

Introduction au Domaine de la Recherche - Les marches aléatoires branchantes

Bastien Mallein

4 juillet 2011

Sous la direction de Zhan Shi

Résumé

Nous allons détailler ici quelques-unes des propriétés les mieux connues des marches aléatoires branchantes, notamment sur la position de l'individu le plus à droite dans ce processus. Nous introduirons également la méthode de décomposition en épine dorsale, un résultat d'importance dans l'étude de ces processus. Pour finir, nous nous intéresserons à une évolution récente de ces modèles, les marches aléatoires branchantes avec sélection.

Table des matières

1	Introduction aux marches aléatoires branchantes	2
1.1	Définition du modèle	3
1.2	Notations et quantités associées aux arbres marqués	3
1.3	Quelques quantités associées au processus de points \mathcal{L}	5
2	Méthode de décomposition en épine dorsale et Lemme de regroupement	5
2.1	Martingales additives exponentielles et propriété de branchement	5
2.2	Changement de probabilité	6
2.3	Le Lemme de regroupement	7
3	La position de l'individu le plus à droite	8
3.1	Vitesse de dispersion vers la droite	8
3.2	Un premier raffinement	9
3.3	Dernières avancées	10
4	Marche aléatoire branchante avec sélection	11

1 Introduction aux marches aléatoires branchantes

De très nombreux modèles de la théorie des probabilités décrivent l'évolution de populations au cours du temps. L'un des plus connus est le processus de Galton-Watson, dont le comportement détaillé est connu de longue date [3]. Dans un arbre de Galton-Watson, chaque individu engendre à la génération suivante un nombre aléatoire A d'enfants, indépendamment du reste du processus. Nous pouvons en particulier citer les résultats très classiques suivants :

- la probabilité d'extinction d'un arbre partant d'un unique individu à l'instant initial est égal à la plus petite solution sur $[0, 1]$ de l'équation $q = \mathbb{E}(q^A)$; en particulier $q = 1$ si et seulement si $m = \mathbb{E}(A) \leq 1$;
- si $m > 1$, alors le nombre d'individus présents à l'instant n divisé par m^n tend vers une constante (aléatoire) W , de plus on a :

$$\mathbb{P}(W = 0) = q \iff \mathbb{E}(A \log^+ A) < +\infty \text{ et } \mathbb{P}(W = 0) = 1 \iff \mathbb{E}(A \log^+ A) = +\infty,$$

où $\log^+ A = \max\{\log A, 0\}$.

Cette seconde propriété donne un ordre de grandeur du nombre d'individus présents dans l'arbre en cas de survie de la population. Remarquons en particulier que dans ce processus, on a soit extinction presque sûre, soit divergence presque sûre vers $+\infty$, à vitesse exponentielle dans la plupart des cas.

Ce processus peut par la suite être enrichi de la façon suivante : à chaque individu x de l'arbre, on associe une position $V(x)$. Cette position s'obtient de la façon suivante : tous les enfants d'un individu donné sont positionnés par rapport à leur parent selon un processus de points \mathcal{L} , indépendamment du reste du processus. Cette construction autorise en particulier que la nouvelle position de deux frères soit corrélée. Ce modèle décrit ainsi une population qui se reproduit tout en se déplaçant sur la droite réelle, colonisant ainsi tout l'espace disponible. Ce processus a été introduit par J.-F. C. Kingman au début des années 1970.

L'ajout de cette composante spatiale augmente la richesse de ce processus, car peut être interprétée de nombreuses façons différentes : comme la position de l'individu au moment de sa naissance bien sûr, mais aussi comme sa date de naissance, ou son degré d'adaptation au milieu. Ces interprétations permettent de relier ce modèle à de nombreux autres, dans des domaines différents : en informatique par exemple, où de nombreux algorithmes utilisent des arbres de recherche assez semblables (L. Devoye, S. Janson, P. Flajolet); en physique théorique, selon les travaux de B. Derrida *et al.*; et bien sûr en biologie, à laquelle ce modèle emprunte son vocabulaire.

Cette composante spatiale peut même être interprétée comme le logarithme de la probabilité d'être sélectionné dans un certain ensemble compact. Ceci forme un lien avec les fractales aléatoires, auto-similaires en loi, introduites par Mandelbrot, au début des années 70 également. Dans ce modèle, on part d'un domaine compact K_0 donné divisé en un certain nombre de sous-domaines de même forme. Pour chacun d'entre eux, on tire une variable aléatoire de Bernoulli indépendante, de paramètre p_i dépendant du sous-domaine. Le compact K_1 est défini comme l'ensemble de ces sous-domaines pour lesquels la variable aléatoire vaut 1. On réalise alors le même découpage pour chacun de ces sous-domaines, et on définit par récurrence une suite de sous-domaines. Le compact $K = \bigcap K_n$ est appelé fractale aléatoire auto-similaire en loi. Ce

modèle, qui permet par exemple d'étudier les lois limites de certains modèles de la physique statistique, a donné naissance à la théorie très féconde des cascades multiplicatives. Les liens entre ces deux domaines sont nombreux, et chaque résultat obtenu dans l'un des deux se traduit automatiquement dans le second.

Nous allons maintenant passer à une définition plus précise d'une marche aléatoire branchante, avec les notations qui y sont liées.

1.1 Définition du modèle

Soit $\mathcal{L} = (l_1, \dots, l_n)$ un processus de points, c'est-à-dire une variable aléatoire donnant un certain nombre n de valeurs réelles (ce nombre n peut également être aléatoire et représente le nombre total d'enfants d'un individu). La marche aléatoire branchante associée à \mathcal{L} est le processus défini comme suit. On commence par un unique individu situé en 0 à l'instant initial. A l'instant 1, cet individu meurt et laisse place à des enfants situés sur les points d'une copie de \mathcal{L} . De même, à chaque instant $n \in \mathbb{N}$, tous les individus de la génération $n - 1$ meurent, en donnant naissance à des individus situés sur les points de copies indépendantes de \mathcal{L} , translatées de la position de leur parent.

Plusieurs variantes de ce modèle existent, comme par exemple le processus CMJ (étudié en particulier par Crump, Mode et Jagers) : dans ce cas, le processus de points \mathcal{L} est à valeurs dans \mathbb{R}^+ , et représente non plus la position de l'individu, mais l'instant auquel celui-ci naît. En d'autres termes le processus \mathcal{L} associé à un individu donné représente l'âge auquel celui-ci aura des enfants. On s'intéressera dans ce cas non plus aux individus présents à la génération n , mais aux individus présents à l'instant t . Ce processus est un analogue à temps continu du processus de Galton-Watson, avec des complications dues au fait que les générations se superposent, un « oncle » peut être plus jeune que son « neveu ».

Une autre variante est le processus de branchement à plusieurs types, dans lequel chaque individu possède un « type » σ . Un individu donne alors naissance à des enfants qui peuvent être, ou non, du même type que lui. Les individus d'un même type se reproduisent suivant la même loi, mais on peut imaginer des différences de reproduction entre les différents types. L'intérêt de ce modèle est d'étudier les interactions entre individus qui peuvent accélérer la colonisation du milieu.

Nous allons maintenant définir un certain nombre de notations relatives aux arbres qui nous permettront par la suite de décrire en détail le comportement de la marche aléatoire branchante.

1.2 Notations et quantités associées aux arbres marqués

Dans toute la suite nous noterons $\mathbb{N} = \{1, 2, \dots\}$ l'ensemble des entiers strictement positifs. La marche aléatoire branchante peut être vue comme une variable aléatoire à valeur dans l'ensemble des arbres marqués. Nous allons donc préciser un certain nombre de notations relatives aux arbres. Nous posons pour commencer $\mathcal{U} = \bigcup_{n \in \mathbb{N}} \mathbb{N}^n \cup \{\emptyset\}$ l'ensemble des suites finies d'entiers, où la suite de longueur nulle est notée \emptyset .

Dans la définition d'un arbre, une suite finie indique le chemin à parcourir depuis la racine \emptyset de l'arbre pour arriver au sommet auquel elle correspond. Le sommet (u_1, \dots, u_n) représente le

$u_n^{\text{ième}}$ fils du $u_{n-1}^{\text{ième}}$ fils ... du $u_1^{\text{ième}}$ fils de l'individu origine. On définit pour commencer un certain nombre de notations sur l'élément $u = (u_1, \dots, u_n) \in \mathcal{U}$:

- on note $|u| = n$ la longueur de la suite, avec la convention $|\emptyset| = 0$ (c'est la génération de l'individu u) ;
- pour $k \leq n$, on note $u|_k = (u_1, \dots, u_k)$ la restriction de u aux k premiers éléments, avec la convention $u|_0 = \emptyset$ ($u|_k$ est l'ancêtre de la génération k de u) ;
- on note $\Pi u = (u_1, \dots, u_{n-1})$ le prédécesseur de la suite u (Πu est le parent de l'individu u) ;
- pour $v = (v_1, \dots, v_m)$ on note $u.v = (u_1, \dots, u_n, v_1, \dots, v_m)$.

Définition 1.1. Un arbre est un sous-ensemble de \mathcal{U} vérifiant les propriétés suivantes :

- $\emptyset \in T$ (l'individu origine est bien dans l'arbre généalogique) ;
- si $u \in T$ alors $\Pi u \in T$ (le parent d'un individu est dans l'arbre généalogique) ;
- si $u \in T$ alors il existe $A_u \in \mathbb{N} \cup \{0, +\infty\}$ tel que pour tout $v \leq A_u$, on a $u.v \in T$ (on numérote les enfants à partir de 1, sans sauter de numéro).

L'ensemble des arbres est noté \mathbb{T} .

Remarquons au passage que nous autorisons un individu donné à avoir une infinité d'enfants, ou à n'en avoir aucun. Il pourra arriver que des restrictions soient nécessaires. Nous allons maintenant définir la notion d'arbre marqué, qui permet de tenir compte de la position de chaque individu.

Après avoir défini la structure d'arbre généalogique de notre population, nous allons maintenant ajouter la composante spatiale de la marche aléatoire branchante. On définit ainsi un arbre marqué : à chaque sommet x d'un arbre T on associe un réel $V(x)$. L'ensemble $\{(x, V(x)), x \in T\}$, noté également (X, V) est appelé arbre marqué. On note \mathcal{T} l'ensemble des arbres marqués.

Il pourra être utile par la suite de regarder les sous-arbres d'un arbre marqué donné. Pour cela, soit (T, V) un arbre marqué et x un sommet de T , on note (T_x, V_x) le sous-arbre marqué de (T, V) issu de x défini de la façon suivante :

- u est un sommet de T_x si $x.u$ est un sommet de T ;
- $V_x(u) = V(u) - V(x)$.

Nous pouvons alors définir la marche aléatoire branchante de façon formelle de la manière suivante : soit $\mathcal{L} = (l_1, \dots, l_n)$ un processus de points et $(\mathcal{L}^{(u)}, u \in \mathcal{U})$ des copies i.i.d. de \mathcal{L} , on pose A_u le nombre de points de $\mathcal{L}^{(u)}$; et T l'arbre défini par ces $(A_u, u \in \mathcal{U})$, c'est-à-dire :

$$x \in T \iff x = \emptyset \text{ ou } \forall 1 < k \leq |x|, x_k \leq A_{x|_{k-1}}.$$

On pose ensuite $V(\emptyset) = 0$, et, pour tout $x \in T$, tout $u \leq A_x$ correspondant à un enfant de x , on définit $V(x)$ comme la somme de $V(\Pi x)$ et de la $u^{\text{ième}}$ valeur de $\mathcal{L}^{(x)}$. La marche aléatoire branchante est alors définie par (T, V) . Notons que tout sous-arbre de la marche aléatoire branchante issue d'un sommet donné est une copie de la marche aléatoire branchante, indépendante du reste du processus.

Par la suite nous aurons également besoin de distinguer un chemin infini particulier dans un arbre, ce chemin sera appelé une épine. Un arbre marqué avec épine est la donnée d'un arbre marqué (T, V) et d'un élément $u \in \mathbb{N}^{\mathbb{N}}$ tel que pour tout $k \in \mathbb{N}$, $u|_k \in T$. L'ensemble des arbres marqués avec épine est alors noté $\tilde{\mathcal{T}}$.

1.3 Quelques quantités associées au processus de points \mathcal{L}

Après avoir défini quelques quantités relatives aux arbres, nous nous intéressons au processus de points \mathcal{L} . La mesure d'intensité de ce processus est l'unique mesure μ vérifiant pour toute fonction ϕ mesurable positive :

$$\mathbb{E} \left[\sum_{|x|=1} \phi(V(x)) \right] = \int \phi d\mu.$$

La transformée de Laplace de la mesure d'intensité de \mathcal{L} joue un rôle important dans l'étude de la marche aléatoire branchante. Pour tout $\theta \in \mathbb{R}$, on pose :

$$\Lambda(\theta) = \mathbb{E} \left[\sum_{|x|=1} \exp(\theta V(x)) \right] = \int_{\mathbb{R}} \mu(dx) e^{\theta x}$$

et $\kappa(\theta) = \log(\Lambda(\theta))$.

De plus on notera $\Lambda'(\theta) = \mathbb{E} \left[\sum_{|x|=1} V(x) \exp(\theta V(x)) \right]$, même si Λ n'est définie qu'au seul point θ et que cette notation n'est donc pas bien définie. Bien entendu, $\kappa'(\theta)$ désignera alors $\frac{\Lambda'(\theta)}{\Lambda(\theta)}$.

Pour finir, on notera $q \in [0, 1]$ la probabilité que la marche aléatoire branchante issue d'un unique individu à l'instant origine s'éteigne. Étant donné que l'arbre généalogique de ce processus est un arbre de Galton-Watson, c'est la plus petite solution de l'équation $\mathbb{E}(q^n) = q$, où n suit la loi du nombre de points de \mathcal{L} .

Nous allons maintenant indiquer quelques méthodes utilisées dans l'étude des marches aléatoires branchantes, à l'origine de nombreux résultats obtenus sur ce modèle.

2 Méthode de décomposition en épine dorsale et Lemme de regroupement

2.1 Martingales additives exponentielles et propriété de branchement

L'étude de la marche aléatoire branchante passe souvent par l'étude fine de certaines martingales qui lui sont associées, notamment les martingales additives exponentielles, définies comme suit. Pour tout $\theta \in \mathbb{R}$ tel que $\kappa(\theta) < +\infty$, on pose :

$$W_n^{(\theta)} = \sum_{|x|=n} \exp(\theta V(x) - n\kappa(\theta)),$$

la propriété de martingale se vérifiant simplement grâce à l'égalité suivante :

$$W_{n+p}^{(\theta)} = \sum_{|x|=n} \exp(\theta V(x) - n\kappa(\theta)) \sum_{y>x, |y|=n+p} \exp(\theta(V(y) - V(x)) - p\kappa(\theta)), \quad (1)$$

où $y > x$ indique que y est un descendant de x . C'est une application directe de la propriété de branchement : les arbres issus de chacun des individus x à la génération n forment des marches aléatoires branchantes indépendantes et de même loi.

Les martingales ainsi définies sont positives, elles convergent donc presque sûrement vers des limites $W^{(\theta)}$. Biggins a déterminé en 1977 quand ces limites étaient triviales ou non, une preuve simple a par la suite été apportée par Lyons [9], utilisant principalement la décomposition en épine dorsale.

Théorème 2.1. *Soit $\theta \in \mathbb{R}$ tel que $\kappa(\theta) < +\infty$ et $\kappa'(\theta)$ existe et est fini. On pose alors*

$$X^{(\theta)} = \sum_{|x|=1} \exp(\theta V(x)).$$

Les assertions suivantes sont équivalentes :

1. $\mathbb{P}(W^{(\theta)} = 0) < 1$;
2. $\mathbb{P}(W^{(\theta)} = 0) = q$;
3. $\mathbb{E}(W^{(\theta)}) = 1$;
4. $\mathbb{E}[X^{(\theta)} \log^+(X^{(\theta)})] < +\infty$ et $\theta \kappa'(\theta) < \kappa(\theta)$.

L'équivalence des deux premières assertions est une conséquence directe de la propriété de branchement et des résultats standards sur les arbres de Galton-Watson. On pose N_n le nombre d'individus présents à l'instant n , en faisant tendre p vers $+\infty$ dans 1, on obtient :

$$W^{(\theta)} = \sum_{|x|=n} \exp(\theta V(x) - \kappa(\theta)) W_x^{(\theta)},$$

où $(W_x^{(\theta)})$ sont des variables aléatoires i.i.d. de même loi que $W^{(\theta)}$.

Par conséquent, on a :

$$\mathbb{P}(W^{(\theta)} = 0) = \mathbb{E} \left[\mathbb{P}(W^{(\theta)} = 0)^{N_n} \right],$$

et $\mathbb{P}(W^{(\theta)} = 0) \leq \mathbb{P}(N_\infty = 0)$, donc $\mathbb{P}(W^{(\theta)} = 0) = q$ ou 0.

Le détail de ce résultat est obtenu en étudiant la marche aléatoire changée de probabilité par rapport à celle que nous avons définie, avec pour dérivée de Radon-Nikodým $M^{(\theta)}$. Nous allons maintenant donner une description de la marche aléatoire branchante sous la loi

$$\mathbb{Q}_{|\mathcal{F}_n}^{(\theta)} = M_n^{(\theta)} \cdot \mathbb{P}_{|\mathcal{F}_n}.$$

2.2 Changement de probabilité

Dans le but de décrire $\mathbb{Q}^{(\theta)}$, nous allons décrire un processus comme un arbre marqué avec épine. En effet, cette loi s'écrit de façon simple en distinguant une lignée particulière, qui se reproduira différemment de tous les autres individus du processus, de la façon suivante.

Théorème 2.2. *On suppose $\kappa(\theta) < +\infty$.*

On pose $\widehat{\mathcal{L}}^\theta$ un processus de points dont la loi admet pour densité $\sum_{x \in \mathcal{L}} e^{\theta l_x - \kappa(\theta)}$ par rapport à la loi de \mathcal{L} .

Sous $\mathbb{Q}^{(\theta)}$, la marche aléatoire branchante se comporte ainsi :

- *à l'instant 0, il y a un unique individu situé en 0 noté w_0 ;*
- *à chaque instant n , les individus de la génération n meurent en laissant place à un certain nombre d'enfants, les enfants des individus différents de w_n sont distribués selon des copies indépendantes de \mathcal{L} , et les enfants de w_n sont distribués selon une copie de $\widehat{\mathcal{L}}^\theta$;*
- *l'individu w_{n+1} est l'enfant x de w_n avec probabilité $\frac{e^{\theta V(w_n, x)}}{\sum_{u=1}^{A_{w_n}} e^{\theta V(w_n, u)}}$.*

En résumé, sous $\mathbb{Q}^{(\theta)}$, la marche aléatoire branchante est donnée par un individu spécial, l'épine, qui donne un certain nombre d'enfants normaux, plus un enfant spécial, les autres individus se reproduisant normalement.

Ce théorème est la décomposition en épine dorsale de la marche aléatoire branchante. Il permet en particulier de démontrer les résultats les plus anciens de convergence de la position de l'individu le plus à droite. Rappelons également des résultats classiques sur les changements de probabilités, qui permettent d'obtenir entre autres les résultats de Biggins.

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_n))$ un espace de probabilité filtré muni de deux probabilités \mathbb{P} et \mathbb{Q} telles que $\mathbb{Q}|_{\mathcal{F}_n} \ll \mathbb{P}|_{\mathcal{F}_n}$. On pose pour $n \in \mathbb{N}$:

$$\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_n} = X_n \text{ et } X = \limsup_{n \rightarrow +\infty} X_n.$$

La proposition suivante permet de préciser le comportement de (X_n) .

Proposition 2.1. *(X_n) est une \mathbb{P} -martingale positive, donc*

$$X_n \xrightarrow[n \rightarrow +\infty]{} X \text{ } \mathbb{P}\text{-p.s. et } X < +\infty \text{ } \mathbb{P}\text{-p.s.}$$

De plus pour tout $A \in \mathcal{F}$, on a :

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}(X \mathbf{1}_A) + \mathbb{Q}(A \cap \{X = +\infty\}).$$

Ce résultat peut également être utilisé pour obtenir des résultats sur la martingale, car on a de plus la dichotomie suivante.

Proposition 2.2.

$$\mathbb{Q} \ll \mathbb{P} \iff X < +\infty \text{ } \mathbb{Q}\text{-p.s.} \iff \mathbb{E}_{\mathbb{P}}(X) = 1 ;$$

$$\mathbb{Q} \perp \mathbb{P} \iff X = +\infty \text{ } \mathbb{Q}\text{-p.s.} \iff \mathbb{E}_{\mathbb{P}}(X) = 0.$$

2.3 Le Lemme de regroupement

Nous allons maintenant nous intéresser à une conséquence intéressante de la décomposition en épine dorsale de la marche aléatoire branchante, le Lemme de regroupement, qui permet de remplacer l'étude de nombreuses fonctionnelles additives sur les lignées de la marche aléatoire

branchante en des fonctionnelles sur une marche aléatoire simple. De nombreux résultats obtenus dans le cadre des marches aléatoires simples peuvent alors être étendues au cas des marches aléatoires branchantes. De façon heuristique, il suffit de remarquer que lorsqu'on suit une branche au hasard de notre marche, le processus des positions est une marche aléatoire.

Lemme 2.1 (Lemme de regroupement). *Soit $\theta \in \mathbb{R}$ tel que $\kappa(\theta) < +\infty$. Il existe une marche aléatoire (S_n) telle que pour tout $n \in \mathbb{N}$, pour toute fonction mesurable $g : \mathbb{R}^n \rightarrow \mathbb{R}^+$, on a*

$$\mathbb{E} \left[\sum_{|x|=n} g(V(x_1), \dots, V(x_n)) \right] = \mathbb{E} [\exp(\theta S_n - n\kappa(\theta)) g(S_1, \dots, S_n)].$$

Ce lemme peut se prouver par récurrence, ou peut être vu comme une conséquence de la décomposition en épine dorsale. Il est utilisé la plupart du temps avec $\theta = 1$ et un processus de points \mathcal{L} renormalisé vérifiant $\kappa(\theta) = 0$, ce qui en simplifie l'expression.

Nous allons maintenant voir quelques résultats qui ont été pour la plupart obtenus en utilisant des variantes de ces méthodes.

3 La position de l'individu le plus à droite

Afin de s'intéresser à la manière dont la population colonise l'espace, on peut s'intéresser à la position de l'individu le plus à droite dans une marche aléatoire branchante (ou bien sûr la particule la plus à gauche, cela revient au même). Les premiers résultats sur ce problème datent des années 75, mais les derniers raffinements n'ont été obtenus que récemment.

On notera dans toute la suite B_n l'individu le plus à droite dans la marche aléatoire branchante, c'est-à-dire :

$$B_n = \max_{|x|=n} V(x), \text{ avec la convention } \max \emptyset = -\infty.$$

3.1 Vitesse de dispersion vers la droite

Le premier résultat que nous citons ici permet de donner la vitesse de dispersion vers la droite v de la marche aléatoire branchante, plus précisément, nous verrons que $\frac{B_n}{n}$ converge vers une constante v , qui peut être calculée.

Théorème 3.1. *Si*

$$\text{il existe } \phi > 0 \text{ tel que } \kappa(\phi) < +\infty, \tag{2}$$

alors il existe une constante $v \geq 0$ telle que

$$\frac{B_n}{n} \xrightarrow[n \rightarrow +\infty]{} v \text{ p.s.,}$$

sur l'ensemble de non-extinction. De plus $v = \inf_{\theta \geq 0} \left\{ \frac{\kappa(\theta)}{\theta} \right\}$.

Ce théorème peut être étendu pour donner l'asymptotique du nombre d'individu situés à droite de na , pour $a < v$, dans l'échelle exponentielle, en fonction de la transformée de Legendre κ^* de κ , définie par :

$$\kappa^*(a) = \sup_{\theta \geq 0} \{\theta a - \kappa(\theta)\}.$$

Observons en particulier que $v = \sup\{a \in \mathbb{R} | \kappa^*(a) < 0\}$.

Remarque 3.1. Sous l'hypothèse d'intégrabilité (2), κ^* est une fonction convexe semi-continue inférieurement croissante dont la valeur minimale est $-\kappa(0) = -\log \mathbb{E}(N)$, qui est négative, car on suppose l'arbre de Galton-Watson associé à la marche aléatoire branchante sur-critique.

Nous pouvons alors formuler le théorème suivant.

Théorème 3.2. *Si l'hypothèse 2 est vérifiée, on a, pour tout $a < v$:*

$$\frac{1}{n} \log \sum_{|x|=n} \mathbf{1}_{\{V(x) > na\}} \xrightarrow[n \rightarrow +\infty]{} -\kappa^*(a) \text{ p.s.}$$

sur l'ensemble de non-extinction.

De plus, pour tout $a > v$:

$$\frac{1}{n} \log \sum_{|x|=n} \mathbf{1}_{\{V(x) > na\}} \xrightarrow[n \rightarrow +\infty]{} -\infty \text{ p.s.}$$

Remarque 3.2. Lorsque le processus de points choisi est la donnée de N points i.i.d. de loi ν , on a alors :

$$\kappa(\theta) = \log \left(\mathbb{E} \left(\sum_{i=1}^N e^{\theta X_i} \right) \right) = \log N - L^*(\theta),$$

où L^* est la transformée de Cramer de ν .

Le résultat précédent peut alors être vu ainsi : lorsqu'on considère une marche aléatoire de loi ν , la probabilité qu'elle soit au dessus de na à l'étape n est de l'ordre de $e^{-n\Lambda^*(a)}$. Nous nous intéressons alors à la marche aléatoire branchante, à l'étape n on a $N^n = e^{n \log N}$ individus, qui ont suivi des marches aléatoires au moins en partie indépendantes donc des événements dont la probabilité était exponentiellement faible peuvent devenir des événements possibles, et ce tant que $\log N - \Lambda^*(\theta) < 0$. C'est d'ailleurs exactement le résultat que nous obtenons.

Le théorème que nous citons est néanmoins bien plus général puisqu'on ne suppose pas a priori l'indépendance des déplacements à l'intérieur d'une même fratrie.

3.2 Un premier raffinement

Il est naturel de s'intéresser à l'écart entre la prédiction de la position de l'individu le plus à droite nv et sa position réelle. Il faudra s'attendre à des hypothèses d'intégrabilité supplémentaires, car si par exemple il y a toujours au moins un enfant placé exactement à distance v de son parent, alors l'approximation obtenue précédemment est exacte.

Théorème 3.3. *S'il existe $0 < \theta < +\infty$ tel que $\theta v - \kappa(\theta) = 0$, alors*

$$B_n - nv \xrightarrow[n \rightarrow +\infty]{} -\infty \text{ p.s.}$$

sur l'ensemble de non-extinction.

On s'interroge ensuite sur la vitesse de cette divergence, de préciser le second ordre de grandeur de la position de l'individu le plus à droite. Ces résultats nécessitent de plus fortes hypothèses d'intégrabilité pour être obtenues. On peut montrer, sous réserve de bonnes conditions d'intégrabilité, que la différence entre B_n et $n\theta$ est d'ordre logarithmique, et qu'à cet ordre, il existe des fluctuations presque sûres, ainsi qu'une limite en probabilité. On obtient le résultat suivant :

Théorème 3.4. *Supposons qu'il existe $\varepsilon > 0$ tel que $\mathbb{E}(N^{1+\varepsilon}) < +\infty$, $\kappa(\theta + \varepsilon) < +\infty$ et $\kappa(-\varepsilon) < +\infty$, alors on a*

$$-\frac{3}{2\theta} = \liminf \frac{B_n - n\Gamma}{\log n} < \limsup \frac{B_n - n\Gamma}{\log n} = -\frac{1}{2\theta} \text{ p.s.}$$

et

$$\frac{B_n - nv}{\log n} \xrightarrow[n \rightarrow +\infty]{(P)} -\frac{3}{2\theta} \text{ p.s.}$$

sur l'ensemble de non-extinction.

Ce résultat peut être obtenu grâce au Lemme de regroupement, en utilisant des estimations finies sur les marches aléatoires, comme dans l'article d'Aïdekon et Shi [2].

3.3 Dernières avancées

Pour finir, les derniers résultats sur l'individu le plus à droite donne également une limite en loi pour la position de l'individu le plus à droite à laquelle on soustrait le présent équivalent. Ce résultat est prouvé dans [1]. Ce théorème est obtenu avec des hypothèses d'intégrabilité moins fortes que le précédent, mais en supposant que la marche aléatoire branchante considérée est critique, c'est-à-dire qui vérifie les conditions suivantes :

$$\kappa(0) > 0, \quad \kappa(-1) = 0 \text{ et } \kappa'(-1) = 0.$$

Il est alors plus naturel de regarder le symétrique du processus, ainsi que la position de l'individu le plus à gauche, que l'on note :

$$M_n = \min_{|x|=n} V(x),$$

qui tend vers $+\infty$ presque sûrement, grâce au Théorème 3.3. Ce théorème reste assez général. En effet, sous des hypothèses suffisantes, une marche aléatoire branchante peut être transformée en marche aléatoire branchante critique, par une transformation affine du processus de points \mathcal{L} .

Nous introduisons la martingale dérivée, définie pour tout $n \geq 0$ par

$$\partial W_n = \sum_{|x|=n} V(x) e^{-V(x)},$$

cette martingale converge presque sûrement vers ∂W_∞ , qui est strictement positive sur l'ensemble de non-extinction de la marche aléatoire branchante, sous des hypothèses suffisantes.

Théorème 3.5. *On suppose que la distribution de \mathcal{L} n'est pas à valeurs dans un réseau, on introduit les variables aléatoires suivantes :*

$$X = \sum_{|x|=1} e^{-V(x)}, \quad \tilde{X} = \sum_{|x|=1} V(x)_+ e^{-V(x)} \quad \text{et} \quad \hat{X} = \sum_{x=1} V(x)^2 e^{-V(x)},$$

et on suppose :

$$\mathbb{E}(X(\ln_+ X)^2) < +\infty, \quad \mathbb{E}(\tilde{X} \ln_+ \tilde{X}) < +\infty, \quad \text{et} \quad \mathbb{E}(\hat{X}) < +\infty.$$

Il existe une constante $c > 0$ telle que pour tout $x \in \mathbb{R}$:

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left(M_n \geq \frac{3}{2} \log n + x \right) = \mathbb{E}(e^{-ce^x \partial W_\infty}).$$

Après ces résultats bien connus, nous allons maintenant introduire une variante importante des marches aléatoires branchantes, qui permet de tenir compte de la sélection naturelle dans les processus de population.

4 Marche aléatoire branchante avec sélection

Dans la plupart des modèles probabilistes de dynamique des populations, on a une dichotomie assez gênante extinction-divergence à vitesse exponentielle vers $+\infty$. Autrement dit, une population ne peut atteindre un équilibre simple, comme ceci est observé dans la nature. Ces observations nous permettent donc de supposer qu'une composante essentielle manque au modèle. Celle-ci pourrait être la sélection naturelle, qui peut être représentée de manière assez simple dans le formalisme de la marche aléatoire branchante. La position V de l'individu, représente alors son degré d'adaptation au milieu, et selon la logique darwinienne, les plus adaptés survivent.

On fixe un entier N représentant la capacité d'accueil du milieu, le nombre d'individus qui peuvent se nourrir. La marche aléatoire branchante avec sélection évolue alors comme une marche aléatoire branchante classique, avec la règle supplémentaire suivante : si à n'importe quel instant donné il y a plus de N individus dans la population, seuls les N plus à droite survivent, les autres sont immédiatement tués sans descendance.

Afin d'éviter l'extinction presque sûre du processus, on supposera $\mathbb{P}(\mathcal{L} = \emptyset) = 0$. Le but devient alors d'étudier la position de l'individu le plus à droite dans ce processus modifié, en fonction du nombre N d'individu conservés. Un couplage trivial montre que $v^N \geq v^{N+1} \geq v$, l'individu le plus à droite d'un processus très sélectif est toujours moins à droite que si on avait gardé d'avantage d'individus, puisqu'on a ainsi éliminé un certain nombre de lignées qui pourraient contenir le « vrai » élément le plus à droite.

Une conjecture de Derrida et Brunet donne la vitesse de convergence de v^N vers v . Cette conjecture prétend que :

$$v^N - v = \frac{c_1}{(\log n)^2} + \frac{c_2 + o(1)}{\log \log n},$$

Bérard et Guéré (c.f. [7]) ont récemment prouvé que l'équivalent de cette suite était bien de cette forme, avec une constante c_1 qui dépend du modèle. On conjecture que la constante c_2 est en revanche universelle.

Ces résultats ont de forts intérêts pratiques : en informatique notamment, ils mesurent l'erreur faite lorsqu'on remplace un arbre de recherche (à croissance exponentielle) par les N meilleures branches à chaque étape, soit une exploration bien plus rapide.

L'intérêt est également évident pour le modèle biologique, bien que l'hypothèse de capacité d'accueil du milieu constante soit assez restrictive. Des modifications du modèle utilisant une capacité d'accueil $\phi(n)$ dépendant du temps, voire de la population présente à l'instant précédent, pourraient être intéressantes à étudier également. Un certain nombre de questions classiques en biologie peuvent alors se poser encore : comment se comporte l'arbre généalogique lorsqu'on remonte dans le temps ? Quel est l'âge du plus récent ancêtre commun ? Existe-t-il des modèles limites d'intérêt dans cette étude ?

Références

- [1] E. Aïdékon. Convergence in law of the minimum of a branching random walk. *arXiv :1101.1810*, 2011.
- [2] Elie Aïdékon and Zhan Shi. Weak convergence for the minimal position in a branching random walk : a simple proof. *Periodica Mathematica Hungarica*, 61(1-2) :43–54, 2010.
- [3] K.B. Athreya and P.E. Ney. *Branching Processes*. Grundlehren der mathematischen Wissenschaften. Springer, 1972.
- [4] J.D. Biggins. Lindley-type equations in the branching random walk. *Stochastic Processes and Applications*, 75(1) :105–133, 1998.
- [5] J.D. Biggins. Branching out. *arXiv :1003.4715*, 2010.
- [6] J.D. Biggins and A.E. Kyprianou. Fixed points of the smoothing transform : the boudary case. *Electronic Journal of Probability*, 10(17) :609–631, 2005.
- [7] Jean Bérard and Jean-Baptiste Gouéré. Brunet-derrida behavior of branching-selection particle systems on the line. *arXiv :0811.2782*, 2010.
- [8] Yueyun Hu and Zhan Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *The Annals of Probability*, 37(2) :742–789, 2009.
- [9] Russel Lyons. A simple path to biggins' martingale convergence for branching random walk. *Classical and Modern Branching Processes*, pages 217–221, 1997.