

Introduction à un domaine de recherche : Grandes composantes d'un graphe aléatoire critique

Guillaume Conchon–Kerjan

24 octobre 2016

Table des matières

1	Introduction	2
2	Graphes aléatoires	3
2.1	Qu'est-ce ?	3
2.2	Comment les étudier ? Exemple d'Erdős-Rényi	3
2.2.1	Un bon paramétrage	3
2.2.2	Processus d'exploration	4
2.2.3	Comparaison avec un processus de branchement	5
2.2.4	La fenêtre critique	6
3	Grandes composantes	7
3.1	Taille et surplus	7
3.2	Arbres continus	7
3.3	Arbres browniens	9
3.4	Du processus d'exploration à la structure d'une composante	9
4	Un modèle de configuration	10
4.1	Définition	10
4.2	Exploration	11
4.3	Résultats	11
4.4	Questions ouvertes	13
5	Références	13

1 Introduction

La notion de "graphe aléatoire" a été introduite il y a quelques décennies, entre les probabilités et la combinatoire. Paul Erdős et Alfred Rényi signent un article fondateur [5] en 1959, s'en suivront d'autres jusqu'en 1968. Il s'agit de regarder des grands réseaux de sommets reliés par des arêtes tracées selon un certain processus aléatoire, et d'en étudier les propriétés asymptotiques, notamment de connexité. Par exemple, quelle proportion des sommets se trouve dans la plus grande ou une des plus grandes composantes connexes ? Combien y en a-t-il ? Quelle est la distance typique entre deux sommets d'une même composante ?

Nous nous intéressons ici à des graphes aléatoires dits "critiques" : ce contexte sous-entend de considérer une suite de graphes indexée par leur nombre croissant de sommets, les arêtes étant tirées selon une procédure aléatoire paramétrée. Lorsqu'un certain paramètre est faible, il y a "peu" d'arêtes et les composantes sont "petites" en général $O(\ln(n))$ pour un graphe à n sommets. La connectivité du graphe augmente avec ce paramètre, et on quitte cette phase sous-critique pour terminer dans une phase sur-critique avec des composantes d'ordre $\Theta(n)$, où les distances typiques voire maximales entre deux sommets peuvent être parfois bornées par une constante finie avec une grande probabilité. Entre ces deux régimes se situe une transition de phase, où les composantes connexes les plus importantes sont de taille intermédiaire, typiquement n^a avec $0 < a < 1$.

Bien sûr, cette notion de est (encore) assez informelle et dépend du modèle. Toutefois, depuis une vingtaine d'années, certains chercheurs cherchent à « universaliser » certaines méthodes et certains résultats. On peut citer Remco Van der Hofstad, Shankar Bhamidi, Nicolas Broutin, Christina Goldschmidt, etc. Cette terminologie peut évoquer la physique, en particulier les changements d'états. En effet, les graphes aléatoires ont des applications en physique (physique statistique, théorie de la percolation, condensat de Bose-Einstein, etc.), mais aussi en chimie ou biologie, lorsqu'il est question de réseaux. Des modélisations dans des domaines plus difficiles à appréhender, comme des phénomènes sociaux (communication, élections...), commencent également à se développer.

Dans le modèle le plus simple, celui d'Erdős-Rényi, Aldous a montré en 1997 dans [1] que les tailles des composantes géantes en régime critique renormalisées par $\Theta(n^{2/3})$ sont données par les longueurs d'excursion au-dessus de zéro d'un mouvement brownien modifié (drifté par une parabole négative puis réfléchi au-dessus de son minimum). Addario-Berry, Broutin et Goldschmidt ont montré dans [2] que ces composantes renormalisées vues en tant qu'espaces métriques ont pour limite des arbres aléatoires browniens dont un nombre fini de sommets sont identifiés par paire. Concernant le modèle de configuration, Joseph a étendu dans [3] le résultat d'Aldous pour une distribution des degrés à queue polynomiale admettant une variance finie. Lorsqu'il existe un moment d'ordre 3, Bhamidi, Shen et van der Hofstad ont obtenu dans [4] un résultat similaire à celui du modèle d'Erdős-Rényi. Un résultat semblable est attendu lorsqu'on retire cette condition.

Contrairement à ce que laisse entendre le titre, nous nous focaliserons surtout sur le modèle d'Erdős-Rényi, dont l'étude est la plus simple parmi les modèles classiques. La dernière section fait le point sur ce qu'il en est du modèle de configuration.

2 Graphes aléatoires

2.1 Qu'est-ce ?

Un graphe est la donnée d'un ensemble de sommets, pouvant être reliés deux-à-deux par des arêtes.

Définition 1 *Un graphe G est la donnée d'un couple (V, E) où $V = \{v_1, \dots, v_n\}$ est un ensemble de sommets et $E = \{(v_{a_1}, v_{b_1}), \dots, (v_{a_l}, v_{b_l})\}$ un ensemble d'arêtes, identifiées par les paires de sommets qu'elles relient. Ainsi, v_i et v_j sont reliées dans G si et seulement si $(v_i, v_j) \in E$.*

On n'oriente pas les arêtes : (v_i, v_j) est la même arête que (v_j, v_i) . Pour des raisons de simplicité, on nommera $V = \{1, \dots, n\}$ les sommets d'un graphe G à n sommets. Un "graphe aléatoire" correspond au tirage, selon une certaine loi, d'un graphe, lequel est constitué d'un ensemble de noeuds et d'arêtes. Plus formellement :

Définition 2 *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Un graphe aléatoire est une application ϕ telle que pour tout $\omega \in \Omega$, $\phi(\omega) = (V(\omega), E(\omega))$ soit un graphe.*

On considère souvent des graphes aléatoires dont le nombre de sommets est déterministe. On considère alors une suite $G_n = (V_n, E_n)$ de tels graphes avec $V_n = \{1, \dots, n\}$ et on étudie les propriétés asymptotiques de G_n . Il existe de nombreuses façons de "tirer un graphe au hasard". L'ouvrage récent de van der Hofstad [6], présentant les principaux modèles avec force détails, est une excellente introduction au sujet. Le graphe aléatoire le plus connu et le mieux étudié aujourd'hui est sans doute celui d'Erdős-Rényi introduit en 1959 :

Exemple 3 *Soit $p \in [0, 1]$ et n un entier positif. Le graphe d'Erdős-Rényi à n sommets $ER_n(p)$ a pour ensemble de sommets $\{1, \dots, n\}$, et pour tous sommets $i \neq j$, l'arête entre i et j existe avec probabilité p , indépendamment de toutes les autres arêtes.*

L'identique distribution des arêtes, et surtout leur indépendance, rendent nombre de calculs plus simples.

2.2 Comment les étudier ? Exemple d'Erdős-Rényi

2.2.1 Un bon paramétrage

Le graphe d'Erdős-Rényi $ER_n(p)$ dépend donc de deux paramètres, n le nombre de sommets et p la probabilité de présence pour chaque arête. Pour éviter d'avoir un graphe ressemblant trop à un amas de points isolés, ou qui soit au contraire presque complet, il semble raisonnable de demander que chaque sommet ait un nombre moyen fini de voisins. Un bon paramétrage de p en fonction de n est

$$p = \frac{\lambda}{n},$$

où $\lambda \in \mathbb{R}_+$ est une constante.

2.2.2 Processus d'exploration

Lorsqu'on a affaire à un modèle discret fini, c'est-à-dire que le graphe aléatoire ne peut prendre qu'un nombre fini de formes avec une probabilité non-nulle, on serait tenté de les classer selon celles qui vérifient la propriété qui nous intéresse, et celles qui ne la vérifient pas. Par exemple, dans le graphe d'Erdős-Rényi, il existe $2^{\binom{n}{2}}$ graphes équiprobables, et la probabilité d'avoir un graphe à k arêtes est

$$\frac{\binom{\binom{n}{2}}{k}}{2^{\binom{n}{2}}}.$$

Cette approche combinatoire un peu brute se révèle vite inefficace, notamment lorsqu'on regarde la probabilité de propriétés plus difficilement accessibles, comme le fait que la plus grande composante connexe soit asymptotiquement de telle ou telle taille.

Un idée clé est d'explorer le graphe à partir d'un sommet, en empruntant des arêtes que l'on n'a pas encore traversées, et ce jusqu'à avoir visité tous les sommets de la composante connexe où l'on se trouvait, puis on recommence en partant d'un nouveau sommet. On y associe un **processus d'exploration** : il s'agit d'un compteur que l'on augmente ou diminue à chaque étape selon une certaine règle, et que l'on peut relier à un objet probabiliste bien connu. En général, il s'agit d'une marche aléatoire sur \mathbb{N} .

Exemple 4 *Exploration du graphe d'Erdős-Rényi*

Sur $ER_n(p)$, on définit le processus suivant (exemple ci-dessous) : tous les sommets sont initialement "endormis", et on en choisit un, v_1 , uniformément dans $\{1, \dots, n\}$. On le réveille à l'étape 0. À l'étape 1, on le tue et on réveille ses voisins. À l'étape 2, on réveille les voisins de v_2 (le voisin de v_1 ayant le plus petit numéro) encore endormis et non décédés, puis on tue v_2 . v_3 est le voisin en vie de plus petit numéro du dernier sommet visité qui en a encore (donc voisin de v_2 si v_2 en a encore au moins un). On procède ainsi étape par étape jusqu'à avoir parcouru toute la composante connexe de v_1 . On choisit alors un nouveau sommet et on recommence. On note v_k le sommet tué à l'étape k .

Le **processus d'exploration** associé est la suite $(S(k))_{k \leq n}$ telle que pour $k \geq 0$, $S(k)$ soit le nombre de sommets réveillés moins le nombre de sommets décédés à l'étape k , $k \geq 1$. Nous allons voir qu'il peut être relativement bien approximé par des objets connus et aisément maniables en probabilité : les marches aléatoires.

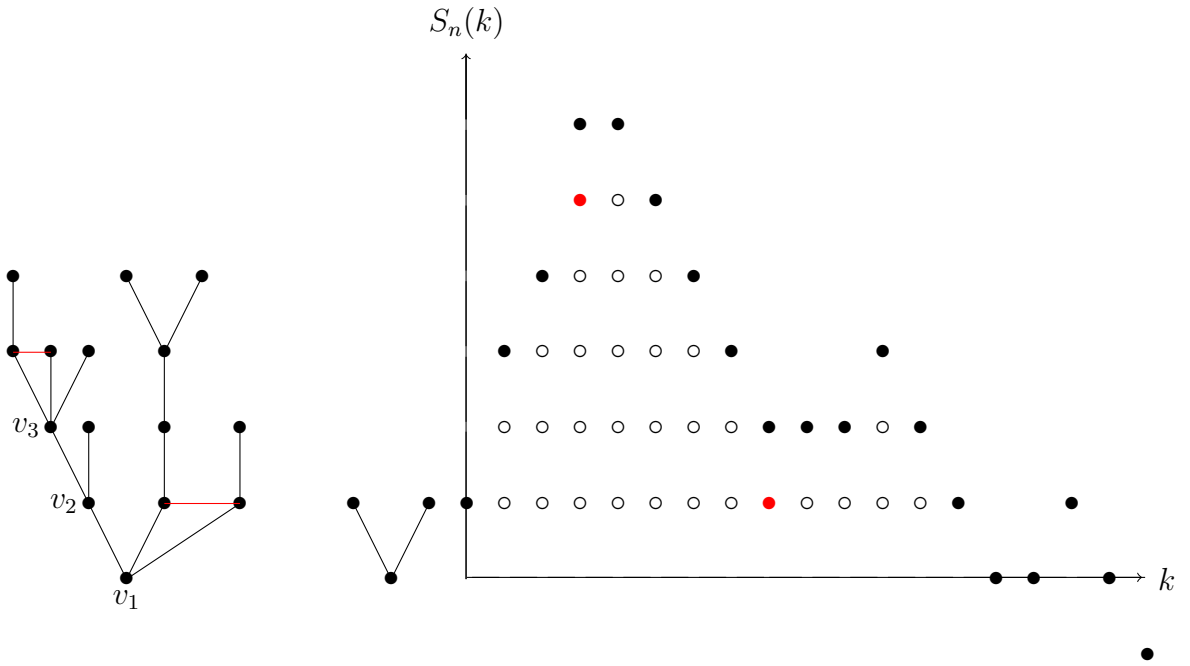


Figure 1 : exploration d'un graphe ayant deux composantes. Les points blancs figurent les sommets réveillés "en stock", donc pas encore explorés à chaque étape. En rouge, les arêtes en surplus, c'est-à-dire ne faisant pas partie de l'arbre d', qui lui comprend les arêtes (v_k, v_{k+1}) . Sur le graphique, les points rouges représentent leur position (i.e. le sommet exploré où on les découvre et le sommet endormi qu'elles rejoignent).

Remarque 5 *Lorsqu'on explore ainsi chaque branche le plus en profondeur possible avant de passer à la suivante (ie, priorité est donnée au fils du premier fils plutôt qu'au deuxième fils), on parle en anglais de "depth-first search", couramment abrégée en DFS, par opposition à une "breadth-first search", qui se fait en largeur. Le processus S associé est nommé "depth-first walk".*

2.2.3 Comparaison avec un processus de branchement

Définition 6 *Un processus de branchement Z de loi ν (sur \mathbb{N}) et de masse initiale $m \in \mathbb{N}$ est une suite de variables aléatoires $(Z_n)_{n \geq 0}$ représentant les différentes générations d'une population, de sorte que le nombre initial d'individus $Z_0 = m$ p.s., et à la génération n , chaque individu donne naissance à un nombre aléatoire de descendants de loi ν , indépendamment de tous les autres.*

Soit $D_{n,p}$ le nombre total de descendants d'un branchement avec masse initiale $m = 1$ et loi de reproduction $\nu = \text{Bin}(n, p)$, C_1 la taille de la première composante connexe rencontrée lors de l'exploration du graphe, et C_{max} celle de la plus grande. Une remarque fondamentale est la suivante :

Proposition 7 *La loi du nombre de sommets réveillés à l'étape k est binomiale de paramètres $n - S(k - 1) - k$ et p , autrement dit,*

$$S(k) - S(k - 1) + 1 \sim \text{Bin}(n - S(k - 1) - k, p).$$

Ainsi, le nombre de sommets ajoutés à l'exploration à chaque étape est dominé stochastiquement par une v.a. de loi ν . De plus, cette domination est assez précise, si la composante est petite, car $n - S(k-1) - k \geq C_1$ pour tout $k \leq C_1$. On en déduit le résultat suivant :

Théorème 8 *Pour tout $k \geq 1$,*

$$\mathbb{P}(D_{n-k,p} \geq k) \leq \mathbb{P}(C_1 \geq k) \leq \mathbb{P}(D_{n,p} \geq k). \quad (1)$$

Il reste à déterminer quelle peut être la taille de la progéniture totale d'un certain branchement binomial.

2.2.4 La fenêtre critique

Rappelons que nous sommes ici intéressés par le cas critique. La fenêtre critique est définie par

$$p = \frac{1}{n} + \frac{\theta}{n^{4/3}}, \quad \theta \in \mathbb{R},$$

autrement dit $\lambda = 1 + \theta/n^{1/3}$, donc $\lambda \rightarrow 1$, et nous resterons désormais dans ce cadre, sauf mention du contraire.

Soit dit en passant, dans le cas sous-critique, où $p \sim \frac{c}{n}$ pour $c < 1$, les plus grandes composantes sont en $\Theta(\ln(n))$ avec grande probabilité, et dans le sur-critique, soit $c > 1$, elles sont en $\Theta(n)$ (voir [6] par exemple pour des détails). On s'attend donc à trouver un ordre de grandeur entre les deux.

Notons que la loi binomiale vue précédemment n'est pas anodine : lorsque $n \rightarrow +\infty$ et $np \rightarrow \lambda$, un petit calcul (passant par la fonction caractéristique) montre que la suite de lois binomiales $Bin(n, p)$ a pour limite une loi de Poisson de paramètre λ . Nous citons le théorème 5.1 du volume I de [6] :

Théorème 9 *Pour tout $\theta \in \mathbb{R}$, il existe $b = b(\theta) \in \mathbb{R}$ tel que pour tout $w > 1$,*

$$\mathbb{P}_{n,p}(w^{-1}n^{2/3} \leq C_{max} \leq wn^{2/3}) \geq 1 - \frac{b}{w}.$$

La démonstration utilise le fait que la probabilité que la descendance totale T^* d'un branchement issu d'une loi de Poisson de paramètre λ vérifie

$$\mathbb{P}_\lambda(T^* = n) = \frac{(\lambda n)^{n-1}}{n!} e^{-\lambda n}$$

pour $n \geq 1$. En appliquant cette formule (dont la preuve peut se faire par récurrence) à $\lambda = 1 + \frac{\theta}{n^{1/3}}$ et en utilisant la formule de Stirling, on peut se convaincre du fait que

$$\mathbb{E}_{n,p}[C_1] = \Theta(n^{1/3}).$$

Et, en imaginant que la contribution majeure à l'espérance de la taille de la plus grande composante soit due à des sommets dans cette plus grande composante, on peut imaginer que pour tout sommet v ,

$$\mathbb{E}_{n,p}[C_1] \simeq \mathbb{E}_{n,p}[C_1 | \mathbf{1}_{1 \in C_{max}}] = \mathbb{E}_{n,p}[C_1 | \mathbf{1}_{v \in C_{max}}] = \mathbb{E}_{n,p}[C_{max}^2]/n.$$

On en tirerait donc $\mathbb{E}_{n,p}[C_{max}^2] = \Theta(n^{4/3})$, cohérent avec l'hypothèse $C_{max} \simeq n^{2/3}$ avec grande probabilité.

3 Grandes composantes

3.1 Taille et surplus

Dans la section précédente, la comparaison avec un processus de branchement permet de montrer que dans le régime critique, la taille typique d'une grande composante est $n^{2/3}$. Aldous a prouvé dans [1] la convergence fonctionnelle en loi du processus d'exploration réfléchi au-dessus de son minimum, soit $R_n(k) = S_n(k) - \min_{j \leq k} S_n(j)$. On définit ainsi pour $t \in \mathbb{R}_+$

$$W_t = B_t + \theta t - \frac{1}{2}t^2 - \min_{0 \leq s \leq t} (B_s + \theta s - \frac{1}{2}s^2), \quad (2)$$

où B est un mouvement brownien standard. Ainsi, W est une fonction continue non dérivable, qui réalise une suite d'excursions au-dessus de 0, dont la longueur tend p.s. vers 0.

Théorème 10 (Aldous 1997) *Soit $\bar{R}_n(t) = n^{-1/3}R_n(tn^{2/3})$ pour $t > 0$, on a la convergence en loi*

$$\bar{R} \rightarrow W.$$

La démonstration s'appuie sur une variation élaborée (théorème 7.1.4 de [7]) du théorème de Donsker selon lequel une marche aléatoire assez régulière renormalisée a pour limite d'échelle un mouvement brownien.

Théorème 11 *Soit Z^n la suite décroissante des tailles des composantes connexes de $ER_n(p)$, et S^n la suite des surplus associés, le surplus d'une composante étant le nombre d'arêtes supplémentaires par rapport à un arbre ayant autant de sommets. On a la convergence en loi suivante :*

$$(n^{-2/3}Z^n, S^n) \rightarrow (Z, S), \quad (3)$$

où Z a la loi de la suite décroissante des excursions de W au-dessus de 0 et S est une suite de v.a. indépendantes de loi de Poisson dont les paramètres sont les aires sous les excursions correspondantes.

La preuve de la convergence pour la première coordonnée découle directement du théorème précédent. Pour la seconde, on fait comme pour la proposition 7 une approximation binomiale, une arête en surplus apparaissant à l'étape i si et seulement si son autre sommet, choisi uniformément au hasard, est réveillé et encore vivant. Le surplus total est donc environ une somme de variables binomiales dont la limite en loi est une loi de Poisson.

3.2 Arbres continus

Le résultat précédent laisse à penser que les grandes composantes d'un graphe d'Erdős-Rényi doivent avoir pour "limite" des arbres dont un nombre fini de paires de sommets sont identifiées, correspondant aux surplus.

Définition 12 On appelle **arbre réel** un espace métrique compact (\mathcal{T}, d) tel que pour tous $x, y \in \mathcal{T}$, il existe un unique plus court chemin $[[x, y]]$ de x à y , qui est l'unique chemin ne s'intersectant pas.

- On appelle sommet tout élément $v \in \mathcal{T}$.
- On dit qu'un arbre est enraciné lorsqu'on en distingue un sommet ρ , appelé racine.
- La hauteur d'un sommet est sa distance à la racine ρ .
- un sommet v est une feuille si et seulement si $v \notin [[\rho, w]]$ pour tout $w \neq v$.

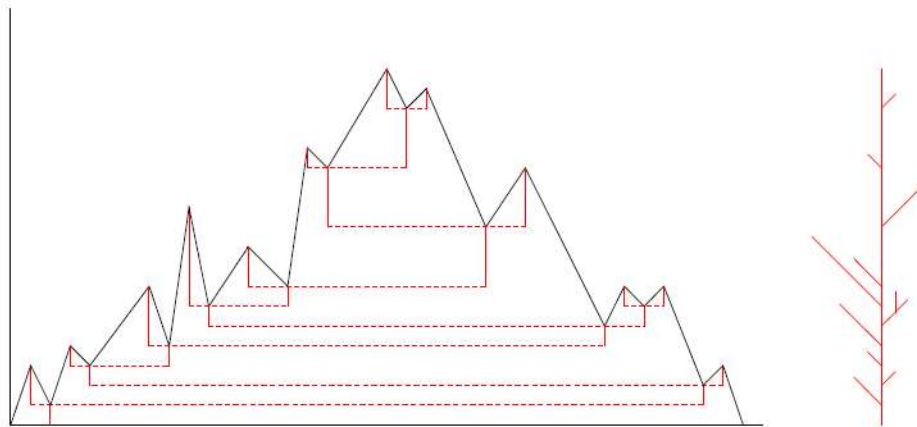
Remarque 13 Ces arbres réels sont plutôt des "arbres continus". Les arbres discrets ont également une structure d'espace topologique sous-jacente, où la distance entre deux sommets est le nombre minimal d'arêtes à parcourir pour passer de l'un à l'autre.

Définition 14 On appelle **excursion** une fonction $f \in C^0(\mathbb{R}_+, \mathbb{R}_+)$ à support compact telle que $f(0) = 0$ et pour tous $0 < x \leq y$, $f(x) = 0$ implique $f(y) = 0$. La longueur l de l'excursion est $l(f) := \sup_{x>0} \{x, f(x) > 0\}$.

Définition 15 Soit f une excursion, et $d_f(x, y) := f(x) + f(y) - 2 \min_{x \leq z \leq y} f(z)$ pour tous $0 < x < y$ une distance sur $[0, +\infty)$. L'espace topologique $\mathcal{T}_f := [0, +\infty) / \sim$ où \sim est la relation d'équivalence définie par $x \sim y$ lorsque $d_f(x, y) = 0$ est appelé **l'arbre réel T encodé par f** . On dit que f est le contour de T .

On enracine cette arbre, ρ étant la classe d'équivalence de 0.

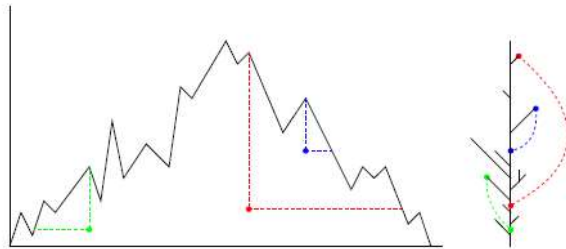
Pour mieux se le représenter, on peut dessiner une excursion affine par morceaux, et coller les deux pentes autour de chaque maximum local, les points identifiés étant ceux à même hauteur autour de ce maximum.



(Image : C. Goldschmidt)

Définition 16 Une excursion marquée est une excursion comportant un nombre fini de points sous sa courbe, appelés "marques".

Le graphe associé à une excursion marquée f est obtenu en traçant l'arbre associé à l'excursion, et en identifiant les classes d'équivalence de $f(x_i)$ et de y_i où (x_i, y_i) sont les coordonnées de la i -ème marque, ce qui crée un "raccourci" par marque.



(Image : C. Goldschmidt)

3.3 Arbres browniens

Définition 17 L'arbre brownien réel continu \mathcal{T}_B est un arbre aléatoire dont la loi est donnée par l'arbre d'une excursion brownienne conditionnée à être de longueur 1.

Il en existe plusieurs constructions (voir [8], très abordable). Nous donnons une propriété importante en tant que limite d'échelle.

Théorème 18 (Aldous 1993, Le Gall 2005) Pour $n \in \mathbb{N}$, \mathcal{T}_n un arbre aléatoire dont la loi est uniforme sur les arbres à n sommets. On a la convergence en loi

$$\frac{1}{\sqrt{n}}\mathcal{T}_n \rightarrow \mathcal{T}_B$$

pour la distance de Gromov-Hausdorff.

Remarque 19 On parle ici d'arbres étiquetés (chaque sommet a un numéro) mais non enracinés.

3.4 Du processus d'exploration à la structure d'une composante

La propriété suivante (valide pour toute valeur de $p \in [0, 1]$) exprime toute l'avantageuse particularité du modèle d'Erdős-Rényi.

Proposition 20 *Une composante de $ER_n(p)$ conditionnée à avoir m sommets et un surplus s est un graphe connexe uniformément choisi parmi ceux ayant m sommets et $m + s - 1$ arêtes.*

Ainsi, la loi d'une composante ne dépend que de sa taille et de son surplus. Remarquons que le processus d'exploration donne un arbre, dont les arêtes sont les (v_{k-1}, v_k) . Chaque arbre à m sommets est ainsi le "squelette" d'un certain nombre de composantes possibles. Dans [2], les auteurs construisent une composante en tirant d'abord son squelette, puis les arêtes supplémentaires.

Lors de l'exploration d'un graphe sommet par sommet, on définit le **processus de hauteur** ("height process" en anglais) $H_n(k)$ comme la distance de graphe entre v_1 et v_k .

Théorème 21 (Marckert et Mokkadem [9]) *Soit e une excursion brownienne de longueur 1, X_n et H_n les processus d'exploration et de hauteur d'un arbre aléatoire uniforme à n sommets, on a la convergence en loi*

$$\frac{1}{\sqrt{n}}(X_n(\lfloor nt \rfloor), H_n(\lfloor nt \rfloor), 0 \leq t \leq 1) \rightarrow (e, 2e)$$

pour la topologie de Skorokhod (plus lâche que celle de la convergence uniforme et adaptée aux fonctions càdlàg - continues à droite et avec une limite à gauche en tout point).

Le processus d'exploration et le processus de hauteur sont donc quasiment identiques dans Erdős-Rényi, avec une limite d'échelle brownienne. Ceci explique qu'une composante, à la limite, doit ressembler à un arbre brownien choisi selon un certain biais et un nombre fini d'identifications de points correspondant aux arêtes en surplus (dont la longueur $n^{-1/3}$ disparaît dans la renormalisation à la limite).

Ce biais correspond, à surplus fixé, aux nombres de composantes dont un arbre peut être le squelette d'exploration; à la limite, il est proportionnel à l'exponentielle de l'aire.

Théorème 22 (Addario-Berry, Broutin, Goldschmidt 2010) *Soit C_1^n, C_2^n, \dots les composantes connexes de $ER_p(n)$ vues comme espaces métriques, et C_1, C_2, \dots les espaces métriques correspondants aux excursions décroissantes de W , marquées par un processus de Poisson uniforme. On a la convergence en loi :*

$$n^{-1/3}(C_1^n, C_2^n, \dots) \rightarrow (C_1, C_2, \dots)$$

pour la métrique $d(A, B) = (\sum_{i \geq 1} d_{GH}(A_i, B_i)^4)^{1/4}$, où d_{GH} est la distance de Gromov-Hausdorff.

4 Un modèle de configuration

4.1 Définition

Le modèle dit "de configuration" est plus complexe que celui d'Erdős-Rényi. Il présente notamment deux sources d'aléatoire.

On commence en effet par choisir le degré de chaque sommet : on se donne $\nu = (\nu_k)_{k \in \mathbb{N}}$ une distribution de probabilité sur \mathbb{N} (donc $\sum \nu_k = 1$), et pour chaque sommet, son degré est tiré au sort selon ν , indépendamment de tous les autres. On peut représenter cela par des "demi-arêtes" sortant de chacun des sommets. Ces demi-arêtes sont ensuite reliées deux par deux, uniformément au hasard. S'il y en avait un nombre impair, la dernière est supprimée.

Remarques 23 *Cette procédure permet de créer des boucles, donc des arêtes dont les extrémités sont confondues en un sommet, ainsi que des arêtes multiples entre deux sommets. En général, on efface les boucles et arêtes surnuméraires, on obtient alors un graphe dit "simple".*

ν est souvent telle qu'il y a peu de telles arêtes erratiques avec grande probabilité, ce qui permet de supposer qu'il n'y en a pas sans modifier les propriétés intéressantes du graphe.

Dans ce modèle, la phase critique est décrite par l'égalité

$$\sum_{k \geq 0} k(k-2)\nu_k = 0, \quad (4)$$

autrement dit, $\mathbb{E}[\nu^2] = 2\mathbb{E}[\nu]$.

On suppose de plus que pour certains $\alpha \in (1, 2)$ et $c > 0$:

$$\nu_k \sim_{k \rightarrow +\infty} ck^{-(\alpha+2)} \quad (5)$$

Ainsi, $\sum_{k=1}^{+\infty} k^2 \nu_k < \infty$ mais $\sum_{k=1}^{+\infty} k^3 \nu_k = \infty$.

4.2 Exploration

Pour $n \in \mathbb{N}$ fixé, on appelle G_n le graphe à n sommets. On explore G_n en choisissant cette fois une demi-arête à chaque étape et en explorant son sommet. Soit r_i le sommet de la i -ème étape et D_i^n son degré.

À l'étape 0, on pose $S_n(0) = 0$. On choisit une demi-arête uniformément au hasard, en nommant r_1 son sommet. On dessine les L_1 éventuelles boucles autour de r_1 , et on pose $S(1) = D_1^n - 2L_1^n$, de sorte que $S_n(1)$ soit le nombre de demi-arêtes découvertes mais pas encore appariées. Celles-ci sont ajoutées en haut d'une pile initialement vide, dans un ordre arbitraire.

À l'étape k , on prend la première demi-arête de la pile, qui appartient donc à r_{k-1} , et on l'apparie à sa demi-arête correspondante dans le graphe, qui appartient donc à un nouveau sommet r_k . On en trace toutes les boucles et arêtes ayant pour extrémité $r_i, i < k$. Les demi-arêtes non appariées de r_k sont ajoutées **en haut** de la pile, et $S_n(k)$ est le cardinal de la pile.

4.3 Résultats

Un fait crucial est qu'après $o(n)$ étapes, les sommets rencontrés l'auront été quasiment selon un **biais par la taille** de ν .

Définition 24 *Biais par la taille* Soit μ une probabilité sur \mathbb{R}_+ , son biais par la taille, noté $\tilde{\mu}$, est la loi définie pour tout borélien $B \in \mathbb{R}_+$ par :

$$\tilde{\mu}(B) = \frac{\int_B x\mu(dx)}{\int_{\mathbb{R}_+} \mu(x)dx}.$$

Dans notre cas, cela revient à dire qu'à chaque étape, chaque sommet a une probabilité d'être appelé environ proportionnelle à son degré, puisqu'on sélectionne les demi-arêtes uniformément. Ce qui nous donne le résultat suivant :

Lemme 25 Soit Z_1, \dots, Z_n des v.a. iid de loi $\tilde{\nu}$ et k_1, \dots, k_n des entiers,

$$\mathbb{P}(\hat{D}_1^n \geq k_1, \dots, \hat{D}_n^n \geq k_n) \leq \mathbb{P}(Z_1 \geq k_1, \dots, Z_n \geq k_n).$$

Adrien Joseph [3] a montré, avec des arguments de semimartingale similaires à ceux d'Aldous, la convergence en loi du processus d'exploration renormalisé vers un processus "presque" α -stable. Ce résultat, prouvé pour $\alpha \in (2, 3)$, peut être étendu à notre cas, moyennant quelques passages fastidieux dus au fait que le biais par la taille de ν a une variance infinie :

Théorème 26 Soit $\bar{S}_n(t) = \frac{1}{n^{1/\alpha+1}} S_n(\lfloor tn^{\alpha/\alpha+1} \rfloor)$. On a la convergence en loi :

$$\bar{S}_n \rightarrow X + A$$

sur tout intervalle compact par rapport à la topologie de Skorokhod, où $A(t) = Ct^\alpha$ pour une constante $C < 0$ et X est l'unique processus à accroissements indépendants tel que pour $t > 0$ et $u \in \mathbb{R}$,

$$\mathbb{E}[\exp(iuX_t)] = \exp\left(\int_0^t ds \int_0^{+\infty} (e^{iux} - 1 - iux) \frac{c}{\eta^{\alpha+1}} e^{-xs/\eta} dx\right) \text{ en notant } \eta = \mathbb{E}[\nu].$$

Ce résultat reste vrai en réfléchissant S_n et le processus limite au-dessus de leur minimum. On en déduit l'échelle des plus grandes composantes et leur taille en fonction des excursions de ce processus.

En approximant de nouveau l'apparition d'arêtes en surplus à chaque étape de l'exploration, qui sont les arêtes repartant vers des sommets déjà explorés, par des lois binomiales dont la somme converge en loi vers une loi de Poisson, on montre le résultat suivant :

Théorème 27 Soit $N_n(t)$ le nombre d'arêtes en surplus rencontrées après $\lfloor tn^{\alpha/\alpha+1} \rfloor$ étapes d'exploration. Soit N un processus ponctuel de Poisson d'intensité (aléatoire) $X(t) + A(t)$ au temps t . On a la convergence en loi :

$$(\bar{S}_n, N_n) \rightarrow (X + A, N)$$

par rapport à la topologie produit, i.e. la topologie de Skorokhod pour chacune des variables.

4.4 Questions ouvertes

On prouve ainsi que dans ce modèle de configuration, les grandes composantes ont un surplus p.s. fini lorsque $n \rightarrow +\infty$. Toutefois, nous n'avons pas d'information sur la structure de l'arbre et la position des "raccourcis" dus aux surplus qui devraient intervenir à la limite.

Dans le cas où $\alpha \in (2, 3)$, Bhamidi, Shen et van der Hofstad ont montré dans [4] la convergence des grandes composantes vers des arbres dont la loi est absolument continue par rapport à des arbres α -stables avec un nombre fini de raccourcis entre des nœuds de degré infini et des feuilles.

On peut de plus se demander si ces résultats sont généralisables à des modèles de configuration ayant seulement une condition sur leurs moments, moins restrictive que d'avoir une queue polynômiale d'un certain degré. Dans un horizon plus lointain, on peut se demander si une certaine "universalité" n'apparaîtrait pas dans les différents modèles comme pour la phase critique : les grandes composantes critiques seraient alors des arbres aléatoires avec un nombre fini de raccourcis, les arbres en question dépendant du modèle et de ses paramètres.

5 Références

- [1] ALDOUS D., Brownian excursions, critical random graphs, and the multiplicative coalescent, *Annals of Probability*, 1997, Vol. 25, No. 2, 812–854.
- [2] ADDARIO-BERRY L., BROUTIN N., GOLDSCHMIDT C., *The continuum limit of critical random graphs*. Springer-Verlag 2010.
- [3] JOSEPH A., The component sizes of a critical random graph with given degree sequence, 11 Sep 2014, arXiv :1012.2352v3.
- [4] BHAMIDI S., SHEN S., VAN DER HOFSTAD R., The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs, arXiv :1508.04645v1, 19 Aug 2015.
- [5] ERDÖS P., RÉNYI, A. (1959). On the evolution of random graphs. *Magy. Tud. Akad. Mat. Kut. Intéz. Közl.* 5 17–61. MR0125031.
- [6] VAN DER HOFSTAD R., *Random graphs and complete networks*, Vol. I and II,
- [7] ETHIER S.N., KURTZ T.G., *Markov Processes : Characterization and Convergence*, Wiley, New York, 1986.
- [8] ALDOUS D., The Continuum Random Tree I., *Ann. of Probab.* (19)1, 1-28 (1991).
- [9] MARCKERT J.-F., MOKKADEM A., The depth-first processes of Galton-Watson trees converge to the same brownian excursion, *Ann. Probab.* 31, 1655–1678 (2003)