



HAL
open science

Marches aléatoires et arbres de Galton-Watson

Aymen Bouaziz

► **To cite this version:**

Aymen Bouaziz. Marches aléatoires et arbres de Galton-Watson. Mathématiques générales [math.GM]. Université d'Orléans; Université de Tunis El Manar, 2017. Français. NNT: 2017ORLE2047. tel-01909729

HAL Id: tel-01909729

<https://theses.hal.science/tel-01909729>

Submitted on 31 Oct 2018

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

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



ÉCOLE DOCTORALE
[MATHEMATIQUES, INFORMATIQUE, PHYSIQUE THEORIQUE ET INGENIERIE DES SYSTEMES

Laboratoire MAPMO et Département de Mathématiques de l'Université de Tunis El Manar

THÈSE EN COTUTELLE INTERNATIONALE présentée par :

Aymen Bouaziz

soutenue le : 09 12 2017

pour obtenir le grade de :

**Docteur de l'université d'Orléans
et de l'Université de Tunis El Manar**

Discipline : Mathématiques

Marches aléatoires et arbres de Galton-Watson

THÈSE dirigée par :

Monsieur Romain Abraham Professeur, Université d'Orléans
Monsieur Mohamed Sifi Professeur, Université de Tunis El Manar

RAPPORTEURS :

Monsieur Thomas Duquesne Professeur, Université Pierre et Marie Curie
Monsieur Afif Masmoudi Professeur, Université de Sfax

JURY (indiquez tous les membres du jury – y compris directeur(s) thèse et rapporteurs) :

Madame Saloua Aouadi Professeur, Université Tunis El Manar, Président du jury
Monsieur Romain Abraham Professeur, Université d'Orléans
Monsieur Mohamed Sifi Professeur, Université de Tunis El Manar
Monsieur Thomas Duquesne Professeur, Université Pierre et Marie Curie
Monsieur Afif Masmoudi Professeur, Université de Sfax
Monsieur Jean François Delmas Professeur, Ecole des ponts

Table des matières

1	Introduction	5
1.1	Définitions et état de l'art.	6
1.1.1	Fonctions harmoniques	6
1.1.2	Limite locale des arbres de Galton Watson	7
1.2	Description des résultats obtenus	18
1.2.1	Chapitre 2 : Fonctions harmoniques discrètes dans un orthant	18
1.2.2	Chapitre 3 : Convergence d'un arbre de Galton Watson critique conditionné à avoir un nombre infini de noeuds protégés	20
1.2.3	Chapitre 4 : Convergence d'un arbre de Galton Watson géométrique conditionné à avoir une génération anormalement grande	21
2	Discrete harmonic functions	23
2.1	Introduction	23
2.2	Proof of Theorem 2.1.1	25
2.2.1	Key technical ingredients	25
2.2.2	A lower estimate of the harmonic measure	27
2.2.3	The Hölder continuity at the boundary	28
2.2.4	The Carleson principle	29
2.2.5	Existence of a positive harmonic function on \mathcal{D}	31
2.3	Uniqueness of the positive harmonic function	31
2.3.1	Boundary Harnack principle	31
2.3.2	The denouement	33
2.3.3	An example	34
3	Local limits of Galton-Watson trees	39
3.1	Introduction	39
3.2	Technical background on GW trees	40
3.2.1	The set of discrete trees	40
3.2.2	Galton Watson trees	42
3.3	Conditioning on the number of marked vertices	42

3.3.1	Definition of the marking procedure	42
3.3.2	Kesten's tree	42
3.3.3	Main theorem	43
3.4	Proof of Theorem 3.3.3	43
3.4.1	Transformation of a subset of a tree onto a tree	45
3.4.2	Distribution of the number of marked nodes	47
3.4.3	Proof of Lemma 3.4.4	47
3.4.4	Proof of (3.7)	49
3.5	Protected nodes	49
4	Very fat geometric Galton-Watson trees	53
4.1	Introduction	2
4.2	Notations	3
4.2.1	The set of discrete trees	4
4.2.2	Convergence of trees	5
4.2.3	Galton-Watson trees	5
4.2.4	Geometric distribution	6
4.3	The geometric GW tree	8
4.4	The Kesten regime or the not so fat case	11
4.4.1	The Kesten tree	11
4.4.2	Convergence of the not so fat geometric GW tree	13
4.5	The Poisson regime or the fat case	14
4.5.1	An infinite Poisson tree	14
4.5.2	Convergence of the fat geometric GW tree	16
4.6	The condensation regime or the very fat case	17
4.6.1	An infinite geometric tree	17
4.6.2	Convergence of the very fat geometric GW tree	23

Chapitre 1

Introduction

Résumé

Dans cette thèse nous nous sommes intéressés à trois types de problèmes :

- Existence et unicité d'une fonction harmonique strictement positive associée à une marche aléatoire inhomogène confinée dans un orthant de \mathbb{Z}^d .
- Étude de la convergence en loi des arbres de Galton Watson critiques conditionnés à avoir un nombre assez grand de noeuds protégés.
- Étude de la convergence en loi des arbres de Galton Watson conditionnés à avoir une génération anormalement grande.

Ce manuscrit de thèse contient quatre chapitres.

Le premier chapitre est réservé à une introduction générale dans laquelle nous présentons les principales motivations et les résultats significatifs obtenus dans ce travail.

Le deuxième chapitre est consacré à l'étude d'une marche aléatoire inhomogène confinée dans un orthant de \mathbb{Z}^d et ses liens avec les fonctions harmoniques associées.

L'objet du troisième chapitre est d'étudier la convergence en loi des arbres de Galton Watson critiques conditionnés à avoir un nombre assez grand des noeuds marqués lorsqu'on marque ces noeuds de façon aléatoire, puis d'appliquer ce résultat à l'étude du même arbre conditionné à avoir un nombre assez grand de noeuds protégés.

Finalement, dans le quatrième chapitre, nous étudions la convergence en loi des arbres de Galton Watson conditionnés à avoir une génération anormalement grande dans le cas d'une loi de probabilité géométrique ayant une masse en 0 dont nous avons étudié tous les cas : critique, sous-critique et

sur-critique.

La suite de cette introduction présente les différents objets étudiés ainsi que les principaux résultats obtenus.

Mots clés : Marches aléatoires, fonctions harmoniques discrètes, orthant, arbres aléatoires, arbres de Galton-Watson, limite locale, noeuds protégés.

Abréviations :

- **i.i.d** : indépendantes et identiquement distribuées.
- **p.s** : Presque sûrement.

1.1 Définitions et état de l'art.

1.1.1 Fonctions harmoniques

On introduit ici la notion de fonction harmonique :

Soient E un ensemble dénombrable et $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov à espace d'états E de matrice de transition Q .

Une fonction $f : E \rightarrow \mathbb{R}_+$ est dite harmonique si pour tout $x \in E$,

$$f(x) = Qf(x) := \sum_{y \in E} Q(x, y)f(y) = \mathbb{E}(f(X_1) \mid X_0 = x).$$

Notons également que

$$f \text{ est harmonique sur } E \text{ si et seulement si } (Lf)(x) = 0 \quad \forall x \in E$$

avec

$$(Lf)(x) = \sum_{y \in E} Q(x, y)f(y) - f(x) : \text{ Laplacien discret de } f.$$

L'étude de l'existence de fonctions harmoniques positives pour les marches aléatoires homogènes a intéressé de nombreux chercheurs depuis plusieurs années.

Dans [14] Doney décrit les fonctions harmoniques et la frontière de Martin associés à une marche aléatoire apériodique $(Z_n)_{n \in \mathbb{N}}$ dans \mathbb{Z} tuée lorsqu'elle franchit la demi-droite réelle négative $\{z \in \mathbb{Z}; z < 0\}$, puis Alily et Doney étendent ce résultat pour la marche aléatoire spatio-temporelle correspondante $((Z_n, n); n \in \mathbb{N})$.

Kurkova et Malyshev [39] décrivent la frontière de Martin pour deux marches aléatoires dont l'une est sur $\mathbb{Z} \times \mathbb{N}$ dont l'ensemble des sauts est $\{(0, 0), (0, 1), (0, -1), (1, 0), (-1, 0)\}$ et l'autre est sur \mathbb{Z}_+^2 avec des conditions

de réflexion sur la frontière.

Plus tard Kurkova et Raschel [40] (2009) étudient les marches aléatoires du quart de plan ayant des sauts à distance au plus un, avec un drift non nul à l'intérieur et absorbées au bord. Plus précisément, ils obtiennent de façon explicite les séries génératrices des probabilités d'absorption au bord, puis leur asymptotique lorsque le site d'absorption tend vers l'infini. Ils calculent également l'asymptotique des fonctions de Green le long de toutes les trajectoires, en particulier selon celles tangentes aux axes.

Notons également que ces résultats semblent peu susceptibles de postuler dans une situation générale; en fait ils reposent de façon essentielle sur la structure unidimensionnelle du processus. En effet Kurkova et Malyshev [39], et Raschel utilisent une méthode analytique dont la structure géométrique de la courbe elliptique définie par la fonction génératrice de la loi de saut de la marche aléatoire joue un rôle crucial.

Notre travail, rédigé au chapitre 2, se situe dans la continuité des résultats présentés dans cette section. En fait nous le considérons comme complémentaire à ceux de Raschel [50] car nous répondons à la conjecture indiquée dans [50] Concernant l'existence et l'unicité de la fonction harmonique associée à une marche aléatoire confinée dans un plan de \mathbb{Z}^2 .

1.1.2 Limite locale des arbres de Galton Watson

Processus de Galton-Watson et arbre de Galton-Watson

Le processus de Galton Watson est considéré comme le premier modèle stochastique pour l'évolution de la population. Ce processus a été étudié par les britanniques Francis Galton et Henry William Watson. En réalité, en 1873, l'anthropologue et statistien Sir Francis Galton, préoccupé par l'extinction des familles nobles et bourgeoises en Angleterre, publiait un article dans la revue "Educational times" posant le problème suivant :

Une grande population de N individus portant tous des noms de famille distincts, dans laquelle nous ne nous intéresserons qu'aux adultes de sexe masculin, colonise une région. Dans cette population, à chaque génération, a_0 pour cent des hommes adultes n'ont aucun fils qui atteindra l'âge adulte, a_1 pour cent en ont 1, a_2 pour cent en ont 2, et ainsi de suite jusqu'à a_5 pour cent qui en ont 5. Trouver

- (1) quelle proportion des noms de famille aura disparu après n générations ;
- (2) combien de noms de famille seront portés par exactement k personnes.

F. Galton suppose que les individus donnent naissance indépendamment l'un de l'autre à un nombre aléatoire d'enfants avec la même loi de reproduction. Après n'avoir reçu aucune réponse précieuse, il commence à travailler

avec H.W. Watson sur ce problème et réussissent à publier un article un an plus tard [22] où ils ont prouvé que la probabilité de l'extinction est un point fixe de la fonction génératrice de la loi de reproduction (ce qui est vrai) et que cette probabilité est toujours égale à 1 (ce qui est faux, voir le paragraphe suivant). Notons également qu'un problème similaire a été étudié en 1845 par le mathématicien Français I.J. Bienaymé, c'est pour cela que ce processus est parfois appelé processus de Bienaymé - Galton - Watson. Nous nous référons notamment à l'article de D.Kendal [33] "The genealogy of genealogy : branching processes before (and after) 1873" et au livre de T. Harris [24] pour des commentaires historiques.

On se donne une suite de réels positifs ou nuls $p = (p(k))_{k \in \mathbb{N}}$ telle que $\sum_{k=0}^{\infty} p(k) = 1$. Le processus de Galton - Watson se déroule de la manière suivante : on commence avec un individu que l'on appelle en général la racine. Elle a k enfants avec probabilité $p(k)$. Chacun d'eux a de même k enfants avec probabilité $p(k)$, indépendamment les uns des autres. De plus, chaque individu vivant à l'instant n meurt à la génération $n + 1$. Cela continue indéfiniment on s'arrête lorsqu'il n'y a plus d'enfant (on parle alors d'extinction). Plus formellement, soit X une variable aléatoire telle que $\mathbb{P}(X = k) = p(k)$ (p est appelée la loi de natalité du processus de Galton - Watson) et soit $(X_i^{(n)})_{n, i \geq 1}$ des copies indépendantes de X . La taille de la génération à l'instant n , notée Z_n , est définie récursivement de la manière suivante :

$$Z_0 = 1, \quad Z_{n+1} = \begin{cases} \sum_{i=1}^{Z_n} X_i^{(n+1)} \\ 0 \end{cases} \quad \text{si } Z_n = 0$$

Notons également que $(Z_n)_{n \in \mathbb{N}}$ est une chaîne de Markov de matrice de transition Q définie par :

$$Q(x, y) = p^{*x}(y)$$

avec

$$p^{*0} = \delta_0 \quad \text{et} \quad p^{*x} = p * \dots * p \quad \text{si } x \geq 1.$$

On note

$$\mu = \sum_{k=0}^{\infty} kp(k).$$

Un processus de Galton-Watson est qualifié de sous-critique quand $\mu < 1$, critique quand $\mu = 1$ et surcritique quand $\mu > 1$.

Deux résultats classiques et particulièrement utiles dans l'étude du processus de Galton-Watson $(Z_n)_{n \in \mathbb{N}}$ sont les suivants :

Proposition 1.1.1. *La fonction génératrice de Z_n est*

$$\mathbb{E}(s^{Z_n}) = f^n(s) := f \circ \dots \circ f(s)$$

l'itéré n fois, où f est la fonction génératrice de la loi p .

De plus on a :

Proposition 1.1.2. *Lorsque $\mu \in]0, +\infty[$, la suite $\frac{Z_n}{\mu^n}$ est une martingale positive relativement à la filtration*

$$\mathcal{F}_0 = \{\emptyset, \Omega\}$$

$$\mathcal{F}_n = \sigma(X_k^{(j)} : k < n, j \in \mathbb{N}) \text{ pour } n \geq 1.$$

et converge p.s vers une variable aléatoire intégrable.

On appelle extinction l'événement $\{\exists n \in \mathbb{N}; Z_n = 0\}$. En notant q la probabilité de l'extinction, on a

$$q = \mathbb{P}(\exists n \in \mathbb{N}; Z_n = 0) = \lim_{n \rightarrow +\infty} \mathbb{P}(Z_n = 0) = \lim_{n \rightarrow +\infty} f^n(0).$$

Alors

- Si $\mu \leq 1$ (critique ou sous-critique) et $p(1) \neq 1$ alors $q = 1$.
- Si $\mu > 1$ alors q est l'unique point fixe de la fonction génératrice f dans $[0, 1[$.

La première détermination complète et donne une version correcte de la probabilité d'extinction a été donné par J.F. Steffensen (1932). Ce problème a également été traité par A. Kolmogorov (1938), qui a réussi à déterminer la forme asymptotique de la probabilité que la famille existe encore après un grand nombre fini de générations.

A.J. Lotka a réalisé l'idée de Galton, en utilisant les données de fertilité américaines, pour déterminer la probabilité d'extinction d'une lignée de descendance masculine.

Notons également qu'il existe un lien intéressant entre le processus de Galton-Watson et le processus de naissance et de mort dans une étude du taux de formation de nouvelles espèces. Il existe également des liens avec la théorie du rayonnement cosmique formulée indépendamment par H.J.Bhabha et W. Heitler (1937) et par F. Carleson et R. Oppenheimer (1937).

Le processus de Galton - Watson intervient dans plusieurs domaines. N. Semonoff a utilisé le modèle de Galton-Watson dans les étapes élémentaires de son traité théorique sur les réactions chimiques, et W. Shockley et J. Pierre (1938) ont utilisé ce modèle pour étudier la multiplication des électrons dans un dispositif de détection électronique (multiplicateur d'électrons).

Après 1940, en raison de l'analogie entre la croissance des familles et les réactions de la chaîne nucléaire et grâce à l'intérêt général accru pour les applications de la théorie des probabilités, l'intérêt pour ce modèle a augmenté. Les premiers travaux stimulés par l'analogie nucléaire sont ceux de Hawkins et Ulam (1944) et Everett. Au cours des dernières années, le modèle de Galton-Watson a fait l'objet de nombreux travaux, nous nous

référons notamment à T. Harris [24], K. Athreya et P. Ney [8] pour des résultats importants autour le processus de Galton-Watson, à M. Kimmel et D. Axelrod [36] et P. Haccou, P. Jagers et V. Vatutin [23] pour des applications en biologie.

Afin d'étudier les processus de Galton-Watson, on peut considérer ce qu'on appelle les arbres géanalogiques ou les arbres de Galton-Watson qui sont considérés actuellement parmi les domaines actifs dans la recherche.

Les arbres de Galton-Watson jouent un rôle crucial dans la description de la généalogie du processus de Galton-Watson, ce dernier étant en fait simplement le nombre d'individus à chaque génération d'un arbre.

En adaptant la notation de Neveu [46] pour le formalisme d'arbres, on considère

$$\mathcal{U} = \bigcup_{n \geq 0} (\mathbb{N}^*)^n$$

l'ensemble des suites finies d'entiers avec la convention $(\mathbb{N}^*)^0 = \{\emptyset\}$. Pour $n \geq 1$ et $u = (u_1, \dots, u_n) \in \mathcal{U}$, on pose $|u| = n$: la longueur de u et $|u|_\infty = \max(|u|, u_1, \dots, u_{|u|})$ avec la convention $|\emptyset| = |\emptyset|_\infty = 0$. Si u et v deux éléments de \mathcal{U} , on désigne par uv la concaténation de u et v , avec la convention $uv = u$ si $v = \emptyset$ et $uv = v$ si $u = \emptyset$, on dit que v est un ancêtre de u et on note $v \prec u$ si : $\exists w \in \mathcal{U} \setminus \emptyset$ tel que $u = vw$. On note

$$A_u = \{v \in \mathcal{U}; v \prec u\}$$

l'ensemble des ancêtres de u .

Un arbre discret est un sous-ensemble de \mathcal{U} obtenu en représentant les individus par des sommets chacun étant lié par une arête à ses enfants. On renvoie le lecteur au chapitre 3 pour la définition exacte. Un sommet de l'arbre sera représenté par une suite finie d'entiers $u = (u_1, \dots, u_n)$. On note $k_u(\mathbf{t}) (\in \mathbb{N} \cup \{+\infty\})$ le nombre d'enfants de u dans l'arbre \mathbf{t} avec la convention $k_u(\mathbf{t}) = -1$ si $u \notin \mathbf{t}$.

Rappelons ici quelques vocabulaires des arbres discrets et les propriétés topologiques de certains espaces d'arbres :

- Le noeud $u \in \mathbf{t}$ est appelé une feuille si $k_u(\mathbf{t}) = 0$ et appelé noeud infini si $k_u(\mathbf{t}) = +\infty$
- Le noeud \emptyset est appelé la racine de \mathbf{t}

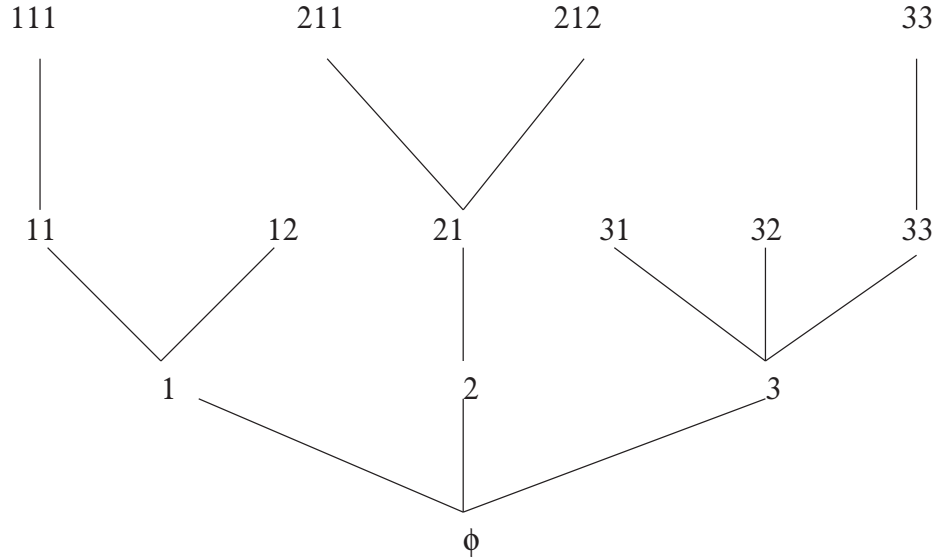


FIGURE 1.1 – Le codage d'un arbre discret

On désigne par :

- \mathbb{T}_∞ l'ensemble des arbres.
- $\mathbb{T} = \{\mathbf{t} \in \mathbb{T}_\infty; k_u(\mathbf{t}) < +\infty, \forall u \in \mathbf{t}\}$ l'ensemble des arbres sans aucun noeud infini.
- $\mathbb{T}_0 = \{\mathbf{t} \in \mathbb{T}; \text{Card}(\mathbf{t}) < +\infty\}$ l'ensemble des arbres finis.
- $\mathcal{L}_0(\mathbf{t}) = \{u \in \mathbf{t}; k_u(\mathbf{t}) = 0\}$ l'ensemble des feuilles de l'arbre \mathbf{t} .
- $H(\mathbf{t}) = \sup\{|u|, u \in \mathbf{t}\}$ la hauteur de l'arbre \mathbf{t} , peut être infinie.
- $H_\infty(\mathbf{t}) = \sup\{|u|_\infty, u \in \mathbf{t}\} = \max(H(\mathbf{t}), \sup\{k_u(\mathbf{t}), u \in \mathbf{t}\})$.
- $Z_n(\mathbf{t}) = \text{Card}\{u \in \mathbf{t}, |u| = n\}$ la taille de la génération n de l'arbre \mathbf{t} , peut être infinie.
- $\mathbb{T}^{(h)} = \{\mathbf{t} \in \mathbb{T}; H(\mathbf{t}) \leq h\}$ l'ensemble des arbres de hauteur inférieur ou égale à h .
- $\mathbb{T}_\infty^{(h)} = \{\mathbf{t} \in \mathbb{T}_\infty; H_\infty(\mathbf{t}) \leq h\}$.
- \mathbb{T}_1 l'ensemble des arbres ayant une unique branche infinie.

Pour $h \in \mathbb{N}$ et $\mathbf{t} \in \mathbb{T}$; on définit la fonction restriction r_h de l'arbre tronqué à la hauteur h par

$$r_h(\mathbf{t}) = \{u \in \mathbf{t}, |u| \leq h\}$$

et on munit \mathbb{T} d'une distance ultramétrique

$$d(\mathbf{t}, \mathbf{t}') = 2^{-\sup\{h \in \mathbb{N}, r_h(\mathbf{t}) = r_h(\mathbf{t}')\}}.$$

Une suite d'arbres $(\mathbf{t}_n, n \in \mathbb{N})$ converge vers un arbre \mathbf{t} dans (\mathbb{T}, d) si et seulement si $\forall h \in \mathbb{N}$, $r_h(\mathbf{t}_n) = r_h(\mathbf{t})$ pour n assez grand, i.e

$$\forall u \in \mathcal{U}, \lim_{n \rightarrow +\infty} k_u(\mathbf{t}_n) = k_u(\mathbf{t}).$$

Notons également que (\mathbb{T}, d) est un espace métrique complet séparable non compact. Plus précisément \mathbb{T}_0 est une partie dénombrable dense dans (\mathbb{T}, d) .

On peut généraliser dans \mathbb{T}_∞ et on définit la fonction restriction $r_{h,\infty}$ de \mathbb{T}_∞ à \mathbb{T}_0 définie par :

$$\forall \mathbf{t} \in \mathbb{T}_\infty, \quad r_{h,\infty}(\mathbf{t}) = \{u \in \mathbf{t}, |u|_\infty \leq h\}$$

qui consiste à restreindre la hauteur de l'arbre au niveau h et à ne garder en chaque sommet que les h premiers enfants.

On munit \mathbb{T}_∞ d'une distance ultramétrique d_∞ :

$$d_\infty(\mathbf{t}, \mathbf{t}') = 2^{-\sup\{h \in \mathbb{N}, r_{h,\infty}(\mathbf{t}) = r_{h,\infty}(\mathbf{t}')\}}.$$

Une suite d'arbres $(\mathbf{t}_n, n \in \mathbb{N})$ converge vers un arbre \mathbf{t} dans $(\mathbb{T}_\infty, d_\infty)$ si $\forall h \in \mathbb{N}$, $r_{h,\infty}(\mathbf{t}_n) = r_{h,\infty}(\mathbf{t})$ pour n assez grand, i.e

$$\forall u \in \mathcal{U}, \quad \lim_{n \rightarrow +\infty} k_u(\mathbf{t}_n) = k_u(\mathbf{t}) \in \mathbb{N} \cup \{-1, +\infty\}.$$

Rappelons la remarque dans [3] page 6 : Soient $(\mathbf{t}_n, n \in \mathbb{N})$ une suite d'arbres dans \mathbb{T} et $\mathbf{t} \in \mathbb{T}$. Alors $(\mathbf{t}_n, n \in \mathbb{N})$ converge vers \mathbf{t} dans (\mathbb{T}, d) si et seulement si $(\mathbf{t}_n, n \in \mathbb{N})$ converge vers \mathbf{t} dans $(\mathbb{T}_\infty, d_\infty)$.

Pour la définition exacte d'un arbre de Galton-Watson τ , nous renvoyons le lecteur au chapitre 3. Cette définition est caractérisée par la relation suivante : $\tau \in \mathbb{T}$ p.s. et, $\forall h \in \mathbb{N}^*$ et $\mathbf{t} \in \mathbb{T}^{(h)}$, on a :

$$\mathbb{P}(r_h(\tau) = \mathbf{t}) = \prod_{u \in r_{h-1}(\mathbf{t})} p(k_u(\mathbf{t})).$$

En particulier, on a pour les arbres finis :

$$\forall \mathbf{t} \in \mathbb{T}_0, \quad P(\tau = \mathbf{t}) = \prod_{u \in \mathbf{t}} p(k_u(\mathbf{t})).$$

Nous nous référons à M. Drmota [16] et S.Evans [20] pour une vision d'ensemble sur les arbres aléatoires discrets y compris les arbres de Galton-Watson.

En utilisant les arbres de Galton-Watson, on peut avoir une vision complète sur la probabilité d'extinction en traitant tous les cas de p . Il y a tout d'abord tous les cas dégénérés :

- Si $p(0) = 0$ alors $q = 0$ i.e $\tau \notin \mathbb{T}_0$.
- Si $p(0) = 1$ alors $\tau = \{\emptyset\}$ et $q = 1$.
- Si $p(0) = 0$ et $p(1) = 1$ alors l'arbre est réduit à une unique branche infinie, en particulier $q = 0$.
- Si $0 < p(0) < 1$ et $p(0) + p(1) = 1$ alors $H(\tau) + 1$ suit la loi géométrique de paramètre $p(1)$ et $\tau = \cup_{0 \leq n \leq H(\tau)} \{1\}^n$. Dans ce cas $q = 1$.

Dans toute la suite on suppose que p vérifie

$$p(0) > 0, \quad p(0) + p(1) < 1, \quad \text{et } \mu := \sum_{n=0}^{+\infty} np(n) < +\infty. \quad (1.1)$$

Convergence en loi des arbres de Galton Watson, limite locale

Il existe de nombreux types de limites qui peuvent être envisagées pour étudier les grands arbres, parmi lesquels sont les limites locales et limites d'échelle. Les limites locales sont celles associées aux distances d et d_∞ introduites précédemment. De façon informelle, on regarde l'arbre jusqu'à une hauteur fixe arbitraire donc on ne voit que ce qui se passe à une distance finie de la racine. Pour les limites d'échelle, on considère les suites d'arbres où les branches sont mises à l'échelle de sorte que tous les sommets restent à une distance bornée de la racine. Ces limites, qui aboutissent aux arbres aléatoires continus dont les branches ont une longueur infinitésimale, ont été intensément étudiées ces dernières années, voir [5], [17] et [18].

Nous nous intéressons ici uniquement aux limites locales des arbres de Galton-Watson critiques ou sous-critiques conditionnés à être anormalement grands. Le fameux conditionnement est le théorème de Kesten qui stipule que les arbres de Galton-Watson critiques ou sous-critiques conditionnés à avoir une hauteur assez grande convergent vers l'arbre de Kesten τ^* . Ce résultat sera rappelé dans le paragraphe suivant.

L'arbre de Kesten associé à une loi de reproduction p est un arbre de Galton-Watson multi-type défini comme suit :

- Il y a deux types de noeuds : les noeuds normaux et les noeuds spéciaux.
- La racine est spéciale.
- Les noeuds normaux se reproduisent selon la loi p et tous les enfants sont normaux.
- Les noeuds spéciaux se reproduisent selon la loi \hat{p} biaisée par la taille définie par

$$\hat{p}(k) = \frac{kp(k)}{\mu}.$$

Tous les enfants d'un noeud spécial sont normaux sauf un, choisi au hasard uniformément parmi tous les descendants (il faut remarquer qu'un noeud spécial a forcément au moins un enfant puisque $\hat{p}(0) = 0$).

Ainsi l'arbre de Kesten est un arbre comportant une épine dorsale infinie sur laquelle se greffent des arbres de Galton-Watson normaux. L'arbre de Kesten est appelé encore l'arbre biaisé par la taille car la loi de l'arbre tronqué au niveau n est absolument continue par rapport à la loi de l'arbre de Galton-Watson initial tronqué au même niveau, la densité étant donnée par la martingale classique Z_n/μ^n . Cet arbre a été étudié par Hawkes [25], Joffe et Waugh [32] et aussi par Lyons, Permantle et Peres [42]. Il peut également être vu comme un arbre de Galton-Watson avec immigration voir Athreya et Ney [8].

Avant de citer les fameux théorèmes de convergence, il est utile de rappeler une conséquence du théorème de Portemanteau qui caractérise la conver-

gence en loi des arbres : Soient $(T_n; n \in \mathbb{N}^*)$ une suite d'arbres aléatoires de \mathbb{T} et T un arbre aléatoire appartenant également à \mathbb{T} . On désigne par $\text{dist}(T)$ la distribution ou la loi de T (qui est parfaitement déterminée par la suite de distributions de $r_h(T)$ pour tout $h \geq 0$) et on note

$$\lim_{n \rightarrow +\infty} \text{dist}(T_n) = \text{dist}(T)$$

la convergence en loi de $(T_n; n \in \mathbb{N}^*)$ vers T .

Theorem 1.1.3. *La suite $(T_n; n \in \mathbb{N}^*)$ converge en loi vers T si et seulement si*

$$\forall h \in \mathbb{N}, \forall \mathbf{t} \in \mathbb{T}^{(h)}, \quad \lim_{n \rightarrow +\infty} \mathbb{P}(r_h(T_n) = \mathbf{t}) = \mathbb{P}(r_h(T) = \mathbf{t}).$$

On peut généraliser dans \mathbb{T}_∞ , si on travaille avec des arbres ayant des noeuds infinis, en utilisant $r_{h,\infty}$ au lieu de r_h et tout reste vrai. Nous nous référons à l'article de Abraham et Delmas [3] page 6 pour plus de détails.

Une caractérisation de convergence locale a été effectuée par Abraham et Delmas [2] afin d'étudier d'autres conditionnements. Avant de citer ce résultat, on introduit quelques notations :

Si $\mathbf{t}, \mathbf{t}' \in \mathbb{T}$ et $x \in \mathcal{L}_0(\mathbf{t})$, on note :

$$\mathbf{t} \otimes_x \mathbf{t}' = \{u \in \mathbf{t}\} \cup \{xv; v \in \mathbf{t}'\},$$

la greffe de \mathbf{t}' sur la feuille x de \mathbf{t} .

$$\mathbb{T}(\mathbf{t}, x) = \{\mathbf{t} \otimes_x \mathbf{t}'; \mathbf{t}' \in \mathbb{T}\}.$$

l'ensemble de tous les arbres obtenus en greffant un arbre sur la feuille x de \mathbf{t} .

Lemma 1.1.4. *Soient T et $T_n, n \in \mathbb{N}$ des arbres aléatoires tels que p.s $T \in \mathbb{T}_0 \cup \mathbb{T}_1$ et $T_n \in \mathbb{T}_0 \cup \mathbb{T}_1$ pour tout $n \in \mathbb{N}$. Alors $(T_n)_{n \in \mathbb{N}}$ converge en loi vers T si et seulement si pour tout arbre fini \mathbf{t} et toute feuille x de \mathbf{t}*

$$\lim_{n \rightarrow +\infty} \mathbb{P}(T_n \in \mathbb{T}(\mathbf{t}, x)) = \mathbb{P}(T \in \mathbb{T}(\mathbf{t}, x)) \quad \text{et} \quad \lim_{n \rightarrow +\infty} \mathbb{P}(T_n = \mathbf{t}) = \mathbb{P}(T = \mathbf{t}).$$

Un résultat analogue dans le cadre des arbres ayant des noeuds infinis est obtenu dans [3] page 7.

Le théorème suivant, considéré parmi les fameux résultats dans la convergence locale des arbres, a été prouvé par Kesten en 1986 en ajoutant la condition "la variance est finie".

Theorem 1.1.5. *(Kesten [35], 1986) Soit τ un arbre aléatoire de Galton-Watson critique ou sous-critique de loi de reproduction p vérifiant (1.1). Soit τ_n un arbre aléatoire dont la loi est celle de τ conditionné par $\{H(\tau) \geq n\}$. Alors*

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau_n) = \text{dist}(\tau^*).$$

Notons également que ce théorème a été prouvé aussi par Janson [29] en toute généralité.

Plus tard, Abraham et Delmas ont donné une preuve de ce résultat comme une application simple d'un résultat plus général :

Theorem 1.1.6. (Abraham, Delmas [2], 2014) *Soit $A : \mathbb{T} \rightarrow \mathbb{N} \cup \{\infty\}$ finie sur \mathbb{T}_0 . On suppose que A vérifie la condition d'additivité suivante : il existe une fonction $D(\mathbf{t}, x)$ telle que*

$$A(\mathbf{t} \otimes_x \mathbf{t}') = A(\mathbf{t}') + D(\mathbf{t}, x)$$

pour $A(\mathbf{t}')$ suffisamment grand. On pose

$$A_n = \{\mathbf{t} \in \mathbb{T}, A(\mathbf{t}) \in [n, n + n_0]\}.$$

Soit τ un arbre de Galton-Watson associé à une loi de reproduction p vérifiant (1.1). Si

- p est critique ou
- p est sous-critique et $D(\mathbf{t}, x) = |x|, \forall \mathbf{t} \in \mathbb{T}_0, x \in \mathcal{L}_0(\mathbf{t})$.

Alors, si

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{P}(\tau \in A_{n+1})}{\mathbb{P}(\tau \in A_n)} = \mu,$$

on a :

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid \tau \in A_n) = \text{dist}(\tau^*).$$

Grâce à ce théorème ; plusieurs résultats deviennent des applications simples, par exemple en posant :

- $A(\mathbf{t}) = H(\mathbf{t})$ et $n_0 = +\infty$, on obtient le théorème de Kesten.
- $A(\mathbf{t}) = \text{Card}(\mathbf{t})$ et $n_0 = 1$, on obtient dans le cas critique

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid \text{Card}(\tau) = n) = \text{dist}(\tau^*)$$

et en prenant $n_0 = +\infty$,

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid \text{Card}(\tau) \geq n) = \text{dist}(\tau^*).$$

Ce résultat a été prouvé par Aldous-Pitman [6], même s'il apparaît plus ou moins explicitement dans [34]. On peut également se référer à S.Janson [29], théorème 7.1.

- $A(\mathbf{t}) = \text{Card}(\mathcal{L}_0(\mathbf{t})) = L_0(\mathbf{t})$: nombre total des feuilles de \mathbf{t} . On obtient dans le cas critique pour $n_0 = 1$

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid L_0(\tau) = n) = \text{dist}(\tau^*)$$

et pour $n_0 = +\infty$

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid L_0(\tau) \geq n) = \text{dist}(\tau^*).$$

Ce résultat a été prouvé par Curien et Kortchemski [12] en 2014 en supposant que la variance est finie.

Nous nous référons à l'article [2] pour plus de détails et les preuves.

Ces résultats ont été généralisé par Abraham et Delmas en 2014 comme suit :

Soit A une partie de \mathbb{N} et \mathbf{t} un arbre. Posons

$$\mathcal{L}_A(\mathbf{t}) = \{u \in \mathbf{t}; k_u(\mathbf{t}) \in A\} \text{ et } L_A(\mathbf{t}) = \text{Card}(\mathcal{L}_A(\mathbf{t})).$$

Theorem 1.1.7. *Soit τ un arbre de Galton-Watson de loi de reproduction p vérifiant (1.1). Alors*

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid L_A(\tau) \geq n) = \text{dist}(\tau^*)$$

Remarquons que si :

- $A = \mathbb{N}$ on obtient $L_A(\mathbf{t}) = \text{Card}(\mathbf{t})$.
- $A = \{0\}$ on obtient L_A le nombre des feuilles de \mathbf{t} . ce qui est déjà fait.

Notons que L_A vérifie la propriété d'additivité :

$$L_A(\mathbf{t} \otimes_x \mathbf{t}') = L_A(\mathbf{t}') + L_A(\mathbf{t}) - \mathbf{1}_{\{0 \in A\}}.$$

D'après le théorème 1.1.6, il reste à prouver que le ratio tend vers 1. Ici ils utilisent le fait que $L_A(\tau)$ suit la loi du cardinal d'un autre arbre de Galton-Watson critique τ_A , propriété introduite par Mimami [43] et Rizzolo [51].

Notre travail, rédigé au chapitre 3, se situe dans la continuité des résultats présentés dans cette section. Nous étudions la convergence en loi d'un arbre de Galton-Watson conditionné à avoir un nombre assez grand de noeuds marqués, puis comme application nous étudions le conditionnement d'avoir un nombre assez grand de noeuds protégés.

Arbres de condensation

La situation est totalement différente dans le cas sous-critique. Ce problème a été étudié par Abraham et Delmas dans [3]. En fait c'est ce qu'on appelle le phénomène de condensation.

Avant d'énoncer le résultat obtenu, il nous faut au préalable introduire quelques objets.

Soit p une probabilité vérifiant (1.1) et $\mu(p) \leq 1$. On définit l'arbre de condensation $\tau^*(p)$ comme étant un arbre de Galton-Watson multi-type :

- Il existe deux types des noeuds : noeud normal et noeud spécial.
- La racine est spéciale.
- Les noeuds normaux se reproduisent suivant la loi p et tous leurs enfants sont normaux.

- Les noeuds spéciaux se reproduisent suivant la loi \tilde{p} : la loi biaisée sur $\mathbb{N} \cup \{+\infty\}$ définie par :

$$\tilde{p}(k) = \begin{cases} kp(k) & \text{si } k \in \mathbb{N} \\ 1 - \mu & \text{si } k = +\infty. \end{cases}$$

- Si un noeud spécial donne naissance à un nombre fini d'enfants, on en choisit 1 au hasard et il est spécial, les autres sont normaux.
- Si un noeud spécial donne naissance à un nombre infini d'enfants, ils sont tous normaux.
- Tous les noeuds se reproduisent d'une façon indépendante.

Notons également que deux cas qui se produisent :

- Si p est critique, alors p.s $\tau^*(p)$ est l'arbre de Kesten.
- Si p est sous-critique, alors p.s $\tau^*(p)$ a exactement un noeud infini et sans branche infinie.

Cet arbre a été introduit dans [31] où il apparaît comme arbre limite d'arbres de Galton-Watson sous-critiques conditionnés par leur population totale dans des cas particuliers. Ce résultat a été généralisé dans [29] à des lois de reproductions générales puis dans [3] à des conditionnements plus généraux et dont nous rappelons le résultat maintenant.

Soit $A \subset \mathbb{N}$. On pose

$$p_{A,\theta}(k) = \begin{cases} c_A(\theta)\theta^k p(k) & \text{si } k \in A \\ \theta^{k-1}p(k) & \text{si } k \notin A. \end{cases}$$

avec

$$c_A(\theta) = \frac{\theta - \mathbb{E}(\theta^X \mathbf{1}_{\{X \notin A\}})}{\theta \mathbb{E}(\theta^X \mathbf{1}_{\{X \in A\}})}$$

où X une variable aléatoire de loi p . On note par I_A l'ensemble des θ strictement positifs tel que $p_{A,\theta}$ est une probabilité.

Rappelons que si p est sous-critique, alors d'après le lemme 5.2 de [3] page 12, ou bien il existe un unique $\theta_A^c \in I_A$ tel que p_{A,θ_A^c} est critique (dans ce cas on dit que p est générique pour A), ou bien p_{A,θ_A^*} est sous-critique avec $\theta_A^* := \max(I_A)$ (dans ce cas on dit que p n'est pas générique pour A).

Definition 1.1.8. *Sous les mêmes notations, on définit p_A^* par :*

- Si p est critique : $p_A^* = p$
- Si p est sous-critique et générique pour A (il existe une unique $\theta_A^c \in I_A$ tel que p_{A,θ_A^c} est critique) : $p_A^* = p_{A,\theta_A^c}$
- Si p est sous-critique et non générique pour A : $p_A^* = p_{A,\theta_A^*}$.

Le résultat peut alors s'écrire :

Theorem 1.1.9. *Soit τ un arbre de Galton-Watson de loi de reproduction p vérifiant (1.1) et $\mu(p) \leq 1$. Alors*

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid L_A(\tau) \geq n) = \text{dist}(\tau^*(p_A^*)).$$

Notons également que ce théorème a été prouvé dans le cas critique et le cas sous-critique générique dans [2].

Conditionnement par la taille de la n -ième génération

Un autre conditionnement intéressant étudié par Abraham et Delmas dans [2] page 16, est le conditionnement sur la taille de la n -ième génération. Le résultat obtenu est le suivant :

Proposition 1.1.10. *Soit τ un arbre de Galton-Watson critique de loi de reproduction p vérifiant (1.1) et soit $(a_n)_{n \in \mathbb{N}}$ une suite d'éléments de \mathbb{N}^* . Si $\forall j \in \mathbb{N}^*$*

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{P}(Z_{n-j} = a_n)}{\mathbb{P}(Z_n = a_n)} = 1,$$

alors

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid Z_n(\tau) = a_n) = \text{dist}(\tau^*).$$

Puis ils s'intéressent à appliquer ce résultat dans le cas d'une loi géométrique critique où les calculs sont explicitement déterminés et prouvent que

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid Z_n(\tau) = a_n) = \text{dist}(\tau^*)$$

si $(a_n)_{n \in \mathbb{N}}$ une suite d'éléments de \mathbb{N}^* vérifiant $\lim_{n \rightarrow +\infty} \frac{a_n}{n^2} = 0$.

Notre travail, rédigé au chapitre 4, se situe dans la continuité de ce résultat. Nous considérons une loi géométrique dans un cadre plus général et nous donnons une description complète de l'arbre limite dans les trois régimes : sous-critique, critique et sur-critique et selon le comportement asymptotique de la taille de la génération.

1.2 Description des résultats obtenus

1.2.1 Chapitre 2 : Fonctions harmoniques discrètes dans un orthant

Ce travail est issu d'une collaboration avec monsieur Sami Mustapha et monsieur Mohamed Sifi et a fait l'objet d'un article [10] publié dans le journal ECP.

L'objectif de ce chapitre est d'établir une approche générale pour étudier le problème d'existence et unicité des fonctions harmoniques strictement positives associés aux marches aléatoires tuées au bord d'un orthant de \mathbb{Z}^d et de généraliser le quart du plan à l'orthant dans \mathbb{Z}^d .

Avant d'énoncer le résultat obtenu il nous faut au préalable introduire quelques objets.

Soit :

- $\Gamma = -\Gamma \subset \mathbb{Z}^d$ une partie finie symétrique de \mathbb{Z}^d contenant tous les vecteurs unitaires de \mathbb{Z}^d . i.e les vecteurs $e_k = (0, \dots, 0, 1, 0, \dots, 0)$, $k \in \{1, \dots, d\}$
- $\pi : \mathbb{Z}^d \times \Gamma \rightarrow [0, 1]$ une application vérifiant :

$$\forall x \in \mathbb{Z}^d, \quad \sum_{e \in \Gamma} \pi(x, e) = 1.$$

- $(S_n)_{n \in \mathbb{N}}$ une chaîne de Markov définie par :

$$\mathbb{P}(S_{n+1} = x + e \mid S_n = x) = \pi(x, e), \forall e \in \Gamma, x \in \mathbb{Z}^d \text{ et } n \in \mathbb{N}.$$

On suppose de plus que :

- $|\Gamma| < +\infty$: l'ensemble des sauts est fini.
- $\sum_{e \in \Gamma} \pi(x, e)e = 0$: marche aléatoire centrée.
- $\exists \alpha > 0$, tel que $\forall x \in \mathbb{Z}^d, |e| = 1, \pi(x, \pm e) \geq \alpha$: condition d'ellipticité.

Notons également que la condition d'ellipticité implique l'irréductibilité de la chaîne $(S_n)_{n \in \mathbb{N}}$.

Notons

$$\mathcal{D} = \{(x_1, \dots, x_d) \in \mathbb{Z}^d; x_1 > 0, \dots, x_d > 0\}$$

l'orthant de \mathbb{Z}^d , et

$$\partial \mathcal{D} = (\mathcal{D} + \Gamma) \cap \mathcal{D}^c = \{x \in \mathcal{D}^c \text{ tel que } x = z + e; z \in \mathcal{D}, e \in \Gamma\}$$

la frontière de \mathcal{D} . Posons $\bar{\mathcal{D}} = \mathcal{D} \cup \partial \mathcal{D}$.

Le résultat que nous obtenons peut alors s'écrire :

Theorem 1.2.1. *Il existe une unique, à une constante multiplicative près, fonction harmonique u strictement positive sur \mathcal{D} et nulle sur la frontière.*

$u : \bar{\mathcal{D}} \rightarrow \mathbb{R}$ vérifiant :

- $\forall x \in \mathcal{D}, u(x) = \sum_{e \in \Gamma} \pi(x, e)u(x + e)$.
- $\forall x \in \mathcal{D}, u(x) > 0$.
- $\forall x \in \partial \mathcal{D}, u(x) = 0$.

La preuve de ce résultat repose sur les outils de la théorie du potentiel discrète à savoir le principe de maximum, inégalités de Harnack, inégalités de Harnack au bord et mesure harmonique discrète. En fait ces outils jouent un rôle crucial dans l'établissement d'estimations optimales pour les fonctions harmoniques.

Notons également que dans le travail de Sami Mustapha [44], ces outils servent à établir aussi des estimations gaussiennes pour le noyau de transition d'une marche aléatoire inhomogène ainsi que pour son noyau de Green.

1.2.2 Chapitre 3 : Convergence d'un arbre de Galton Watson critique conditionné à avoir un nombre infini de noeuds protégés

Ce travail est issu d'une collaboration avec Romain Abraham et Jean-François Delmas et a fait l'objet d'un article [1] publié dans le journal "Journal of applied probability".

Soit \mathbf{t} un arbre fixé ; on place aléatoirement des marques sur les noeuds de \mathbf{t} indépendamment les uns des autres avec une probabilité qui dépend uniquement du nombre d'enfants du noeud. Plus précisément, on considère une fonction de marquage $q : \mathbb{N} \rightarrow [0, 1]$ et une famille de variables aléatoires $(Z_u(\mathbf{t}), u \in \mathbf{t})$ telle que $\mathbb{P}(Z_u(\mathbf{t}) = 1) = q(k_u(\mathbf{t}))$. On dit qu'un noeud u de \mathbf{t} est marqué si $Z_u(\mathbf{t}) = 1$.

On désigne par $M(\mathbf{t})$ le nombre des noeuds marqués de \mathbf{t} . Alors on a le résultat suivant :

Theorem 1.2.2. *Soit τ un arbre de Galton-Watson critique de loi de reproduction p vérifiant (1.1). On suppose qu'il existe $k \in \mathbb{N}$ tel que $p(k)q(k) > 0$. Alors*

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid M(\tau) = n) = \text{dist}(\tau^*).$$

En posant $q(k) = \mathbf{1}_{\{k \in A\}}$ on obtient le théorème 1.1.7.

La preuve de ce résultat contient deux parties : la première partie consiste en une adaptation de la preuve du théorème 1.1.6 de [2]. Dans la deuxième partie, la preuve repose sur que le fait $M(\tau)$ suit la même loi que le cardinal d'un arbre de Galton-Watson critique dont la construction est identique à celle de Rizzolo.

Le reste de notre article est réservé à l'étude du conditionnement par les noeuds protégés. Un noeud protégé est un noeud qui n'est pas une feuille et dont aucun des ses enfants est une feuille. Le comportement asymptotique du nombre de noeuds protégés dans un arbre de Galton-Watson conditionné à sa population totale a par exemple été étudié par Devroye et Janson [13].

Le résultat que nous obtenons peut alors s'écrire :

Theorem 1.2.3. *Soit τ un arbre de Galton-Watson critique de loi de reproduction p vérifiant (1.1). On note $A(\tau)$ le nombre des noeuds protégés de τ . Alors*

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau \mid A(\tau) = n) = \text{dist}(\tau^*).$$

Notons tout d'abord que A vérifie la propriété d'additivité du théorème 1.1.6 de [2]. Pour finir la preuve de ce résultat, il suffit donc de démontrer que :

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{P}(A(\tau) = n + 1)}{\mathbb{P}(A(\tau) = n)} = 1.$$

Dans cette étape, nous avons utilisé le résultat sur les noeuds marqués.

Tout d'abord, soit τ^0 un arbre aléatoire ayant la même loi que $\tau \mid \{k_\emptyset(\tau) > 0\}$. En utilisant le théorème précédent, avec $q(k) = \mathbf{1}_{\{k > 0\}}$, (ou bien Theorem 6 and Corollary 2 dans [51] avec $A = \mathbb{N}^*$) on a que l'arbre $\tau_{\mathbb{N}^*}^0$ (qui est l'arbre τ^0 auquel on a retiré toutes les feuilles) est un arbre de Galton-Watson critique. Sachant $\{\tau_{\mathbb{N}^*}^0 = \mathbf{t}\}$, on ajoute des enfants aux noeuds de $\tau_{\mathbb{N}^*}^0$ de telle sorte qu'on obtient un arbre $\hat{\tau}$ ayant la même loi que l'arbre initial τ^0 .

Dans cette procédure, on marque les noeuds de $\tau_{\mathbb{N}^*}^0$ dont le nombre d'enfants ajoutés est 0. Alors $M(\tau_{\mathbb{N}^*}^0)$, le nombre des noeuds marqués, est égale en loi à $A(\hat{\tau})$, le nombre de noeuds protégés, ce qui termine la preuve.

1.2.3 Chapitre 4 : Convergence d'un arbre de Galton Watson géométrique conditionné à avoir une génération anormalement grande

Ce travail, rédigé au chapitre 4, est issu d'une collaboration avec Romain Abraham et Jean-François Delmas et fait l'objet d'un article qui sera prochainement soumis.

L'objectif de ce chapitre est d'étudier le problème de conditionnement d'un arbre de Galton-Watson géométrique à avoir une génération anormalement grande dont nous traitons tous les cas sous-critique, critique et sur-critique. Plus précisément

Soit $\eta \in (0, 1]$ et $q \in (0, 1)$. On considère une loi géométrique $p = (p(k), k \in \mathbb{N})$ définie par :

$$\begin{cases} p(0) = 1 - \eta, \\ p(k) = \eta q(1 - q)^{k-1} \quad \text{pour } k \in \mathbb{N}^*. \end{cases} \quad (1.2)$$

Nous avons étudié la limite en loi de $\tau \mid Z_n = a_n$ et les résultats que nous obtenons se résument dans le tableau suivant :

	régime de Kesten	régime de Poisson	régime de Condensation
Sous-critique	$a_n \ll \mu^{-n}$	$a_n \sim \theta \mu^{-n}$	$a_n \gg \mu^{-n}$
Critique	$a_n \ll n^2$	$a_n \sim \theta n^2$	$a_n \gg n^2$
Sur-critique	$a_n \ll \mu^n$	$a_n \sim \theta \mu^n$	$a_n \gg \mu^n$

Notons également que :

- Dans le régime de Kesten, l'arbre limite est l'arbre de Kesten. Nous renvoyons le lecteur à chapitre 4, section 4 pour la description complète de l'arbre de Kesten.
- Dans le régime de Poisson, l'arbre limite est un arbre multi-type mais qui ne vérifie par la propriété de branchement. Il est constitué

d'un squelette infini obtenu par immigration poissonnienne à chaque génération sur lequel se greffent des arbres de Galton-Watson normaux conditionnés à s'éteindre. Nous renvoyons le lecteur à chapitre 4, section 5 pour plus de détails.

- Dans le régime de condensation, l'arbre limite obtenu est un arbre multi-type dont la racine est un noeud infini et sur les individus de la première génération sont greffés des arbres de Galton-Watson indépendants dont la loi de reproduction dépend de la génération du noeud parent. Cet arbre vérifie à nouveau la propriété de branchement. La section 6 est consacrée pour la description complète de cet arbre. Nous avons donné aussi une représentation avec les martingales qui sera très utile d'établir le résultat de convergence.

Les preuves de ces résultats reposent sur les deux caractérisations de convergence en loi suivantes :

- Soient $(T_n, n \in \mathbb{N})$ et T des arbres aléatoires sans noeuds infinis. On suppose que $H(T) = +\infty$ p.s. Alors :

$$T_n \xrightarrow[n \rightarrow \infty]{(d)} T \iff \lim_{n \rightarrow \infty} \mathbb{P}(r_h(T_n) = \mathbf{t}) = \mathbb{P}(r_h(T) = \mathbf{t}) \quad \text{pour tout } h \in \mathbb{N}^*, \mathbf{t} \text{ arbre fini d'ordre } h$$
(1.3)

- Soient $(T_n, n \in \mathbb{N})$ et T des arbres aléatoires dans $\mathbb{T}_f^* = \{\mathbf{t} \in \mathbb{T}_\infty; \forall u \in \mathbf{t}^*, k_u(\mathbf{t}) < +\infty\}$, on suppose que p.s $H(T) = +\infty$ et $k_\emptyset(T) = +\infty$ alors :

$$T_n \xrightarrow[n \rightarrow \infty]{(d)} T \iff \lim_{n \rightarrow \infty} \mathbb{P}(r_{h,k}(T_n) = \mathbf{t}) = \mathbb{P}(r_{h,k}(T) = \mathbf{t}) \quad \text{pour tout } h, k \in \mathbb{N}^*, \text{ et pour tout } \mathbf{t} \in \mathbb{T}_f^*$$
(1.4)

Chapitre 2

Discrete harmonic functions on an orthant in \mathbb{Z}^d

This work comes from a collaboration with Sami Mustapha and Mohamed Sifi and has been the object of an article [10] published in the journal *Electronic Communications in Probability*.

Abstract. We give a positive answer to a conjecture on the uniqueness of harmonic functions stated in [50]. More precisely we prove the existence and uniqueness of a positive discrete harmonic function for a random walk satisfying finite range, centering and ellipticity conditions, killed at the boundary of an orthant in \mathbb{Z}^d .

2.1 Introduction

An explicit description of the Martin compactification for random walks is usually a non-trivial problem and the most of the existing results are obtained for so-called homogeneous random walks, when the transition probabilities of the process are invariant with respect to the translations over the state space E (see [15], [19], [48], [53]). In [11] Cartier identified the Martin compactification for random walks on non-homogeneous trees. Doney [14] described the harmonic functions and the Martin boundary of a random walk $(Z(n))$ on \mathbb{Z} killed on the negative half-line $\{z : z < 0\}$. Alili and Doney [7] extend this result for the corresponding space-time random walk $S(n) = (Z(n), n)$. Kurkova and Malyshev [39] described the Martin boundary for nearest neighbor random walks on $\mathbb{Z} \times \mathbb{N}$ and on \mathbb{Z}_+^2 with reflected conditions on the boundary. The recent results of Raschel [50] and Kurkova and Raschel [40] identify the Martin compactification for random walks in \mathbb{Z}_+^2 with jumps at distance at most 1 and absorbing boundary. All these results use methods that seem to be unlikely to apply in a general situation. The methods of Doney [14] and Alili and Doney [7] rely in an es-

quential way on the one-dimensional structure of the process $(Z(n))$. Kurkova and Malyshev [39], Raschel [50] and Kurkova and Raschel [40] use an analytical method where the geometrical structure of the elliptic curve defined by the jump generating function of the random walk plays a crucial role. For small steps walks the methods use the fact that the corresponding elliptic curve is homeomorphic to the torus. In a very recent paper Fayolle and Raschel [21] show how to extend the method for random walks with arbitrary big jumps (in this case the algebraic curve is no more homeomorphic to the torus).

The aim of our chapter is to offer a general approach allowing to study positive harmonic functions for random walks killed at the boundary of an orthant in \mathbb{Z}^d without assuming invariance with respect the translations over the state space.

More precisely let $\Gamma = -\Gamma \subset \mathbb{Z}^d$, be a symmetric finite subset of \mathbb{Z}^d containing all unit vectors in \mathbb{Z}^d , and $\pi : \mathbb{Z}^d \times \Gamma \rightarrow [0, 1]$ be such that

$$\sum_{e \in \Gamma} \pi(x, e) = 1, \quad x \in \mathbb{Z}^d.$$

Then, we let $(S_j)_{j \in \mathbb{N}}$, be the Markov chain defined by

$$\mathbb{P}[S_{j+1} = x + e / S_j = x] = \pi(x, e), \quad e \in \Gamma, x \in \mathbb{Z}^d, j = 0, 1, \dots$$

We shall impose the following three conditions on the step set Γ and the probabilities $\pi(x, e)$:

(Finite range) $|\Gamma| < \infty$,

(Centering) $\sum_{e \in \Gamma} \pi(x, e)e = 0$,

(Ellipticity) there exists $\alpha > 0$ such that $\pi(x, \pm e_i) \geq \alpha$; $x \in \mathbb{Z}^d$, $i = 1, \dots, d$.

We shall denote by :

$$\mathcal{D} = \{(x_1, \dots, x_d) \in \mathbb{Z}^d; \quad x_1 > 0, \dots, x_d > 0\}$$

and we shall be interested in characterizing the positive harmonic functions for the random walk killed at the boundary of the orthant \mathcal{D} , i.e in functions $u : \overline{\mathcal{D}} \rightarrow \mathbb{R}$ such that :

i) For any $x \in \mathcal{D}$, $u(x) = \sum_{e \in \Gamma} \pi(x, e)u(x + e)$,

ii) if $x \in \mathcal{D}$, $u(x) > 0$,

iii) if $x \in \partial\mathcal{D}$, $u(x) = 0$,

where we denote by

$$\partial\mathcal{D} = \{x \in \mathcal{D}^c / x = z + e, \quad \text{for some } z \in \mathcal{D}, \text{ and } e \in \Gamma\}; \quad \text{and } \overline{\mathcal{D}} = \mathcal{D} \cup \partial\mathcal{D}.$$

Theorem 2.1.1. *Let $(S_j)_{j \in \mathbb{N}}$ be a spatially inhomogeneous random walk satisfying the above centering, finite range and ellipticity conditions. Then, up to a multiplicative constant, there exists a unique positive harmonic function for the random walk killed at the boundary.*

Note that general results on harmonic functions for killed random walks in half spaces $\mathbb{Z}^+ \times \mathbb{Z}^{d-1}$ and orthants were obtained in the non-zero drift case in [28, 40]. In this case, the Martin boundary is composed of infinitely many harmonic functions. Regarding random walks with zero drift, important results were obtained recently by Raschel in [50] in the quarter plane. In a sense our results can be seen as complementing those of [50] because they provide an answer to the conjecture stated in [50] in the case of random walks with finite support (Cf. Conjecture 1, §4.1 of [50]). Our methods, using ideas introduced in [44, 45], allow on the other hand to generalize from the quarter plane to orthants in higher dimensions and to treat the spatially inhomogeneous walks.

The chapter is organized as follows. Section 2 gives in its first part the main technical ingredients in the proof of Theorem 1. A lower estimate of the harmonic measure providing a Hölder continuity property at the boundary is established. Specifically, Proposition 1 gives a control of the growth of positive harmonic functions vanishing on the boundary. A Carleson-type estimate is then derived. In section 2.2 we establish the existence of a positive harmonic function on \mathcal{D} . Section 3 is devoted to prove the uniqueness of such positive harmonic function on \mathcal{D} .

2.2 Proof of Theorem 2.1.1

2.2.1 Key technical ingredients

The proof of Theorem 2.1.1 relies on a systematic use of two fundamental principles of potential theory, maximum principle and Harnack principle.

Let $A \subset \mathbb{Z}^d$ denote a bounded domain (i.e. a finite connected set of vertices of \mathbb{Z}^d). We let

$$\partial A = \{x \in A^c / x = z + e, \text{ for some } z \in A, \text{ and } e \in \Gamma\}; \quad \text{and } \bar{A} = A \cup \partial A.$$

We say that a function $u : \bar{A} \rightarrow \mathbb{R}$ is harmonic on A if

$$u(x) = \sum_{e \in \Gamma} \pi(x, e) u(x + e), \quad x \in A.$$

The following maximum principle is immediate

Theorem 2.2.1. (Maximum principle) *Let $A \subset \mathbb{Z}^d$ be a bounded domain in \mathbb{Z}^d and $u : \bar{A} \rightarrow \mathbb{R}$ a harmonic function on A . Assume that $u \geq 0$ on ∂A . Then $u \geq 0$ in A .*

The following theorem is a centered version of the Harnack inequality established by Lawler [41] for spatially inhomogeneous random walk with bounded symmetric increments.

Theorem 2.2.2. (*Harnack principle*) (Cf. [37]) *Let $u : B_{2R}(y) \rightarrow \mathbb{R}$ a nonnegative harmonic function on a discrete Euclidean ball centered at y and with radius $2R$ ($R = 1, 2, \dots$ and $y \in \mathbb{Z}^d$). Then*

$$\max_{x \in B_R(y)} u(x) \leq C \min_{x \in B_R(y)} u(x),$$

where $C = C(d, \alpha, \Gamma)$ is independent of $y \in \mathbb{Z}^d$, $R \geq 1$ and u .

In the proof of Theorem 2.1.1, together with Theorems 2.2.1 and 2.2.2, we need their parabolic versions.

Let $B = A \times \{a \leq k \leq b\} \subset \mathbb{Z}^d \times \mathbb{Z}$ where $A \subset \mathbb{Z}^d$ is a bounded domain of \mathbb{Z}^d and $a < b \in \mathbb{Z}$. We let

$$\begin{aligned} \partial_\ell B &= \cup_{a < k < b} \partial A \times \{k\} \\ \partial_p B &= \partial_\ell B \cup (\bar{A} \times \{a\}) \\ \bar{B} &= B \cup \partial_p B \end{aligned}$$

$\partial_p B$ is the parabolic boundary of B and $\partial_\ell B$ its lateral boundary. We say that $u : \bar{B} \rightarrow \mathbb{R}$ is caloric in B if

$$u(x, k+1) = \sum_{e \in \Gamma} \pi(x, e) u(x+e, k), \quad (x, k) \in A \times \{a \leq k < b\}.$$

Caloric functions share with harmonic functions the fundamental property that there exist quantitative relations between their values at different points. Maximum principle extends so that we have a control of the values of a caloric function inside the cylinder B by its values on the parabolic boundary $\partial_p B$.

Harnack principle generalizes for nonnegative caloric functions on two conditions : being at distance R - in the parabolic sense- from the parabolic boundary and observing a waiting period of time $\simeq R^2$. This requires to be at a distance R^2 from the bottom and R from the lateral boundary in addition to the waiting time. We then obtain a control of the values of the function in the lower sub-cylinder $B_R(y) \times \{s-3R^2 \leq k \leq s-2R^2\}$ by those in the upper sub-cylinder $B_R(y) \times \{s-R^2 \leq k \leq s\}$.

Theorem 2.2.3. (*Parabolic maximum principle*) *Let $B = A \times \{a \leq k \leq b\} \subset \mathbb{Z}^d \times \mathbb{Z}$, where $a < b \in \mathbb{Z}$ and $A \subset \mathbb{Z}^d$ is a bounded domain in \mathbb{Z}^d , and $u : \bar{B} \rightarrow \mathbb{R}$ a caloric function in B . If $u \geq 0$ on $\partial_p B$ then $u \geq 0$ on B .*

The following theorem is a parabolic version of Harnack principle

Theorem 2.2.4. (Cf. [38]) Let u be a nonnegative caloric function on $B_{2R}(y) \times [s - 4R^2, s]$, where $(y, s) \in \mathbb{Z}^d \times \mathbb{Z}$ and $R \geq 1$. Then

$$\begin{aligned} \max \{u(x, k); \quad x \in B_R(y), \quad s - 3R^2 < k < s - 2R^2\} \\ \leq C \min \{u(x, k); \quad x \in B_R(y), \quad s - R^2 < k < s\}, \end{aligned}$$

where $C = C(d, \alpha, \Gamma)$ is independent of $(y, s) \in \mathbb{Z}^d \times \mathbb{Z}$, $R \geq 1$ and u .

2.2.2 A lower estimate of the harmonic measure

For a finite set $A \subset \mathbb{Z}^d$ we denote by τ_A the exit time from A . We define the harmonic measure $h_A^x(E)$ ($E \subset \partial A$) of $(S_j)_{j \in \mathbb{N}}$ in A at the point $x \in A$, by

$$h_A^x(E) = \mathbb{P}_x [S_{\tau_A} \in E].$$

Lemma 2.2.5. Let $y \in \partial \mathcal{D}$ and $R \geq 1$ denote a large integer. Then

$$h_{B_{2R}(y) \cap \mathcal{D}}^x [\partial \mathcal{D} \cap B_{2R}(y)] \geq \theta, \quad x \in B_R(y) \cap \mathcal{D} \quad (2.1)$$

where $\theta = \theta(d, \alpha, \Gamma) > 0$ is independent of y and R .

The proof of (2.1) we give below is the only place where caloric functions are used. It may seem strange that an issue regarding harmonic functions requires going through caloric functions. Indeed, it would have been more natural to estimate the left hand side of (2.1) by $h_{B_{2R}(y)}^x(\partial B_{2R}(y) \cap \mathcal{D}^c)$ (using the maximum principle) and then establish the desired lower estimate for $h_{B_{2R}(y)}^x(\partial B_{2R}(y) \cap \mathcal{D}^c)$. Such an approach may lead to the conclusion provided we have good estimates of the Green Kernel as in [27] p. 511; or by using a barrier argument as in [9] p.157, but this is difficult to implement for inhomogeneous walks. A way to get round the difficulty is to apply Harnack principle to a harmonic extension of $h_{B_{2R}(y)}^x(\partial B_{2R}(y) \cap \mathcal{D}^c)$, but harmonic functions are difficult to extend; hence the idea of going through a caloric function defined on a cylinder outside of $\mathcal{D} \times \mathbb{Z}$ and much easier to extend so that it is possible to apply parabolic Harnack principle. The idea of such a construction is inspired by the proof of Lemma 4.1 in [52].

Démonstration. Let $z = y - \frac{R}{2}e_i$ ($i = 1, \dots, d$) (the direction e_i to be chosen in an appropriate way depending on the piece of the boundary to which y belongs) and u_1, u_2 and u_3 the three caloric functions defined by :

- $u_1 : (B_{2R}(y) \cap \mathcal{D}) \times \{-4R^2 \leq k \leq 0\} \rightarrow \mathbb{R}$,

$$u_1(x, k) = h_{B_{2R}(y) \cap \mathcal{D}}^x (\partial \mathcal{D} \cap B_{2R}(y)).$$

- $u_2 : \Omega = B_{2R}(y) \times \{-4R^2 \leq k \leq 0\} \rightarrow \mathbb{R}$, the solution of the parabolic boundary problem

$$\left\{ \begin{array}{l} u_2(x, k+1) = \sum_{e \in \Gamma} \pi(x, e) u_2(x+e, k), \quad \text{if } (x, k) \in B_{2R}(y) \times \{-4R^2 \leq k < 0\} \\ u_2(x, -4R^2) = 1_{B_{R/8}(z)}(x) \\ u_2|_{\partial_p \Omega \setminus B_{R/8}(z)} = 0 \end{array} \right.$$

- $u_3 : \Omega' = B_{R/8}(z) \times \{-4R^2 \leq k \leq 0\} \rightarrow \mathbb{R}$, the solution of the parabolic boundary problem

$$\left\{ \begin{array}{l} u_3(x, k+1) = \sum_{e \in \Gamma} \pi(x, e) u_3(x+e, k), \quad \text{if } (x, k) \in B_{R/8}(z) \times \{-4R^2 \leq k < 0\} \\ u_3(x, -4R^2) = 1_{B_{R/8}(z)}(x) \\ u_3|_{\partial_t \Omega'} = 0 \end{array} \right.$$

The parabolic maximum principle implies :

$$u_1(x, k) \geq u_2(x, k), \quad (x, k) \in (\mathcal{D} \cap B_{2R}(y)) \times \{-4R^2 \leq k \leq 0\}$$

$$u_2(x, k) \geq u_3(x, k), \quad (x, k) \in B_{R/8}(z) \times \{-4R^2 \leq k \leq 0\}$$

Combining with parabolic Harnack principle gives that for $x \in B_R(y) \cap \mathcal{D}$

$$h_{B_{2R}(y) \cap \mathcal{D}}^x(\partial \mathcal{D} \cap B_{2R}(y)) \geq c u_3(z, -R^2) \geq \theta > 0,$$

simply because the caloric function u_3 can be extended to a larger domain (e.g. $B_{R/8}(z) \times \{-10R^2 \leq k \leq 0\}$ so that $u_3 = 1$ on $B_{R/8}(z) \times \{-10R^2 \leq k \leq -4R^2\}$. \square

2.2.3 The Hölder continuity at the boundary

An important consequence of the lower bound of Lemma 2.2.5 is that :

Proposition 2.2.6. *Let $y \in \partial \mathcal{D}$ and $R \geq 1$ denote a large integer. Let u a non negative harmonic function in $\mathcal{D} \cap B_{2R}(y)$ which vanishes on $\partial \mathcal{D} \cap B_{2R}(y)$. Assume that $R \geq C$. Then*

$$\max_{x \in \mathcal{D} \cap B_R(y)} u(x) \leq \rho \max_{x \in \mathcal{D} \cap B_{2R}(y)} u(x) \quad (2.2)$$

with $0 < \rho = \rho(d, \alpha, \Gamma) < 1$.

Démonstration. Let $x \in \mathcal{D} \cap B_R(y)$. Let τ denote the exit time from $\mathcal{D} \cap B_{2R}(y)$.

We have

$$\begin{aligned} u(x) &= \sum_{z \in \partial(\mathcal{D} \cap B_{2R}(y))} u(z) \mathbb{P}_x [S_\tau = z] \\ &= \sum_{z \in \partial(\mathcal{D} \cap B_{2R}(y)) \setminus (\partial \mathcal{D} \cap B_{2R}(y))} u(z) \mathbb{P}_x [S_\tau = z]. \end{aligned}$$

Hence

$$\begin{aligned} u(x) &\leq \mathbb{P}_x [S_\tau \in \partial(\mathcal{D} \cap B_{2R}(y)) \setminus (\partial \mathcal{D} \cap B_{2R}(y))] \max_{x \in \mathcal{D} \cap B_{2R}(y)} u(x) \\ &= (1 - \mathbb{P}_x [S_\tau \in \partial \mathcal{D} \cap B_{2R}(y)]) \max_{x \in \mathcal{D} \cap B_{2R}(y)} u(x). \end{aligned}$$

Using Lemma 2.2.5 and taking supremum on $x \in \mathcal{D} \cap B_R(y)$ we deduce that :

$$\max_{x \in \mathcal{D} \cap B_R(y)} u(x) \leq (1 - \theta) \max_{x \in \mathcal{D} \cap B_{2R}(y)} u(x),$$

which implies (2.2) with $\rho = 1 - \theta$. \square

2.2.4 The Carleson principle

This principle asserts that

$$\sup_{x \in \mathcal{D} \cap B_R} u(x) \leq C u(Re), \quad R \geq C, \quad (2.3)$$

where u is a positive harmonic function on \mathcal{D} , vanishing on $\partial \mathcal{D} \cap B_{2R}$ and where B_R denote the R - discrete ball centered at the origin and $e = (1, \dots, 1)$. The positive constant C in (2.3) depends on d, α , and Γ , i.e $C = C(d, \alpha, \Gamma)$ but is independent of R and u .

Démonstration. To prove (2.3) we first observe that positive harmonic functions satisfy (thanks to the ellipticity condition) an obvious local Harnack principle. This local Harnack principle allows us to assume that to distance of $x = (x_1, \dots, x_d) \in \mathcal{D} \cap B_{2R}$ to the boundary $\partial \mathcal{D}$ is sufficiently large. Without loss of generality we may assume that $\min_{1 \leq i \leq d} x_i = x_1 \geq C$. On the other hand, the Harnack principle implies (by the classical chain argument) that

$$u(x) \leq C \left(\frac{R}{x_1} \right)^\gamma u(Re); \quad x \in B_{2R} \cap \mathcal{D} \quad (2.4)$$

where γ and C are positive constants depending only on d, α and Γ . Let $x_0 = (0, x_2, \dots, x_d) \in \partial \mathcal{D}$. By proposition 2.2.6 applied to the positive harmonic function in $B_{2x_1}(x_0) \cap \mathcal{D}$, we have

$$u(x) \leq \rho u(z),$$

where $z \in \overline{\mathcal{D} \cap B_{2x_1}(x_0)}$ satisfies

$$\max_{\mathcal{D} \cap B_{2x_1}(x_0)} u(x) = u(z).$$

We distinguish two cases : (i) where $x_1 \leq \frac{1}{2^N}(2R - |x|)$ and (ii) where $x_1 > \frac{1}{2^N}(2R - |x|)$. The constant N that appears in the definition of these conditions is appropriately large and will be chosen later.

Case (i) : $x_1 \leq \frac{1}{2^N}(2R - |x|)$.

In this case we observe that $z \in \mathcal{D} \cap B_{2R}$. This because :

$$\begin{aligned} |z| &\leq |z - x| + |x| \leq 8x_1 + |x| \\ &\leq \frac{1}{2^{N-3}}(2R - |x|) + |x| \\ &\leq \frac{R}{2^{N-4}} + (1 - \frac{1}{2^{N-3}})|x| \\ &= \frac{R}{2^{N-4}} + (1 - \frac{1}{2^{N-3}})2R = 2R, \end{aligned}$$

and that

$$2R - |z| \geq 2R - |x| - 8x_1 \geq (1 - \frac{1}{2^{N-3}})(2R - |x|)$$

From which it follows that

$$\begin{aligned} (2R - |x|)^\gamma u(x) &\leq (1 - \frac{1}{2^{N-3}})^{-\gamma} (2R - |z|)^\gamma \rho u(z) \\ &\leq (1 - \frac{1}{2^{N-3}})^{-\gamma} \rho \max_{x \in \mathcal{D} \cap B_{2R}} (2R - |x|)^\gamma u(x) \end{aligned}$$

and we choose N large enough such that $(1 - \frac{1}{2^{N-3}})^{-\gamma} \rho = \rho' < 1$.

Case (ii) : $x_1 > \frac{1}{2^N}(2R - |x|)$.

In this case we observe that :

$$(2R - |x|)^\gamma u(x) \leq 2^{\gamma N} x_1^\gamma u(x) \leq C 2^{\gamma N} R^\gamma u(Re)$$

where the last inequality follows from (2.4). Combining case (i) and case (ii) we deduce that for all $x \in \mathcal{D} \cap B_{2R}$

$$(2R - |x|)^\gamma u(x) \leq \rho' \max_{x \in \mathcal{D} \cap B_{2R}} (2R - |x|)^\gamma u(x) + C 2^{\gamma N} R^\gamma u(Re).$$

Using the fact that $\rho' < 1$ and $(2R - |x|) \simeq R$ for $x \in \mathcal{D} \cap B_R$ we deduce the estimate (2.3). \square

2.2.5 Existence of a positive harmonic function on \mathcal{D}

In order to establish the existence of a positive harmonic function on \mathcal{D} vanishing on $\partial\mathcal{D}$, let :

$$u_l(x) = \alpha_l(G_{l+1}(x, e) - G_l(x, e)), \quad x \in B_{2^l} \cap \mathcal{D},$$

where $G_l(x, e); l = 1, \dots$ is the Green function of $(S_n)_{n \in \mathbb{N}}$ in the discrete ball B_{2^l} , with pole at $e = (1, \dots, 1)$ and where the α_l are chosen so that $u_l(e) = 1; l = 1, 2, \dots$. Harnack principle combined with Carleson estimate (2.3) implies that the u_l satisfy

$$u_l(x) \leq C, \quad x \in B_{2^k} \cap \mathcal{D}; \quad l > k,$$

with a constant $C = C(k)$ depending only on k . The diagonal process then allows us to deduce the existence of a positive harmonic function defined globally in \mathcal{D} and vanishing on $\partial\mathcal{D}$. More precisely, for each $k = 1, 2, \dots$ there exists an extraction φ_k so that $(u_{\varphi_k(n)})_{n \geq 1}$ converges on $B_{2^k} \cap \mathcal{D}$. An easy induction on k gives a sequence of harmonic functions $(u_{\varphi_1 \circ \varphi_2 \circ \dots \circ \varphi_k(n)})_{n \geq 1}$ converging on $B_{2^k} \cap \mathcal{D}$. Taking $\psi(n) = \varphi_1 \circ \dots \circ \varphi_n(n)$ and setting $u(y) = \lim_{n \rightarrow \infty} u_{\psi(n)}(y)$ gives the required harmonic function.

2.3 Uniqueness of the positive harmonic function

2.3.1 Boundary Harnack principle

The previous section provides us a positive harmonic function in \mathcal{D} vanishing on $\partial\mathcal{D}$. It remains to show that up to a multiplicative constant, this function is unique. One way that this can be seen is by using boundary Harnack principle.

Theorem 2.3.1. *Assume that u and v are two non-negative harmonic functions in \mathcal{D} . Assume that $u = 0$ on $\partial\mathcal{D}$. Then*

$$\sup_{x \in \mathcal{D} \cap B_{R/2}} \frac{u(x)}{v(x)} \leq C \frac{u(Re)}{v(Re)}, \quad R \geq C, \quad (2.5)$$

where $C = C(d, \alpha, \Gamma) > 0$.

Démonstration. Using (2.3) we see that

$$u(x) \leq Cu(Re), \quad x \in \partial B_R \cap \mathcal{D}.$$

Hence the maximum principle implies that

$$u(x) \leq Cu(Re)h_{B_R \cap \mathcal{D}}^x[\partial B_R \cap \mathcal{D}] \quad (2.6)$$

for all $x \in B_R \cap \mathcal{D}$. For $x \in (\partial B_R \cap \mathcal{D}) \cap \{x_i \geq \epsilon R, i = 1, \dots, d\}$ (for an appropriate small $\epsilon > 0$), Harnack principle gives

$$v(x) \geq cv(Re),$$

and so the maximum principle implies that

$$v(x) \geq cv(Re)h_{B_R \cap \mathcal{D}}^x [(\partial B_R \cap \mathcal{D}) \cap \{x_i \geq \epsilon R, i = 1, \dots, d\}] \quad (2.7)$$

for all $x \in B_R \cap \mathcal{D}$. The proof will follow from the following claim.

Claim. Let $x = (x_1, \dots, x_d) \in B_{R/2} \cap \mathcal{D}$ with $0 < x_i < \epsilon R$ for some $i = 1, \dots, d$. Then

$$h_{B_R \cap \mathcal{D}}^x [\partial B_R \cap \mathcal{D}] \leq C h_{B_R \cap \mathcal{D}}^x [(\partial B_R \cap \mathcal{D}) \cap \{x_i \geq \epsilon R, i = 1, \dots, d\}] \quad (2.8)$$

provided that ϵ is small enough.

Let us take the claim for granted and show how to finish the proof. We can assume, in the proof of (2.5), that $\text{dist}(x, \partial \mathcal{D}) \leq \epsilon R$ (otherwise (2.5) is an immediate consequence of Harnack principle). Combining (2.6) and (2.8) we deduce that

$$\frac{u(x)}{u(Re)} \leq C h_{B_R \cap \mathcal{D}}^x [(\partial B_R \cap \mathcal{D}) \cap \{x_i \geq \epsilon R, i = 1, \dots, d\}].$$

Since

$$h_{B_R \cap \mathcal{D}}^x [(\partial B_R \cap \mathcal{D}) \cap \{x_i \geq \epsilon R, i = 1, \dots, d\}] \leq C \frac{v(x)}{v(Re)},$$

by (2.7), our inequality is verified. \square

Proof of the claim. Let $x = (x_1, \dots, x_d) \in B_{R/2} \cap \mathcal{D}$. Without loss of generality we can assume that $x_1 = \text{dist}(x, \partial \mathcal{D})$ and suppose that $0 < x_1 < \epsilon R$. It follows from Harnack principle (using, as in §2.4, a sequence of balls, connecting x to $\frac{R}{2}e$) that

$$1 \leq C_1 \left(\frac{R}{x_1} \right)^\gamma h_{B_R \cap \mathcal{D}}^x [\partial B_R \cap \mathcal{D}]$$

Hence

$$h_{B_R \cap \mathcal{D}}^x [\partial B_R \cap \mathcal{D}] \geq \frac{1}{C_1} \left(\frac{x_1}{R} \right)^\gamma.$$

On the other hand, the same considerations, as in §2.2.3, show that

$$h_{B_R \cap \mathcal{D}}^x [(\partial B_R \cap \mathcal{D}) \cap \{x_i \leq \epsilon R, i = 1, \dots, d\}] \leq C_2 \rho^{\frac{R}{x_1}},$$

where $0 < \rho < 1$. Setting $\rho = e^{-r}$, for an appropriate $r > 0$ and choosing ϵ sufficiently small and satisfying

$$C_2 \epsilon^{-\gamma} e^{-\frac{r}{\epsilon}} < \frac{1}{C_1}$$

gives the inequality (2.8).

2.3.2 The denouement

Let us consider u_1 and u_2 , two positive harmonic functions on \mathcal{D} vanishing on $\partial\mathcal{D}$. By the boundary Harnack principle, we have

$$\frac{1}{C} \frac{u_1(x)}{u_1(Re)} \leq \frac{u_2(x)}{u_2(Re)} \leq C \frac{u_1(x)}{u_1(Re)}, \quad x \in B_{R/2} \cap \mathcal{D}, \quad R \geq C,$$

where $C > 1$ is an appropriate large positive constant. Assume that u_1 and u_2 are normalized so that $u_1(e) = u_2(e) = 1$. It follows then that

$$\frac{1}{C} u_1(Re) \leq u_2(Re) \leq C u_1(Re), \quad R \geq C.$$

Then by letting $R \rightarrow +\infty$ we deduce that :

$$\frac{1}{C^2} \leq \frac{u_2(x)}{u_1(x)} \leq C^2, \quad x \in \mathcal{D}.$$

Define

$$u_3(x) = u_2(x) + \frac{1}{C^2 - 1} (u_2(x) - u_1(x)), \quad x \in \mathcal{D}.$$

We have

$$u_3(x) = \frac{C^2}{C^2 - 1} \left(u_2(x) - \frac{1}{C^2} u_1(x) \right) \geq 0, \quad x \in \mathcal{D}.$$

and clearly u_3 is harmonic function, vanishing on the boundary $\partial\mathcal{D}$ and satisfies the normalization condition $u_3(e) = 1$. We iterate and define (for $p \geq 4$)

$$u_p(x) = u_{p-1}(x) + \frac{1}{C^2 - 1} (u_{p-1}(x) - u_1(x)).$$

An easy computation shows that u_p is a non-negative harmonic function on \mathcal{D} , vanishing on the boundary $\partial\mathcal{D}$, satisfying the normalization condition $u_p(e) = 1$ and such that

$$\frac{1}{C^2} \leq \frac{u_p(x)}{u_1(x)} \leq C^2, \quad x \in \mathcal{D}.$$

On the other hand, combining

$$\begin{aligned} u_p(x) &= \frac{C^2}{C^2 - 1} u_{p-1}(x) - \frac{1}{C^2 - 1} u_1(x) \\ u_1(x) &= \frac{C^2}{C^2 - 1} u_1(x) - \frac{1}{C^2 - 1} u_1(x). \end{aligned}$$

We deduce that

$$u_p(x) - u_1(x) = \frac{C^2}{C^2 - 1} (u_{p-1}(x) - u_1(x)), \quad x \in \mathcal{D}, \quad p \geq 3.$$

From which it follows that

$$u_p(x) = \left(\frac{C^2}{C^2 - 1} \right)^{p-2} (u_2(x) - u_1(x)) + u_1(x).$$

Dividing by $u_1(x)$ we deduce that :

$$\frac{1}{C^2} \leq \left(\frac{C^2}{C^2 - 1} \right)^{p-2} \left(-1 + \frac{u_2(x)}{u_1(x)} \right) + 1 \leq C^2, \quad x \in \mathcal{D}, p \geq 4.$$

Letting $p \rightarrow +\infty$ we deduce that we necessarily have

$$u_1(x) = u_2(x), \quad x \in \mathcal{D}.$$

2.3.3 An example

Here we illustrate our main result by an example of a spatially inhomogeneous random walk in the quarter plane \mathbb{Z}_+^2 .

Example 1 We consider a random walk with alternating transition probabilities defined as follows. On the lines $\mathcal{L}_k = \{(x, x + 2k); x \in \mathbb{Z}\}$ ($k \in \mathbb{Z}$) the transition probabilities are those of the simple random walk, see Figure 2.1,

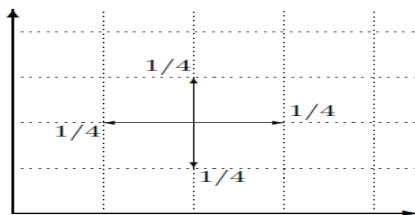


FIGURE 2.1 – The simple random walk in the quarter plane

On the lines $\mathcal{L}'_k = \{(x, x + 2k + 1); x \in \mathbb{Z}\}$ the walk is allowed to visit its eight nearest neighbors with probabilities $p_{i,j} = \frac{1}{8}$, $-1 \leq i, j \leq 1$, see Figure 2.2.

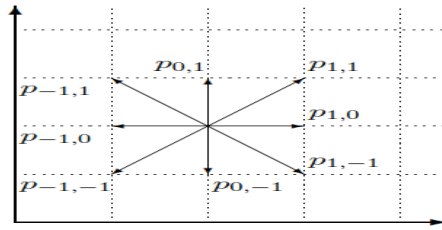


FIGURE 2.2 – Walks with small steps in the quarter plane

The resulting random walk with killing at $\partial\mathbb{Z}_+^2$ is represented on Figure 2.3.

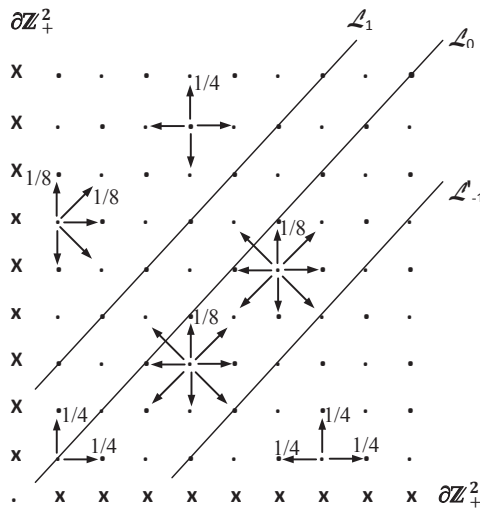


FIGURE 2.3 – The walk has alternating transition probabilities inside the quarter plane and is killed on the boundary of \mathbb{Z}_+^2 which consists of $\{(x, 0); x \in \mathbb{N}^*\} \cup \{(0, y); y \in \mathbb{N}^*\}$.

Easy computation shows that the function

$$u(x, y) = xy \tag{2.9}$$

satisfy, as in the case of the simple random walk, conditions i), ii) and iii) of § 1 with respect to our random walk and is then the unique harmonic function found in §2.

The previous example can be generalized as follows. Let $(S_n)_{n \in \mathbb{N}}$ denote a random walk as in Figure 2, but where the transition probabilities $p_{ij} = p_{ij}(x, y)$ depend on the position (x, y) . Assume that

$$p_{1,0}; \quad p_{-1,0}; \quad p_{0,-1}; \quad p_{-1,-1} \geq \alpha > 0$$

(i.e the ellipticity condition is satisfied) and that $(S_n)_{n \in \mathbb{N}}$ is centered :

$$\sum_{-1 \leq i, j \leq 1} i p_{ij} = \sum_{-1 \leq i, j \leq 1} j p_{ij} = 0. \quad (2.10)$$

Then u as defined by (2.9) is harmonic for the random walk $(S_n)_{n \in \mathbb{N}}$, if it satisfies the finite difference equation

$$\sum_{-1 \leq i, j \leq 1} p_{ij}(x, y) u(x+i, y+j) = u(x, y). \quad (2.11)$$

We observe that the right hand side of (2.11) is given by

$$\begin{aligned} \sum_{-1 \leq i, j \leq 1} p_{ij}(x, y) (x+i)(y+j) &= \left(\sum_{-1 \leq i, j \leq 1} p_{ij}(x, y) \right) xy \\ &+ \left(\sum_{-1 \leq i, j \leq 1} j p_{ij}(x, y) \right) x + \left(\sum_{-1 \leq i, j \leq 1} i p_{ij}(x, y) \right) y \\ &+ \sum_{-1 \leq i, j \leq 1} ij p_{ij}(x, y). \end{aligned}$$

Using (2.10) and the fact that $\sum_{-1 \leq i, j \leq 1} p_{ij} = 1$, we see that equation(2.11) is satisfied if and only if

$$\sum_{-1 \leq i, j \leq 1} ij p_{ij}(x, y) = 0.$$

This condition is obviously satisfied by the random walk represented on Figure 3.

A second example : we now give a second example, more sophisticated than the previous one and whose harmonic function is not of the product xy . We start from the random walk in the weyl chamber of the lie algebra $SL_3(\mathbb{C})$ (c.f [50]) and we reflect this walk with respect to the diagonal $x = y$. The obtained walk is assigned to the even lines \mathcal{L}_k where is allowed to visit its six neighbors with equal probabilities

$$p_{1,0} = p_{-1,0} = p_{0,1} = p_{0,-1} = p_{1,-1} = p_{-1,1} = \frac{1}{6}$$

on the odd lines \mathcal{L}'_k we modify these probabilities as follows

$$p_{1,0} = p_{-1,1} = p_{0,-1} = \frac{2 - \alpha}{6}$$

et

$$p_{0,1} = p_{-1,0} = p_{1,-1} = \frac{\alpha}{6}$$

where $0 < \alpha < 1$ is fixed. Easy computations show that

$$u(x, y) = \frac{xy(x + y)}{2}$$

(The same function as in [50]) is harmonic for inhomogeneous random walk, we observe that we can allow α to take different values on the odd lines as it varies within a range ensuring ellipticity.

Chapitre 3

Local limits of Galton-Watson trees conditioned on the number of protected nodes

This work comes from a collaboration with Romain Abraham and Jean François Delmas and has been the object of an article in "Journal of Applied Probability".

Abstract. We consider a marking procedure of the vertices of a tree where each vertex is marked independently from the others with a probability that depends only on its out-degree. We prove that a critical Galton-Watson tree conditioned on having a large number of marked vertices converges in distribution to the associated size-biased tree. We then apply this result to give the limit in distribution of a critical Galton-Watson tree conditioned on having a large number of protected nodes.

3.1 Introduction

In [35], Kesten proved that a critical or sub-critical Galton-Watson (GW) tree conditioned on reaching at least height h converges in distribution (for the local topology on trees) as h goes to infinity toward the so-called size-biased tree (that we call here Kesten's tree and whose distribution is described in Section 4.4.1). Since then, other conditionings have been considered, see [2, 3, 26] and the references therein for recent developments on the subject.

A protected node is a node that is not a leaf and none of its offsprings is a leaf. Precise asymptotics for the number of protected nodes in a conditioned

GW tree have already been obtained in [13, 30] for instance. Let $A(\mathbf{t})$ be the number of protected nodes in the tree \mathbf{t} . Remark that this functional A is clearly monotone in the sense of [26] (using for instance (3.13)); therefore, using Theorem 2.1 of [26], we immediately get that a critical GW tree τ conditioned on $\{A(\tau) > n\}$ converges in distribution toward Kesten's tree as n goes to infinity. Conditioning on $\{A(\tau) = n\}$ needs extra work and is the main objective of this paper. Using the general result of [2], if we have the following limit

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{P}(A(\tau) = n + 1)}{\mathbb{P}(A(\tau) = n)} = 1, \quad (3.1)$$

then the critical GW tree τ conditioned on $\{A(\tau) = n\}$ converges in distribution also toward Kesten's tree, see Theorem 3.5.1.

In fact, the limit (3.1) can be seen as a special case of a more general problem : conditionally given the tree, we mark the nodes of the tree independently of the rest of the tree with a probability that depends only on the number of offsprings of the nodes. Then we prove that a critical GW tree conditioned on the total number of marked nodes being large converges in distribution toward Kesten's tree, see Theorem 3.3.3.

The chapter is then organized as follows : we first recall briefly the framework of discrete trees, then we consider in Section 3.3 the problem of a marked GW tree and the proofs of the results are given in Section 3.4. In particular, we prove the limit (3.1) when A is the number of marked nodes in Lemma 3.4.2 and we deduce the convergence of a critical GW tree conditioned on the number of marked nodes toward Kesten's tree in Theorem 3.3.3. We finally explain in Section 3.5 how the problem of protected nodes can be viewed as a problem on marked nodes and deduce the convergence in distribution of a critical GW tree conditioned on the number of protected nodes toward Kesten's tree in Theorem 3.5.1.

3.2 Technical background on GW trees

3.2.1 The set of discrete trees

We denote by $\mathbb{N} = \{0, 1, 2, \dots\}$ the set of non-negative integers and by $\mathbb{N}^* = \{1, 2, \dots\}$ the set of positive integers.

If E is a subset of \mathbb{N}^* , we call the span of E the greatest common divisor of E . If X is an integer-valued random variable, we call the span of X the span of $\{n > 0; \mathbb{P}(X = n) > 0\}$.

We recall Neveu's formalism [46] for ordered rooted trees. Let $\mathcal{U} = \bigcup_{n \geq 0} (\mathbb{N}^*)^n$ be the set of finite sequences of positive integers with the convention $(\mathbb{N}^*)^0 = \{\emptyset\}$. For $u \in \mathcal{U}$, its length or generation $|u| \in \mathbb{N}$ is defined by $u \in (\mathbb{N}^*)^{|u|}$. If u and v are two sequences of \mathcal{U} , we denote by uv the conca-

tenation of the two sequences, with the convention that $uv = u$ if $v = \emptyset$ and $uv = v$ if $u = \emptyset$. The set of ancestors of u is the set

$$\text{An}(u) = \{v \in \mathcal{U}; \exists w \in \mathcal{U} \text{ such that } u = vw\}.$$

Notice that u belongs to $\text{An}(u)$. For two distinct elements u and v of \mathcal{U} , we denote by $u < v$ the lexicographic order on \mathcal{U} i.e. $u < v$ if $u \in \text{An}(v)$ and $u \neq v$ or if $u = wi u'$ and $v = wj v'$ for some $i, j \in \mathbb{N}^*$ with $i < j$. We write $u \leq v$ if $u = v$ or $u < v$.

A tree \mathbf{t} is a subset of \mathcal{U} that satisfies :

- $\emptyset \in \mathbf{t}$.
- If $u \in \mathbf{t}$, then $\text{An}(u) \subset \mathbf{t}$.
- For every $u \in \mathbf{t}$, there exists $k_u(\mathbf{t}) \in \mathbb{N}$ such that, for every $i \in \mathbb{N}^*$, $ui \in \mathbf{t}$ iff $1 \leq i \leq k_u(\mathbf{t})$.

The vertex \emptyset is called the root of \mathbf{t} . The integer $k_u(\mathbf{t})$ represents the number of offsprings of the vertex $u \in \mathbf{t}$. The set of children of a vertex $u \in \mathbf{t}$ is given by :

$$C_u(\mathbf{t}) = \{ui; 1 \leq i \leq k_u(\mathbf{t})\}. \quad (3.2)$$

By convention, we set $k_u(\mathbf{t}) = -1$ if $u \notin \mathbf{t}$.

A vertex $u \in \mathbf{t}$ is called a leaf if $k_u(\mathbf{t}) = 0$. We denote by $\mathcal{L}_0(\mathbf{t})$ the set of leaves of \mathbf{t} . A vertex $u \in \mathbf{t}$ is called a *protected node* if $C_u(\mathbf{t}) \neq \emptyset$ and $C_u(\mathbf{t}) \cap \mathcal{L}_0(\mathbf{t}) = \emptyset$, that is u is not a leaf and none of its children is a leaf. For $u \in \mathbf{t}$, we define $F_u(\mathbf{t})$, the fringe subtree of \mathbf{t} above u , as

$$F_u(\mathbf{t}) = \{v \in \mathbf{t}; u \in \text{An}(v)\} = \{uv; v \in S_u(\mathbf{t})\}$$

with $S_u(\mathbf{t}) = \{v \in \mathcal{U}; uv \in \mathbf{t}\}$.

Notice that $S_u(\mathbf{t})$ is a tree. We denote by \mathbb{T} the set of trees and by $\mathbb{T}_0 = \{\mathbf{t} \in \mathbb{T}; \text{Card}(\mathbf{t}) < +\infty\}$ the subset of finite trees.

We say that a sequence of trees $(\mathbf{t}_n, n \in \mathbb{N})$ converges locally to a tree \mathbf{t} if and only if $\lim_{n \rightarrow \infty} k_u(\mathbf{t}_n) = k_u(\mathbf{t})$ for all $u \in \mathcal{U}$. Let $(T_n, n \in \mathbb{N})$ and T be \mathbb{T} -valued random variables. We denote by $\text{dist}(T)$ the distribution of the random variable T and write

$$\lim_{n \rightarrow +\infty} \text{dist}(T_n) = \text{dist}(T)$$

for the convergence in distribution of the sequence $(T_n, n \in \mathbb{N})$ to T with respect to the local topology.

If $\mathbf{t}, \mathbf{t}' \in \mathbb{T}$ and $x \in \mathcal{L}_0(\mathbf{t})$ we denote by

$$\mathbf{t} \otimes_x \mathbf{t}' = \{u \in \mathbf{t}\} \cup \{xv; v \in \mathbf{t}'\} \quad (3.3)$$

the tree obtained by grafting the tree \mathbf{t}' on the leaf x of the tree \mathbf{t} . For every $\mathbf{t} \in \mathbb{T}$ and every $x \in \mathcal{L}_0(\mathbf{t})$, we shall consider the set of trees obtained by grafting a tree on the leaf x of \mathbf{t} :

$$\mathbb{T}(\mathbf{t}, x) = \{\mathbf{t} \otimes_x \mathbf{t}'; \mathbf{t}' \in \mathbb{T}\}.$$

3.2.2 Galton Watson trees

Let $p = (p(n), n \in \mathbb{N})$ be a probability distribution on \mathbb{N} . We assume that

$$p(0) > 0, p(0) + p(1) < 1, \text{ and } \mu := \sum_{n=0}^{+\infty} np(n) < +\infty. \quad (3.4)$$

A \mathbb{T} -valued random variable τ is a GW tree with offspring distribution p if the distribution of $k_\emptyset(\tau)$ is p and it enjoys the branching property : for $n \in \mathbb{N}^*$, conditionally on $\{k_\emptyset(\tau) = n\}$, the subtrees $(S_1(\tau), \dots, S_n(\tau))$ are independent and distributed as the original tree τ .

The GW tree and the offspring distribution are called critical (resp. sub-critical, super-critical) if $\mu = 1$ (resp. $\mu < 1$, $\mu > 1$).

3.3 Conditioning on the number of marked vertices

3.3.1 Definition of the marking procedure

We begin with a fixed tree \mathbf{t} . We add marks on the vertices of \mathbf{t} in an independent way such that the probability of adding a mark on a node u depends only on the number of children of u . More precisely, we consider a mark function $q : \mathbb{N} \rightarrow [0, 1]$ and a family of independent Bernoulli random variables $(Z_u(\mathbf{t}), u \in \mathbf{t})$ such that for all $u \in \mathbf{t}$:

$$\mathbb{P}(Z_u(\mathbf{t}) = 1) = 1 - \mathbb{P}(Z_u(\mathbf{t}) = 0) = q(k_u(\mathbf{t})).$$

The vertex u is said to have a mark if $Z_u(\mathbf{t}) = 1$. We denote by $\mathcal{M}(\mathbf{t}) = \{u \in \mathbf{t}; Z_u(\mathbf{t}) = 1\}$ the set of marked vertices and by $M(\mathbf{t})$ its cardinal. We call $(\mathbf{t}, \mathcal{M}(\mathbf{t}))$ a marked tree.

A marked GW tree with offspring distribution p and mark function q is a couple $(\tau, \mathcal{M}(\tau))$, with τ a GW tree with offspring distribution p and conditionally on $\{\tau = \mathbf{t}\}$ the set of marked vertices $\mathcal{M}(\tau)$ is distributed as $\mathcal{M}(\mathbf{t})$.

Remark 3.3.1. Notice that for $\mathcal{A} \subseteq \mathbb{N}$, if we set $q(k) = \mathbf{1}_{\{k \in \mathcal{A}\}}$, then the set $\mathcal{M}(\mathbf{t})$ is just the set of vertices with out-degree (i.e. number of offsprings) in \mathcal{A} considered in [2, 51]. Hence, the above construction can be seen as an extension of this case.

3.3.2 Kesten's tree

Let p be an offspring distribution satisfying Assumption (3.4) with $\mu \leq 1$ (i.e. the associated GW process is critical or sub-critical). We denote by $p^* = (p^*(n) = np(n)/\mu, n \in \mathbb{N})$ the corresponding size-biased distribution.

We define an infinite random tree τ^* (the size-biased tree that we call Kesten's tree in this paper) whose distribution is described as follows :

There exists a unique infinite sequence $(v_k, k \in \mathbb{N}^*)$ of positive integers such that, for every $h \in \mathbb{N}$, $v_1 \cdots v_h \in \tau^*$, with the convention that $v_1 \cdots v_h = \emptyset$ if $h = 0$. The joint distribution of $(v_k, k \in \mathbb{N}^*)$ and τ^* is determined recursively as follows. For each $h \in \mathbb{N}$, conditionally given (v_1, \dots, v_h) and $\{u \in \tau^*; |u| \leq h\}$ the tree τ^* up to level h , we have :

- The number of children $(k_u(\tau^*), u \in \tau^*, |u| = h)$ are independent and distributed according to p if $u \neq v_1 \cdots v_h$ and according to p^* if $u = v_1 \cdots v_h$.
- Given $\{u \in \tau^*; |u| \leq h + 1\}$ and (v_1, \dots, v_h) , the integer v_{h+1} is uniformly distributed on the set of integers $\{1, \dots, k_{v_1 \cdots v_h}(\tau^*)\}$.

Remark 3.3.2. Notice that by construction, a.s. τ^* has a unique infinite spine. And following Kesten [35], the random tree τ^* can be viewed as the tree τ conditioned on non extinction.

For $\mathbf{t} \in \mathbb{T}_0$ and $x \in \mathcal{L}_0(\mathbf{t})$, we have :

$$\mathbb{P}(\tau^* \in \mathbb{T}(t, x)) = \frac{\mathbb{P}(\tau = t)}{\mu^{|x|} p(0)}.$$

3.3.3 Main theorem

Theorem 3.3.3. Let p be a critical offspring distribution that satisfies Assumption (3.4). Let $(\tau, \mathcal{M}(\tau))$ be a marked GW tree with offspring distribution p and mark function q such that $p(k)q(k) > 0$ for some $k \in \mathbb{N}$. For every $n \in \mathbb{N}^*$, let τ_n be a tree whose distribution is the conditional distribution of τ given $\{M(\tau) = n\}$. Let τ^* be a Kesten's tree associated with p . Then we have :

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau_n) = \text{dist}(\tau^*),$$

where the limit has to be understood along a subsequence for which $\mathbb{P}(M(\tau) = n) > 0$.

Remark 3.3.4. If for every $k \in \mathbb{N}$, $0 < q(k) < 1$, then $\mathbb{P}(M(\tau) = n) > 0$ for every $n \in \mathbb{N}$, hence the above conditioning is always valid.

3.4 Proof of Theorem 3.3.3

Set $\gamma = \mathbb{P}(M(\tau) > 0)$. Since there exists $k \in \mathbb{N}$ with $p(k)q(k) > 0$, we have $\gamma > 0$. A sufficient condition (but not necessary) to have $\mathbb{P}(M(\tau) = n) > 0$ for every n large enough is to assume that $\gamma < 1$ (see Lemma 3.4.3 and Section 3.4.4). Taking $q = \mathbf{1}_{\mathcal{A}}$, see Remark 3.3.1 for $0 \in \mathcal{A} \subset \mathbb{N}$ implies $\gamma = 1$ and some periodicity may occur.

The following result is the analogue in the random case of Theorem 3.1 in [2] and its proof is in fact a straightforward adaptation of the proof in [2] by using :

- (i) $M(\mathbf{t}) \leq \text{Card}(\mathbf{t})$.
- (ii) For every $\mathbf{t} \in \mathbb{T}_0$, $x \in \mathcal{L}_0(\mathbf{t})$ and $\mathbf{t}' \in \mathbb{T}$, we have that $M(\mathbf{t} \otimes_x \mathbf{t}')$ is distributed as $\hat{M}(\mathbf{t}') + M(\mathbf{t}) - \mathbf{1}_{\{Z_x(\mathbf{t})=1\}}$, where $\hat{M}(\mathbf{t}')$ is distributed as $M(\mathbf{t}')$ and is independent of $\mathcal{M}(\mathbf{t})$.

Proposition 3.4.1. *Let $n_0 \in \mathbb{N} \cup \{\infty\}$. Assume that $\mathbb{P}(M(\tau) \in [n, n+n_0]) > 0$ for n large enough. Then, if*

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{P}(M(\tau) \in [n+1, n+1+n_0])}{\mathbb{P}(M(\tau) \in [n, n+n_0])} = 1, \quad (3.5)$$

we have :

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau | M(\tau) \in [n, n+n_0]) = \text{dist}(\tau^*).$$

Démonstration. According to Lemma 2.1 in [2], a sequence $(T_n, n \in \mathbb{N})$ of finite random trees converges in distribution (with respect to the local topology) to some Kesten's tree τ^* if and only if, for every finite tree $\mathbf{t} \in \mathbb{T}_0$ and every leaf $x \in \mathcal{L}_0(\mathbf{t})$,

$$\lim_{n \rightarrow +\infty} \mathbb{P}((T_n \in \mathbb{T}(\mathbf{t}, x)) = \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) \quad \text{and} \quad \lim_{n \rightarrow +\infty} \mathbb{P}(T_n = \mathbf{t}) = 0. \quad (3.6)$$

Let $\mathbf{t} \in \mathbb{T}_0$ and $x \in \mathcal{L}_0(\mathbf{t})$. We set $D(\mathbf{t}, x) = M(\mathbf{t}) - \mathbf{1}_{\{Z_x(\mathbf{t})=1\}}$. Notice that $D(\mathbf{t}, x) \leq \text{Card}(\mathbf{t}) - 1$. Elementary computations give for every $\mathbf{t}' \in \mathbb{T}_0$ that :

$$\mathbb{P}(\tau = \mathbf{t} \otimes_x \mathbf{t}') = \frac{1}{p(0)} \mathbb{P}(\tau = \mathbf{t}) \mathbb{P}(\tau = \mathbf{t}') \quad \text{and} \quad \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) = \frac{1}{p(0)} \mathbb{P}(\tau = \mathbf{t}).$$

As τ is a.s. finite, we have :

$$\begin{aligned} & \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, x), M(\tau) \in [n, n+n_0]) \\ &= \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}(\tau = \mathbf{t} \otimes_x \mathbf{t}', M(\tau) \in [n, n+n_0]) \\ &= \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}(\tau = \mathbf{t} \otimes_x \mathbf{t}') \mathbb{P}(M(\mathbf{t} \otimes_x \mathbf{t}') \in [n, n+n_0]) \\ &= \sum_{\mathbf{t}' \in \mathbb{T}_0} \frac{\mathbb{P}(\tau = \mathbf{t}) \mathbb{P}(\tau = \mathbf{t}')}{p(0)} \mathbb{P}(\hat{M}(\mathbf{t}') + D(\mathbf{t}, x) \in [n, n+n_0]) \\ &= \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)) \mathbb{P}(\hat{M}(\tau) + D(\mathbf{t}, x) \in [n, n+n_0]). \end{aligned}$$

Notice that :

$$\begin{aligned}
& \mathbb{P}(\hat{M}(\tau) + D(\mathbf{t}, x) \in [n, n + n_0]) \\
&= \sum_{k=0}^{\text{Card}(\mathbf{t})-1} \mathbb{P}(\hat{M}(\tau) + D(\mathbf{t}, x) \in [n, n + n_0] \mid D(\mathbf{t}, x) = k) \mathbb{P}(D(\mathbf{t}, x) = k) \\
&= \sum_{k=0}^{\text{Card}(\mathbf{t})-1} \mathbb{P}(M(\tau) \in [n - k, n + n_0 - k]) \mathbb{P}(D(\mathbf{t}, x) = k).
\end{aligned}$$

Then we obtain using Assumption (3.5) that :

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{P}(\hat{M}(\tau) + D(\mathbf{t}, x) \in [n, n + n_0])}{\mathbb{P}(M(\tau) \in [n, n + n_0])} = 1,$$

that is

$$\lim_{n \rightarrow +\infty} \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, x) \mid M(\tau) \in [n, n + n_0]) = \mathbb{P}(\tau^* \in \mathbb{T}(\mathbf{t}, x)).$$

This proves the first limit of (3.6).

The second limit is immediate since, for every $n \geq \text{Card}(\mathbf{t})$,

$$\mathbb{P}(\tau = \mathbf{t} \mid M(\tau) \in [n, n + n_0]) = 0.$$

□

The main ingredient for the proof of Theorem 3.3.3 is then the following lemma.

Lemma 3.4.2. *Let d be the span of the random variable $M(\tau) - 1$. We have*

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{P}(M(\tau) \in [n + 1, n + 1 + d])}{\mathbb{P}(M(\tau) \in [n, n + d])} = 1. \quad (3.7)$$

The end of this section is devoted to the proof of Lemma 3.4.2, see Section 3.4.4, which follows the ideas of the proof of Theorem 5.1 of [2].

3.4.1 Transformation of a subset of a tree onto a tree

We recall Rizzolo's map [51] which from $\mathbf{t} \in \mathbb{T}_0$ and a non-empty subset A of \mathbf{t} builds a tree \mathbf{t}_A such that $\text{Card}(A) = \text{Card}(\mathbf{t}_A)$. We will give a recursive construction of this map $\phi : (\mathbf{t}, A) \mapsto \mathbf{t}_A = \phi(\mathbf{t}, A)$. We will check in the next section that this map is such that if τ is a GW tree then τ_A will also be a GW tree for a well chosen subset A of τ . Figure 1 below shows an example of a tree \mathbf{t} , a set A and the associated tree \mathbf{t}_A which helps to understand the construction.

For a vertex $u \in \mathbf{t}$, recall $C_u(\mathbf{t})$ is the set of children of u in \mathbf{t} . We define for $u \in \mathbf{t}$:

$$R_u(\mathbf{t}) = \bigcup_{w \in \text{An}(u)} \{v \in C_w(\mathbf{t}); u < v\}$$

the vertices of \mathbf{t} which are larger than u for the lexicographic order and are children of u or of one of its ancestors. For a vertex $u \in \mathbf{t}$, we shall consider A_u the set of elements of A in the fringe subtree above u :

$$A_u = A \cap F_u(\mathbf{t}) = A \cap \{uv; v \in S_u(\mathbf{t})\}. \tag{3.8}$$

Let $\mathbf{t} \in \mathbb{T}_0$ and $A \subset \mathbf{t}$ such that $A \neq \emptyset$. We shall define $\mathbf{t}_A = \phi(\mathbf{t}, A)$ recursively. Let u_0 be the smallest (for the lexicographic order) element of A . Consider the fringe subtrees of \mathbf{t} that are rooted at the vertices in $R_{u_0}(\mathbf{t})$ and contain at least one vertex in A , that is $(F_u(\mathbf{t}); u \in R_{u_0}^A(\mathbf{t}))$, with

$$R_{u_0}^A(\mathbf{t}) = \{u \in R_{u_0}(\mathbf{t}); A_u \neq \emptyset\} = \{u \in R_{u_0}(\mathbf{t}); \exists v \in A \text{ such that } u \in \text{An}(v)\}.$$

Define the number of children of the root of tree \mathbf{t}_A as the number of those fringe subtrees :

$$k_\emptyset(\mathbf{t}_A) = \text{Card}(R_{u_0}^A(\mathbf{t})).$$

If $k_\emptyset(\mathbf{t}_A) = 0$ set $\mathbf{t}_A = \{\emptyset\}$. Otherwise let $u_1 < \dots < u_{k_\emptyset(\mathbf{t}_A)}$ be the ordered elements of $R_{u_0}^A(\mathbf{t})$ with respect to the lexicographic order on \mathcal{U} . And we define $\mathbf{t}_A = \phi(\mathbf{t}, A)$ recursively by :

$$F_i(\mathbf{t}_A) = \phi(F_{u_i}(\mathbf{t}), A_{u_i}) \quad \text{for } 1 \leq i \leq k_\emptyset(\mathbf{t}_A). \tag{3.9}$$

Since $\text{Card}(A_{u_i}) < \text{Card}(A)$, we deduce $\mathbf{t}_A = \phi(\mathbf{t}, A)$ is well defined and it is a tree by construction. Furthermore, we clearly have that A and \mathbf{t}_A have the same cardinal :

$$\text{Card}(\mathbf{t}_A) = \text{Card}(A). \tag{3.10}$$

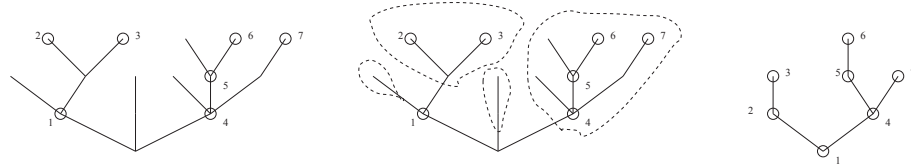


FIGURE 3.1 – Left : A tree \mathbf{t} and the set A . Center : The fringe subtrees rooted at the vertices in $R_{u_0}(\mathbf{t})$. Right : the tree \mathbf{t}_A . The labels have no signification, they only show which node of \mathbf{t} corresponds to a node of \mathbf{t}_A

3.4.2 Distribution of the number of marked nodes

Let $(\tau, \mathcal{M}(\tau))$ be a marked GW tree with critical offspring distribution p satisfying (3.4) and mark function q . Recall $\gamma = \mathbb{P}(M(\tau) > 0) = \mathbb{P}(\mathcal{M}(\tau) \neq \emptyset)$.

Let $((X_i, Z_i), i \in \mathbb{N}^*)$ be i.i.d. random variables such that X_i is distributed according to p and Z_i is conditionally on X_i Bernoulli with parameter $q(X_i)$. We define :

- $G = \inf\{k \in \mathbb{N}^*; \sum_{i=1}^k (X_i - 1) = -1\}$.
- $N = \inf\{k \in \mathbb{N}^*; Z_k = 1\}$.
- \tilde{X} a random variable distributed as $1 + \sum_{i=1}^N (X_i - 1)$ conditionally on $\{N \leq G\}$.
- Y a random variable which is conditionally on \tilde{X} binomial with parameter (\tilde{X}, γ) .

We say that a probability distribution on \mathbb{N} is aperiodic if the span of its support restricted to \mathbb{N}^* is 1. The following result is immediate as the distribution p of X_1 satisfies (3.4).

Lemma 3.4.3. *The distribution of Y satisfies (3.4) and if $\gamma < 1$ then it is aperiodic.*

Recall that for a tree $\mathbf{t} \in \mathbb{T}_0$, we have :

$$\sum_{u \in \mathbf{t}} (k_u(\mathbf{t}) - 1) = -1 \quad (3.11)$$

and $\sum_{u \in \mathbf{t}, u < v} (k_u(\mathbf{t}) - 1) > -1$ for any $v \in \mathbf{t}$. We deduce that G is distributed according to $\text{Card}(\tau)$ and thus N is distributed like the index of the first marked vertex along the depth-first walk of τ . Then, we have :

$$\gamma = \mathbb{P}(N \leq G). \quad (3.12)$$

We denote by $(\tau^0, \mathcal{M}(\tau^0))$ a random marked tree distributed as $(\tau, \mathcal{M}(\tau))$ conditioned on $\{\mathcal{M}(\tau) \neq \emptyset\}$. By construction, $\text{Card}(\tau^0)$ is distributed as G conditioned on $\{N \leq G\}$.

Lemma 3.4.4. *Under the hypothesis of this section, we have that $\tau_{\mathcal{M}(\tau^0)}^0 = \phi(\tau^0, \mathcal{M}(\tau^0))$ is a critical GW tree with the law of Y as offspring distribution.*

3.4.3 Proof of Lemma 3.4.4

In order to simplify notation, we write $\tilde{\tau}$ for $\tau_{\mathcal{M}(\tau^0)}^0 = \phi(\tau^0, \mathcal{M}(\tau^0))$ and for $u \in \tau^0$, we set R_u for $R_u(\tau^0)$.

Lemma 3.4.5. *The random tree $\tilde{\tau}$ is a GW tree with offspring distribution the law of Y .*

Démonstration. Let u_0 be the smallest (for the lexicographic order) element of $\mathcal{M}(\tau^0)$. The branching property of GW trees implies that, conditionally given u_0 and R_{u_0} , the fringe subtrees of τ^0 rooted at the vertices in R_{u_0} , $(S_u(\tau^0), u \in R_{u_0})$ are independent and distributed as τ . Recall notation (3.8) so that the set of marked vertices of the fringe subtree rooted at u is $\mathcal{M}_u(\tau^0) = \mathcal{M}(\tau^0) \cap F_u(\tau^0)$. Define $\tilde{\mathcal{M}}_u(\tau^0) = \{v; uv \in \mathcal{M}_u(\tau^0)\}$ the corresponding marked vertices of $S_u(\mathbf{t})$. Then, the construction of the marks $\mathcal{M}(\tau)$ implies that the corresponding marked trees $((S_u(\tau^0), \tilde{\mathcal{M}}_u(\tau^0)), u \in R_{u_0})$ are independent and distributed as $(\tau, \mathcal{M}(\tau))$. Notice that for $u \in R_{u_0}$, the fringe subtree $F_u(\tau^0)$ contains at least one mark iff u belongs to

$$R_{u_0}^{\mathcal{M}(\tau^0)} = \{u \in R_{u_0}; \exists v \in \mathcal{M}(\tau^0) \text{ such that } u \in \text{An}(v)\}.$$

Then by considering only the fringe subtrees containing at least one mark, we get that, conditionally on $R_{u_0}^{\mathcal{M}(\tau^0)}$, the subtrees $((S_u(\tau^0), \tilde{\mathcal{M}}_u(\tau^0)), u \in R_{u_0}^{\mathcal{M}(\tau^0)})$ are independent and distributed as $(\tau^0, \mathcal{M}(\tau^0))$. We deduce from the recursive construction of the map ϕ , see (3.9), that $\tilde{\tau}$ is a GW tree. Notice that the offspring distribution of $\tilde{\tau}$ is given by the distribution of the cardinal of $R_{u_0}^{\mathcal{M}(\tau^0)}$. We now compute the corresponding offspring distribution. We first give an elementary formula for the cardinal of $R_u(\mathbf{t})$. Let $\mathbf{t} \in \mathbb{T}_0$ and $u \in \mathbf{t}$. Consider the tree $\mathbf{t}' = R_u(\mathbf{t}) \cup \{v \in \mathbf{t}; v \leq u\}$. Using (3.11) for \mathbf{t}' , we get :

$$-1 = \sum_{v \in \mathbf{t}'} (k_v(\mathbf{t}') - 1) = \sum_{v \in \mathbf{t}; v \leq u} (k_v(\mathbf{t}') - 1) + \sum_{v \in R_u(\mathbf{t})} (-1).$$

This gives $\text{Card}(R_u(\mathbf{t})) = 1 + \sum_{v \in \mathbf{t}; v \leq u} (k_v(\mathbf{t}') - 1)$. We deduce from the definition of \tilde{X} that $\text{Card}(R_{u_0})$ is distributed as \tilde{X} . We deduce from the first part of the proof that conditionally on $\text{Card}(R_{u_0})$, the distribution of $\text{Card}(R_{u_0}^{\mathcal{M}(\tau^0)})$ is binomial with parameter $(\text{Card}(R_{u_0}(\tau^0)), \gamma)$. This gives that the offspring distribution of $\tilde{\tau}$ is given by the law of Y . \square

Lemma 3.4.6. *The GW tree $\tilde{\tau}$ is critical.*

Démonstration. Since the offspring distribution is the law of Y we need to check that $\mathbb{E}[Y] = 1$ that is $\gamma \mathbb{E}[\tilde{X}] = 1$ since Y is conditionally on \tilde{X} binomial with parameter (\tilde{X}, γ) .

Recall N has finite expectation as $\mathbb{P}(Z_1 = 1) > 0$, is not independent of $(X_i)_{i \in \mathbb{N}^*}$ and is a stopping time with respect to the filtration generated by $((X_i, Z_i), i \in \mathbb{N}^*)$. Using Wald's equality and $\mathbb{E}[X_i] = 1$, we get $\mathbb{E} \left[\sum_{i=1}^N (X_i - 1) \right] = 0$ and thus using the definition of \tilde{X} as well as (3.12) :

$$\gamma \mathbb{E}[\tilde{X}] = \gamma + \mathbb{E} \left[\sum_{i=1}^N (X_i - 1) \mathbf{1}_{\{N \leq G\}} \right] = \gamma - \mathbb{E} \left[\sum_{i=1}^N (X_i - 1) \mathbf{1}_{\{N > G\}} \right].$$

We have :

$$\begin{aligned} \mathbb{E} \left[\sum_{i=1}^N (X_i - 1) \mathbf{1}_{\{N > G\}} \right] &= \mathbb{E} \left[\sum_{i=1}^G (X_i - 1) \mathbf{1}_{\{N > G\}} \right] + \mathbb{P}(N > G) \mathbb{E} \left[\sum_{i=1}^N (X_i - 1) \right] \\ &= -\mathbb{P}(N > G) \\ &= \gamma - 1, \end{aligned}$$

where we used the strong Markov property of $((X_i, Z_i), i \in \mathbb{N}^*)$ at the stopping time G for the first equation, the definition of T and Wald's equality for the second, and (3.12) for the third. We deduce that $\mathbb{E}[Y] = \gamma \mathbb{E}[\tilde{X}] = 1$, which ends the proof. \square

3.4.4 Proof of (3.7)

According to Lemma 3.4.4 and (3.10), we have that $M(\tau^0)$ is distributed as the total size of a critical GW whose offspring distribution satisfies (3.4). The proof of Proposition 4.3 of [2] (see Equation (4.15) in [2]) entails that if τ' is a critical GW tree, then, if d denotes the span of the random variable $\text{Card}(\tau') - 1$, we have

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}(\text{Card}(\tau') \in [n+1, n+1+d])}{\mathbb{P}(\text{Card}(\tau') \in [n, n+d])} = 1.$$

3.5 Protected nodes

Recall that a node of a tree \mathbf{t} is protected if it is not a leaf and none of its offsprings is a leaf. We denote by $A(\mathbf{t})$ the number of protected nodes of the tree \mathbf{t} .

Theorem 3.5.1. *Let τ be a critical GW tree with offspring distribution p satisfying (3.4) and let τ^* be the associated Kesten's tree. Let τ_n be a random tree distributed as τ conditionally given $\{A(\tau) = n\}$. Then :*

$$\lim_{n \rightarrow +\infty} \text{dist}(\tau_n) = \text{dist}(\tau^*).$$

Démonstration. Notice that $\mathbb{P}(A(\tau) = n) > 0$ for all $n \in \mathbb{N}$. Notice that the functional A satisfies the additive property of [2], namely for every $\mathbf{t} \in \mathbb{T}$, every $x \in \mathcal{L}_0(\mathbf{t})$ and every $\mathbf{t}' \in \mathbb{T}$ that is not reduced to the root, we have

$$A(\mathbf{t} \otimes_x \mathbf{t}') = A(\mathbf{t}) + A(\mathbf{t}') + D(\mathbf{t}, x) \quad (3.13)$$

where $D(\mathbf{t}, x) = 1$ if x is the only child of its first ancestor which is a leaf (therefore this ancestor becomes a protected node in $\mathbf{t} \otimes_x \mathbf{t}'$) and $D(\mathbf{t}, x) = 0$ otherwise. According to Theorem 3.1 of [2], to end the proof it is enough to check that

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{P}(A(\tau) = n+1)}{\mathbb{P}(A(\tau) = n)} = 1. \quad (3.14)$$

For a tree $\mathbf{t} \neq \{\emptyset\}$, let $\mathbf{t}_{\mathbb{N}^*} = \phi(\mathbf{t}, \mathbf{t} \setminus \mathcal{L}_0(\mathbf{t}))$ be the tree obtained from \mathbf{t} by removing the leaves. Let τ^0 be a random tree distributed as τ conditioned to $\{k_\emptyset(\tau) > 0\}$. Using Theorem 6 and Corollary 2 of [51] with $A = \mathbb{N}^*$ (or Lemma 3.4.4 with $q(k) = \mathbf{1}_{\{k > 0\}}$), we have that $\tau_{\mathbb{N}^*}^0$ is a critical GW tree with offspring distribution :

$$p_{\mathbb{N}^*}(k) = \sum_{n=\max(k,1)}^{+\infty} p(n) \binom{n}{k} (p(0))^{n-k} (1-p(0))^{k-1}, \quad k \in \mathbb{N}.$$

Conditionally given $\{\tau_{\mathbb{N}^*}^0 = \mathbf{t}\}$, we consider independent random variables $(W(u), u \in \mathbf{t})$ taking values in \mathbb{N}^* whose distributions are given for all $u \in \mathbf{t}$ by $\mathbb{P}(W(u) = 0) = 0$ for $k_u(\mathbf{t}) = 0$ and otherwise for $k_u(\mathbf{t}) + n > 0$ (remark that $p_{\mathbb{N}^*}(k_u(\mathbf{t})) > 0$), by

$$\mathbb{P}(W(u) = n) = \frac{p(k_u(\mathbf{t}) + n)}{p_{\mathbb{N}^*}(k_u(\mathbf{t}))} \binom{k_u(\mathbf{t}) + n}{n} p(0)^n (1-p(0))^{k_u(\mathbf{t})-1}.$$

In particular for $k_u(\mathbf{t}) > 0$, we have :

$$\mathbb{P}(W(u) = 0) = \frac{p(k_u(\mathbf{t}))}{p_{\mathbb{N}^*}(k_u(\mathbf{t}))} (1-p(0))^{k_u(\mathbf{t})-1}. \tag{3.15}$$

Then, we define a new tree $\hat{\tau}$ by grafting, on every vertex u of $\tau_{\mathbb{N}^*}^0$, $W(u)$ leaves in a uniform manner, see Figure 3.2.

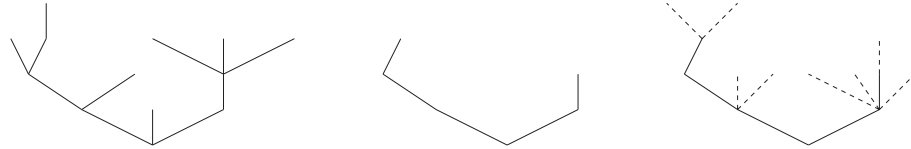


FIGURE 3.2 – The trees τ^0 , $\tau_{\mathbb{N}^*}^0$ and $\hat{\tau}$

More precisely, given $\tau_{\mathbb{N}^*}^0$ and $(W(u), u \in \tau_{\mathbb{N}^*}^0)$, we define a tree $\hat{\tau}$ and a random map $\psi : \tau_{\mathbb{N}^*}^0 \mapsto \hat{\tau}$ recursively in the following way. We set $\psi(\emptyset) = \emptyset$. Then, given $k_\emptyset(\tau_{\mathbb{N}^*}^0) = k$, we set $k_\emptyset(\hat{\tau}) = k + W(\emptyset)$. We also consider a family (i_1, \dots, i_k) of integer-valued random variables such that $(i_1, i_2 - i_1, \dots, i_k - i_{k-1}, W(u) + k + 1 - i_k)$ is a uniform positive partition of $W(u) + k + 1$. Then, for every $j \leq k$ such that $j \notin \{i_1, \dots, i_k\}$, we set $k_j(\hat{\tau}) = 0$ i.e. these are leaves of $\hat{\tau}$. For every $1 \leq j \leq k$, we set $\psi(j) = i_j$ and we apply to them the same construction as for the root and so on.

Lemma 3.5.2. *The new tree $\hat{\tau}$ is distributed as the original tree τ^0 .*

Démonstration. Let $\mathbf{t} \in \mathbb{T}_0$. As $\mathbb{P}(\hat{\tau} = \{\emptyset\}) = 0$, we assume that $k_\emptyset(\mathbf{t}) > 0$. Let $\mathbf{t}_{\mathbb{N}^*}$ be the tree obtained from \mathbf{t} by removing the leaves. Using (3.11),

we have :

$$\begin{aligned}\mathbb{P}(\hat{\tau} = \mathbf{t}) &= \prod_{u \in t_{\mathbb{N}^*}} p_{\mathbb{N}^*}(k_u(\mathbf{t}_{\mathbb{N}^*})) \mathbb{P}(W(u) = k_u(\mathbf{t}) - k_u(\mathbf{t}_{\mathbb{N}^*})) \frac{1}{\binom{k_u(\mathbf{t})}{k_u(\mathbf{t}) - k_u(\mathbf{t}_{\mathbb{N}^*})}} \\ &= \frac{\mathbb{P}(\tau = t)}{1 - p(0)} \\ &= \mathbb{P}(\tau^0 = t).\end{aligned}$$

□

Notice that the protected nodes of $\hat{\tau}$ are exactly the nodes of $\tau_{\mathbb{N}^*}^0$ on which we did not add leaves i.e. for which $W(u) = 0$. If we set $\mathcal{M}(\tau_{\mathbb{N}^*}^0) = \{u \in \tau_{\mathbb{N}^*}^0, W(u) = 0\}$, we have $M(\tau_{\mathbb{N}^*}^0) = A(\hat{\tau})$.

Using (3.15), we get that the corresponding mark function q is given by :

$$q(k) = \frac{p(k)(1 - p(0))^{k-1}}{p_{\mathbb{N}^*}(k)} \mathbf{1}_{\{k \geq 1\}}.$$

As $\hat{\tau}$ is distributed as τ^0 , we have :

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{P}(A(\tau^0) = n + 1)}{\mathbb{P}(A(\tau^0) = n)} = \lim_{n \rightarrow +\infty} \frac{\mathbb{P}(A(\hat{\tau}) = n + 1)}{\mathbb{P}(A(\hat{\tau}) = n)} = \lim_{n \rightarrow +\infty} \frac{\mathbb{P}(M(\tau_{\mathbb{N}^*}^0) = n + 1)}{\mathbb{P}(M(\tau_{\mathbb{N}^*}^0) = n)}.$$

As $\tau_{\mathbb{N}^*}^0$ is a critical GW tree, we deduce from Lemma 3.4.2 that

$$\lim_{n \rightarrow +\infty} \frac{\mathbb{P}(M(\tau_{\mathbb{N}^*}^0) = n + 1)}{\mathbb{P}(M(\tau_{\mathbb{N}^*}^0) = n)} = 1.$$

As $\mathbb{P}(A(\tau) = n) = \mathbb{P}(A(\tau) = n | k_\emptyset(\tau) > 0) \mathbb{P}(k_\emptyset(\tau) > 0)$ and $\mathbb{P}(A(\tau) = n | k_\emptyset(\tau) > 0) = \mathbb{P}(A(\tau^0) = n)$ for $n \geq 2$, we obtain (3.14) and hence end the proof. □

Chapitre 4

Very fat geometric Galton-Watson trees

This work comes from a collaboration with Romain Abraham and Jean François Delmas and has been the object of an article that will be shortly submitted.

Abstract. Let τ_n be a random tree distributed as a Galton-Watson tree with geometric offspring distribution conditioned on $\{Z_n = a_n\}$ where Z_n is the size of the n -th generation and (a_n) is a deterministic increasing sequence. We study the local limit of these trees τ_n as $n \rightarrow +\infty$ and observe three distinct regime : if (a_n) increases slowly, the limit is the sized-biased tree which consists in an infinite spine decorated with finite trees, in an intermediate regimes, the limiting tree is composed of an infinite skeleton (that does not satisfy the branching property) still decorated with finite trees and, if the sequence (a_n) increases rapidly, a condensation phenomenon appears and the root of the limiting tree has an infinite number of offspring.

Marches aléatoires et arbres aléatoires de Galton
Watson

Aymen Bouaziz

4.1 Introduction

A Galton-Watson (GW for short) process $(Z_n, n \geq 0)$ describes the size of an evolving population where, at each generation, every extant individual reproduces according to the same offspring distribution p independently of the rest of the population. The associated genealogical tree τ is called a Galton-Watson tree. A classical result is, if we denote by μ the mean number of offspring per individual, then if $\mu < 1$ (sub-critical case) or $\mu = 1$ (critical case), then the population becomes a.s. extinct (i.e. $Z_n = 0$ for some $n \geq 0$ a.s.) whereas if $\mu > 1$ (super-critical case), the population has a positive probability for living forever.

Another now classical result is Kesten's work [35] in which the limit in distribution of a critical or subcritical GW tree conditioned on $\{Z_n > 0\}$ is studied. The limiting tree is the so-called sized-biased tree which can also be viewed as a two-type GW tree and which therefore can be seen as a critical or sub-critical GW tree conditioned on non-extinction.

There are other ways of conditioning the tree of being large : conditioning on having a large total population size, or a large number of leaves... In the critical case, all these conditionings lead to the same limit (see [2]). In the sub-critical case, a condensation phenomenon (i.e. a vertex with an infinite number of offspring at the limit) may happen (see [3]), but even there, there can be only two different limiting trees, a size-biased GW tree or a condensation tree.

In order to have different limits, an idea is to condition the tree to be even bigger, i.e. to consider conditionings of the form $\{Z_n = a_n\}$ for some deterministic sequence $a_n \rightarrow +\infty$.

Some results on branching processes conditioned on their limit behaviour already appeared in previous works, see for instance [47] where the distributions of the conditioned Yule process (which corresponds to a super-critical branching process) or a critical binary branching are described via an infinitesimal generator and a martingale problem.

The first study of local limits for GW trees with such a conditioning appears in [2] where it is proven that, if p is a critical offspring distribution with finite variance, then the tree conditioned on $\{Z_n = a_n\}$ converges in distribution to the associated sized-biased tree if and only if $\lim a_n n^{-2} = 0$.

The goal of this paper is to study what happens beyond that condition and in the sub-critical and super-critical cases. We give a complete description of all the cases when the offspring distribution is a geometric distribution with a Dirac mass at 0 (in that case, the distribution of Z_n is explicit).

For this geometric distribution, we observe three regimes according to (in the critical case) the limit of $\mathbb{P}(Z_{n-1} = a_n)/\mathbb{P}(Z_n = a_n)$ is equal to 1 (Kesten regime), belongs to $(0, 1)$ (Poisson regime), and is equal to 0 (condensation regime).

- In the Kesten regime, the limiting tree is the so called Kesten tree, which is a two-type GW tree, with an infinite spine corresponding to the individuals having an infinite progeny (called the survivor type), on which are grafted independent GW trees distributed as τ conditioned to be finite (notice this conditioning is useless in the critical or sub-critical cases as τ is a.s. finite) corresponding to individuals having a finite progeny (called extinction type).
- In the Poisson regime, the limiting tree is no more a GW tree, but it still has two types, with a backbone without leaves corresponding to individuals having an infinite progeny (also called the survivor type), on which are grafted independent GW trees distributed as τ conditioned to be finite corresponding to individuals having a finite progeny (also called extinction type). However, the backbone can not be seen as a GW tree, as it lacks the branching property. This is more like a random tree with a Poissonian immigration at each generation, with all the configuration having the same probability. (The Kesten tree appears as the limit of the Poisson regime, when the Poissonian immigration goes down to zero.)
- In the condensation regime, the limiting tree is again a two-type GW tree, with a backbone without leaves corresponding to individuals having an infinite progeny (also called the survivor type), on which are grafted independent GW trees distributed as τ (notice there is no more conditioning on τ to be finite, and the corresponding individuals are called normal type). The backbone can be seen as an inhomogeneous GW tree with the root having an infinite number of children (condensation regime), and super-critical offspring distribution at level $h > 0$ with finite mean μ_h which decreases to 1 as h goes to infinity.

We summarize all these cases in the following table :

	Kesten regime	Poisson regi
Sub-critical case	$a_n \ll \mu^{-n}$	$a_n \sim \theta \mu^{-n}$
Critical case	$a_n \ll n^2$	$a_n \sim \theta n^2$
Super-critical case	$a_n \ll \mu^n$	$a_n \sim \theta \mu^n$

Remark 4.1.1. *In the super-critical case, the distribution of τ conditionally on non-extinction that is on the event $\mathcal{E}^c = \{H(\tau) = +\infty\}$ is given in Corollary 1.2.7 of [4]. Even if on \mathcal{E}^c a.s. $\lim_{n \rightarrow \infty} \mu^{-n} Z_n$ exists and belongs to $(0, +\infty)$, we notice that conditioning on the non extinction event \mathcal{E}^c is not equivalent to condition on $\{\mu^{-n} Z_n = \theta\}$ for some $\theta \in (0, +\infty)$ and let n goes to infinity.*

4.2 Notations

We denote by $\mathbb{N} = \{0, 1, 2, \dots\}$ the set of non-negative integers, by $\mathbb{N}^* = \{1, 2, \dots\}$ the set of positive integers and $\bar{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$. For any finite set

E , we denote by $\sharp E$ its cardinal.

4.2.1 The set of discrete trees

We recall Neveu's formalism [46] for ordered rooted trees. Let $\mathcal{U} = \bigcup_{n \geq 0} (\mathbb{N}^*)^n$ be the set of finite sequences of positive integers with the convention $(\mathbb{N}^*)^0 = \{\emptyset\}$. We also set $\mathcal{U}^* = \bigcup_{n \geq 1} (\mathbb{N}^*)^n = \mathcal{U} \setminus \{\emptyset\}$.

For $u \in \mathcal{U}$, let $|u|$ be the length or the generation of u defined as the integer n such that $u \in (\mathbb{N}^*)^n$. If u and v are two sequences of \mathcal{U} , we denote by uv the concatenation of two sequences, with the convention that $u\emptyset = u$ if $v = \emptyset$ and $u\emptyset = \emptyset$ if $u = \emptyset$.

The set of strict ancestors of $u \in \mathcal{U}^*$ is defined by :

$$\text{An}(u) = \{v \in \mathcal{U}, \exists w \in \mathcal{U}^*, u = vw\},$$

and for $\mathcal{S} \subset \mathcal{U}^*$, being non-empty, we set $\text{An}(\mathcal{S}) = \bigcup_{u \in \mathcal{S}} \text{An}(u)$.

A tree \mathbf{t} is a subset of \mathcal{U} that satisfies :

- $\emptyset \in \mathbf{t}$.
- If $u \in \mathbf{t}$, then $\text{An}(u) \subset \mathbf{t}$.
- For every $u \in \mathbf{t}$, there exists $k_u(\mathbf{t}) \in \bar{\mathbb{N}}$ such that, for every positive integer i , $ui \in \mathbf{t} \iff 1 \leq i \leq k_u(\mathbf{t})$.

We denote by \mathbb{T}_∞ the set of trees. Let $\mathbf{t} \in \mathbb{T}_\infty$ be a tree. The vertex \emptyset is called the root of the tree \mathbf{t} and we denote by $\mathbf{t}^* = \mathbf{t} \setminus \{\emptyset\}$ the tree without its root. For a vertex $u \in \mathbf{t}$, the integer $k_u(\mathbf{t})$ represents the number of offspring (also called the out-degree) of the vertex $u \in \mathbf{t}$. By convention, we shall write $k_u(\mathbf{t}) = -1$ if $u \notin \mathbf{t}$. The height $H(\mathbf{t})$ of the tree \mathbf{t} is defined by :

$$H(\mathbf{t}) = \sup\{|u|, u \in \mathbf{t}\} \in \bar{\mathbb{N}}.$$

For $n \in \mathbb{N}$, the size of the n -th generation of \mathbf{t} is defined by :

$$z_n(\mathbf{t}) = \sharp\{u \in \mathbf{t}, |u| = n\}.$$

We denote by \mathbb{T}_f^* the subset of trees with finite out-degrees except the root's :

$$\mathbb{T}_f^* = \{\mathbf{t} \in \mathbb{T}_\infty; \forall u \in \mathbf{t}^*, k_u(\mathbf{t}) < +\infty\}$$

and by $\mathbb{T}_f = \{\mathbf{t} \in \mathbb{T}_f^*; k_\emptyset(\mathbf{t}) < +\infty\}$ the subset of trees with finite out-degrees.

Let $h, k \in \mathbb{N}^*$. We define $\mathbb{T}^{(h)}$ the subset of finite trees with height h :

$$\mathbb{T}_f^{(h)} = \{\mathbf{t} \in \mathbb{T}_f; H(\mathbf{t}) = h\}$$

and $\mathbb{T}_k^{(h)} = \{\mathbf{t} \in \mathbb{T}_f^{(h)}; k_\emptyset(\mathbf{t}) = k\}$ the subset of finite trees with height equal to h and out-degree of the root equal to k . We also define the restriction operators r_h and $r_{h,k}$ defined by, for every $\mathbf{t} \in \mathbb{T}_\infty$:

$$r_h(\mathbf{t}) = \{u \in \mathbf{t}; |u| \leq h\} \quad \text{and} \quad r_{h,k}(\mathbf{t}) = \{\emptyset\} \cup \{u \in r_h(\mathbf{t})^*; \text{An}(u) \cap \{1, \dots, k\} \neq \emptyset\},$$

so that, for $\mathbf{t} \in \mathbb{T}_f$, if $H(\mathbf{t}) \geq h$ then $r_h(\mathbf{t}) \in \mathbb{T}_f^{(h)}$ and if furthermore $k_\emptyset(\mathbf{t}) \geq k$ then $r_{h,k}(\mathbf{t}) \in \mathbb{T}_k^{(h)}$.

4.2.2 Convergence of trees

Set $\mathbb{N}_1 = \{-1\} \cup \bar{\mathbb{N}}$, endowed with the usual topology of the one-point compactification of the discrete space $\{-1\} \cup \mathbb{N}$. For a tree $\mathbf{t} \in \mathbb{T}$, recall that by convention the out-degree $k_u(\mathbf{t})$ of u is set to -1 if u does not belong to \mathbf{t} . Thus a tree $\mathbf{t} \in \mathbb{T}_\infty$ is uniquely determined by the sequence $(k_u(\mathbf{t}), u \in \mathcal{U})$ and then \mathbb{T}_∞ is a subset of $\mathbb{N}_1^{\mathcal{U}}$. By Tychonoff theorem, the set $\mathbb{N}_1^{\mathcal{U}}$ endowed with the product topology is compact. Since \mathbb{T}_∞ is closed it is thus compact. In fact, the set \mathbb{T}_∞ is a Polish space (but we don't need any precise metric at this point). The convergence of sequences of trees is then characterized as follows. Let $(\mathbf{t}_n, n \in \mathbb{N})$ and \mathbf{t} be trees in \mathbb{T}_∞ . We say that $\lim_{n \rightarrow \infty} \mathbf{t}_n = \mathbf{t}$ if and only if $\lim_{n \rightarrow \infty} k_u(\mathbf{t}_n) = k_u(\mathbf{t})$ for all $u \in \mathcal{U}$. It is easy to see that :

- If $(\mathbf{t}_n, n \in \mathbb{N})$ and \mathbf{t} are trees in \mathbb{T}_f , then we have $\lim_{n \rightarrow \infty} \mathbf{t}_n = \mathbf{t}$ if and only if $\lim_{n \rightarrow \infty} r_h(\mathbf{t}_n) = r_h(\mathbf{t})$ for all $h \in \mathbb{N}^*$.
- If $(\mathbf{t}_n, n \in \mathbb{N})$ and \mathbf{t} are trees in \mathbb{T}_f^* , then we have $\lim_{n \rightarrow \infty} \mathbf{t}_n = \mathbf{t}$ if and only if $\lim_{n \rightarrow \infty} r_{h,k}(\mathbf{t}_n) = r_{h,k}(\mathbf{t})$ for all $h, k \in \mathbb{N}^*$.

Let T be a \mathbb{T}_f -valued (resp. \mathbb{T}_f^* -valued) random variable. It is easy to get that if a.s. $H(T) = +\infty$ (resp. a.s. $H(T) = +\infty$ and $k_\emptyset(T) = +\infty$), then the distribution of T is characterized by $(\mathbb{P}(r_h(T) = \mathbf{t}); h \in \mathbb{N}^*, \mathbf{t} \in \mathbb{T}_f^{(h)})$ (resp. $(\mathbb{P}(r_{h,k}(T) = \mathbf{t}); h, k \in \mathbb{N}^*, \mathbf{t} \in \mathbb{T}_k^{(h)})$). Using the Portmanteau theorem, we deduce the following results :

- Let $(T_n, n \in \mathbb{N})$ and T be \mathbb{T}_f -valued random variables. Then we have the following characterization of the convergence in distribution if a.s. $H(T) = +\infty$:

$$T_n \xrightarrow[n \rightarrow \infty]{(d)} T \iff \lim_{n \rightarrow \infty} \mathbb{P}(r_h(T_n) = \mathbf{t}) = \mathbb{P}(r_h(T) = \mathbf{t}) \quad \text{for all } h \in \mathbb{N}^*, \mathbf{t} \in \mathbb{T}_f^{(h)}. \quad (4.1)$$

- Let $(T_n, n \in \mathbb{N})$ and T be \mathbb{T}_f^* -valued random variables. Then we have the following characterization of the convergence in distribution if a.s. $H(T) = +\infty$, $k_\emptyset(T) = +\infty$:

$$T_n \xrightarrow[n \rightarrow \infty]{(d)} T \iff \lim_{n \rightarrow \infty} \mathbb{P}(r_{h,k}(T_n) = \mathbf{t}) = \mathbb{P}(r_{h,k}(T) = \mathbf{t}) \quad \text{for all } h, k \in \mathbb{N}^*, \mathbf{t} \in \mathbb{T}_k^{(h)}. \quad (4.2)$$

4.2.3 Galton-Watson trees

Let $p = (p(n), n \in \mathbb{N})$ be a probability distribution on \mathbb{N} . A \mathbb{T}_f -valued random variable τ is called a GW tree with offspring distribution p if for all

$h \in \mathbb{N}^*$ and $\mathbf{t} \in \mathbb{T}_f$ with $H(\mathbf{t}) \leq h$:

$$\mathbb{P}(r_h(\tau) = \mathbf{t}) = \prod_{u \in r_{h-1}(\mathbf{t})} p(k_u(\mathbf{t})).$$

The generation size process defined by $(Z_n = z_n(\tau), n \in \mathbb{N})$ is the so called GW process. We refer to [8] for a general study of GW processes. We set \mathbb{P}_k the probability under which the GW process $(Z_n)_{n \geq 0}$ starts with $Z_0 = k$ individuals and write \mathbb{P} for \mathbb{P}_1 so that

$$\mathbb{P}_k(Z_n = a) = \mathbb{P}(Z_n^{(1)} + \dots + Z_n^{(k)} = a)$$

where the $(Z^{(i)}, 1 \leq i \leq k)$ are i.i.d. copies of Z under \mathbb{P} .

We consider a sequence $(a_n, n \in \mathbb{N}^*)$ of elements in \mathbb{N}^* and, when $\mathbb{P}(Z_n = a_n) > 0$, τ_n a random tree distributed as the GW tree τ conditionally on $\{Z_n = a_n\}$. Let $n \geq h \geq 1$ and $\mathbf{t} \in \mathbb{T}_f^{(h)}$. We have, with $k = z_h(\mathbf{t})$:

$$\mathbb{P}(r_h(\tau_n) = \mathbf{t}) = \mathbb{P}(r_h(\tau) = \mathbf{t}) \frac{\mathbb{P}_k(Z_{n-h} = a_n)}{\mathbb{P}(Z_n = a_n)}. \quad (4.3)$$

4.2.4 Geometric distribution

Let $\eta \in (0, 1]$ and $q \in (0, 1)$. We define the geometric $\mathcal{G}(\eta, q)$ distribution $p = (p(k), k \in \mathbb{N})$ by

$$\begin{cases} p(0) = 1 - \eta, \\ p(k) = \eta q (1 - q)^{k-1} \quad \text{for } k \in \mathbb{N}^*. \end{cases} \quad (4.4)$$

We shall always consider that τ is a GW tree with geometric offspring distribution $\mathcal{G}(\eta, q)$.

The mean of $\mathcal{G}(\eta, q)$ is given by $\mu = \eta/q$ and its generating function φ is given by :

$$\varphi(s) = \frac{(1 - \eta) - s(1 - q - \eta)}{1 - s(1 - q)}, \quad s \in [0, 1/(1 - q)).$$

We set :

$$\gamma = \frac{1}{1 - q} \quad \text{and} \quad \kappa = \frac{1 - \eta}{1 - q} \quad (4.5)$$

where γ is the radius of convergence of φ and κ and 1 are the only fixed points of φ on $[0, \gamma)$. If $\mu = 1$ then there is only one fixed point as $\kappa = 1$. We shall use frequently the following relations :

$$\gamma - \kappa = \mu(\gamma - 1) \quad \text{and, if } \mu \neq 1, \quad \gamma - 1 = \frac{\kappa - 1}{1 - \mu}. \quad (4.6)$$

Notice that $\kappa \in [0, +\infty)$ and $\gamma \in (1, +\infty)$ allow to recover η and q as :

$$\eta = 1 - \frac{\kappa}{\gamma} \quad \text{and} \quad q = 1 - \frac{1}{\gamma}. \quad (4.7)$$

For this reason, we shall also write $\mathcal{G}[\kappa, \gamma]$ for $\mathcal{G}(\eta, q)$. Notice that if $\mu < 1$, then $q > \eta$ and $\gamma > \kappa > 1$; and if $\mu > 1$, then $\eta > q$ and $\gamma > 1 > \kappa \geq 0$.

Since φ is an homography, we get for $s \in [0, \gamma) \setminus \{1\}$:

$$\frac{\varphi(s) - \kappa}{\varphi(s) - 1} = \frac{1}{\mu} \frac{s - \kappa}{s - 1}. \quad (4.8)$$

We set $\varphi_1 = \varphi$ and, for $n \in \mathbb{N}^*$, $\varphi_{n+1} = \varphi \circ \varphi_n$. Notice that κ is a fixed point of φ_n as it is a fixed point of φ . We deduce from (4.8) and the second equality of (4.6) if $\mu \neq 1$ and by direct recurrence if $\mu = 1$, that φ_n , for $n \in \mathbb{N}^*$, is the generating function of the geometric distribution $\mathcal{G}[\kappa, \gamma_n] = \mathcal{G}(\eta_n, q_n)$ with mean $\mu_n = \mu^n$ and, thanks to (4.7) :

$$\eta_n = 1 - \frac{\kappa}{\gamma_n}, \quad q_n = 1 - \frac{1}{\gamma_n} \quad \text{with} \quad \gamma_n = \begin{cases} \frac{\kappa - \mu^n}{1 - \mu^n} = 1 + (\gamma - 1) \frac{q^{n-1}(q - \eta)}{q^n - \eta^n} & \text{if } \mu \neq 1, \\ 1 + (\gamma - 1) \frac{1}{n} & \text{if } \mu = 1. \end{cases} \quad (4.9)$$

By convention, we set φ_0 the identity function defined on $[0, +\infty)$ and $\gamma_0 = +\infty$ so that for all $n \in \mathbb{N}$, we have $\gamma_n = \lim_{r \rightarrow +\infty} \varphi_n^{-1}(r)$ that is in short $\gamma_n = \varphi_n^{-1}(\infty)$. We deduce that for all $n \geq \ell \geq 0$:

$$\varphi_\ell(\gamma_n) = \gamma_{n-\ell}. \quad (4.10)$$

We derive some asymptotics for γ_n for large n . It is easy to deduce from (4.9) that :

$$\lim_{n \rightarrow \infty} \gamma_n = \max(1, \kappa) = \begin{cases} \kappa & \text{if } \mu \leq 1, \\ 1 & \text{if } \mu \geq 1. \end{cases} \quad (4.11)$$

Using (4.6), we get for large n :

$$(\gamma_n - \kappa)(\gamma_n - 1) = \begin{cases} \mu^n (\kappa - 1)^2 + O(\mu^{2n}) & \text{if } \mu < 1, \\ (\gamma - 1)^2 n^{-2} & \text{if } \mu = 1, \\ \mu^{-n} (\kappa - 1)^2 + O(\mu^{-2n}) & \text{if } \mu > 1. \end{cases} \quad (4.12)$$

We derive from (4.9) the logarithm asymptotics of γ_n / γ_{n-h} for given $h \in \mathbb{N}^*$ and large n :

$$\log(\gamma_{n-h} / \gamma_n) = \log(\gamma_{n-h}) - \log(\gamma_n) = \begin{cases} \mu^{n-h} (1 - \mu^h) (\kappa - 1) / \kappa + O(\mu^{2n}) & \text{if } \mu < 1, \\ (\gamma - 1) h n^{-2} + O(n^{-3}) & \text{if } \mu = 1, \\ \mu^{-n} (\mu^h - 1) (1 - \kappa) + O(\mu^{-2n}) & \text{if } \mu > 1. \end{cases} \quad (4.13)$$

We recall the following well-known equality which holds for all $k \in \mathbb{N}^*$ and $r \in (0, 1)$:

$$\sum_{\ell \geq k} \binom{\ell-1}{k-1} r^\ell = \left(\frac{r}{1-r} \right)^k. \quad (4.14)$$

And we end this section with an elementary lemma.

Lemma 4.2.1. *Let $(X_\ell, \ell \in \mathbb{N}^*)$ be independent random variables with distribution $\mathcal{G}(\eta, q) = \mathcal{G}[\kappa, \gamma]$. For $a \geq k \geq 1$:*

$$\mathbb{P} \left(\sum_{\ell=1}^k X_\ell = a \right) = \sum_{i=1}^k \binom{k}{i} \binom{a-1}{i-1} \kappa^{k-i} (\gamma - \kappa)^i (\gamma - 1)^i \gamma^{-a-k}.$$

Démonstration. We have :

$$\begin{aligned} \mathbb{P} \left(\sum_{\ell=1}^k X_\ell = a \right) &= \sum_{i=1}^k \binom{k}{i} \mathbb{P}(X_1 = 0)^{k-i} \mathbb{P} \left(\sum_{\ell=1}^i X_\ell = a, X_\ell \geq 1 \text{ for } \ell \in \{1, \dots, i\} \right) \\ &= \sum_{i=1}^k \binom{k}{i} (1-\eta)^{k-i} \binom{a-1}{i-1} (\eta q)^i (1-q)^{a-i}. \end{aligned} \quad (4.15)$$

Then use (4.7) to conclude. \square

4.3 The geometric GW tree

Let τ be a GW tree with geometric $\mathcal{G}(\eta, q)$ offspring distribution p given by (4.4), with $\eta \in (0, 1]$ and $q \in (0, 1)$. We denote by $(Z_n, n \in \mathbb{N})$ the associated Galton-Watson process defined for every $n \in \mathbb{N}$ by $Z_n = z_n(\tau)$.

For $k \in \mathbb{N}^*$, we denote by \mathbb{P}_k the distribution of the geometric Galton-Watson forest composed of k i.i.d. geometric Galton-Watson trees, and write \mathbb{P} for \mathbb{P}_1 . For convenience, we shall under \mathbb{P} denote by $Z^{(k)} = (Z_n^{(k)}, n \in \mathbb{N})$ a GW process distributed as $Z = (Z_n, n \in \mathbb{N})$ under \mathbb{P}_k .

For $n \in \mathbb{N}^*$, we set :

$$M_n = \gamma_1^{-Z_1} \gamma_n^{Z_n}. \quad (4.16)$$

Since Z_n has generating function φ_n under \mathbb{P} , we deduce from (4.10) that $(M_n, n \in \mathbb{N}^*)$ is a martingale with $M_1 = 1$.

For $n \geq h \geq 1$, we set :

$$b_{n,h} = \left(\frac{\gamma_n}{\gamma_{n-h}} \right)^{a_n}. \quad (4.17)$$

We shall use the following formula when $\lim_{n \rightarrow \infty} b_{n,h}$ exists and belongs to $(0, \infty)$.

Lemma 4.3.1. *Let $n \geq h \geq 1$ and $k \in \mathbb{N}^*$. We have :*

$$\frac{\mathbb{P}_k(Z_{n-h} = a_n)}{\mathbb{P}(Z_n = a_n)} = b_{n,h} \sum_{i=1}^k \binom{k}{i} \kappa^{k-i} G_{n,h}(k, i), \quad (4.18)$$

with

$$G_{n,h}(k, i) = \binom{a_n - 1}{i - 1} \frac{\gamma_n}{\gamma_{n-h}^k} \frac{(\gamma_{n-h} - \kappa)^i (\gamma_{n-h} - 1)^i}{(\gamma_n - \kappa)(\gamma_n - 1)}. \quad (4.19)$$

Démonstration. Let $n \geq h \geq 1$. Since Z_n has distribution $\mathcal{G}[\kappa, \gamma_n]$, we obtain thanks to (4.5) :

$$\mathbb{P}(Z_n = a_n) = \eta_n q_n (1 - q_n)^{a_n - 1} = (\gamma_n - \kappa)(\gamma_n - 1) \gamma_n^{-a_n - 1}.$$

Using that Z_{n-h} is under \mathbb{P}_k distributed as the sum of k independent random variables with distribution $\mathcal{G}[\kappa, \gamma_{n-h}]$, we deduce from Lemma 4.2.1 that :

$$\begin{aligned} \frac{\mathbb{P}_k(Z_{n-h} = a_n)}{\mathbb{P}(Z_n = a_n)} &= \sum_{i=1}^k \binom{k}{i} \binom{a_n - 1}{i - 1} \kappa^{k-i} \frac{(\gamma_{n-h} - \kappa)^i (\gamma_{n-h} - 1)^i}{\gamma_{n-h}^{a_n + k}} \frac{\gamma_n^{a_n + 1}}{(\gamma_n - \kappa)(\gamma_n - 1)} \\ &= b_{n,h} \sum_{i=1}^k \binom{k}{i} \kappa^{k-i} G_{n,h}(k, i). \end{aligned}$$

This gives the result. \square

We shall use the following formula when $\lim_{n \rightarrow \infty} b_{n,h} = 0$ and $\lim_{n \rightarrow \infty} a_n = +\infty$.

Lemma 4.3.2. *Let $n > h \geq 1$, $k_0 \in \mathbb{N}^*$ and $\mathbf{t} \in \mathbb{T}_{h,k_0}$. We have, with $a_n \geq k = z_h(\mathbf{t})$:*

$$\mathbb{P}(r_{h,k_0}(\tau_n) = \mathbf{t}) = \frac{1 - q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t}) \left(\gamma_h^k - R_{n,h}^1(k) - R_{n,h}^2(k) \right), \quad (4.20)$$

with $\alpha_n = (\gamma_{n-h} - \kappa)(\gamma_{n-h} - 1)$, $x_n = \gamma_n / \gamma_{n-h}$ and :

$$0 \leq R_{n,h}^1(k) \leq b_{n,h} \frac{\alpha_n}{1 - x_n} \max(1, \kappa)^{k-1} 2^{2k-1} \left(2 + \left(\frac{\alpha_n}{1 - x_n} \right)^{k-1} + (\alpha_n a_n)^{k-1} \right), \quad (4.21)$$

$$R_{n,h}^2(k) = (\kappa + 1 - \gamma) \frac{\mathbb{P}_k(Z_{n-h} = a_n)}{\mathbb{P}(Z_n = a_n)}. \quad (4.22)$$

Démonstration. Let $n > h \geq 1$, $k_0 \in \mathbb{N}^*$ and $\mathbf{t} \in \mathbb{T}_{h,k_0}$. We set $k = z_h(\mathbf{t})$. For every $1 \leq j \leq k_0$, we denote by \mathbf{t}_j the subtree rooted at the j -th offspring of the root i.e.

$$u \in \mathbf{t}_j \iff ju \in \mathbf{t}.$$

In what follows, we denote by $\tilde{Z}^{(i)}$ a process distributed as $Z^{(i)}$ and independent of $Z^{(k)}$. We have :

$$\begin{aligned} \mathbb{P}(r_{h,k_0}(\tau_n) = \mathbf{t}) &= \sum_{i=0}^{+\infty} p(i+k_0) \left[\prod_{j=1}^{k_0} \mathbb{P}(r_{h-1}(\tau) = \mathbf{t}_j) \right] \frac{\mathbb{P}(Z_{n-h}^{(k)} + \tilde{Z}_{n-1}^{(i)} = a_n)}{\mathbb{P}(Z_n = a_n)} \\ &= \mathbb{P}(r_h(\tau) = \mathbf{t}) \sum_{i=0}^{+\infty} (1-q)^i \frac{\mathbb{P}(Z_{n-h}^{(k)} + \tilde{Z}_{n-1}^{(i)} = a_n)}{\mathbb{P}(Z_n = a_n)} \\ &= \frac{1-q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t})(A+B), \end{aligned}$$

where we used the branching property for the first and second equalities, the independence of $Z^{(k)}$ and $\tilde{Z}^{(i)}$ for the third, where

$$A = \sum_{\ell=0}^{a_n} \mathbb{P}(Z_{n-h}^{(k)} = \ell) \sum_{i=0}^{+\infty} p(i) \frac{\mathbb{P}(Z_{n-1}^{(i)} = a_n - \ell)}{\mathbb{P}(Z_n = a_n)} \text{ and } B = \left(\frac{\eta q}{1-q} - (1-\eta) \right) \frac{\mathbb{P}(Z_{n-h}^{(k)} = a_n)}{\mathbb{P}(Z_n = a_n)}.$$

We have :

$$A = \sum_{\ell=0}^{a_n} \mathbb{P}(Z_{n-h}^{(k)} = \ell) \frac{\mathbb{P}(Z_n = a_n - \ell)}{\mathbb{P}(Z_n = a_n)} = \sum_{\ell=0}^{a_n} \mathbb{P}(Z_{n-h}^{(k)} = \ell) \gamma_n^\ell = \left(\varphi_{n-h}(\gamma_n)^k - R_{n,h}^1(k) \right),$$

where we used that $k_\emptyset(\tau)$ has distribution p for the first equality, that Z_n has distribution $\mathcal{G}[\kappa, \gamma_n]$ for the second one and thus $\mathbb{P}(Z_n = k) = \eta_n q_n \gamma_n^{-(k-1)}$, and for the last one that :

$$R_{n,h}^1(k) = \sum_{\ell > 0} \mathbb{P}(Z_{n-h}^{(k)} = \ell + a_n) \gamma_n^{\ell + a_n}.$$

We have, with $\alpha_n = (\gamma_{n-h} - \kappa)(\gamma_{n-h} - 1)$ and $x_n = \gamma_n / \gamma_{n-h}$:

$$\begin{aligned} \mathbb{P}(Z_{n-h}^{(k)} = \ell + a_n) \gamma_n^{\ell + a_n} &= b_{n,h} \sum_{i=1}^k \binom{k}{i} \binom{\ell + a_n - 1}{i-1} \kappa^{k-i} (\gamma_{n-h} - \kappa)^i (\gamma_{n-h} - 1)^i \gamma_{n-h}^{-\ell-k} \gamma_n^\ell \\ &\leq b_{n,h} x_n^\ell \max(1, \kappa)^{k-1} \sum_{i=1}^k \binom{k}{i} \binom{\ell + a_n - 1}{i-1} \alpha_n^i, \end{aligned}$$

where we used Lemma 4.2.1 for the first equality and $\gamma_{n-h} \geq \max(1, \kappa)$ for the last. Using that $(x+y)^j \leq 2^{j-1}(x^j + y^j)$ for $j \in \mathbb{N}^*$ and $x, y \in (0, +\infty)$, we deduce that :

$$\binom{\ell + a_n - 1}{i-1} \leq \frac{2^{i-1}}{(i-1)!} (\ell^{i-1} + a_n^{i-1}).$$

We have the following rough bounds :

$$\begin{aligned}
0 \leq R_{n,h}^1(k) &\leq b_{n,h} \max(1, \kappa)^{k-1} 2^{k-1} \sum_{i=1}^k \alpha_n^i \binom{k}{i} \sum_{\ell>0} \left(\frac{\ell^{i-1}}{(i-1)!} x_n^\ell + a_n^{i-1} x_n^\ell \right) \\
&\leq b_{n,h} \frac{x_n \alpha_n}{1-x_n} \max(1, \kappa)^{k-1} 2^{k-1} \sum_{i=1}^k \binom{k}{i} \left(\left(\frac{\alpha_n}{1-x_n} \right)^{i-1} + (\alpha_n a_n)^{i-1} \right) \\
&\leq b_{n,h} \frac{\alpha_n}{1-x_n} \max(1, \kappa)^{k-1} 2^{2k-1} \left(2 + \left(\frac{\alpha_n}{1-x_n} \right)^{k-1} + (\alpha_n a_n)^{k-1} \right)
\end{aligned}$$

where we used that $x_n \in (0, 1)$ as the sequence $(\gamma_m, m \in \mathbb{N}^*)$ is non-increasing and that $\sum_{\ell>0} \ell^{i-1} x_n^\ell / (i-1)! \leq x_n(1-x_n)^{i-1}$ for the last inequality but one. Then use (4.10), which gives $\varphi_{n-h}(\gamma_n) = \gamma_h$, to get $A = \gamma_h^k - R_{n,h}^1(k)$ as well as (4.21).

We can rewrite the constant in B as $\left(\frac{\eta q}{1-q} - (1-\eta) \right) = -(\kappa + 1 - \gamma)$, so that $B = -R_{n,h}^2(k)$, see (4.22), and thus $A + B = \gamma_h^k - R_{n,h}^1(k) - R_{n,h}^2(k)$. This ends the proof. \square

4.4 The Kesten regime or the not so fat case

4.4.1 The Kesten tree

In this section, we denote by τ a GW tree with geometric $p = \mathcal{G}(\eta, q)$ with $\eta, q \in (0, 1)$. Recall that the extinction event $\mathcal{E} = \{H(\tau) < +\infty\}$ has probability $\mathfrak{c} = \min(1, \kappa)$. Moreover, as we assume $\eta < 1$, we have $\mathfrak{c} > 0$.

We define the probability distribution $\mathfrak{p} = (\mathfrak{p}(n), n \in \mathbb{N})$ by :

$$\mathfrak{p}(n) = \mathfrak{c}^{n-1} p(n) \quad \text{for } n \in \mathbb{N}. \quad (4.23)$$

The tree τ conditionally on the extinction event \mathcal{E} is distributed as a GW tree with offspring distribution \mathfrak{p} . We denote by \mathfrak{m} the mean of \mathfrak{p} . If $\mu \leq 1$, then we have $\mathfrak{p} = p$, $\mathfrak{m} = \mu$ and $\mathfrak{c} = 1$.

Let $k \in \mathbb{N}^*$. We define the k -th order size-biased probability distribution of p as $p_{[k]} = (p_{[k]}(n), n \in \mathbb{N})$ defined by :

$$p_{[k]}(n) = \frac{n!}{(n-k)! \varphi^{(k)}(1)} p(n) \quad \text{for } n \in \mathbb{N} \text{ and } n \geq k. \quad (4.24)$$

The generating function of $p_{[k]}$ is $\varphi_{[k]}(s) = s^k \varphi^{(k)}(s) / \varphi^{(k)}(1)$. The probability distribution $p_{[1]}$ is the so-called size-biased probability distribution of p .

For the distribution $\mathcal{G}(\eta, q)$, we have $\varphi^{(k)}(1) = k!\eta q^{-k}(1-q)^{k-1}$, so the k -th order size-biased probability distribution of p is given by :

$$p_{[k]}(n) = \binom{n}{k} q^{k+1} (1-q)^{n-k} \quad \text{for } n \in \mathbb{N} \text{ and } n \geq k. \quad (4.25)$$

We now define the so-called Kesten tree $\hat{\tau}^0$ associated with the offspring distribution p as a two-type GW tree where the vertices are either of type s (for survivor) or of type e (for extinction). It is then characterized as follows.

- The number of offspring of a vertex depends, conditionally on the vertices of lower or same height, only on its own type (branching property).
- The root is of type s.
- A vertex of type e produces only vertices of type e with offspring distribution \mathbf{p} .
- The random number of children of a vertex of type s has the size-biased distribution of \mathbf{p} that is $\mathbf{p}_{[1]}$ defined by (4.24) with $k = 1$. Furthermore, all of the children are of type e but one, uniformly chosen at random, which is of type s.

Informally the individuals of type s in $\hat{\tau}^0$ form an infinite spine on which are grafted independent GW trees distributed as τ conditionally on the extinction event \mathcal{E} .

We define $\tau^0 = \text{Ske}(\hat{\tau}^0)$ as the tree $\hat{\tau}^0$ when one forgets the types of the vertices. The distribution of τ^0 is given in the following classical result.

Lemma 4.4.1. *Let $p = \mathcal{G}(\eta, q)$ with $\eta, q \in (0, 1)$. The distribution of τ^0 is characterized by : for all $n \geq h \geq 1$ and $\mathbf{t} \in \mathbb{T}_f^{(h)}$ with $k = z_h(\mathbf{t})$:*

$$\mathbb{P}(r_h(\tau^0) = \mathbf{t}) = k \mathbf{c}^{k-1} \mathbf{m}^{-h} \mathbb{P}(r_h(\tau) = \mathbf{t}). \quad (4.26)$$

We give a short proof of this well-known result.

Démonstration. Since τ^0 belongs to \mathbb{T}_f and has infinite height, its distribution is indeed characterized by (4.26) for all $n \geq h \geq 1$ and $\mathbf{t} \in \mathbb{T}_f^{(h)}$ with $k = z_h(\mathbf{t})$.

Let $n \geq h \geq 1$, $\mathbf{t} \in \mathbb{T}_f^{(h)}$ and $v \in \mathbf{t}$ such that $|v| = h$. Let V be the vertex of type s at level h in $\hat{\tau}^0$. We have, with $k = z_h(\mathbf{t})$:

$$\begin{aligned} \mathbb{P}(r_h(\tau^0) = \mathbf{t}, V = v) &= \prod_{u \in \mathbf{t} \setminus \text{An}(\{v\}); |u| < h} \mathbf{p}(k_u(\mathbf{t})) \prod_{u \in \text{An}(\{v\})} \frac{1}{k_u(\mathbf{t})} \mathbf{p}_{[1]}(k_u(\mathbf{t})) \\ &= \mathbf{m}^{-h} \mathbf{c}^{\sum_{u \in r_{h-1}(\mathbf{t})} (k_u(\mathbf{t}) - 1)} \prod_{u \in r_{h-1}(\mathbf{t})} p(k_u(\mathbf{t})) \\ &= \mathbf{m}^{-h} \mathbf{c}^{k-1} \mathbb{P}(r_h(\tau) = \mathbf{t}), \end{aligned}$$

where we used (4.24) (with $k = 1$, $n = k_u(\mathbf{t})$ and p replaced by \mathbf{p}) and (4.23) (with $n = k_u(\mathbf{t})$) for the second equality and that $\sum_{u \in r_{h-1}(\mathbf{t})} (k_u(\mathbf{t}) - 1) = k - 1$ for the last one. Summing over all $v \in \mathbf{t}$ such that $|v| = h$ gives the result. \square

4.4.2 Convergence of the not so fat geometric GW tree

We consider a sequence $(a_n, n \in \mathbb{N}^*)$ with $a_n \in \mathbb{N}^*$ and a random tree τ_n distributed as the GW tree τ with offspring distribution $p = \mathcal{G}(\eta, q)$ conditionally on $\{Z_n = a_n\}$. We have the following result.

Proposition 4.4.2. *Let $\eta \in (0, 1)$ and $q \in (0, 1)$. Assume that $\lim_{n \rightarrow \infty} a_n \mu^n = 0$ if $\mu < 1$, $\lim_{n \rightarrow \infty} a_n n^{-2} = 0$ if $\mu = 1$ or $\lim_{n \rightarrow \infty} a_n \mu^{-n} = 0$ if $\mu > 1$. Then we have the following convergence in distribution :*

$$\tau_n \xrightarrow[n \rightarrow \infty]{(d)} \tau^0.$$

The critical case, $\mu = 1$, appears in Corollary 6.2 of [2] for general offspring distribution with second moment.

Démonstration. Let $h \in \mathbb{N}^*$ and $k \in \mathbb{N}^*$. Recall the definitions of $b_{n,h}$ in (4.17) and of $G_{n,h}$ in (4.19). According to Lemma 4.3.1, we have for $n \geq h \geq 1$ and $k \in \mathbb{N}^*$:

$$\frac{\mathbb{P}_k(Z_{n-h} = a_n)}{\mathbb{P}(Z_n = a_n)} = b_{n,h} \sum_{i=1}^k \binom{k}{i} \kappa^{k-i} G_{n,h}(k, i).$$

According to (4.17), we have $b_{n,h} = \exp(-a_n \log(\gamma_{n-h}/\gamma_n))$. We deduce from (4.13) and the hypothesis on $(a_n, n \in \mathbb{N}^*)$ that $\lim_{n \rightarrow \infty} a_n \log(\gamma_{n-h}/\gamma_n) = 0$ and thus $\lim_{n \rightarrow \infty} b_{n,h} = 1$. We deduce from (4.19), (4.11) and (4.12) that, for $k \geq i > 1$, $\lim_{n \rightarrow \infty} G_{n,h}(k, i) = 0$ and for $k \geq 1$:

$$\lim_{n \rightarrow \infty} G_{n,h}(k, 1) = \begin{cases} \kappa^{1-k} \mu^{-h} & \text{if } \mu < 1, \\ 1 & \text{if } \mu = 1, \\ \mu^h & \text{if } \mu > 1. \end{cases}$$

We deduce that :

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}_k(Z_{n-h} = a_n)}{\mathbb{P}(Z_n = a_n)} = \begin{cases} k \mu^{-h} & \text{if } \mu < 1 \\ k & \text{if } \mu = 1 \\ k \kappa^{k-1} \mu^h & \text{if } \mu > 1 \end{cases} = k \mathbf{c}^{k-1} \mathbf{m}^{-h}.$$

Then, as a.s. $H(\tau^0) = +\infty$, we can use the characterization (4.1) of the convergence in \mathbb{T}_f , as well as (4.3) and Lemma 4.4.1 to conclude. \square

4.5 The Poisson regime or the fat case

4.5.1 An infinite Poisson tree

Let $\theta \in (0, +\infty)$. We consider a two-type random tree $\hat{\tau}^\theta$ where the vertices are either of type s (for survivor) or of type e (for extinction). We define $\tau^\theta = \text{Ske}(\hat{\tau}^\theta)$ as the tree $\hat{\tau}^\theta$ when one forgets the types of the vertices of $\hat{\tau}^\theta$. We denote by $\mathcal{S}_h = \{u \in \tau^\theta; |u| = h \text{ and } u \text{ is of type s in } \hat{\tau}^\theta\}$ the set of vertices of $\hat{\tau}^\theta$ with type s at level $h \in \mathbb{N}$. Notice that $\hat{\tau}^\theta$ is completely characterized by τ^θ and $(\mathcal{S}_h, h \in \mathbb{N})$. Recall \mathbf{p} defined by (4.23) and the k -th order size-biased distribution, $p_{[k]}$, defined by (4.24). The random tree $\hat{\tau}^\theta$ is defined as follows.

- The root is of type s (i.e. $\mathcal{S}_0 = \{\emptyset\}$).
- The number of offspring of a vertex of type e does not depend on the vertices of lower or same height (branching property only for individuals of type e).
- A vertex of type e produces only vertices of type e with offspring distribution \mathbf{p} (as in the Kesten tree).
- For $h \in \mathbb{N}$, let $\Delta_h = \#\mathcal{S}_{h+1} - \#\mathcal{S}_h$ be the increase of number of vertices of type s between generations h and $h + 1$. Conditionally on $r_h(\tau^\theta)$ and $(\mathcal{S}_\ell, 0 \leq \ell \leq h)$, Δ_h is distributed as a Poisson random variable with mean $\theta \zeta_h$, where :

$$\zeta_h = \begin{cases} \mu^{-h-1}(1-\mu)(\kappa-1)/\kappa & \text{if } \mu < 1, \\ (\gamma-1) & \text{if } \mu = 1, \\ \mu^h(\mu-1)(1-\kappa) & \text{if } \mu > 1. \end{cases} \quad (4.27)$$

The vertex $u \in \mathcal{S}_h$ has $\kappa^s(u) \geq 1$ children of type s, with all the configurations $(\kappa^s(u), u \in \mathcal{S}_h)$ having the same probability, that is $1/\binom{\#\mathcal{S}_{h+1}-1}{\#\mathcal{S}_h-1} = 1/\binom{\#\mathcal{S}_{h+1}-1}{\Delta_h}$. (This breaks the branching property!) Furthermore, conditionally on $\kappa^s(u) = s_u \geq 1$, the vertex u has $\kappa^e(u)$ vertices of type e such that $k_u(\tau^\theta) = \kappa^s(u) + \kappa^e(u)$ has distribution $\mathbf{p}_{[s_u]}$ and the s_u individuals of type s are chosen uniformly at random among the $k_u(\tau^\theta)$ children.

More precisely, for $h \in \mathbb{N}$, $n \in \mathbb{N}$, $u \in \mathcal{S}_h$, $k_u \geq s_u \geq 1$, $A_u \subset \{1, \dots, k_u\}$ with $\#A_u = s_u$ and $\sum_{u \in \mathcal{S}_h} s_u = n + \#\mathcal{S}_h$, we have with

$$k = \sum_{u \in \mathcal{S}_h} k_u :$$

$$\begin{aligned} & \mathbb{P} \left(\kappa^s(u) + \kappa^e(u) = k_u \text{ and } \mathcal{S}_{h+1} \cap \{u1, \dots, uk_u\} = uA_u \quad \forall u \in \mathcal{S}_h \mid r_h(\tau^\theta), \mathcal{S}_h \right) \\ &= \frac{(\theta \zeta_h)^n}{n!} e^{-\theta \zeta_h} \frac{1}{\binom{\#\mathcal{S}_h + n - 1}{n}} \prod_{u \in \mathcal{S}_h} \frac{1}{\binom{k_u}{s_u}} \mathfrak{p}_{[s_u]}(k_u) \\ &= \frac{(\#\mathcal{S}_h - 1)!}{(\#\mathcal{S}_h + n - 1)!} (\theta(\gamma - 1)\zeta_h)^n e^{-\theta \zeta_h} \prod_{u \in \mathcal{S}_h} \mathfrak{p}^{(k_u)} \begin{cases} \mu^{-\#\mathcal{S}_h} & \text{if } \mu \leq 1, \\ \mu^{\#\mathcal{S}_h} \left(\frac{\mu}{\kappa}\right)^n & \text{if } \mu > 1, \end{cases} \end{aligned} \quad (4.28)$$

where we used (4.25) and (4.23) as well as (4.7) for the last equality.

By construction, a.s. individuals of type s have a progeny which does not suffer extinction whereas individuals of type e have a progeny which suffers extinction. Since the individuals of type s do not satisfy the branching property, the random tree $\hat{\tau}^\theta$ is not a multi-type GW tree. We stress out that $\hat{\tau}^\theta$ truncated at level h can be recovered from $r_h(\tau^\theta)$ and \mathcal{S}_h as all the ancestors of a vertex of type s are also of type s and a vertex of type s has at least one children of type s.

We have the following result.

Lemma 4.5.1. *Let $\eta \in (0, 1]$ and $q \in (0, 1)$. Let $\theta \in (0, +\infty)$. Let $n \geq h \geq 1$ and $\mathbf{t} \in \mathbb{T}_h$. We have, with $k = z_h(\mathbf{t})$:*

$$\mathbb{P}(r_h(\tau^\theta) = \mathbf{t}) = \mathcal{H}(h, k, \theta) \mathbb{P}(r_h(\tau) = \mathbf{t}),$$

where $\mathcal{H}(h, k, \theta)$ is equal to

$$\begin{aligned} & \mu^{-h} e^{-\theta(\mu^{-h}-1)(\kappa-1)/\kappa} \sum_{i=1}^k \binom{k}{i} \frac{(\theta \mu^{-h} (\kappa - 1)^2 / \kappa)^{i-1}}{(i-1)!} & \text{if } \mu < 1, \\ & e^{-\theta(\gamma-1)h} \sum_{i=1}^k \binom{k}{i} \frac{(\theta(\gamma-1)^2)^{i-1}}{(i-1)!} & \text{if } \mu = 1, \\ & \mu^h e^{-\theta(\mu^h-1)(1-\kappa)} \sum_{i=1}^k \binom{k}{i} \kappa^{k-i} \frac{(\theta \mu^h (1-\kappa)^2)^{i-1}}{(i-1)!} & \text{if } \mu > 1. \end{aligned}$$

Remark 4.5.2. *We deduce from Lemma 4.4.1 that $\tau^\theta \xrightarrow[\theta \rightarrow 0]{(d)} \tau^0$. Therefore the trees τ^θ appear as a generalization of the Kesten tree. We will also prove in Proposition 4.6.3 that a limit also exists when $\theta \rightarrow +\infty$.*

Démonstration. We consider only the super-critical case. The sub-critical case and the critical case can be handled in a similar way.

Let $h \in \mathbb{N}^*$, $\mathbf{t} \in \mathbb{T}_h$ and $S_h \subset \{u \in \mathbf{t}; |u| = h\}$ be non empty. In order to shorten the notations, we set $\mathcal{A} = \text{An}(S_h)$. Notice that \mathcal{A} is a tree. We

set, for $\ell \in \{0, \dots, h-1\}$, $S_\ell = \{u \in \mathcal{A}, |u| = \ell\}$ the vertices at level ℓ which have at least one descendant in S_h and $\Delta_\ell = \#S_{\ell+1} - \#S_\ell$. We recall that $\hat{\tau}^\theta$ truncated at level h can be recovered from $r_h(\tau^\theta)$ and S_h as all the ancestors of a vertex of type s is also of type s and a vertex of type s has at least one children of type s . We compute $\mathcal{C}_{S_h} = \mathbb{P}(r_h(\tau^\theta) = \mathbf{t}, S_h = S_h)$. We have, using (4.28) and (4.27) :

$$\begin{aligned} \mathcal{C}_{S_h} &= \left[\prod_{u \in \mathbf{t}, u \notin \mathcal{A}, |u| < h} \mathbf{p}(k_u(\mathbf{t})) \right] \\ &\quad \prod_{\ell=0}^{h-1} \left[\frac{(\#S_\ell - 1)!}{(\#S_{\ell+1} - 1)!} (\theta(\gamma - 1)\zeta_\ell)^{\Delta_\ell} e^{-\theta\zeta_\ell} \left[\prod_{u \in S_\ell} \mathbf{p}(k_u(\mathbf{t})) \right] \mu^{\#S_\ell} \left(\frac{\mu}{\kappa}\right)^{\Delta_\ell} \right] \\ &= \left[\prod_{u \in r_{h-1}(\mathbf{t})} \mathbf{p}(k_u(\mathbf{t})) \right] \frac{\left(\frac{\theta(\gamma-1)(\mu-1)(1-\kappa)}{\kappa}\right)^{\sum_{\ell=0}^{h-1} \Delta_\ell}}{(\#S_h - 1)!} e^{-\theta \sum_{\ell=1}^{h-1} \zeta_\ell} \prod_{\ell=0}^{h-1} \mu^{(\ell+1)\Delta_\ell + \#S_\ell} \\ &= \left[\prod_{u \in r_{h-1}(\mathbf{t})} \kappa^{k_u(\mathbf{t})-1} \right] \left[\prod_{u \in r_{h-1}(\mathbf{t})} p(k_u(\mathbf{t})) \right] \frac{\left(\frac{\theta(1-\kappa)^2}{\kappa}\right)^{\#S_h-1}}{(\#S_h - 1)!} e^{-\theta(\mu^h-1)(1-\kappa)} \mu^{h\#S_h} \\ &= \kappa^{z_h(\mathbf{t})-\#S_h} \mathbb{P}(r_h(\tau) = \mathbf{t}) \frac{\mu^h (\theta\mu^h(1-\kappa)^2)^{\#S_h-1}}{(\#S_h - 1)!} e^{-\theta(\mu^h-1)(1-\kappa)}, \end{aligned}$$

where we used for the third equality that $\sum_{\ell=0}^{h-1} \Delta_\ell = \#S_h - 1$, $\sum_{\ell=1}^{h-1} \zeta_\ell = (\mu^h - 1)(1 - \kappa)$ and $\sum_{\ell=0}^{h-1} (\ell+1)\Delta_\ell + \#S_\ell = \sum_{\ell=0}^{h-1} (\ell+1)\#S_{\ell+1} - \ell\#S_\ell = h\#S_h$. Since \mathcal{C}_{S_h} depends only of $\#S_h$, we shall write $\mathcal{C}_{\#S_h}$ for \mathcal{C}_{S_h} . Set $k = z_h(\mathbf{t}) = \#\{u \in \mathbf{t}; |u| = h\}$. Since $\#S_h \geq 1$ as the root if of type s , we obtain :

$$\mathbb{P}(r_h(\tilde{\tau}^\theta) = \mathbf{t}) = \sum_{i=1}^k \sum_{S_h \subset \{u \in \mathbf{t}; |u|=h\}} \mathbf{1}_{\{\#S_h=i\}} \mathcal{C}_{S_h} = \sum_{i=1}^k \binom{k}{i} \mathcal{C}_i = \mathbb{P}(r_h(\tau) = \mathbf{t}) \mathcal{H}(h, k, \theta),$$

where we used the definition of \mathcal{H} for the last equality. \square

4.5.2 Convergence of the fat geometric GW tree

We consider a sequence $(a_n, n \in \mathbb{N}^*)$, with $a_n \in \mathbb{N}^*$ and τ_n a random tree distributed as the GW tree τ conditionally on $\{Z_n = a_n\}$. We have the following result.

Proposition 4.5.3. *Let $\eta \in (0, 1]$, $q \in (0, 1)$ and $\theta \in (0, +\infty)$. Assume that $\lim_{n \rightarrow \infty} a_n \mu^n = \theta$ if $\mu < 1$ or $\lim_{n \rightarrow \infty} a_n n^{-2} = \theta$ if $\mu = 1$ or $\lim_{n \rightarrow \infty} a_n \mu^{-n} = \theta$ if $\mu > 1$. Then we have the following convergence in distribution :*

$$\tau_n \xrightarrow[n \rightarrow \infty]{(d)} \tau^\theta.$$

Démonstration. Recall the definitions of $b_{n,h}$ in (4.17) and of $G_{n,h}$ in (4.19). According to Lemma 4.3.1, we have for $n \geq h \geq 1$ and $k \in \mathbb{N}^*$:

$$\frac{\mathbb{P}_k(Z_{n-h} = a_n)}{\mathbb{P}(Z_n = a_n)} = b_{n,h} \sum_{i=1}^k \binom{k}{i} \kappa^{k-i} G_{n,h}(k, i).$$

According to Definition (4.17), we have $b_{n,h} = \exp(-a_n \log(\gamma_{n-h}/\gamma_n))$. We deduce from (4.13) and the hypothesis on $(a_n, n \in \mathbb{N}^*)$ that

$$\lim_{n \rightarrow \infty} -\log(b_{n,h}) = \begin{cases} \theta(\mu^{-h} - 1)(\kappa - 1)/\kappa & \text{if } \mu < 1, \\ \theta(\gamma - 1)h & \text{if } \mu = 1, \\ \theta(\mu^h - 1)(1 - \kappa) & \text{if } \mu > 1. \end{cases}$$

We deduce from (4.19), (4.11) and (4.12), that for $h \in \mathbb{N}^*$, $k \geq i \geq 1$:

$$\lim_{n \rightarrow \infty} (i-1)! G_{n,h}(k, i) = \begin{cases} (\theta\mu^{-h}(\kappa - 1)^2)^{i-1} \mu^{-h} \kappa^{1-k} & \text{if } \mu < 1, \\ (\theta(\gamma - 1)^2)^{i-1} & \text{if } \mu = 1, \\ (\theta\mu^h(1 - \kappa)^2)^{i-1} \mu^h & \text{if } \mu > 1. \end{cases}$$

Using definition of \mathcal{H} in Lemma 4.5.1, we obtain that :

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}_k(Z_{n-h} = a_n)}{\mathbb{P}(Z_n = a_n)} = \mathcal{H}(h, k, \theta).$$

Then use the characterization of the convergence in \mathbb{T}_f , (4.3) and Lemma 4.5.1 to conclude. \square

4.6 The condensation regime or the very fat case

4.6.1 An infinite geometric tree

Recall γ_n defined in (4.9). For $n \in \mathbb{N}^*$, we define the probability $\tilde{p}_n = (\tilde{p}_n(k), k \in \mathbb{N})$ by :

$$\tilde{p}_n(k) = \frac{\gamma_{n+1}^k}{\gamma_n} p(k).$$

Thanks to (4.10), we get $\sum_{k \in \mathbb{N}} \tilde{p}_n(k) = \varphi(\gamma_{n+1})\gamma_n^{-1} = 1$, so that \tilde{p} is indeed a probability distribution on $\bar{\mathbb{N}}$. For $n = 0$, we set \tilde{p}_0 the Dirac mass at $+\infty$, which is a degenerate probability measure on $\bar{\mathbb{N}}$.

We define τ^∞ as a Galton-Watson tree with reproduction distribution \tilde{p}_h at generation $h \in \mathbb{N}$. In particular the root has an infinite number of children, whereas all the other individuals have a finite number of children. More precisely, for all $h \in \mathbb{N}^*$, $k_0 \in \mathbb{N}^*$ and $\mathbf{t} \in \mathbb{T}_{h,k_0}$, we have :

$$\mathbb{P}(r_{h,k_0}(\tau^\infty) = \mathbf{t}) = \prod_{u \in r_{h-1}(\mathbf{t})^*} \tilde{p}_{|u|}(k_u(\mathbf{t})), \quad (4.29)$$

where we recall that $\mathbf{t}^* = \mathbf{t} \setminus \{\emptyset\}$. Remark that a.s. $\tau^\infty \in \mathbb{T}_f^*$.

We give a representation of the distribution of τ^∞ as the distribution of τ with a martingale weight.

Lemma 4.6.1. *Let $\eta \in (0, 1]$ and $q \in (0, 1)$. For all $h \in \mathbb{N}^*$, $k_0 \in \mathbb{N}^*$ and F a non-negative function on \mathbb{T}_∞ , we have :*

$$\mathbb{E}[F(r_{h,k_0}(\tau^\infty))] = \frac{\mathbb{E}[F(r_h(\tau)) M_h \mathbf{1}_{\{k_\emptyset(\tau)=k_0\}}]}{\mathbb{P}(k_\emptyset(\tau) = k_0)},$$

where $(M_h, h \in \mathbb{N}^*)$ is the martingale defined by (4.16). Equivalently, for all $h \in \mathbb{N}^*$, $k_0 \in \mathbb{N}^*$ and $\mathbf{t} \in \mathbb{T}_{k_0}^{(h)}$, we have with $k = z_h(\mathbf{t})$:

$$\mathbb{P}(r_{h,k_0}(\tau^\infty) = \mathbf{t}) = \frac{1-q}{\eta q} \gamma_h^k \mathbb{P}(r_h(\tau) = \mathbf{t}). \quad (4.30)$$

Démonstration. Let $h \in \mathbb{N}^*$, $k_0 \in \mathbb{N}^*$ and $\mathbf{t} \in \mathbb{T}_{h,k_0}$. Set $k = z_h(\mathbf{t})$. We have :

$$\begin{aligned} \frac{1-q}{\eta q} \gamma_h^k \mathbb{P}(r_h(\tau) = \mathbf{t}) &= \frac{1-q}{\eta q} \left[\prod_{u \in \mathbf{t}, |u|=h-1} \gamma_h^{k_u(\mathbf{t})} \right] \left[\prod_{u \in r_{h-1}(\mathbf{t})} p(k_u(\mathbf{t})) \right] \\ &= \frac{1-q}{\eta q} \gamma_1^{k_0} \left[\prod_{u \in r_{h-1}(\mathbf{t})^*} \gamma_{|u|}^{-1} \gamma_{|u|+1}^{k_u(\mathbf{t})} \right] \left[\prod_{u \in r_{h-1}(\mathbf{t})} p(k_u(\mathbf{t})) \right] \\ &= \frac{1-q}{\eta q} \gamma_1^{k_0} p(k_0) \left[\prod_{u \in r_{h-1}(\mathbf{t})^*} \tilde{p}_{|u|}(k_u(\mathbf{t})) \right] \\ &= \mathbb{P}(r_{h,k_0}(\tau^\infty) = \mathbf{t}), \end{aligned}$$

where we used that $\sum_{u \in \mathbf{t}, |u|=\ell} k_u(\mathbf{t}) = \sum_{u \in \mathbf{t}, |u|=\ell+1} 1$ for the second equality and the definition of $p(k_0)$ and $\gamma_1 = \gamma$ as well as (4.29) for the last one. To conclude, notice also that thanks to the definition of $p(k_0)$ and $\gamma_1 = \gamma$ as well as (4.16), we have on $\{k_\emptyset(\tau) = k_0\}$:

$$\frac{1-q}{\eta q} \gamma_h^{z_h(\tau)} = \frac{M_h}{p(k_0)}.$$

□

We give an alternative description of τ^∞ as the skeleton of a two-type GW tree. We set for $n \in \mathbb{N}$:

$$\nu_n = 1 - \frac{\gamma_{n+1} - 1}{\gamma_1 - 1} = \begin{cases} \mu(1 - \mu^n)(1 - \mu^{n+1})^{-1} & \text{if } \mu \neq 1, \\ n(n+1)^{-1} & \text{if } \mu = 1. \end{cases}$$

We have $\nu_n \in [0, 1)$. It is easy to check (using the first expression of ν_{n-1} for the first equality and the second expression for ν_{n-1} and ν_n for the second equality) that for all $n \in \mathbb{N}^*$:

$$\frac{1 - q\nu_{n-1}}{1 - q} = \gamma_n \quad \text{and} \quad \frac{1}{\mu}(1 - \nu_{n-1})\frac{\nu_n}{1 - \nu_n} = 1. \quad (4.31)$$

We consider a two-type GW tree $\hat{\tau}^\infty$ where the vertices are either of type s (for survivor) or of type n (for normal). We define $\text{Ske}(\hat{\tau})$ as the tree $\hat{\tau}^\infty$ when one forgets the types of the vertices of $\hat{\tau}^\infty$. We denote by $\mathcal{S}_h = \{u \in \text{Ske}(\hat{\tau}^\infty); |u| = h \text{ and } u \text{ is of type s in } \hat{\tau}^\infty\}$ the set of vertices of $\hat{\tau}$ with type s at level $h \in \mathbb{N}$. The random tree $\hat{\tau}^\infty$ is defined as follows :

- The number of offspring of a vertex depends, conditionally on the vertices of lower or same height, only on its own type (branching property).
- The root is of type s (i.e. $\mathcal{S}_0 = \{\emptyset\}$).
- A vertex of type n produces only vertices of type n with offspring distribution $\mathcal{G}(\eta, q)$. Notice the difference with the Kesten tree in Section 4.4.1 and the tree described in Section 4.5.1 for $\mu > 1$.
- A vertex $u \in \hat{\tau}^\infty$ at level h of type s produces $\kappa^s(u)$ vertices of type s with probability distribution $\mathcal{G}(1, \nu_h)$ (with the convention that if $h = 0$, then $\kappa^s(\emptyset) = +\infty$) and $\kappa^n(u)$ vertices of type n such that the type of the vertices $(ui, 1 \leq i \leq \kappa^s(u) + \kappa^n(u))$ is a sequence of heads (type s) and tails (type n) where the probability to get an head is q and a tail is $1 - q$, stopped just before the $(\kappa^s(u) + 1)$ -th head. Equivalently, for $|u| \geq 1$, conditionally on $\kappa^s(u) = s_u \geq 1$, the vertex u has $\kappa^n(u)$ vertices of type n such that $k_u(\text{Ske}(\hat{\tau}^\infty)) = \kappa^s(u) + \kappa^n(u)$ has distribution $p_{[s_u]}$, defined in (4.25), and the s_u individuals of type s are chosen uniformly at random among the $k_u(\text{Ske}(\hat{\tau}^\infty))$ children. More precisely, we have for $k_0 \in \mathbb{N}^*$ and $S_1 \subset \{1, \dots, k_0\}$:

$$\mathbb{P}(\mathcal{S}_1 \cap \{1, \dots, k_0\} = S_1) = q^{\#S_1} (1 - q)^{k_0 - \#S_1},$$

and for $h \geq 2$, $k \in \mathbb{N}^*$, $u \in \mathcal{U}$ with $|u| = h$, $s_u \in \{1, \dots, k\}$, and $A \subset \{1, \dots, k\}$ such that $\#A = s_u$:

$$\begin{aligned} \mathbb{P}(\kappa^s(u) + \kappa^n(u) = k, \mathcal{S}_{h+1} \cap \{u1, \dots, uk\} = uA \mid u \in \mathcal{S}_h) \\ = \nu_h (1 - \nu_h)^{s_u - 1} q^{s_u + 1} (1 - q)^{k - s_u}. \end{aligned}$$

By construction individuals of type s have a progeny which does not suffer extinction. Notice that if $\mu \leq 1$, then individuals of type n have a.s. a finite progeny. However in the super-critical case $\mu > 1$, the GW trees which are grafted have the same distribution as the initial GW tree τ , and thus individuals of type n (for normal!) have a probability $1 - \kappa$ of having an infinite progeny. This is a major difference with Sections 4.4.1 and 4.5.1,

where the grafted trees had the same distribution as the initial GW tree τ conditioned to be extinct.

We stress out that $\hat{\tau}^\infty$, truncated at level h and when considering only the first k_0 children of the root, can be recover from $r_{h,k_0}(\text{Ske}(\hat{\tau}^\infty))$ and \mathcal{S}_h as all the ancestors of a vertex of type s is also of a type s and a vertex of type s has at least one children of type s .

We have the following result.

Lemma 4.6.2. *Let $\eta \in (0, 1]$ and $q \in (0, 1)$. We have that τ^∞ is distributed as $\text{Ske}(\hat{\tau}^\infty)$.*

Démonstration. Let $h \in \mathbb{N}^*$, $k_0 \in \mathbb{N}^*$, $\mathbf{t} \in \mathbb{T}_{h,k_0}$ and $S_h \subset \{u \in \mathbf{t}; |u| = h\}$ which might be empty. In order to shorten the notations, we set $\mathcal{A} = \text{An}(S_h)$ which is a tree if S_h is non-empty. For $u \in \mathcal{A}$, we set $s_u = \#\{i \in \mathbb{N}; ui \in \mathcal{A} \cup S_h\}$ the number of children of u which have at least one descendant in S_h . We set, for $\ell \in \{0, \dots, h-1\}$, $S_\ell = \{u \in \mathcal{A}, |u| = \ell\}$ the vertices at level ℓ which have at least one descendant in S_h . Notice that $\sum_{u \in S_\ell} s_u = \#S_{\ell+1}$. Set $k = z_h(\mathbf{t})$. We compute $\mathcal{C}_{S_h} = \mathbb{P}(r_{h,k_0}(\text{Ske}(\hat{\tau}^\infty)) = \mathbf{t}, \mathcal{S}_h = S_h)$. If S_h is non-empty, we have :

$$\begin{aligned} \mathcal{C}_{S_h} &= \left[\prod_{u \in r_{h-1}(\mathbf{t}), u \notin \mathcal{A}} p(k_u(\mathbf{t})) \right] q^{\#S_1} (1-q)^{k_0 - \#S_1} \prod_{u \in \mathcal{A}^*} \nu_{|u|} (1 - \nu_{|u|})^{s_u - 1} q^{s_u + 1} (1-q)^{k_u(\mathbf{t}) - s_u} \\ &= \left[\prod_{u \in r_{h-1}(\mathbf{t})^*} p(k_u(\mathbf{t})) \right] q^{\#S_1} (1-q)^{k_0 - \#S_1} \prod_{u \in \mathcal{A}^*} \frac{\nu_{|u|}}{1 - \nu_{|u|}} \frac{1-q}{\eta} \left(\frac{q}{1-q} (1 - \nu_{|u|}) \right)^{s_u} \\ &= \frac{1-q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t}) \left(\frac{q}{1-q} \right)^{\#S_1} \prod_{\ell=1}^{h-1} \left(\frac{\nu_\ell}{1 - \nu_\ell} \frac{1-q}{\eta} \right)^{\#S_\ell} \left(\frac{q}{1-q} (1 - \nu_\ell) \right)^{\#S_{\ell+1}} \\ &= \frac{1-q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t}) \left(\frac{\nu_1}{1 - \nu_1} \frac{q}{\eta} \right)^{\#S_1} \left(\frac{q}{1-q} (1 - \nu_{h-1}) \right)^{\#S_h} \prod_{\ell=2}^{h-1} \left(\frac{\nu_\ell}{1 - \nu_\ell} \frac{q}{\eta} (1 - \nu_{\ell-1}) \right)^{\#S_\ell} \\ &= \frac{1-q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t}) \left(\frac{q}{1-q} (1 - \nu_{h-1}) \right)^{\#S_h}, \end{aligned}$$

where we used for the second equality that if $u \in \mathcal{A}$ and $\mathcal{S}_h = S_h$, then $k_{\text{Ske}(\hat{\tau}^\infty)}(u) \geq 1$; and for the fifth the second equation from (4.31) as well as $\nu_1/(1 - \nu_1) = \mu = \eta/q$ (which comes also from the second equation in (4.31) with $n = 0$). If S_h is empty, then we have :

$$\mathcal{C}_\emptyset = (1-q)^{k_0} \prod_{u \in r_{h-1}(\mathbf{t})^*} p(k_u(\mathbf{t})) = \frac{1-q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t}).$$

Notice that \mathcal{C}_{S_h} depends on S_h only through $\#S_h$. We deduce that :

$$\begin{aligned}
 \mathbb{P}(r_{h,k_0}(\text{Ske}(\hat{\tau}^\infty)) = \mathbf{t}) &= \sum_{i=0}^k \sum_{S_h \subset \{u \in \mathbf{t}; |u|=h\}} \mathbf{1}_{\{\#S_h=i\}} \mathcal{C}_{S_h} \\
 &= \frac{1-q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t}) \sum_{i=0}^k \binom{k}{i} \left(\frac{q}{1-q} (1 - \nu_{h-1}) \right)^i \\
 &= \frac{1-q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t}) \left(1 + \frac{q}{1-q} (1 - \nu_{h-1}) \right)^k \\
 &= \frac{1-q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t}) \left(\frac{1 - q\nu_{h-1}}{1-q} \right)^k \\
 &= \frac{1-q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t}) \gamma_h^k,
 \end{aligned}$$

where we used the first equation from (4.31) for the last equality. Then we conclude using (4.30) from Lemma 4.6.1. \square

As in Remark 4.5.2, we also have the convergence of the trees τ^θ introduced in Section 4.5.1 to the infinite geometric tree τ^∞ as $\theta \rightarrow +\infty$.

Proposition 4.6.3. *Let $\eta \in (0, 1]$ and $q \in (0, 1)$. Then we have the following convergence in distribution :*

$$\tau^\theta \xrightarrow[\theta \rightarrow \infty]{(d)} \tau^\infty.$$

Démonstration. We only deal with the supercritical case, the subcritical and critical cases can be handled in a similar way.

For $\mathbf{t}, \mathbf{t}' \in \mathbb{T}_f$, let us denote by $\mathbf{t} * \mathbf{t}'$ the tree obtained by grafting \mathbf{t} and \mathbf{t}' on the same root i.e. :

$$\mathbf{t} * \mathbf{t}' = \mathbf{t} \cup \{(u_1 + k_\emptyset(\mathbf{t}), u_2, \dots, u_n), (u_1, \dots, u_n) \in \mathbf{t}'^*\},$$

with the convention $\mathbf{t} * \mathbf{t}' = \mathbf{t}$ if $\mathbf{t}' = \{\emptyset\}$.

We denote by $\mathbb{T}_f^{(\leq h)}$ the subset of \mathbb{T}_f of trees with height less than or equal to h . Let $h, k_0 > 0$ and let $\mathbf{t} \in \mathbb{T}_{k_0}^{(h)}$. Then using Lemma 4.5.1 with $k = z_h(\mathbf{t})$ and $k' = z_h(\mathbf{t}')$, we have :

$$\begin{aligned}
 &\mathbb{P}(r_{h,k_0}(\tau^\theta) = \mathbf{t}) \\
 &= \sum_{\mathbf{t}' \in \mathbb{T}_f^{(\leq h)}} \mathbb{P}(r_h(\tau^\theta) = \mathbf{t} * \mathbf{t}') \\
 &= \sum_{\mathbf{t}' \in \mathbb{T}_f^{(\leq h)}} \mu^h e^{-\theta(\mu^h - 1)(1 - \kappa)} \sum_{i=1}^{k+k'} \binom{k+k'}{i} \kappa^{k+k'-i} \frac{(\theta \mu^h (1 - \kappa)^2)^{i-1}}{(i-1)!} \mathbb{P}(r_h(\tau) = \mathbf{t} * \mathbf{t}').
 \end{aligned}$$

Let us remark that, if $\mathbf{t}' \neq \{\emptyset\}$, then

$$\begin{aligned} \mathbb{P}(r_h(\tau) = \mathbf{t} * \mathbf{t}') &= \frac{\mathbb{P}(r_h(\tau) = \mathbf{t})}{p(k_\emptyset(\mathbf{t}))} \frac{\mathbb{P}(r_h(\tau) = \mathbf{t}')}{p(k_\emptyset(\mathbf{t}'))} p(k_\emptyset(\mathbf{t}) + k_\emptyset(\mathbf{t}')) \\ &= \frac{1-q}{\eta q} \mathbb{P}(r_h(\tau) = \mathbf{t}) \mathbb{P}(r_h(\tau) = \mathbf{t}'). \end{aligned}$$

Since $\mathbb{P}(r_h(\tau^\theta) = \mathbf{t})$ converges to 0 as θ increases to infinity, we deduce that for $\theta \rightarrow +\infty$:

$$\mathbb{P}(r_{h,k_0}(\tau^\theta) = \mathbf{t}) = \frac{1-q}{\eta q} \mu^h \mathbb{P}(r_h(\tau) = \mathbf{t}) e^{-\theta(\mu^h-1)(1-\kappa)} A_1 + o(1),$$

with

$$A_1 = \sum_{\mathbf{t}' \in \mathbb{T}_f^{(\leq h)} \setminus \{\emptyset\}} \sum_{i=1}^{k+k'} \binom{k+k'}{i} \kappa^{k+k'-i} \frac{(\theta \mu^h (1-\kappa)^2)^{i-1}}{(i-1)!} \mathbb{P}(r_h(\tau) = \mathbf{t}').$$

We have :

$$\begin{aligned} A_1 &= \sum_{k'=0}^{+\infty} \sum_{i=1}^{k+k'} \binom{k+k'}{i} \kappa^{k+k'-i} \frac{(\theta \mu^h (1-\kappa)^2)^{i-1}}{(i-1)!} \sum_{\{\mathbf{t}' \in \mathbb{T}_f^{(\leq h)}, z_h(\mathbf{t}')=k'\}} \mathbb{P}(r_h(\tau) = \mathbf{t}') \\ &= \sum_{k'=0}^{+\infty} \sum_{i=1}^{k+k'} \binom{k+k'}{i} \kappa^{k+k'-i} \frac{(\theta \mu^h (1-\kappa)^2)^{i-1}}{(i-1)!} \mathbb{P}(Z_h = k') \\ &= \sum_{k'=0}^{+\infty} \sum_{i=1}^{k+k'} \binom{k+k'}{i} \kappa^{k+k'-i} \frac{(\theta \mu^h (1-\kappa)^2)^{i-1}}{(i-1)!} \left(1 - \frac{1}{\gamma_h}\right) \left(1 - \frac{\kappa}{\gamma_k}\right) \frac{1}{\gamma_h^{k'-1}} \\ &= \left(1 - \frac{1}{\gamma_h}\right) \left(1 - \frac{\kappa}{\gamma_k}\right) (A_2 + A_3), \end{aligned}$$

where

$$A_2 = \sum_{i=k+1}^{+\infty} \left(\sum_{k'=i-k}^{+\infty} \binom{k+k'}{i} \left(\frac{\kappa}{\gamma_h}\right)^{k'-1} \right) \frac{(\theta \mu^h (1-\kappa)^2)^{i-1}}{(i-1)!} \kappa^{k-i+1}$$

and

$$A_3 = \sum_{i=1}^k \left(\sum_{k'=0}^{+\infty} \binom{k+k'}{i} \left(\frac{\kappa}{\gamma_h}\right)^{k'-1} \right) \frac{(\theta \mu^h (1-\kappa)^2)^{i-1}}{(i-1)!} \kappa^{k-i+1}.$$

Using (4.14), it is easy to check that $\lim_{\theta \rightarrow +\infty} e^{-\theta(\mu^h-1)(1-\kappa)} A_3 = 0$. We

have using (4.14) that :

$$\begin{aligned}
 A_2 &= \sum_{i=k+1}^{+\infty} \left(\sum_{k'=i-k}^{+\infty} \binom{k+k'}{i} \left(\frac{\kappa}{\gamma_h} \right)^{k'-1} \right) \frac{(\theta\mu^h(1-\kappa)^2)^{i-1}}{(i-1)!} \kappa^{k-i+1} \\
 &= \sum_{i=k+1}^{+\infty} \frac{1}{\left(1 - \frac{\kappa}{\gamma_h}\right)^{i+1}} \left(\frac{\kappa}{\gamma_h} \right)^{i-k-1} \frac{(\theta\mu^h(1-\kappa)^2)^{i-1}}{(i-1)!} \kappa^{k-i+1} \\
 &= \gamma_h^k \frac{\gamma_h - 1}{\gamma_h - \kappa} e^{\frac{\theta\mu^h(1-\kappa)^2}{\gamma_h - \kappa}}.
 \end{aligned}$$

Then, as $(\gamma_h - 1)/(\gamma_h - \kappa) = \mu^{-h}$ and $(1 - \kappa)/(\gamma_h - \kappa) = 1 - \mu^{-h}$, we get that :

$$\lim_{\theta \rightarrow +\infty} e^{-\theta(\mu^h-1)(1-\kappa)} A_1 = \lim_{\theta \rightarrow +\infty} e^{-\theta(\mu^h-1)(1-\kappa)} A_2 = \mu^{-h} \gamma_h^k.$$

We deduce that :

$$\lim_{\theta \rightarrow +\infty} \mathbb{P}(r_{h,k_0}(\tau^\theta) = \mathbf{t}) = \frac{1-q}{\eta q} \gamma_h^k \mathbb{P}(r_h(\tau) = \mathbf{t}).$$

Using (4.30), this gives the result. \square

4.6.2 Convergence of the very fat geometric GW tree

We consider a sequence $(a_n, n \in \mathbb{N}^*)$, with $a_n \in \mathbb{N}^*$ and τ_n a random tree distributed as the GW tree τ conditionally on $\{Z_n = a_n\}$. We have the following result.

Proposition 4.6.4. *Let $\eta \in (0, 1]$ and $q \in (0, 1)$. Assume that $\lim_{n \rightarrow \infty} a_n \mu^n = +\infty$ if $\mu < 1$ or $\lim_{n \rightarrow \infty} a_n n^{-2} = +\infty$ if $\mu = 1$ or $\lim_{n \rightarrow \infty} a_n \mu^{-n} = +\infty$ if $\mu > 1$. Then we have the following convergence in distribution :*

$$\tau_n \xrightarrow[n \rightarrow \infty]{(d)} \tau^\infty.$$

Démonstration. First notice that a.s. $H(\tau^\infty) = +\infty$. Then, using the characterization (4.2) for the convergence in distribution in \mathbb{T}_f^* , the result is a direct consequence of (4.20) in Lemma 4.3.2 and of (4.30) in Lemma 4.6.1, provided that $\lim_{n \rightarrow \infty} R_{n,h}^i(k) = 0$ for $i \in \{1, 2\}$, $h \geq 2$ and $k \in \mathbb{N}^*$, where $R_{n,h}^i$ are defined in (4.21) and (4.22).

According to (4.17) and the definitions in Lemma 4.3.2, we have $b_{n,h} = \exp(-a_n \log(\gamma_{n-h}/\gamma_n))$, $\alpha_n = (\gamma_{n-h} - \kappa)(\gamma_{n-h} - 1)$ and $x_n = \gamma_n/\gamma_{n-h}$. Since $\kappa > 1$ (resp. $\gamma > 1$, resp. $\kappa < 1$) if $\mu < 1$ (resp. $\mu = 1$, resp. $\mu > 1$), and since $h \geq 1$, we deduce from (4.9), (4.12) and (4.13) that $\log(\gamma_{n-h}/\gamma_n)$, α_n and $1 - x_n$ are of the same order μ^{-n} (resp. n^{-2} , resp. μ^n). In particular

$\lim_{n \rightarrow \infty} \alpha_n / (1 - x_n)$ exists and is finite. Because of the hypothesis on $(a_n, n \in \mathbb{N}^*)$, we deduce that $\lim_{n \rightarrow \infty} a_n \log(\gamma_{n-h} / \gamma_n) = +\infty$ and thus $\lim_{n \rightarrow \infty} b_{n,h} = 0$ as well as $\lim_{n \rightarrow \infty} b_{n,h} (\alpha_n a_n)^{k-1} = 0$ as $a_n \log(\gamma_{n-h} / \gamma_n)$ and $\alpha_n a_n$ are of the same order. This gives $\lim_{n \rightarrow \infty} R_{n,h}^1(k) = 0$

Since $p(k) \mathbb{P}_k(Z_{n-h} = a_n) \leq \sum_{i \in \mathbb{N}} p(i) \mathbb{P}_i(Z_{n-h} = a_n) = \mathbb{P}(Z_{n-h+1} = a_n)$, we deduce that :

$$\begin{aligned} \frac{\mathbb{P}_k(Z_{n-h} = a_n)}{\mathbb{P}(Z_n = a_n)} &\leq \frac{1}{p(k)} \frac{\mathbb{P}(Z_{n-h+1} = a_n)}{\mathbb{P}(Z_n = a_n)} \\ &= \frac{1}{p(k)} b_{n,h-1} \frac{(\gamma_{n-h+1} - \kappa)(\gamma_{n-h+1} - 1)}{(\gamma_n - \kappa)(\gamma_n - 1)} \frac{\gamma_n}{\gamma_{n-h+1}}, \end{aligned}$$

where we used that Z_ℓ has distribution $\mathcal{G}[\kappa, \gamma_\ell]$ and (4.9) for the last equality. According to the previous paragraph, we have $\lim_{n \rightarrow \infty} b_{n,h-1} = 0$ as $h \geq 2$. Furthermore, using (4.13), we get that :

$$\lim_{n \rightarrow \infty} \frac{(\gamma_{n-h+1} - \kappa)(\gamma_{n-h+1} - 1)}{(\gamma_n - \kappa)(\gamma_n - 1)} \frac{\gamma_n}{\gamma_{n-h+1}} = \mu^{-h+1}.$$

This implies that $\lim_{n \rightarrow \infty} \mathbb{P}_k(Z_{n-h} = a_n) / \mathbb{P}(Z_n = a_n) = 0$ and thus $\lim_{n \rightarrow \infty} R_{n,h}^2(k) = 0$. This finishes the proof. \square

Bibliographie

- [1] R. Abraham, A. Bouaziz and J. F. Delmas, *Local limits of Galton-Watson trees conditioned of the number of the protected nodes*, Appl. J. of Probab., 54 (01).
- [2] R. Abraham and J. F. Delmas, *Local limits of conditioned Galton-Watson trees : the infinite spine case*, Elec. J. of Probab., **19** : Article 2, 1-19, 2014.
- [3] R. Abraham and J. F. Delmas, *Local limits of conditioned Galton-Watson trees : the condensation case*, Elec. J. of Probab., **19** : Article 56, 1-29, 2014.
- [4] R. Abraham and J. F. Delmas, *An introduction to Galton-Watson trees and their locals limits*, arxiv ; 1506. 05571,2015.
- [5] D. Aldous. *The continuum random tree 3*, Ann. of Probab., **21** : 248-289, 1993.
- [6] D. Aldous and J. Pitman, *Tree-valued Markov chains derived from Galton-Watson processes*. Ann. de l'Inst. Henri Poincaré, 34 :637-686, 1998.
- [7] L. Alili and R. A. Doney, *Martin boundaries associated with a killed random walk*, Ann. Inst. H. Poincaré Probab. Statist. **37** (2001), no. 3, 313-338.
- [8] K. Athreya and P. Ney, *Branching processes*, Springer-Verlag, New York-Heidelberg,1972. Die Grundlehren der mathematischen Wissenschaften, Band **196**.
- [9] P. Bauman : *Positive solutions of elliptic equations in non-divergence form and their adjoints*. Arkiv Math 22, (1984), 153-173. MR-0765409.
- [10] A. Bouaziz, S. Mustapha and M. Sifi, *Discrete harmonic functions on an orthant in \mathbb{Z}^d* , Electron. Commun. Probab. **20** (2015), no. 52, 1-13.
- [11] P. Cartier, *Fonctions harmoniques sur un arbre*, Symposia Mathematica **9** (1972), 203-270.
- [12] N. Curien and I. Kortchemski, *Random non crossing plane configurations : a conditioned Galton-Watson tree approach*, Random Struct. and Alg.

- [13] L. Devroye and S. Janson, *Protected nodes and fringe subtrees in some random trees*, Elec, Comm. in Probab. **19**, 1-10
- [14] R. A. Doney, *The Martin boundary and ratio limit theorems for killed random walks*, J.London Math.Soc. **2** (1998), no. 58, 761-768.
- [15] J. L. Doob, *Discrete potential theory and boundaries*, J.Math. and Mech. **8** (1959), 433-458.
- [16] M. Drmota, *Random trees*, Springer Wien New York, Vienna, 2009. An interplay between combinatorics and probability.
- [17] T. Duquesne and J.F. Le Gall, *Random trees, Levy processes and spatial branching processes*, Number 281. Astérisque, 2002.
- [18] T. Duquesne, *A limit theorem for the contour process of conditioned Galton-Watson trees*, Ann. of Probab., **31** : 996-1027, 2003.
- [19] E. Dynkin, *The boundary theory of Markov processes (discrete case)*, Uspehi Mat. Nauk **24** (1969), 3-42.
- [20] S.N. Evans, *Probability and real trees*, volume 1920 of Lecture Notes in Mathematics. Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6-23, 2005.
- [21] G. Fayolle and K. Raschel, *Text about a possible analytic approach for walks in the quarter plane with arbitrary big jumps*. C. R. Math. Acad. Sci. Paris **353**, (2015), no. 2, 89-94. MR-3300936.
- [22] F. Galton and H. Watson, *On the probability of extinction of families*, J.Anthropol. Inst., 4 :138-144, 1874.
- [23] P. Haccou, P. Jagers, and V.A. Vatutin, *Branching processes : variation, growth, and extinction of populations*, volume **5** of Cambridge Studies in Adaptive Dynamics. Cambridge University Press, Cambridge ; Laxenburg, 2007.
- [24] T.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.
- [25] J. Hawkes, *Trees generated by a simple branching process*, J.London Math. Soc.(2) **24** (1981), no. 2, 373-384.
- [26] X. He, *Local convergence of critical random trees and continuum condensation tree*, (2015) , arXiv :1503.00951.
- [27] R. A. Hunt and R. L. Wheeden : *Positive harmonic functions on Lipschitz domains*. Trans. Amer. Math. Soc. 132, (1970), 507, MR-02074787.
- [28] I. Ignatiouk-Robert, *Martin boundary of a killed random walk on \mathbb{Z}^d* , Preprint arXiv :0909.3921 (2009), 1-49.
- [29] S. Janson, *Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation*, Probab. Surv., **9** (2012), 103-252.

- [30] S. Janson, *Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees*, Random Struct. and Alg. To appear.
- [31] T. Jonnsson and S. Stefansson. Condensation in nongeneric trees. J. Stat. Phys., 142 :227-313,2011.
- [32] A. Joffe and W.A.O. N. Waugh, *Exact distributions of kin numbers in a Galton-Watson process*, J. Appl. Prob. **19** (1982), no. 4, 767-775.
- [33] D.G. Kendall, *The genealogy of genealogy : branching processes before (and after) 1873*. Bull. London Math. Soc , **7** (3) : 225-253, 1975.
- [34] D. Kennedy, *The Galton-Watson process conditioned on the total progeny*. J. Appl. Probability, **12**(4) : 800–806, 1975.
- [35] H. Kesten, *Subdiffusive behavior of random walk on a random cluster*, Ann. Inst. H. Poincaré Probab. Statist. **22**(1986), no. 4, 425-487.
- [36] M. Kimmel and D.E. Axelrod, *Branching processes in biology*, volume 19 of interdisciplinary Applied Mathematics. Springer-Verlag, New York, 2002.
- [37] H.J. Kuo and N.S. Trudinger, *Positive difference operators on general meshes*, Duke Math. J. **83**, (1996), no. 2, 415-433.
- [38] H.J. Kuo and N.S. Trudinger, *Evolving monotone difference operators on general space-time meshes*, Duke Math.J. **91**, (1998), 587-607.
- [39] I.A. Kurkova and V.A. Malyshev, *Martin boundary and elliptic curves*, Markov Processes Related Fields **4** (1998), 203-272.
- [40] I. Kurkova and K. Raschel, *Random walks in \mathbb{Z}_+^2 with non-zero drift absorbed at the axes*, Bulletin de la Société Mathématique de France **139**, (2011), 341-387.
- [41] G.F. Lawler, *Estimates for differences and Harnack inequality for difference operators coming from random walks with symmetric, spatially inhomogeneous increments*, Proc London Math.Soc.(3)**63** (1991), no.3, 552-568.
- [42] R. Lyons, R. Pemantle, and Y. Peres, *Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes*, Ann. Probab. **23** (1995), no. 3, 1125-1138.
- [43] N. Minami, *On the number of vertices with a given degree in a Galton-Watson tree*, Adv. in Appl. Probab.**37** (2005), no. 1, 229-264.
- [44] S. Mustapha, *Gaussian estimates for stability inhomogeneous random walks on \mathbb{Z}^d* , The Annals of Probability 2006, Vol. **34**, No. 1, 264-283.
- [45] S. Mustapha, *Gambler's ruin estimates for random walks with symmetric spatially inhomogeneous increments*, Bernoulli **13** (2007), no. 1, 131-147.

- [46] J. Neveu, *Arbres et processus de Galton-Watson*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 199-207.
- [47] L. Overbeck *Martin boundary of some branching processes*. Ann. Inst. H. Poincaré Probab. Statist., **30** (2) : 181-195, 1994.
- [48] M. Picardello, and W. W. Woess, *Martin boundaries of Cartesian products of Markov chains*, Nagoya Math. J. **128** (1992), 153-169.
- [49] K. Raschel, *Green functions and Martin compactification for killed random walks related to $SU(3)$* , Electronic Communications in Probability **15** (2010), 176-190.
- [50] K. Raschel, *Random walks in the quarter plane, discrete harmonic functions and conformal mappings, with an appendix by Sandro Franceschi*, Stochastic Processes and their Applications **124** (2014), 3147-3178.
- [51] D. Rizzolo, *Scaling limits of Markov branching trees and Galton-Watson trees conditioned on the number of vertices with out-degree in a given set*, Ann.Inst.H.Poincaré Probab.Statist., (2014).
- [52] M. V. Safonov and Y. Yuan : *Doubling properties for second order parabolic equations*. Annals of Mathematics 150. (1999), 313-327. MR-1715327.
- [53] W. Woess, *Random walks on infinite graphs and groups*, Cambridge University Press, Cambridge, 2000.

Aymen Bouaziz

Marches aléatoires et arbres de Galton-Watson

Dans cette thèse nous nous sommes intéressés de trois types de problèmes :

-Existence et unicité d'une fonction harmonique strictement positive associée à une marche aléatoire inhomogène confinée dans un orthant.

-Etude de la convergence en loi des arbres de Galton Watson critiques conditionnés à avoir un nombre assez grand de noeuds protégés.

-Etude de la convergence en loi des arbres de Galton Watson conditionnés à avoir une génération anormalement grande.

Mots clés : Marches aléatoires, fonctions harmoniques discrètes, orthant, arbres aléatoires, arbres de Galton-Watson, limite locale, noeuds protégés.

Random Walk and Galton-Watson trees

In this thesis we are interested in three types of problems:

-Existence and uniqueness of a positive harmonic function associated with an inhomogeneous random walk confined in an orthant.

-Study of convergence in distribution of critical Galton Watson trees conditioned to have a large enough number of protected nodes.

-Study of the convergence in distribution of Galton Watson trees conditioned to have a large generation.

Key words: Random walks, discrete harmonic functions, orthant, random trees, Galton-Watson trees, local limit, protected nodes.