



# Sur la destruction de grands arbres aléatoires récursifs

Erich Baur, Jean Bertoin

► **To cite this version:**

Erich Baur, Jean Bertoin. Sur la destruction de grands arbres aléatoires récursifs. 2016. <hal-01354254>

**HAL Id: hal-01354254**

**<https://hal.archives-ouvertes.fr/hal-01354254>**

Submitted on 18 Aug 2016

**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.

# Sur la destruction de grands arbres aléatoires récurrents

Erich Baur\* et Jean Bertoin†  
ENS Lyon et Universität Zürich

## Résumé

L'objet de ce texte est de présenter aussi simplement que possible les principaux résultats d'un programme récent portant sur l'étude des différents régimes observés lors de la destruction d'un arbre aléatoire uniforme en supprimant ses arêtes les une après les autres, uniformément au hasard.

## 1 Arbres aléatoires récurrents et modèles afférents

Comme son nom l'indique, un *arbre récurrent*, au sens de la théorie des graphes, est un arbre, c'est-à-dire un graphe connexe et sans cycles, qu'on peut construire de façon récurrente en incorporant les sommets les uns après les autres. Autrement dit, on part d'un ensemble totalement ordonné, disons  $\{0, 1, \dots, n\}$ , le sommet 0 étant en quelque sorte distingué et pouvant être vu comme la racine. Pour chaque  $j = 1, \dots, n$ , on relie le sommet  $j$  par une arête à l'un des sommets précédents  $0, 1, \dots, j - 1$ . Une propriété immédiate d'une telle construction est que la suite des sommets le long d'une branche quelconque de l'arbre qui part de la racine, est toujours croissante. Pour cette raison, les arbres récurrents sont parfois nommés également *arbres croissants*, voir [6].

On parle d'*arbre aléatoire récurrent uniforme*, souvent en omettant l'adjectif uniforme, pour désigner un arbre  $T_n$  choisi uniformément au hasard parmi les  $n!$  arbres récurrents ayant  $n$  sommets, la racine non comprise. Un tel arbre aléatoire est construit par un algorithme stochastique très simple : le sommet auquel  $j$  est rattaché est choisi uniformément au hasard parmi les précédents, et ce, indépendamment des arêtes déjà créées. On peut ainsi décrire la construction en termes d'urnes : on considère  $n$  boules numérotées de 1 à  $n$ , et une urne qui contient initialement une seule boule numérotée 0. À la  $j$ -ème étape, on tire une boule uniformément au hasard parmi les  $j$  boules déjà dans l'urne, on note son numéro, disons  $i$ . On la remet ensuite dans l'urne ainsi que la boule numérotée  $j$ , et on représente ce tirage par une arête  $(i, j)$ . Quand les  $n$  boules ont été mises dans l'urne, l'arbre ainsi construit est une version de  $T_n$ .

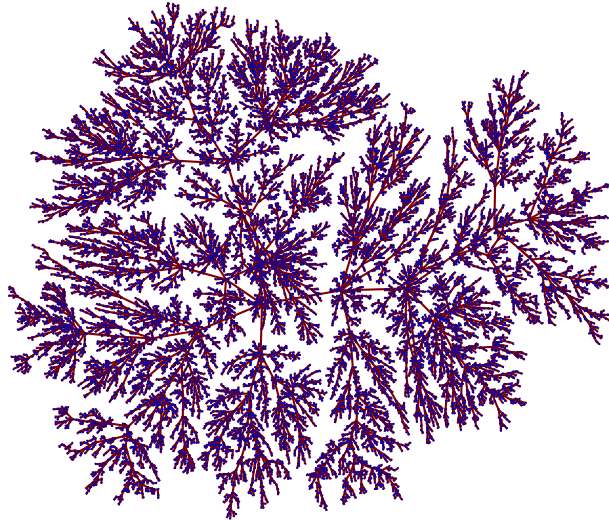
La figure 1, réalisée par Igor Kortchemski, est une simulation de  $T_{15000}$ . Soulignons la distorsion nécessaire à l'obtention d'une représentation plane qui induit un aspect touffu de l'arbre. Certaines arêtes sont beaucoup plus longues que d'autres afin de permettre le plongement du graphe dans le

---

\*erich.baur@ens-lyon.fr

†jean.bertoin@math.uzh.ch

plan. Si la structure combinatoire est respectée, la géométrie quant à elle ne l'est pas : par exemple, on sait que la plupart des sommets d'un arbre aléatoire récursif de taille  $n \gg 1$  sont à distance environ  $\ln n$  de la racine (voir e.g. [9]), et deux sommets typiques sont à distance d'environ  $2 \ln n$  l'un de l'autre, ce qui n'est visiblement pas le cas sur la figure.



Représentation planaire d'un arbre aléatoire récursif de taille 15000 (Igor Kortchemski)

Les arbres aléatoires récursifs forment l'une des familles les plus simples et les plus naturelles d'arbres aléatoires ; voir Drmota [10]. Ils sont intimement liés à d'autres processus stochastiques également bien connus. Tout d'abord, la construction récursive est essentiellement une variante de celle du *restaurant chinois* qui permet de simuler une permutation aléatoire uniforme, et dont nous rappelons brièvement la dynamique (voir Pitman [21]). Imaginons un restaurant ayant un nombre infini de tables, chacune de capacité infinie. Au temps initial, le restaurant est vide ; un premier client entre et prend place à une première table. Lorsque le  $j$ -ème client arrive, il s'assoit à la droite du  $i$ -ème client avec probabilité  $1/j$  pour chaque  $i = 1, \dots, j - 1$  (on crée alors une arête reliant  $i$  à  $j$ ), et s'assoit seul à une nouvelle table avec probabilité  $1/j$  (on crée alors une arête reliant la racine 0 à  $j$ ). Les tables du restaurant correspondent aux cycles de la permutation aléatoire, et peuvent être interprétées comme les sous-arbres obtenus en supprimant la racine 0 d'un arbre aléatoire récursif.

On peut également voir les arbres aléatoires récursifs comme ceux décrivant la généalogie d'un *processus de Yule*. Rappelons tout d'abord qu'un processus de Yule est un modèle élémentaire d'évolution de populations qui s'accroissent au cours du temps, pour lequel chaque individu donne naissance à un enfant à taux 1 (c'est-à-dire après un temps aléatoire qui suit une loi exponentielle de moyenne 1), indépendamment des autres individus. Si on attribue le numéro 0 à l'ancêtre et énumère les enfants dans l'ordre de leur naissance, l'arbre généalogique des  $n$  premiers enfants est un arbre aléatoire récursif sur  $\{0, 1, \dots, n\}$ . Cette interprétation généalogique est particulièrement importante, de nombreuses propriétés des arbres aléatoires récursifs trouvant alors des justifications très simples en termes

de processus de branchement. Il est important de noter que la structure des processus de Yule est plus riche que celle des arbres aléatoires récursifs, car l'arbre généalogique ne permet pas de retrouver les temps de naissance des enfants. Cette information additionnelle contenue dans un processus de Yule peut s'avérer très utile pour l'étude des arbres récursifs.

Nous nous intéresserons dans ce texte aux effets d'un algorithme de destruction sur des arbres aléatoires récursifs de grandes tailles, notamment en identifiant différents régimes de percolation remarquables. Commençons par rappeler l'origine de ces questions.

## 2 Aux origines de la destruction : isolation de la racine

Au début des années 1970, Meir et Moon [18, 19] se sont intéressés à un processus de destruction d'arbres par suppression de leurs arêtes les unes après les autres, uniformément au hasard. Ils ont étudié plus particulièrement le nombre  $X_n$  d'arêtes appartenant au sous-arbre contenant la racine qu'on doit retirer jusqu'à ce que celle-ci soit isolée. Autrement dit, dans le processus de destruction, on ne compte que les arêtes qui jouent un rôle dans l'isolation de la racine, et on ignore celles qui, au moment où elles sont retirées, appartenaient à un sous-arbre déjà déconnecté de la racine.

Meir et Moon ont déterminé le comportement asymptotique au premier ordre de  $X_n$  :

$$\lim_{n \rightarrow \infty} \frac{\ln n}{n} X_n = 1 \quad \text{en probabilité.} \quad (1)$$

Pour cela, ils ont d'abord noté la *propriété fractale* suivante des arbres aléatoires récursifs. Si on retire une arête de  $T_n$  uniformément au hasard, disons la  $j$ -ème avec  $1 \leq j \leq n$ , on obtient deux sous-arbres aléatoires, l'un contenant la racine 0, et l'autre le sommet  $j$  qu'on peut voir comme la racine de ce second sous-arbre. Alors conditionnellement à leurs tailles respectives, ces deux sous-arbres sont indépendants et ont eux-mêmes une structure d'arbres aléatoires récursifs. Cette observation élémentaire est très utile car elle permet évidemment d'itérer le processus. Pour obtenir des informations quantitatives, on a besoin de connaître la loi des tailles des deux sous-arbres, et Meir et Moon ont montré que la loi de la taille du sous-arbre qui ne contient pas la racine est particulièrement simple : c'est la même que celle d'une variable  $\xi$  conditionnée à être au plus égale à  $n$ , où

$$\mathbb{P}(\xi = k) = \frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}, \quad k \geq 1. \quad (2)$$

À l'aide de ces deux résultats, Meir et Moon obtiennent des équations de récurrence pour les premiers et seconds moments de  $X_n$ , desquelles l'estimation au premier ordre (1) découle facilement.

Le problème de caractériser les fluctuations de  $X_n$ , c'est-à-dire de l'approximation au second ordre, est resté ouvert pendant près de 35 ans avant que Drmota, Iksanov, Möhle et Rösler [11] ne parviennent à établir leur caractère non-gaussien. Plus précisément, ils ont montré que

$$\lim_{n \rightarrow \infty} \left( \frac{\ln^2 n}{n} X_n - \ln n - \ln \ln n \right) = X \quad \text{en distribution,} \quad (3)$$

où  $X$  désigne une variable de Cauchy complètement asymétrique ayant pour fonction caractéristique

$$\mathbb{E}(\exp(itX)) = \exp\left(it \ln |t| - \frac{\pi}{2}|t|\right), \quad t \in \mathbb{R}. \quad (4)$$

C'est un résultat *a priori* plutôt surprenant, car on aurait pu s'attendre à ce que  $X_n - n/\ln n$  convenablement renormalisé converge en loi vers une variable gaussienne comme c'est très souvent le cas. Non seulement il n'en est rien, mais on peut aussi souligner qu'en plus du recentrage par  $n/\ln n$  auquel on s'attend, il est nécessaire de faire une seconde correction déterministe en  $n(\ln \ln n)/\ln^2 n$  avant de renormaliser. L'approche initiale de Drmota et al. s'inscrit dans la lignée des méthodes de combinatoire analytique développées notamment par Philippe Flajolet : on utilise la propriété fractale des arbres aléatoires récursifs pour obtenir une équation sur la fonction génératrice  $\varphi(n, s) = \sum_{n=0}^{\infty} \mathbb{E}(s^{X_n})$  de  $X_n$ , puis des techniques d'analyse complexe pour étudier le comportement asymptotique de  $\varphi(n, \cdot)$  quand  $n \rightarrow \infty$ .

Iksanov et Möhle [16] ont proposé peu de temps après une élégante explication probabiliste, qui nécessite très peu de calculs. Elle repose sur le couplage suivant, qui est une conséquence facile de la propriété fractale. Introduisons une suite  $(\xi_j : j \in \mathbb{N})$  de variables aléatoires indépendantes et toutes distribuées selon la loi (2). On considère d'abord la marche aléatoire  $S_k = \xi_1 + \dots + \xi_k$ , pour  $k \geq 0$ . Puis, pour  $n$  fixé, on modifie  $S$  en supprimant tous les pas qui lui feraient dépasser le niveau  $n$ . Plus précisément, on pose  $S_0^* = 0$  et  $k_0 = 0$ , et on définit par itération  $S_1^* = S_0^* + \xi_{k_1}$  avec  $k_1 = \inf\{j \geq k_0 : S_0^* + \xi_j \leq n\}$ ,  $S_2^* = S_1^* + \xi_{k_2}$  avec  $k_2 = \inf\{j \geq k_1 : S_1^* + \xi_j \leq n\}$ , ... L'algorithme se termine lorsque  $S^*$  atteint la valeur  $n$ . Iksanov et Möhle ont observé que la marche modifiée  $S^*$  a la même loi que la suite du nombre de sommets qui sont déconnectés de la racine, lorsqu'on supprime une à une les arêtes du sous-arbre qui contient la racine, et cela uniformément au hasard. Ainsi, la loi de  $X_n$  est la même que celle du nombre de pas de la marche modifiée  $S^*$ , et on peut établir que cette dernière est proche quand  $n$  est grand, du premier temps de passage de la marche aléatoire  $S$  au-dessus du niveau  $n$ . La formule (1) montre que la marche aléatoire  $S$  appartient au domaine d'attraction d'un processus de Cauchy asymétrique, et le théorème limite (3) découle enfin de résultats classiques sur les premiers temps de passages de marches aléatoires positives.

Pour conclure cette section, mentionnons encore que le problème initial d'isolation de la racine a été par la suite étendu dans plusieurs directions ; voir notamment l'article récent de Kuba et Panholzer [17], ainsi que [3] et les références qui y sont citées.

### 3 Régimes de percolation

Lorsqu'on travaille avec le graphe complet  $G_n$  avec  $n$  sommets (et donc  $n(n-1)/2$  arêtes) au lieu de l'arbre récursif  $T_n$ , le processus de destruction doit être vu comme une version dynamique du modèle du *graphe aléatoire* introduit par Erdős et Rényi [12, 13] et Gilbert [14]. Rappelons que pour  $0 < p < 1$ ,  $G(n, p)$  désigne le graphe aléatoire obtenu à partir de  $G_n$  après une percolation de paramètre  $p$ , c'est-à-dire en conservant chaque arête avec probabilité  $p$  et en la supprimant avec

probabilité  $1 - p$ , indépendamment des autres arêtes. Parmi les résultats asymptotiques classiques dans ce domaine quand  $n \rightarrow \infty$ , l'un des plus importants concerne les régimes où le paramètre de percolation  $p = p(n)$  dépend de  $n$  avec  $p(n) \sim c/n$ , pour une constante  $c \in (0, \infty)$  :

- Lorsque  $c < 1$ , avec grande probabilité pour  $n \gg 1$ , la plus grande composante connexe de  $G(n, p)$  est de l'ordre de  $\ln n$  ; on parle du régime *sous-critique*.
- Lorsque  $c > 1$ , avec grande probabilité pour  $n \gg 1$ , il existe une unique composante connexe géante, plus précisément de taille proche de  $\theta(c)n$ , où  $\theta(c) \in (0, 1)$  est l'unique solution de l'équation  $x + e^{-cx} = 1$ , et les composantes connexes suivantes sont d'ordre  $\ln n$  seulement ; on parle du régime *sur-critique*.
- Une transition de phase a donc lieu pour  $c = 1$ , et c'est dans la *fenêtre critique*

$$p(n) = n^{-1} + tn^{-4/3}(1 + o(1)),$$

avec  $t \in \mathbb{R}$ , que la composante connexe géante est en train de se former, les composantes connexes les plus grandes étant alors de l'ordre de  $n^{2/3}$ .

Les résultats asymptotiques présentés dans la section précédente pour la destruction de grands arbres récursifs aléatoires concernent seulement le sous-arbre qui contient la racine, et il est naturel de s'intéresser plus généralement à l'évolution des tailles de tous les sous-arbres lors du processus de destruction. Comme dans le cadre du graphe aléatoire, on peut distinguer trois régimes principaux que nous allons décrire plus en détails dans la section suivante.

Lorsqu'on interprète un arbre aléatoire récursif comme l'arbre généalogique d'un processus de Yule, la percolation peut être vue comme modélisant des mutations<sup>1</sup> neutres (c'est-à-dire sans effets sur la reproduction), qu'on superpose à la généalogie. Ceci conduit à deux points de vue conceptuellement assez différents. On peut soit construire d'abord l'arbre aléatoire récursif de taille  $n$ ,  $T_n$ , puis ensuite effectuer une percolation de paramètre  $p = p(n)$ , soit réaliser ces deux opérations simultanément : à chaque étape où l'on incorpore un nouveau sommet, on décide avec probabilité  $p$  de rattacher ce sommet à un des précédents choisi uniformément au hasard, et avec probabilité  $1 - p$  de ne créer aucune arête. Ces deux points de vue sont complémentaires ; le premier est mieux adapté à une approche dynamique de la percolation, notamment grâce au couplage d'Iksanov et Möhle, le second, plus robuste dans le sens où des arguments analogues restent valables pour d'autres familles d'arbres aléatoires, permet de mieux utiliser la propriété de branchement.

### 3.1 Régime sur-critique

Comme nous l'avons rappelé précédemment, la hauteur d'un sommet typique de  $T_n$ , c'est-à-dire la distance à la racine d'un sommet choisi uniformément au hasard, est proche de  $\ln n$  avec grande probabilité lorsque  $n \gg 1$ . Il s'ensuit que si le paramètre de percolation  $p = p(n)$  est tel que

$$1 - p(n) \sim c/n, \quad \text{pour un certain } c > 0, \quad (5)$$

---

1. Il est intéressant de signaler qu'à l'origine, Yule s'était intéressé précisément à un modèle d'évolution de populations avec mutations ; voir [23].

alors la probabilité qu'un sommet typique reste connecté à la racine est proche de

$$p(n)^{\ln n} \sim (1 - c/\ln n)^{\ln n} \sim e^{-c}.$$

De même, la probabilité que deux sommets choisis uniformément au hasard soient toujours connectés après percolation est voisine de  $p(n)^{2\ln n} \sim e^{-2c}$ . Des calculs simples de premier et deuxième moments montrent alors qu'avec grande probabilité pour  $n \gg 1$ , la taille du sous-arbre contenant la racine<sup>2</sup>,  $C_0$ , reste alors de l'ordre de  $n$  :

$$C_0 \sim e^{-c}n. \quad (6)$$

La composante qui contient la racine est donc géante, et il est facile de vérifier que c'est la seule. Plus précisément, si on note  $C_1 \geq C_2 \geq \dots$  la suite ordonnée des tailles des sous-arbres ne contenant pas la racine, alors il est établi dans [7] que dans le régime (5), pour tout  $j \geq 1$ , le vecteur aléatoire

$$\left( \frac{\ln n}{n} C_1, \dots, \frac{\ln n}{n} C_j \right)$$

converge en loi vers  $(x_1, \dots, x_j)$ , où  $x_1 > x_2 > \dots$  désigne la suite des atomes d'une mesure aléatoire de Poisson sur  $(0, \infty)$  avec intensité  $ce^{-c}x^{-2}dx$ . La loi limite peut également être décrite de la façon suivante :  $1/x_1, 1/x_2 - 1/x_1, \dots, 1/x_j - 1/x_{j-1}$  forment une suite de  $j$  variables i.i.d. de loi exponentielle de paramètre  $ce^{-c}$ .

On notera que la taille de ces sous-arbres est de l'ordre de  $n/\ln n$ . De façon informelle, les plus grandes composantes qui ne contiennent pas la racine sont donc *presque géantes*, ce qui contraste nettement avec le cas du graphe aléatoire. Le couplage d'Iksanov et Möhle qui a été décrit à la fin de la section précédente, joue un rôle fondamental pour les preuves.

Ce résultat a été étendu dans [2] à tout le régime sur-critique

$$1/n \ll 1 - p(n) \ll 1. \quad (7)$$

La borne inférieure  $1/n$  est naturelle, car pour  $1 - p(n) = o(1/n)$ , avec grande probabilité, aucune arête n'est supprimée. Comme nous l'avons indiqué plus haut, le sous-arbre contenant la racine est de taille  $C_0 \sim n^{p(n)}$ , et pour tout  $j \geq 1$ , la suite des  $j$  sous-arbres suivants les plus grands convenablement normalisés

$$\left( \frac{n^{-p(n)}}{1 - p(n)} C_1, \dots, \frac{n^{-p(n)}}{1 - p(n)} C_j \right)$$

converge en loi quand  $n \rightarrow \infty$  vers

$$(y_1, \dots, y_j)$$

où  $y_1 > y_2 > \dots$  désigne la suite des atomes d'une mesure aléatoire de Poisson sur  $(0, \infty)$  avec intensité  $y^{-2}dy$ . On observera que cette mesure est l'intensité des sauts d'un processus de Cauchy totalement asymétrique, ce qui est bien sûr très lié à (3). On notera également que dans tout le régime

---

2. Pour alléger la notation, la dépendance en la taille  $n$  de l'arbre sera systématiquement omise dans la notation.

sur-critique (7), avec grande probabilité quand  $n \gg 1$ , le sous-arbre contenant la racine est le plus gros et les suivants sont plus petits par un facteur  $1 - p(n)$ .

Avant de nous pencher sur les autres régimes, revenons brièvement sur (5). L'estimation du premier ordre (6) peut être vue comme une loi des grands nombres ; les fluctuations, autrement dit l'estimation au second ordre, sont données comme suit. Dans le régime (5), lorsque  $n \rightarrow \infty$ , la convergence en loi suivante a lieu :

$$(n^{-1}C_0 - e^{-c}) \ln n - ce^{-c} \ln \ln n \implies -ce^{-c} (\ln c - X) ,$$

où la variable  $X$  suit une loi de Cauchy complètement asymétrique dont la fonction caractéristique est donnée par (4). Ce résultat a été établi initialement par Schweinsberg [22], dans une étude du coalescent de Bolthausen et Sznitman (des liens profonds entre ce dernier et les arbres aléatoires récursifs avaient été observés en premier par Goldschmidt et Martin [15]) ; une preuve plus simple en a ensuite été donnée dans [8], en s'appuyant à la fois sur le couplage d'Iksanov et Möhle et les liens avec les processus de Yule. Les fluctuations non-gaussiennes de  $C_0$  rappellent bien sûr celles de (3) relatives à l'isolation de la racine, même si il n'y a probablement pas de façon simple de rapprocher ces deux résultats. Elles contrastent nettement avec les fluctuations pour la composante géante du graphe aléatoire dans le régime sur-critique, qui elles sont gaussiennes.

### 3.2 Régime sous-critique

Nous avons vu que le régime sur-critique était donné par (7). Il n'est alors pas surprenant que le régime sous-critique, qui a été étudié très récemment dans [5], corresponde à

$$1/n \ll p(n) \ll 1.$$

À nouveau, la borne inférieure  $1/n$  s'explique par le fait que pour  $p(n) = o(1/n)$ , avec grande probabilité, toutes les arêtes sont supprimées. Nous allons nous intéresser aux effets de la percolation dans certains sous-régimes naturels plus spécifiques.

L'une des premières questions concerne le seuil pour lequel les plus grandes composantes connexes ont une taille donnée. La réponse est simple : pour tous  $a > 0$  et  $\ell \in \mathbb{N}$  fixés, considérons le régime

$$p(n) \sim an^{-1/\ell}. \tag{8}$$

Alors avec grande probabilité lorsque  $n \gg 1$ , il n'existe pas de sous-arbres de taille strictement plus grande que  $\ell + 1$ , et la probabilité de trouver exactement  $k$  sous-arbres de taille  $\ell + 1$  converge vers

$$\exp(-\ell!a^\ell) \frac{(\ell!a^\ell)^k}{k!}, \quad k = 0, 1, \dots$$

Il est intéressant de rappeler que dans un arbre récursif de taille  $n$ , le degré de la racine est voisin de  $\ln n$  avec grande probabilité quand  $n \gg 1$ . En conséquence, dans le régime (8), toutes les arêtes contiguës à la racine sont supprimées et la racine est isolée, avec grande probabilité. Il est dès lors



naturel de considérer également le régime pour lequel la racine cesse d'être isolée, c'est-à-dire

$$p(n) \sim a/\ln n, \tag{9}$$

où  $a > 0$  est une constante. Il est facile de montrer, par exemple en utilisant les liens avec les processus de Yule, qu'asymptotiquement, le sous-arbre contenant la racine suit une loi géométrique de paramètre  $e^{-a}$ . Le plus grand sous-arbre a quant à lui une taille de l'ordre de  $\ln n$ . Plus précisément, il est établi dans [5] que dans le régime (9), avec grande probabilité pour  $n \gg 1$ , le plus grand sous-arbre a pour taille

$$t^* \ln n - \frac{\ln \ln n}{2(\ln(at^*) - \ln(1 + at^*))} + O(1)$$

où  $t^* > 0$  désigne l'unique solution de l'équation

$$1 + t^* \ln(at^*) - \frac{1 + at^*}{a} \ln(1 + at^*) = 0.$$

### 3.3 Fenêtre critique

Nous venons de voir qu'avec grande probabilité quand  $n \gg 1$ , le sous-arbre qui contient la racine est le plus grand de tous dans le régime de percolation sur-critique où  $p(n)$  converge vers 1, alors que quand  $p(n)$  converge vers 0, ce n'est plus le cas. Nous nous intéressons maintenant à la transition de phase, correspondant au cas où le paramètre  $p$  ne dépend pas de  $n$ . Plus précisément, nous considérons une percolation dynamique en assignant à chaque arête une variable aléatoire de loi exponentielle qui correspond au temps en lequel cette arête est supprimée, et ce bien sûr, indépendamment des autres arêtes. Ainsi, au temps  $t \geq 0$ , nous obtenons une percolation de Bernoulli, avec paramètre  $p = e^{-t}$ .

Rappelons que  $C_0$  désigne la taille du sous-arbre qui contient la racine, et  $C_1 \geq C_2 \geq \dots$  la suite ordonnée des tailles des autres composantes. Les résultats suivants ont été obtenus dans [4].

Tout d'abord, il est facile de montrer qu'au temps  $t$ ,  $C_0$  est de l'ordre de  $n^p = n^{e^{-t}}$ . Plus précisément, on a

$$\lim_{n \rightarrow \infty} n^{-e^{-t}} C_0 = X_0(t) \quad \text{en probabilité, pour tout } t \geq 0,$$

où  $(X_0(t))_{t \geq 0}$  est un processus dit de Mittag-Leffler. Ce dernier a notamment pour loi marginale

$$\mathbb{P}(X_0(t) \in dx)/dx = \frac{e^t}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{k!} \Gamma(ke^{-t} + 1) x^{k-1} \sin(\pi ke^{-t});$$

et a été étudié indépendamment par Möhle [20].

Plus généralement, pour tout  $i = 0, 1, 2, \dots$ , on a

$$\lim_{n \rightarrow \infty} n^{-e^{-t}} C_i = X_i(t) \quad \text{en probabilité,} \tag{10}$$

où  $(X_0(t), X_1(t), \dots)$  est une famille de processus dont nous allons maintenant décrire la loi, en notant

tout d'abord  $Y_i(t) = \ln X_i(t)$ , puis en introduisant la famille

$$\mathcal{Y}_t = \{Y_i(t) : i \geq 0\}.$$

Le processus  $\mathcal{Y} = (\mathcal{Y}_t : t \geq 0)$  peut alors être vu comme un *processus d'Ornstein-Uhlenbeck branchant*. C'est un processus de Markov tel que la loi conditionnelle de  $\mathcal{Y}_{s+t}$  sachant  $\mathcal{Y}_s = \{y_i : i \geq 0\}$  est la même que celle de

$$\{e^{-t}y_i + \zeta_j^{(i)} : i \geq 0, j \geq 0\},$$

où les familles  $\{\zeta_j^{(i)} : j \geq 0\}$  sont indépendantes et toutes de même loi que  $\mathcal{Y}_t$ . À un retournement du temps près, il s'agit de l'analogie de la description en termes de processus de coalescence due à Aldous [1], de l'évolution du graphe aléatoire dans la fenêtre critique.

On peut ainsi voir  $\mathcal{X}_t = \{X_i(t) : i \geq 0\}$ ,  $t \geq 0$ , comme un exemple remarquable de processus de *croissance-fragmentation*. Le terme de croissance peut être surprenant au premier abord (puisqu'après tout on considère une dynamique purement de destruction), il s'explique par le fait que la normalisation (10) dépend du temps  $t$ .

## Références

- [1] Aldous, D. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** (1997), 812-854.
- [2] Baur, E. Percolation on random recursive trees. *Random Structures Algorithms* **48-4** (2016), 655-680.
- [3] Baur, E. et Bertoin, J. Cutting edges at random in large recursive trees. Dans : *Stochastic Analysis and Applications*, Springer Proc. Math. Stat., **100**, (2014), pp. 51-76.
- [4] Baur, E. et Bertoin, J. The fragmentation process of an infinite recursive tree and Ornstein-Uhlenbeck type processes. *Electron. J. Probab.* **20-98** (2015), 1-20.
- [5] Baur, E. et Bertoin, J. Weak limits for the largest subpopulations in Yule processes with high mutation probabilities. ArXiv : 1603.06564.
- [6] Bergeron, F., Flajolet, Ph., et Salvy, B. Varieties of increasing trees. Dans : *CAAP '92 (Rennes, 1992)*, Lecture Notes in Comput. Sci. **581**, pp. 24-48, Springer, Berlin, 1992.
- [7] Bertoin, J. Sizes of the largest clusters for supercritical percolation on random recursive trees. *Random Structures Algorithms* **44-1**, (2014), 1098-2418 .
- [8] Bertoin, J. On the non-Gaussian fluctuations of the giant cluster for percolation on random recursive trees. *Electron. J. Probab.* **19-24** (2014), 1-15.
- [9] Devroye, L. Applications of the theory of records in the study of random trees. *Acta Inform.* **26** (1988), 123-130.
- [10] Drmota, M. *Random trees*. Springer. New York, Vienna (2009).

- [11] Drmota, M., Iksanov, A., Möhle, M. et Rösler, U. A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree. *Random Structures Algorithms* **34-3** (2009), 319-336.
- [12] Erdős, P. et Rényi, A. On random graphs. I. *Publ. Math. Debrecen* **6** (1959), 290-297.
- [13] Erdős, P. et Rényi, A. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17-61.
- [14] Gilbert, E.N. Random graphs. *Ann. Math. Statist.* **30** (1959), 1141-1144.
- [15] Goldschmidt, C. et Martin, J. B. Random recursive trees and the Bolthausen-Sznitman coalescent. *Electron. J. Probab.* **10** (2005), 718-745.
- [16] Iksanov, A. et Möhle, M. A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree. *Electron. Comm. Probab.* **12** (2007), 28-35.
- [17] Kuba, M. et Panholzer, A. Multiple isolation of nodes in recursive trees. *Online Journal of Analytic Combinatorics* **9** (2014).
- [18] Meir, A. et Moon, J. W. Cutting down random trees. *J. Austral. Math. Soc.* **11** (1970), 313-324.
- [19] Meir, A. et Moon, J. W. Cutting down recursive trees. *Mathematical Biosciences* **21** (1974), 173-181.
- [20] Möhle, M. The Mittag-Leffler process and a scaling limit for the block counting process of the Bolthausen-Sznitman coalescent. *Lat. Am. J. Probab. Math. Stat.* **12-1** (2015), 35-53.
- [21] Pitman, J. *Combinatorial Stochastic Processes*. École d'été de Probabilités de St. Flour. Lecture Notes in Mathematics **1875**, Springer (2006).
- [22] Schweinsberg, J. Dynamics of the evolving Bolthausen-Sznitman coalescent. *Electron. J. Probab.* **17** (2012), 1-50.
- [23] Yule, G. U. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London (series B)* **213** (1925), 21-87.