Rapport de stage 3A

University of Bath

Gabriel Athènes

June 27, 2022

Contents

Plagiarism statement

I, Gabriel Athenes, certify that this report is my own work, based on my personal study and research and that I have acknowledged all material and sources used in its preparation [1] [2] [3] [4] [5] [6] [7].

Appreciation

I would like to thank Cecile Mailler for proposing the subject of this internship to me. Cécile was always available to help and direct me on the best avenues of research that suited my advancements. The weekly 1 hour zoom meeting that we had enabled me to stay motivated during the whole internship. She was always nice and encouraging and made me understand the philosophy of a researcher in aspects that go beyong work or Mathematics.

1 Introduction

During my internship at the University of Bath from April to July 2021, I studied the local and scaling limit of the critical Galton-Watson Tree (GWT), in fixed, random, and varying environments.

1.1 Introduction of the formalism

We introduce in this subsection the formalism needed to understand the mathematics of each section of this report. We first begin by introducing some notations. The formalism used is the same as the one used in Kersting and vatutin's book [4].

Let $\mathcal{P}(\mathbb{N})$ be the set of probability measures on $\mathbb{N}$. $v = f_1, ...f_\infty$ is called a **varying environment**. Given a varying environment, let Z_n be a Markov process defined by $Z_0 = 1$ and

$$Z_n = \begin{cases} \sum_{i=1}^{Z_{n-1}} & \text{if } Z_{n-1} > 0 \\ 0 & \text{otherwise} \end{cases}$$

Then Z_n is called a Galton-Watson process in a varying environment. $\overline{f_n} := \sum_{k=0}^{\infty} f_n[k]$ denotes the mean of the distribution f_n. The case in which $f_n = f \ \forall n \in \mathbb{N}$ is the simplest GW process and is called GW process with offspring distribution f. A critical GW process corresponds to the case in which $\overline{f} = 1$. We use the notation $\mathcal{P}_v$ for probability measures in a varying environment v.

The other kind of environment we deal with in this report is the **random environment**. If we endow the space $\mathcal{P}(\mathbb{N})$ of a Borel-σ-algebra, we can consider random probability measures on $\mathbb{N}$ i.e random variables F taking values in $\mathcal{P}(\mathbb{N})$. A sequence $\mathcal{V} = \{F_1, ..., F_\infty\}$ where all F_i are random probability measures is a random environment. We use the notation $\mathbb{P}$ for probability measures in random environment. Now a GW process in random environment is a process such as once a random environment $\mathcal{V} = \{F_1, ..., F_\infty\}$ is simulated, it evolves in its value $v = \{f_1, ..., f_\infty\}$: $\mathbb{P}(Z_n = z_n) = \mathcal{P}(Z_n = z_n | \mathcal{V} = v) \ \mathbb{P} - a.s.$

Informally, a Galton-Watson process can be associated to a tree. Indeed Z_{n-1} represents the number of individuals at generation $n-1$, and $Y_{i,n}$ the number of children of each individual $i \in [\![1, Z_{n-1}]\!]$ (whose sum gives the number of individuals at generation n). We can therefore associate a tree to this process by attributing nodes to individuals and drawing arcs when one individual is the child of another.

In Section 2, we discuss about probabilities of such trees in the case of a given varying environment $v = f_1, ...f_\infty$. We therefore need to define the probability space $(\mathcal{T}, \mathcal{F}, \mathcal{P})$ in which the random Tree T lives. We therefore define the sample space $\mathcal{T}$, event space $\mathcal{F}$ and probability function $\mathcal{P}$ defining the distribution of T.

First, let's define the sample space, which is a set of trees. In order to define a tree, we first need to define the individuals that constitute the tree. For this, we use the Ulam-Harris labelling. Individuals are elements i of $\mathcal{I} := \bigcup_{n=0}^{+\infty} \mathbb{N}^n$, with $\mathbb{N}^0$ being $\emptyset$. $i \in \mathcal{I}$ is written as a string $j_1...j_n$.

Then a tree is a subset of $\mathcal{I}$ satisfying 4 criteria :

1. a tree has a root : $\emptyset \in t$,
2. an individual that is not a root has a parent in t: $i \in t, j \in \mathbb{N}$ *and* $ij \in t \Rightarrow i \in t$
3. If an individual i of t has y children they are labelled $i1, ..., iy$: for $i \in \mathcal{I},\ j \in \mathbb{N}$ and $ij \in t,\ ij' \in t$ for $1 \leq j' \leq j$
4. individuals have a finite number of children : for $i \in t,\ \exists j \in \mathbb{N}\ s.t\ ij \notin t$

The sample space $\mathcal{T}$ is the set of all such trees.

Second, let's define the event space. In order to endow the event space of a sigma-field, $\mathcal{F}$, we introduce

$$\begin{array}{cccc} g & : & \mathcal{I} & \rightarrow & \mathbb{N} \\ & & j_1...j_n & \mapsto & n \end{array}$$

the function returning the generation of an individual, as well as

$$\begin{array}{ccccc} h & : & \mathcal{T} & \rightarrow & \mathbb{N} \\ & & t & \mapsto & \max\{g(i), i \in t\} \end{array}$$

the function returning the height of a tree. Now, $t \overset{h}{=} t'$ if they coincide up to height h. We can now endow $\mathcal{T}$ of the sigma-field of all sets of the form $\{t' \in \mathcal{T}, t' \overset{h}{=} t\}$ for $h \geq 0$ and $t \in \mathcal{T}$ to create an event space.

Finally, let $y : \mathcal{I} \rightarrow \mathbb{N}$ be the function giving the number of children of an individual. The random tree T is a measurable function taking values in $\mathcal{T}$ and with distribution $P_v(T \overset{h}{=} t)$=$\prod_{i \in t, g(i) < h} f_{g(i)+1}[y(i)]$ for all $h \geq 0$ and $t \in \mathcal{T}$. We notice that the generation sizes $Z_n = \#\{i \in T, g(i) = n\}$ gives us the Galton-Walton process previously defined.

In Section 3 and 4, we talk about different encodings of Galton Watson processes. What we call an encoding is a function associating to each individual a real value. More precisely, given a tree t of n individuals, we provide to the individuals of the tree an order from 1 to n, and we attribute to each of these integers a value. In the two encodings we use in Section 3 and 4, the ordering used is the Depth-First ordering.

1.2 Plan of the report

The first kind of limit I study in Section 2 of this report is the Local Limit. The Local Limit looks at the behaviour of the tree up to a finite height. We show in this part that up to a height n, the probability of a critical GWT is in fact the deformation of the probability of a GWT conditionned on having infinite height, called Geiger-Tree or Size-Biaised-Tree. We also give some applications of this relation, particularly in order to understand the behaviour of a martingale associated to the critical GWT.

The other kind of limit I study in Sections 3 and 4 is the Scaling Limit. This limit looks at the GWT in which the size of the branches are rescaled by the square root of the total number of nodes. This way, "seen from afar", the tree remains in a bounded compact and converges towards a fractal object called the Continuous Random Tree (CRT).

Results already exist in the case of an environment in which the offspring distribution is constant. In Section 3, we synthesize different articles treating this case in order to exhibit clearly the different steps leading to the convergence of the GWT to the CRT. Explaining these different steps highlights the key results needed in order to prove convergence in a varying envionment and on which we focus in Section 4. Indeed, convergence in a varying environment has not been subject to much research, and so one of the key interests of my internship was to help researcher Cécile Mailler investigate this topic. In Section 3, we first study the limit of different encodings of sequences of critical GWT. We then define a distance between trees as well as the limit object of GWT according to this distance (the Continuous Random Tree), and show, based on studied articles, how the limit of these encodings enable us to prove the convergence of the critical GWT towards the CRT.

Finally, in Section 4, we study the scaling limits of a GWT in a varying environment, in which the progeny laws have been drawn with some assumptions. Since Section 3 dissects the different steps necessary for the convergence of the GWT, we therefore study the same steps but under the new assumptions implied by a varying environment.

2 Local limit

In this section, we first define the size biased tree associated to an offspring distribution ν. We then show that this size biased tree is the local limit of the GW tree of offspring distribution conditioned to survive. This subsection is very inspired of the section 1.4 of Kersting and vatutin's book [4].

2.1 Size Biaised Tree

A tree t (resp a random variable T) can be divided into subtrees $t_1, ..., t_y$ (resp $T_1, ..., T_y$) growing from its children $1, ..., y$ (resp $1, ..., Y$). Then noting v_1 the shifted environment $f_2, ..., f_\infty$,

$$\mathcal{P}_v(T \overset{n}{=} t) = f_1[y] \prod_{j=1}^{y} \mathcal{P}_{v_1}(T_j \overset{n-1}{=} t_j).$$

We now write $\mathcal{P}_v(T \overset{n}{=} t|Z_n > 0)$ using the structure of the trees stemming from the children of the root, in order to construct the size biaised tree:

$$\mathcal{P}_v(T \overset{n}{=} t|Z_n > 0) = \frac{\mathcal{P}_v(T \overset{n}{=} t, Z_n > 0)}{\mathcal{P}(Z_n > 0)} = \frac{\mathcal{P}_v(T \overset{n}{=} t,\ t \text{ has height at least } n)}{\mathcal{P}(Z_n > 0)}.$$

If t is of height at least n, we let $d = d(t)$ denote the largest integer such that for all $i \leq d$, t_i has height at least $n-1$, and on the right the trees are unconditionned. Therefore,

$$\begin{aligned}
\mathcal{P}_v(T \overset{n}{=} t|Z_n > 0) &= \frac{f_1[y]}{\mathcal{P}(Z_n > 0)} \prod_{j=1}^{d-1} \mathcal{P}_{v_1}(T_j = t_j) \mathcal{P}_{v_1}(T_d \overset{n-1}{=} t_d) \prod_{j=d+1}^{y} \mathcal{P}_{v_1}(T_j \overset{n-1}{=} t_j) \\
&= \left(\frac{f_1[y]}{1 - f_{0,n}[0]} (1 - f_{1,n}[0]^{d-1}) \right) \times \left(\prod_{j=1}^{d-1} \mathcal{P}_{v_1}(T_j = t_j | Z_{n-1} = 0) \right) \\
&\times \left(\mathcal{P}_{v_1}(T_d \overset{n-1}{=} t_d | Z_{n-1} > 0) \right) \times \left(\prod_{j=d+1}^{y} \mathcal{P}_{v_1}(T_j \overset{n-1}{=} t_j) \right)
\end{aligned}$$

and so we have $\mathcal{P}_v(T \overset{n}{=} t|Z_n > 0) = g_n[d, y] \times p_{d-1} \times p_d \times p_{d+1}$ with $g_n[d, y] = \frac{f_1[y]}{1-f_{0,n}[0]}(1 - f_{1,n}[0]^{d-1})$.

We can see that in order to simulate a tree T equal to t up to height at least n and conditionned on having height n, we can execute Algo 1 :

1. Simulate a variable (D, Y) with probability $g_n[d, y]$
2. simulate $T_1, ..., T_y$ trees where $T_1, ..., T_{D-1}$ are conditionned on having heights $< n-1$, T_D has height at least $n-1$, and the T_{D+i} are unconditionned.

Since T_D is conditionned on having height at least $n-1$, we can reiter the process, this time in environment $f_2, ..., f_\infty$ meaning that (D, Y) will be simulated according to law $\frac{1-f_{2,n}[0]}{1-f_{1,n}[0]} \times f_2[y] \times f_{2,n}[0]^{d-1}$ which we call $g_{2,n}$ with $g_{m,n}[d, y] = \frac{1-f_{m,n}[0]}{f_m(1)-f_m(f_{m,n}[0])} \times f_m[y] \times f_{m,n}[0]^{d-1}$.

We can now simulate T by executing Algo 1 with $T_{D_1}, ...T_{D_n}$, simulating $(Y_i, D_i)_{i=1,...,n}$

and then simulating conditionned processes on the left, and unconditionned processes on the right. Now, under the condition that $Z_n \underset{n\to+\infty}{\to} 0$ (the underlying process goes extinct), $\mathcal{P}_{v_m}(T \overset{h}{=} t | Z_n = 0) \underset{n\to+\infty}{\to} \mathcal{P}_{v_m}(T \overset{h}{=} t)$.

In other words, the trees on the left get unconditionned in the limit. We can now build a tree conditionned on having infinite height in another way. Indeed $g_{m,n}[d,y] = \frac{1-f_{m,n}[0]}{f_m(1)-f_m(f_{m,n}[0])} \times f_m[y] \times f_{m,n}[0]^{d-1} \underset{n\to+\infty}{\to} \frac{f_m[y]}{\overline{f_m}}$ because $\frac{1-f_{m,n}[0]}{f_m(1)-f_m(f_{m,n}[0])} \underset{n\to+\infty}{\to} \frac{1}{f_m'} = \frac{1}{\overline{f_m}}$ and $f_{m,n}[0]^{d-1} \underset{n\to+\infty}{\to} 1$

Now, D_m, Y_m is obtained by simulating Y_m which has law $\frac{f_m[y]}{\overline{f_m}} y$ and by simulating D_m uniformly. **We call T^* this tree, also called Geiger Tree or Size Biaised Tree.** Thanks to Gotz Kersting and Vladimir Vatutin, we have a lemma uniting the behaviour of T^* and T, the size biaised and non biaised GWT.

Lemma 2.1. *Let t be a tree of a height at least $n \geq 1$ and let $d \in t$ be an individual in generation n. Then for a tree T and a size-biaised tree T^*, both in the same environment $v = f_1..., f_\infty$,*

$$\mathcal{P}(T^* \overset{n}{=} t, \Delta_n = d) = \frac{1}{\overline{f_1}, ..., \overline{f_n}} \mathcal{P}(T \overset{n}{=} t)$$

or equivalently,

$$\mathcal{P}(T^* \overset{n}{=} t) = \frac{z_n(t)}{\overline{f_1}, ..., \overline{f_n}} \mathcal{P}(T \overset{n}{=} t) = W_n(t) \mathcal{P}(T \overset{n}{=} t)$$

Proof. By induction, $\mathcal{P}(T^* \overset{1}{=} t, \Delta_1 = d) = \frac{f_1[y]}{\overline{f_1}} = \frac{\mathcal{P}(T \overset{1}{=} t)}{\overline{f_1}}$

$$\begin{aligned} \textit{Then, } \mathcal{P}(T^* \overset{n+1}{=} t, \Delta_{n+1} = d) &= \mathcal{P}(T^* \overset{n}{=} t, \Delta_n = d) \frac{f_{n+1}[y(d')]}{\overline{f_{n+1}}} \prod_{i \in t, i \neq d', g(i) = n} f_{n+1}[y(i)] \\ &= \frac{1}{\overline{f_1} ... \overline{f_{n+1}}} \mathcal{P}(T \overset{n+1}{=} t) \end{aligned}$$

□

By noting $\mathcal{P}_{T^*}(A) := \mathcal{P}(T^* \overset{n}{\in} A)$ and respectively $\mathcal{P}_T(A) := \mathcal{P}(T \overset{n}{\in} A)$, this lemma means that $\mathcal{P}_{T^*}(A) = \int_A W_n(t) dP_T(t) = E_T(1_A W_n)$.

Reciprocitly, let's suppose that there is extinction with probability 1. Then the Size-Biaised GWT exists. In all that follows, let $\mathcal{F}_n$ be the σ-fields generated by the equivalence classes related to the equivalence relations $\overset{n}{=}$, i.e the coincidence of trees up to height n.

Theorem 2.2. *Suppose that $Z_n \xrightarrow[n\to\infty]{} 0$ with probability 1, define $\forall n \in \mathbb{N}$, $\mathcal{P}^*_{\mathcal{F}_n} := W_n \mathcal{P}_{\mathcal{F}_n}$ and let $A_1, ..., A_y$ be set of trees, with $y \geq 1$. Then*

$$\mathcal{P}^*(Y = y, T_1 \in A_1, ..., T_y \in A_y) = \frac{y f_1[y]}{\overline{f_1}} \times \frac{1}{y} \times \sum_{i=1}^{y} \mathcal{P}(\mathcal{A}_\infty) ... \mathcal{P}(A_{i-1}) \mathcal{P}^*(A_i) \mathcal{P}(A_{i+1}) ... \mathcal{P}(A_y)$$

In other words, the root has probability $\frac{yf_1[y]}{\overline{f_1}}$.
Furthermore, we can see that $\mathcal{P}^*(Y = y, T_1 \in A_1, ..., T_y \in A_y)$ is stable by permutation of the A_i, which means that $\mathcal{P}^*(Y = y, T_i \in B) = \mathcal{P}^*(Y = y, T_j \in B)$ and that the vertex from which the subtree rooted at is size biaised is chosen uniformly. Therefore $\mathcal{P}^*$ is unique and $\mathcal{P}^*(A) = \mathcal{P}(T^* \in A)$

Proof. Assume $A_1, ..., A_n$ have height $\leq n$.

$$\begin{aligned}
\mathcal{P}^*(Y = y, T_1 \in A_1, ..., T_y \in A_y) &= E(\frac{Z_n}{\overline{f_1}...\overline{f_n}} \mathbb{1}_{\{Y=y,T_1\in A_1,...,T_y\in A_y\}}) \\
&= \frac{1}{\overline{f_1}...\overline{f_n}} E(\sum_{i=1}^{y} Z_{n-1}^{(i)} \cdot \mathbb{1}_{\{Y=y\}} \cdot \mathbb{1}_{\{Y=y,T_1\in A_1,...,T_y\in A_y\}}) \\
&= \frac{\mathcal{P}(Y = y)}{\overline{f_1}...\overline{f_n}} \cdot \sum_{i=1}^{y} E(Z_{n-1}^{(i)} \cdot \mathbb{1}_{\{T_1\in A_1,...,T_n\in A_n\}}) \\
&= \frac{\mathcal{P}(Y = y)}{\overline{f_1}...\overline{f_n}} \sum_{i=1}^{y} E(Z_{n-1}^{(i)} \cdot \mathbb{1}_{T_i\in A_i}) \prod_{j\neq i} P(A_j) \\
&= \frac{f_1[y]}{\overline{f_1}} \sum_{i=1}^{y} \mathcal{P}^*(A_i) \prod_{j\neq i} \mathcal{P}(A_j)
\end{aligned}$$

$\square$

2.2 Application

We notice that $\mathcal{P}(T^* \overset{n}{=} t) = W_n(t)\mathcal{P}(T \overset{n}{=} t) \Rightarrow \mathbb{P}(T^* \overset{n}{=} t|\mathcal{V}) = W_n(t)\mathbb{P}(T \overset{n}{=} t|\mathcal{V})\ \mathcal{V} - a.s.$ Therefore Lemma 2.1 still holds in random environment as by integrating,

$$\mathbb{P}(T^* \overset{n}{=} t) = \mathbb{E}_{\mathcal{V}}(\mathbb{P}(T^* \overset{n}{=} t|\mathcal{V})) = \mathbb{E}_{\mathcal{V}}(\mathbb{P}(T \overset{n}{=} t|\mathcal{V})W_n(t)) = \mathbb{P}(T \overset{n}{=} t)$$

. Therefore $\mathbb{P}^*_{\mathcal{F}_n} = W_n\mathbb{P}_{\mathcal{F}_n}$. $\mathbb{P}^*$ is a good tool to analyse the limit probability of $\mathbb{P}$. Indeed, $\mathbb{P}^*$ is a deformation of $\mathbb{P}$, and the information of the deformation is contained in W_n. Let's consider the two cases, extinction and non extinction.

Extinction case : In this case, the size biaised tree exists, but it does not give us any information on $\mathbb{P}$. Indeed, $\mathbb{P}^*$ is concentrated on a set that goes to 0 almost surely when $n \to +\infty$. Since W_n is a deformation of weight 1, $W_n \to +\infty$ on this set and $\mathbb{P}^*$ takes all its value on a set that is insignificant to $\mathbb{P}$. It's normal, T^* is infinite size and $Z_n \to 0$ almost surely, meaning that $\mathbb{P}$ and $\mathcal{P}^*$ have very little in common. $\mathbb{P}^*$ cannot be used to infer anything on $\mathbb{P}$.

Non extinction case : $\mathbb{P}(W_n > 0) = \mathbb{P}(Z_n = 0) > 0\ \forall n \in \mathbb{N}$. $\mathbb{P}^*$ is $\mathbb{P}$ concentrated on $W_n > 0$ on each $\sigma - field\ \mathcal{F}_n$. If W>0 with probability >0, this gives information on $\mathbb{P}$ since $\mathbb{P} = \mathbb{1}_{W>0} \cdot \frac{1}{W}\mathbb{P}^*$ and $\mathbb{P}^*$ is the probability of a Geiger tree. It's also normal, since the population goes to $+\infty$ with probability >0, the pseudo geiger tree and the Galton Walton are more correlated. Let's sum up this by the following proposition.

Proposition 2.1. *Let (F_n) be a filtration, and let $\mathcal{P}$ and $\hat{\mathcal{P}}$ be probabilities on (Ω, F_∞). Assume that for any n, $\hat{\mathcal{P}}_{F_n}$ is absolutely continuous with respect to $\mathcal{P}_{F_n}$. Let $W_n := \frac{\hat{\mathcal{P}}_{F_n}}{\mathcal{P}_{F_n}}$,* *and let* $W := \underset{n \to \infty}{lim\ sup}\, W_n$.

$$\hat{\mathcal{P}} << \mathcal{P} \iff W < 0,\ \hat{\mathcal{P}} - almost\ surely \iff E(W) = 1$$

$$\hat{\mathcal{P}} \perp \mathcal{P} \iff W = 0,\ \hat{\mathcal{P}} - almost\ surely \iff E(W) = 0$$

We now understand the whole interest of $\mathbb{P}^*$, which is to deduce properties on W. Indeed, theorem 2.4 of [4] uses this tool to show that $\mathbb{E}(\frac{Z_1 log^+(Z_1)}{\overline{F_1}})$to caracterize W.

Theorem 2.3 (Theorem 2.4 of Vatutin's book)**.** *Assume that* $0 < \mathbb{E}(X) < +\infty$ *and* $q < 1$*. Then, the following conditions are equivalent :* $\mathbb{P}(W = 0) = q \iff \mathbb{E}(W) = 1 \iff \mathbb{E}(\frac{Z_1 log^+(Z_1)}{\overline{F_1}})$

The idea of the proof uses $\mathbb{P}^*$ by proving that $E(\frac{Z_1 log^+(Z_1)}{\overline{F_1}}) = +\infty \Rightarrow W = +\infty,\ \mathbb{P}^* - almost\ surely \Rightarrow \mathbb{E}(W) = 0$ and that $\mathbb{E}(\frac{Z_1 log^+(Z_1)}{\overline{F_1}}) < +\infty \Rightarrow W < +\infty,\ \mathbb{P}^* - almost\ surely \Rightarrow \mathbb{E}(W) = 1$

3 Scaling Limits in the case of a critical GW tree with constant offspring distribution

In this section, we study the convergence of a critical GW tree with offspring distribution given by $(p_i)_{i\geq 0}$. The plan of this section is as follows : Subsection 3.2 defines the Random Continuous Tree towards which the critical GWT converges, as well as a distance in which the convergence takes place. In order to prove the convergence, a necessary result is the convergence of what we call the conditionned height encoding of a GWT. We therefore first work with limits of encodings in Subsection 3. More precisely, using the existing litterature we show that the convergence of the Lukasiewicz encoding implies the convergence of the unconditionned Height encoding 3.4 which implies the convergence of the conditionned Height encoding using a variant of Donsker's theorem [3].

3.1 Limits of different encodings of a critical Galton Watson tree

In this subsection, we focus our research on scaling limits of different encodings of GWTs, an encoding being a sequence of integers describing a unique GWT. As the GWT is a random variable so is the encoding. The two encodings we study are the Lukasiewicz encoding, that we call X, and the Height encoding, that we call H. Let T be a random critical Galton Watson tree with offspring distribution $(p_i)_{i\geq 0}$. We order the individuals of T according to the Depth-First ordering. The nodes are now labelled by integers from 1 to the size of T.

The Lukasiewicz encoding is defined by $X(i) = \sum_{j=1}^{i}(Y(j)-1)$; $X(0) = 0$. At each step, we add the number of children-1 of the i-th node. If T has n nodes, we therefore have $X(n) = -1$ since $total\ number\ of\ children - n = -1$.

The Height encoding H simply associates i to it's generation in the tree, which is geometrically the height of the node in the tree. According to the notations of the introduction 1.1 we should call this process G, as h is used to talk about the height of a tree, but I kept the notation of the majority of articles I read as it is more intuitive geometrically.

3.1.1 Lukasiewicz encoding

We begin with an observation that will be very useful in what follows.

Theorem 3.1. *Let t be a tree. X encodes t in the sense that $H(k) = Card(\{0 \leq i \leq k-1 : X(i) = \min_{i\leq j\leq k} X(j)\})$*

Proof. $H(k)$ corresponds to the number of subtrees that we started exploring hence to the number of roots of subtrees that we haven't finished exploring before time k. For each root r of a subtree, $X(r) = \min_{r\leq j\leq k} X(j)$. X becomes strictly smaller when it exits the subtree. Hence for r the root of a subtree that we finished exploring, $X(r) > \min_{r\leq j\leq k} X(j)$, whereas for r the root of a tree we haven't finished exploring, $X(r) = \min_{r\leq j\leq k} X(j)$. We conclude. □

We now focus on Galton Watson trees with an off-spring distribution given by $(p_i)_{i\geq 0}$. We alternatively call ϵ the variable of the offspring law. The variable of the total progeny is called N. Let $R(k), k \geq 0$ be a Random Walk with $R(0) = 0$ and step distribution given by $(p_{i+1})_{i \geq -1}$ and $M = inf\{k \geq 0 : R(k) = -1\}$, then $(X(k), 0 \leq k \leq N) = R(k), 0 \leq k \leq M$. In other words, X is a random walk with step distribution the variable $\gamma = \epsilon - 1$. Now in the critical case, as $E(\gamma) = 0$ and $Var(\gamma) = \sigma^2$, Donsker's theorem applies to the process X.

Theorem 3.2 (Application of Donsker's theorem)**.** *Let X be a Lukasiewicz encoding of a forest of GWTs, then*

$$(\frac{X[nt]}{\sqrt{n}\sigma}, t \geq 0) \xrightarrow[n\to\infty]{d} (W(t),\ t \geq 0)$$

where W is the Brownian motion.

3.1.2 Height encoding

From theorem 3.2 we deduce a similar limit theorem for the height encoding.

Theorem 3.3. *Let $t_1 \geq 0, ..., t_m \geq 0$ and H be a Height encoding of a forest of GWTs, then*

$$\left(\frac{H[nt]}{\sqrt{n}},\ t \geq 0\right) \xrightarrow[n\to\infty]{d} \frac{2}{\sigma}\left(W(t) - \inf_{0 \leq s \leq t} W(s),\ t \geq 0\right)$$

where W is the Brownian motion.

In order to prove the convergence in distribution of the height process, we need to proceed in two steps. First, we need to show the convergence in distribution of finite dimension marginals $\left(\frac{H[nt_1]}{\sqrt{n}\sigma}, ..., \frac{H[nt_m]}{\sqrt{n}\sigma}\right)$. Second, we need to prove the tightness of the laws of the processes $H^n \equiv \left(\frac{H[nt]}{\sqrt{n}\sigma},\ t \geq 0\right)$ in the set of all probability measures on the Skorokhod space $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$.

Theorem 3.4 (Convergence of finite-dimension marginals)**.** *Let $t_1 \geq 0, ..., t_m \geq 0,$ then*

$$\left(\frac{H[nt_1]}{\sqrt{n}}, ..., \frac{H[nt_m]}{\sqrt{n}}\right) \xrightarrow[n\to\infty]{} \frac{2}{\sigma}\left(W(t_1) - \inf_{0 \leq s \leq t_1} W(s), ..., W(t_m) - \inf_{0 \leq s \leq t_m} W(s)\right)$$

where W is the Brownian motion.

We begin with the proof of an intermediate lemma.

Lemma 3.5. *Let $T := \{inf\ k \geq 0, X_k \geq 0\}$, $v(i) = p_{i+1}$ the distribution of $\gamma = \epsilon - 1$ and $\overline{v}(i) = \sum_{k=i}^{+\infty} v(k)$ Then $E(X(T)) = \frac{\sigma^2}{2}$*

Proof. We follow exercice 4 page 9 of Goldschmidt's minicourse [2].

$$\begin{aligned} P(X(T) = k) &= P(X(T) = k | X_1 \geq 0) \cdot P(X_1 \geq 0) + P(X(T) = k | X_1 < 0) \cdot P(X_1 < 0) \\ &= P(X(T) = k, X_1 \geq 0) + v(-1) \cdot P(X(T) = k | X_1 < 0) \\ &= v(k) + v(-1) P(X(T) = k | X_1 < 0) \end{aligned}$$

We have that $P(X(T) = k | X_1 < 0) = P(X(T) = k | X_1 = -1)$. Let Y=X+1 and $T_y = inf\{k \geq 1, Y(k) \geq 0\}$. We also have that $T = \{k \geq 0, Y_k > 0\}$, and that

$$\begin{aligned} P(X(T) = k | X_1 < 0) &= P(Y(T) = k + 1 | Y_1 = 0, Y(T) > 0) \\ &= P(Y(T_y) = k + 1 | Y_1 = 0, Y(T) > 0) \\ &= P(X(T) = k + 1 | X_1 = 0, X(T) > 0) \\ &= P(X(T) = k + 1 | X(T) > 0) \end{aligned}$$

So by telescopy, and letting $c = \frac{v(-1)}{P(X(T)>0)}$,

$$P(X(T) = k) = v(k) + v(-1)\frac{P(X(T) = k+1)}{P(X(T) > 0)} = \sum_{j=0}^{+\infty} c^j v(k+j)$$

Then from $\sum_{k\geq -1} kv(k) = 0$ we deduce $v(-1) = \sum_{k\geq 1} kv(k)$ and so

$$\sum_{k=0} \overline{v}(k) = \overline{v}(0) + \sum_{k=1}^{+\infty} \overline{v}(k) = 1 - v(-1) + v(-1) = 1$$

$\sum_{k=0}^{+\infty} P(X(T) = k) = \sum_{k=0}^{+\infty} \sum_{j=0}^{+\infty} c^j v(k+j) = \sum_{j=0}^{+\infty} c^j \overline{v}(j) = 1$
and since $\overline{v}(0) < 1$, we deduce $c = 1$ and so $P(X(T) = k) = \overline{v}(k)$. Finally,

$$\begin{aligned}
E(X(T)) &= \sum_{k\geq 0} k\overline{v}(k) = \sum_{k\geq 1} k \sum_{j=k}^{+\infty} v(j) = \sum_{j=1}^{+\infty} (\sum_{k=1}^{j} k) v(j) \\
&= \sum_{j=1}^{+\infty} \frac{j(j+1)}{2} v(j) = \frac{1}{2}(\sum_{j=1}^{+\infty} j^2 v(j) + \sum_{j=1}^{+\infty} jv(j)) \\
&= \frac{1}{2}(\sigma^2 - v(-1) + v(-1)) = \frac{\sigma^2}{2}.
\end{aligned}$$

□

We can now prove the theorem :

Proof. We let $S(n) = \underset{0\leq k\leq n}{sup} X(k)$ and $I(n) = \underset{0\leq k\leq n}{inf} X(k)$. We introduce $\hat{X}^n(k) = X(n) - X(n-k)$

$$(*)(\hat{X}^n(k), 0 \leq k \leq n) \stackrel{d}{=} (X(k), 0 \leq k \leq n)$$

Indeed,

$$\begin{aligned}
P(\hat{X}^n(k) = i | \hat{X}^n(k-1) = j) &= P(X(k) = X(n) - i | X(k-1) = X(n) - j) \\
&= p_{i-j+1} = P(X(k) = i | X(k-1) = j)
\end{aligned}$$

Then,

$$\begin{aligned}
H(n) &= \#\{0 \leq k \leq n-1 : X(k) = \underset{k\leq j\leq n}{inf} X(j)\} \\
&= \#\{1 \leq i \leq n : X(n-i) = \underset{0\leq l\leq i}{inf} X(n-l)\} \\
&= \#\{1 \leq i \leq n : \hat{X}^n(i) = \underset{0\leq l\leq i}{sup} \hat{X}^n(l)\}
\end{aligned}$$

We define $J(n) = \#\{1 \leq i \leq n : X(i) = \underset{0\leq l\leq i}{sup} X(l)\}$
Note that $\underset{0\leq k\leq n}{sup} \hat{X}^n(k) = X(n) - I(n)$.
Then from $(*)$, by replacing X with $\hat{X}^n$, we have that $(S(n), J(n)) \stackrel{d}{=} (X(n) - I(n), H(n))$.

We now define $T_0 = 0,\ T_k = inf\{i > T_{k-1} : X(i) = S(i)\}$. $X(T_{k+1}) - X(T_k), k \geq 0$ are i.i.d and have mean equal to $\frac{\sigma^2}{2}$.

Then $\frac{H(n)}{X(n)-I(n)} \underset{n\to+\infty}{\longrightarrow} \frac{2}{\sigma^2}$. Indeed,

$$\begin{aligned}\frac{S(n)}{J(n)} &= \frac{\sum_{k\geq 1, T_k \leq n} S(T_k) - S(T_{k-1})}{J(n)} \\ &= \frac{\sum_{k\geq 1}^{J(n)} S(T_k) - S(T_{k-1})}{J(n)} \\ &= \frac{\sum_{k\geq 1}^{J(n)} X(T_k) - X(T_{k-1})}{J(n)} \\ &\underset{n\to+\infty}{\longrightarrow} \frac{\sigma^2}{2}\end{aligned}$$

Finally, using theorem 3.2, and the fact that

$$g : X \to \left(X(\lfloor nt_1 \rfloor) - \underset{0\leq k\leq \lfloor nt_1 \rfloor}{inf}(X(k)), ..., X(\lfloor nt_m \rfloor) - \underset{0\leq k\leq \lfloor nt_m \rfloor}{inf}(X(k)) \right)$$

is continuous measurable, we have that

$$\frac{1}{\sqrt{n}} \left(X(\lfloor nt_1 \rfloor) - \underset{0\leq k\leq \lfloor nt_1 \rfloor}{inf}(X(k)), ..., X(\lfloor nt_m \rfloor) - \underset{0\leq k\leq \lfloor nt_m \rfloor}{inf}(X(k)) \right) \underset{n\to\infty}{\longrightarrow} \sigma(g_{t_1}(W), ..., g_{t_m}(W))$$

hence the result by multiplying each component i by $\frac{H(\lfloor nt_i \rfloor)}{X(\lfloor nt_i \rfloor) - I(\lfloor nt_i \rfloor)}$ □

Theorem 3.6 (Tightness)**.** *The laws of the processes* $H^n \equiv \left(\frac{H[nt]}{\sqrt{n}\sigma},\ t \geq 0\right)$ *in the set of all probability measures on the Skorokhod space* $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ *are tight.*

Proof. The proof is long and detailed in Le Gall's article page 13 [5] □

3.1.3 Encoding limits of Conditionned Galton Watson trees

Let's call X^n (resp. H^n) the process X (resp. H) conditionned to $N = n$, meaning that the population goes extinct at time n. Then a conditionned version of Donsker's theorem gives us two theorems :

Theorem 3.7.

$$(\frac{X^n[nt]}{\sqrt{n}\sigma}, t \geq 0) \xrightarrow[n\to\infty]{d} e(t),\ t \geq 0$$

where $e(t)$ is a standard brownian excursion.

Theorem 3.8.

$$(\frac{H^n[nt]}{\sqrt{n}}, t \geq 0) \xrightarrow[n\to\infty]{d} \frac{2}{\sigma} e(t),\ t \geq 0$$

where $e(t)$ is a standard brownian excursion.

3.2 Continuous Random Tree

3.2.1 Definition of the CRT

In this subsection, we define $\mathbb{R}$-trees and show that given a certain distance, Galton Watson trees converge towards a particular $\mathbb{R}$-tree, the Continuous Random Tree. We use this convergence to give an alternative definition of the Browinan Continuous Random Tree.

We first start by giving the definition of an $\mathbb{R}$-tree.

Definition 3.1. *A compact metric space (T,d) is an $\mathbb{R}$-tree if the following conditions are fulfilled for every pair x,y $\in$ T :*

1. *There exists a unique isometric map $f_{x,y} : [0, d(x, y)] \to T$ such that $f_{x,y}(0) = x$ and $f_{x,y}(d(x, y)) = y$. We write $[[s, y]] := f_{x,y}([0, d(x, y)])$.*
2. *If g is a continuous injective map $[0, 1] \to T$ such that $g(0) = x$ and $g(1) = y$ then g([0,1])=[[x,y]].*

We now define a continuous excursion as well as the pseudo-metric associated in order to give a constructive definition of $\mathbb{R}$-trees.

Definition 3.2.
1. *A continuous excursion is a continuous function $h : [0, \sigma] \to \mathbb{R}_+$ with strictly positive values on $]0, \sigma[$ for some $0 < \sigma < \infty$*
2. *For a continuous excursion h defined on $[0, \sigma]$, the pseudo-metric d_h associated is defined on $[0, \sigma]$ by $d_h(x, y) = h(x) + h(y) - 2 \inf_{\min(x,y) \le z \le \max(x,y)} h(z)$*

This pseudo-metric defines an equivalence relation by $x \sim y$ if and only if $d_h(x, y) = 0$, and let T_h be given by the quotient $[0, \sigma]/\sim$. Then theorem 10 page 12 of Goldschmidt's minicourse [2] states that T_h is an $\mathbb{R}$-tree.

Theorem 3.9. *For any continuous excursion h, (T_h, d_h) is an $\mathbb{R}$-tree.*

We can now define the continuous random tree (CRT).

Definition 3.3. *The CRT is (T_{2e}, d_{2e}) with e the brownian excursion.*

3.2.2 Convergence of Galton Watson Trees towards the CRT

In order to talk about convergence of GWTs toward the CRT, we finally have to define a distance between trees. We now define the Gromov-Hausdorff distance. Let M be the quotient space of compact metric spaces isometries. Let (X, d) and $(X', d') \in M$.

Definition 3.4.
1. *A correspondance R is a subset of $X \times X'$ such that $\forall x \in X, \exists x' \in X'$ such that $(x, x') \in R$ and vice versa.*
2. *The distorsion of a correspondance R is $dis(R) := sup\{|d(x, y) - d'(x', y')| : (x, x'), (y, y') \in R\}$*
3. *The Gromov-hausdorff distance between X and X' is $d_{GH}((X, d), (X', d')) = \frac{1}{2} \underset{R}{inf}\, dis(R)$ with the infimum taken over all correspondances between X and X'.*

We now give a very important theorem, showing that given certain asumptions on the GWT, it converges towards a CRT for the Gromov-Hausdorff distance.

Theorem 3.10. *Let T_n be a a critical GWT conditionned to have size n, with offspring variance $\sigma^2 > 0$, and let d_n be the natural distance on the graph of size n ($d_n(i,j) = k$ if the shortest path between i and j has k arcs).*

$$Then\ (T_n, \frac{\sigma}{\sqrt{n}} d_n) \xrightarrow[n\to\infty]{d} (T_{2e}, d_{2e})$$

Proof. From the Representation theorem of Skorokhod, there iexists a probability space in which the convergence happens almost surely with the uniform norm :

$$\frac{\sigma}{\sqrt{n}}(H^n(\lfloor nt \rfloor, 0 \le t \le 1) \xrightarrow[n\to\infty]{a.s} 2(e(t), 0 \le t \le 1)$$

Calling $v_0, ..., v_{n-1}$ the vertices of T_n in the lexicographic order, $(T_n, \frac{\sigma}{\sqrt{n}} d_n)$ is isometric to $\{0, 1, ..., n-1\}$ with the distance $d^n(i,j) = \frac{\sigma}{\sqrt{n}}(v_i, v_j)$. We define now a correspondance between $\{0, 1, ..., n-1\}$ and $[0,1]$ by letting $(i, s) \in R_n$ if $i = \lfloor ns \rfloor$. We also say that $(n-1, 1) \in R_n$. We provide $[0,1]$ of the pseudo-metric d_{2e}. We are going to bound $dis(R_n)$.

$d_n(v_i, v_j) = d_n(v_0, v_i) + d_n(v_0, v_j) - 2d_n(v_0, min(v_i, v_j))$. We also have $d_n(v_0, v_i) = H^n(i)$ and $|d_n(v_0, min(v_i, v_j) - \underset{i \le k \le j}{min} H^n(k)| \le 1$

Let $(i, s), (j, t) \in R_n$ with $s \le t$, then the previous results give us

$$|d_n(i,j) - d_{2e}(s,t)| = \frac{\sigma}{\sqrt{n}}|(H^n\lfloor ns \rfloor + H^n\lfloor nt \rfloor - 2d_n(v_0, min(v_i, v_j)) - 2e(s) + 2e(t) - 4\underset{s \le u \le t}{min} e(u))|$$

$$\le |\frac{\sigma}{\sqrt{n}}(H^n\lfloor ns \rfloor + H^n\lfloor nt \rfloor - 2\underset{s \le u \le t}{min} H^n\lfloor nu \rfloor) - 2e(s) + 2e(t) - 4\underset{s \le u \le t}{min} e(u)| + 2\frac{\sigma}{\sqrt{n}}$$

et donc $dis(R_n) \underset{n\to+\infty}{\to} 0$ from the conditionned version of Donsker's theorem.

Since $d_{GH}((T_n, \frac{\sigma}{\sqrt{n}}), (T_{2e}, d_{2e})) \le \frac{1}{2} dis(R_n) \underset{n\to+\infty}{\longrightarrow} 0$, we have the result. ☐

3.2.3 Alternative constructive definition of the CRT

We now use the results from 3.2.2 to give a constructive definition of the CRT. Indeed, we use theorem 3.10 to show that if we simulate a uniform tree on the set of unordered trees with a root, n vertices and whose vertices are seperated by a distance of $\sqrt{n}$, then when $n \to +\infty$ we simulate a CRT.

We start with a variant of the Aldous-Broder algorithm. This algorithm builds a tree whose law is uniform on $\mathbb{T}_n^*$:=the set of unordered rooted trees. Here are the steps of the algorithm :

1. Start from a graph of n unconnected vertices labelled by $[1, ..., n]$
2. Start from vertex 1. For $2 \le i \le n$, connect vertex i to V_i by an edge, where

$$V_i = \begin{cases} i-1 & \text{with probability } 1 - \frac{i-2}{n-1} \\ k & \text{with probability } \frac{1}{n-1} \text{ for } 1 \le k \le i-2 \end{cases}$$

3. permute the vertices randomly.

This algorithm can be seen as building paths of consecutive labelled vertices, with such a path ending whenever we reach a vertex labelled connected to $V_i \neq i-1$. Once we stop a path, we pick a uniform random point along it and start growing another path. Call $\mathcal{T}_n$ the tree produced by the algorithm, and provide $\mathcal{T}_n$ of the natural distance d_n on the graph. Then $\mathcal{T}_n$ converges in distribution to the CRT.

Proposition 3.1.

$$(\mathcal{T}_n, \frac{d_n}{\sqrt{n}}) \underset{n\to+\infty}{\longrightarrow} (\mathcal{T}_{2e}, d_{2e})$$

Proof. As $\mathcal{T}_n$ is uniform on $\mathbb{T}_n^*$ (see [2] page 2), exercice 2 of Goldschmidt's minicourse [2] shows that $\mathcal{T}_n$ can be seen as a Galton Watson tree with a progeny whose mean and variance is equal to 1. We now use theorem 3.10 to conclude. □

Finally, this gives us a new way to build the CRT. Indeed the two variables used to simulate the random variable produced by Aldous algorithm are $C_1^n := inf\{i \geq 2 : V_i \neq i-1\}$ and then a uniform variable $J_1^n := \mathcal{U}[0, C_1^n]$.
We can therefore build $\underset{n\to+\infty}{lim}(\mathcal{T}_n, \frac{d_n}{\sqrt{n}})$ by simulating $C_1 := \underset{n\to+\infty}{lim}\frac{C_1^n}{\sqrt{n}}$, then simulating $J_1 := \mathcal{U}[0, C]$ and the resimulating another path C_2 with same law as C_1 from J_1 and so on.
The distance used on the tree is the length of the shortest path between two points. Finally, we can compute the law of C_1.

Proposition 3.2.

$$\frac{C_1^n}{\sqrt{n}} \xrightarrow{d} C_1 \ with \ P(C_1 > x) = exp(-x^2/2) \ for \ x \geq 0$$

Proof. $P(\frac{C_1^n}{\sqrt{n}} > x) = P(C_1^n \geq \lfloor \frac{x}{\sqrt{n}} \rfloor + 1) = \prod_{i=1}^{\lfloor x\sqrt{n}-2 \rfloor}(1 - \frac{i}{n-1})$
therefore $-log(P(\frac{C_1^n}{\sqrt{n}} > x)) = -\sum_{i=1}^{\lfloor x\sqrt{n}-2 \rfloor} log(1 - \frac{i}{n-1}) = \sum_{i=1}^{\lfloor x\sqrt{n}-2 \rfloor} \frac{i}{n-1} + o(1) \underset{n\to+\infty}{\longrightarrow} \frac{x^2}{2}$ □

3.3 Summury

In order to study the convergence of the critical GWT in varying environment towards the CRT in Section 4, let's synthesize and label the 3 important steps of this section.

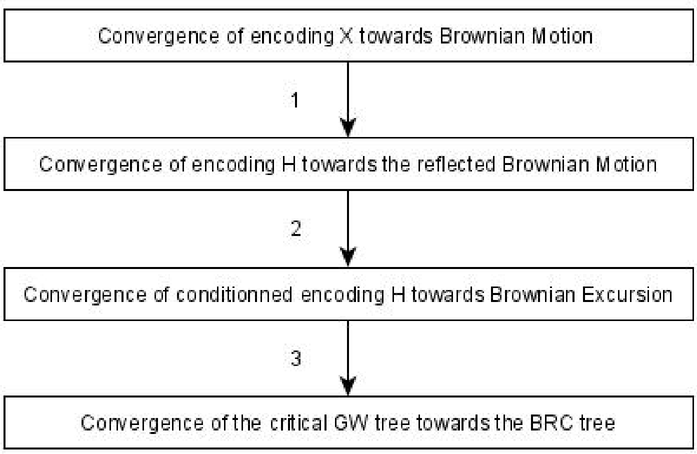

Figure 1: Summury of Section 3

4 Scaling Limit in a varying environment

Let's start by explaining the reasons leading us to study the Scaling limits of a GWT in varying environment. The question of the limits of a varying environment comes from the convergence of a GWT in a special case of random environment, the i.i.d environment. A sequence $\mathcal{V} = \{F_1, ..., F_\infty\}$ where all F_i have same law F is called an i.i.d environment.

Then, in an i.i.d environment, $X(i) = \sum_{j=1}^{i} \xi_j$ with $\xi_j \sim \xi \ for \ all \ j \in \mathbb{N}$ and with $P(\xi = k) = \int_{\mathcal{P}(\mathbb{N})} f[k-1]dQ(f)$ with Q the law of F on $\mathcal{P}(\mathbb{N})$. We are back to the case of a constant offspring distribution. Indeed, $\mathbb{E}(\xi) = \mathbb{E}(E(\xi|F)) = \mathbb{E}(\overline{F} - 1)$. Therefore, in the case in which $\mathbb{E}(\overline{F}) = 1$, $\mathbb{E}(\xi) = 0$ and Donsker's theorem applies, meaning that

$$(\frac{X[nt]}{\sqrt{n}\sigma}, t \geq 0) \xrightarrow[n\to\infty]{d} (W(t),\ t \geq 0)$$

with $\sigma = Var(\xi) = \mathbb{E}(\overline{F}^2)$

Thus we know that convergence occurs in average over all the possible varying environments when $\mathbb{E}(\overline{F}) = 1$. On the other hand, we do not understand what happens for a given varying environment. This is exactly what we study in this section. We need to make the assumption $\mathbb{E}(\overline{F}) = 1$, but we make one even stronger in order to start with an easier case : $\overline{F} = 1$ almost surely. More precisely, given the law of F and G the set of all possible varying environments, we would like to show that there exists H with G\H negligeable such that the Galton Watson tree converges towards the CRT on H.

Finally, recall the three important steps of the figure 1. In this section we study how these steps translate in the case of a varying environment. We first notice that the theorems used in step 2 and 3 do not require a specific environment. They therefore apply in the case of a varying environment. What we need to prove is this section is the convergence of X and the first step. Because of time constraints we focus on the convergence of the Lukasiewicz encoding in this report.

4.1 Lukasiewicz encoding in a varied Environment

We now focus on the behaviour of the Lukasiewicz encoding in the case of a fixed environment $f_1, ..., f_\infty$. We want to study the case in which the $f_1, ..., f_\infty$ are drawn according to the law of a variable F$\in \mathcal{P}(\mathbb{N})$ with $E(\overline{F}) = 1$. We focus our study on the particular case in which $\overline{F} = 1$ almost surely. More precisely, given G the set of all possible varying environments, we would like to show that given a fixed environment belonging to a set H with G\H negligeable, the Lukasiewicz encoding converges towards a Brownian Motion not depending on the $f_1, ..., f_\infty$.

Conjecture 4.1. *For $f_1, ..., f_\infty$ in a certain space,*

$$(\frac{X[nt]}{\sqrt{n}\sigma}, t \geq 0) \xrightarrow[n\to\infty]{d} (W(t),\ t \geq 0)$$

with $\sigma = \mathbb{E}(Var(F))$

We first show how our computer simulations support this evidence of a convergence towards a Brownian Motion when the $f_1, ..., f_\infty$ are chosen i.i.d.

4.1.1 Computer Simulations

We have chosen to implement our simulations using a python Notebook in order to give detailed explanations on the used algorithms.

Case 1 : For the first part of our simulations, to simulate each law $f \in \{f_1, ..., f_\infty\}$, we draw α according to $\mathcal{U}([0,1])$ and we set $f = \frac{(1-\alpha)}{2}\delta_0 + \alpha\delta_1 + \frac{(1-\alpha)}{2}\delta_2$.
Notice that $E(f) = 1$ for any α. The first image shows the distribution of $\frac{X[nt]}{\sqrt{n}}$ for different values of t and big values of n.

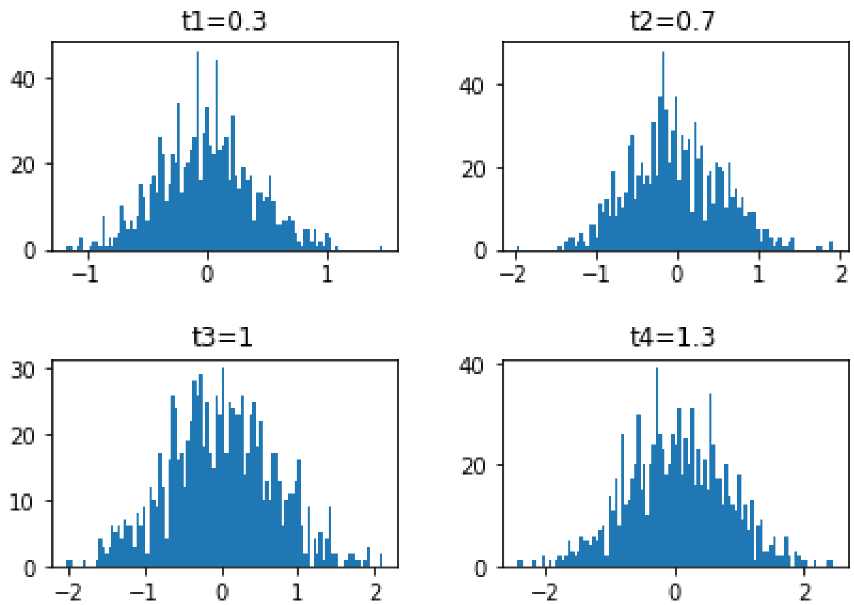

Figure 2: Histogram of a 1000 samples of $\frac{X[nt]}{\sqrt{n}}$ with t=0.3, 0.7, 1, 1.3, n=3000 for the Lukasiewicz encoding.

Remark : We notice that the histogram ressembles a lot to the histogram of a Gaussian variable, which confirms our conjecture. Furthermore, the variance of the sample seems to increase with t.

We also conducted three different normality tests available in a Python library on the sample corresponding to $t = 1$ to check if the distribution was likely Gaussian or not. It turns out that all three tests affirm that the sample is drawn by a Gaussian variable.

```
================================================================================
Marche de Luka : résultats de trois différents tests pour la valeur 1
================================================================================
stat_luka_shapiro=0.998, p_luka_shapiro=0.286
Probably Gaussian according to Shapiro test
-----------------------------------------------------------------
stat_luka_agostino=3.039, p_luka_agostino=0.219
Probably Gaussian according to Agostino test
-----------------------------------------------------------------
stat=0.341
Probably Gaussian at the 15.0% level
Probably Gaussian at the 10.0% level
Probably Gaussian at the 5.0% level
Probably Gaussian at the 2.5% level
Probably Gaussian at the 1.0% level
```

Figure 3: Results of three different tests for the sample corresponding to t=1. The tests used are the Shapiro test, the Agostino test as well as the Anderson test.

We also plotted complete trajectories of the Lukasiewicz encoding, which seems to behave like a brownian motion.

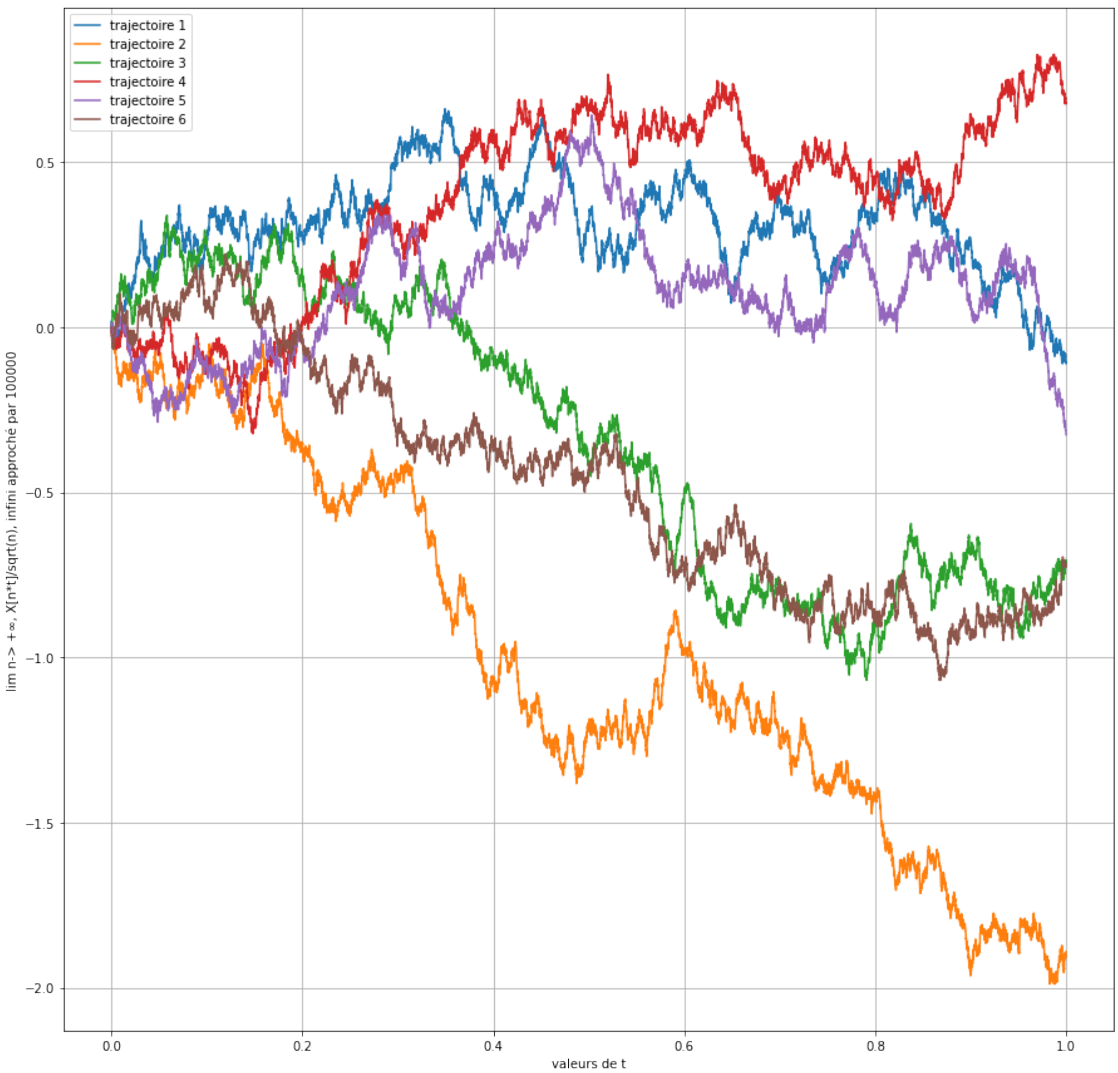

Figure 4: 6 trajectories of a Lukasiewicz encoding with n=100 000 and step of 0.001.

Finally, we investigated the variance of the sample $\frac{X[nt]}{\sqrt{n}}$ in order to have an idea of the parameter σ in our conjecture 3.1.

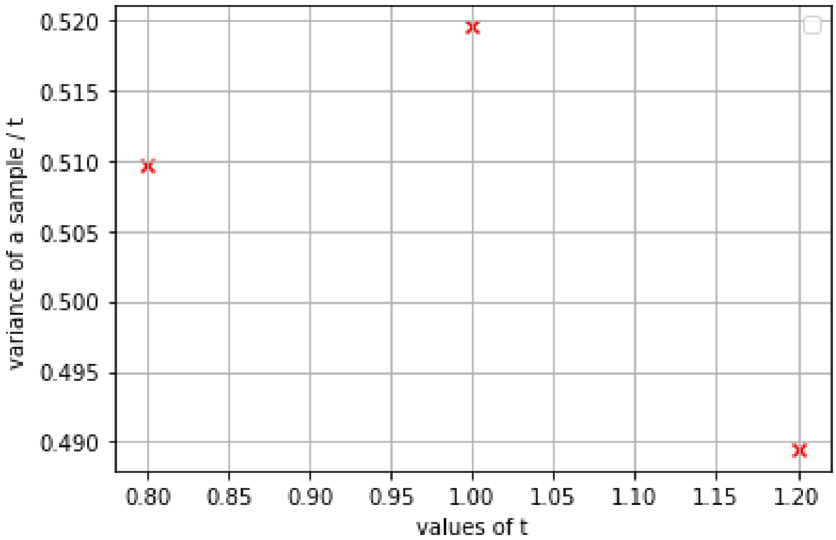

Figure 5: Plot of the variance divided by t of a sample of size 2000 of the variable $\frac{X[nt]}{\sqrt{n}}$ for n=3000, and $t = 0.8, 1, 1.2$

We can make the conjecture that in the case of the law of case 1, we have

$$(\frac{X[nt]}{\sqrt{n}\sigma}, t \geq 0) \xrightarrow[n \to \infty]{d} (W(t),\ t \geq 0)$$

with $\sigma = 0.5$

Furthermore, $\mathbb{E}(Var(F)) = \int_0^1 Var(f_\alpha)d\sigma = \int_0^1 (1-\alpha)d\alpha = 0.5$ therefore our computations confirm the conjecture 3.1.

5 Reflections and conclusion

This internship enabled me to understand the everyday problematics of a Researcher more precisely. I first started reading various articles concerning different aspects of the research I was meant to conduct. As my understanding of the subject improved, I learned to target articles that concerned my problematic the most. As the topic was very new I could not find much litterature about the scaling limits of GWTs in varying environments. I therefore analysed a simpler case in order to understand clearly what was left to do in a different environment. We made the conjecture that GWT encodings had a similar behaviour in a varying environment than in an environment with constant offspring probability law, and studied the conjecture by conducting computer simulations. Our simulations then confirmed the pertinence of this conjecture. With more time, I wish to be able to write down the mathematical ideas proving the conjecture, as well as the other steps missing to prove the convergence of the scaling limits of GWTs in varying environment. I believe that going through these different steps in my research confronted me to a multitude of challenges specific to the world of Research. To cite a few of them, I learned the necessity to learn about other researchers' work in the domain concerned as well as the importance of communication with researchers knowing the subject better, especially during times in which I found myself locked in my research and did not know how to advance.

References

[1] Thomas Duquesne and Jean-François Le Gall. *Random trees, Lévy processes and spatial branching processes*. Vol. 281. Société mathématique de France, 2002.

[2] Christina Goldschmidt. "Scaling limits of random trees and random graphs". In: *PIMS-CRM Summer School in Probability*. Springer. 2017, pp. 1–33.

[3] William D Kaigh. "An invariance principle for random walk conditioned by a late return to zero". In: *The Annals of Probability* (1976), pp. 115–121.

[4] Götz Kersting and Vladimir Vatutin. *Discrete time branching processes in random environment*. John Wiley & Sons, 2017.

[5] Jean-François Le Gall. "Random trees and applications". In: *Probability surveys* 2 (2005), pp. 245–311.

[6] Gregory L McColm. "An introduction to random trees". In: *Research on Language and Computation* 1.3 (2003), pp. 203–227.

[7] Zhan Shi. *Branching random walks*. Springer, 2015.