. Probab. **31** (2018), no. 2, 757–788. MR 3803914 20, 99

[AJ97] Krishna B. Athreya and Peter Jagers (eds.), *Classical and modern branching processes*, The IMA Volumes in Mathematics and its Applications, vol. 84, Springer-Verlag, New York, 1997, Papers from the IMA Workshop held at the University of Minnesota, Minneapolis, MN, June 13–17, 1994. MR 1601681 12, 20, 98

[Ald91a] David Aldous, *The continuum random tree. I*, Ann. Probab. **19** (1991), no. 1, 1–28. MR 1085326 12, 99

[Ald91b] ______, *The continuum random tree. II. An overview*, Stochastic analysis (Durham, 1990), London Math. Soc. Lecture Note Ser., vol. 167, Cambridge Univ. Press, Cambridge, 1991, pp. 23–70. MR 1166406 12, 17, 52, 61, 85, 99

[Ald93] ______, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. MR 1207226 12

[All11] Linda J. S. Allen, *An introduction to stochastic processes with applications to biology*, second ed., CRC Press, Boca Raton, FL, 2011. MR 2560499 12, 20, 98

[AMP04] David Aldous, Grégory Miermont, and Jim Pitman, *The exploration process of inhomogeneous continuum random trees, and an extension of Jeulin's local time identity*, Probab. Theory Related Fields **129** (2004), no. 2, 182–218. MR 2063375 17, 72

[AN04] K. B. Athreya and P. E. Ney, *Branching processes*, Dover Publications, Inc., Mineola, NY, 2004, Reprint of the 1972 original [Springer, New York; MR0373040]. MR 2047480 22, 90, 102

[AP98] David Aldous and Jim Pitman, *Tree-valued Markov chains derived from Galton-Watson processes*, Ann. Inst. H. Poincaré Probab. Statist. **34** (1998), no. 5, 637–686. MR 1641670 17

[AP11] Josh Abramson and Jim Pitman, *Concave majorants of random walks and related Poisson processes*, Combin. Probab. Comput. **20** (2011), no. 5, 651–682. MR 2825583 134, 151

[APRUB11] Josh Abramson, Jim Pitman, Nathan Ross, and Gerónimo Uribe Bravo, *Convex minorants of random walks and Lévy processes*, Electron. Commun. Probab. **16** (2011), 423–434. MR 2831081 151

[AS95] Laurent Alonso and René Schott, *Random generation of trees*, Kluwer Academic Publishers, Boston, MA, 1995, Random generators in computer science. MR 1331596 99

[BA99] Albert-László Barabási and Réka Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509–512. MR 2091634 12, 52

[Bai00] Jinho Baik, *Random vicious walks and random matrices*, Comm. Pure Appl. Math. **53** (2000), no. 11, 1385–1410. MR 1773413 35

[BD94] J. Bertoin and R. A. Doney, *On conditioning a random walk to stay nonnegative*, Ann. Probab. **22** (1994), no. 4, 2152–2167. MR 1331218 11, 30, 34, 36

[Ber93] Jean Bertoin, *Splitting at the infimum and excursions in half-lines for random walks and Lévy processes*, Stochastic Process. Appl. **47** (1993), no. 1, 17–35. MR 1232850 7, 8, 10, 30, 31, 34, 35, 134, 147, 148

[Ber96] ———, *Lévy processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996. MR MR1406564 138, 141, 150

[Ber97] ———, *Regularity of the half-line for Lévy processes*, Bull. Sci. Math. **121** (1997), no. 5, 345–354. MR 1465812 136

[Ber01] ———, *Eternal additive coalescents and certain bridges with exchangeable increments*, Ann. Probab. **29** (2001), no. 1, 344–360. MR 1825153 26, 27, 65, 146

[Ber02] ———, *Some aspects of additive coalescents*, Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002), Higher Ed. Press, Beijing, 2002, pp. 15–23. MR 1957515 27, 135

[BFP+09] Alexei Borodin, Patrik L. Ferrari, Michael Prähofer, Tomohiro Sasamoto, and Jon Warren, *Maximum of Dyson Brownian motion and non-colliding systems with a boundary*, Electron. Commun. Probab. **14** (2009), 486–494. MR 2559098 35

[BGT89] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation*, Encyclopedia of Mathematics and its Applications, vol. 27, Cambridge University Press, Cambridge, 1989. MR 1015093 18, 95, 122, 153, 154

[Bie45] Irénée-Jules Bienaymé, *De la loi de multiplication et de la durée des familles*, Soc. Philomat. Paris Extraits, Sér **5** (1845), 37–39. 12

[Bil99] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999, A Wiley-Interscience Publication. MR 1700749 64, 69

[BK00] Jürgen Bennies and Götz Kersting, *A random walk approach to Galton-Watson trees*, J. Theoret. Probab. **13** (2000), no. 3, 777–803. MR 1785529 156

[BM14a] Olivier Bernardi and Alejandro H. Morales, *Counting trees using symmetries*, J. Combin. Theory Ser. A **123** (2014), 104–122. MR 3157803 132

[BM14b] Nicolas Broutin and Jean-François Marckert, *Asymptotics of trees with a prescribed degree sequence and applications*, Random Structures Algorithms **44** (2014), no. 3, 290–316. MR 3188597 12, 13, 14, 17, 20, 57, 64, 65, 98, 99, 122, 123, 125, 128, 134, 156

[BNB63] O. Barndorff-Nielsen and Glen Baxter, *Combinatorial lemmas in higher dimensions*, Trans. Amer. Math. Soc. **108** (1963), 313–325. MR 0156261 134

[BO18] Gabriel Hernán Berzunza Ojeda, *On scaling limits of multitype Galton-Watson trees with possibly infinite variance*, ALEA Lat. Am. J. Probab. Math. Stat. **15** (2018), no. 1, 21–48. MR 3748121 20

[Bre92] Leo Breiman, *Probability*, Classics in Applied Mathematics, vol. 7, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992, Corrected reprint of the 1968 original. MR 1163370 153

[BS07] Jinho Baik and Toufic M. Suidan, *Random matrix central limit theorems for nonintersecting random walks*, Ann. Probab. **35** (2007), no. 5, 1807–1834. MR 2349576 35

[CC08] Francesco Caravenna and Loï c Chaumont, *Invariance principles for random walks conditioned to stay positive*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), no. 1, 170–190. MR 2451576 30

[CD05] L. Chaumont and R. A. Doney, *On Lévy processes conditioned to stay positive*, Electron. J. Probab. **10** (2005), no. 28, 948–961. MR 2164035 7, 8, 30, 31, 34, 42

[CDS11] Sourav Chatterjee, Persi Diaconis, and Allan Sly, *Random graphs with a given degree sequence*, Ann. Appl. Probab. **21** (2011), no. 4, 1400–1435. MR 2857452 12, 20, 52

[Cha94] L. Chaumont, *Sur certains processus de Lévy conditionnés à rester positifs*, Stochastics Stochastics Rep. **47** (1994), no. 1-2, 1–20. MR 1787140 30

[Cha97] ______, *Excursion normalisée, méandre et pont pour les processus de Lévy stables*, Bull. Sci. Math. **121** (1997), no. 5, 377–403. MR 1465814 86

[CKB$^+$] A. Ciampi, L. Kates, R. Buick, Y. Kriukov, and J. E. Till, *Multi-type galton-watson process as a model for proliferating human tumour cell populations derived from stem cells: Estimation of stem cell self-renewal probabilities in human ovarian carcinomas*, Cell Proliferation **19**, no. 2, 129–140. 20, 98

[CL16] Loïc Chaumont and Rongli Liu, *Coding multitype forests: application to the law of the total population of branching forests*, Trans. Amer. Math. Soc. **368** (2016), no. 4, 2723–2747. MR 3449255 7, 20, 21, 22, 25, 98, 99, 101, 102, 105, 111, 112, 114, 132

[CM15] Arrigo Coen and Ramsés H. Mena, *Ruin probabilities for Bayesian exchangeable claims processes*, J. Statist. Plann. Inference **166** (2015), 102–115. MR 3390137 137

[CPGUB13] M. Emilia Caballero, José Luis Pérez Garmendia, and Gerónimo Uribe Bravo, *A Lamperti-type representation of continuous-state branching processes with immigration*, Ann. Probab. **41** (2013), no. 3A, 1585–1627. MR 3098685 19, 55, 66, 68, 69, 71, 78, 79, 90, 93, 94

[CPGUB17] M. Emilia Caballero, José Luis Pérez Garmendia, and Gerónimo Uribe Bravo, *Affine processes on $\mathbb{R}_+^m \times \mathbb{R}^n$ and multiparameter time changes*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 3, 1280–1304. MR 3689968 28, 136

[CUB11] Loïc Chaumont and Gerónimo Uribe Bravo, *Markovian bridges: weak continuity and pathwise constructions*, Ann. Probab. **39** (2011), no. 2, 609–647. MR 2789508 134

[CUB15] Loïc Chaumont and Gerónimo Uribe Bravo, *Shifting processes with cyclically exchangeable increments at random*, XI Symposium on Probability and Stochastic Processes, Progr. Probab., vol. 69, Birkhäuser/Springer, Cham, 2015, pp. 101–117. MR 3558138 8, 27, 28, 133, 135, 136, 147, 150

[Dev12] Luc Devroye, *Simulating size-constrained Galton-Watson trees*, SIAM J. Comput. **41** (2012), no. 1, 1–11. MR 2888318 8, 98, 100, 122, 125

[DG97] Michael Drmota and Bernhard Gittenberger, *On the profile of random trees*, Random Structures Algorithms **10** (1997), no. 4, 421–451. MR 1608230 7, 18, 52, 53, 61, 85

[DIM77] Richard T. Durrett, Donald L. Iglehart, and Douglas R. Miller, *Weak convergence to Brownian meander and Brownian excursion*, Ann. Probability **5** (1977), no. 1, 117–129. MR 0436353 8, 28, 133, 136

[DJ08] S. Dallaporta and A. Joffe, *The Q-process in a multitype branching processes*, Int. J. Pure Appl. Math. **42** (2008), no. 2, 235–240. MR 2383922 99

[DLG02] Thomas Duquesne and Jean-François Le Gall, *Random trees, Lévy processes and spatial branching processes*, Astérisque (2002), no. 281, vi+147. MR 1954248 12, 20, 98

[Don07] Ronald A. Doney, *Fluctuation theory for Lévy processes*, Lecture Notes in Mathematics, vol. 1897, Springer, Berlin, 2007, Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005, Edited and with a foreword by Jean Picard. MR 2320889 31, 34

[Drm09] Michael Drmota, *Random trees*, SpringerWienNewYork, Vienna, 2009, An interplay between combinatorics and probability. MR 2484382 12, 52, 122

[Duq03] Thomas Duquesne, *A limit theorem for the contour process of conditioned Galton-Watson trees*, Ann. Probab. **31** (2003), no. 2, 996–1027. MR 1964956 86

[Dur14a] Jetlir Duraj, *On harmonic functions of killed random walks in convex cones*, Electron. Commun. Probab. **19** (2014), no. 80, 10. MR 3283611 34, 50

[Dur14b] ______, *Random walks in cones: the case of nonzero drift*, Stochastic Process. Appl. **124** (2014), no. 4, 1503–1518. MR 3163211 9, 31

[DW10] Denis Denisov and Vitali Wachtel, *Conditional limit theorems for ordered random walks*, Electron. J. Probab. **15** (2010), no. 11, 292–322. MR 2609589 33, 50

[DW15] ______, *Random walks in cones*, Ann. Probab. **43** (2015), no. 3, 992–1044. MR 3342657 9, 31, 33, 34, 50

[Dwa69] Meyer Dwass, *The total progeny in a branching process and a related random walk.*, J. Appl. Probability **6** (1969), 682–686. MR 0253433 7, 20, 98, 99, 104, 127

[Dys62] Freeman J. Dyson, *A Brownian-motion model for the eigenvalues of a random matrix*, J. Mathematical Phys. **3** (1962), 1191–1198. MR 0148397 8, 30, 36

[EK86] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986, Characterization and convergence. MR 838085 55, 69, 79

[EK08] Peter Eichelsbacher and Wolfgang König, *Ordered random walks*, Electron. J. Probab. **13** (2008), no. 46, 1307–1336. MR 2430709 9, 30, 31, 33, 34

[EPW06] Steven N. Evans, Jim Pitman, and Anita Winter, *Rayleigh processes, real trees, and root growth with re-grafting*, Probab. Theory Related Fields **134** (2006), no. 1, 81–126. MR 2221786 12

[ER60] P. Erdős and A. Rényi, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **5** (1960), 17–61. MR 0125031 12, 99

[Fel71] William Feller, *An introduction to probability theory and its applications. Vol. II.*, Second edition, John Wiley & Sons Inc., New York, 1971. MR 0270403 91, 154

[FHN06] Michael Fuchs, Hsien-Kuei Hwang, and Ralph Neininger, *Profiles of random trees: limit theorems for random recursive trees and binary search trees*, Algorithmica **46** (2006), no. 3-4, 367–407. MR 2291961 52

[Fis84] Michael E. Fisher, *Walks, walls, wetting, and melting*, J. Statist. Phys. **34** (1984), no. 5-6, 667–729. MR 751710 8

[Fri72] Bert Fristedt, *Upper functions for symmetric processes with stationary, independent increments*, Indiana Univ. Math. J. **21** (1971/1972), 177–185. MR 0288841 27, 135, 138, 139, 146

[FS09] Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009. MR 2483235 52

[GK12] Bernhard Gittenberger and Veronika Kraus, *The degree profile of random Pólya trees*, J. Combin. Theory Ser. A **119** (2012), no. 7, 1528–1557. MR 2925941 52

[GLP16] Ion Grama, Ronan Lauvergnat, and Émile Le Page, *Limit theorems for affine markov walks conditioned to stay positive*, arXiv preprint arXiv:1601.02991 (2016). 30

[GMU18a] Jorge González Cázares, Aleksandar Mijatović, and Gerónimo Uribe Bravo, *Geometrically Convergent Simulation of the Extrema of L\'{e}vy Processes*, arXiv e-prints (2018), arXiv:1810.11039. 137

[GMU18b] Jorge Ignacio González Cázares, Aleksand ar Mijatović, and Gerónimo Uribe Bravo, *Exact Simulation of the Extrema of Stable Processes*, arXiv e-prints (2018), arXiv:1806.01870. 137

[GP75] G. L. Ghai and E. Pollak, *On some results for a bivariate branching process*, Biometrics **31** (1975), no. 3, 761–763. MR 0392022 20, 98

[GR16] Rodolphe Garbit and Kilian Raschel, *On the exit time from a cone for random walks with drift*, Rev. Mat. Iberoam. **32** (2016), no. 2, 511–532. MR 3512425 31, 51

[Gra99] David J. Grabiner, *Brownian motion in a Weyl chamber, non-colliding particles, and random matrices*, Ann. Inst. H. Poincaré Probab. Statist. **35** (1999), no. 2, 177–204. MR 1678525 35, 37

[Har63] Theodore E. Harris, *The theory of branching processes*, Die Grundlehren der Mathematischen Wissenschaften, Bd. 119, Springer-Verlag, Berlin; Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963. MR 0163361 20, 98

[Hir01] Katsuhiro Hirano, *Lévy processes with negative drift conditioned to stay positive*, Tokyo J. Math. **24** (2001), no. 1, 291–308. MR 1844435 30

[HJV07] Patsy Haccou, Peter Jagers, and Vladimir A. Vatutin, *Branching processes: variation, growth, and extinction of populations*, Cambridge Studies in Adaptive Dynamics, vol. 5, Cambridge University Press, Cambridge; IIASA, Laxenburg, 2007. MR 2429372 12

[HM12] Hamed Hatami and Michael Molloy, *The scaling window for a random graph with a given degree sequence*, Random Structures Algorithms **41** (2012), no. 1, 99–123. MR 2943428 20

[Ign18] I. Ignatiouk-Robert, *Harmonic functions of random walks in a semigroup via ladder heights*, ArXiv e-prints (2018). 9, 31, 34

[IK11] Minami Izumi and Makoto Katori, *Extreme value distributions of noncolliding diffusion processes*, Spectra of random operators and related topics, RIMS Kôkyûroku Bessatsu, B27, Res. Inst. Math. Sci. (RIMS), Kyoto, 2011, pp. 45–65. MR 2885254 35

[Jag75] Peter Jagers, *Branching processes with biological applications*, Wiley-Interscience [John Wiley & Sons], London-New York-Sydney, 1975, Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics. MR 0488341 12, 20, 98

[Jan06] Svante Janson, *Random cutting and records in deterministic and random trees*, Random Structures Algorithms **29** (2006), no. 2, 139–179. MR 2245498 57

[Jan12] ______, *Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation*, Probab. Surv. **9** (2012), 103–252. MR 2908619 89

[Joh00] Kurt Johansson, *Shape fluctuations and random matrices*, Comm. Math. Phys. **209** (2000), no. 2, 437–476. MR 1737991 35

[Joh02] ______, *Non-intersecting paths, random tilings and random matrices*, Probab. Theory Related Fields **123** (2002), no. 2, 225–280. MR 1900323 35

[Jos14] Adrien Joseph, *The component sizes of a critical random graph with given degree sequence*, Ann. Appl. Probab. **24** (2014), no. 6, 2560–2594. MR 3262511 20

[Joy81] André Joyal, *Une théorie combinatoire des séries formelles*, Adv. in Math. **42** (1981), no. 1, 1–82. MR 633783 52

[Jr58] Miloslav Ji ˇ rina, *Stochastic branching processes with continuous state space*, Czechoslovak Math. J. **8 (83)** (1958), 292–313. MR 0101554 93

[JY85] Th. Jeulin and M. Yor (eds.), *Grossissements de filtrations: exemples et applications*, Lecture Notes in Mathematics, vol. 1118, Springer-Verlag, Berlin, 1985, Papers from the seminar on stochastic calculus held at the Université de Paris VI, Paris, 1982/1983. MR 884713 72

[Kac54] M. Kac, *Toeplitz matrices, translation kernels and a related problem in probability theory*, Duke Math. J. **21** (1954), 501–509. MR 0062867 8, 29, 134

[Kal73] Olav Kallenberg, *Canonical representations and convergence criteria for processes with interchangeable increments*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **27** (1973), 23–36. MR 0394842 15, 55, 64, 133, 134

[Kal02] ______, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR MR1876169 47, 92, 139

[Kal05] ______, *Probabilistic symmetries and invariance principles*, Probability and its Applications (New York), Springer, New York, 2005. MR 2161313 135, 138, 147, 150

[Kee92] Robert W. Keener, *Limit theorems for random walks conditioned to stay positive*, Ann. Probab. **20** (1992), no. 2, 801–824. MR 1159575 30

[Ker11] Götz Kersting, *On the height profile of a conditioned galton-watson tree*, arXiv preprint arXiv:1101.3656 (2011). 7, 18, 52, 53, 61, 62, 65, 71, 85, 90

[Kes86] Harry Kesten, *Subdiffusive behavior of random walk on a random cluster*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 4, 425–487. MR 871905 94

[KL05] Andreas E. Kyprianou and R. Loeffen, *Lévy processes in finance distinguished by their coarse and fine path properties*, Exotic option pricing and advanced Lévy models, Wiley, Chichester, 2005, pp. 1–28. MR 2343206 28, 136

[KM59] Samuel Karlin and James McGregor, *Coincidence probabilities*, Pacific J. Math. **9** (1959), 1141–1164. MR 0114248 31, 37

[Kni96] F. B. Knight, *The uniform law for exchangeable and Lévy process bridges*, Astérisque (1996), no. 236, 171–188, Hommage à P. A. Meyer et J. Neveu. MR 1417982 28, 135, 144, 149, 150

[KNT04] Makoto Katori, Taro Nagao, and Hideki Tanemura, *Infinite systems of non-colliding Brownian particles*, Stochastic analysis on large scale interacting systems, Adv. Stud. Pure Math., vol. 39, Math. Soc. Japan, Tokyo, 2004, pp. 283–306. MR 2073337 35

[Knu98] Donald E. Knuth, *The art of computer programming. Vol. 3*, Addison-Wesley, Reading, MA, 1998, Sorting and searching, Second edition [of MR0445948]. MR 3077154 52

[Knu06] ———, *The art of computer programming. Vol. 4, Fasc. 4*, Addison-Wesley, Upper Saddle River, NJ, 2006, Generating all trees—history of combinatorial generation. MR 2251473 52

[KO01] Wolfgang König and Neil O'Connell, *Eigenvalues of the Laguerre process as non-colliding squared Bessel processes*, Electron. Comm. Probab. **6** (2001), 107–114 (electronic). MR 1871699 35

[Kön05] Wolfgang König, *Orthogonal polynomial ensembles in probability theory*, Probab. Surv. **2** (2005), 385–447. MR 2203677 8, 35

[KOR02] Wolfgang König, Neil O'Connell, and Sébastien Roch, *Non-colliding random walks, tandem queues, and discrete orthogonal polynomial ensembles*, Electron. J. Probab. **7** (2002), no. 5, 24 pp. (electronic). MR 1887625 30, 35

[Kor13] Igor Kortchemski, *A simple proof of Duquesne's theorem on contour processes of conditioned Galton-Watson trees*, Séminaire de Probabilités XLV, Lecture Notes in Math., vol. 2078, Springer, Cham, 2013, pp. 537–558. MR 3185928 154, 156

[Kor15a] I. Kortchemski, *Sub-exponential tail bounds for conditioned stable Bienaym\'e-Galton-Watson trees*, ArXiv e-prints (2015). 87

[Kor15b] Igor Kortchemski, *Limit theorems for conditioned non-generic Galton-Watson trees*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 2, 489–511. MR 3335012 89

[KS10] Wolfgang König and Patrick Schmid, *Random walks conditioned to stay in Weyl chambers of type C and D*, Electron. Commun. Probab. **15** (2010), 286–296. MR 2670195 35

[KT03a] Makoto Katori and Hideki Tanemura, *Functional central limit theorems for vicious walkers*, Stoch. Stoch. Rep. **75** (2003), no. 6, 369–390. MR 2029612 37

[KT03b] ———, *Noncolliding Brownian motions and Harish-Chandra formula*, Electron. Comm. Probab. **8** (2003), 112–121 (electronic). MR 2042750 35

[KT04] ______, *Symmetry of matrix-valued stochastic processes and noncolliding diffusion particle systems*, J. Math. Phys. **45** (2004), no. 8, 3058–3085. MR 2077500 35

[KVZ17a] Zakhar Kabluchko, Vladislav Vysotsky, and Dmitry Zaporozhets, *Convex hulls of random walks: expected number of faces and face probabilities*, Adv. Math. **320** (2017), 595–629. MR 3709116 134

[KVZ17b] ______, *Convex hulls of random walks, hyperplane arrangements, and Weyl chambers*, Geom. Funct. Anal. **27** (2017), no. 4, 880–918. MR 3678504 134

[KW71] Kiyoshi Kawazu and Shinzo Watanabe, *Branching processes with immigration and related limit theorems*, Teor. Verojatnost. i Primenen. **16** (1971), 34–51. MR 0290475 93

[Kyp14] Andreas E. Kyprianou, *Fluctuations of Lévy processes with applications*, second ed., Universitext, Springer, Heidelberg, 2014, Introductory lectures. MR 3155252 27, 136, 139

[Lam67a] John Lamperti, *Continuous state branching processes*, Bull. Amer. Math. Soc. **73** (1967), 382–386. MR 0208685 14, 55

[Lam67b] ______, *The limit of a sequence of branching processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **7** (1967), 271–288. MR 0217893 93

[Lam00] A. Lambert, *Completely asymmetric Lévy processes confined in a finite interval*, Ann. Inst. H. Poincaré Probab. Statist. **36** (2000), no. 2, 251–274. MR 1751660 35

[Lam02] Amaury Lambert, *The genealogy of continuous-state branching processes with immigration*, Probab. Theory Related Fields **122** (2002), no. 1, 42–70. MR 1883717 93

[Lam07] ______, *Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct*, Electron. J. Probab. **12** (2007), no. 14, 420–446. MR 2299923 90

[Law13] Averill M. Law, *Simulation modeling and analysis*, fifth ed., McGraw-Hill Education, 2013. 132

[Lei17] T. Lei, *Scaling limit of random forests with prescribed degree sequences*, ArXiv e-prints (2017). 17, 20, 99, 107

[LG05] Jean-François Le Gall, *Random trees and applications*, Probab. Surv. **2** (2005), 245–311. MR 2203728 13, 17, 63, 99

[LGLJ98] Jean-Francois Le Gall and Yves Le Jan, *Branching processes in Lévy processes: the exploration process*, Ann. Probab. **26** (1998), no. 1, 213–252. MR 1617047 12, 20, 98

[LW16] Karl Liechty and Dong Wang, *Nonintersecting Brownian motions on the unit circle*, Ann. Probab. **44** (2016), no. 2, 1134–1211. MR 3474469 35

[LW17] ______, *Nonintersecting Brownian bridges between reflecting or absorbing walls*, Adv. Math. **309** (2017), 155–208. MR 3607275 35

[Mar19] Cyril Marzouk, *On the growth of random planar maps with a prescribed degree sequence*, arXiv e-prints (2019), arXiv:1902.04539. 20

[Mie08] Grégory Miermont, *Invariance principles for spatial multitype Galton-Watson trees*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), no. 6, 1128–1161. MR 2469338 20, 98, 99

[Mil77] P. W. Millar, *Zero-one laws and the minimum of a Markov process*, Trans. Amer. Math. Soc. **226** (1977), 365–391. MR 0433606 8, 27, 28, 133, 135

[MR95] Michael Molloy and Bruce Reed, *A critical point for random graphs with a given degree sequence*, Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993), vol. 6, 1995, pp. 161–179. MR 1370952 20

[MR98] ______, *The size of the giant component of a random graph with a given degree sequence*, Combin. Probab. Comput. **7** (1998), no. 3, 295–305. MR 1664335 20

[MSM94] Servet Martínez and Jaime San Martín, *Quasi-stationary distributions for a Brownian motion with drift and associated limit laws*, J. Appl. Probab. **31** (1994), no. 4, 911–920. MR 1303922 34

[Nak78] Tetsuo Nakagawa, *The Q-process associated with a multitype Galton-Watson process and the additional results*, Bull. Gen. Ed. Dokkyo Univ. School Medicine **1** (1978), 21–32. MR 653216 20, 99

[Nev86] J. Neveu, *Arbres et processus de Galton-Watson*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 199–207. MR 850756 17, 53

[O'C03] Neil O'Connell, *Random matrices, non-colliding processes and queues*, Séminaire de Probabilités, XXXVI, Lecture Notes in Math., vol. 1801, Springer, Berlin, 2003, pp. 165–182. MR 1971584 35

[O'C12] ______, *Directed polymers and the quantum Toda lattice*, Ann. Probab. **40** (2012), no. 2, 437–458. MR 2952082 35

[Ott49] Richard Otter, *The multiplicative process*, Ann. Math. Statistics **20** (1949), 206–224. MR 0030716 7, 20, 98, 99, 104, 127

[OY02] Neil O'Connell and Marc Yor, *A representation for non-colliding random walks*, Electron. Comm. Probab. **7** (2002), 1–12 (electronic). MR 1887169 36

[PÍ1] Sophie Pénisson, *Continuous-time multitype branching processes conditioned on very late extinction*, ESAIM Probab. Stat. **15** (2011), 417–442. MR 2870524 99

[PÍ6] Sophie Pénisson, *Beyond the Q-process: various ways of conditioning the multitype Galton-Watson process*, ALEA Lat. Am. J. Probab. Math. Stat. **13** (2016), no. 1, 223–237. MR 3476213 20, 99

[Pak03] Anthony G. Pakes, *Biological applications of branching processes*, Stochastic processes: modelling and simulation, Handbook of Statist., vol. 21, North-Holland, Amsterdam, 2003, pp. 693–773. MR 1973557 12

[Par08] J. C. Pardo, *On the rate of growth of Lévy processes with no positive jumps conditioned to stay positive*, Electron. Commun. Probab. **13** (2008), 494–506. MR 2447836 30

[Pit95] Jim Pitman, *Exchangeable and partially exchangeable random partitions*, Probab. Theory Related Fields **102** (1995), no. 2, 145–158. MR 1337249 152

[Pit98] ______, *Enumerations of trees and forests related to branching processes and random walks*, Microsurveys in discrete probability (Princeton, NJ, 1997), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 41, Amer. Math. Soc., Providence, RI, 1998, pp. 163–180. MR 1630413 12, 25, 52, 100, 105, 118, 119, 120

[Pit06] J. Pitman, *Combinatorial stochastic processes*, Lecture Notes in Mathematics, vol. 1875, Springer-Verlag, Berlin, 2006, Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. MR 2245368 13, 63, 101, 107, 152, 156

[PR69] E. A. Pečerskiĭ and B. A. Rogozin, *The combined distributions of the random variables connected with the fluctuations of a process with independent increments*, Teor. Verojatnost. i Primenen. **14** (1969), 431–444. MR 0260005 136

[PR04] Zbigniew Palmowski and Tomasz Rolski, *Markov processes conditioned to never exit a subspace of the state space*, Probab. Math. Statist. **24** (2004), no. 2, Acta Univ. Wratislav. No. 2732, 339–353. MR 2157210 35

[PUB12] Jim Pitman and Gerónimo Uribe Bravo, *The convex minorant of a Lévy process*, Ann. Probab. **40** (2012), no. 4, 1636–1674. 8, 28, 29, 133, 136, 150, 151

[PV91] Ileana Popescu and Ion Văduva, *A survey on computer generation of some classes of stochastic processes*, Math. Comput. Simulation **33** (1991), no. 3, 223–241. MR 1136990 99

[Ras14] Kilian Raschel, *Random walks in the quarter plane, discrete harmonic functions and conformal mappings*, Stochastic Process. Appl. **124** (2014), no. 10, 3147–3178, With an appendix by Sandro Franceschi. MR 3231615 37

[RFW17] Julien Randon-Furling and Florian Wespi, *Facets on the convex hull of d-dimensional Brownian and Lévy motion*, Phys. Rev. E **95** (2017), no. 3, 032129, 5. MR 3782474 134

[Rog68] B. A. Rogozin, *The local behavior of processes with independent increments*, Teor. Verojatnost. i Primenen. **13** (1968), 507–512. MR 0242261 8, 27, 133, 135

[San71] David Sankoff, *Branching processes with terminal types. Application to context-free grammars*, J. Appl. Probability **8** (1971), 233–240. MR 0286196 20, 98

[Sat99] Ken-iti Sato, *Lévy processes and infinitely divisible distributions*, Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 1999, Translated from the 1990 Japanese original, Revised by the author. MR 1739520 27, 135, 138, 139

[Saw97] Stanley A. Sawyer, *Martin boundaries and random walks*, Harmonic functions on trees and buildings (New York, 1995), Contemp. Math., vol. 206, Amer. Math. Soc., Providence, RI, 1997, pp. 17–44. MR 1463727 31

[Sil68] M. L. Silverstein, *A new approach to local times*, J. Math. Mech. **17** (1967/1968), 1023–1054. MR 0226734 93

[Spi56] Frank Spitzer, *A combinatorial lemma and its application to probability theory*, Trans. Amer. Math. Soc. **82** (1956), 323–339. MR 0079851 134

[Ste30] Johan Frederik Steffensen, *Om sandsynligheden for at afkommet uddør*, Matematisk tidsskrift. B (1930), 19–23. 12

[Şte98] Cătălina Ştefănescu, *Simulation of a multitype galton-watson chain*, Simulation Practice and Theory **6** (1998), no. 7, 657 – 663. 99

[Ste18] Robin Stephenson, *Local convergence of large critical multi-type Galton-Watson trees and applications to random maps*, J. Theoret. Probab. **31** (2018), no. 1, 159–205. MR 3769811 20, 99

[Tan89] Hiroshi Tanaka, *Time reversal of random walks in one-dimension*, Tokyo J. Math. **12** (1989), no. 1, 159–174. MR 1001739 7, 8, 30, 43

[Tan04] ______, *Lévy processes conditioned to stay positive and diffusions in random environments*, Stochastic analysis on large scale interacting systems, Adv. Stud. Pure Math., vol. 39, Math. Soc. Japan, Tokyo, 2004, pp. 355–376. MR 2073341 30

[TW07] Craig A. Tracy and Harold Widom, *Nonintersecting Brownian excursions*, Ann. Appl. Probab. **17** (2007), no. 3, 953–979. MR 2326237 35

[UB14] Gerónimo Uribe Bravo, *Bridges of Lévy processes conditioned to stay positive*, Bernoulli **20** (2014), no. 1, 190–206. MR 3160578 8, 133, 134

[Ver79] Wim Vervaat, *A relation between Brownian bridge and Brownian excursion*, Ann. Probab. **7** (1979), no. 1, 143–149. MR MR515820 7, 14, 20, 64, 98, 99

[Vig02] Vincent Vigon, *Lévy processes and Wiener-hopf factorization*, Theses, INSA de Rouen, April 2002. 27, 135

[VL05] Fabien Viger and Matthieu Latapy, *Efficient and simple generation of random simple connected graphs with prescribed degree sequence*, Computing and combinatorics, Lecture Notes in Comput. Sci., vol. 3595, Springer, Berlin, 2005, pp. 440–449. MR 2190867 20

[VZ18] Vladislav Vysotsky and Dmitry Zaporozhets, *Convex hulls of multidimensional random walks*, Trans. Amer. Math. Soc. **370** (2018), no. 11, 7985–8012. MR 3852455 134

[Wan14] Hua Ming Wang, *On total progeny of multitype Galton-Watson process and the first passage time of random walk on lattice*, Acta Math. Sin. (Engl. Ser.) **30** (2014), no. 12, 2161–2172. MR 3285942 99

[Wat71] Takesi Watanabe, *An integro-differential equation for a compound Poisson process with drift and the integral equation of H. Cramer*, Osaka J. Math. **8** (1971), 377–383. MR 0303602 27, 135

[WG75] Henry William Watson and Francis Galton, *On the probability of the extinction of families*, The Journal of the Anthropological Institute of Great Britain and Ireland **4** (1875), 138–144. 12

[Whi80] Ward Whitt, *Some useful functions for functional limit theorems*, Mathematics of Operations Research (1980). 69, 71, 79

[Wu08] Biao Wu, *On the weak convergence of subordinated systems*, Statist. Probab. Lett. **78** (2008), no. 18, 3203–3211. MR 2479479 71

[Zol86] V. M. Zolotarev, *One-dimensional stable distributions*, Translations of Mathematical Monographs, vol. 65, American Mathematical Society, Providence, RI, 1986, Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver. MR 854867 95