Arbres couvrants minimums aléatoires inhomogènes, propriétés et limite

Othmane Safsafi

Laboratoire de probabilités statistique & modélisation

T H È S E

pour obtenir le titre de

Docteur en Sciences

de l'Université de

Mention : Mathématiques

Présentée et soutenue par

Othmane SAFSAFI

Arbres couvrants minimums aléatoires inhomogènes, propriétés et limite.

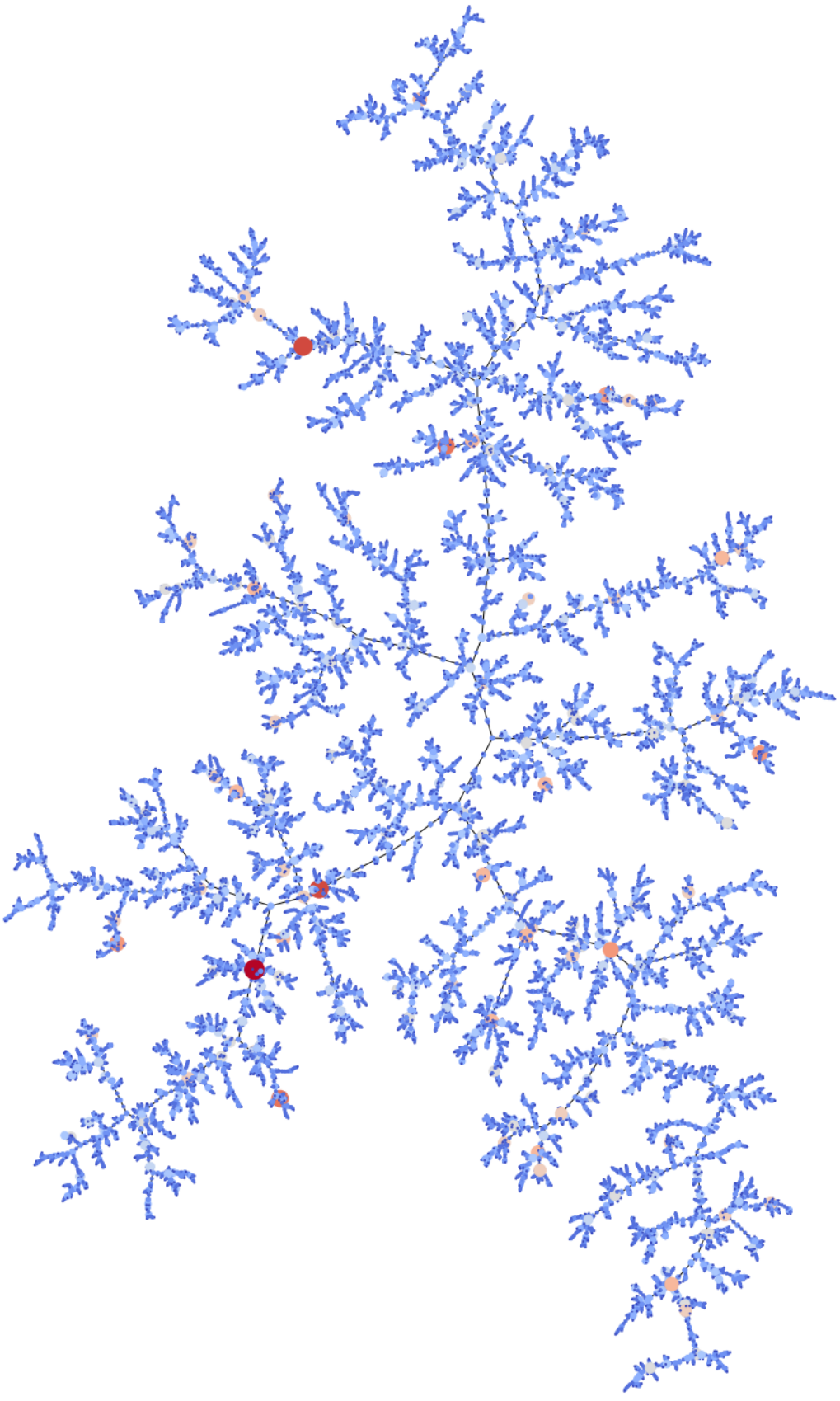
Thèse dirigée par Nicolas BROUTIN

préparée à Sorbonne université, Laboratoire LPSM

soutenue le 12 mars 2021

Jury :

Rapporteurs :	-Philippe Chassaing
	-Peter Mörters
Examinateurs :	-Thomas Duquesne
	-Christina Goldschmidt
	-Jean-Francois Marckert
	-Anna Ben-Hamou



Remerciements

Je voudrais commencer par exprimer ma plus profonde gratitude et mes remerciements à mon directeur de thèse Nicolas BROUTIN. Il m'a proposé un sujet de thèse passionnant et prometteur. Merci d'avoir eu la patience et la bienveillance de lire et de relire mes travaux à maintes reprises. Merci aussi de m'avoir donné l'opportunité de partager mes travaux à travers le monde.

Phillipe Chassaing et Peter Mörters ont accepté de rapporter cette thèse, je les remercie chaleureusement pour leur temps et leur précieuses remarques. Je tiens aussi à remercier Thomas Duquesne, Jean-Francois Marckert, Christina Goldschmidt, et Anna Ben-Hamou d'avoir accepté de faire partie du jury de cette thèse.

J'ai eu la chance de vivre cette thèse au côté de personnes formidables dans les laboratoires du LPSM. Merci à mes petits frères de thèse, Lucas Iziquel et Arthur Blanc-Renandie pour les discussions mathématiques intéressantes. Merci à tous mes autres amis du laboratoires, que ce soit à la cantine, autour d'un café en sale de repos ou durant le goûter post-séminaire des doctorants, discuter avec vous a toujours été un plaisir que j'attendais avec impatience.

Je tiens aussi à remercier les secrétaires du laboratoire, ils m'ont aidé plus que je ne l'imaginais en entreprenant cette thèse. Je reste infiniment reconnaissant aux personnes qui m'ont guidé et soutenues. Stéphane Boucheron mon tuteur à l'ENS, pour son soutien et son aide au début de cette thèse, et Thomas Duquesne pour le soutien moral en fin de thèse.

Un grand merci à mes amis, Simon, Pierre, Guillaume, Ludovic, Houssam, Marouane et bien d'autres. Que ce soit nos discussions de mathématiques, de vulgarisation scientifiques, de poker ou de tout autre sujet. Grâce à vous, les heures passées loin de la thèse ont été tout aussi intéressantes que celle passées à explorer mes arbres aléatoires.

Finalement, j'aimerais transmettre mes remerciements ineffables à mes plus grands soutiens. Mes parents, pour m'avoir continuellement encouragé et soutenu, mes petites soeurs pour avoir toujours su me remettre le sourire au lèvres, et ma femme, sans qui ce travail n'aurait jamais pu se concrétiser. Elle est mon pilier, mon amour et ma meilleure amie.

Table des matières

1 Introduction **7**
- 1.1 Sujet principal de la thèse . . . 8
  - 1.1.1 Introduction de l'introduction . . . 8
  - 1.1.2 Une brève histoire de l'arbre couvrant minimum . . . 10
  - 1.1.3 Constructions et propriétés du MST . . . 12
- 1.2 Rappels et exemples d'inégalités de concentration . . . 17
- 1.3 Les graphes aléatoires . . . 19
  - 1.3.1 Arbres de Galton-Watson . . . 19
  - 1.3.2 Comment explorer les graphes . . . 21
  - 1.3.3 Codages d'arbres de Galton-Watson . . . 25
  - 1.3.4 Graphes aléatoires uniformes et inhomogènes . . . 32
- 1.4 Retour vers l'arbre couvrant minimum . . . 36
  - 1.4.1 Le rapport entre MST et graphes aléatoires . . . 36
  - 1.4.2 Aparté sur la physique statistique . . . 39
- 1.5 Limite d'échelle du MST renormalisé . . . 39
  - 1.5.1 Convergence d'espaces métriques . . . 40
  - 1.5.2 Convergence de graphes inhomogènes discrets . . . 41
- 1.6 Résultats de cette thèse et questions futures . . . 43

2 Exponential bounds for inhomogeneous random graphs **53**
- 2.1 Introduction . . . 55
  - 2.1.1 The model . . . 55
  - 2.1.2 Definition of the exploration process . . . 55
  - 2.1.3 Conditions and main theorem . . . 57
  - 2.1.4 Motivation and previous work . . . 60
- 2.2 Bounding the weights . . . 62
  - 2.2.1 First concentration result and the mean . . . 63
  - 2.2.2 A more precise concentration inequality . . . 66
- 2.3 Bounds on the exploration process . . . 75
- 2.4 The structure of the giant component . . . 84
  - 2.4.1 The size of the giant component . . . 84
  - 2.4.2 The excess of the giant component. . . . 88
  - 2.4.3 The excess of the components discovered before the largest connected component. . . . 91
- 2.5 The structure of the tail's components . . . 93
  - 2.5.1 Preliminaries . . . 93
  - 2.5.2 The size of connected components discovered after the largest connected component . . . 97
  - 2.5.3 The excess of the tail . . . 102

3 Diameter of inhomogeneous minimum spanning tree **109**
- 3.1 Introduction . . . 111
  - 3.1.1 The model . . . 111
  - 3.1.2 Notations and definition of the exploration process . . . 112
  - 3.1.3 Further discussion and related work in statistical physics . . . 115
- 3.2 First ingredients of the proof . . . 118
  - 3.2.1 The phase transition . . . 118

3.2.2 The supercritical phase . . . 118
3.2.3 Known results . . . 120
3.3 Dealing with the small components . . . 121
3.3.1 Construction of the growth of small components . . . 122
3.3.2 Coupling with Galton-Watson trees . . . 126
3.4 Bounding the length of the longest path of the giant component . . . 132
3.4.1 A new pruning procedure for the giant component . . . 132
3.4.2 Bounding the longest path . . . 133
3.5 Proofs of Theorems 43, 45, 48 and 49 . . . 137

4 Scaling limit of inhomogeneous minimum spanning tree 141
4.1 introduction . . . 143
4.1.1 Defnitions and main results . . . 143
4.1.2 Related work . . . 144
4.1.3 Organization of the chapter . . . 145
4.2 Metric space notions and convergence . . . 146
4.2.1 Gromov-Haussdorf distance . . . 146
4.2.2 General definitions . . . 147
4.3 Cycle breaking algorithm in the continuous and discrete settings . . . 149
4.3.1 Definition of discrete and continuous cycle-breaking . . . 149
4.3.2 Relation between discrete and continuous cycle breaking . . . 150
4.4 Convergence of the discrete minimum spanning tree to its scaling limit . . . 151
4.4.1 The scaling limit of inhomogeneous random graphs . . . 151
4.4.2 Preliminary results and convergence of minimum spanning trees in the critical window . . . 154
4.4.3 Convergence of the largest minimum spanning tree in the supercritical regime 155
4.4.4 Properties of the scaling limit . . . 157

Bibliography 159

Chapitre 1

Introduction

Contents

1.1 Sujet principal de la thèse 8
1.1.1 Introduction de l'introduction 8
1.1.2 Une brève histoire de l'arbre couvrant minimum 10
1.1.3 Constructions et propriétés du MST 12
1.2 Rappels et exemples d'inégalités de concentration 17
1.3 Les graphes aléatoires 19
1.3.1 Arbres de Galton-Watson 19
1.3.2 Comment explorer les graphes 21
1.3.3 Codages d'arbres de Galton-Watson 25
1.3.4 Graphes aléatoires uniformes et inhomogènes 32
1.4 Retour vers l'arbre couvrant minimum 36
1.4.1 Le rapport entre MST et graphes aléatoires 36
1.4.2 Aparté sur la physique statistique 39
1.5 Limite d'échelle du MST renormalisé 39
1.5.1 Convergence d'espaces métriques 40
1.5.2 Convergence de graphes inhomogènes discrets 41
1.6 Résultats de cette thèse et questions futures 43

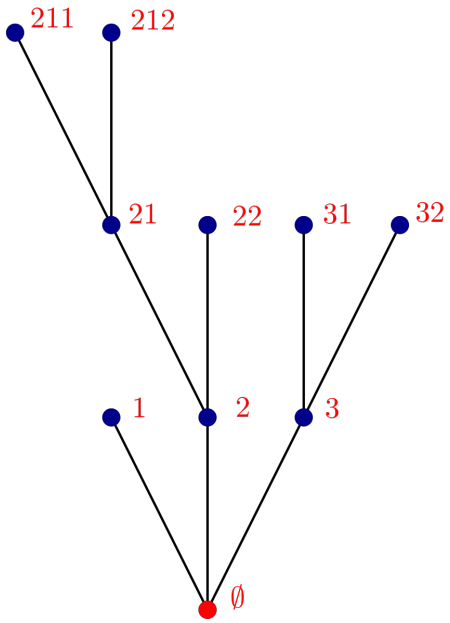
Introduction

1.1 Sujet principal de la thèse

1.1.1 Introduction de l'introduction

Dans cette introduction, nous présenterons les principaux objets auxquels on s'intéresse dans cette thèse ainsi qu'un résumé rapide de l'histoire de ces objets et de leurs importance de différents points de vue. Cette introduction permettra de bien comprendre les prochains chapitres de la présente thèse. En plus de cette introduction, cette thèse se divisent en trois chapitres. Ils sont complémentaires, chaque chapitre utilise les résultats des chapitres qui le précèdent afin d'obtenir de nouveaux théorèmes.

Durant la majeure partie de mon travail de thèse, je me suis intéressé à des questions autour des graphes aléatoires. Et c'est le sujet de cette thèse. Cependant, durant ma dernière année de thèse, je me suis intéressé à des applications en informatique d'outils développés pour les graphes aléatoires. Ceci m'a permis de travailler sur plusieurs collaborations, l'article Laroche, Safsafi, Broutin, and Féraud [2020] en est une. Dans ce dernier, nous appliquons des méthodes issues de l'étude des processus de branchement afin de résoudre un problème issue de la théorie des bandits. D'autres collaborations concernent des articles en cours de rédaction. J'ai aussi eu la chance d'appliquer certains de ces outils pour construire un algorithme de prédiction de l'évolution de la pandémie du Covid-19. Vu le temps restreint, je n'ai pas pu effectuer une analyse théorique poussée de cet algorithme. Cependant les résultats obtenus étaient assez concluants pour être utilisés par le gouvernement marocain dans son processus de décision concernant les politiques de confinement dans le pays. Tout ce travail n'est pas présent dans ce document de thèse qui est consacré à la partie mathématique de mes travaux de thèse. Cette introduction est donc elle aussi consacrée à cette partie.



Les principaux objets de cette thèse sont les graphes. Nous donnons ici une définition intuitive de ces objets.

Définition : Graphe

Un graphe est un couple (V, E). V étant l'ensemble des noeuds, généralement de la forme $\{1, 2, ..., n\}$ où $n \geq 1$, ou $V = \mathbb{N}^*$ et dans ce cas le graphe est dit infini. Et E l'ensemble des arêtes qui sont des ensembles contenant deux noeuds distincts.

On représente généralement visuellement les graphes comme dans l'exemple de l'Illustration 1.1.

Voici quelques notions importantes pour la suite de cette thèse.

- On appelle extrémités d'une arête les noeuds la composant.
- Un chemin entre deux noeuds a et b dans un graphe est une suite d'arêtes $(\{i_1, j_1\}, \{i_2, j_2\}, ..., \{i_k, j_k\})$ avec $k \geq 1$, telle que $i_1 = a$, $j_k = b$, et pour tout $1 \leq l \leq k-1$, $j_l = i_{l+1}$.
- Un graphe est dit connexe s'il existe un chemin entre n'importe quelle paire de noeuds de ce graphe.
- On dit qu'un graphe G' est un sous-graphe, ou est contenu dans un autre graphe G si les ensembles de noeuds et d'arêtes de G' sont respectivement des sous-ensembles des ensembles de noeuds et d'arêtes de G.
- Une composante connexe d'un graphe est un sous-graphe connexe qui n'est strictement contenu dans aucun autre sous-graphe connexe.
- Un cycle dans un graphe est un chemin dont les noeuds d'arrivée et de départ sont les mêmes.
- La distance de graphe entre deux noeuds dans un graphe est le nombre minimum d'arêtes composant un chemin entre ces deux noeuds.
- Un arbre est un graphe connexe sans cycles.

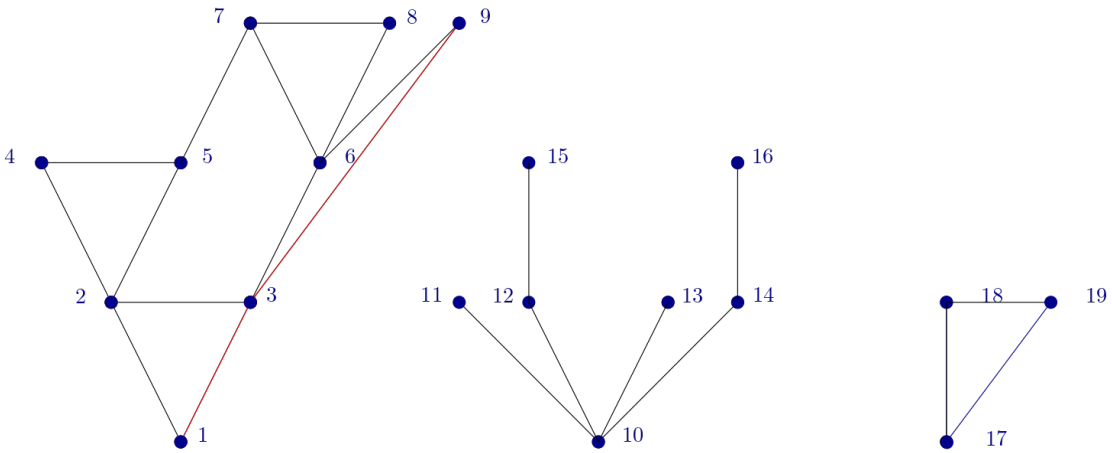


FIGURE 1.1 – Un exemple de graphe à 19 noeuds et 3 composantes connexes. La deuxième composante connexe en partant de la gauche est un arbre. La distance entre le noeud 1 et 9 est 2, et en rouge nous avons un chemin de longueur 2 entre ces deux noeuds.

1.1.2 Une brève histoire de l'arbre couvrant minimum

Dans les chapitres suivants de cette thèse, on s'intéresse à un type de graphe bien particulier : L'arbre couvrant de poids minimum (MST pour minimum spanning tree). Avant de présenter les variantes, aléatoires, et inhomogènes que nous traitons dans cette thèse. Nous donnerons dans cette partie une définition de cet arbre ainsi qu'une brève présentation de son histoire.

Définition : Arbre couvrant minimum

Soit $G = (V, E)$ un graphe connexe, on associe à chaque arête de ce graphe un poids qui est un réel positif. Un arbre couvrant minimum est un sous-graphe connexe de G ayant le même ensemble de noeuds V et qui minimise la somme des poids des arêtes.

Si G n'est pas connexe on parle alors de forêt couvrante minimum, c'est à dire un ensemble d'arbres couvrants minimums de ses différentes composantes connexes.

La première mention explicite des arbres couvrants minimums remonte à Borůvka [1926]. L'auteur de ce dernier considéra le problème de la construction d'un réseau électrique dans une ville de façon efficace. Plus généralement, le problème de l'arbre couvrant minimum a des applications immédiates dans les domaines des réseaux électriques, de communication, routiers, etc. Mais l'étude de cet objet a des intérêts supplémentaires. Comme on le verra dans la sous-section suivante, il existe des algorithmes simples et rapides qui permettent de trouver l'arbre couvrant minimum d'un graphe donné. Ceci en fait un candidat idéal pour approximer les solutions de problèmes plus compliqués, comme le problème du voyageur de commerce, ou encore celui du matching dans les graphes. Nous renvoyons à Graham and Hell [1985] pour plus de détails historiques.

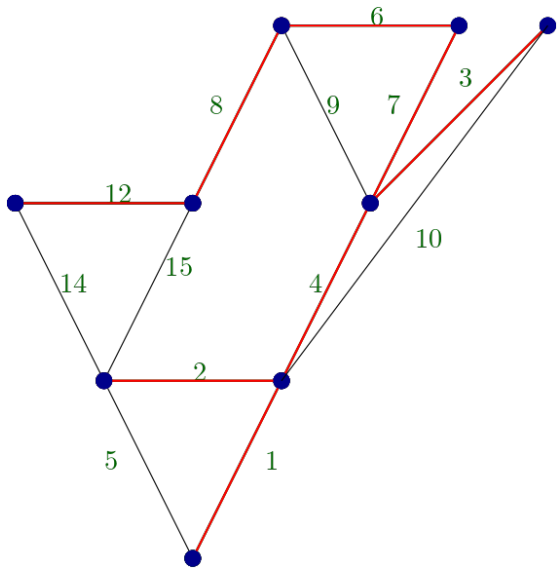


FIGURE 1.2 – Un exemple de graphe à 9 noeuds avec des poids (en vert) entre les arêtes. L'arbre couvrant minimum de ce graphe correspond aux arêtes en rouge.

C'est pour cela que l'arbre couvrant minimum était et reste encore aujourd'hui l'objet d'une recherche foisonnante, avec plusieurs algorithmes, variantes et problématiques étudiées. Parmi les problématiques abordées dès le début de ces recherches se trouve celle des arbres couvrants minimums aléatoires. Afin de mieux cerner la structure des arbres couvrants minimums, l'idée est apparue assez tôt d'introduire de l'aléa à certains niveaux de la construction. Le premier exemple d'une telle méthode remonte à Beardwood, Halton, and Hammersley [1959]. Les auteurs y ont

étudié ce qu'on appelle aujourd'hui **l'arbre couvrant minimum euclidien**. Considérons une mesure de probabilité μ absolument continue et à support borné sur $\mathbb{R}^d$. On tire indépendemment n points suivant cette mesure et on considère le graphe à n sommets qui contient toutes les arêtes possibles, et tel que le poids d'une arête $\{i, j\}$ est la distance euclidienne entre le i-ème et le j-ème point tiré. Il existe alors une constante c qui dépend de μ, telle que si X_n est le poids de l'arbre couvrant minimum alors :

$$\frac{X_n}{n^{(d-1)/d}} \to c,$$

où la convergence a lieu presque sûrement. Ce résultat donna lieu à un grand nombre d'articles autour de l'arbre couvrant minimum euclidien. Nous renvoyons ici le lecteur au livre de Yukich [1998] pour une liste plus exhaustive des résultats et méthodes autour des fonctionnelles d'ensembles de points plongés dans un espace euclidien.

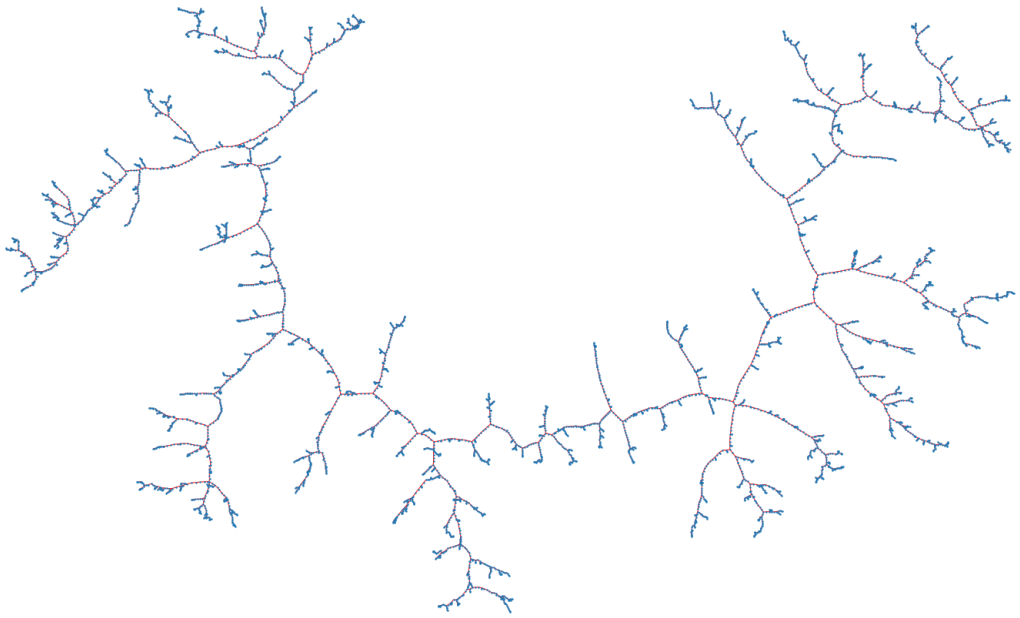


FIGURE 1.3 – Exemple d'arbre couvrant minimal euclidien correspondant à 5000 noeuds tiré uniformément dans $[0, 1]^2$.

Un autre objet tout aussi naturel qui sera étudié un peu plus tard est ce qu'on appellera simplement ici **l'arbre couvrant minimum aléatoire**. Considérons K_n le graphe complet à n noeuds. C'est le graphe qui contient toutes les arêtes possibles. Soit μ une mesure sans atomes sur $\mathbb{R}^+$. On associe aux arêtes de K_n des poids indépendants identiquement distribués (i.i.d.) suivant la mesure μ. L'arbre couvrant minimum de K_n associé à μ est ce qu'on appellera ici un arbre couvrant minimum aléatoire. Cet objet fut lui aussi largement étudié. Tout d'abord Frieze [1985] démontra que l'espérance du poids total $\mathbb{E}[X_n]$ de cet arbre converge vers une constante qui ne dépend que de μ à partir du moment où la fonction de répartition des poids est dérivable à droite de 0 et de dérivée positive. Ce résultat, surprenant en premier abord (on s'attendrait intuitivement à ce que le poids de l'arbre diverge), donna lui aussi lieu à des théorèmes similaires pour d'autres types de graphes de base au lieu du graphe complet (Beveridge, Frieze, and McDiarmid [1998], Frieze, Ruszinkó, and Thoma [2000], Frieze and McDiarmid [1989]). Pour ce même objet, Aldous (Aldous [1990]) donna la distribution asymptotique du degré du noeud 1 dans l'arbre couvrant minimum aléatoire. Janson (Janson [1995]) a quant à lui démontré un théorème central limite pour le poids total de l'arbre dans le cas où μ est la loi uniforme sur

$[0, 1]$. Plus précisément, il a démontré que $n^{1/2}(X_n - \zeta(3))$ converge en loi vers une loi normale (ζ étant la fonction zéta de Riemann).

Plus tard, un autre type de problèmes a été étudié autour du même objet. Ces problèmes concernent la structure "géométrique" de l'arbre couvrant minimum aléatoire. Addario-Berry, Broutin, and Reed [2006] démontrent que le diamètre, c'est à dire le plus long chemin en nombre d'arêtes dans l'arbre couvrant minimum est en moyenne de l'ordre de $n^{1/3}$. Puis cette idée est poussée plus loin par Addario-Berry, Broutin, Goldschmidt, and Miermont [2017b], où il est démontré que l'arbre couvrant minimum, vu comme espace métrique avec comme distance la distance de graphe habituelle renormalisée par $n^{1/3}$, converge dans un certain sens vers un espace métrique compact qui correspond intuitivement à un arbre continu. Nous expliquerons plus en détail dans la suite à quoi ressemble un tel arbre et nous donnerons le sens exact de cette convergence. Pour l'instant nous plaçons le travail de cette thèse dans la continuité historique de ces travaux. L'**arbre couvrant minimum inhomogène** qui nous intéresse dans cette thèse correspond à une généralisation naturelle de l'arbre couvrant minimum aléatoire. Dans la suite nous donnerons la construction de cet objet et expliquerons les motivations qui nous ont poussé à l'étudier.

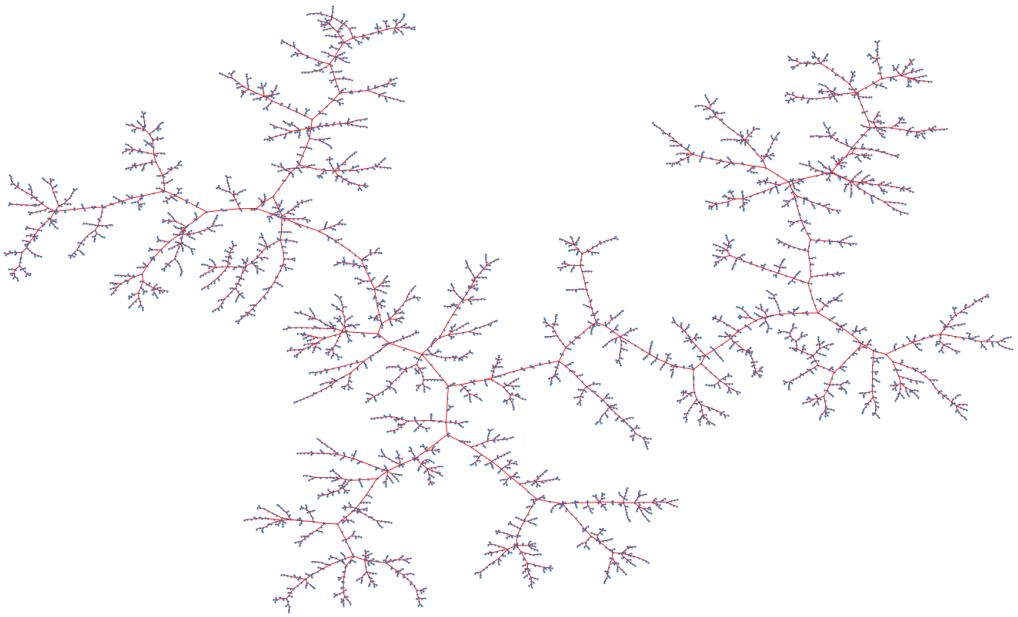


FIGURE 1.4 – Exemple d'arbre couvrant minimal aléatoire à 5000 noeuds.

1.1.3 Constructions et propriétés du MST

Dans la suite de cette partie nous allons présenter plusieurs algorithmes permettant de construire l' arbre couvrant minimum. Nous démontrons ici la correction de ces différents algorithmes. Ainsi nous ne referons pas ces démonstrations dans les chapitres ultérieurs. Mais tout d'abord, donnons une propriété importante des arbres couvrants minimums.

Unicité de l'arbre couvrant minimum

Si les poids d'un graphe G sont deux à deux distincts alors il existe une unique forêt couvrante minimum de G.

L'existence est une conséquence directe des algorithmes de construction de l'arbre couvrant minimum que nous présenterons après. Pour l'unicité, voici une preuve classique :

Preuve: Il suffit de démontrer le résultat pour un graphe connexe, on peut étendre l'unicité aux graphes non connexes simplement en considérant chacune de leurs composantes connexes à part. Soit $G = (V, E)$ un graphe connexe avec un ensemble de poids deux à deux distincts associés à ses arêtes. Supposons qu'il existe deux arbres couvrants minimums T_1 et T_2 de G. Soit e l'arête de poids minimal qui appartient à un des deux arbres mais pas à l'autre. Sans perte de généralité supposons que e appartient à T_1. Dans ce cas, ajouter e à T_2 créera un cycle C. Soit f l'arête de poids maximal dans ce cycle, si $e = f$ alors toutes les autres arêtes de C ont un poids plus petit que e. Dans ce cas, par définition de e, toutes les autres arêtes de C appartiennent aussi à T_1. Donc C tout entier appartient à T_1, ce qui est impossible. Alors, forcément $e \neq f$. Remplacer f par e dans T_2 donne un nouvel arbre T_3 couvrant de G, et dont la somme des poids des arêtes est strictement plus petite que celle de T_2. Ceci contredit le fait que T_2 est un arbre couvrant minimum. □

Ce théorème est intéressant car les graphes que nous considérerons dans cette thèse auront des poids deux à deux distincts presque sûrement. Passons maintenant à l'existence de tels arbres, nous détaillerons ici trois algorithmes de création d'arbres couvrants minimums. Ces trois algorithmes seront tous utilisés dans la suite de cette thèse. Chacun des trois algorithmes présente des avantages différents et une manière différente de voir l'arbre couvrant minimum. Dans la suite de cette sous-section on supposera donné un graphe connexe $G = (V, E)$ avec des poids deux à deux distincts associés à ses arêtes.

Algorithme de Kruskal (Kruskal [1956]) : L'algorithme de Kruskal construit l'arbre couvrant minimum par étapes. D'abord, on ordonne les arêtes de G par ordre croissant de leurs poids $(e_1, e_2, ...)$. On commence par un graphe $T_0 = (V, \emptyset)$. A l'étape $i \geq 1$ l'arête e_i est ajoutée au graphe T_{i-1} si et seulement si son ajout ne crée pas de cycle. On s'arrête après avoir considéré toutes les arêtes, le graphe obtenu à la fin est alors l'arbre couvrant minimum. Ainsi, à chaque pas, soit on ajoute une arête au graphe, soit on n'ajoute rien. On démontre de façon assez simple que cet algorithme produit bien l'arbre couvrant minimum. (Exemple dans l'Illustration 1.5)

Preuve: On démontre d'abord que l'algorithme de Kruskal produit bien un arbre couvrant. Le graphe produit par l'algorithme ne contient pas de cycles par construction. De plus, si à l'étape i, e_i relie deux composantes connexes disjointes dans T_{i-1}, alors e_i sera pris dans T_i. Ainsi l'arbre obtenu à la fin est connecté et il est clairement couvrant.

Maintenant supposons que cet arbre, noté T_∞, n'est pas l'arbre couvrant minimum. Il existe donc une arête e de poids minimal qui appartient à T_∞ mais pas à l'arbre couvrant minimum. L'ajout de e à l'arbre couvrant minimum crée donc un cycle C. A partir de là, la démonstration est la même que pour l'unicité. Soit f l'arête de poids maximal dans C, si $e = f$ alors toutes les autres arêtes de C ont un poids plus petit que e. Dans ce cas, par définition de e, toutes les autres arêtes de C appartiennent aussi à T_∞. Donc C tout entier appartient à T_∞, ce qui est impossible. Alors, forcément $e \neq f$. Remplacer f par e dans l'arbre couvrant minimum donne un nouvel arbre T couvrant de G, et dont la somme des poids des arêtes est strictement plus petite que celle de l'arbre couvrant minimum. □

L'algorithme de Kruskal a la particularité de construire l'arbre couvrant minimum en ajoutant ses arêtes par ordre croissant de leurs poids. Cependant, une autre propriété importante qui n'est pas garantie par l'algorithme de Kruskal est le fait de garder un arbre à chaque étape de la construction. Un autre algorithme classique et qui garantit cette propriété est le suivant.

Algorithme de Prim (Prim [1957]) : L'algorithme de Prim construit l'arbre couvrant minimum par étapes. On commence par un arbre T'_1 ne contenant que le noeud 1. On y ajoute l'arête de poids minimal ayant 1 comme une de ses extrémités pour obtenir T'_2. Ainsi on obtient un arbre composé de deux noeuds. Puis à chaque étape $i > 1$, on obtient T'_i en ajoutant l'arête

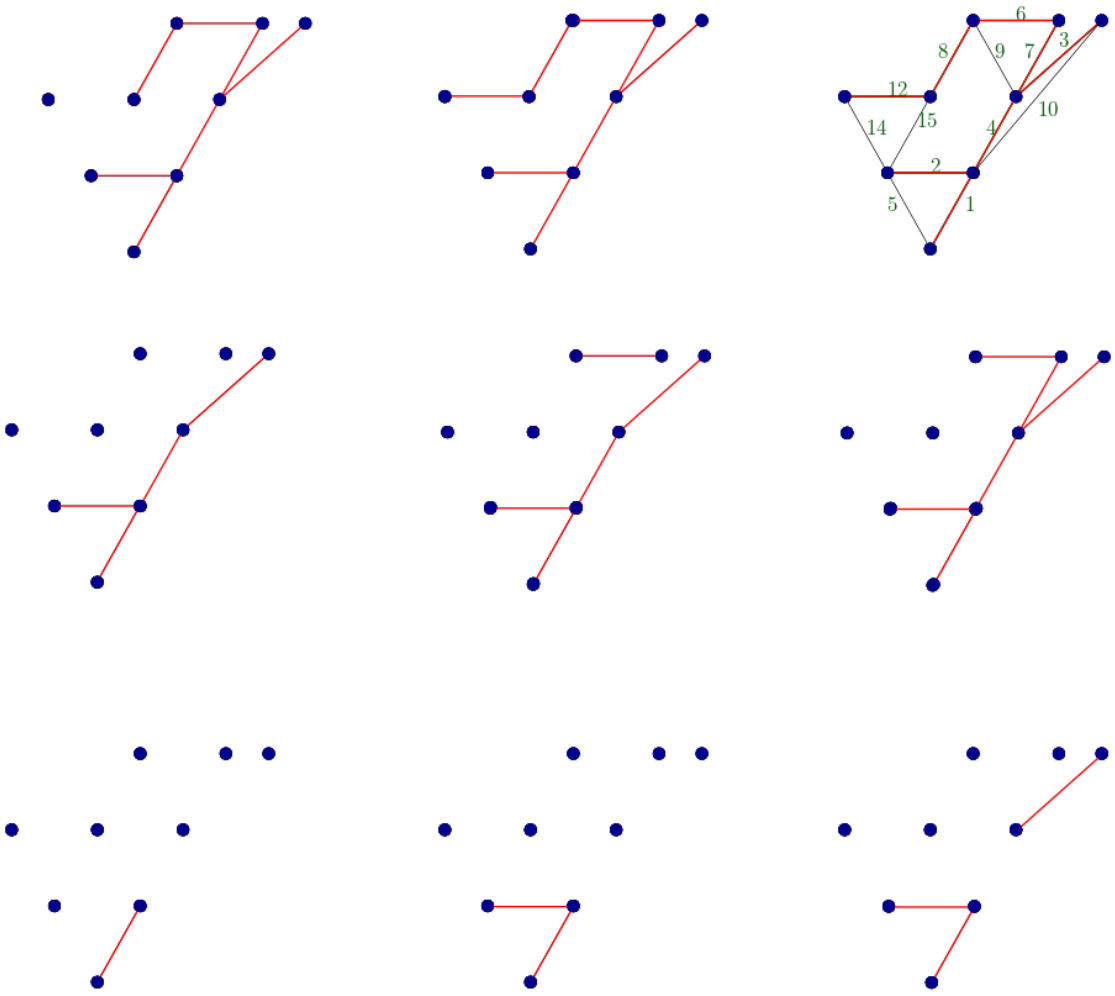


FIGURE 1.5 – Exemple d'application de l'algorithme de Kruskal. Seules les étapes où une arête est effectivement ajoutée sont représentées. Ordre de lecture de bas en haut et de gauche à droite.

de poids minimal ayant exactement une extrémité dans T'_{i-1} à ce dernier. Remarquez que T'_i est toujours un arbre de i noeuds. (Exemple dans l'Illustration 1.6)

Preuve: Il est clair que l'algorithme de Prim produit bien un arbre couvrant. Nous procédons par l'absurde pour montrer qu'il est minimum. Supposons que cet arbre ne soit pas l'arbre couvrant minimum. Soit $k \geq 2$ le plus petit entier tel que T'_k n'est pas un sous-arbre de l'arbre couvrant minimum. T'_{k-1} est un sous-arbre de l'arbre couvrant minimum, et donc l'arête e'_k ajoutée à l'étape k de l'algorithme n'est pas dans l'arbre couvrant minimum. Soit j l'extrémité de cette arrête qui n'est pas dans T'_{k-1}. Il existe un chemin dans l'arbre couvrant minimum qui relie un certain noeud l de T'_{k-1} à j. La première arête f de ce chemin a, par construction, un poids strictement plus grand que celui de e'_k. Soit T' l'arbre obtenu en remplaçant f par e'_k dans l'arbre couvrant minimum. Alors T' est aussi un arbre couvrant, et son poids total est strictement plus petit que celui de l'arbre couvrant minimum. □

Finalement on décrit un algorithme qui construit l'arbre couvrant minimum en enlevant des arêtes au graphe de base au lieu de partir d'un graphe sans arêtes et d'en ajouter de façon incrémentale.

Algorithme de suppression d'arêtes : L'algorithme de suppression d'arête part du graphe

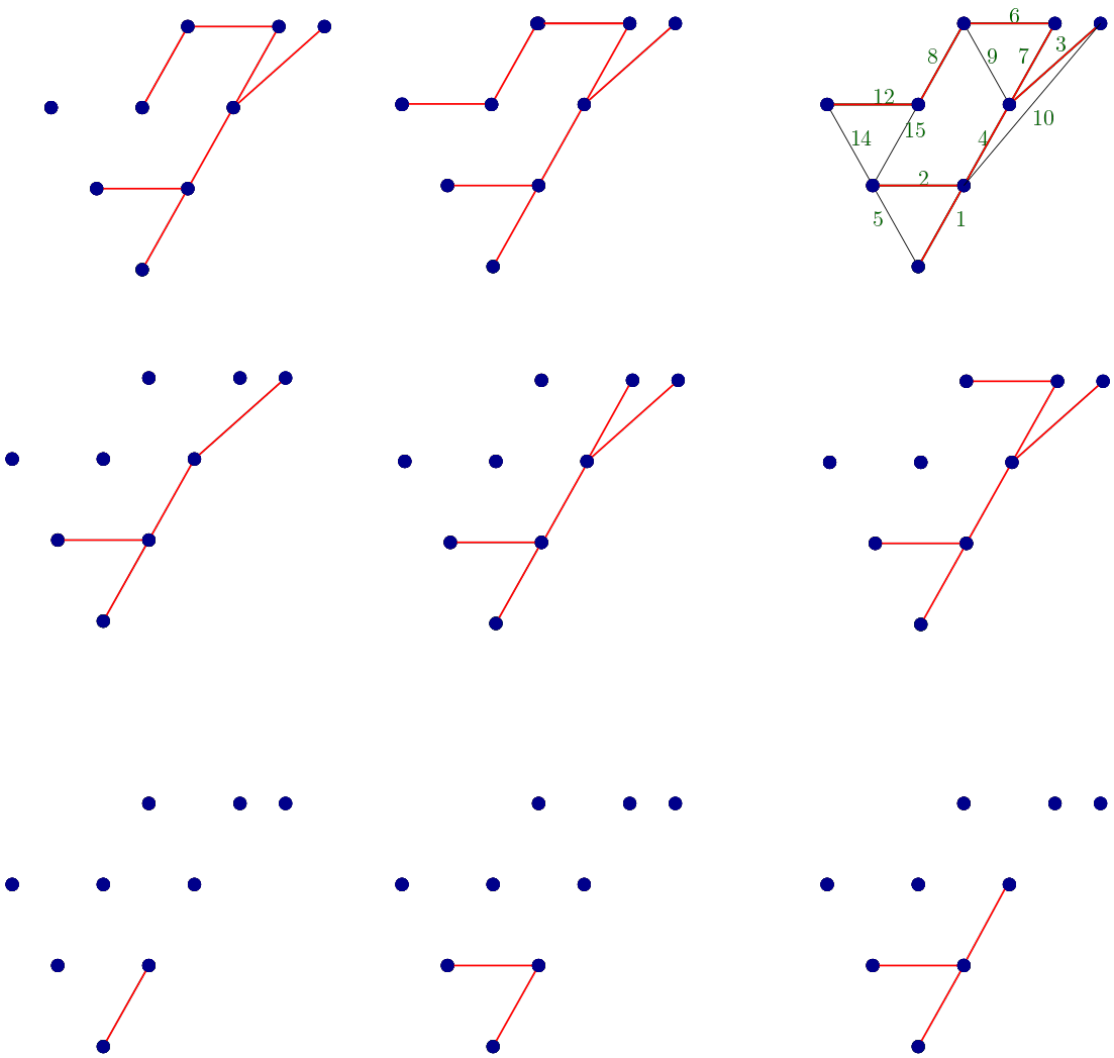


FIGURE 1.6 – Exemple d'application de l'algorithme de Prim. Ordre de lecture de bas en haut et de gauche à droite.

$\tilde{T}_1 = G$. On ordonne les arêtes de G par ordre décroissant de leurs poids $(\tilde{e}_1, \tilde{e}_2, ...)$. Pour passer de $\tilde{T}_i$ à $\tilde{T}_{i+1}$ on enlève l'arête $\tilde{e}_i$ de $\tilde{T}_i$ si et seulement si $\tilde{T}_i$ reste connecté sans l'arête $\tilde{e}_i$. (Exemple dans l'Illustration 1.1.3)

Preuve: Par construction le graphe obtenu par l'algorithme de suppression d'arêtes est connexe. De plus, si ce graphe devait contenir un cycle C, soit $\tilde{e}_j$ l'arête de poids maximal dans C. L'algorithme de suppression d'arête va supprimer $\tilde{e}_j$ et "casser" le cycle, car la suppression de cette arrête ne déconnecte pas le graphe. Ainsi nous avons démontré que le graphe obtenu par cet algorithme est bien un arbre couvrant. Nous procédons par l'absurde pour montrer qu'il est minimum. Supposons que cet arbre ne soit pas l'arbre couvrant minimum. Soit $\tilde{e}_i = \{a, b\}$ une arête qui se trouve dans l'arbre obtenu par l'algorithme de suppression d'arêtes mais qui n'appartient pas à l'arbre couvrant minimum. Si on ajoute $\tilde{e}_i$ au chemin entre a et b dans l'arbre couvrant minimum, on obtient un cycle C . Soit f l'arête de poids maximal dans ce cycle. Forcément $f \neq \tilde{e}_i$ sinon l'algorithme aurait supprimé $\tilde{e}_i$. L'arbre obtenu en enlevant f de l'arbre couvrant minimum et en y ajoutant $\tilde{e}_i$ est alors un arbre couvrant de poids strictement

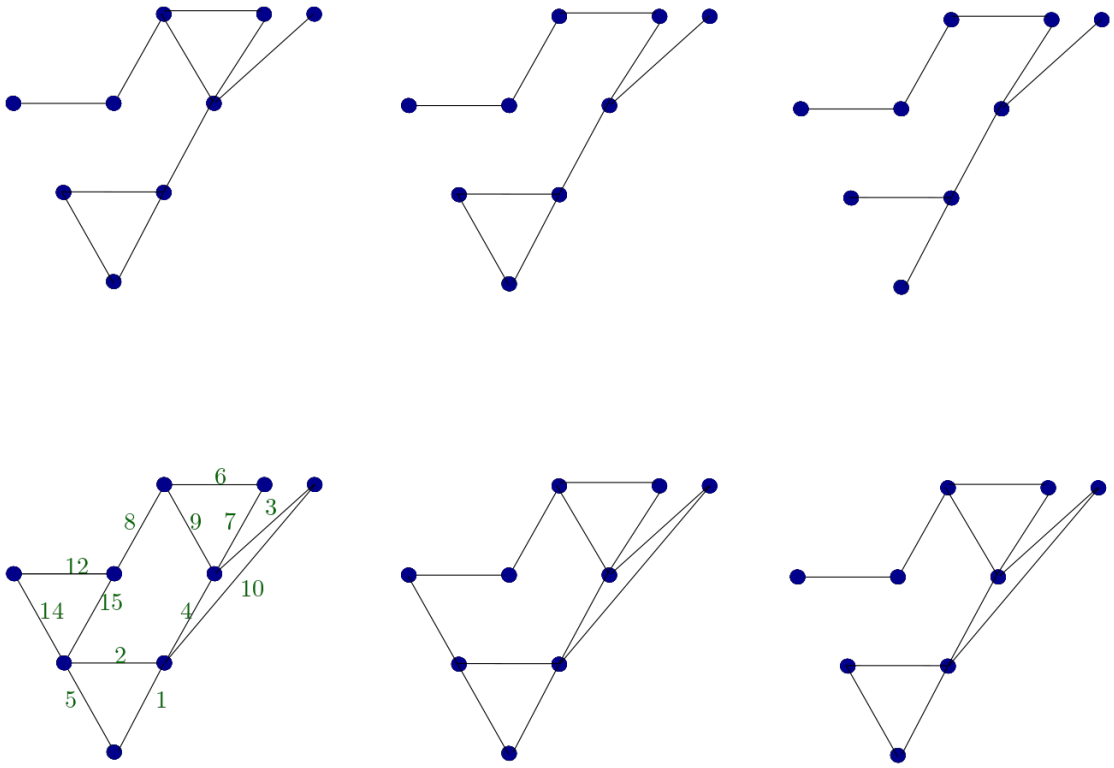


FIGURE 1.7 – Exemple d'application de l'algorithme de suppressions d'arêtes. Seules les étapes où une arête est supprimée sont illustrées. Ordre de lecture de bas en haut et de gauche à droite.

plus petit que celui de l'arbre couvrant minimum. □

1.2 Rappels et exemples d'inégalités de concentration

Le travail de cette thèse repose en grande partie sur de nouvelles utilisations des inégalités de concentrations. En effet, comme on le verra plus tard dans cette introduction, l'étude des graphes aléatoires peut souvent être ramenée à l'étude de processus aléatoires, discrets ou continus. C'est ainsi que plusieurs de nos résultats pourront être rapportés à l'étude d'inégalité de concentration. L'inégalité de concentration la plus connue est probablement l'inégalité de Markov.

Inégalité de Markov

Soit X une variable aléatoire positive, on a pour tout $\lambda > 0$:

$$\mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}[X]}{\lambda}.$$

Grâce à l'inégalité de Markov, on obtient la fameuse borne de Chernoff (Chernoff [1952]) :

Borne de Chernoff

Soit $t \geq 0$, et X une variable aléatoire réelle ayant une fonction génératrice des moments finie :

$$\phi(t) = \mathbb{E}[\exp(tX)] < \infty.$$

Alors pour tout $\lambda \geq 0$:

$$\mathbb{P}(X \geq \lambda) \leq \phi(t)e^{-t\lambda}.$$

En optimisant le paramètre t dans cette borne, on obtient les inégalités de concentration classiques comme l'inégalité de Hoeffding (Hoeffding [1963a]), l'inégalité d'Azuma (Azuma [1967]), on encore l'inégalité de Bernstein (Bernstein [1924]). Cette dernière sera utilisée intensivement dans ce document, nous en donnant donc l'énoncé exacte et une démonstration ici. Pour une revue globale sur les inégalités de concentration, nous renvoyons le lecteur vers le livre de Boucheron, Lugosi, and Massart [2013].

Inégalité de Bernstein

Soit $X_1, X_2, ..., X_n$ des variables aléatoires i.i.d. d'espérance m et de variance σ. Supposons de plus qu'il existe un $c > 0$ tel que $\mathbb{P}(|X_1| \leq c) = 1$. Posons

$$\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_i,$$

Alors pour tout $a \geq 0$:

$$\mathbb{P}(|\bar{X}_n - m| \geq a) \leq 2\exp\left(\frac{-na^2}{2\sigma^2 + 2ca/3}\right).$$

Preuve: Sans perte de généralité, supposons que $m = 0$. Pour $t \geq 0$, posons

$$F(t) = \sum_{r=2}^{\infty} \frac{t^{r-2}\mathbb{E}[X_1^r]}{r!\sigma^2}.$$

On a alors :

$$\mathbb{E}[e^{tX_1}] = \mathbb{E}\left[1 + tX_1 + \sum_{r=2}^{\infty} \frac{t^r X_1^r}{r!}\right] = 1 + t^2\sigma^2 F(t) \leq e^{t^2\sigma^2 F(t)}.$$

De plus, pour tout $r \geq 2$ on a $\mathbb{E}[X_1^r] = \mathbb{E}[X_1^{r-2}X^2] \leq c^{r-2}\sigma^2$. Par conséquent :

$$F(t) \leq \sum_{r=2}^{\infty} \frac{t^{r-2}c^{r-2}\sigma^2}{r!\sigma^2} = \frac{1}{(tc)^2}\sum_{r=2}^{\infty}\frac{(tc)^r}{r!} = \frac{e^{tc}-1-tc}{(tc)^2}.$$

On obtient donc l'inégalité suivante :

$$\mathbb{E}[e^{tX_1}] \leq \exp\left(t^2\sigma^2\frac{e^{tc}-1-tc}{(tc)^2}\right). \tag{1.1}$$

Par la borne de Chernoff et l'équation (1.1) on a :

$$\begin{aligned}\mathbb{P}(\bar{X}_n \geq a) &\leq e^{-tna}\mathbb{E}\left[\exp\left(t\sum_{k=1}^{n}X_k\right)\right]\\ &\leq e^{-tna}\prod_{k=1}^{n}\left(\mathbb{E}[e^{tX_i}]\right)\\ &\leq e^{-tna}\exp\left(nt^2\sigma^2\frac{e^{tc}-1-tc}{(tc)^2}\right).\end{aligned}$$

En prenant $t = (1/c)\log(1+ac/\sigma^2)$ on obtient :

$$\mathbb{P}(\bar{X}_n \geq a) \leq \exp\left(\frac{-n\sigma^2}{c^2}h\left(\frac{ca}{\sigma^2}\right)\right),$$

avec $h(u) = (1+u)\ln(1+u) - u$. Finalement, une étude rapide de la fonction $u \mapsto h(u)$ permet de voir que

$$h(u) \geq \frac{u^2}{2+2u/3},$$

pour tout $u \geq 0$. Ceci termine la preuve de la borne supérieure. La borne inférieure se démontre de façon similaire en considérant $-\bar{X}_n$ au lieu de $\bar{X}_n$. □

Remarquez le rôle important joué par la borne de Chernoff dans cette démonstration. Grâce à ce constat nous pouvons obtenir plusieurs variantes de cette inégalité, la première que nous présentons ici sous le nom d'inégalité de Bernstein généralisée est la suivante.

Inégalité de Bernstein généralisée

Soit $X_1, X_2, \ldots$ des variables aléatoires i.i.d. d'espérance m et de variance σ. Supposons de plus qu'il existe un $c > 0$ tel que $\mathbb{P}(|X_1| \leq c) = 1$. Pour tout $n \geq 1$, posons

$$\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n}X_i,$$

Alors pour tout $a \geq 0$:

$$\mathbb{P}\left(\max_{k\leq n}\left(|k\bar{X}_k - km|\right) \geq na\right) \leq 2\exp\left(\frac{-na^2}{2\sigma^2+2ca/3}\right).$$

Preuve: Sans perte de généralité, supposons que $m = 0$. Dans ce cas le processus $(n\bar{X}_n)_{n\geq 1}$ est une martingale. Pour $t \geq 0$, par inégalité de Jensen le processus $(e^{tn\bar{X}_n})_{n\geq 1}$ est donc une sous-martingale. Par inégalité de Doob pour les martingales (Revuz and Yor [1999] Theorème *II*.1.7) on a alors :

$$\mathbb{P}\left(\max_{k\leq n}\left(e^{tk\bar{X}_k}\right) \geq e^{nta}\right) \leq e^{-nta}\mathbb{E}\left[e^{tn\bar{X}_n}\right].$$

A partir de là, il suffit de reprendre la démonstration de l'inégalité de Bernstein pour finir la preuve de la borne supérieure. Celle de la borne inférieure se démontre de la même façon en remplaçant $(X_n)_{n\geq 1}$ par $(-X_n)_{n\geq 1}$. □

Malheureusement pour nous, les variables aléatoires que nous étudions dans cette thèse ne sont pas i.i.d.. Elles correspondent à des tirages sans remise parmi une liste donnée de réels positifs. Il existe plusieurs résultats de concentration pour ce type de tirages, par exemple Hoeffding [1963b] et Serfling [1974]. Mais, double peine, ces résultats ne sont vrais que pour des tirages uniformes. Hors, les tirages que nous étudions dans cette thèse ne le sont pas. Il existe cependant un résultat dû à Ben-Hamou, Peres, and Salez [2018] qui permet de déduire des inégalités de concentration pour de tels tirages.

Borne de Chernoff pour les tirages biaisés sans remises

Soit $w_1 \geq w_2 \geq ... \geq w_n \geq 0$ des "poids" positifs. Soit $1 \geq p_1 \geq p_2 \geq ... \geq p_n \geq 0$ des probabilités de tirages, on a donc $\sum_{k=1}^{n} p_k = 1$. Soit $m \leq n$ et $(X_1, X_2, ..., X_m)$ des poids tirés avec remise parmi les w_i suivant les probabilités p_i, et soit $(Y_1, Y_2, ..., Y_m)$ des poids tirés sans remises parmi les w_i suivant les probabilités p_i. Pour tout fonction croissante et convexe f on a alors :

$$\mathbb{E}\left[f\left(\sum_{i=1}^{m} Y_i\right)\right] \leq \mathbb{E}\left[f\left(\sum_{i=1}^{m} X_i\right)\right].$$

Comme la fonction exponentielle est croissante et convexe, ce théorème montre que l'inégalité de Bernstein qui est vraie pour les tirages avec remise va rester vraie pour les tirages sans remise. Cependant, ce résultat pose deux problèmes. Tout d'abord, il ne permet d'avoir qu'une borne supérieure, car en effet si on veut considérer les tirages opposés pour avoir la borne inférieure, la condition de décroissance des probabilités n'est plus vraie. Ensuite, il n'est en général pas vrai que

$$\mathbb{E}\left[\sum_{i=1}^{m} Y_i\right] = \mathbb{E}\left[\sum_{i=1}^{m} X_i\right].$$

Comme démontré dans le Lemme 35, on aura généralement

$$\mathbb{E}\left[\sum_{i=1}^{m} Y_i\right] < \mathbb{E}\left[\sum_{i=1}^{m} X_i\right],$$

et donc même la borne supérieure donnée par le théorème de Ben-Hamou et al. [2018] n'est pas assez précise vu qu'elle sert à comparer des variables aléatoires qui n'ont pas la même moyenne. C'est pour cela qu'une partie de cette thèse sera consacrée à la démonstration d'inégalités de concentration plus fines pour ces tirages biaisés sans remises dans le cas particulier qui nous intéresse.

1.3 Les graphes aléatoires

1.3.1 Arbres de Galton-Watson

Bien que le point de vue qui consiste à voir les arbres comme des graphes sans cycles soit le plus "naturel". L'utilisation des notation de Neveu (Neveu [1986]) est plus commode quand il s'agit de traiter des arbres de Galton-Watson. Nous introduisons ici cette notation, avant de définir les arbres de Galton-Watson.

La notation de Neveu permet de définir les arbres plans enracinés. Un tel arbre est un sous-ensemble T de l'ensemble des suites finies d'entiers :

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

qui vérifie les trois conditions suivantes :

- La suite vide est dans T, $\emptyset \in T$.
- Si $(v_1, v_2, ..., v_n) \in T$ alors $(v_1, v_2, ..., v_{n-1}) \in T$.
- Pour tout $u = (v_1, v_2, ..., v_n) \in T$ il existe $c(u) \in \mathbb{N} \cup \{\infty\}$ tel que pour tout $1 \leq j \leq c(u)$ on a $(v_1, v_2, ..., v_n, j) \in T$.

Ainsi, on peut voir les éléments de T comme des noeuds. $\emptyset$ est la racine. Et si $u = (v_1, v_2, ..., v_n) \in T$ est un noeud, alors l'ensemble $((v_1, v_2, ..., v_n, j))_{1\leq j\leq c(u)}$ correspond aux enfants de u. Pour deux éléments $(u, v) \in \mathcal{U}^2$, on écrit uv pour la concaténation des deux suites. Soit $T_u := \{v \in \mathcal{U}/uv \in T\}$ l'arbre décalé de u. Remarquez que T_u est lui même un arbre plan enraciné. La génération d'un noeud $u = (v_1, v_2, ..., v_n) \in T$, dénoté par $|u|$, est n. La taille de T est son nombre de noeuds. La hauteur de T est la plus grande génération d'un noeud de T moins un. Concrètement, la notation de Neveu établi un ordre partiel sur les noeuds d'un arbre vu comme un graphe. Cet ordre respecte la hiérarchie parent-enfant, un enfant vient toujours après son père. La racine de l'arbre vient donc avant tous les autres noeuds. Notons $\mathbb{T}$ l'ensemble des arbres plans enracinés.

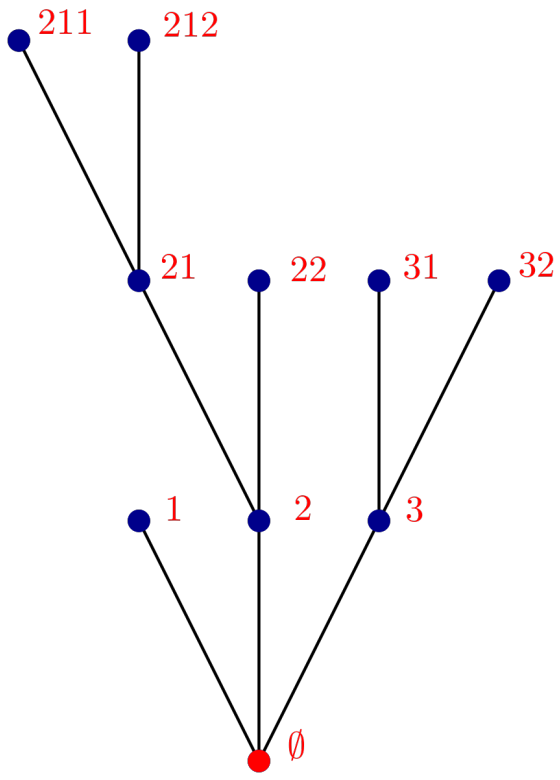


FIGURE 1.8 – Un arbre de taille 10 et de hauteur 3 avec sa notation de Neveu. Le noeud rouge représente la racine, et on ordonne par convention les noeuds d'une même hauteur de gauche à droite.

Nous pouvons maintenant donner la définition d'un arbre de Galton-Watson.

Arbres de Galton-Watson

Soit μ une loi de probabilité sur $\mathbb{N}$. Il existe une unique loi de probabilité $\mathbb{P}_\mu$ sur $\mathbb{T}$, telle que :

- Pour tout $j \in \mathbb{N}$, $\mathbb{P}_\mu(c(\emptyset) = j) = \mu(\{j\})$.
- Conditionnellement à $c(\emptyset) = j$, les arbres $(T_{\{k\}})_{1 \leq k \leq j}$ sont i.i.d. de loi $\mathbb{P}_\mu$.

Un arbre aléatoire généré suivant $\mathbb{P}_\mu$ est appelé arbre de Galton-Watson de loi de reproduction μ.

Généralement on ne considère pas le cas trivial $\mu(\{1\}) = 1$. Cette définition signifie que chaque noeud de l'arbre a un nombre d'enfants aléatoire suivant la loi μ indépendamment de tout le reste.

Les arbres de Galton-Watson revêtent une importance historique pour plusieurs raisons. Tout d'abord, c'est le premier modèle d'arbres aléatoires qui a été étudié (Watson and Galton [1875], même si ces derniers n'ont étudié que le processus des tailles des générations). De plus, ce modèle bien que simple se rapporte à plusieurs autres modèles d'arbres aléatoires en apparence plus compliqués (par exemple Camarri and Pitman [2000]). La simplicité du modèle de Galton-Watson, et surtout l'indépendance qu'il y a entre les noeuds permet une étude exhaustive de plusieurs propriétés qu'il serait difficile d'étudier pour des modèles d'arbres plus compliqués. Finalement, les arbres de Galton-Watson apparaissent naturellement dans l'étude de différents modèles de graphes aléatoires. Nous reviendrons sur cette remarque plus tard.

La première question étudiée pour les arbres de Galton-Watson est celle de l'extinction. Sous quelles conditions sur la loi μ l'arbre de Galton-Watson peut-il être infini avec probabilité positive ? Cette question a rapidement été résolue. Nous renvoyons le lecteur à Athreya and Ney [1972] pour une preuve du résultat suivant ainsi que de plusieurs théorèmes élémentaires sur les processus de branchement.

Extinction de l'arbre de Galton-Watson

Soit μ une loi de probabilité sur $\mathbb{N}$ et T un arbre de Galton-Watson de loi de reproduction μ. Soit $m = \sum_{k=0}^{\infty} k\mu(\{k\})$. Alors :

- Si $m < 1$, alors avec probabilité 1, T a une taille finie. On appelle cela le régime sous-critique.
- Si $m > 1$, alors avec probabilité strictement positive, T a une taille infinie. On appelle cela le régime sur-critique.
- Si $m = 1$, alors avec probabilité 1, T a une taille finie, sauf si $\mu(\{1\}) = 1$. On appelle cela le régime critique.

Nous voyons ainsi qu'il existe une transition de phase en $m = 1$. En cette transition de phase, les propriétés géométriques de l'arbre de Galton-Watson changent. Par exemple, si $\sigma^2 = \sum_{k=0}^{\infty} k^2\mu(\{k\}) < \infty$, alors si $m < 1$ la hauteur moyenne de T est finie. Mais si $m = 1$, l'espérance de la hauteur de T devient infinie. Cette transition de phase est d'autant plus importante qu'elle apparaît aussi dans les graphes aléatoires que nous étudierons, pour des raisons similaires que nous expliquerons plus tard.

1.3.2 Comment explorer les graphes

Afin d'étudier les graphes et arbres aléatoires, nous allons passer par des fonctions qui les codent. Cette idée a été est utilisée intensivement dans l'étude des graphes et arbres aléatoires (Aldous [1993], Aldous [1997], Aldous and Limic [1998], Duquesne and Le Gall [2002], Stegehuis, van der Hofstad, Janssen, and van Leeuwaarden [2017], Addario-Berry et al. [2017b],

Addario-Berry [2019] ...). Coder un graphe par une fonction permet l'utilisation de tout l'arsenal probabiliste connu pour les processus aléatoires, propriété de Markov, théorie des martingales, inégalités de concentration etc ... Ce changement de point de vue n'est cependant pas gratuit, chaque type de fonction codante présente des avantages et des inconvénients, et il est souvent nécessaire de jongler entre les différents codages pour arriver au résultat escompté. Nous présenterons ici deux types de codage, et une fonction qui n'est pas formellement un codage mais qui s'en rapproche. Ainsi que les deux types d'exploration utilisés principalement avec ces codages. Cette partie ne suit pas les normes et conventions d'usage général. En effet, nous estimons qu'il est plus naturel et utile de dissocier le codage de l'algorithme d'exploration. Un algorithme d'exploration est une méthode qui permet d'ordonner les noeuds d'un graphe. Un codage est une fonction qui à un graphe avec un ordre sur les noeuds renvoie une fonction réelle des noeuds du graphe. Commençons par les explorations. Dans cette partie l'ordre d'un ensemble sera représenté par une suite finie composée des éléments de cet ensemble sans répétitions. L'ordre total sous-jacent s'obtient alors facilement en considérant l'ordre des éléments dans la suite.

Exploration en largeur, breadth-first search

Etant donné un graphe $G = (V, E)$ connexe, l'exploration en largeur construit une suite finie constituée des noeuds de G de la façon suivante. On définit trois ensembles :

— $\mathcal{U}$ l'ensemble des noeuds non découverts.
— $\mathcal{D}$ l'ensemble des noeuds découverts mais non explorés.
— $\mathcal{E}$ l'ensemble des noeuds explorés.

- Au début, on choisit un noeud de manière quelconque pour être le premier noeud $v(1)$. On commence avec $\mathcal{D} = \{v(1)\}$, $\mathcal{U} = V \setminus \{v(1)\}$ et $\mathcal{E} = \emptyset$.
- A l'étape 1, on déplace les voisins de $v(1)$ dans $\mathcal{U}$ vers $\mathcal{D}$ et on déplace $v(1)$ de $\mathcal{D}$ vers $\mathcal{E}$. L'ordre d'ajout des voisins de $v(1)$ peut lui aussi être quelconque.
- On continue ainsi : A l'étape i, on déplace le i-ème noeud ajouté à $\mathcal{D}$ vers $\mathcal{E}$ et on déplace ses voisins présents dans $\mathcal{U}$ vers $\mathcal{D}$.
- L'algorithme se termine quand $\mathcal{E} = V$ et on obtient une suite ordonnée $(v(1), v(2), ..., v(n))$ des noeuds du graphe.

Dans les chapitres suivants nous considérerons l'exploration en largeur en ajoutant les noeuds $\mathcal{D}$ suivant un ordre donné par des poids associés aux noeuds. Remarquez que la manière dont les noeuds sont retirés de l'ensemble $\mathcal{D}$ correspond à une liste FIFO (first in first out). Si au lieu on considère une list LIFO (Last in first out), on obtient l'exploration en profondeur.

Exploration en profondeur, depth-first search

Etant donné un graphe $G = (V, E)$ connexe, l'exploration en profondeur construit un ordre sur les noeuds de G de la façon suivante. On commence avec trois ensembles :

— $\mathcal{U}$ l'ensemble des noeuds non découverts.
— $\mathcal{D}$ l'ensemble des noeuds découverts mais non explorés.
— $\mathcal{E}$ l'ensemble des noeuds explorés.

- Au début, on choisit un noeud de manière quelconque pour être le premier noeud $v(1)$. On commence avec $\mathcal{D} = \{v(1)\}$, $\mathcal{U} = V \setminus \{v(1)\}$ et $\mathcal{E} = \emptyset$.
- A l'étape 1, on déplace les voisins de $v(1)$ dans $\mathcal{U}$ vers $\mathcal{D}$ et on déplace $v(1)$ de $\mathcal{D}$ vers $\mathcal{E}$. L'ordre d'ajout des voisins de $v(1)$ peut lui aussi être quelconque.
- On continue ainsi : A l'étape i, on déplace le dernier noeud ajouté à $\mathcal{D}$ (donc dans l'ordre inverse des ajouts) vers $\mathcal{E}$ et on déplace ses voisins dans $\mathcal{U}$ vers $\mathcal{D}$.
- L'algorithme se termine quand $\mathcal{E} = V$ et on obtient une suite ordonnée $(v(1), v(2), ..., v(n))$ des noeuds du graphe.

Il est utile de remarquer que dans le cas des arbres plans enracinés, avec la notation de Neveu, l'ordre de l'exploration en profondeur peut être obtenu en considérant l'ordre lexicographique sur

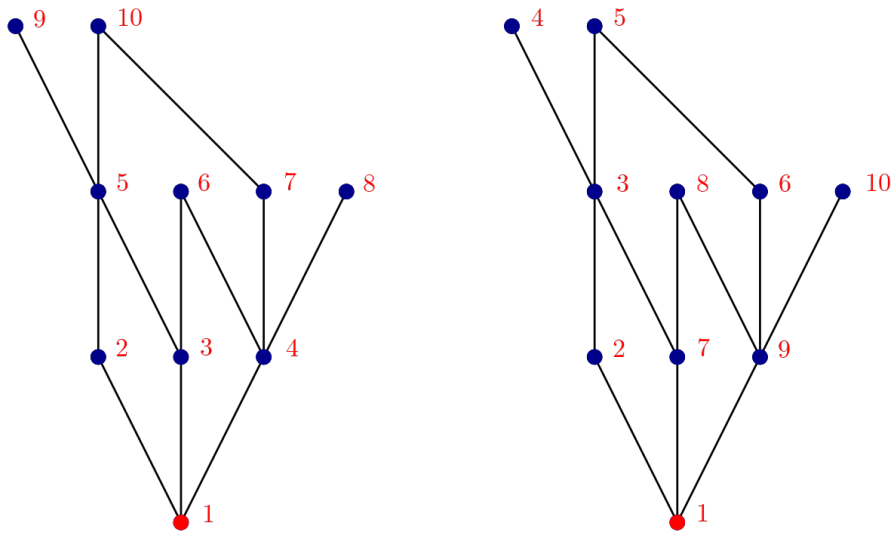


FIGURE 1.9 – Un graphe de taille 10 et de hauteur 3 avec, à gauche son parcours en largeur partant du noeud rouge. A droite son parcours en profondeur partant du noeud rouge. On ajoute par convention les noeuds de gauche à droites.

les noeuds. Ces deux explorations peuvent être généralisées au cas où G n'est pas connexe en ordonnant d'abord les composantes connexes de G de façon quelconque, puis en ordonnant chaque composante connexe seule et en concaténant les ordres des différentes composantes connexes. Passons maintenant aux codages. Pour tous les codages, nous supposerons donné un graphe $G(V, E)$ avec un réordonnement $(v(1), v(2), ..., v(n))$ des noeuds. On appelle codage une fonction ou "méthode" qui à un couple (graphe, ordre), renvoie une suite de réels. (Exemple dans l'illustration 1.9)

Codage par comptage d'enfants

On pose $L(0) = 1$. Le codage par comptage d'enfants se base lui aussi sur les ensembles suivants :

— $\mathcal{U}$ l'ensemble des noeud non découverts.
— $\mathcal{D}$ l'ensemble des noeuds découverts mais non explorés.
— $\mathcal{E}$ l'ensemble des noeuds explorés.

On procède ainsi :

- L'élément 1 du codage, $L(1)$, est égal au nombre de voisins du noeuds $v(1)$ moins 1.
- On déplace les voisins de $v(1)$ dans $\mathcal{U}$ vers $\mathcal{D}$ et on déplace $v(1)$ de $\mathcal{D}$ vers $\mathcal{E}$.
- On continue ainsi : A chaque étape, i, on traite $v(i)$. On pose $L(i) - L(i-1)$ égale au nombre de voisins de $v(i)$ dans $\mathcal{U}$ moins 1. On déplace $v(i)$ vers $\mathcal{E}$ et on ajoute ses voisins dans $\mathcal{U}$ à $\mathcal{D}$.
- On traite ainsi tous les noeuds suivant l'ordre de G jusqu'à avoir $\mathcal{E} = V$.

Voir l'Illustration 1.10 pour un exemple de ce codage. Pour un graphe connexe, le codage par comptage d'enfants se termine toujours en 0. Historiquement, le codage par comptage d'enfants utilisé avec un ordre issu du parcours en profondeur est appelé marche de Łukasiewicz (souvent dans la litérature cette marche est décalée pour commencer en 0 au lieu de 1). Dans les chapitres 2 et 3, nous utilisons ce codage avec un parcours en largeur particulier, et nous l'appelons simplement "exploration process" ou "exploration process associated to a breadth-first search". Le codage suivant est plus naturel. Dans la littérature, il est exclusivement utilisé pour les arbres.

Codage par fonction de hauteur

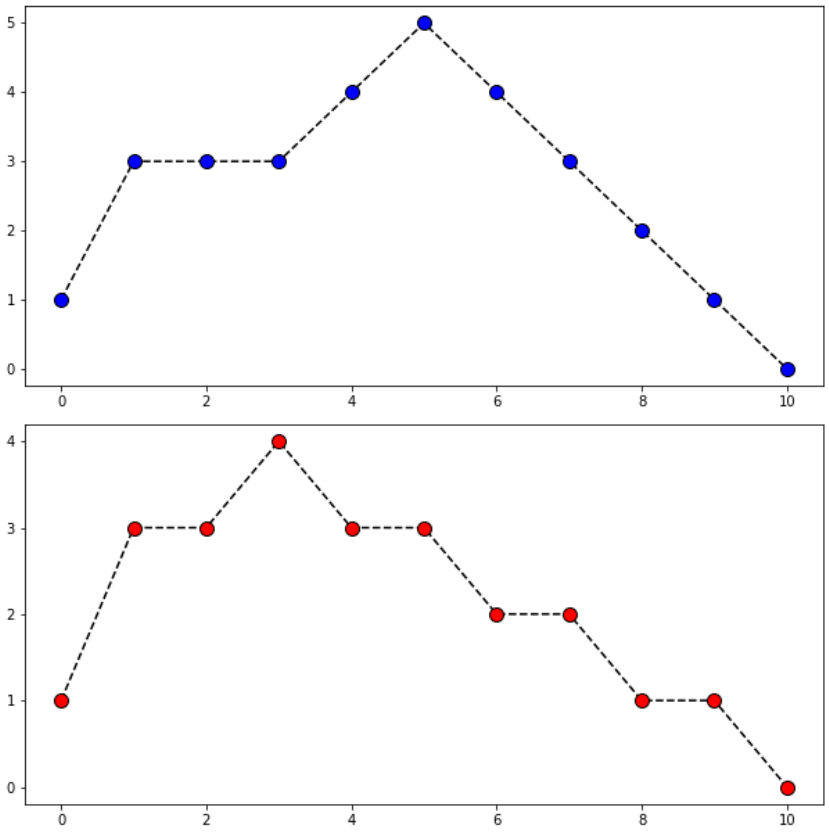


FIGURE 1.10 – Codage par comptage d'enfants associé aux parcours du graphe de la Illustration 1.9. Le premier codage correspond donc à un parcours en largeur, tandis que le second correspond à un parcours en profondeur.

Le codage par fonction de hauteur est définit par, pour $i \geq 1$, $H(i)$ est égale à la distance dans le graphe G entre $v(i)$ et $v(1)$. Voir les Illustrations 1.11 et 1.12 pour un exemple de ce codage.

Le codage par fonction de hauteur est généralement utilisée en association avec l'ordre du parcours en profondeur. Dans ce cas, il est simplement appelé fonction de hauteur. Il existe un lien direct entre la fonction de hauteur et la marche de Łukasiewicz.

Relation entre la fonction de hauteur et la marche de Łukasiewicz

Pour tout i tel que $n \geq i \geq 1$, on a dans le cas de la fonction de hauteur et de la marche de Łukasiewicz :

$$H(i) = \text{Card}\left\{ j < i-1,\, L(j) = \inf_{j \leq k \leq i-1} (L(k)) \right\}.$$

Nous renvoyons le lecteur à Le Gall and Le Jan [1998] (Corollaire 2.2), pour une démonstration de ce théorème.

Finalement il existe une autre fonction associée aux arbres, cette fonction ne correspond pas à un codage avec nos définitions. Mais elle est généralement étudiée en parallèle de la fonction de hauteur et de la marche de Łukasiewicz.

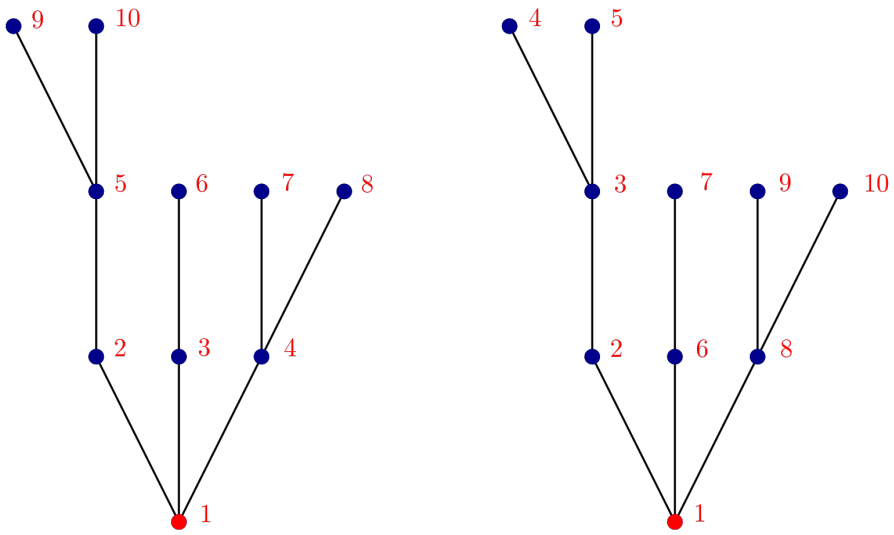


FIGURE 1.11 – Un arbre de taille 10 et de hauteur 3 avec, à gauche son parcours en largeur partant du noeud rouge. A droite son parcours en profondeur partant du noeud rouge. On ajoute par convention les noeuds de gauche à droites.

La fonction de contour

Nous ne considérons la fonction de contour que pour les arbres plans enracinés finis. Soit T un arbre plan enraciné. La fonction de contour se construit en suivant une sorte d'exploration en profondeur. La fonction de contour C note les hauteurs des noeuds de l'arbre à chaque étape. Comme pour la fonction de hauteur, si un noeud u a des enfants non découverts alors on explore un de ces enfants. Cependant, dans la fonction de contour, si le noeud u qui vient d'être exploré n'a pas d'enfants non découverts, alors on revient au parent de u dans l'arbre et on note la hauteur de ce dernier.

Plus formellement, on dit que le noeud a est ancêtre de b s'il existe une suite de noeuds $(u_1, u_2, ..., u_k)$ telle que $u_1 = a$, $u_k = b$, et pour tout i, $i \leq k-1$, u_{i+1} est un enfant de u_i. Notons $u \cap v$ l'ancêtre commun des noeuds u et v qui a la plus grande hauteur. Soit $(f_1, f_2, ..., f_k)$ les feuilles de T. C'est à dire les noeuds de T qui n'ont pas d'enfants. La fonction de contour de T est la seule fonction de $[0, 2(n-1)]$ à valeurs dans $\mathbb{R}^+$, affine par morceaux, dont les pentes sont dans $\{-1, 1\}$, et dont les valeurs des extremas locaux sont successivement $0, |f_1|, |f_1 \cap f_2|, |f_2|, |f_2 \cap f_3|, ..., |f_{k-1} \cap f_k|, |f_k|, 0$. Voir l'Illustration 1.13 pour un exemple de ce codage.

La marche de Łukasiewicz, la fonction de hauteur et la fonction de contour sont reliées. D'abord car chacune de ces fonctions est une bijection avec l'ensemble des arbres plans enracinés vu comme des graphes ordonnés. Mais aussi grâce au théorème de passage de la marche de Łukasiewicz vers la fonction de hauteur présenté plus haut. Il est aussi possible de passer de la fonction de hauteur à la fonction de contour, nous renvoyons ici vers les résultats de Marckert and Mokkadem [2003] pour une explication des différentes relations entre ces trois fonctions.

1.3.3 Codages d'arbres de Galton-Watson

Dans le cas des arbres de Galton-Watson. Le codage par comptage d'enfants associé à une exploration en largeur est une marche aléatoire à incréments indépendants. Cette marche aléatoire a la même loi que la marche de Łukasiewicz. Plus formellement, soit μ une loi de probabilité

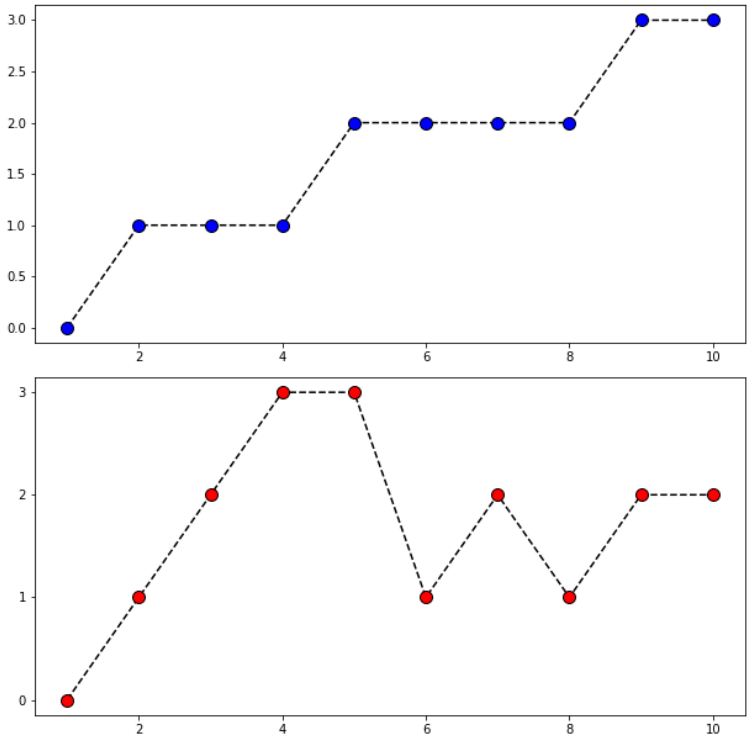


FIGURE 1.12 – Codages par les fonctions de hauteur associées aux parcours du graphe de l'Illustration 1.11. Le premier codage correspond donc à un parcours en largeur, tandis que le second correspond à un parcours en profondeur.

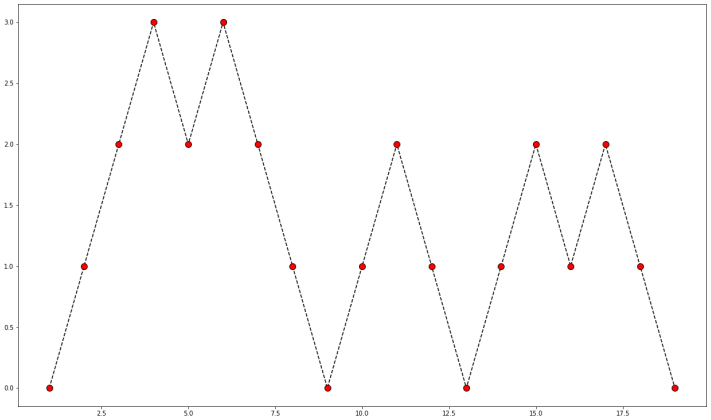


FIGURE 1.13 – La fonction de contour associé à l'arbre de la figure 1.11 munit de son parcours en profondeur présent dans cette même figure.

sur $\mathbb{N}$. Soit T un arbre de Galton-Watson de loi de reproduction μ, vu en tant que graphe, et soit L sa marche de Łukasiewicz. Soit μ^* la loi décalée de 1, c'est à dire pour tout $i \geq -1$:

$$\mu^*(\{i\}) = \mu(\{i+1\}).$$

Soit $(\tilde{L}(n))_{n\geq 0}$ une marche aléatoire sur $\mathbb{Z}$, issue de 0 et dont la loi des sauts est μ^*. Soit $\tau = \inf\{n \geq 1, \tilde{L}(n) = -1\}$. Alors la loi de L est la même que celle de $\tilde{L}$ arrêtée au temps τ. Grâce à cette propriété, nous pouvons utiliser toutes les méthodes et théorèmes connus pour les marches aléatoires afin d'obtenir des résultats sur les arbres de Galton-Watson. Donnons quelques exemples.

Loi limite pour la taille d'un arbre de Galton-Watson Le théorème suivant a d'abord été démontré par Otter [1949] en utilisant une méthode de fonction génératrice :

Théorème : Formule d'Otter-Dwass.

Soit T un arbre de Galton-Watson de loi de reproduction μ, vu en tant que graphe, et soit L sa marche de Łukasiewicz. Soit $\tilde{L}$ la marche aléatoire définie plus haut. Alors, pour tout $n \geq 1$:

$$\mathbb{P}(|T| = n) = \frac{1}{n}\mathbb{P}(\tilde{L}(n) = -1),$$

avec $|T|$ la taille de l'arbre T.

Nous renvoyons à Pitman [1998] pour une démonstration utilisant des comptages d'arbres de Galton-Watson de ce résultat. Grâce à la formule d'Otter-Dwass, nous pouvons obtenir une approximation assez précise de la probabilité qu'un arbre de Galton Watson ait une taille n. En effet supposons que μ est d'espérance finie $m \geq 0$ et de variance finie σ^2. De plus supposons qu'il n'existe pas d'entiers a et $b > 1$ tels que μ est portée par l'ensemble $\{z \in \mathbb{Z}, a + bz\}$. Cette dernière condition assure que la marche de Łukasiewicz L ne sera pas cantonnée à un sous-ensemble strict de $\mathbb{Z}$. Sous ces conditions, le théorème central limite local discret dit que :

$$\mathbb{P}(\tilde{L}(n) = -1) = \frac{1}{\sqrt{2\pi n}\sigma} \exp\left(\frac{-(n(m-1))^2}{2n\sigma^2}\right) + o(n^{-1/2}). \tag{1.2}$$

Nous renvoyons le lecteur à Davis and McDonald [1995] pour une preuve concise du théorème central limite local discret, ainsi qu'une brève mise en contexte historique de ce théorème.

Grâce à la formule d'Otter-Dwass et à l'Equation (1.2), nous obtenons le théorème suivant :

Limite locale de la taille des arbres de Galton-Watson

Soit T un arbre de Galton-Watson de loi de reproduction μ vérifiant les conditions de l'Equation (1.2). Alors, pour tout $n \geq 1$:

$$\mathbb{P}(|T| = n) = \frac{1}{\sqrt{2\pi}n^{3/2}\sigma} \exp\left(\frac{-(1+n(m-1))^2}{2n\sigma^2}\right) + o(n^{-3/2}).$$

On déduit de ce résultat deux choses intéressantes. D'abord, quand $m = 1$, c'est à dire quand l'arbre de Galton-Watson est critique. On a :

$$\mathbb{P}(|T| = n) = \frac{1}{\sqrt{2\pi}n^{3/2}\sigma} + o(n^{-3/2}).$$

Et aussi, en sommant, vu qu'on sait que T est fini avec probabilité 1 :

$$\mathbb{P}(|T| \geq n) = \frac{2}{\sqrt{2\pi}n^{1/2}\sigma} + o(n^{-1/2}). \tag{1.3}$$

Cependant, quand $m \neq 1$, c'est à dire dans les cas sous-critique et sur-critique, la probabilité $\mathbb{P}(|T| \geq n \cap |T| < \infty)$ décroît exponentiellement. On voit ainsi, que dans le cas sur-critique l'arbre n'est quasiment jamais grand et fini. Donc il est soit infini, soit relativement petit. Plus formellement, un petit calcul (voir par exemple Lemme 1.2.5 dans Abraham and Delmas [2014]) montre qu'un arbre de Galton-Watson sur-critique conditionné à mourir a la même distribution qu'un arbre de Galton-Watson sous-critique avec une distribution qui s'écrit en fonction de la distribution de l'arbre sur-critique. Rappelons finalement que tout ce qui est dit ici pour la marche de Łukasiewicz reste vrai pour le codage par comptage de nombre d'enfants associé à l'exploration en largeur, car ces deux processus ont la même loi dans le cas des arbres de Galton-Watson.

Le théorème d'Aldous et l'arbre brownien continu

Aldous (Aldous [1991a], Aldous [1991b], Aldous [1993]) a étudié les arbres de Galton-Watson en utilisant les codages, et a démontré la convergence du processus de contour associé aux arbres de Galton-Watson continues. Ce résultat lui permis ensuite de démontrer la convergence des arbres de Galton-Watson en tant qu'espaces métriques compacts vers un espace limite. Nous commençons ici par définir la limite du processus de contour avant de passer à celle des arbres de Galton-Watson.

Soit $(B_t)_{t\geq 1}$ un mouvement brownien réel. Notons $g_1 = \sup\{t < 1, B_t = 0\}$ et $d_1 = \inf\{t > 1, B_t = 0\}$. Clairement $d_1 - g_1 > 0$ presque sûrement. On définit alors l'excursion Brownienne

$$(e(t))_{0\leq t\leq 1} = \left(\frac{B_{g_1+t(d_1-g_1)}}{\sqrt{d_1-g_1}}\right)_{0\leq t\leq 1}.$$

Soit μ une loi de probabilité critique (de moyenne 1) et de variance finie $\sigma^2 < \infty$. Notons T_n pour un arbre de Galton-Watson de loi de reproduction μ conditionné à avoir n noeuds. Soit $(C_{T_n}(k))_{0\leq k\leq 2n-1}$ la fonction de contour de T_n, et soit $(\tilde{C}_{T_n}(t))_{0\leq t\leq 1}$ la normalisation à l'intervalle de temps $[0,1]$ de cette fonction. Pour I un intervalle de $\mathbb{R}$, on note $\mathscr{C}(I,\mathbb{R})$ l'espace des fonctions continues de I dans $\mathbb{R}$ muni de la topologie de la convergence uniforme sur tous les compacts de I. $\mathscr{C}([0,1],\mathbb{R})$ est un espace métrique complet et séparable (Billingsley [1968]). Un tel espace est généralement appelé espace polonais. Avec ces notations, nous avons le théorème suivant de Aldous [1993].

Convergence de la fonction de contour d'un arbre de Galton-Watson critique.

Avec les conditions citées plus haut, la convergence suivante

$$\left(\frac{\sigma}{2\sqrt{n}}\tilde{C}_{T_n}(t)\right)_{0\leq t\leq 1} \xrightarrow{d} (e(t))_{0\leq t\leq 1}$$

tient dans l'espace $\mathscr{C}([0,1],\mathbb{R})$ des fonctions continues de $[0,1]$ dans $\mathbb{R}$.

Notons que ce théorème donne directement comme corollaire la convergence de la hauteur de T_n normalisée par racine de n vers $\frac{2}{\sigma}\sup_{0\leq t\leq 1}(e(t))$. En fait, on peut obtenir bien plus grâce à ce théorème. On peut démontrer que la suite des arbres $(T_n)_{n\geq 1}$ converge vers un arbre réel. Pour cela nous devons introduire quelques définitions.

Arbres réels compacts

Un espace métrique compact $(\mathscr{T}, d)$ est appelé arbre réel, ou $\mathbb{R}$-arbre, si pour tout couple de points $(u, v) \in \mathscr{T}$ les deux conditions suivantes sont vérifiées :

- Il existe une unique isométrie $c_{u,v} : [0, d(u, v)] \to \mathscr{T}$ telle que $c_{u,v}(0) = u$ et $c_{u,v}(d(u, v)) = v$.
- Si $f : [0, 1] \to \mathscr{T}$ est une application injective continue telle que $f(0) = u$ et $f(1) = v$ alors $f([0, 1]) = c_{u,v}([0, d(u, v)])$.

Pour faire le parallèle avec les arbres discrets, la première condition assure la connexité, tandis que la seconde assure qu'il n'y a pas de cycles dans l'arbre. Un $\mathbb{R}$-arbre enraciné est un triplet $(\mathscr{T}, \rho, d)$ où $(\mathscr{T}, d)$ est un $\mathbb{R}$-arbre compact et ρ un point distingué dans $\mathscr{T}$. On considère aussi ces arbres à bijections isométriques préservant la racine prêt, et on note ainsi $\mathscr{T}_{\mathbb{R}}$ l'espace des classes d'équivalences de $\mathbb{R}$-arbres compacts enracinés.

Soit $g : [0, 1] \to \mathbb{R}$ une fonction continue telle que $g(x) > 0$ pour $0 < x < 1$, $g(0) = 0$ et $g(1) = 0$. Pour $(a, b) \in [0, 1]^2$, avec $a \leq b$, on pose

$$d_g(a, b) = g(a) + g(b) - 2 \min_{[a,b]}(g(t)),$$

et $d_g(b, a) = d_g(a, b)$. Il est clair que d_g est une pseudo-distance. Soit T_g l'ensemble des classes d'équivalences associées à d_g, la relation d'équivalence étant $a \sim b$ si $d_g(a, b) = 0$. T_g muni de la projection canonique $\tilde{d}$ de d_g est un espace métrique. Soit $\tilde{0}$ la classe d'équivalence de 0 dans T_g, alors $(T_g, \tilde{0}, \tilde{d})$ est un $\mathbb{R}$-arbre compact enraciné (Voir l'Illustration 1.14 pour un exemple). Voici donc une définition possible de l'arbre Brownien continu (présente dans Aldous [1993]).

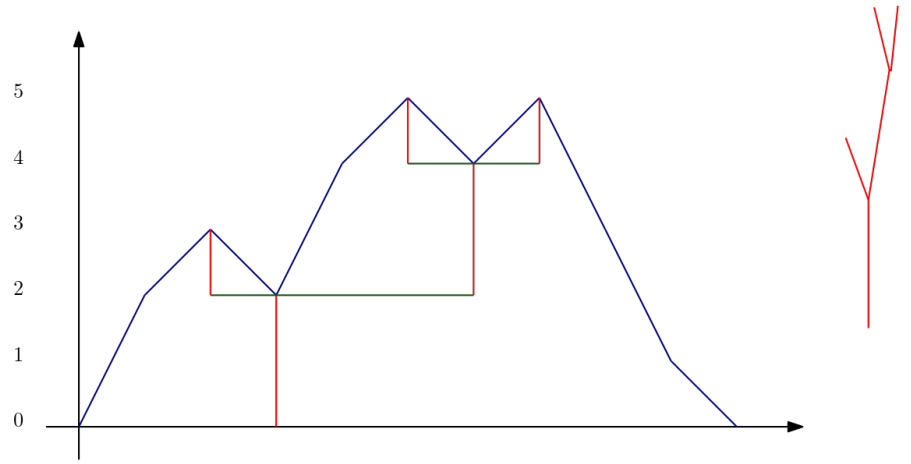


FIGURE 1.14 – Une excursion (en bleu) et un plongement dans le plan du $\mathbb{R}$-arbre associé (en rouge).

Arbre Brownien continu (CRT)

L'arbre Brownien continu est le $\mathbb{R}$-arbre compact obtenu comme classe d'équivalence associée à la pseudo distance issue de l'excursion Brownienne multipliée par deux.

On voit ainsi que deux fois l'excursion Brownienne, qui est la limite des fonctions de contours normalisées des arbres T_n code elle même un $\mathbb{R}$-arbre continue. De plus, on peut munir l'espace $\mathscr{T}_{\mathbb{R}}$ d'une topologie pour laquelle la convergence des fonctions de contours implique la convergence

des arbres associés. Pour un espace métrique (M, d), on note d_H la distance de Hausdorff entre les compacts de M. Si M_1 et M_2 sont deux compacts de M alors :

$$d_H(M_1, M_2) = \inf\{\varepsilon > 0 : M_1 \subset F_\varepsilon(M_2) \text{ et } M_2 \subset F_\varepsilon(M_1)\},$$

avec $F_\varepsilon(M') = \{x \in M : d(x, M') \leq \varepsilon\}$ pour tout M' compact de M. Si $(\mathscr{T}, \rho, d)$ et $(\mathscr{T}', \rho', d')$ sont deux $\mathbb{R}$-arbres compacts enracinés, on définit la distance de Gromov-Hausdorff entre les deux arbres compacts enracinés

$$d_{GH}(\mathscr{T}, \mathscr{T}') = \inf\{d_H(\Phi(\mathscr{T}), \Phi'(\mathscr{T})') \vee \delta(\Phi(\rho), \Phi'(\rho'))\},$$

où l'infimum est pris sur tous les choix d'espaces (M, d) et les injections isométriques $\Phi : \mathscr{T} \to M$ et $\Phi' : \mathscr{T}' \to M$. Ici $\delta(\Phi(\rho), \Phi'(\rho')) = +\infty$ si $\Phi(\rho) \neq \Phi'(\rho')$ et 0 sinon.

D'après (Duquesne and Le Gall [2005], Lemme 2.3), si g et g' sont deux excursions alors

$$d_{GH}(T_g, T_{g'}) \leq 2||g - g'||_\infty.$$

Grace à ce résultat et au théorème d'Aldous sur la convergence de la fonction de contour, on obtient :

Convergence des arbres de Galton-Watson conditionnés vers le CRT

Soit μ une loi de probabilité critique et de variance finie $\sigma^2 < \infty$. Soit $\frac{\sigma}{\sqrt{n}}T_n$ l'arbre de Galton-Watson vu en tant que $\mathbb{R}$-arbre compact, enraciné en un point distingué et dont la métrique correspond à la distance de graphe habituelle normalisé par $\frac{\sigma}{\sqrt{n}}$. On a alors la convergence suivante

$$\frac{\sigma}{\sqrt{n}}T_n \xrightarrow{d} T_{2e},$$

où T_{2e} est l'arbre brownien et la convergence a lieu pour la distance de Gromov-Hausdorff.

Ce résultat n'est que le premier d'une longue série de résultats similaires, pour les arbres de Lévy stables (Le Gall and Duquesne [2002]), les graphes d'Erdős-Rényi (Addario-Berry, Broutin, and Goldschmidt [2012]), les graphes inhomogènes (Bhamidi, van der Hofstad, and Sen [2018]), et l'arbre couvrant minimum aléatoire (Addario-Berry et al. [2017b]). Nous présentons ici une dernière application des codages pour démontrer la concentration d'une quantité associée aux arbres de Galton-Watson.

Concentration de la largeur des arbres de Galton-Watson

Soit T un arbre de Galton Watson critique de loi de reproduction μ à support bornée par $M > 0$, de moyenne $m = 1$ et de variance σ. Nous savons d'après l'Equation (1.3) que

$$\mathbb{P}(|T| \geq n) = \frac{2}{\sqrt{2\pi}n^{1/2}\sigma} + o(n^{-1/2}).$$

On définit la largeur $W(T)$ de T comme le nombre maximal de noeuds de T ayant une même hauteur. Considérons L le codage par comptage de nombre d'enfants associé à une exploration en profondeur de T. Comme ce codage compte les noeuds à différence de hauteur au plus 1. Il est facile de voir que, pour tout $l \geq 0$ et $n \geq 0$:

$$\mathbb{P}(W(T) \geq l \cap |T| \leq n) \leq \mathbb{P}\left(\sup_{i \leq n}(L(i)) \geq l \cap |T| \leq n\right).$$

Ici nous utilisons une technique qui sera souvent réutilisée, avec plusieurs variations, dans les chapitres 2 et 3. Comme L a la même loi que la marche aléatoire centrée $\tilde{L}$, et que $\tilde{L}$ est une martingale, son exponentielle $\exp(\tilde{L})$ est une sous-martingale. Par l'inégalité de Doob pour les sous-martingales (Revuz and Yor [1999]) :

$$\mathbb{P}\left(\sup_{i \leq n}(L(i)) \geq l\right) \leq \frac{\mathbb{E}[\exp(\tilde{L}(n))]}{e^l}.$$

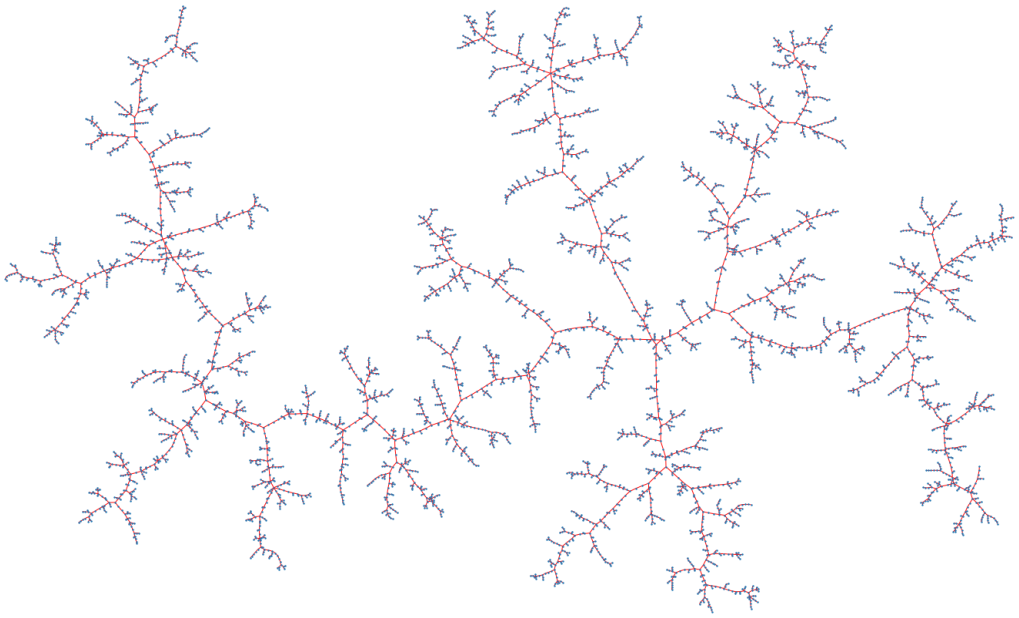


FIGURE 1.15 – Un arbre de Galton-Watson de loi de reproduction Poisson de paramètre 1 et conditionné à avoir taille 5000. Cet arbre est une bonne approximation de T_e.

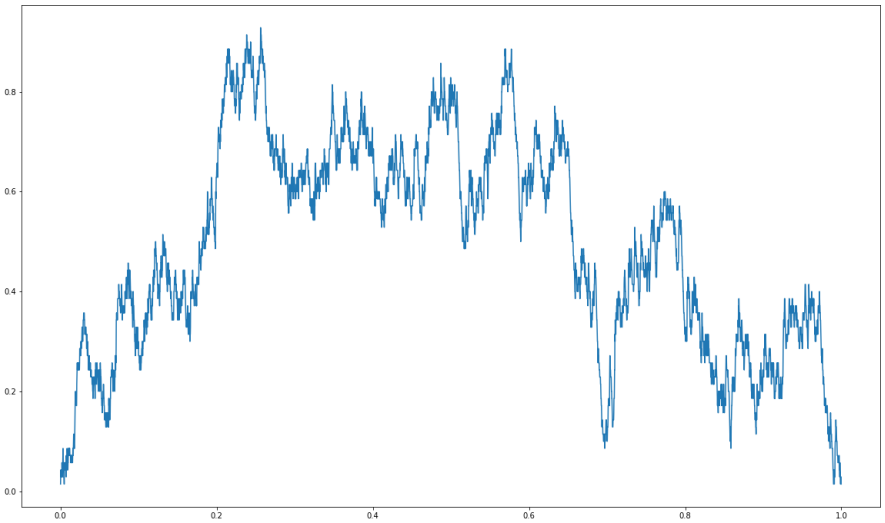


FIGURE 1.16 – La fonction de contour associée à l'arbre de la Illustration 1.15 renormalisé en temps par 5000 et en espace par $2\sqrt{5000}$. Elle représente une bonne simulation de e.

Ainsi, toutes les bornes supérieures de concentration qui découlent de l'inégalité de Chernoff pour $\tilde{L}$ sont aussi vraies pour $\sup_{i\leq n}(L(i))$. En particulier l'inégalité de Bernstein tient elle aussi (Bernstein [1924]), et on obtient

$$\begin{aligned}\mathbb{P}(W(T)\geq l) &\leq \mathbb{P}\left(\sup_{i\leq n}(L(i))\geq l\cap |T|\leq n\right)+\mathbb{P}(|T|\geq n)\\ &\leq \mathbb{P}\left(\sup_{i\leq n}(L(i))\geq l\right)+\frac{2}{\sqrt{2\pi}n^{1/2}\sigma}+o(n^{-1/2})\\ &\leq \frac{\mathbb{E}[\exp(\tilde{L}(n))]}{e^l}+\frac{2}{\sqrt{2\pi}n^{1/2}\sigma}+o(n^{-1/2})\\ &\leq \exp\left(\frac{-l^2}{2(n(m^2+\sigma^2)+lM/3)}\right)+\frac{2}{\sqrt{2\pi}n^{1/2}\sigma}+o(n^{-1/2}),\end{aligned}$$

le terme $o(n^{1/2})$ étant bien sûr uniforme en l. On peut ensuite choisir $l=O(\sqrt{n})$ pour avoir une borne de concentration sur la largeur de l'arbre en fonction de n. On voit aussi avec cette démonstration que la largeur devrait être en général de l'ordre de grandeur de la racine de la taille de l'arbre. Ce résultat est démontré formellement dans Addario-Berry [2019] pour les arbres de Galton-Watson non-conditionnés, et dans Addario-Berry et al. [2013] pour les arbres conditionnés avec une loi de reproduction à second moment fini, puis par Kortchemski [2017] pour les arbres conditionnés avec une loi de reproduction dans le domaine d'attraction d'une loi stable.

1.3.4 Graphes aléatoires uniformes et inhomogènes

Erdős-Rényi

Le modèle des graphes d'Erdős-Rényi a été introduit dans Erdős and Rényi [1960]. Depuis, il a donné lieu à un énorme corpus de recherche. Dans cette introduction, suivant les conventions habituelles, nous appelerons modèle d'Erdős-Rényi le modèle introduit dans Gilbert [1959] simultanément et qui est très similaire au véritable modèle introduit par Erdős et Rényi.

Graphe d'Erdős-Rényi

Soit $n\geq 1$ et $1\geq p\geq 0$. Un graphe d'Erdős-Rényi, noté $G(n,p)$, est un graphe à n noeuds $1,2,...,n$ dont chaque arrête potentielle est présente indépendamment des autres avec probabilité p.

Dans leur article fondateur Erdős and Rényi [1960], Erdős et Rényi on démontré qu'il existe un changement de structure (changement de phase) dans le graphe d'Erdős-Rényi pour n grand et $p=p(n)=\frac{c}{n}$ avec $c>0$. En effet :

- Quand $c<1$ la plus grande composante connexe est de taille $O(\log(n))$ (cas sous-critique).
- Quand $c>1$ il existe une composante connexe géante qui contient une proportion strictement positive des noeuds (sur-critique).
- Quand $c=1$ les plus grandes composantes connexes de $G(n,p(n))$ sont de taille $\Theta(n^{2/3})$ (cas critique).

Cette transition de phase ressemble a celle observée pour les arbres de Galton-Watson, et cette ressemblance n'est pas fortuite. En voici une explication intuitive : une composante connexe de $G(n,p)$ de taille n' et avec m arêtes correspond un graphe connexe choisi uniformément parmi l'ensemble des graphes connexes de taille n' et ayant m arêtes. Si $s=m-n'=-1$ alors ce graphe est en fait un arbre, la quantité s est appelé surplus (parfois on appelle plutôt $s+1$ surplus) ; $s+1$ correspond au nombre minimal d'arêtes qu'il faudrait retirer au graphe si on voulait en faire un arbre connexe. Dans le cas où $s=-1$ la composante connexe correspond à un

arbre uniforme sur l'ensemble des arbres (vus comme graphes) à n' sommets. D'un autre côté, un arbre de Galton-Watson de loi de reproduction μ de Poisson de paramètre 1 conditionné à être de taille n' est aussi un arbre uniforme à n' sommet. Ceci constitue un premier lien direct entre les graphes d'Erdős-Rényi et les arbres de Galton-Watson. De plus, dans le cas où $s > 1$, la marche de Łukasiewicz associée à la composante connexe est en bijection naturelle avec un arbre couvrant de cette même composante connexe. Cet arbre correspond à une version "biaisée" par s des arbres uniformes (nous renvoyons ici à Addario-Berry et al. [2017b] pour une description de ce biais).

Ce rapport direct est encore plus visible dans le théorème de convergence des tailles des composantes connexes d'Aldous. Soit

$$B^\lambda(t) = B(t) + t\lambda - \frac{t^2}{2},$$

avec $(B(t))_{t\geq 0}$ un mouvement Brownien standard. Soit $\mathbf{Z} = (Z_1, Z_2, ...)$ la suite décroissante des longueurs des excursions de la réflection $(B^\lambda(t) - \min_{0\leq s\leq t}(B^\lambda(s)))_{t\geq 0}$ par rapport à 0 de $B^\lambda(t)$. Aldous a démontré dans Aldous [1997] que $\mathbf{Z}$ est bien définie et que :

Théorème d'Aldous pour la convergence des tailles de composantes connexes de graphes critiques

Soit $\lambda \in \mathbb{R}$ et pour tout $n \geq 1$ soit $p(n) = \frac{1}{n} + \frac{\lambda}{n^{4/3}}$. Notons $\mathbf{Z}^n = (Z_1^n, Z_2^n, ...)$ la suite des tailles des composantes connexes d'un graphe $G(n, p(n))$, complétée avec des 0 après la taille de la dernière composante connexe pour en faire une suite infinie. La convergence suivante

$$n^{-2/3}\mathbf{Z}^n \xrightarrow{d} \mathbf{Z}$$

a lieu en loi dans ℓ^2_{dec}, l'espace des suites décroissantes $(x_1, x_2, ...)$ à termes positifs et telles que $\sum_{i\geq 1} x_i^2 < \infty$.

Pour $p(n) = \frac{1}{n} + \frac{\lambda}{n^{4/3}}$ on dit que $G(n, p(n))$ est dans la fenêtre critique. Cette fenêtre critique avait déjà suscité de l'intérêt avant ce résultat d'Aldous (Łuczak [1990], Łuczak, Pittel, and Wierman [1994]). L'apparition des excursions d'un mouvement Brownien réfléchi à drift parabolique dans ce théorème n'est pas sans rappeler le théorème de convergence de la marche de Łukasiewicz pour des arbres de Galton-Watson conditionnés. Ce même parallèle apparaît dans l'article Addario-Berry et al. [2012], qui démontre la convergence des graphes d'Erdős-Rényi critiques vers des suites d'espaces métriques compacts pour la topologie de Gromov-Hausdorff. Dans Aldous [1997], Aldous démontre un résultat encore plus général qui s'applique aux graphes inhomogènes que nous définirons par la suite. Depuis, le théorème d'Aldous a été étendu à d'autres fonctionnelles et modèles de graphes aléatoires (Bhamidi, van der Hofstad, and van Leeuwaarden [2010], Dhara, van der Hofstad, van Leeuwaarden, and Sen [2017]).

Les graphes inhomogènes

Une extension naturelle du modèle d'Erdős-Rényi est le modèle de graphes aléatoires communément appelé graphes inhomogènes. Ce nom englobe plusieurs modèles similaires, nous nous intéresserons dans cette thèse au modèle de processus de graphes Poissonien, aussi appelé modèle de Norros-Reittu Norros and Reittu [2006]. Bien que notre modèle diffère légèrement de ce dernier en cela qu'il ne permet pas d'arêtes multiples. Le premier modèle qu'on pourrait qualifier de graphe inhomogène est en effet celui introduit par Aldous Aldous [1997] dans son coalescent multiplicatif, puis étudié plus précisément dans Aldous and Limic [1998]. Mais nous suivons les conventions en vigueur et appellerons donc notre modèle simplement modèle de Norros-Reittu. Notons cependant que pour les régimes qui nous intéressent dans cette thèse nos résultats s'étendent naturellement aux autres modèles similaires (Chung and Lu [2006], Britton, Deijfen, and Martin-Löf [2006]). Ceci est dû au fait que les ordres de grandeur des différences

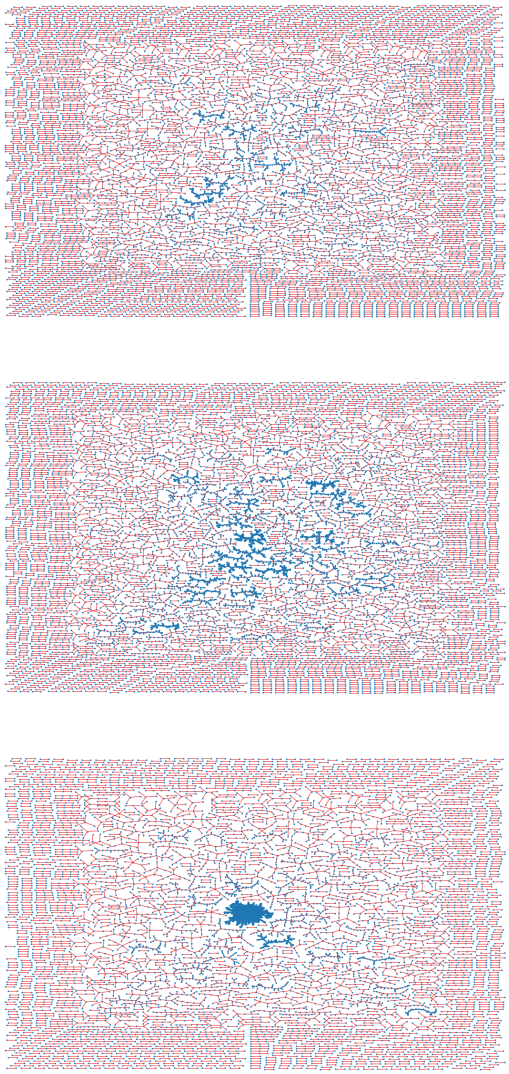


FIGURE 1.17 – Trois graphes d'Erdős-Rényi à $n = 20000$ noeuds. Avec comme paramètres $p_1 = \frac{0.8}{n}$, $p_2 = \frac{1}{n}$ et $p_3 = \frac{1.2}{n}$. Les trois graphes sont construits à partir des mêmes poids uniformes sur $[0, 1]$ attribués aux arêtes du graphe complet. Pour construire chaque graphe on ne garde que les arêtes ayant un poids plus petit que p_i pour $i \in \{1, 2, 3\}$.

entre les différents modèles sont en général négligeables devant les ordres de grandeur qui apparaissent typiquement dans nos résultats.

Graphe inhomogène

Soit $p \geq 0$, $n \geq 1$ et $\mathbf{W} = (w_1, w_2, ..., w_n)$ une suite de poids positifs. Le graphe inhomogène $G(\mathbf{W}, p)$ est un graphe à n sommets $\{1, 2, ..., n\}$ est dans lequel chaque arête $\{i, j\}$ est présente indépendamment de tout le reste avec probabilité

$$1 - \exp(w_i w_j p).$$

Si tous les poids sont égaux, nous obtenons un graphe d'Erdős-Rényi. Remarquez que $\mathbf{W}$ dépend implicitement de n. Posons

$$\nu = \frac{\sum_{k=1}^{n} w_k^2}{\ell_n},$$

avec $\ell_n = \sum_{k=1}^{n} w_k$. De plus supposons que $\nu = 1 + o(n^{-1/3})$ avec probabilité tendant vers 1 quand n tend vers l'infini et que la liste des poids $\mathbf{W}$ correspond à un tirage i.i.d. de loi μ ayant un troisième moment fini. Soit W une variable aléatoire de loi μ. Sous cette condition nous avons une transition de phase similaire à celle des graphes d'Erdős-Rényi (Bollobás, Janson, and Riordan [2007]).

Transition de phase pour les graphes inhomogènes

Soit $(G(\mathbf{W}, \frac{c}{\ell_n}))_{n\geq 1}$ une suite de graphes inhomogènes. Les évènements suivants sont vrais avec probabilité tendant vers 1 quand n tend vers l'infini :

- **Régime sous-critique** Si $c < 1$, alors la plus grande composante connexe est de taille $o(n)$.
- **Régime sur-critique** Si $c > 1$, alors la plus grande composante connexe contient une proportion positive des noeuds, et pour tout $i > 1$ la i-ième plus grande composante connexe est de taille $o(n)$.
- **Régime critique** Si $c = 1$, alors pour tout $i \geq 1$ la i-ième plus grande composante connexe est de taille $\Theta(n^{2/3})$.

Il y a généralement deux types de conditions différents qui sont étudiées. Celui où μ a un troisième moment fini, et celui où μ est à queue en loi de puissance, c'est à dire il existe $3 < \tau < 4$ tel que $\mu(\{k\})$ est d'ordre $\frac{1}{k^\tau}$ pour tout k assez grand. On appelle ce second cas le modèle à échelle libre (scale free). Et dans ce cas il existe un théorème similaire pour la transition de phase, sauf que dans le régime critique les composantes connexes ont une taille d'ordre $n^{\frac{\tau-2}{\tau-1}}$ (Bhamidi, van der Hofstad, and Sen [2018]). Notons finalement que la condition sur les poids sera remplacée dans les chapitres 2 et 3 par une condition un peu plus générale, mais ici nous nous contenterons de supposer un tirage i.i.d. par simplicité.

Tout comme le parallèle entre les arbres de Galton-Watson et le graphe d'Erdős-Rényi, il existe un parallèle entre les graphes inhomogènes et un autre type d'arbres. Pour un arbre enraciné, on dit qu'un noeud a est enfant de b si b est le premier noeud après a dans le chemin entre a et la racine.

p-arbres

Soit $n \geq 1$ et $\mathbf{p} = (p_1, p_2, ..., p_n)$ un vecteur de probabilité avec $p_i > 0$ pour tout $i \geq 1$. La loi d'un $\mathbf{p}$-arbre T est la suivante, pour tout arbre $\mathbf{t}$ à n noeuds $(1, 2, ..., n)$ et enraciné :

$$\mathbb{P}(T = \mathbf{t}) = \prod_{n \geq i \geq 1} p_i^{d_i(\mathbf{t})},$$

où $d_i(\mathbf{t})$ est le nombre d'enfants de i dans $\mathbf{t}$.

Le cas particulier $p_1 = p_2 = p_3 ... = p_n = \frac{1}{n}$ correspond à un arbre uniforme enraciné, qui est un cas particulier d'arbre de Galton-Watson enraciné et conditionné à avoir taille n. Tout comme les graphes inhomogènes sont des généralisations des graphes d'Erdős-Rényi, les $\mathbf{p}$-arbres sont des généralisations des arbres de Galton-Watson. De façon similaire, l'article Aldous and Pitman [2000] démontre que les $\mathbf{p}$-arbres convergent vers des arbres réels, les ICRT, qui peuvent être vus comme des généralisations du CRT. Nous ne nous attarderons pas sur ces arbres et renvoyons le lecteur vers ledit article pour plus d'informations. Les $\mathbf{p}$-arbres apparaissent de façon biaisés dans l'étude de processus associés aux graphes inhomogènes (voir par exemple Bhamidi et al. [2018]). Cependant, dans la suite de cette thèse ces arbres ne seront pas utilisés. En effet, nous avons préféré utiliser des approximations par arbres de Galton-Watson afin d'avoir pleinement accès à la multitude de résultats connus pour ces derniers.

Finalement, un travail initié dans Bhamidi, van der Hofstad, and van Leeuwaarden [2010] et Bhamidi et al. [2018], et terminé récemment dans Broutin, Duquesne, and Wang [2020] démontre que les graphes inhomogènes convergent vers des suites d'espaces métriques compacts. Ces espaces sont fondamentalement différents suivant que la loi des poids μ est à troisième moment fini où à queue en loi de puissance. Dans le cas du troisième moment fini, la limite d'échelle des graphes inhomogènes est la même en loi que celle des graphes d'Erdős-Rényi à normalisation par des constantes prêt. En un sens, l'inhomogénéité s'efface à la limite. Par contre dans le cas d'échelle libre, les lois limites sont mutuellement singulières. Dans la suite de cette thèse nous étudierons le cas du troisième moment fini.

1.4 Retour vers l'arbre couvrant minimum

1.4.1 Le rapport entre MST et graphes aléatoires

L'arbre couvrant minimum aléatoire de la partie 1.2.4 peut être construit avec n'importe quel distribution μ de base. Cependant, quand on s'intéresse à des propriétés géométriques de ce dernier, la distribution μ n'est pas importante. En effet, que ce soit pour l'algorithme de Prim ou de Kruskal, μ permet simplement de donner un ordre aléatoire des arêtes, du plus petit poids jusqu'au plus grand. Cet ordre est uniforme quelque soit la distribution μ de base. On peut donc sans perte de généralité choisir pour μ la loi uniforme sur $[0, 1]$.

D'un autre côté, pour n fixé. On peut construire un processus croissant de graphes (pour l'inclusion) avec p tel que chaque élément du processus est un graphe d'Erdős-Rényi en procédant ainsi :

A chaque arête e du graphe complet à n noeuds K_n on associe un poids uniforme dans $[0, 1]$ indépendamment de tout le reste. Soit $0 \leq p \leq 1$. On construit ensuite un sous-graphe $G(n, p)$ de K_n en ne gardant que les arêtes e telles que $W_e \leq p$. Il est clair que $G(n, p)$ comme défini ici est effectivement un graphe d'Erdős-Rényi. De plus, le processus $(G(n, p))_{0 \leq p \leq 1}$ est croissant pour l'inclusion et fournit donc un couplage naturel entre des graphes d'Erdős-Rényi pour des valeurs de p différentes. Notons $\mathbf{T}_n$ l'arbre couvrant minimum aléatoire associé aux poids (W_e), et pour $p \geq 0$ notons $T(n, p)$ le sous-graphe de $\mathbf{T}_n$ composé des arêtes de $\mathbf{T}_n$ ayant un poids plus

petit que p. Alors $T(n,p)$ est une forêt, le processus $(T(n,p))_{1\geq p\geq 0}$ est lui aussi croissant pour l'inclusion, et pour p fixé les arbres dans $T(n,p)$ sont les arbres couvrants minimums associés au poids (W_e) des composantes connexes de $G(n,p)$. Donc, l'étude des graphes $G(n,p)$ devrait donner des informations sur les arbres $T(n,p)$ qui à leur tour vont donner des informations sur $\mathbf{T}_n$. C'est en effet ce qui a été fait dans Addario-Berry et al. [2006] pour étudier le diamètre de l'arbre couvrant minimum aléatoire. Si une composante connexe de $G(n,p)$ est un arbre, alors elle va correspondre à une composante connexe de $T(n,p)$. De façon générale, tant que les composantes connexes de $G(n,p)$ ont un petit surplus, les longueurs de leurs plus longs chemins simples ne devraient pas être significativement différentes des diamètres de leurs arbres couvrants minimums. Cette intuition peut être formalisée par le lemme suivant, prouvé dans le Chapitre 3.

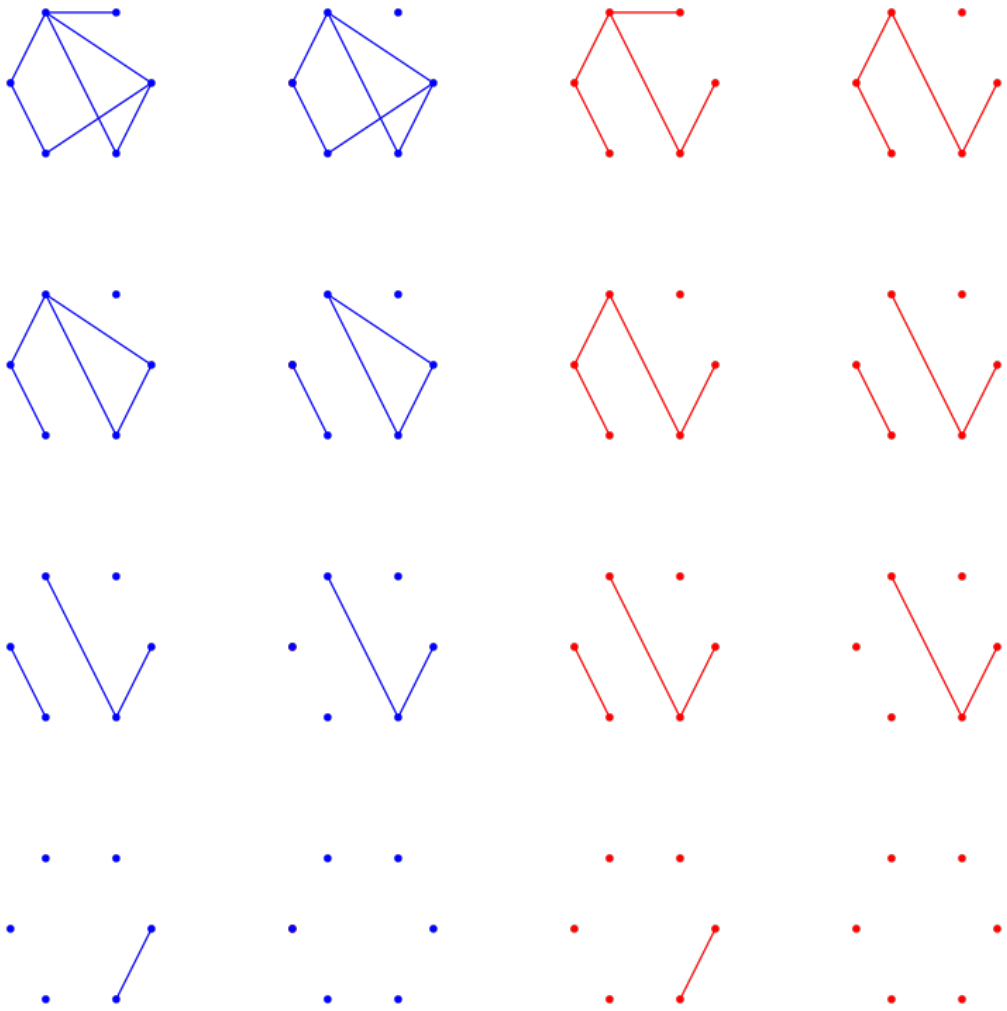


FIGURE 1.18 – Exemple de constructions simultanées d'un processus de graphes d'Erdős-Rényi (en bleu à gauche) et d'un arbre couvrant minimal (en rouge à droite). Les poids des noeuds ne sont pas représentés et on s'arrête quand l'arbre est construit.

Soit G un graphe connexe ayant surplus q et et ayant un arbre couvrant de hauteur h, alors le plus long chemin sans cycles de G est de taille au plus

$$2h(q+1)+q.$$

Addario-Berry, Broutin, and Reed [2006] utilise le fait que la moyenne du plus grand diamètre dans $T(n,p)$ atteint une taille de l'ordre de $n^{1/3}$ dans le régime critique $p_n = \frac{1}{n}$, en effet il était déjà connu avant (par exemple dans Janson, Łuczak, and Rucinski [2000]) qu'avec probabilité strictement positive $\varepsilon > 0$ il existe une composante connexe de $G(n,p_n)$ d'ordre $n^{2/3}$ et qui est un arbre. Comme mentionné précédemment, une telle composante connexe sera donc un arbre aléatoire uniforme et il est connu que de tels arbres ont avec grande probabilité une hauteur de l'ordre de la racine de leur taille (voir l'Article Flajolet and Odlyzko [1982] et Addario-Berry, Devroye, and Janson [2013]). Les auteurs de Addario-Berry et al. [2006] ont ensuite démontré que le diamètre de $T(n,p)$ n'augmente pas significativement pour $p > p_n$. C'est cette partie qui constitua le gros du travail et qui demande une analyse approfondie du graphe d'Erdős-Rényi dans la fenêtre critique. Cette étude a été ensuite reprise et raffinée dans Addario-Berry et al. [2017b] afin de démontrer la convergence au sense de Gromov-Hausdorff-Prokhorov de l'arbre couvrant minimum aléatoire, vu en tant qu'espace métrique et avec des distances normalisées par $n^{1/3}$, vers un arbre continu dont la loi est mutuellement singulière avec celle du CRT. Encore une fois, l'utilisation du parallèle avec le graphe d'Erdős-Rényi s'est avérée cruciale. Il a fallut notamment utiliser une version continue de l'algorithme de suppression d'arêtes pour obtenir la limite continue de l'arbre couvrant minimum aléatoire.

Dans cette thèse, nous reprenons les idées de ces articles afin d'étudier une généralisation de l'arbre couvrant minimum aléatoire. Considérons encore une fois le graphe complet K_n, et un vecteur $\mathbf{W} = (w_1, w_2, ..., w_n)$ de poids réels positifs. A chaque arête $\{i,j\}$ de K_n, on associe une capacité $E_{\{i,j\}}$ qui est une variable aléatoire exponentielle de paramètre $w_i w_j$, et ce, indépendamment de tout le reste. L'arbre couvrant minimum inhomogène associé à $\mathbf{W}$ est l'arbre couvrant minimum de K_n pour les capacités $(E_{\{i,j\}})$. Si tous les poids sont égaux, l'arbre couvrant minimum inhomogène sera un arbre couvrant minimum aléatoire. En ce sens, l'arbre couvrant minimum inhomogène est une généralisation de l'arbre couvrant minimum aléatoire de Addario-Berry et al. [2006]. Cependant, dans l'arbre couvrant minimum inhomogène certains noeuds auront des degrés bien plus grands que d'autres en moyenne. Plus le poids associé à un noeud est grand, plus il aura tendance à avoir un gros degré. Cette inhomogénéité introduit plusieurs difficultés dans l'étude de ces arbres. Par exemple il y a plusieurs résultats importants sur les arbres de Galton-Watson qui ne sont plus utilisables dans le cas des arbres couvrants minimums inhomogènes. Comme dit précedemment, la marche de Łukasiewicz d'un arbre de Galton-Watson est une marche aléatoire à incréments indépendants. Mais quand on étudie les graphes inhomogènes on a affaire à des arbres dont la marche de Łukasiewicz ne vérifie pas la propriété de Markov. Ces différences, et bien d'autres, font que l'étude des arbres couvrants minimums aléatoires demande plusieurs résultats nouveaux.

Afin d'attaquer ce problème nous allons utiliser un codage par comptage de nombres d'enfants associé à une exploration en largeur des graphes inhomogènes. Quand on finit d'explorer une composante connexe, on choisit le noeud à explorer avec probabilité proportionnelle à son poids parmi les noeuds restants à explorer. L'intérêt principal d'un tel codage est que les noeuds du graphe apparaissent dans un ordre $(v(1), v(2), ..., v(n))$ biaisé par leurs poids dans l'exploration en largeur. Plus précisément, si on note $\mathcal{V}_i = \{v(1), v(2), ..., v(i)\}$ pour $i \geq 1$. Alors le i-ième noeud découvert dans l'exploration en largeur vérifie

$$\mathbb{P}(v(i+1) = j | \mathcal{V}_i) = \frac{\mathbb{1}(j \notin \mathcal{V}_i) w_j}{\ell_n - \sum_{k \in \mathcal{V}_i} w_k}$$

pour tout $j \geq 1$. Ceci permet donc d'avoir un peu d'indépendance conditionnellement aux

poids déjà découverts. Ce résultat est discuté dans les Chapitres 2 et 3. Dans le Chapitre 2, nous donnons plusieurs résultats sur ces noeuds tirés sans remise et de façon biaisée. Nous démontrons notamment des inégalités de concentration nouvelles pour ce genre de tirages.

1.4.2 Aparté sur la physique statistique

Bien que peu étudiés en tant que tel dans le monde des mathématiques, l'intérêt porté aux arbres couvrants minimums inhomogènes n'est pas une nouveauté. Dans le monde de la physique statistique ces arbres sont utilisés dans des modélisations de plusieurs problèmes (nous renvoyons à l'abstract de Chen, López, Havlin, and Stanley [2006] pour une liste non exhaustive de telles modélisations). Ainsi, une conjecture posée il y a une quinzaine d'années, et appuyée par des simulations est la suivante : Considérons un graphe inhomogène sur-critique de taille $n \geq 1$, on sait que la plus grande composante connexe de ce graphe a une taille proportionelle à n avec grande probabilité. Posons des poids i.i.d. sur les arêtes de cette plus grande composante connexe, et considérons l'arbre couvrant minimum associé à ces poids. On a les deux conjectures suivantes. Quelque soit la distribution des poids sur les arêtes :

- Dans le cas où le graphe inhomogène est construit avec une loi à troisième moment fini, les distances typiques de l'arbre couvrant minimum de la plus grande composante connexe ont une moyenne de l'ordre de $n^{1/3}$.
- Dans le cas où le graphe inhomogène est construit avec une loi à échelle libre de paramètre $4 > \tau > 3$, les distances typiques de l'arbre couvrant minimum de la plus grande composante connexe ont une moyenne de l'ordre de $n^{(\tau-3)/(\tau-1)}$.

Les distances typiques sont les distances entre deux points choisis uniformément dans l'arbre. Ces deux conjectures sont présentes sous des formes un peu différentes dans plusieurs articles (Braunstein, Buldyrev, Cohen, Havlin, and Stanley [2003], Chen et al. [2006], Braunstein, Wu, Chen, Buldyrev, Kalisky, Sreenivasan, Cohen, Lopez, Havlin, and Stanley [2007]). Cette même conjecture est aussi parfois posée pour les distances typiques dans la composante connexe géante dans le régime critique des graphes inhomogènes (Bhamidi, van der Hofstad, and Sen [2018]).

Dans cette thèse, nous allons répondre à la conjecture concernant le cas troisième moment fini par l'affirmative, et ce, pour l'arbre couvrant minimum associé à la composante connexe géante dans les cas critique et sur-critique. Nous donnerons une démonstration de ce résultat dans le Chapitre 3. Notons que cette démonstration est un corollaire de notre travail, ainsi que du travail de effectué dans Broutin, Duquesne, and Wang [2020]. Notre preuve présente aussi un cheminement possible vers une preuve pour la conjecture dans le cas de l'échelle libre.

Notre preuve dans le Chapitre 3 sera d'abord faite pour les arbres $T(n,p)$ définis plus haut, et pour la composante connexe géante dans le modèle de Norros-Reittu (Norros and Reittu [2006]). Or, en général dans la littérature de physique statistique le modèle utilisé est plutôt proche de celui de Britton-Deijfen-Martin-Löf (Britton et al. [2006]). Pour ce qui est de la composante connexe géante dans le régime critique, comme dit précédemment, les deux modèles sont similaires. Et pour ce qui est de l'arbre couvrant minimum de la composante connexe géante dans le régime sur-critique, nous démontrons dans le Chapitre 3 par un rapide calcul que le modèle de physique statistique est aussi assez proche de nos arbres $T(n,p)$ pour que les résultats démontrés dans un modèle restent valides dans l'autre.

1.5 Limite d'échelle du MST renormalisé

Dans le Chapitre 4, nous démontrons un résultat concernant la limite d'échelle d'un type de MST inhomogènes. Dans cette section, nous présentons les notions nécessaires à la compréhension de ce résultat.

1.5.1 Convergence d'espaces métriques

Nous noterons (X, d) un espace métrique muni de sa distance d, et $[X, d]$ pour la classe d'équivalence correspondant à tous les espaces isométriques à (X, d). De plus quand il n'y a pas d'ambiguïté, nous noterons simplement X à la place de (X, d) ou $[X, d]$. Quand il s'agit de limite d'échelles, il y a généralement deux notions de distance entre espaces métriques qui sont utilisées. La distance de Gromov-Hausdorff, et celle de Gromov-Hausdorff-Prokhorov. La seconde distance incorpore en plus dans les espaces métriques une mesure de masse. Dans cette thèse, nous n'utiliserons que la distance de Gromov-Hausdorff. Nous renvoyons le lecteur intéressé à Abraham et al. [2013] pour plus d'information sur la distance de Gromov-Hausdorff-Prokhorov ainsi que son utilité à la définition d'un nouveau type d'arbre continu.

Rappelons rapidement la définition de cette distance déjà évoquée plus haut. Commençons par la distance de Hausdorff, cette dernière mesure à quel point deux sous-espaces métriques d'un même grand espace métrique sont loin d'être parfaitement superposés. Soit X et X' deux sous-espaces métriques d'un même grand espace métrique compact (M, d). La distance de Hausdorff entre X et X' se définie ainsi :

$$d_{\mathrm{H}}(X, X') = \max\left\{ \sup_{x\in X} \inf_{x'\in Y} d(x, x'),\ \sup_{x'\in Y} \inf_{x\in X} d(x, x') \right\}.$$

De manière équivalente, cette distance peut être calculée en utilisant la notion d'epsilon-grossissement :

$$d_{\mathrm{H}}(X, Y) = \inf\{\varepsilon \geq 0\,;\ X \subseteq Y_\varepsilon \text{ and } Y \subseteq X_\varepsilon\},$$

avec

$$X_\varepsilon := \bigcup_{x\in X} \{z \in M\,;\ d(z, x) \leq \varepsilon\},$$

l'epsilon-grossissement de X. Le problème se pose si X et X' ne sont pas des sous-espaces métriques du même espace métrique. C'est dans ce cas que nous utilisons la distance de Gromov-Hausdorff

Distance de Gromov-Hausdorff

Soit (X, d) et (X', d) deux espaces métriques compacts, on définit la distance de Gromov-Hausdorff entre $[X, d]$ et $[X', d']$, $d_{\mathrm{GH}}(X, X')$, comme l'infimum sur tous les nombres $d_{\mathrm{H}}(f(X), g(X'))$ pour tous les espaces métriques M et les plongements isométriques $f : X \to M$ et $g : X' \to M$.

Notons que ceci définit une distance entre $[X, d]$ et $[X', d']$ mais pas entre (X, d) et (X', d). Cette définition de la distance de Gromov-Hausdorff n'est que rarement utilisée en pratique, car peu maniable. Nous donnons ici une autre définition plus pratique pour les démonstrations. Avec les même notation. Soit C un sous-espace de $X \times X'$, La **distortion** $\mathrm{dis}(C)$ de C est définit par :

$$\mathrm{dis}(C) = \sup\left\{|d(x, y) - d'(x', y')| : (x, x') \in C, (y, y') \in C\right\}.$$

Une **correspondence** C entre X et X' est un sous-ensemble mesurable de $X \times X'$ tel que pour tout $x \in X$ il existe $x' \in X'$ tel que $(x, x') \in C$ et vice versa. Soit $C(X, X')$ l'ensemble de toutes les correspondances entre X et X', dans ce cas la distance de Gromov-Hausdorff entre $[X, d]$ et (X, d) peut aussi être définie par :

$$d_{\mathrm{GH}}(X, Y) = \frac{1}{2} \inf\{\mathrm{dis}(C)\ :\ C \in C(X, X')\}.$$

De plus, il existe une correspondance dont la distortion est égale à cet infimum. Pour une démonstration de cette proposition et d'autres propriétés de cette distance, nous redirigeons le

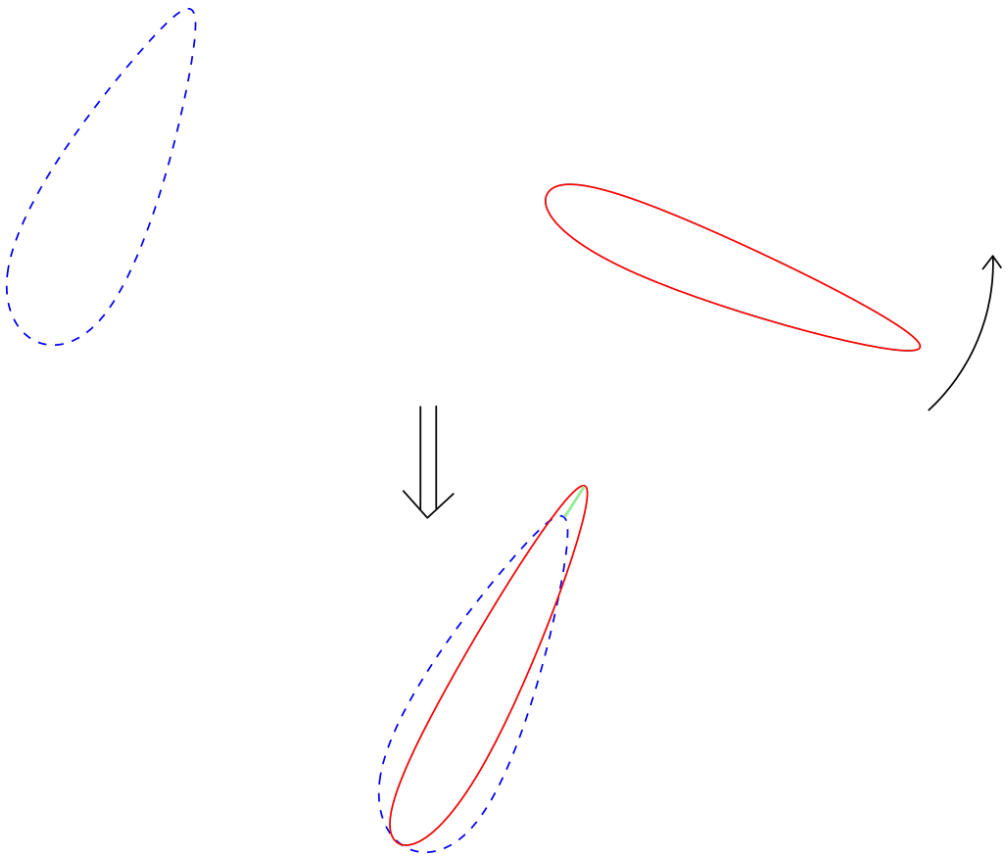


FIGURE 1.19 – Exemple de calcul de distance de Gromov-Hausdorff entre deux espaces (en bleu en pointillés et en rouge). Le segment vert représente la distance maximale entre les deux objets.

lecteur vers Kalton and Ostrovskii [1999]. Finalement, si on dénote par $\bar{\mathcal{M}}$ l'ensemble des classes d'isométries d'espaces métriques compacts, on peut vérifier que $(\bar{\mathcal{M}}, d_{\mathrm{GH}})$ est un espace polonais.

Rappelons que pour un produit d'espaces topologiques $\prod_{i\in I} X_i$, la topologie produit est la topologie la moins fine pour laquelle toutes les projections canoniques $p_i : X \to X_i$ sont continues. De plus, une suite d'éléments dans un produit d'espaces topologiques converge pour la topologie produit si et seulement si chacune des suites obtenues par les projections canoniques de cette suite converge dans l'espace projeté (Willard [1970]).

1.5.2 Convergence de graphes inhomogènes discrets

Les $\mathbb{R}$-graphes sont des espaces métriques compacts qui étendent naturellement la notion de graphes discrets au cas continu. Ils apparaissent aussi comme limite d'échelles de graphes aléatoires discrets renormalisés.

$\mathbb{R}$-graphes

Soit (X, d) un espace métrique compact. Pour $x \in X$, et pour $r > 0$, on dénote par $B_r(x) = \{y \in X : d(x, y) < r\}$ la boule ouverte de centre x et de rayon r. On dit que (X, d) est un $\mathbb{R}$-graphe, si pour tout $x \in X$ il existe $\varepsilon > 0$ tel que $B_\varepsilon(x)$ est un $\mathbb{R}$-arbre avec la distance induite par X.

Autrement dit, localement un $\mathbb{R}$-graphe ressemble à un $\mathbb{R}$-arbre. Nous avons vu comment

les $\mathbb{R}$-arbres peuvent être obtenus à partir d'excursions. De façon similaire, on peut obtenir des $\mathbb{R}$-graphes. Soit $a > 0$ un réel et h une excursion sur $[0, a]$, soit $\mathscr{P} \in \mathbb{R}^+ \times \mathbb{R}^+$ un sous-ensemble dénombrable, posons :

$$h \cap \mathscr{P} = \{(x, y) \in \mathscr{P};\ 0 \leq x \leq a,\ 0 \leq y < h(x)\}\,.$$

Supposososns de plus que $|h \cap \mathscr{P}| < \infty$, dans ce cas, il existe $k \in \mathbb{N}$ tel que $h \cap \mathscr{P} = \{(x_i, y_i);\ 1 \leq i \leq k\}$. Pour tout $1 \leq i \leq k$ on pose

$$r(x_i, y_i) = \inf\{x \geq x_i;\ h(x) \leq y_i\},$$

et on identifie la classe d'équivalence de x_i et celle de $r(x_i, y_i)$ dans $\mathscr{T}_h$, le $\mathbb{R}$-arbre correspondant à h. On obtient ainsi un espace métrique $(\mathscr{G}(h, \mathscr{P}), \tilde{d}_h)$, avec $\tilde{d}_h$ qui est la distance naturelle induite par celle de $\mathscr{T}_h$ sur $\mathscr{G}(h, \mathscr{P})$. Cette espace métrique est un $\mathbb{R}$-graphe (Voir l'Illustration 1.20 pour un exemple de cette construction). Soit B un mouvement Brownien standard, nous

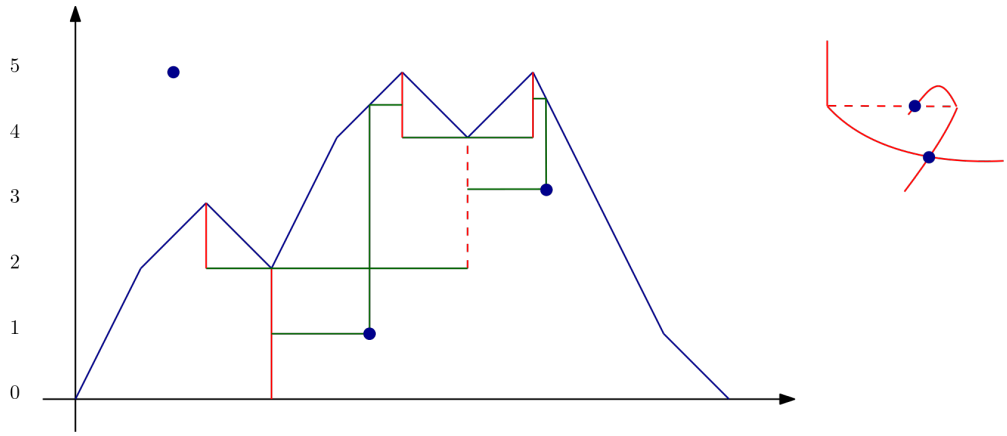


FIGURE 1.20 – En bleu, une excursion. Les petits disques bleus correspondent à des points du plans. Deux de ces points sont en dessous de l'excursion. En rouge, un plongement dans le plan du $\mathbb{R}$-graphe associé. Nous avons dessiné en traits pointillés un segment particulier du $\mathbb{R}$-graphe pour bien préciser le plongement.

définissons :

$$F(W)^\lambda(s) = \sqrt{\frac{\mathbb{E}[W^3]}{\mathbb{E}[W]}} B(s) + \lambda s - \frac{s^2 \mathbb{E}[W^3]}{2\mathbb{E}[W]^2}.$$

Aldous (Aldous [1997]) a montré que les excursions du processus réfléchi $(F(W)^\lambda(s) - \min_{u\in[0,s]} F(W)^\lambda(u))_{s\geq 0}$ peuvent être ordonnées de façon décroissante et que la suite des tailles de ces excursion est dans $\ell_2^\downarrow$, l'espace des suites décroissantes ayant une norme ℓ^2 finie. De plus, ce même article pour les graphes d'Erdős-Rényi, puis dans l'Article Bhamidi et al. [2010] pour les graphes inhomogènes avec un troisième moment fini, il est démontré que si l'on écrit les tailles des excursions comme $\mathbf{Z}(W) = (Z_1(W), Z_2(W), ...)$ et qu'on écrit $\mathscr{G}(\mathbf{W}_n, p(\lambda)) = (\mathscr{G}(\mathbf{W}_n, p(\lambda), i))_{i\geq 1}$., avec $p(\lambda) = \frac{1}{\ell_n} + \frac{\lambda}{\ell_n^{4/3}}$, pour les composantes connexes de $G(\mathbf{W}_n, p(\lambda))$ prisent dans l'ordre décroissant, alors on a quand $n \to \infty$:

$$\frac{|G(\mathbf{W}_n, p(\lambda))|}{n^{2/3}} \xrightarrow{\mathrm{d}} \mathbf{Z}\,(W)\,,$$

pour la topologie ℓ^2.

On dénote par $\delta(1)$ une variable aléatoire ayant comme distribution le Dirac en 1, si $W = \delta(1)$ alors $(G(\mathbf{W}_n, p(\lambda)))_{n\geq 1}$ est un graphe d'Erdős-Rényi. Prenons les excursions au dessus de 0

du processus réfléchit $(F(W)^\lambda(s) - \min_{u\in[0,s]} F(W)^\lambda(u))_{s\geq 0}$. Translatons ces excursions afin qu'elles commencent en $(0,0)$ et prenons les dans l'ordre décroissant $(\zeta^1, \zeta^2, ...,)$. Soit $(\mathcal{P}_i)_{i\geq 1}$ une suite i.i.d. de processus de Poisson dans $\mathbb{R}^+ \times \mathbb{R}^+$. Considérons la suite de $\mathbb{R}$-graphes $\mathcal{G}_\lambda(W) = \big(\mathcal{G}(\zeta^1, \mathcal{P}_1), \mathcal{G}(\zeta^2, \mathcal{P}_2), ...\big)$ construite par la procédure que l'on vient de décrire. Addario-Berry et al. [2012] a démontré que :

Limite d'échelle des graphes d'Erdős-Rényi

Soit $\lambda > 0$, et soit $(G(\mathbf{W}_n, p(\lambda))_{n\geq 1}$ une suite i.i.d. de graphes d'Erdős-Rényi. Dans ce cas nous avons $W = \delta(1)$. Posons $\mathscr{G}(\mathbf{W}_n, p(\lambda)) = (\mathscr{G}(\mathbf{W}_n, p(\lambda), i))_{i\geq 1}$ pour la suite des composantes connexes de $G(\mathbf{W}_n, p(\lambda))$ prisent dans l'ordre décroissant de leurs tailles et vues comme $\mathbb{R}$-graphes avec des distances d'arêtes renormalisées par $n^{-1/3}$. Alors nous avons la convergence suivante quand $n \to \infty$

$$\mathscr{G}(\mathbf{W}_n, p(\lambda)) \xrightarrow{\mathrm{d}} \mathcal{G}_\lambda(\delta(1))$$

pour la topologie produit.

En fait ce résultat est démontré pour une topologie plus fine que la topologie produit, nous renvoyons au Chapitre 4 pour une brève description de cette topologie. Ce résultat a été étendu dans Broutin et al. [2020] pour les graphes inhomogènes, dans le cas où W a un troisième moment fini, leur résultat implique le théorème suivant.

Limite d'échelle des graphes inhomogènes à troisième moment fini

Soit $\lambda > 0$, et soit $(G(\mathbf{W}_n, p_\lambda))_{n\geq 1}$ une suite i.i.d. de graphes aléatoires inhomogènes. Supposons de plus que W a un troisième moment fini, et posons

$$\kappa = \frac{\mathbb{E}[W]}{\mathbb{E}[W^3]^{2/3}}.$$

Soit $\kappa\mathcal{G}_\lambda(\kappa\delta(1))$ l'espace métrique $\mathcal{G}_\lambda(\kappa\delta(1))$ dont toutes les distances ont été multipliées par κ. On a alors la convergence suivante quand $n \to \infty$

$$\mathscr{G}(\mathbf{W}_n, p(\lambda)) \xrightarrow{\mathrm{d}} \kappa\mathcal{G}_\lambda(\kappa\delta(1))$$

pour la topologie produit.

Dans ce théorème le facteur κ signifie que les distances dans $\mathcal{G}_\lambda(\kappa\delta(1))$ sont renormalisées par κ.

1.6 Résultats de cette thèse et questions futures

Chapitre 2

Dans le Chapitre 2 nous démontrons plusieurs résultats concernant la structure des graphes inhomogènes dans la fenêtre critique. Ce chapitre traite le cas où la loi des poids μ a un troisième moment fini. L'étude qu'on fait dans ce chapitre a une double utilité. Elle est utilisée dans le chapitre suivant pour l'étude de l'arbre couvrant minimum inhomogène. De plus, elle démontre que des propriétés qui sont connues pour être vraies dans le cas du graphe d'Erdős-Rényi restent vraies dans ce cas là aussi.

Plus précisément nous démontrerons que :

Inégalités de concentration pour les graphes inhomogènes

Soit $f = f(n) = o(n^{1/3})$, et $p_f = \frac{1}{\ell_n} + \frac{f}{\ell_n^{4/3}}$. Il existe $F > 0$ et $A > 0$ assez grands tels que pour tout $1 > \varepsilon > 0$ si $f \geq F$, dans le graphe $G(\mathbf{W}, p_f)$ les propriétés suivantes

- la plus grande composante connexe a une taille de l'ordre de $f\ell_n^{2/3}$,
- la somme des poids des noeuds de la plus grande composante connexe est aussi de l'ordre de $f\ell_n^{2/3}$,
- le surplus de la plus grande composante connexe est un $O(f^3)$,
- la taille de n'importe quelle autre composante connexe ne dépasse pas $\frac{\ell_n^{2/3}}{f^{1-\varepsilon}}$,
- le poids de n'importe quelle autre composante connexe ne dépasse pas $\frac{\ell_n^{2/3}}{f^{1-\varepsilon}}$,
- le surplus de n'importe quelle autre composante connexe ne dépasse pas f^{ε},

sont vraies avec probabilité au moins

$$1 - A\left(\exp\left(\frac{-f_n^{\varepsilon/2}}{A}\right) + \exp\left(\frac{-\sqrt{f_n}}{A}\right) + \exp\left(\frac{-n^{1/12}}{A}\right)\right).$$

Un des points focaux du Chapitre 2 est la démonstration d'une inégalité de concentration qui ressemble à l'inégalité de Bernstein (Bernstein [1924]) pour les tirages biaisés sans remise. Rappelons que $(v(1), v(2), ..., v(n))$ correspond à un ordre aléatoire sur les poids des noeuds obtenus par un tirage biaisé par la taille sans remise. On dit que $(a(n), b(n))_{n\geq 1}$ vérifie les Conditions 1 si :

— Pour tout n assez grand on a

$$\exp\left(\frac{-b(n)^2}{(b(n)w_{\max} + a(n))}\right) < 1/4.$$

— Les deux limites suivantes sont vraies

$$\text{liminf}_{n\to\infty} a(n) = \text{liminf}_{n\to\infty} b(n) = +\infty$$

— $a(n) = o(n)$.
— $b(n) = O(m)$
— $a(n) = O\left(b(n)n^{1/3}\right)$.
— $a(n)^2 = O\left(b(n)n\right)$.

Nous démontrons dans le Chapitre 3 l'inégalité de concentration suivante

Inégalité de concentration pour les tirages biaisés sans remises

Soit μ une loi de probabilité ayant un troisème moment fini. Pour $n \geq 1$ soit $(w_1^n, w_2^n, ..., w_n^n)_{n\geq 1}$ une suite de vecteurs de poids i.i.d. de loi μ. Notons $w(n)_{\max} = \max_{1\leq i\leq n}(w_k^n)$. Supposong que $(m(n), y(n))_{n\geq 1}$ verifie les Conditions 1. Définissons l'événement suivant :

$$\mathcal{B}(n) := \left\{ \sup_{1\leq i\leq j\leq m(n)} \left| \sum_{k=i}^{j} w_{v(k)}^n - \mathbb{E}\left[\sum_{k=i}^{j} w_{v(k)}^n\right] \right| \geq y(n) \right\}.$$

Alors, il existe une constante $A > 0$ qui ne dépend que de la lois de μ telle que pour tout $n \geq 1$:

$$\mathbb{P}\left[\mathcal{B}(n)\right] \leq A \exp\left(\frac{-y(n)^2}{A(y(n)w(n)_{\max} + m(n))}\right).$$

Les Conditions 1 sont façonnées aux besoins du Chapitre 3, mais la méthode utilisée pour démontrer l'inégalité de concentration avec ces conditions peut être réutilisée pour démontrer des inégalités similaires avec des conditions différentes. A notre connaissance, ce résultat est nouveau. Ben-Hamou, Peres, and Salez [2018] avait effectivement démontré des inégalités de concentration pour les tirages biaisés sans remise, cependant, ces inégalités ne concernent que la borne supérieure. La seule inégalité qu'ils obtiennent et qui donne une borne supérieure et inférieure est une modification de l'inégalité de Hoeffding. Cette dernière n'est malheureusement pas assez forte pour notre application.

Afin de démontrer cette inégalité de concentration, nous utilisons un passage par un temps aléatoire. Au lieu de tirer les poids un par un de manière discrète, on associe à chaque poids w_i une variable exponentielle T_i de paramètre $\frac{w_i}{\ell_n}$ indépendamment de tout le reste. Puis, on considère qu'au temps t, on a tiré tous les poids i pour lesquels $T_i \leq t$. Si on note $N(t)$ le nombre de poids tirés au temps t, par les propriétés des variables exponentielles, conditionnellement à $N(t)$, les poids tirés de cette manière on la même distribution que les $(w_{v(i)})_{i\leq N(t)}$. Ainsi on ramène l'étude de la somme sur des temps discrets à une étude sur une somme à des temps aléatoires. L'intérêt de ce changement de point de vue est qu'on peut maintenant utiliser pleinement des théorèmes connus sur les sommes de variables aléatoires indépendantes et sur les martingales.

Nous attirons l'attention du lecteur sur deux points. Tout d'abord, cette astuce a déjà été utilisée auparavant, par exemple par Aldous dans Aldous [1997], pour des problèmes différents de celui là. Cependant, afin de prouver notre inégalité nous utilisant la méthode de chaînage (voir Boucheron et al. [2013] chapitres 12 et 13), qui à notre connaissance n'a pas été utilisé auparavant dans ce cadre. Nous pensons aussi que cette méthode en particulier, et les inégalités de concentration pour des tirages biaisés sans remise de façon générale mériteraient plus d'attention. Le champ d'application de telles inégalités dépasse le domaine des graphes aléatoires (voir la discussion dans Ben-Hamou et al. [2018] par exemple), et pour l'instant, aussi loin qu'on sache, il n'y a très peu de littérature dessus.

Le régime dans lequel $f = f(n)$ tend vers l'infini tout en ayant $f(n) = o(n^{1/3})$ est en général appelé régime quasi-surcritique. Un des intérêts des résultats du Chapitre 2 est qu'on obtient une description assez précise de ce régime important qui marque le passage entre régime critique et régime surcritique dans les graphes inhomoghènes. Par exemple un corollaire direct des théorèmes démontrés dans le Chapitre 2 est le suivant :

Convergence des tailles des composantes connexes dans le régime quasi-surcritique.

Supposons que μ, la loi correspondant au vecteur $\mathbf{W}$, a un troisème moment fini. Soit W une variable aléatoire de loi μ. Soit $(f(n))_{n\geq 1}$ une fonction qui tend vers l'infini et telle que $f(n) = o(n^{1/3})$. Notons par $(|C_1|, |C_2|, |C_3|, ...)$ la suite infinie des tailles des composantes connexes de $G(\mathbf{W}, p_{f(n)})$ prises dans l'ordre décroissant, avec la convention $|C_i| = 0$ si il n'existe pas de i-ième plus grande composante connexe. Cette suite dépend implicitement de n. Alors, la convergence suivante

$$\left(\frac{|C_1|}{2f(n)\ell_n^{2/3}}, \frac{|C_2|}{\ell_n^{2/3}}, \frac{|C_3|}{\ell_n^{2/3}}, \frac{|C_4|}{\ell_n^{2/3}}, ...\right) \to \left(\frac{\mathbb{E}[W^3]}{\mathbb{E}[W]}, 0, 0, ...\right)$$

a lieu en probabilité dans l'espace $\mathbb{R}^{\mathbb{N}}$ munit de la norme ℓ^p pour tout $p > 2$.

La démonstration de ce théorème se trouve dans le Chapitre 2.

Chapitre 3

Le Chapitre 3 utilise les résultats du Chapitre 2 en plus de couplages originaux pour démontrer que l'espérance du diamètre de l'arbre couvrant minimum inhomogène est de l'ordre de $n^{1/3}$ avec les conditions de troisième moment fini. Le théorème principal de ce chapitre est donc le

suivant.

La moyenne du diamètre de l'arbre couvrant minimum inhomogène

Supposons que μ est une loi de probabilité sur $\mathbb{R}^+$ de troisème moment fini. Pour $n \geq 1$, soit $\mathbf{W}(n)$ un vecteur de poids aléatoires i.i.d. de loi μ. Et soit T_n un arbre couvrant minimum inhomogène associé au vecteur $\mathbf{W}(n)$. Il existe des constantes $a > 0$ et $A > 0$ qui ne dépendent que de μ et telles que, pour tout $n \geq 1$, $\mathrm{diam}(T_n)$, le diamètre de T_n vérifie :

$$an^{1/3} \leq \mathbb{E}[\mathrm{diam}(T_n)] \leq An^{1/3}.$$

Une partie des résultats nécessaires à la démonstration de ce théorème se trouve dans le Chapitre 2. En effet, ce résultat, tout comme celui pour les arbres couvrants minimums aléatoires de Addario-Berry et al. [2006], est démontré en utilisant le couplage présenté plus tôt entre processus de graphes aléatoires et processus d'arbres couvrants aléatoires fait grâce à l'algorithme de Kruskal. Et la partie importante dans ce couplage est le passage du régime critique au régime quasi-surcritique. En effet, le diamètre des arbres couvrants minimums des composantes connexes de $G(\mathbf{W}, p)$ atteint l'ordre de grandeur $n^{1/3}$ en moyenne dans la fenêtre critique. Il est assez facile de démontrer que dans le passage entre le régime surcritique et le graphe complet (i.e pour $p = +\infty$) l'ordre de grandeur ne change pas. Il suffit alors pour finir la démonstration de prouver que le passage entre régime critique et quasi-surcritique se fait sans changement d'ordre. C'est cette partie qui est délicate, et qui demande une étude fine des graphes inhomogènes. Dans les grandes lignes, on donne des bornes fortes sur les évènements suivants :

- Le plus long chemin sans cycles dans la composante connexe géante ne dépasse pas un ordre de grandeur de $f^5 \ell_n^{1/3}$, où f est la constante de la fenêtre critique définie plus haut.
- Toutes les autres composantes connexes ont des plus longs chemins sans cycles dont la taille ne dépasse pas $\frac{5\ell_n^{1/3}}{f^{1/4}}$.

En utilisant ces bornes, et les bornes déjà obtenues au Chapitre 2 on arrive à démontrer le théorème. Nous renvoyons le lecteur au Chapitre 3 pour une preuve complète.

Afin d'avoir des bornes sur les plus longs chemins des composantes connexes, on utilise les bornes du Chapitre 2 sur le surplus des composantes et des bornes sur les hauteurs d'arbres couvrants des dites composantes. Ceci afin de pouvoir d'utiliser le Théorème 1.4.1. Ce sont ces bornes sur les hauteurs d'arbres couvrants qui posent problème. En effet, comme dit précédemment, il n'y a pas d'arbres "simples" qui recouvrent les composantes connexes de graphes inhomoghènes critiques. Le mieux qu'on puisse faire, c'est voir les arbres correspondant au codage par comptage de nombre d'enfants associé à un parcours en largeur comme des p-arbres biaisés. Hors, il n'y a, à notre connaissance, pas de résultats assez forts donnant des bornes sur les hauteurs des p-arbres, et encore moins sur celles des p-arbres biaisés. Nous avons donc opté pour des couplages en loi qui permettent de trouver ces p-arbres biaisés comme des sous-arbres d'arbres de Galton-Watson. Des couplages similaires ont déjà été utilisé par Broutin et al. [2020] pour leurs théorèmes de limite d'échelle générale. Mais dans notre cas, on utilise en plus un algorithme d'élagage et un argument récursif original.

L'idée principale consiste à créer des arbres de Galton-Watson enracinés avec des lois de Poisson de paramètre ayant la loi de $\tilde{W}$, où $\tilde{W}$ correspond à un poids tiré de manière biaisée par la taille parmi les poids dans $\mathbf{W}$. La construction de cet arbre peut être faite en largeur en tirant d'abord le poids d'un noeud, puis en lui associant un nombre d'enfants de loi de Poisson de paramètre ce poids. Ainsi, à chaque noeuds on peut associer une étiquette, qui correspond à l'indice du poids qu'on lui a octroyé. On colorie ensuite les noeuds de cet arbre en faisant un parcours en largeur. Un noeud est colorié en rouge si et seulement si son père est rouge et au moment où il est découvert dans le parcours il n'existe pas d'autres noeuds rouges ayant la même étiquette que lui. Un exemple de ce coloriage est fournit dans l'Illustration 1.21, et nous

renvoyons vers le Chapitre 3 pour une description précise de cet algorithme de coloriage. L'arbre rouge obtenu sera un p-arbre biaisé ayant la même loi que le premier p-arbre biaisé exploré dans le parcours en largeur d'un graphe inhomogène. Cette construction s'étend facilement au reste des composantes connexes. Au lieu de borner la hauteur des arbres rouges, on borne celle des arbres de Galton-Watson qui les contiennent. On peut donc utiliser tous les résultats connus sur les hauteurs d'arbres de Galton-Watson (par exemple dans Łuczak [1990] et Luczak [1991]). Cette construction n'est quand même pas suffisante, car dans le régime qu'on considère, il y aura certains arbres de Galton-Watson infinis qui seront créés par cette construction. Notamment l'arbre de Galton-Watson qui contient une copie de l'arbre couvrant de la composante connexe géante du graphe inhomogène en loi aura de grandes chances d'être infini. Pour remédier à celà, on fait un élagage du p-arbre biaisé qui recouvre la composante connexe géante. On lui enlève, de façon i.i.d., avec une certaine probabilité certaines arêtes. On obtient donc une forêt et on fait le couplage en loi sur les arbres élagués. Tout ceci est expliqué de façon formelle dans le Chapitre 3.

Chapitre 4

Dans ce chapitre on s'attaque à la question de la limite d'échelle de l'arbre couvrant minimum. Pour une suite de graphes inhomogènes surcritiques, posons des poids i.i.d. sur leurs arêtes et dénotons par $(T'(\mathbf{W}_n))_{n\geq 1}$ la suite des arbre couvrant minimums de leurs composantes connexes géantes. On définit les $\mathbb{R}$-arbres $(\mathscr{T}'(\mathbf{W}_n))_{n\geq 1}$ comme étant les espaces métriques obtenus à partir des $(T'(\mathbf{W}_n))_{n\geq 1}$ en considérant chaque arête comme un segment de longueur 1 et en divisant toutes les distances obtenues par $n^{1/3}$. Nous avons alors le théorème suivant :

Convergence de l'arbre couvrant minimum inhomogène

Supposons que W a un troisième moment fini. Alors il existe un $\mathbb{R}$-arbre compact aléatoire $\mathscr{M}(W)$ tel que, quand $n \to \infty$

$$\mathscr{T}'(\mathbf{W}_n) \xrightarrow{\mathrm{d}} \mathscr{M}(W),$$

dans l'espace $(\mathcal{M}, d_{\mathrm{GH}})$.

De plus nous démontrons que $\mathscr{M}(W)$ est universel dans le sens où, si W' a aussi un troisème moment fini, alors la distribution de $\mathscr{M}(W')$ est la même (à une constante multiplicative prêt [1]) que celle de $\mathscr{M}(W)$. En particulier, elle est identique à celle obtenue par Addario-Berry et al. [2017b] pour le cas des graphes d'Erdős-Rényi. Ainsi, certaines des propriétés déjà connues pour cette dernière resteront vraies pour notre limite. Par exemple, les arbres $\mathscr{M}(W)$ sont presque sûrement binaires.

Afin de démontrer ce théorème nous suivons la méthode utilisée dans Addario-Berry et al. [2017b] pour la limite d'échelle de l'arbre couvrant minimum aléatoire. Afin de garder ce chapitre le plus concis possible, nous ne redémontrons pas les résultats concernant les espaces métriques et renvoyons le lecteur directement à Addario-Berry et al. [2017b].

Nous commençons ce chapitre par des rappels des objets utilisés, puis nous discutons nos résultats de convergence et expliquons le rapport qu'ils ont avec des résultats précédents, comme celui de Addario-Berry and Sen [2019]. Ensuite nous introduisons toutes les notions d'espaces métriques dont nous aurons besoin, ainsi que certains théorèmes utiles. Nous passons ensuite à une description formelle de la procédure de découpage qui permet d'obtenir l'arbre couvrant minimum discret, ainsi que son homologue continu à partir des graphes inhomogènes. Cette procédure correspond dans le cas discret à une version randomisée de l'algorithme de suppression d'arêtes discuté dans cette introduction. Dans le cas continue, elle correspond à une version intuitive de ce même algorithme qui permet d'obtenir des $\mathbb{R}$-arbres à partir de $\mathbb{R}$-graphes. Nous

1. Nous renvoyons à 4 pour une description formelle de cette distribution

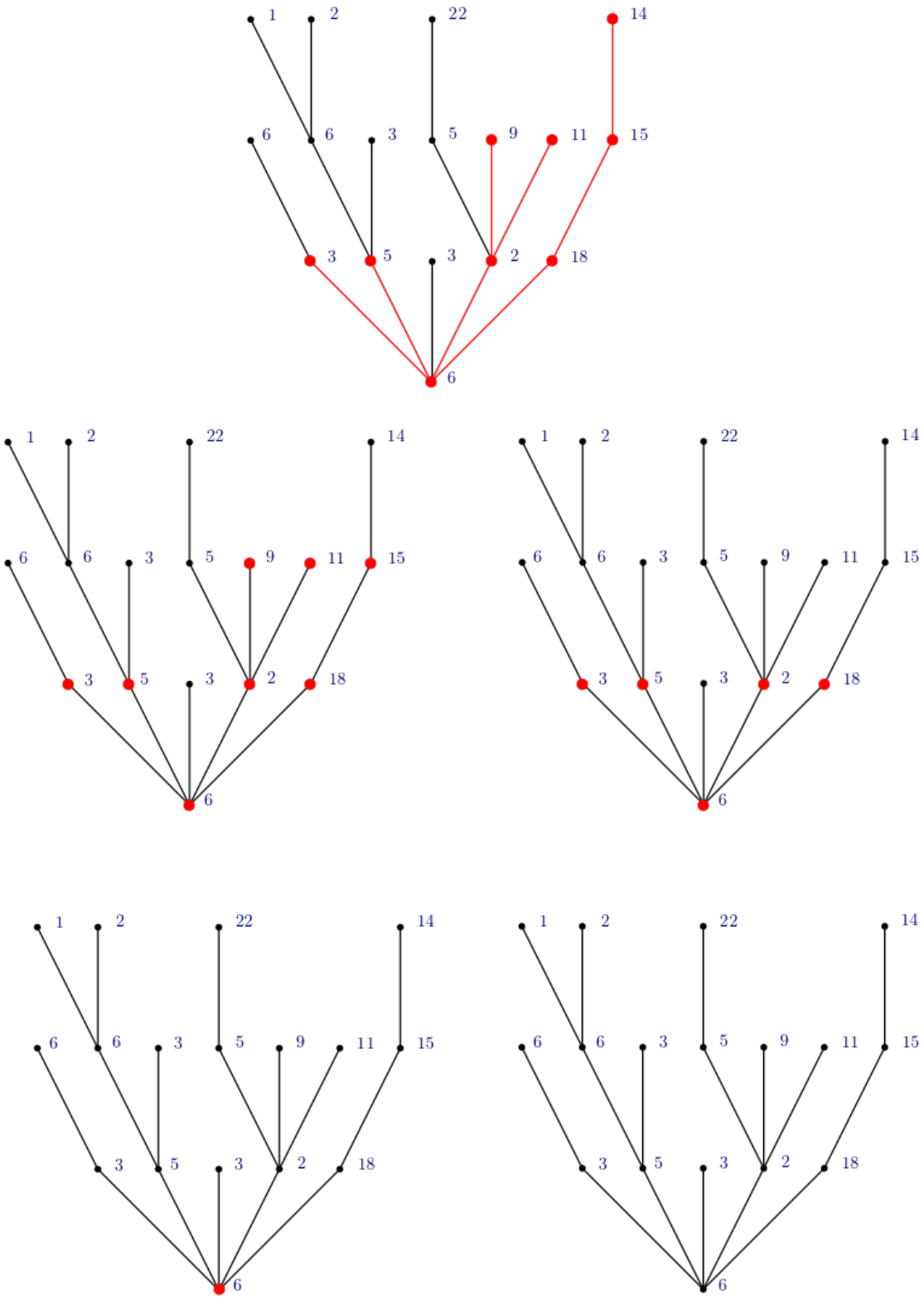


FIGURE 1.21 – Un exemple de la procédure de coloriage. L'ordre pris ici est implicite dans le dessin, de bas en haut et de gauche à droite. Chaque noeud à son étiquette directement à sa droite. On a dessiné les étapes génération par génération.

démontrons aussi comment il est possible de relier la procédure dans le cas discret à celle dans le cas continu soit en discrétisant les $\mathbb{R}$-graphes, ou en transformant les graphes discrets en $\mathbb{R}$-graphes.

Dans la suite de ce chapitre nous rappelons les résultats déjà connus de convergence des graphes inhomogènes, et en utilisant ces résultats nous présentons un premier théorème de convergence des arbres couvrant minimaux de graphes inhomogènes mais seulement dans la fenêtre critique. Ensuite, nous nous attaquons à deux problèmes de tension. Le premier concerne le problème connu sous le nom de "problème du leader". Ce dernier consiste à montrer que, si on regarde le processus croissant de graphes inhomogènes, quand le paramètre critique λ croit, à partir d'un certain moment les composantes connexes géantes des différents graphes inhomogènes pour différents λ deviennent toutes inclues les unes dans les autres avec grande probabilité. Ce résultat a été démontré par Łuczak [1990] pour le cas des graphes d'Erdős-Rényi, puis par Addario-Berry, Bhamidi, and Sen [2017a] dans le cas des graphes inhomogènes. Le deuxième problème de tension consiste à démontrer que, quand le paramètre critique λ croit, la distance de Hausdorff entre les composantes connexes géantes pour des paramètres λ différents ne croit pas trop vite avec une bonne probabilité. Ce dernier résultat utilise de façon cruciale les résultats des chapitres précédents. Le théorème de convergence finale est obtenu comme conséquence directe de tous ces résultats. Nous renvoyons le lecteur au Chapitre 4 pour des énoncés formels de tous ces résultats.

Questions futures

Ce travail ouvre la voie à plusieurs résultats possibles. Tout d'abord autre direction de recherche concerne les tirages sans remise. Soit $n \geq 1$ un entier et $a_1 \geq a_2 \geq ... \geq a_n)$ une suite finie décroissante de réels. De plus soit $(p_1, p_2, ..., p_n)$ un vecteur de probabilité :

$$\sum_{k=1}^{n} p_i = 1.$$

Soit $(V(1), V(2), ..., V(n))$ un mélange aléatoires des indices $(1, 2, ..., n)$ qui correspond à un tirage biaisé sans remise :

$$\begin{aligned}
&\forall i \leq n-1, \quad \forall j \leq n \\
&\mathbb{P}\left(V(1) = j\right) = p_j. \\
&\mathbb{P}\left(V(i+1) = j \quad |(V(1), ..., V(i))\right) = \frac{p_j \mathbb{1}(V(j) \notin (V(1), ..., V(i)))}{\sum_{k=1}^{n} p_k - \sum_{k=1}^{i} p_{V(k)}}.
\end{aligned}$$

Notons $(J(1), J(2), ..., J(n))$ pour un vecteur de variables aléatoires i.i.d. ayant la même loi que $V(1)$. Remarquons que les tirages sans remise biaisés par la taille présentés plus tôt dans cette introduction sont des cas particuliers de tirages biaisés. Nous posons ici deux questions concernant ces tirages biaisés sans remise. Nous renvoyons le lecteur au Chapitre 2 pour une explication un peu plus détaillée de l'intuition derrière ces questions. Premièrement, sous quels conditions sur les probabilités et sur les poids a-t-on pour tous $n \geq m \geq l$ et $x \geq 0$:

$$\mathbb{P}\left(\left|\sum_{k=1}^{m} a_{V(i)} - \mathbb{E}[a_{V(i)}]\right| \geq x\right) \leq \mathbb{P}\left(\left|\sum_{k=1}^{m} a_{J(i)} - \mathbb{E}[a_{J(i)}]\right| \geq x\right).$$

Cette conjecture signifie que le tirage sans remise est plus concentré autour de sa moyenne que le tirage avec remise. L'intuition derrière vient du fait que les tirages sans remise ont tendance à s'auto-réguler. Si on tire un a_i qui est extrêmement grand ou petit on ne va plus le retirer après. Par contre pour les tirages avec remise, les évènements "extrêmes" peuvent survenir plusieurs fois. Nous avons testé cette conjecture avec plusieurs simulations, en changeant le biais et en testant des cas extrêmes. Elle a tenue à chaque fois. Si la conjecture n'est pas vraie, il serait encore

plus intéressant de savoir sous quelles hypothèses sur le tirage elle est vraie. Pour l'instant, le seul résultat qui se rapproche de cette conjecture est le cas des tirages uniformes. Par exemple Serfling [1974] démontre que dans ce cas les tirages sans remise vérifient des inégalités de concentration plus fortes que celle issues de la borne de Chernoff pour les tirages avec remise. Par exemple pour l'inégalité de Hoeffding classique on a :

$$\mathbb{P}\left(\left|\sum_{k=1}^{m} a_{J(i)} - \mathbb{E}[a_{J(i)}]\right| \geq x\right) \leq \exp\left(\frac{-2x^2}{m(a_{\max} - a_{\min})^2}\right),$$

où $a_{\max}$ et $a_{\min}$ sont respectivement les maximums et minimums des $(a_i)_{i\leq n}$. Pour les tirages sans remise on a l'inégalité de Hoeffding-Serfling (Serfling [1974]) si le tirage sans remise est uniforme alors on a :

$$\mathbb{P}\left(\left|\sum_{k=1}^{m} a_{V(i)} - \mathbb{E}[a_{V(i)}]\right| \geq x\right) \leq \exp\left(\frac{-2x^2}{m(1-\frac{m-1}{n})(a_{\max} - a_{\min})^2}\right),$$

ici le terme $(1 - \frac{m-1}{n})$ vient se rajouter, et il rend la borne beaucoup plus petite quand m se rapproche de n. Ce qui est naturel, vu que si on fait n tirages sans remise alors on aura tiré chaque a_i exactement une fois.

Pour la deuxième question supposons que $p_1 \geq p_2 \geq ... \geq p_n$. Les plus grands a_i ont le plus de chance d'être tirés. Est-il vrai que pour tout $n - 1 \geq m \geq 1$, et vecteur de nombres réels $(x_1, x_2, x_3, ...x_n)$

$$\mathbb{P}\left(a_{V(1)} \geq x_1 \cap ... \cap a_{V(m)} \geq x_m\right) \geq \mathbb{P}\left(a_{V(2)} \geq x_1 \cap a_{V(3)} \geq x_2 ... \cap a_{V(m+1)} \geq x_m\right),$$

et aussi

$$\mathbb{P}\left(a_{J(1)} \geq x_1 \cap ... \cap a_{J(m)} \geq x_m\right) \geq \mathbb{P}\left(a_{V(1)} \geq x_1 \cap ... \cap a_{V(m)} \geq x_m\right).$$

Encore une fois, l'intuition derrière cette conjecture est que plus on tire sans remise de façon biaisée "croissante" moins on aura de chances de tirer les gros poids. Car, ces gros poids ont de grandes chances d'avoir été tirés au tout début. Nous avons aussi effectué plusieurs simulations pour cette conjecture, et elle semble être vraie. Ce résultat est démontré dans un cas très particulier dans le Chapitre 3. Le cas général reste ouvert.

Répondre à ces questions de manière générale ouvrira aussi la voie vers d'autres résultats concernant les arbres couvrants minimums inhomogènes. Notamment la question des conditions d'échelle libre. Quand la loi μ avec laquelle on génère les poids des noeuds n'a plus un troisième moment fini mais a une queue similaire à une loi de puissance de paramètre $3 < \tau < 4$. Dans ce cas quel est le diamètre de l'arbre couvrant minimum inhomogène issu de cette loi ? Si on part du principe que le diamètre de l'arbre couvrant minimum dans ce cas est lui aussi du même ordre de grandeur que les distances typiques dans les composantes connexes du graphe inhomogène critique. Alors, suivant Bhamidi et al. [2018], la moyenne du diamètre de l'arbre couvrant minimum associé à une loi "d'échelle libre" de paramètre τ devrait être de l'ordre de $n^{\frac{\tau-3}{\tau-1}}$. Remarquons que si c'est le cas, $\tau = 4$ donnerait un résultat similaire au Chapitre 3.

D'un autre côté il reste plusieurs questions ouvertes concernant la limite d'échelle obtenue dans le Chapitre 4. Tout d'abord, la méthode utilisée dans ce chapitre, et avant par Addario-Berry et al. [2017b], pour démontrer la convergence de l'arbre couvrant minimum aléatoire vers un arbre continu repose sur un argument de tension. L'arbre limite n'est pas bien compris. Par exemple, on ne connaît pas la loi des distances typiques dans cet arbre, ni celle de son diamètre. On ne sait pas non plus quel type d'excursion correspond à ce $\mathbb{R}$-arbre. Un point de vue différent sur cette convergence semble nécessaire afin d'obtenir une définition plus constructive de cette limite d'échelle, et de calculer explicitement certaines de ses fonctionnelles. Il serait aussi intéressant d'étendre notre convergence à d'autres types de graphes, on pense intuitivement que tous les graphes exhibants un comportement d'échelle moyenne (mean-field) devrait avoir le même arbre

minimum limite si on leur associe des poids i.i.d.. Finalement, il reste la question de la limite d'échelle de l'arbre couvrant minimum inhomogène dans le cas des lois d'échelle libre. Dans ce cas, on estime que la limite sera complètement différente de celle obtenue pour le troisème moment fini.

Chapitre 2

Exponential bounds for inhomogeneous random graphs

Contents

2.1 Introduction 55
2.1.1 The model 55
2.1.2 Definition of the exploration process 55
2.1.3 Conditions and main theorem 57
2.1.4 Motivation and previous work 60
2.2 Bounding the weights 62
2.2.1 First concentration result and the mean 63
2.2.2 A more precise concentration inequality 66
2.3 Bounds on the exploration process 75
2.4 The structure of the giant component 84
2.4.1 The size of the giant component 84
2.4.2 The excess of the giant component. 88
2.4.3 The excess of the components discovered before the largest connected component. 91
2.5 The structure of the tail's components 93
2.5.1 Preliminaries 93
2.5.2 The size of connected components discovered after the largest connected component 97
2.5.3 The excess of the tail 102

Exponential bounds for inhomogeneous random graphs in the Gaussian case

Rank-1 inhomogeneous random graphs are a natural generalization of Erdős-Rényi random graphs. In this generalization each node is given a weight. Then the probability that an edge is present depends on the product of the weights of the nodes it is connecting. In this chapter, we give precise and uniform exponential bounds on the size, weight and surplus of rank-1 inhomogeneous random graphs where the weight of the nodes behave like a random variable with finite third moments. We focus on the case where the mean degree of a random node is equal to 1 (critical regime), or slightly larger than 1 (barely supercritical regime). These bounds will be used in follow up chapters to study a general class of random minimum spanning trees. They are also of independent interest since they show that these inhomogeneous random graphs behave like Erdős-Rényi random graphs even in a barely supercritical regime. The proof relies on novel concentration bounds for sampling without replacement and a careful study of the exploration process.

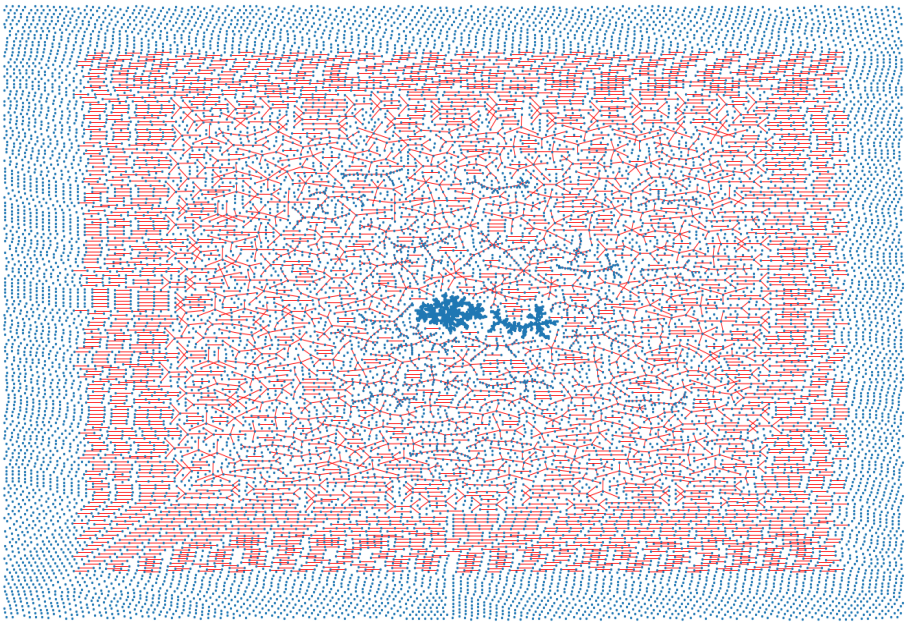


FIGURE 2.1 – An inhomogeneous random graph of size $n = 20000$. The node weights are i.i.d with Pareto distribution of parameters $2/3, 4$, and $p = \frac{5}{4n}$. These paremeters correspond to typical graphs that will be studied in this chapter.

2.1 Introduction

2.1.1 The model

Consider $n \in \mathbb{N}$ vertices labeled $1, 2, ..., n$. For a vector of weights $\mathbf{W} = (w_1, w_2, ..w_n)$, where $0 < w_n \leq w_{n-1} \leq ... \leq w_1$, we create the inhomogeneous random graph associated to $\mathbf{W}$ and to $p \leq +\infty$ in the following way :

Each potential edge $\{i, j\}$ is in the graph with probability $1 - e^{-w_i w_j p}$ independently from everything else. This gives a random graph that we call the rank-1 inhomogeneous random graph associated to $\mathbf{W}$ and $p \leq +\infty$.

One can couple the graphs for the different values of p as follow : Let K_n be the complete graph of size n. To every potential edge $\{i, j\}$, associate independently the random capacity $E_{\{i,j\}}$ which is an exponential random variable of rate $w_i w_j$. The weights are then used to create a sequence of graphs. For each $p \in [0, +\infty]$ let $G(\mathbf{W}, p)$ be the graph on $\{1, 2...., n\}$ containing the edges of weight at most p. So the edge set of $G(\mathbf{W}, p)$ is :

$$\left\{\{i, j\} | E_{\{i,j\}} \leq p\right\}.$$

Then $(G(\mathbf{W}, p))_{p\in[0,+\infty]}$ is an increasing sequence of graphs for inclusion, and for each fixed value of p, this constructions matches the first one. We will use both construction interchangeably in this chapter.

2.1.2 Definition of the exploration process

Before stating our main theorems, we define the exploration process of $G(\mathbf{W}, p)$ seen as a graph from the sequence $(G(\mathbf{W}, p))_{p\in[0,+\infty]}$ for a fixed p. All the results of this chapter are proven by a careful study of this process. It is based on an "horizontal" exploration of the graph, called the breadth-first walk (BFW). The BFW constructs the spanning forest of $G(\mathbf{W}, p)$, called the exploration forest. This is a forest consisting of spanning trees of all the connected components of $G(\mathbf{W}, p)$, constructed in a particular way.

For each potential edge $\{i, j\}$ recall the definition of $E_{\{i,j\}}$ from the model presentation. The BFW operates by steps, define the following sets of vertices. A vertex is always in exactly one of those sets.

— $(\mathcal{U}(i))_{n\geq i\geq 1}$ is the sequence of sets of undiscovered vertices at each step.
— $(\mathcal{D}(i))_{n\geq i\geq 1}$ is the sequence of sets of discovered but not yet explored vertices at each step.
— $(\mathcal{F}(i))_{n\geq i\geq 1}$ is the sequence of sets of explored vertices at each step.

First, choose a vertex i with probability :

$$\mathbb{P}(v(1) = i) = \frac{w_i}{\ell_n},$$

and call it $v(1)$. Let $\mathcal{V}'$ be the set of all vertices labels, and $\mathcal{U}(1) = \mathcal{V}' \setminus \{v(1)\}$, $\mathcal{D}(1) = \{v(1)\}$. At step 2, $v(1)$ is explored. It is thus not present in $\mathcal{D}(2)$ and moved to $\mathcal{F}(2)$. We call children of $v(1)$ the vertices j that are unexplored at step 1 and such that $E_{\{j,v(1)\}} \leq p$. Those children are moved to $\mathcal{D}(2)$ and become discovered but not yet explored. Let $c(1)$ be the number of children of $v(1)$. Call them $(v(2), v(3), ..., v(c(1) + 1)$ in increasing order of their $E_{\{j,v(1)\}}$'s. For $i \geq 1$, denote the set $\{v(1), v(2), ..., v(i)\}$ by $\mathcal{V}_i$. Hence, at step 2 we have :

— $\mathcal{U}(2) = \mathcal{V} \setminus \mathcal{V}_{c(1)+1}$.
— $\mathcal{D}(2) = \mathcal{V}_{c(1)+1} \setminus \mathcal{V}_1$.
— $\mathcal{F}(2) = \mathcal{V}_1$.

Now, at step 3, $v(2)$ becomes explored and its children $\{v(c(1)+2), v(c(1)+3)..., v(c(1)+c(2)+3)\}$ become discovered but not yet explored. The BFW continues like this, node $v(i)$ becomes

explored at step $i+1$, and its children are discovered at the same step. If the set of discovered nodes becomes empty at some step i, this means that the exploration of a connected component is finished. In that case, move on to the next step by choosing a vertex j with probability proportional to its weight w_j among the unexplored vertices and calling it $v(i)$ (like we did for $v(1)$) and exploring it. This construction ensures that a child has exactly one parent, since a child is always discovered while the process is exploring its parent. This ensures that we are constructing a forest. It is the exploration forest. We call the trees in that forest the exploration trees. By construction, exploration trees are spanning trees of the connected components of $G(\mathbf{W}, p)$. We say that a connected component is discovered at step i if its first node discovered by the BFW is $v(i)$. Similarly, we say that a connected component is explored at step i if its last node discovered by the BFW is $v(i-1)$. Generally, let $c(i)$ be the number of children of the node labeled $v(i)$. The exploration process associated to the BFW above is defined as follow for $n-1 \geq i \geq 0$:

$$\begin{aligned} L'_0 &= 1, \\ L'_{i+1} &= L'_i + c(i+1) - 1. \end{aligned}$$

The reflected exploration process is defined by

$$\begin{aligned} L_0 &= 1, \\ L_{i+1} &= \max(L_i + c(i+1) - 1, 1). \end{aligned}$$

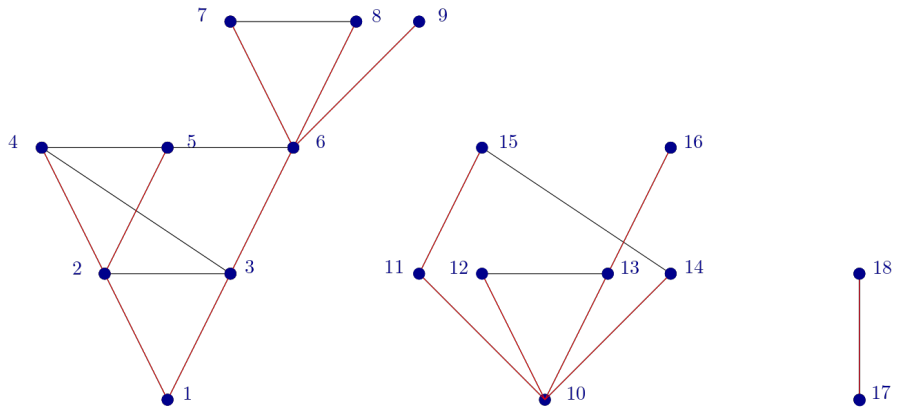


FIGURE 2.2 – Example of a graph with ordered nodes. The integers correspond to the order in the exploration process. The edges in red correspond to the edges of the exploration trees. The labels of the nodes are not represented.

The increment of the process L' at step i is the number of nodes added to the set of discovered nodes in the BFW after exploring node i. This number is at least -1 if the node being explored has no children. The process L' contains a lot of information about $G(\mathbf{W}, p)$. For example, each time a connected component is explored L' attains a new minimum. Using L' transforms geometrical questions about the graph, such as "Is there a connected component of size proportional to n ?" into questions regarding random walks such as "Is there an excursion of L' above its past minimum of size proportional to n ?".

Moreover, the order of appearance of the nodes in the exploration process corresponds to a size-biased sampling. Formally, we have for $i \in \{1, 2, ..., n-1\}$ and $j \in \{1, 2, ..., n\}$,

$$\begin{aligned} \mathbb{P}\left(v(1) = j\right) &= \frac{w_j}{\ell_n}. \\ \mathbb{P}\left(v(i+1) = j \quad | \mathcal{V}_i\right) &= \frac{w_j \mathbb{1}(j \notin \mathcal{V}_i)}{\ell_n - \sum_{k=1}^{i} w_{v(k)}}. \end{aligned}$$

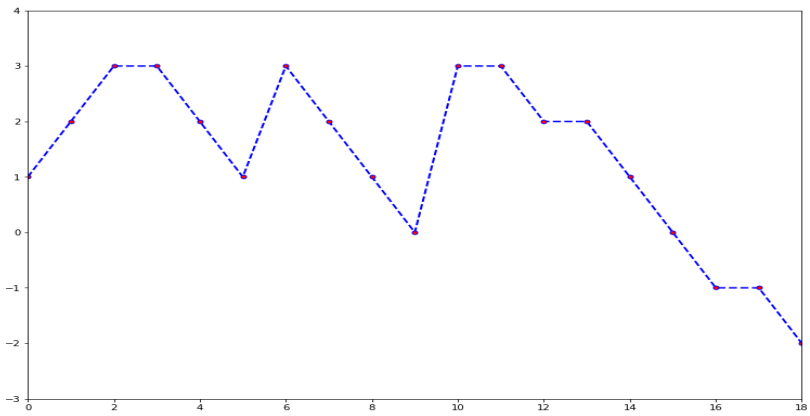


FIGURE 2.3 – The exploration process of the graph in Figure 2.2.

The proof of this fact uses only elementary results on exponential random variables. It is a widely known and used result (Aldous [1997], Bhamidi et al. [2010], Broutin et al. [2020] ...). We sketch the proof here.

Proof. By construction :

$$\mathbb{P}\left(v(1)=j\right)=\frac{w_j}{\ell_n}.$$

Then for $v(2)$, if $c(1)$, the number of children of $v(1)$, is 0 then, by definition, for any $j \geq 1$:

$$\mathbb{P}\left(v(2)=j|\mathcal{V}_1,\, c(i)=0\right)=\frac{w_j \mathbb{1}(j\notin\mathcal{V}_1)}{\ell_n - w_{v(1)}}.$$

Moreover if $c(1) \geq 1$, this means that there exists at least one $j \geq 1$ such that $j \neq v(1)$ and $E_{\{j,v(1)\}} \leq p$. By the absence of memory property of exponential random variables, for any $j \geq 1$:

$$\begin{aligned}
&\mathbb{P}(v(2)=j,\, c(1)\geq 1|\mathcal{V}_1)\\
&=\mathbb{P}(v(2)=j|\mathcal{V}_1)-\mathbb{P}(v(2)=j,\, c(1)=0|\mathcal{V}_1)\\
&=\mathbb{P}(\mathrm{argmin}_{k\neq v(1)}(E_{\{k,v(1)\}})=j|\mathcal{V}_1)-\mathbb{P}(\mathrm{argmin}_{k\neq v(1)}(E_{\{k,v(1)\}})=j|\mathcal{V}_1)\mathbb{P}(c(1)=0|\mathcal{V}_1)\\
&=\mathbb{P}(\mathrm{argmin}_{k\neq v(1)}(E_{\{k,v(1)\}})=j|\mathcal{V}_1)\mathbb{P}(c(1)\geq 1|\mathcal{V}_1).
\end{aligned}$$

By well known properties of exponential random variables, since conditionally on $\mathcal{V}_1$ the $(E_{\{k,v(1)\}})_{k\neq v(1)}$'s are independent, we have :

$$\mathbb{P}(\mathrm{argmin}_{k\neq v(1)}(E_{\{k,v(1)\}})=j|\mathcal{V}_1)=\frac{w_j \mathbb{1}(j\notin\mathcal{V}_1)}{\ell_n - w_{v(1)}}.$$

This shows the statement for $v(2)$, and we can move to subsequent nodes by induction. □

2.1.3 Conditions and main theorem

The weights in $\mathbf{W}$ depend implicitly on n. We will assume the following conditions on $\mathbf{W}$ in the entire chapter.

Conditions 1. *There exists some positive random variable W such that :*

(i) *The distribution of a uniformly chosen weight w_X converges weakly to W.*

(ii) $\mathbb{E}[W^3] < \infty$.

(iii) $\mathbb{E}[W^2] = \mathbb{E}[W]$.

(iv) $\ell_n = \mathbb{E}[W]n + o(n^{2/3})$.

(v) $\sum_{k=1}^{n} w_k^2 = \mathbb{E}[W^2]n + o(n^{2/3})$.

(vi) $\sum_{k=1}^{n} w_k^3 = \mathbb{E}[W^3]n + o(1)$.

(vii) $\max_{i\leq n} w_i = o(n^{1/3})$.

Conditions *i*,*ii* and *iii* ensure that the weak limit of $w_{v(1)}$ has a finite variance and mean 1. Condition *iii* can be ensured by changing the value of p.

Conditions *iv*,*v* and *vi* ensure that asymptotically the sum of the weights behave like the sum of independent identically distributed (i.i.d.) copies of W. Moreover, to further ease notations, as $n^{1/3} \geq w_1$, we will always use $n^{1/3}$ in our inequalities, even when w_1 would be sufficient. An important case to keep in mind is when $(w_1, w_2, ..., w_n)$ are realizations of random variables $(W_1, W_2, ..., W_n)$ which are i.i.d. with distribution W. In that case Conditions *iv*,*v* and *vi* are consequences of Conditions *ii* and *iii* (see Bhamidi et al. [2010] for a proof [1]). The node weights in Figure 2.1 verify Condition 1.

We define the size of a connected component $\mathcal{C}$, with vertices set $V(\mathcal{C})$, of $G(\mathbf{W}, p)$ as the number of vertices in $\mathcal{C}$. The distance between two vertices of $\mathcal{C}$ is the number of edges in the smallest (in number of edges) path between them. We also define the weight of $\mathcal{C}$ as :

$$\sum_{i\in\mathcal{V}(C)} w_i.$$

We call surplus (or excess) of $\mathcal{C}$ the number of edges that have to be removed from it in order to make it a tree. For instance, the surplus of a tree is 0, and the surplus of a cycle is 1.

Write $C = \frac{\mathbb{E}[W^3]}{\mathbb{E}[W]}$, and $p_{f_n} = \frac{\ell_n^{1/3}+f_n}{\ell_n^{4/3}}$. We can now state the main theorems of this chapter. Of course, these theorems hold only under Conditions 1.

Theorem 1 (**Size and weight of the giant component**). *Let $1 \geq \varepsilon' > 0$. Then for $f_n = o(n^{1/3})$ large enough. Consider the following event :*
The largest connected component of $G(n, \boldsymbol{W})$ has its size in the interval

$$\left[\frac{2(1-\varepsilon'/2)f_n\ell_n^{2/3}}{C} - \frac{\ell_n^{2/3}}{C}, \frac{2(1+\varepsilon'/2)f_n\ell_n^{2/3}}{C}\right],$$

and its weight in the interval

$$\left[\frac{2(1-\varepsilon')f_n\ell_n^{2/3}}{C}, \frac{2(1+\varepsilon')f_n\ell_n^{2/3}}{C}\right],$$

Then if Conditions 1 hold, there exists a positive constant $A > 0$ that only depend on the distribution of W, and such that the probability of this event not happening is at most :

$$A\exp\left(\frac{-f_n}{A}\right).$$

1. Bhamidi et al. [2010] shows that in that case the probability that the conditions hold tend to 1 when n tend to infinity. However, since we need concentration bounds, our weights need to verify these conditions deterministically.

Theorem 2 (**The excess of the giant component**). *Let Exc be the excess of the largest connected component of $G(n, \boldsymbol{W})$. Then if Conditions 1 hold, there exists a positive constant $A > 0$ that only depends on the distribution of W such that :*

$$\mathbb{P}(Exc \geq A f_n^3) \leq A \exp\left(\frac{-f_n}{A}\right).$$

Theorem 3 (**The sizes and weights of the small components**). *Let $1 > \varepsilon' > 0$ then for $f_n = o(n^{1/3})$ large enough, for any $1 \geq \varepsilon > 0$ Consider the following events :*

— *All the connected components discovered before the largest connected component in the exploration process of $G(n, \boldsymbol{W})$ have size smaller than*

$$\frac{\ell_n^{2/3}}{f_n^{1-\varepsilon}},$$

and weight smaller than

$$\frac{(1+\varepsilon')\ell_n^{2/3}}{f_n^{1-\varepsilon}}.$$

— *All the connected components discovered after the largest connected component in the exploration process of $G(n, \boldsymbol{W})$ have size smaller than*

$$\frac{\ell_n^{2/3}}{f_n},$$

and weight smaller than

$$\frac{(1+\varepsilon')\ell_n^{2/3}}{f_n}.$$

Then if Conditions 1 hold, there exists a positive constant $A > 0$ that only depends on the distribution of W such that the probability of one of those events not happening is at most :

$$A\left(\exp\left(\frac{-f_n^{\varepsilon}}{A}\right) + \exp\left(\frac{-\sqrt{f_n}}{A}\right) + \exp\left(\frac{-n^{1/8}}{A}\right)\right).$$

Theorem 4 (**The excess of the small components**). *Let Exc_0 be the the sum of the excesses of the connected components discovered before the largest connected component in the exploration process of $G(n, \boldsymbol{W})$. And let Exc_1 be the maximal excess of the connected component discovered after the largest connected component.*

Then if Conditions 1 hold, there exists a positive constant $A > 0$ that only depends on the distribution of W such that, for any $1 \geq \varepsilon > 0$:

$$\mathbb{P}\left(\mathrm{Exc}_0 \geq A f_n^{\varepsilon}\right) \leq A \exp\left(\frac{-f_n^{\varepsilon/2}}{A}\right),$$

and

$$\mathbb{P}\left(\mathrm{Exc}_1 \geq A f_n^{\varepsilon}\right) \leq A\left(\exp\left(\frac{-f_n^{\varepsilon}\ln(\sqrt{f_n})}{A}\right) + \exp\left(\frac{-\sqrt{f_n}}{A}\right) + \exp\left(\frac{-n^{1/8}}{A}\right)\right).$$

These theorems give precise bounds on the size, weight and excess of not only the largest connected component but also the other small connected components of the graph $G(n, \mathbf{W})$ in the barely supercritical regime, and in the critical regime when f_n is a large enough constant. As a direct corollary of those theorems, we also obtain convergence results when $f_n \to +\infty$ (see Corollary 38.2). Statements concerning the largest connected component and the connected components discovered before it are proven in Section 4. While statements concerning the

connected components discovered after the largest one are proven in Section 5. Moreover, at the cost of heavier notations, Theorem 38 provides a more precise statement than the one we presented in Theorem 3.

Notation : In the remainder of the chapter we drop the n from f_n. f will always be the critical parameter. Moreover we will always assume $f = o(n^{1/3})$ and $f \geq F$, where $F > 0$ is a constant independent of n which is large enough for all our theorems to hold. Similarly the variables $m = m_n$, $l = l_n$, $h = h_n$ and $y = y_n$ will always depend on n. The letters $A, A', A''...$ will be used for large positive constants that may only depend on the distribution of W.

2.1.4 Motivation and previous work

If $w_i = 1$ for all i, then the edge capacities $(E_{\{i,j\}})$ are i.i.d.. In that case $G(\mathbf{W}, p)$ is an Erdős-Rényi random graph. This is why the rank-1 inhomogeneous random graph model is a natural generalization of Erdős-Rényi random graphs. There are several variations of inhomogeneuous random graphs. The original inhomogeneous graph model was introduced by Aldous in his pioneer work on the multiplicative coalescent (Aldous [1997]), in this article he proved convergence of the component weights to a suitable limit. Then this model was further studied in Aldous and Limic [1998]. The model we study here is closely related to the so called Norros-Reittu model (Norros and Reittu [2006]). The difference between their model and ours being that their model allows for multi-edges. This, however, has no incidence on our proofs. And everything we show here still holds for their model. Other models of inhomogeneuous random graphs include the Britton-Deijfen-Martin-Löf (Section 3 in Addario-Berry et al. [2006]) model, where edge $\{i, j\}$ is present with probability :

$$\frac{w_i w_j}{n + w_i w_j}.$$

And the Chung-Lu model (Chapter 5, Section 3 in Chung and Lu [2006]) , where edge $\{i, j\}$ is present with probability :

$$\frac{w_i w_j}{\ell_n}.$$

This definition supposes that $\max_{i,j}(w_i w_j) \leq \ell_n$. we could have chosen some other representation of the edge probabilities. However, under our conditions and regime, all the results that we will prove are also true for those models. Generally, it is easy to see that all the theorems we prove here under Conditions 1 will still hold for any of the models above. The choice of $p_f = \frac{\ell_n^{1/3}+f}{\ell_n^{4/3}}$, with $f = o(n^{1/3})$ is motivated by the phase transition that appears in the following theorem (proved in Bollobás et al. [2007]).

Theorem 5. *Take $G(\boldsymbol{W}, \frac{c}{\ell_n})$ and suppose that Conditions 1 are verified, then the following results hold with high probability*[2] *:*

- ***Subcritical regime*** *If $c < 1$ then the largest connected component is of size $o(n)$.*
- ***Supercritical regime*** *If $c > 1$ then the largest connected component is of size $\Theta(n)$ and for any $i > 1$ the i-th largest connected component is of size $o(n)$.*
- ***Critical regime*** *If $c = 1$ then for any $i \geq 1$ the i-th largest connected component is of size $\Theta(n^{2/3})$.*

From this theorem it appears that there is a phase transition at $c = 1$. Just as in the Erdős-Rényi model, the right scale to look at the phase transition is for $c_n = 1 + \frac{\lambda}{\ell_n^{1/3}}$, with $\lambda > 0$ a constant. Which explains our choice of p_f. This is the so called critical window. In Theorems 1, 2, 3, and 4 we look at $c \sim 1$ and f that is either a large constant, or that goes to infinity but stays $o(n^{1/3})$. The latter is what we call the barely supercritical regime. The graph in Figure 2.1 corresponds to an inhomogeneous random graph approximately in the critical window.

2. We say that a sequence of events E_n holds with high probability if $\lim_{n\to\infty} \mathbb{P}(E_n) = 1$

Plenty of work was done on $G(\mathbf{W}, \lambda)$ with λ constant. The most recent and comprehensive one being in Broutin, Duquesne, and Wang [2018] and Broutin et al. [2020]. Aldous was the first to study the closely related multiplicative coalescent in Aldous [1997]. In Bhamidi et al. [2010] it is shown, under Conditions 1, that the sequence of sizes of the connected components, properly rescaled, converges to a random vector. In Bhamidi et al. [2017] this result is further extended, under stronger conditions than Conditions 1, by showing that the sequence of connected components of the whole graph, seen as metric spaces, when properly rescaled, converge to a limit sequence of compact metric spaces. Moreover, under Conditions 1, up to a multiplicative constant, this limit object has the distribution of the scaling limit of Erdős-Rényi random graphs (presented in Addario-Berry et al. [2012]). This shows that there is an invariance principle, although we have a generalization of Erdős-Rényi random graphs the limit objects are just rescaled versions of one another.

However, unlike the Erdős-Rényi case (see Addario-Berry, Broutin, and Reed [2009]), there is no uniform study when f moves through the critical window. For instance, there are no known concentration results that depend on f for the size of the largest component of rank-1 inhomogeneous random graphs. Moreover, there are no known concentration results for the barely supercritical regime. These are the cases that we treat in this chapter.

This study has other implications for another object. For $n \in \mathbb{N}$, assign i.i.d., uniform random variables on $(0, 1)$, that we call weights, to the edges of a complete graph of size n. Then the random minimum spanning tree (random MST) is the (almost surely unique) connected subgraph with n vertices that minimizes the sum of the weights. It is a tree. In Article Addario-Berry et al. [2017b] it is proven that when rescaling the distances by $n^{-1/3}$, the random MST converges to a compact tree-like metric space called. The proof in Addario-Berry et al. [2017b] relies heavily on a uniform study of the critical Erdős-Rényi graph through the critical window and in the barely supercritical regime (done before in Article Addario-Berry et al. [2009]).

In order to do the same for the rank-1 inhomogeneous random graphs, instead of putting i.i.d. weights on a complete graph, put capacity $E_{\{i,j\}}$ on edge $\{i, j\}$ and construct the minimum spanning tree for those capacities. Call such a tree the inhomogeneous random MST. Clearly, this tree can be coupled with rank-1 inhomogeneous random graphs in the same fashion as in Addario-Berry et al. [2017b]. One can ask whether that tree, when properly rescaled, also converges to a continuous random tree-like metric space. And if the answer is yes, will this metric space be a rescaled version of the scaling limit of the random MST in Addario-Berry et al. [2017b] ? A positive answer would show that there is still an invariance principle for those trees.

We intend on answering these questions in follow up chapters, and the bounds we prove in this chapter will be crucial in our future proofs.

The biggest difficulty in proving our theorems is that the weight discovered at step i of the exploration process depend on the weights discovered before it. Those weights appear in a size-biased fashion. This is why we show new concentration inequalities for size-biased sampling without replacement. We also make use of the note Ben-Hamou et al. [2018] in order to estimate the deviations of the sum of variables sampled without replacement. Another difficulty is that we cannot rely on known results (for example results in Article Łuczak [1990]) that were proved for Erdős-Rényi graphs. Everything has to be done separately for inhomogeneuous random graphs.

There are other interesting problems that require more work. For instance there is the case of power law distributions for the node weights. Conditions 1 ensure that a uniform node weight behaves like a random variable with finite third moment. One can change those conditions, and allow the variable to follow a power law distribution of parameter $\tau > 3$. If $\tau > 4$, then we are in the case of finite third moments treated here. However, when $\tau \leq 4$, we expect the results to be vastly different. Informal arguments show that in that case the scaling limit of the minimum spanning tree should be mutually singular with the scaling limit of random MST. This intuition is due to the appearance of Levy trees when studying those graphs (see van der Hofstad, Kliem,

and van Leeuwaarden [2018] for further discussion of this model).

Finally another totally different set of questions regard biased sampling without replacement. Let $n \geq 1$ be an integer and $(a_1, a_2, ...a_n)$ be decreasing real number. Moreover let $(p_1, p_2, ..., p_n)$ be positive real numbers such that :

$$\sum_{k=1}^{n} p_i = 1.$$

Let $(V(1), V(2), ..., V(n))$ be a vector random variables that correspond to indices sampled without replacement in the following way, for any $i \in \{1, 2, ..., n-1\}$ and $j \in \{1, 2, ..., n-1\}$:

$$\mathbb{P}(V(1) = j) = p_j,$$
$$\mathbb{P}(V(i+1) = j \quad |(V(1), V(2)..., V(i))) = \frac{p_j \mathbb{1}(V(j) \notin (V(1), ..., V(i)))}{\sum_{k=1}^{n} p_k - \sum_{k=1}^{i} p_{V(k)}}.$$

Consider also $(J(1), J(2), ..., J(n))$ that is a vector of independent random variables with the same distribution as $V(1)$. The $J(i)$'s correspond to indices sampled with replacement. Remark that size-biased sampling is a special case of biased sampling. While working on this chapter two questions arose regarding these two samplings. First, under which set of conditions do we have the following inequality for any $n \geq m \geq l$ and real number $x \geq 0$:

$$\mathbb{P}\left(\left|\sum_{k=l}^{m} a_{V(i)} - \mathbb{E}[a_{V(i)}]\right| \geq x\right) \leq \mathbb{P}\left(\left|\sum_{k=l}^{m} a_{J(i)} - \mathbb{E}[a_{J(i)}]\right| \geq x\right).$$

This inequality means that biased sampling without replacement is more concentrated around its mean than biased sampling with replacement. The main idea behind this conjecture is that sampling without replacement tends to auto-concentrate itself around its mean. For instance, if for some $i \geq 1$, $V(i) = j$ and a_j is very large, then we will not draw the same index j in subsequent rounds. But in biased sampling with replacement, the same "bad" event can keep happening.

We were not able to find any trivial counter example to this inequality, so it could be true that it holds without any further assumptions. If not, then under which set of assumptions does it hold ? With such an inequality it would be easy to answer the question regarding inhomogeneous random graphs with power law distribution presented in the paragraph above.

Another question is for the ordered case. Suppose now that $p_1 \geq p_2 \geq ... \geq p_n$. This means that larger a_i's have larger probabilities of being drawn first. This is again a general case of size-biased sampling. Is it true then that for any $n - 1 \geq m \geq 1$, and real numbers $(x_1, x_2, x_3, ...x_n)$

$$\mathbb{P}\left(a_{V(1)} \geq x_1,\, a_{V(2)} \geq x_2, ...,\, a_{V(m)} \geq x_m\right) \geq \mathbb{P}\left(a_{V(2)} \geq x_1,\, a_{V(3)} \geq x_2, ...,\, a_{V(m+1)} \geq x_m\right),$$

and also

$$\mathbb{P}\left(a_{J(1)} \geq x_1,\, a_{J(2)} \geq x_2, ...,\, a_{J(m)} \geq x_m\right) \geq \mathbb{P}\left(a_{V(1)} \geq x_1,\, a_{V(2)} \geq x_2, ...,\, a_{V(m)} \geq x_m\right).$$

In Lemma 35, we prove those inequalities for $m = 1$. With some more work, we can prove them for $m = 2$ also. We conjecture that they are in fact true for all $1 \leq m \leq n - 1$.

2.2 Bounding the weights

A well known fact is that the sum of weights sampled uniformly without replacement verifies slightly better Chernoff concentration inequalities as the sum of weights sampled uniformly with replacement (See Serfling [1974]). No such general result is available for size-biased sampling.

In this section we will always assume that Conditions 1 are verified. We will prove concentration bounds for the weights sampled in size-biased order and without replacement under some conditions.

2.2.1 First concentration result and the mean

The following theorem, from Article Ben-Hamou et al. [2018], is a first important step in comparing the sum of the $(w_{v(i)})_i$'s with the sum of i.i.d. copies of a random variable.

Theorem 6. *Let $0 < l \leq m \leq n$ be two integers, and $J(1), J(2)..., J(n)$ be i.i.d. random variables with the distribution of $v(l)$, then for any convex function g :*

$$\mathbb{E}\left[g\left(\sum_{i=l}^{m} w_{v(i)}\right)\right] \leq \mathbb{E}\left[g\left(\sum_{i=l}^{m} w_{J(i)}\right)\right].$$

Generally, concentration bounds that use Chernoff's inequality are based on the fact that :

$$\mathbb{E}\left[\exp\left(\sum_{i=l}^{m} w_{J(i)}\right)\right] = \mathbb{E}\left[\exp\left(w_{J(1)}\right)\right]^m.$$

Hence, taking g to be the exponential function in Theorem 6 shows a Chernoff type inequality. This means that upper bounds that use Chernoff's inequality (first used in Bernstein [1924]) and which hold for size-biased sampling with replacement are still true for size-biased sampling without replacement. This fact will be used later in the proofs. This is true in particular for Bernstein's inequality (Bernstein [1924]) which stems from Chernoff's bound.

The following lemmas give an estimation of the mean of $w_{v(i)}$. This first Lemma is already shown in one of the proofs that appear in Bhamidi et al. [2010], we prove it here again for clarity.

Lemma 7. *Suppose that Conditions 1 hold. Then for any $0 < l = o(n)$, and $i \in \{1, 2, 3\}$ we have :*

$$\sum_{k=1}^{l} w_k^i = o(n).$$

Proof. We do the proof for $i = 3$, the other cases can be proved similarly or deduced easily from this case. Recall that the weights $(w_1, w_2, ...w_n)$ are taken in decreasing order. For any $K > 0$:

$$\begin{aligned}\sum_{k=1}^{l} \frac{w_k^3}{\ell_n} &\leq \sum_{k=1}^{l} \frac{w_k^3 \mathbb{1}(w_k \leq K)}{\ell_n} + \sum_{k=1}^{n} \frac{w_k^3 \mathbb{1}(w_k > K)}{\ell_n} \\ &\leq \frac{lK^3}{\ell_n} + \sum_{k=1}^{n} \frac{w_k^3 \mathbb{1}(w_k > K)}{\ell_n}.\end{aligned} \tag{2.1}$$

By the weak convergence in Conditions 1 :

$$\lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{w_k^3 \mathbb{1}(w_k \leq K)}{n}\right) = \mathbb{E}[W^3 \mathbb{1}(W \leq K)],$$

and by the fact that :

$$\sum_{k=1}^{n} w_k^3 = \mathbb{E}[W^3]n + o(n),$$

it follows that :

$$\begin{aligned}\lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{w_k^3 \mathbb{1}(w_k > K)}{\ell_n}\right) &= \frac{1}{\mathbb{E}[W]} \left(\mathbb{E}[W^3] - \mathbb{E}[W^3 \mathbb{1}(W \leq K)]\right) \\ &= \frac{\mathbb{E}[W^3 \mathbb{1}(W > K)]}{\mathbb{E}[W]}.\end{aligned}$$

Since $\mathbb{E}[W^3] < \infty$:

$$\lim_{K\to\infty}\left(\lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{w_k^3\mathbb{1}(w_k>K)}{\ell_n}\right)\right)=0.$$

Together with the fact that and $l = o(n)$, letting n go to infinity then K go to infinity in Equation (2.1) yields :

$$\sum_{k=1}^{l}\frac{w_k^3}{\ell_n}=o(1). \tag{2.2}$$

□

Lemma 8. *Suppose that Conditions 1 hold. Recall that* $C=\frac{\mathbb{E}[W^3]}{\mathbb{E}[W]}$. *For any* $l = o(n)$ *:*

$$\mathbb{E}[w_{v(l)}^2]=C+o(1).$$

Proof. We have using Lemma 7 :

$$\sum_{k\in\mathcal{V}_{l-1}}\frac{w_k}{\ell_n}\le\sum_{k=1}^{l-1}\frac{w_k}{\ell_n}=o(1).$$

Hence :

$$\begin{aligned}\mathbb{E}[w_{v(l)}^2]&=\mathbb{E}\left[\sum_{k\notin\mathcal{V}_{l-1}}\frac{w_k^3}{\ell_n-\sum_{k'\in\mathcal{V}_{l-1}}w_k'}\right]\\&=\mathbb{E}\left[\sum_{k\notin\mathcal{V}_{l-1}}\frac{w_k^3}{\ell_n}\right](1+o(1))\\&=C(1+o(1))-\mathbb{E}\left[\sum_{k\in\mathcal{V}_{i-1}}\frac{w_k^3}{\ell_n}\right](1+o(1))+o(1).\end{aligned} \tag{2.3}$$

In order to finish the proof we use Lemma 7 again :

$$\mathbb{E}\left[\sum_{k\in\mathcal{V}_{l-1}}\frac{w_k^3}{\ell_n}\right]\le\sum_{k=1}^{l-1}\frac{w_k^3}{\ell_n}=o(1). \tag{2.4}$$

From Equations (2.3) and(2.4) we obtain :

$$\mathbb{E}(w_{v(l)}^2)=C+o(1), \tag{2.5}$$

which finishes the proof. □

Lemma 9. *Suppose that Conditions 1 hold. Let* $l = o(n)$*, we have :*

$$\mathbb{E}[w_{v(l)}]=1+o(1).$$

Proof. As in the proof of Lemma 8 we have :

$$\begin{aligned}\mathbb{E}(w_{v(l)})&=\frac{\mathbb{E}[W^2]}{\mathbb{E}[W]}(1+o(1))-\mathbb{E}\left[\sum_{k\in\mathcal{V}_{l-1}}\frac{w_k^2}{\ell_n}\right](1+o(1))\\&=\frac{\mathbb{E}[W^2]}{\mathbb{E}[W]}(1+o(1)).\end{aligned}$$

Recalling that $\frac{\mathbb{E}[W^2]}{\mathbb{E}[W]}=1$ ends the proof. □

By the same argument we also have :

Lemma 10. *Suppose that Conditions 1 hold. Let $l = o(n)$. For any $0 < i < l$ we have :*

$$\mathbb{E}(w_{v(i)}w_{v(l)}) = 1 + o(1).$$

Proof. We have using Lemma 7 :

$$\begin{aligned}\mathbb{E}(w_{v(i)}w_{v(l)}) &= \mathbb{E}\left[w_{v(i)}\sum_{k\notin\mathcal{V}_{l-1}}\frac{w_k^2}{\ell_n - \sum_{k'\in\mathcal{V}_{l-1}}w'_k}\right] \\ &= \mathbb{E}\left[w_{v(i)}\sum_{k\notin\mathcal{V}_{l-1}}\frac{w_k^2}{\ell_n}\right](1+o(1)) \\ &= 1 + o(1),\end{aligned}$$

which ends the proof. □

Thanks to these lemmas, we obtain a more precise estimation of the mean of $w_{v(l)}$.

Lemma 11. *Suppose that Conditions 1 hold. For any $l = o(n)$, we have :*

$$\mathbb{E}[w_{v(l)}] = 1 + \frac{l}{\ell_n}(1-C) + o\left(\frac{l+n^{2/3}}{n}\right).$$

Proof. By definition :

$$\mathbb{E}[w_{v(l)}] = \mathbb{E}\left[\sum_{i\notin\mathcal{V}_{l-1}}\frac{w_i^2}{\ell_n - \sum\limits_{i'\in\mathcal{V}_{l-1}}w_{i'}}\right].$$

Moreover, by Lemma 7 :

$$\begin{aligned}\mathbb{E}\left[w_{v(l)}\right] &= \mathbb{E}\left[\sum_{i\notin\mathcal{V}_{l-1}}\frac{w_i^2}{\ell_n\left(1-\frac{\sum_{i'\in\mathcal{V}_{l-1}}w_{i'}}{\ell_n}\right)}\right] \\ &= \mathbb{E}\left[\sum_{i\notin\mathcal{V}_{l-1}}\frac{w_i^2}{\ell_n}\left(1+\frac{\sum_{i'\in\mathcal{V}_{l-1}}w_{i'}}{\ell_n}\right)\right] + o\left(\frac{l}{n}\right).\end{aligned}$$

By Lemmas 7, 8 and 9 it follows that :

$$\begin{aligned}
\mathbb{E}\left[w_{v(l)}\right] &= \mathbb{E}\left[\sum_{i\notin\mathcal{V}_{l-1}} \frac{w_i^2}{\ell_n}\left(1+\frac{\sum_{i'\in\mathcal{V}_{l-1}} w_{i'}}{\ell_n}\right)\right] + o\left(\frac{l}{n}\right) \\
&= \frac{\sum_{i=1}^n w_i^2}{\ell_n} + \mathbb{E}\left[\frac{(\sum_{i'\in\mathcal{V}_{l-1}} w_{i'})(\sum_{i=1}^n w_i^2)}{\ell_n^2}\right] - \mathbb{E}\left[\frac{\sum_{i\in\mathcal{V}_{l-1}} w_i^2}{\ell_n}\right] \\
&- \mathbb{E}\left[\frac{\left(\sum_{i\in\mathcal{V}_{l-1}} w_i^2\right)\left(\sum_{i'\in\mathcal{V}_{l-1}} w_{i'}\right)}{\ell_n^2}\right] + o\left(\frac{l}{n}\right) \\
&= \frac{\sum_{i=1}^n w_i^2}{\ell_n} + \mathbb{E}\left[\frac{(\sum_{i'\in\mathcal{V}_{l-1}} w_{i'})(\sum_{i=1}^n w_i^2)}{\ell_n^2}\right] - \mathbb{E}\left[\frac{\sum_{i\in\mathcal{V}_{l-1}} w_i^2}{\ell_n}\right] \\
&- o\left(\mathbb{E}\left[\frac{\sum_{i\in\mathcal{V}_{l-1}} w_i^2}{\ell_n}\right]\right) + o\left(\frac{l}{n}\right) \\
&= 1 + \mathbb{E}\left[\frac{(\sum_{i'\in\mathcal{V}_{l-1}} w_{i'})(\sum_{i=1}^n w_i^2)}{\ell_n^2}\right] - \mathbb{E}\left[\frac{\sum_{i\in\mathcal{V}_{l-1}} w_i^2}{\ell_n}\right] + o\left(\frac{l+n^{2/3}}{n}\right). \\
&= 1 + \frac{l}{\ell_n}(1-C) + o\left(\frac{l+n^{2/3}}{n}\right).
\end{aligned}$$

□

Observe that with the assumption that $\mathbb{E}[W^2] = \mathbb{E}[W]$, the Cauchy-Schwarz inequality implies that :

$$1 - C = \left(1 - \frac{\mathbb{E}[W^3]}{\mathbb{E}[W]}\right) \leq 0,$$

so asymptotically $\mathbb{E}(w_{v(i)})$ decreases with i. Lemma 35 shows that in fact, it decreases all the time.

2.2.2 A more precise concentration inequality

In order to obtain concentration inequalities for size-biased sampling without replacement, we will use a randomization trick. The main idea here is that taking weights without replacement is the same as putting exponential "clocks" on each weight and taking a weight when its clock rings.

More precisely let $(T_i)_{i\leq n}$ be a sequence of independent exponential random variables with respective rates $(w_i/\ell_n)_{i\leq n}$. Define the following quantities for $x \geq 0$:

$$N(x) = \sum_{k=1}^n \mathbb{1}(T_k \leq x),$$

$$X(x) = \sum_{k=1}^n w_k \mathbb{1}(T_k \leq x).$$

By basic properties of exponential random variables, $(v'(1), v'(2), ..., v'(n))$, the distinct random indices of the T_i's taken in increasing order, i.e :

$$T_{v'(1)} \leq T_{v'(2)} \leq ... \leq T_{v'(n)},$$

are distributed as a size-biased sample taken without replacement.

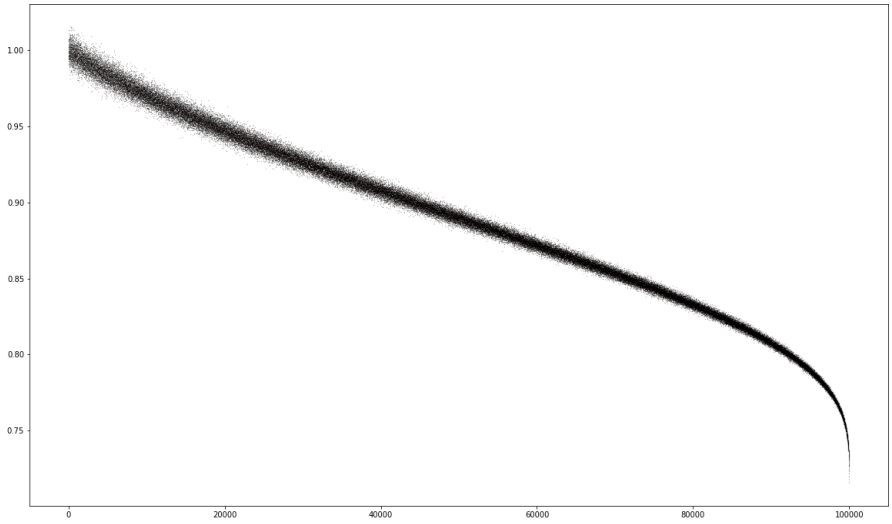


FIGURE 2.4 – A simulation of the values of the $\mathbb{E}[w_{v(i)}]$'s for $n \geq i \geq 1$. This simulation is done on $n = 100000$ weights verifying Conditions 1 by doing $m = 10000$ rounds of biased sampling without replacement and averaging the result.

Moreover the following equality holds :

$$X(x) = \sum_{k=1}^{n} w_{v'(k)} \mathbb{1}(N(x) \geq k).$$

Since $N(x)$ and $X(x)$ are sums of independent random variables, we can apply Bernstein's inequality (Bernstein [1924]) in order to obtain the following lemma. We let $w_{v(0)} = 0$.

Lemma 12. *Suppose that Conditions 1 hold. For any $x \geq 0$ and $t \geq 0$, the following holds :*

$$\mathbb{P}(|X(x) - \mathbb{E}[X(x)]| \geq t) \leq 2 \exp\left(\frac{-t^2}{2(tn^{1/3} + x)}\right),$$

and

$$\mathbb{P}(|N(x) - \mathbb{E}[N(x)]| \geq t) \leq 2 \exp\left(\frac{-t^2}{2(t + x)}\right).$$

The following conditions will always be verified in this section. They give a regime where our concentration bounds hold.

Conditions 2. *We say that $(a(n), b(n))$ verifies Conditions 2 if for all n large enough :*

$$\exp\left(\frac{-b(n)^2}{\bar{A}(b(n)n^{1/3} + a(n))}\right) < 1/4,$$

$$\lim_{n\to\infty} a(n) = \lim_{n\to\infty} b(n) = +\infty$$

$$a(n) = o(n),$$

$$b(n) = O(a(n)),$$

$$a(n) = O\left(b(n)\ell_n^{1/3}\right),$$

and :

$$(a(n))^2 = O\left(b(n)\ell_n\right),$$

where $\bar{A} > 0$ is independent of n and larger than all the other constants $A, A', A''...$ that appear in this chapter.

The condition $b(n) = O(a(n))$ is not necessary, but it makes some computations easier and will be true in all the practical cases in this chapter. Moreover, notice that if $(a(n), b(n))$ verifies Conditions 2 then for any $A > 0$ the couple $(a(n), Ab(n))$ will also verify those conditions. We want to prove that there exists an $A > 0$ such that if (m, y) verify Conditions 2 then :

$$\mathbb{P}\left[\sup_{i\leq m}\left|\sum_{k=1}^{i} w_{v(k)} - \mathbb{E}\left[\sum_{k=1}^{i} w_{v(k)}\right]\right| \geq y\right] \leq A\exp\left(\frac{-y^2}{A(yn^{1/3}+m)}\right).$$

In order to do so, we will use the fact that if $N(u_n) \geq m$ for some $u_n > 0$ then :

$$\sup_{i\leq m}\left|\sum_{k=1}^{i} w_{v'(i)} - \mathbb{E}\left[\sum_{k=1}^{i} w_{v'(i)}\right]\right| \leq \sup_{x\leq u_n}\left|X(x) - \sum_{k=1}^{N(x)}\mathbb{E}\left[w_{v'(i)}\right]\right|.$$

Then we will show concentration of the right-hand side of the above inequality. The following fact will be used through this whole section. For any $x \geq 0$:

$$x \geq 1 - e^{-x} \geq x - \frac{x^2}{2}. \tag{2.6}$$

We start by showing the following lemma :

Lemma 13. *Suppose that Conditions 1 hold. Let $(a(n), b(n))$ verify Conditions 2. Then there exists a constant $A > 0$ such that for all n large enough :*

$$\mathbb{P}\left[\sup_{x\leq a(n)}\mathbb{E}\left[X(x)\right] - \sum_{k=1}^{N(x)}\mathbb{E}\left[w_{v(i)}\right] \geq b(n)\right] \leq \mathbb{P}\left[\inf_{x\leq a(n)} N(x) - \mathbb{E}[N(x)] \leq \frac{-b(n)}{A} + 1\right],$$

and :

$$\mathbb{P}\left[\inf_{x\leq a(n)}\mathbb{E}\left[X(x)\right] - \sum_{k=1}^{N(x)}\mathbb{E}\left[w_{v(i)}\right] \leq -b(n)\right] \leq \mathbb{P}\left[\sup_{x\leq a(n)} N(x) - \mathbb{E}[N(x)] \geq \frac{b(n)}{A} - 1\right],$$

and the same inequalities hold without the sup *and* inf*.*

Proof. Let $x \leq a(n)$. By Equation (2.6) and Conditions 1 :

$$\begin{aligned}
\mathbb{E}[X(x)] &= \sum_{k=1}^{n} w_k \mathbb{P}(T_k \leq x) \\
&= \sum_{k=1}^{n} w_k\left(1 - \exp\left(\frac{-w_k x}{\ell_n}\right)\right) \\
&\leq \sum_{k=1}^{n}\frac{w_k^2 x}{\ell_n} \\
&= x(1 + o(n^{-1/3})).
\end{aligned} \tag{2.7}$$

For any $b'(n)$ such that $(a(n), b'(n))$ verify Conditions 2, there exists $A' > 0$ such that :

$$x^2 \leq a(n)^2 \leq A' b'(n) \ell_n.$$

Denote $\lceil \mathbb{E}[N(x)] - b'(n) \rceil$ by u. By Conditions 1 and Equation (2.6) we obtain :

$$\begin{aligned} u &\geq x - b'(n) - \sum_{k=1}^{n} \frac{w_k^2 x^2}{2\ell_n^2} \\ &\geq x - b'(n) - \frac{x^2}{2\ell_n} + o\left(\frac{x^2}{\ell_n}\right) \\ &\geq x - b'(n) - \frac{A' b'(n)}{2} + o\left(\frac{x^2}{\ell_n}\right). \end{aligned} \tag{2.8}$$

Moreover by Condition 2 :

$$\begin{aligned} u^2 &\leq (x + b'(n))^2 \\ &\leq 2x^2 + 2b'(n)^2 \\ &\leq 2A' \ell_n b'(n) + 2b'(n)^2 \\ &\leq A'' \ell_n b'(n), \end{aligned} \tag{2.9}$$

where $A'' > 0$ is some large constant. By Equations (2.8), (2.9), Conditions 2 and Lemma 11 we have :

$$\begin{aligned} \sum_{k=1}^{u} \mathbb{E}\left[w_{v(i)}\right] &= \sum_{k=1}^{u} \left(1 + \frac{k}{\ell_n}(1 - C)\right) + o\left(\frac{u^2 + un^{1/3}}{n}\right) \\ &= u + \frac{u^2}{2\ell_n}(1 - C) + o\left(\frac{u^2 + un^{1/3}}{n}\right) \\ &\geq x - A''' b'(n), \end{aligned} \tag{2.10}$$

where $A''' > 0$ is a large constant. Inequalities (2.7) and (2.10) and Conditions 2 yield :

$$\mathbb{E}[X(x)] - \sum_{k=1}^{u} \mathbb{E}\left[w_{v(i)}\right] \leq A''' b'(n) + o(xn^{-1/3}).$$

And of course, since $\mathbb{E}\left[w_{v(i)}\right]$ is positive for all $i \leq n$, the same inequality holds if we replace u by $u' \geq u$. This show that :

$$\left(\mathbb{E}\left[X(x)\right] - \sum_{k=1}^{N(x)} \mathbb{E}\left[w_{v(i)}\right] \geq 2A''' b'(n)\right) \Rightarrow (N(x) \leq \mathbb{E}[N(x)] - b'(n) + 1)$$

Taking $b(n) = 2A''' b'(n)$ proves the first inequality of the lemma, the second inequality is proved similarly. □

Similarly we have the following lemma for which we omit the proof

Lemma 14. *Suppose that Conditions 1 hold. Let $(a(n), b(n))$ verify Conditions 2. Then there exists a constant $A > 0$ such that for all n large enough :*

$$\begin{aligned} &\mathbb{P}\left[\sup_{x \leq a(n)} \mathbb{E}\left[X(a(n)) - X(x)\right] - \sum_{k=N(x)}^{N(a(n))} \mathbb{E}\left[w_{v(i)}\right] \geq b(n)\right] \\ \leq &\mathbb{P}\left[\inf_{x \leq a(n)} N(a(n)) - N(x) - \mathbb{E}[N(a(n)) - N(x)] \leq \frac{-b(n)}{A} + 1\right], \end{aligned}$$

and :

$$\mathbb{P}\left[\inf_{x\leq a(n)} \mathbb{E}\left[X(a(n)) - X(x)\right] - \sum_{k=N(x)}^{N(a(n))} \mathbb{E}\left[w_{v(i)}\right] \leq -b(n)\right]$$
$$\leq \mathbb{P}\left[\sup_{x\leq a(n)} N(a(n)) - N(x) - \mathbb{E}[N(a(n)) - N(x)] \geq \frac{b(n)}{A} - 1\right],$$

and the same inequalities hold without the sup *and* inf*.*

These lemmas will allow us to prove the following concentration inequality. Recall that $m = m(n)$ and $y = y(n)$ depend implicitly on n.

Lemma 15. *Suppose that Conditions 1 hold. There exist a constant $A > 0$ such that if (x_n, y) verifies Conditions 2, then :*

$$\mathbb{P}\left(\left|X(x_n) - \sum_{k=1}^{N(x_n)} \mathbb{E}[w_{v(i)}]\right| \geq y\right) \leq A\exp\left(\frac{-y^2}{A(yn^{1/3} + x_n)}\right),$$

Proof. By the union bound :

$$\begin{aligned}&\mathbb{P}\left[\left|X(x_n) - \sum_{k=1}^{N(x_n)} \mathbb{E}\left[w'_{v(i)}\right]\right| \geq y\right] \\ &\leq \left(\mathbb{P}\left[\left|X(x_n) - \mathbb{E}\left[X(x_n)\right]\right| \geq \frac{y}{2}\right] + \mathbb{P}\left[\left|\mathbb{E}\left[X(x_n)\right] - \sum_{k=1}^{N(x_n)} \mathbb{E}\left[w'_{v(i)}\right]\right| \geq \frac{y}{2}\right]\right).\end{aligned} \tag{2.11}$$

We bound separately each term of the right-hand side of Equation (2.11). Lemma 12 states that :

$$\mathbb{P}\left[\left|X(x_n) - \mathbb{E}\left[X(x_n)\right]\right| \geq \frac{y}{2}\right] \leq 2\exp\left(\frac{-y^2}{8(yn^{1/3} + x_n)}\right). \tag{2.12}$$

Using Equations (2.12), Lemma 13 on $(x_n, y/2)$ and Lemma 12 to bound the second expression in the right-hand side of Equation (2.11) expression in Equation (2.11) shows that :

$$\mathbb{P}\left(\left|X(x_n) - \sum_{k=1}^{N(x_n)} \mathbb{E}[w_{v(i)}]\right| \geq y\right) \leq A'\exp\left(\frac{-y^2}{A'(yn^{1/3} + x_n)}\right),$$

where $A' > 0$ is a large constant. □

In order to prove concentration inequalities on the $N(t)$ and $X(t)$ for all t in some interval, we use the chaining method. This method consists of crafty discretizations of the "time" parameter space in order to derive general bounds for all "times". The method is explained in Chapter 13 of Boucheron et al. [2013]. It is attributed to Kolmogorov, and it has been vastly used and improved by Dudley (Dudley [1973]) and Talagrand (for instance Talagrand [2005]).

Lemma 16. *Suppose that Conditions 1 hold. There exist a constant $A > 0$ such that, for any (m, y) that verify Conditions 2 :*

$$\mathbb{P}\left(\sup_{0\leq t\leq m} \left(X(t) - \mathbb{E}\left[X(t)\right]\right) \geq y\right) \leq A\exp\left(\frac{-y^2}{A(yn^{1/3} + m)}\right).$$

Proof. Recall that $w_1 \geq w_2 \geq w_3 \geq ...$ and for $i \leq n$ write :

$$X_i(t) = \sum_{k=i+1}^{n} w_k \mathbb{1}(T_k \leq t).$$

By Bernstein's inequality and basic computations, for any $u > 0$ and $s < t$:

$$\mathbb{P}\left(|X_i(t) - X_i(s) - \mathbb{E}[X_i(t) - X_i(s)]| \geq \sqrt{2(t-s)\sum_{k=i+1}^{n} \frac{w_k^3}{\ell_n} u} + u w_{i+1} \right) \leq 2\exp(-u). \quad (2.13)$$

For $i \geq 0$ let :

$$\Gamma_i = \left\{ m\frac{k}{2^i}, 0 \leq k < 2^i \right\} \cup \left\{ T_k, 1 \leq k < 2^i \right\}.$$

Let $f_i : t \in [0, m] \mapsto \max\{z \in \Gamma_i, t > z\}$. We have, by definition of f_i and Γ_i, for any $t \leq m$:

$$\begin{aligned} X(t) - X(f_i(t)) &= \sum_{k=1}^{n} w_k \mathbb{1}(f_i(t) < T_k \leq t) \\ &= \sum_{k=2^i}^{n} w_k \mathbb{1}(f_i(t) < T_k \leq t) \\ &= X_{2^i-1}(t) - X_{2^i-1}(f_i(t)). \end{aligned}$$

Since $f_i(t)$ is measurable with respect to the $(T_k)_{k<2^i}$'s. And conditionally on $f_i(t)$, $X(t) - X(f_i(t))$ is a sum of independent random variables. We can apply Bernstein's inequality to obtain similarily to Equation (2.13) :

$$\mathbb{P}\left(|X(t) - X(f_i(t)) - \mathbb{E}[X(t)] - X(f_i(t))]| \geq \sqrt{2(t - f_i(t))\sum_{k=2^i}^{n} \frac{w_k^3}{\ell_n} u} + u w_{2^i} \right) \quad (2.14)$$
$$\leq 2\exp(-u).$$

Let :

$$\rho_i = \sqrt{3\frac{m}{2^i} C(u(i+1))} + u(i+1) w_{2^i}.$$

Since $(t - f_i(t)) \leq \frac{m}{2^i}$ and $\sum_{k=1}^{n} \frac{w_k^3}{\ell_n} = C(1 + o(1))$. Inequality (2.14) with $u' = u(i+1)$ yields :

$$\mathbb{P}\left(|X_i(f_i(t)) - X_i(t) - \mathbb{E}[X_i(f_i(t)) - X_i(s)]| \geq \rho_i \right) \leq 2\exp(-u(i+1)).$$

The classical chaining argument is that for any $0 \leq t \leq m$ can be written as :

$$t = \sum_{i=0}^{\infty} (f_{i+1}(t) - f_i(t)),$$

This gives us by union bound, if we suppose that $u > \ell_n(4)$:

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\le t\le m}|X(t)-\mathbb{E}[X(t)]|\ge\sum_{i=0}^{\infty}\rho_i\right)\\
&\le\sum_{i=0}^{\infty}\sum_{t\in\Gamma_{i+1}}\mathbb{P}\left(|X_i(t)-X_i(f_i(t))-\mathbb{E}[X_i(t)-X_i(f_i(t))]|\ge\rho_i\right)\\
&\le\sum_{i=0}^{\infty}\sum_{t\in\Gamma_{i+1}}2\exp(-u(i+1))\\
&\le\sum_{i=0}^{\infty}2^{i+3}\exp(-u(i+1))\\
&\le\frac{8e^{-u}}{1-e^{-(u-\ell_n(2))}}\\
&\le Ae^{-u},
\end{aligned}\tag{2.15}$$

where $A>0$ is some large constant and with the convention that $w_k=0$ if $k\ge n$. Now notice that as $\sum_{k=1}^{n}w_k^3\le An$ for some constant A, we have for any $i\ge 0$, $w_{2^i}\le\frac{A^{1/3}n^{1/3}}{2^{i/3}}$. Hence :

$$\begin{aligned}
\sum_{i=1}^{\log(n)}(i+1)w_{2^i}&\le\sum_{i=1}^{+\infty}\frac{A^{1/3}(i+1)n^{1/3}}{2^{i/3}}\\
&\le A'n^{1/3},
\end{aligned}\tag{2.16}$$

where $A'>0$ is some large constant. With Equation (2.16), a simple computation shows that there exists $A>0$ such that :

$$\sum_{i=0}^{\infty}\rho_i=A'\left(\sqrt{mu}+un^{1/3}\right),$$

Replacing in Equation (2.15) give just another way of writing Bernstein's inequality, we finish by taking for instance :

$$u=\frac{y^2}{2A'^2(n^{1/3}y+m)},$$

which also ensures that $u>\ln(4)$ by Conditions 2. □

The following three lemmas have similar proofs, and their proofs are thus omitted.

Lemma 17. *Suppose that Conditions 1 hold. There exists $A>0$ such that, for any (m,y) that verifies Conditions 2 :*

$$\mathbb{P}\left(\sup_{0\le t\le m}|N(m)-N(m-t)-\mathbb{E}\left[N(m)-N(m-t)\right]|\ge y\right)\le A\exp\left(\frac{-y^2}{A(y+m)}\right).$$

Lemma 18. *Suppose that Conditions 1 hold. There exists $A>0$ such that, for any (m,y) that verify Conditions 2 :*

$$\mathbb{P}\left(\sup_{0\le t\le m}|N(t)-\mathbb{E}\left[N(t)\right]|\ge y\right|\le A\exp\left(\frac{-y^2}{A(y+m)}\right)$$

Lemma 19. *Suppose that Conditions 1 hold. There exists $A>0$ such that, for any (m,y) that verifies Conditions 2 :*

$$\mathbb{P}\left(\sup_{0\le t\le m}|X(m)-X(m-t)-\mathbb{E}\left[X(m)-X(m-t)\right]|\ge y\right)\le A\exp\left(\frac{-y^2}{A(yn^{1/3}+m)}\right).$$

Now we can prove the concentration of the size-biased sum of weights sampled without replacement.

Theorem 20. *Suppose that Conditions 1 hold. There exists a constant $A > 0$ that satisfies the following, for (m, y) that verifies Conditions 2, we have :*

$$\mathbb{P}\left[\sup_{0\leq i\leq j\leq m}\left|\sum_{k=i}^{j} w_{v(k)} - \mathbb{E}\left[\sum_{k=i}^{j} w_{v(k)}\right]\right| \geq y\right] \leq A\exp\left(\frac{-y^2}{A(yn^{1/3}+m)}\right).$$

Proof. Let $l(m)$ be such that $\mathbb{E}[N(l(m))] = m$. If $E = \{N(3(l(m)+y)) \geq m\}$ holds, then :

$$\sup_{0\leq i\leq j\leq m}\left|\sum_{k=i}^{j} w_{v'(k)} - \mathbb{E}\left[\sum_{k=i}^{j} w_{v'(k)}\right]\right| \leq \sup_{0\leq x\leq z\leq 3(l(m)+y))}\left|X(z) - X(x) - \sum_{k=N(x)}^{N(z)} \mathbb{E}\left[w_{v'(k)}\right]\right|.$$

We only bound :

$$\mathbb{P}\left[\inf_{i\leq j\leq m}\sum_{k=i}^{j} w_{v(k)} - \mathbb{E}\left[\sum_{k=i}^{j} w_{v(k)}\right] \leq -y\right],$$

as the argument for bounding the other part is the same. By union bound with the event E :

$$\begin{aligned}
&\mathbb{P}\left(\inf_{i\leq j\leq m}\sum_{k=i}^{j} w_{v'(k)} - \mathbb{E}\left[\sum_{k=i}^{j} w_{v'(k)}\right] \leq -y\right)\\
\leq&\mathbb{P}\left(E, \inf_{i\leq j\leq m}\sum_{k=i}^{j} w_{v'(k)} - \mathbb{E}\left[\sum_{k=i}^{j} w_{v'(k)}\right] \leq -y\right) + P(\bar{E})\\
\leq&\mathbb{P}\left(\inf_{0\leq x\leq z\leq 3(l(m)+y))} X(z) - X(x) - \sum_{k=N(x)}^{N(z)} \mathbb{E}\left[w_{v'(i)}\right] \leq -y\right) + P(\bar{E}).
\end{aligned} \tag{2.17}$$

Note that by Conditions 1, for n large enough :

$$\begin{aligned}
\mathbb{E}\left[N\left(\frac{\ell_n}{9}\right)\right] &\geq \sum_{k=1}^{n}\left(\frac{w_i}{9} - \frac{w_i^2}{162}\right)\\
&\geq \frac{\ell_n}{11}(1+o(1))\\
&\geq \frac{\ell_n}{12}.
\end{aligned} \tag{2.18}$$

Since $(\mathbb{E}[N(x)])_{x\geq 0}$ is an increasing function, by Equation (2.18), $l(m) \leq \ell_n/9$. Hence, by Equation (2.6) :

$$\begin{aligned}
\mathbb{E}[N(l(m)] &= m\\
&\geq l(m) - \frac{\sum_{k=1}^{n} w_k^2 l(m)^2}{2\ell_n^2}\\
&\geq l(m) - \frac{l(m)}{18}(1+o(1))\\
&\geq \frac{8l(m)}{9}.
\end{aligned} \tag{2.19}$$

By Lemma 12 and Equation (2.19) :

$$\mathbb{P}(\bar{E}) \leq A\exp\left(\frac{-y^2}{A(y+m)}\right), \tag{2.20}$$

for some large constant $A > 0$. Now we need to prove that :

$$\mathbb{P}\left(\inf_{0\le x\le z\le 3(l(m)+y))} X(z) - X(x) - \sum_{k=N(x)}^{N(z)} \mathbb{E}\left[w_{v(i)}\right] \le -y\right) \le A\exp\left(\frac{-y^2}{A(yn^{1/3}+x)}\right). \tag{2.21}$$

By equation (2.19) :

$$3(l(m)+y) \le 4m+3y. \tag{2.22}$$

Let :

$$\mathcal{C} = \left\{\inf_{0\le x\le z\le 4m+3y} X(z) - X(x) - \sum_{k=N(x)}^{N(z)} \mathbb{E}\left[w_{v(i)}\right] \le -y\right\},$$

and :

$$\mathcal{B} = \left\{X(4m+y) - \sum_{k=0}^{N(4m+3y)} \mathbb{E}\left[w_{v(i)}\right] \le -y/2\right\}.$$

Also, write

$$(x^*, z^*) = \inf\left\{0 \le x \le z \le 4m+3y :\ X(z) - X(x) - \sum_{k=N(x)}^{N(z)} \mathbb{E}\left[w_{v(i)}\right] \le -y\right\},$$

where the infimum is taken in lexicographical order. And, by convention, $\inf(\emptyset) = (0, 4m+3y)$. Let :

$$\mathcal{D} := \left\{X(x^*) - \sum_{k=1}^{N(x^*)} \mathbb{E}\left[w_{v(k)}\right] \ge y/4\right\} \text{ or } \left\{X(4m+y) - X(z^*) - \sum_{k=N(z^*)}^{N(4m+3y)} \mathbb{E}\left[w_{v(k)}\right] \ge y/4\right\}.$$

If $\mathcal{C}$ happens then one of the events $\mathcal{B}$ or $\mathcal{D}$ happens. By Lemma 15 :

$$\mathbb{P}(\mathcal{B}) \le A\exp\left(\frac{-y^2}{A(yn^{1/3}+m)}\right). \tag{2.23}$$

By Lemma 13 and union bound :

$$\begin{aligned}\mathbb{P}\left(X(x^*) - \sum_{k=1}^{N(x^*)} \mathbb{E}\left[w_{v(k)}\right] \ge y/4\right) \le& \mathbb{P}\left(\sup_{t\le 4m+3y} X(t) - \sum_{k=1}^{N(t)} \mathbb{E}\left[w_{v(k)}\right] \ge y/4\right)\\ \le& \mathbb{P}\left(\sup_{t\le 4m+3y} X(t) - \mathbb{E}\left[X(t)\right] \ge y/8\right)\\ &+\mathbb{P}\left(\sup_{t\le 4m+3y} N(t) - \mathbb{E}\left[N(t)\right] \ge \frac{y}{A}+1\right),\end{aligned} \tag{2.24}$$

where $A > 0$ is the positive constant that appears in Lemma 13. And by the same arguments, using Lemma 14 gives :

$$\begin{aligned}&\mathbb{P}\left(X(4m+3y) - X(z^*) - \sum_{k=N(z^*)}^{N(4m+3y)} \mathbb{E}\left[w_{v(k)}\right] \ge y/4\right)\\ \le&\mathbb{P}\left(\sup_{t\le 4m+3y} X(4m+3y) - X(t) - \sum_{k=N(t)}^{N(4m+3y)} \mathbb{E}\left[w_{v(k)}\right] \ge y/4\right)\\ \le&\mathbb{P}\left(\sup_{t\le 4m+3y} X(4m+3y) - X(t) - \mathbb{E}\left[X(4m+3y) - X(t)\right] \ge y/8\right)\\ &+\mathbb{P}\left(\sup_{t\le 4m+3y} N(4m+3y) - N(t) - \mathbb{E}\left[N(4m+3y) - N(t)\right] \ge \frac{y}{A'}+1\right).\end{aligned} \tag{2.25}$$

the union bound using Inequality (2.24) and (2.25) alongside Lemmas 16,17, 18 and 19 yield :

$$\mathbb{P}(\mathcal{D}) \leq A'' \exp\left(\frac{-y^2}{A''(yn^{1/3}+m)}\right) \tag{2.26}$$

Hence, from Equations (2.23) and (2.26) we obtain :

$$\begin{aligned}\mathbb{P}(\mathcal{C}) &\leq \mathbb{P}(\mathcal{B}) + \mathbb{P}(\mathcal{D}) \\ &\leq A''' \exp\left(\frac{-y^2}{A'''(yn^{1/3}+m)}\right).\end{aligned} \tag{2.27}$$

This proves Equation (2.21). We can then bound Equation (2.17) by using Equation (2.20) and Equation (2.27) which finishes the proof. □

In the above theorems we started the sums from one for the sake of clarity. The following general theorem has a similar proof.

Theorem 21. *Suppose that Conditions 1 hold. There exists a constant $A > 0$ such that, if $1 \leq l \leq m$ is such that $\left(\sqrt{m(m-l)}, y\right)$ verify Conditions 2, $m - l \to \infty$ and $y = O(m-l)$ then :*

$$\mathbb{P}\left[\sup_{l\leq i\leq j\leq m}\left|\sum_{k=i}^{j} w_{v(k)} - \mathbb{E}\left[\sum_{k=i}^{j} w_{v(k)}\right]\right| \geq y\right] \leq A\exp\left(\frac{-y^2}{A(yn^{1/3}+(m-l))}\right).$$

2.3 Bounds on the exploration process

In this section we prove concentration inequalities for the exploration process and related processes. These various inequalities will be used in the following sections. Recall that $f = o(n)$ is the critical parameter and $p_f = \frac{1}{\ell_n} + \frac{f}{\ell_n^{4/3}}$. In the rest of this section we consider the BFW of $G(\mathbf{W}, p_f)$.

For $0 \leq i \leq n$ and $0 \leq j \leq n$ define :

$$Y(i,j) = \mathbb{1}(\text{There is an edge between nodes } i \text{ and } j).$$

Then by definition of the BFW we have :

$$\begin{aligned}L_0 &= 1, \\ X_{i+1} &= \sum_{j\notin\mathcal{V}(i+L_i)} Y(v(i+1),j) - 1, \\ L_{i+1} &= \max(L_i + X_{i+1}, 1).\end{aligned} \tag{2.28}$$

Recall also that :

$$\begin{aligned}L'_0 &= 1, \\ L'_{i+1} &= L'_i + X_{i+1}.\end{aligned} \tag{2.29}$$

When seen as processes of i, L' is equal to L until we finish discovering the first connected component. After that $L' = L - 1$ until the second connected component is discovered, then $L' = L - 2$ and so on. Generally L' is equal to L minus the number of connected components fully discovered. We say that the process L visits 0 in i if $L'_i = \min_{j\leq i} L'_j$.

One of the difficulties in studying this process lies in the fact that X_{i+1} depends on L_i. In the case of simple Erdős–Rényi random graphs, Addario-Berry et al. [2009] use a different exploration process where the children of a node being explored are taken uniformly. This allows

them to use a simpler and close enough process in order to circumvent this problem. If we want to do like them, in our case the naive way to define such a process would be as follows, for $h \geq 0$:

$$\begin{aligned} L_0^h &= 1, \\ X_{i+1}^h &= \sum_{j \notin \mathcal{V}(i+1+h)} Y(v(i+1), j) - 1, \\ L_{i+1}^h &= L_i^h + X_{i+1}^h. \end{aligned}$$

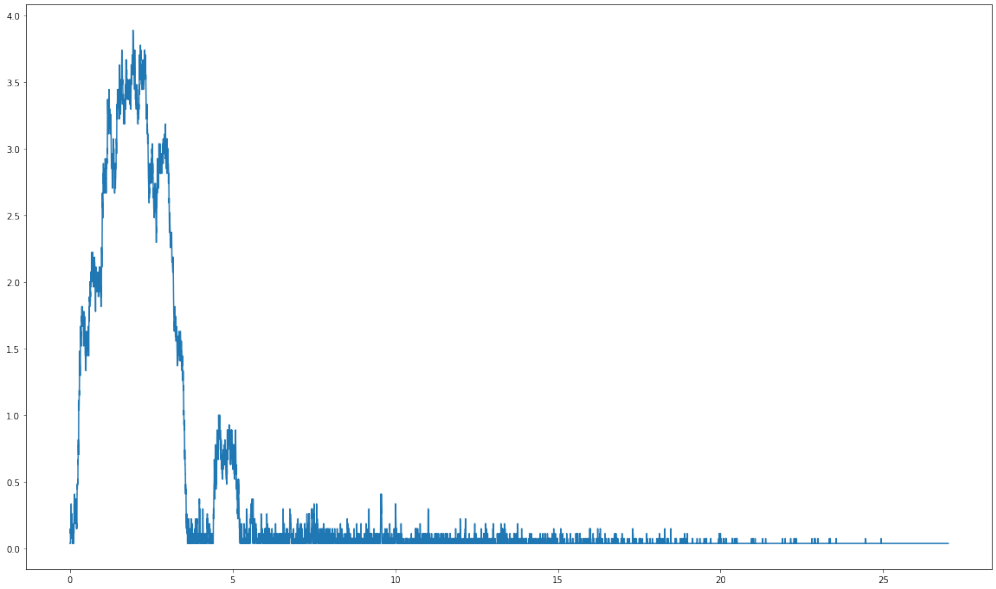


FIGURE 2.5 – The reflected exploration process of the graph in Figure 2.1 with time rescaled by $20000^{2/3}$ and space is rescaled by $20000^{1/3}$.

In that case L^0 is always above L' and in general $L_i^h \leq L_i'$ as long as $L_i \leq h+1$. L^0 is used to bound L' (and thus L) from above while L^h for h large enough would be used to bound it from below. However, in our case we sort the discovered children of a node by the weights of their edges. Hence, it is very likely that the indicator functions present in L_i' but not in L_i^h for $h > L_i$ will be equal to 1 and hence L_i^h would be too far away from L_i'. This is why we will use a martingale technique that we present now.

Note that for $i \geq 1$, L_i is $\sigma(X_1, X_2, ..., X_i)$ measurable. Let $(\mathcal{F}_i)_{i \geq 1}$ be the increasing sequence of σ-fields such that $\mathcal{F}_i$ is the σ-field generated by $\mathcal{V}(i+L_i)$ and the $(X_k)_{k \leq i}$'s, with the convention that $\mathcal{V}(k) = \mathcal{V}$ when $k \geq n$. Then for any $i \geq 1$, X_i is measurable with respect to $\mathcal{F}_i$ and moreover we have :

$$\mathbb{E}[X_i | \mathcal{F}_{i-1}] + 1 = \sum_{k > i + L_{i-1} - 1} (1 - e^{-w_{v(i)} w_{v(k)} p_f}).$$

And we have the following fact

Fact 22. *Let*

$$\tilde{L}_i = \sum_{k=0}^{i} \mathbb{E}[X_k | \mathcal{F}_{k-1}],$$

with the convention that $X_0 = 1$. Then for any $l \geq 0$, the process $(L'(i) - L'(l) - (\tilde{L}_i - \tilde{L}_l))_{i \geq l}$ is a martingale with respect to $(\mathcal{F}_i)_{i \geq l}$.

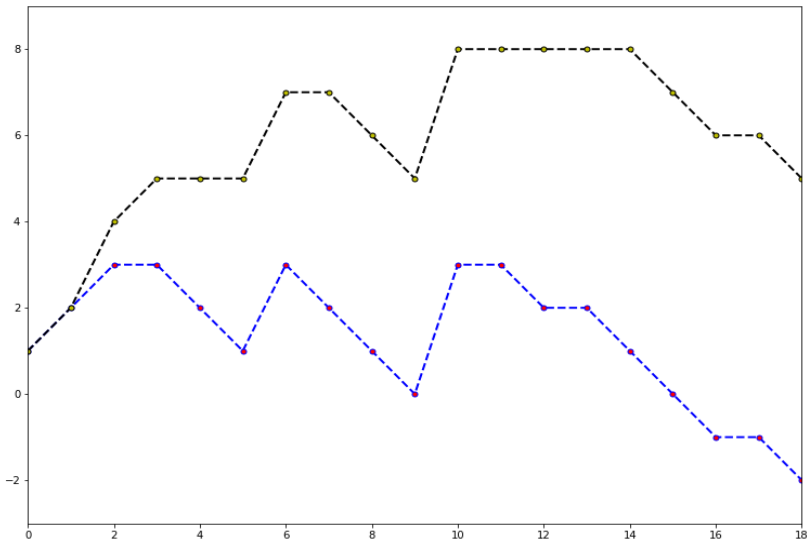


FIGURE 2.6 – In red with blue dashes, the exploration process of the graph in Figure 2.2. In yellow with black dashes, the process L^0 for the same graph. L^0 is always above L'.

This fact allows us to use Bernstein's inequality for martingales (Freedman [1975]). Then in order to bound L'_i from below, we will use the fact that $(L'_i - \tilde{L}_i)_{i\geq 1}$ is a martingale, and for $i \geq 1$ as long as $L_i \leq h$ we have :

$$\mathbb{E}[X_i|\mathcal{F}_{i-1}] + 1 \geq \sum_{k>i+h} (1 - e^{-w_{v(i)}w_{v(k)}p_f}).$$

This is why we define the following process, for $i \geq 1$ and $h \geq 0$:

$$\tilde{L}^h_m + m - 1 = \sum_{i=1}^{m} \sum_{k>i+h} (1 - e^{-w_{v(i)}w_{v(k)}p_f}),$$

then $\tilde{L}^h_m$ will be close to, and greater than $\tilde{L}_m$ as long as $h \geq L_i$ and h is not too large. A second important fact is that while constructing the exploration process, we never inspect the potential surplus edges, namely the $Y(v(i), v(j))$'s where $i \geq 1$ and $i + 1 \leq j \leq i + L_i - 1$. This means that :

Fact 23. *Conditionally on $\mathcal{F}_n$, the σ-field generated by $\mathcal{V}$ and the $(X_k)_{k\leq n}$'s, the random variables*

$$Y(v(i), v(j))_{1\leq i, \leq j\leq i+L_i-1},$$

are independent Bernoulli random variables of parameters

$$\left(1 - e^{-w_{v(i)}w_{v(j)}p_f}\right)_{1\leq i,\, i+1\leq j\leq i+L_i-1}.$$

Moreover, for $h \geq 0$ and $i \geq 0$ define

$$\bar{L}'(k) = \sum_{i=0}^{k} X_i \mathbb{1}(X_i \leq 2n^{1/3}),$$

and if we write $d(i)$ for the degree of node i, then $d(i)$ is a sum of independent Bernoulli variables. Hence, when Conditions 1 hold, by the classical Bernstein inequality (Bernstein [1924]) we have :

$$\mathbb{P}(d(i) \geq w_i + n^{1/3}) \leq \exp\left(\frac{-(n^{1/3})^2}{2(n^{1/3}+w_i)}\right).$$

By using Conditions 1 we have for n large enough :

$$\begin{aligned}\mathbb{P}(\exists(k,h),\, \bar{L}'(k) \neq L'(k)) &\leq \sum_{i=1}^{n} \mathbb{P}(d(i) \geq w_i + n^{1/3}) \\ &\leq \sum_{i=1}^{n} \exp\left(\frac{-(n^{1/3})^2}{2(n^{1/3}+w_i)}\right) \\ &\leq A\exp\left(\frac{-n^{1/3}}{A}\right),\end{aligned} \tag{2.30}$$

where A is some large constant, this probability is smaller than the ones we will get in this section and the one following it. It is also clear that Fact 22 also holds if we replace $L'(k)$ by $\bar{L}'(k)$ and X_i by $X_i\mathbb{1}(X_i \leq 2n^{1/3})$. Hence, we will assume that the increments of L' are smaller than $2n^{1/3}$. And we will assume the same of L^0. This will make computations lighter, as Bernstein's inequality requires a bound on the maximal increment of the process. We will not have to do a union bound in each calculation and consider the case where $\bar{L}'(k) \neq L'(k)$. This convention will be used up to Section 5, after that we will use the fact that the increments of $L'(k)$ are even smaller when k is large enough.

A direct corollary of Lemma 11 is the following :

Corollary 23.1. *For all $m \geq l \geq 1$ such that $m = o(n)$, and $h = o(n)$:*

$$\mathbb{E}(\tilde{L}^h_m - \tilde{L}^h_{l-1}) = (m-l)\left(f\ell_n^{-1/3} - \frac{C(m+l)+2h}{2\ell_n}\right) + 1 + o\left(\frac{m^2-l^2+(m-l)(h+n^{2/3})}{n}\right).$$

Proof. For any $l-1 \leq i \leq m$, let :

$$\tilde{X}^h_{i+1} = \sum_{j \notin \mathcal{V}(i+1+h)} (1 - e^{-w_{v(i+1)}w_j p_f}) - 1,$$

then :

$$\tilde{L}^h_{i+1} = \tilde{L}^h_i + \tilde{X}^h_{i+1},$$

and :

$$\mathbb{E}[\tilde{X}^h_i] + 1 = \mathbb{E}\left[\sum_{j \geq i+1+h} 1 - \exp\left(-w_{v(i)}w_{v(j)}p_f\right)\right]. \tag{2.31}$$

By Conditions 1, $w_{v(i)}w_{v(j)}p_f = o(1)$ deterministically for any (i,j). The bounds giving O and o in the following expectations can thus be chosen to be deterministic. By Equation (2.6) we have :

$$\begin{aligned}\mathbb{E}[\tilde{X}^h_i] + 1 &= \mathbb{E}\left[\sum_{j \geq i+1+h} w_{v(i)}w_{v(j)}p_f(1 + O(w_{v(i)}w_{v(j)}p_f))\right] \\ &= \mathbb{E}\left[w_{v(i)}\left(1 + f\ell_n^{-1/3} + O\left(\sum_{j=1}^{n} w_{v(i)}w^2_{v(j)}p_f^2\right)\right) - \sum_{j<i+1+h} w_{v(i)}w_{v(j)}p_f(1+o(1))\right] \\ &= \mathbb{E}\left[w_{v(i)}\left(1 + f\ell_n^{-1/3} + o\left(n^{-2/3}\right)\right) - \sum_{j<i+1+h} w_{v(i)}w_{v(j)}p_f(1+o(1))\right].\end{aligned} \tag{2.32}$$

We use Lemmas 8 and 11 to do the proper replacements in Equation (2.32) :

$$\mathbb{E}[\tilde{X}_i^h] = -1 + \left(1 + \frac{i(1-C)}{\ell_n} + o\left(\frac{i+n^{2/3}}{n}\right)\right)\left(1 + \frac{f}{\ell_n^{1/3}}\right)(1+o(1)) - \mathbb{E}\left[\sum_{j<i+1+h} w_{v(i)}w_{v(j)}p_f(1+o(1))\right].$$

Finally, Lemma 10 yields :

$$\mathbb{E}[\tilde{X}_i^h] = -1 + \left(1 + \frac{i(1-C)}{\ell_n} + o\left(\frac{i+n^{2/3}}{n}\right)\right)\left(1 + \frac{f}{\ell_n^{1/3}}\right)(1+o(1)) - \frac{i+h}{\ell_n}(1+o(1)).$$

Summing over i ends the proof. □

We will first show concentration results for $\tilde{L}^h$ before moving to L. We start by stating a set of conditions that will ensure the theorems holds.

Conditions 3. *We say that $(a(n), b(n), c(n), d(n))$ verifies Conditions 3 if :*

$$a(n) + c(n) = o(n),$$

and :

$$\lim_n (a(n) - b(n)) = +\infty,$$

and

$$d(n) = O(a(n) - b(n)),$$

and

$$\left(\sqrt{(a(n)-b(n))(a(n)+c(n))}, d(n)\right)$$

verify Conditions 2.

We start with the following technical lemma. Concentration follows here from the concentration of the ordered weights proved in the previous section.

Lemma 24. *Suppose that Conditions 1 hold. There exists a constant $A > 0$ such that, if (m, l, h, y) verifies Conditions 3, then the following is true :*

$$\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\left|\tilde{L}_j^h - \tilde{L}_i^h - \mathbb{E}\left[\tilde{L}_j^h - \tilde{L}_i^h\right]\right|\right) \ge y\right) \le A\exp\left(\frac{-y^2}{A(yn^{1/3}+m-l)}\right),$$

Proof. Let :

$$D = \mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\left|\tilde{L}_j^h - \tilde{L}_i^h - \mathbb{E}\left[\tilde{L}_j^h - \tilde{L}_i^h\right]\right|\right) \ge y\right)$$

Since $p_f \ge 1/n$ and $m - l = o(n)$. Conditions 1 and Equation (2.6) yield :

$$\sum_{k=i+1}^{j}\sum_{k'>k+h}\left(1 - e^{-w_{v(k)}w_{v(k')}p_f} - \mathbb{E}\left[1 - e^{-w_{v(k)}w_{v(k')}p_f}\right]\right) = \sum_{k=i+1}^{j}\left(\sum_{k'>k+h} w_{v(k)}w_{v(k')}p_f - \mathbb{E}\left[w_{v(k)}w_{v(k')}p_f\right]\right) + O(1).$$

Moreover, recall, by our conditions, that $y = y(n)$ and $\lim_{n\to\infty} y(n) = +\infty$. Since

$$\sum_{k'>k+h} w_{v(k)}w_{v(k')}p_f = \sum_{k'=1}^{n} w_{v(k)}w_{v(k')}p_f - \sum_{k'\le k+h} w_{v(k)}w_{v(k')}p_f,$$

we obtain by the union bound for n large enough :

$$\begin{aligned} D &\leq \mathbb{P}\left(\sup_{l\leq i\leq j\leq m}\left|\sum_{k=i+1}^{j}\left(\sum_{k'>k+h} w_{v(k)}w_{v(k')}p_f\right) - \mathbb{E}\left[\sum_{k=i+1}^{j}\left(\sum_{k'>k+h} w_{v(k)}w_{v(k')}p_f\right)\right]\right| \geq y/2\right) \\ &\leq \mathbb{P}\left(\sup_{l\leq i\leq j\leq m}\left|\sum_{k=i+1}^{j}\left(\sum_{k'\leq k+h} w_{v(k)}w_{v(k')}p_f\right) - \mathbb{E}\left[\sum_{k=i+1}^{j}\left(\sum_{k'\leq k+h} w_{v(k)}w_{v(k')}p_f\right)\right]\right| \geq y/4\right) \\ &+ \mathbb{P}\left(\sup_{l\leq i\leq j\leq m}\left|\sum_{k=i+1}^{j} w_{v(k)} - \mathbb{E}\left[\sum_{k=i+1}^{j} w_{v(k)}\right]\right| \geq \frac{y}{4\ell_n p_f}\right). \end{aligned} \tag{2.33}$$

Since $\ell_n p_f \leq 2$, by Conditions 3 we can apply Theorem 21 with (m, l, y) to obtain :

$$\mathbb{P}\left(\sup_{l\leq i\leq j\leq m}\left|\sum_{k=i+1}^{j} w_{v(k)} - \mathbb{E}\left[\sum_{k=i+1}^{j} w_{v(k)}\right]\right| \geq \frac{y}{4\ell_n p_f}\right) \leq A\exp\left(\frac{-y^2}{A(yn^{1/3}+m-l)}\right). \tag{2.34}$$

By injecting Inequality (2.34) in Inequality (2.33), bounding D amounts to bounding :

$$\mathbb{P}\left(\sup_{l\leq i\leq j\leq m}\left|\sum_{k=i+1}^{j}\left(\sum_{k'\leq k+h} w_{v(k)}w_{v(k')}p_f\right) - \mathbb{E}\left[\sum_{k=i+1}^{j}\left(\sum_{k'\leq k+h} w_{v(k)}w_{v(k')}p_f\right)\right]\right| \geq y/4\right).$$

We focus on proving a one-sided version of this inequality, the other half of the inequality is proven similarly :

$$\mathbb{P}\left(\sup_{l\leq i\leq j\leq m}\left(\sum_{k=i+1}^{j}\left(\sum_{k'\leq k+h} w_{v(k)}w_{v(k')}p_f\right) - \mathbb{E}\left[\sum_{k=i+1}^{j}\left(\sum_{k'\leq k+h} w_{v(k)}w_{v(k')}p_f\right)\right]\right) \geq y/4\right).$$

By Lemmas 8 and 10, for any $l \leq i \leq j \leq m$:

$$\mathbb{E}\left[\sum_{k=i+1}^{j}\left(\sum_{k'\leq k+h} w_{v(k)}w_{v(k')}p_f\right)\right] = \frac{j^2-i^2+2(j-i)h}{2\ell_n}(1+o(1)). \tag{2.35}$$

By a simple computation, Conditions 3 imply that $\left(m+h, \frac{y(m+h)}{16(m-l)}\right)$ verify Conditions 2. Using this with Theorem 21 yields for n large enough :

$$\begin{aligned} &\mathbb{P}\left(\sup_{l\leq k\leq m}\left|\sum_{k'\leq k+h} w_{v(j)} - \mathbb{E}\left[\sum_{k'\leq k+h} w_{v(j)}\right]\right| \geq \frac{y}{16p_f(m-l)}\right) \\ &\leq \mathbb{P}\left(\sup_{1\leq k\leq m+h}\left|\sum_{k'=1}^{k} w_{v(j)} - \mathbb{E}\left[\sum_{k'=1}^{k} w_{v(j)}\right]\right| \geq \frac{y(m+h)}{16(m-l)}\right) \\ &\leq A\exp\left(\frac{-y^2(m+h)^2}{A(y(m+h)(m-l)n^{1/3}+(m+h)(m-l)^2}\right) \\ &\leq A\exp\left(\frac{-y^2}{A(yn^{1/3}+(m-l)}\right). \end{aligned}$$

Hence, by the above inequality and Equation (2.35) the union bound yields :

$$\begin{aligned}
&\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\sum_{k=i+1}^{j} w_{v(k)}\left(\sum_{k'\le k+h} w_{v(k')}p_f\right) - \mathbb{E}\left[\sum_{k=i+1}^{j}\left(\sum_{k'\le k+h} w_{v(k)}w_{v(k')}p_f\right)\right]\right) \ge y/4\right)\\
\le&\mathbb{P}\left(\sup_{l\le k\le m}\left|\sum_{k'\le k+h} w_{v(k')} - \mathbb{E}\left[\sum_{k'\le k+h} w_{v(k')}\right]\right| \ge \frac{y}{16p_f(m-l)}\right)\\
&+\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\sum_{k=i+1}^{j} w_{v(k)}\left(\frac{y}{16(m-l)}\right)\right) \ge y/8\right)\\
&+\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\sum_{k=i+1}^{j} w_{v(k)}\mathbb{E}\left[\sum_{k'\le k+h} w_{v(k')}p_f\right] - \frac{j^2-i^2+2(j-i)h}{2\ell_n}(1+o(1))\right) \ge y/8\right)\\
\le&A\exp\left(\frac{-y^2}{A(yn^{1/3}+(m-l)}\right) + \mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\sum_{k=i+1}^{j} w_{v(k)}\left(\frac{y}{16(m-l)}\right)\right) \ge y/8\right)\\
&+\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\sum_{k=i+1}^{j} w_{v(k)}\mathbb{E}\left[\sum_{k'\le k+h} w_{v(k')}p_f\right] - \frac{j^2-i^2+2(j-i)h}{2\ell_n}(1+o(1))\right) \ge y/8\right).
\end{aligned} \tag{2.36}$$

By Corollary 9, for any $k \le m$:

$$\mathbb{E}\left[\sum_{k'\le k+h} w_{v(k')}p_f\right] = \frac{(k+h)(1+o(1))}{\ell_n}. \tag{2.37}$$

Moreover, notice that for any $l \le i \le j \le m$:

$$\sum_{k=i+1}^{j} w_{v(k)}\frac{k+h}{\ell_n} = \frac{i}{\ell_n}\sum_{k=i+1}^{j} w_{v(k)} + \frac{h}{\ell_n}\sum_{k=i+1}^{j} w_{v(k)} + \left(\frac{1}{\ell_n}\right)\sum_{k=i+1}^{j}\sum_{k'=k}^{j} w_{v(k')}.$$

By Conditions 3, we have for any for any $l \le i \le j \le m$:

$$\frac{j^2-i^2+2(j-i)h}{2\ell_n} = O(y). \tag{2.38}$$

Moreover, by Conditions 3 $y \le A(m-l)$, for some large constant $A > 0$. Hence, by the union bound, Equation (2.36) becomes :

$$\begin{aligned}
&\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\sum_{k=i+1}^{j} w_{v(k)}\left(\sum_{k'\le k+h} w_{v(k')}p_f\right) - \mathbb{E}\left[\sum_{k=i+1}^{j}\left(\sum_{k'\le k+h} w_{v(k)}w_{v(k')}p_f\right)\right]\right) \ge y/4\right)\\
\le&\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\frac{i}{\ell_n}\sum_{k=i+1}^{j}(w_{v(k)} - \mathbb{E}[w_{v(k)}])(1+o(1))\right) \ge \frac{y}{48}\right)\\
&+\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\left(\frac{h}{\ell_n}\right)\sum_{k=i+1}^{j}(w_{v(k)} - \mathbb{E}[w_{v(k)}])(1+o(1))\right) \ge \frac{y}{48}\right)\\
&+\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\left(\frac{1}{\ell_n}\right)\sum_{k=i+1}^{j}\sum_{k'=k}^{j}(w_{v(k')} - \mathbb{E}[w_{v(k')}])(1+o(1))\right) \ge \frac{y}{48}\right)\\
&+A\exp\left(\frac{-y^2}{A(yn^{1/3}+(m-l)}\right) + \mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\sum_{k=i+1}^{j} w_{v(k)}\right) \ge \frac{2y}{A}\right).
\end{aligned} \tag{2.39}$$

Notice that we implicitly use Equation (2.38) in the above Inequality in order to make the o factors match and at the cost of taking $y/48$. This is why we are able to write :

$$\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\frac{i}{\ell_n}\sum_{k=i+1}^{j}(w_{v(k)}-\mathbb{E}[w_{v(k)}])(1+o(1))\right)\ge\frac{y}{48}\right),$$

instead of :

$$\mathbb{P}\left(\sup_{l\le i\le j\le m}\left(\frac{i}{\ell_n}\sum_{k=i+1}^{j}\big(w_{v(k)}(1+o(1))-\mathbb{E}[w_{v(k)}](1+o(1))\big)\right)\ge\frac{y}{24}\right).$$

By Conditions 3 we can apply Theorem 21 with $(m-l, y/48)$ to obtain :

$$\mathbb{P}\left(\sup_{l\le i\le j\le m}\left|\sum_{k=i+1}^{j}w_{v(k)}-\mathbb{E}\left[\sum_{k=i+1}^{j}w_{v(k)}\right]\right|\ge\frac{y}{48}\right)\le A\exp\left(\frac{-y^2}{A(yn^{1/3}+m-l)}\right). \tag{2.40}$$

We finish by noticing that the first three probabilities in the right-hand side of Inequality 2.39 are all smaller than the left hand-side of 2.40. □

Theorem 25. *Suppose that Conditions 1 hold. There exists a constant $A>0$ such that, if $(m,l,0,y)$ verifies Conditions 3, then the following holds :*

$$\mathbb{P}\left(\sup_{l\le u\le w\le m}\left|L_w^0-L_u^0-\mathbb{E}[L_w^0-L_u^0]\right|\ge y\right)\le A\exp\left(\frac{-y^2}{A(yn^{1/3}+m-l)}\right).$$

Proof. For $i\ge 0$ let $\mathcal{F}_i^0$ be the sigma-field generated by $\mathcal{V}_{i+1}$ and the random variables $(X_k^0)_{k\le i}$. Write :

$$D_1=\mathbb{P}\left(\sup_{l\le u\le w\le m}\left|L_w^0-L_u^0-\sum_{i=u+1}^{w}\mathbb{E}\left[X_i^0|\mathcal{F}_{i-1}^0\right]\right|\ge y/2\right),$$

and

$$D_2=\mathbb{P}\left(\sup_{l\le u\le w\le m}\left|\sum_{i=u+1}^{w}\mathbb{E}\left[X_i^0|\mathcal{F}_{i-1}^0\right]-\mathbb{E}[L_w^0-L_u^0]\right|\ge y/2\right).$$

Then, by the union bound :

$$\mathbb{P}\left(\sup_{l\le u\le w\le m}\left|L_w^0-L_u^0-\mathbb{E}[L_w^0-L_u^0]\right|\ge y\right)\le D_1+D_2.$$

We start by bounding D_1. We have by the union bound :

$$\begin{aligned}D_1\le&\mathbb{P}\left(\sup_{l\le u\le m}\left|L_u^0-L_l^0-\sum_{i=l+1}^{u}\mathbb{E}\left[X_i^0|\mathcal{F}_{i-1}^0\right]\right|\ge y/4\right)\\&+\mathbb{P}\left(\sup_{l\le w\le m}\left|L_w^0-L_l^0-\sum_{i=l+1}^{w}\mathbb{E}\left[X_i^0|\mathcal{F}_{i-1}^0\right]\right|\ge y/4\right).\end{aligned}$$

Notice that

$$\left(L_w^0-L_l^0-\sum_{i=l+1}^{w}\mathbb{E}\left[X_i^0|\mathcal{F}_{i-1}^0\right]\right)_{w\ge l},$$

is a martingale with respect to the $(\mathcal{F}_i^0)_{i\ge l}$'s. Moreover :

$$\mathbb{E}[(X_i^0)^2|\mathcal{F}_{i-1}^0]=\sum_{k\notin\mathcal{V}_i}\sum_{k'\notin\mathcal{V}_i}\mathbb{E}[Y(v(i),k)Y(v(i),k')|\mathcal{F}_{i-1}^0]\le w_{v(i)}+w_{v(i)}^2.$$

Applying Theorem 6 to the $(w^2_{v(i)})$'s, let $J(1), J(2), ...$ be i.i.d copies of $v(l+1)$. We have by Lemma 8 the following Bernstein inequality :

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=l+1}^{m} w^2_{v(i)} \geq \sum_{i=l+1}^{m} 2\mathbb{E}[w^2_{v(i)}] + 2yn^{1/3}\right) \\
&\leq \mathbb{E}\left[\mathbb{E}\left[\frac{\exp\left(\sum_{i=l+1}^{m} w^2_{v(i)}\right)}{\exp\left(\sum_{i=l+1}^{m} 2C\mathbb{E}[w^2_{v(i)}] + 2yn^{1/3}\right)}\right]\right] \\
&\leq \mathbb{E}\left[\mathbb{E}\left[\frac{\exp\left(\sum_{i=l+1}^{m} w^2_{J(i)}\right)}{\exp\left(2C(m-l)(1+o(1)) + 2yn^{1/3}\right)}\right]\right] \\
&\leq \mathbb{E}\left[\exp\left(\frac{-\left|2C(m-l)(1+o(1)) + 2yn^{1/3} - \sum_{i=l+1}^{m}\mathbb{E}[w^2_{J(i)}]\right|_+^2}{\left(Ayn + A(m-l)n^{2/3} + \sum_{i=l+1}^{m}\mathbb{E}[w^4_{J(i)}]\right)}\right)\right] \\
&\leq \mathbb{E}\left[\exp\left(\frac{-\left|2C(m-l)(1+o(1)) + 2yn^{1/3} - \sum_{i=l+1}^{m}\mathbb{E}[w^2_{J(i)}]\right|_+^2}{An^{2/3}\left(yn^{1/3} + (m-l) + \sum_{i=l+1}^{m}\mathbb{E}[w^2_{J(i)}]\right)}\right)\right],
\end{aligned} \tag{2.41}$$

where line 3 of the equation is a Chernoff bound which yields Bernstein's inequality in line 4 (as in the original proof of Bernstein [1924]), and we used the fact that, by Conditions 1, we have $\mathbb{E}[w^4_{J(i)}] \leq n^{2/3}\mathbb{E}[w^2_{J(i)}]$. Now notice that by definition, and by Lemma 8, for any $i \geq l+1$:

$$\mathbb{E}[w^2_{J(i)}] = C(1+o(1)). \tag{2.42}$$

Since $(m, l, 0, y)$ verifies Conditions 3, we can apply Theorem 21 on $(m, l, 2y)$ to obtain :

$$\mathbb{P}\left(\sum_{i=l+1}^{m} w_{v(i)} \geq \sum_{i=l+1}^{m}\mathbb{E}[w_{v(i)}] + 2y\right) \leq A\exp\left(\frac{-y^2}{A\left(yn^{1/3} + (m-l)\right)}\right), \tag{2.43}$$

where $A > 0$ is a large enough constant. By Equations (2.41), (2.42) and (2.43) and by Bernstein's inequality for martingales (Theorem 2.1 in Freedman [1975]) we obtain :

$$D_1 \leq A'\exp\left(\frac{-y^2}{A'(yn^{1/3} + m - l)}\right). \tag{2.44}$$

In order to bound D_2 notice that the sum inside D_2 is equal to the one in Lemma 24 when $h = 0$ by definition. This finishes the proof. □

Since L^0 is always greater than L' deterministically, Theorem 25 gives us the following theorem.

Theorem 26. *Suppose that Conditions 1 hold. Let* $\frac{4f\ell_n^{2/3}}{C} \geq m \geq \frac{f\ell_n^{2/3}}{C}$, *then there exists* $A > 0$ *and* $A' > 0$ *such that for any* $\varepsilon > 0$ *:*

$$\mathbb{P}\left(\sup_{1\leq i\leq m}(L_i) \geq \frac{10f^2\ell_n^{1/3}}{C}\right) \leq A\exp\left(\frac{-f}{A}\right).$$

Proof. By definition :

$$\sup_{1\leq i\leq m}(L_i) \leq \sup_{1\leq u\leq v\leq m}(L^0_v - L^0_u).$$

Hence

$$\mathbb{P}\left(\sup_{1\le i\le m}(L_i)\ge\frac{10f^2\ell_n^{1/3}}{C}\right)\le\mathbb{P}\left(\sup_{1\le u\le v\le m}(L^0_v-L^0_u)\ge\frac{10f^2\ell_n^{1/3}}{C}\right). \tag{2.45}$$

From Corollary 23.1 :

$$\begin{aligned}\min_{1\le u\le v\le m}(\mathbb{E}[L^0_v-L^0_h])&=\min_{1\le u\le v\le m}\left((v-u)\left(f\ell_n^{-1/3}-\frac{C(v+u)}{2\ell_n}\right)(1+o(1))+1\right)\\&\ge\min_{1\le u\le v\le m}\left(-\frac{C(v+u)(v-u)}{2\ell_n}(1+o(1))+1\right)\\&\ge\frac{-9f^2\ell_n^{1/3}}{C}.\end{aligned} \tag{2.46}$$

We finish by injecting Equation (2.46) in (2.45) and using Theorem 25 with $\left(m,1,0,\frac{f^2\ell_n^{1/3}}{C}\right)$. □

The same method that we used to bound the term D_1 in the proof of Theorem 25 directly yields

Theorem 27. *Suppose that Conditions 1 hold. There exists a constant $A>0$ such that, if $(m,l,0,y)$ verifies Conditions 3, then the following holds :*

$$\mathbb{P}\left(\sup_{l\le u\le v\le m}\left|L'_v-L'_u-\mathbb{E}[\tilde{L}_v-\tilde{L}_u]\right|\ge y\right)\le A\exp\left(\frac{-y^2}{A(yn^{1/3}+m-l)}\right).$$

2.4 The structure of the giant component

The bounds in the previous section will allow us to determine the structure of the giant component of $G(\mathbf{W},p_f)$. We write H^*_f for the component of $G(\mathbf{W},p_f)$ being explored at time $\frac{f\ell_n^{2/3}}{C}$. We will prove that this component is the largest one with high enough probability. Informally, the BFW has a random unbiased part plus a drift (its expectation). Corollary 23.1 shows that the drift of L^0 is a parabola that has its maximum at $\frac{f\ell_n^{2/3}}{C}$. Given the concentration of L^0, and if we also assume that it behaves like L, it follows that L also has its maximum around $\frac{f\ell_n^{2/3}}{C}$. Now recall that L corresponds to the number of nodes discovered but not yet explored. It is then naturally maximal when the exploration process is in a large connected component. Hence H^*_f should be the largest component. In this section we will prove this rigorously. Then we will prove in the following section that the other connected components are small enough. [3]

2.4.1 The size of the giant component

Theorem 28. *Suppose that Conditions 1 hold. Let $1>\varepsilon'>0$. For f large enough and for any $1\ge\varepsilon>0$ consider the following event :*
The exploration of H^*_f starts before time $\frac{\ell_n^{2/3}}{f^{1-\varepsilon}C}$ and ends between times $\frac{2(1-\varepsilon')f\ell_n^{2/3}}{C}$ and $\frac{2(1+\varepsilon')f\ell_n^{2/3}}{C}$.
Then there exists a positive constant $A>0$ such that the probability of this event not happening is at most

$$A\exp\left(\frac{-f^\varepsilon}{A}\right).$$

3. In the rest of the proof, and in order to ease notations we do not use integer part notations for the indices and instead abuse notation by using real indices in our sums sometimes.

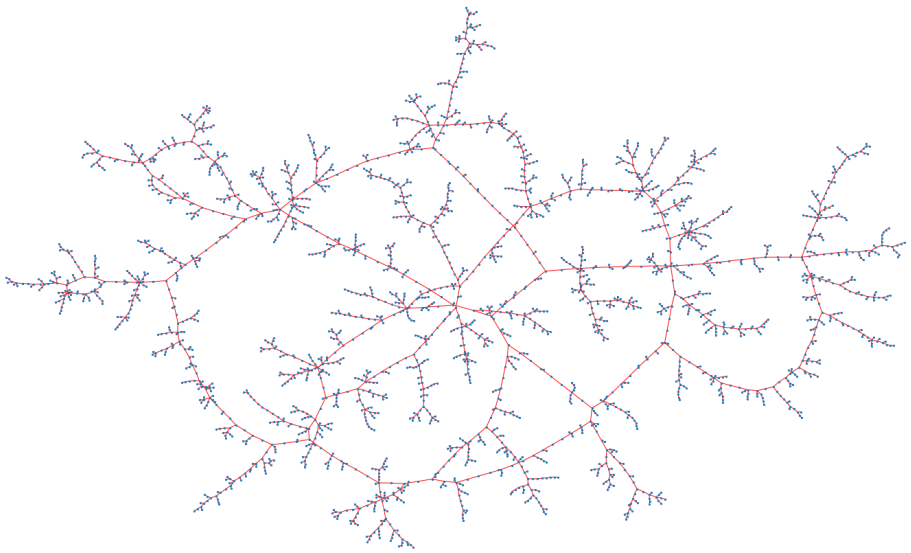


FIGURE 2.7 – The largest connected component of the graph in Figure 2.1. Its size is 2654.

Proof. Let $t_1 = \frac{\ell_n^{2/3}}{f^{1-\varepsilon}C}$, $t_2 = \frac{2(1-\varepsilon')f\ell_n^{2/3}}{C}$ and $t_3 = \frac{2(1+\varepsilon')f\ell_n^{2/3}}{C}$.
In order to prove this theorem we need to bound the probability that L visits zero between times t_1 and t_2 and also the probability that L does not visit 0 between times t_2 and t_3. Recall that for any i :

$$\tilde{L}_i = \sum_{k=1}^{i} \mathbb{E}[X_k|\mathcal{F}_{k-1}].$$

We start by the probability of the first event. Recall that by definition $L \geq L'$. We will thus focus on L'. For any $h > 0$, $\tilde{L}$ is at least $\tilde{L}^h$ until the first time i when $L_i \geq h$.

Let $h = \frac{10f^2\ell_n^{1/3}}{C}$. Then by Theorem 26 and Conditions 1 :

$$\mathbb{P}\left(\sup_{1\leq j\leq t_2} L_j \geq h\right) \leq A\exp\left(\frac{-f}{A}\right). \tag{2.47}$$

Now divide the interval $[t_1, t_2]$ by introducing intervals of the form $[t'_i, t'_{i+1}]$ with

$$t'_i = t_1 + \frac{2^{i+1}\ell_n^{2/3}}{f^{1-\varepsilon}C}.$$

This subdivision is necessary in order to respect Conditions 3 when we apply our concentration theorems. We stop at $t'_{\bar{i}} = t_2$ by truncating the last interval. By Corollary 23.1 and a straightforward calculation, for $i < \bar{i} - 1$:

$$\min_{t'_i\leq j\leq t'_{i+1}} \mathbb{E}(\tilde{L}_j^h) \geq \frac{2^i\varepsilon' f^\varepsilon \ell_n^{1/3}}{2C}, \tag{2.48}$$

and :

$$\min_{t'_{\bar{i}-1}\leq j\leq t'_{\bar{i}}} \mathbb{E}(\tilde{L}_j^h) \geq \frac{\varepsilon' f^2 \ell_n^{1/3}}{2C}. \tag{2.49}$$

A simple computation shows that we can apply Theorem 24 to $\tilde{L}^h$ between 1 and t_{i+1} in order to obtain the following inequalities for $i < \bar{i} - 1$ and for $\bar{i}$:

$$\begin{aligned}
&\mathbb{P}\left(\inf_{t'_i \le j \le t'_{i+1}} (\tilde{L}^h_j - \mathbb{E}(\tilde{L}^h_j)) \le -\frac{2^{i-1}\varepsilon' f^\varepsilon \ell_n^{1/3}}{2C}\right) \le A\exp\left(\frac{-2^{i-1}f^\varepsilon}{A}\right),\\
&\mathbb{P}\left(\inf_{t'_{\bar{i}-1} \le j \le t'_{\bar{i}}} (\tilde{L}^h_j - \mathbb{E}(\tilde{L}^h_j)) \le -\frac{\varepsilon' f^2 \ell_n^{1/3}}{4C}\right) \le A\exp\left(\frac{-f}{A}\right).
\end{aligned} \tag{2.50}$$

By the union bound using Equations (2.48), (2.49) and (2.50), we get :

$$\begin{aligned}
&\mathbb{P}\left(\inf_{t_1 \le j \le t_2} L_j \le 0\right)\\
&\le \sum_{i=0}^{\bar{i}-1} \mathbb{P}\left(\inf_{t'_i \le j \le t'_{i+1}} \tilde{L}_j \le \frac{2^{i-1}\varepsilon' f^\varepsilon \ell_n^{1/3}}{2C}\right) + \sum_{i=0}^{\bar{i}-1} \mathbb{P}\left(\inf_{t'_i \le j \le t'_{i+1}} (L'_j - \tilde{L}_j) \le -\frac{2^{i-1}\varepsilon' f^\varepsilon \ell_n^{1/3}}{2C}\right)\\
&+ \mathbb{P}\left(\inf_{t'_{\bar{i}-1} \le j \le t'_{\bar{i}}} \tilde{L}_j \le \frac{\varepsilon' f^2 \ell_n^{1/3}}{4C}\right) + \mathbb{P}\left(\inf_{t'_{\bar{i}-1} \le j \le t'_{\bar{i}}} (L'_j - \tilde{L}_j) \le -\frac{\varepsilon' f^2 \ell_n^{1/3}}{4C}\right)\\
&\le \sum_{i=0}^{\bar{i}-1} \mathbb{P}\left(\inf_{t'_i \le j \le t'_{i+1}} \tilde{L}^h_j \le \frac{2^{i-1}\varepsilon' f^\varepsilon \ell_n^{1/3}}{2C}\right) + \sum_{i=0}^{\bar{i}-1} \mathbb{P}\left(\inf_{t'_i \le j \le t'_{i+1}} (L'_j - \tilde{L}_j) \le -\frac{2^{i-1}\varepsilon' f^\varepsilon \ell_n^{1/3}}{2C}\right)\\
&+ \mathbb{P}\left(\inf_{t'_{\bar{i}-1} \le j \le t'_{\bar{i}}} \tilde{L}^h_j \le \frac{\varepsilon' f^2 \ell_n^{1/3}}{4C}\right) + \mathbb{P}\left(\inf_{t'_{\bar{i}-1} \le j \le t'_{\bar{i}}} (L'_j - \tilde{L}_j) \le -\frac{\varepsilon' f^2 \ell_n^{1/3}}{4C}\right)\\
&+ \mathbb{P}\left(\sup_{1 \le j \le t_2} L_j \ge h\right)\\
&\le \sum_{i=0}^{\bar{i}-1} \mathbb{P}\left(\inf_{t'_i \le j \le t'_{i+1}} (\tilde{L}^h_j - \mathbb{E}[\tilde{L}^h_j]) \le -\frac{2^{i-1}\varepsilon' f^\varepsilon \ell_n^{1/3}}{2C}\right) + \sum_{i=0}^{\bar{i}-1} \mathbb{P}\left(\inf_{t'_i \le j \le t'_{i+1}} (L'_j - \tilde{L}_j) \le -\frac{2^{i-1}\varepsilon' f^\varepsilon \ell_n^{1/3}}{2C}\right)\\
&+ \mathbb{P}\left(\inf_{t'_{\bar{i}-1} \le j \le t'_{\bar{i}}} (\tilde{L}^h_j - \mathbb{E}[\tilde{L}^h_j]) \le -\frac{\varepsilon' f^2 \ell_n^{1/3}}{4C}\right) + \mathbb{P}\left(\inf_{t'_{\bar{i}-1} \le j \le t'_{\bar{i}}} (L'_j - \tilde{L}_j) \le -\frac{\varepsilon' f^2 \ell_n^{1/3}}{4C}\right)\\
&+ \mathbb{P}\left(\sup_{1 \le j \le t_2} L_j \ge h\right)\\
&\le \sum_{i=0}^{\infty} A\exp\left(\frac{-2^{i-1}f^\varepsilon}{A}\right) + A\exp\left(\frac{-f}{A}\right)\\
&\le A'\exp\left(\frac{-f^\varepsilon}{A'}\right),
\end{aligned} \tag{2.51}$$

here the constant $A' > 0$ is large enough and of course these inequalities only hold for n large enough.

We now show that L visits 0 between times t_2 and t_3. Recall that $(Z(i))_{i \le n}$ is defined by $Z(i) = L_i - L'_i$. Then if $L'_{t_3} \le -Z(t_2)$, it means that L' attained a new minimum between t_2 and t_3 i.e L visited 0 between t_2 and t_3. Also, by construction, $Z(i) = -\min_{j \le i}(L'_j) + 1$. Since L' is deterministically smaller than L^0, if $L'_{t_3} \ge -Z(t_2)$ then $L^0_{t_3} \ge -Z(t_2)$. Therefore, it is sufficient to bound $\mathbb{P}(L^0_{t_3} \ge -Z(t_2))$. We do so by introducing an intermediate term :

$$\begin{aligned}
\mathbb{P}(L^0_{t_3} \ge -Z_{t_2}) &\le \mathbb{P}\left(L^0_{t_3} \ge -\frac{\varepsilon' f^2 \ell_n^{1/3}}{C}\right) + \mathbb{P}\left(Z(t_2) \ge \frac{\varepsilon' f^2 \ell_n^{1/3}}{C}\right)\\
&\le \mathbb{P}\left(L^0_{t_3} \ge -\frac{\varepsilon' f^2 \ell_n^{1/3}}{C}\right) + \mathbb{P}\left(Z(t_2) \ge \frac{\varepsilon' f^\varepsilon \ell_n^{1/3}}{C}\right),
\end{aligned} \tag{2.52}$$

we bound each one of the two terms of the right-hand side of (2.52) separately. First :

$$\mathbb{P}\left(Z(t_2) \geq \frac{\varepsilon' f^\varepsilon \ell_n^{1/3}}{C}\right) \leq \mathbb{P}\left(Z(t_1) \geq \frac{\varepsilon' f^\varepsilon \ell_n^{1/3}}{C}\right) + \mathbb{P}\left(Z(t_2) > Z_{t_1}\right).$$

Since $Z(t_2) > Z(t_1)$ occurs precisely if L visits 0 between t_1 and t_2 we already know by Equation (2.51) that :

$$\mathbb{P}(Z(t_2) > Z(t_1)) \leq A' \exp\left(\frac{-f^\varepsilon}{A'}\right). \tag{2.53}$$

By definition $Z(t_1) \geq r$ precisely if $L'_i < 1 - r$ for some $i \leq t_1$. By Corollary 23.1, for any $i \leq t_1$:

$$\mathbb{E}(L_i^h) \geq 0.$$

Using this inequality alongside Inequality (2.47) and Theorems 24 and 27 yields :

$$\begin{aligned}
&\mathbb{P}\left(Z(t_1) \geq \frac{\varepsilon' f^\varepsilon \ell_n^{1/3}}{C}\right) \\
&= \mathbb{P}\left(\inf_{i \leq t_1}(L'_i) \leq 1 - \frac{\varepsilon' f^\varepsilon \ell_n^{1/3}}{C}\right) \\
&\leq \mathbb{P}\left(\inf_{1 \leq j \leq t_1}(\tilde{L}_j^h) \leq -\frac{\varepsilon' f^\varepsilon \ell_n^{1/3}}{4C}\right) + \mathbb{P}\left(\inf_{1 \leq j \leq t_1}(L'_j - \tilde{L}_j) \leq -\frac{\varepsilon' f^\varepsilon \ell_n^{1/3}}{4C}\right) + \mathbb{P}\left(\sup_{1 \leq j \leq t_1} L_j \geq h\right) \\
&\leq A \exp\left(\frac{-f^\varepsilon}{A}\right).
\end{aligned} \tag{2.54}$$

By the union bound between Equations (2.53) and (2.54) we get :

$$\mathbb{P}\left(Z(t_2) \geq \frac{\varepsilon' f^\varepsilon \ell_n^{1/3}}{C}\right) \leq A \exp\left(\frac{-f^\varepsilon}{A}\right). \tag{2.55}$$

Furthermore, by Corollary 23.1 :

$$\mathbb{E}[L^0_{t_3}] \leq -\frac{2\varepsilon' f^2 \ell_n^{1/3}}{C}.$$

By this fact and Theorem 25 we obtain :

$$\begin{aligned}
\mathbb{P}\left(L^0(t_3) \geq -\frac{\varepsilon' f^2 \ell_n^{1/3}}{C}\right) &\leq \mathbb{P}\left(L^0(t_3) - \mathbb{E}[L^0_{t_3}] \geq \frac{\varepsilon' f^2 \ell_n^{1/3}}{C}\right) \\
&\leq A' \exp\left(\frac{-f}{A'}\right).
\end{aligned} \tag{2.56}$$

Injecting Inequalities (2.55) and (2.56) in Inequality (2.52) yields :

$$\mathbb{P}(L^0_{t_3} \geq -Z(t_2)) \leq A \exp\left(\frac{-f^\varepsilon}{A}\right),$$

and this finishes the proof. □

The following theorem gives a lower and upper bound on the total weight of H_f^*.

Theorem 29. *Suppose that Conditions 1 hold. Let* $1 > \varepsilon' > 0$. *For* f *large enough and for any* $1 \geq \varepsilon > 0$, *let* $t_1 = \frac{\ell_n^{2/3}}{f^{1-\varepsilon}C}$, $t_2 = \frac{2(1-\varepsilon')f\ell_n^{2/3}}{C}$ *and* $t_3 = \frac{2(1+\varepsilon')f\ell_n^{2/3}}{C}$.
There exists a constant $A > 0$ *such that the probability that the total weight of* H_f^* *is less than* $t_2 - t_1 - \varepsilon'(t_2 - t_1)$ *or more than* $t_3 + \varepsilon' t_3$ *is at most*

$$A \exp\left(\frac{-f^\varepsilon}{A}\right).$$

Proof. Let E be the event that L_i visists 0 for an $t_1 \leq i \leq t_2$ or L_i does not visit 0 for any $t_2 \leq i \leq t_3$. For n large enough, Theorem 28 states that there exists $A > 0$ such that :

$$\mathbb{P}(E) \leq A \exp\left(\frac{-f^\varepsilon}{A}\right).$$

If E does not hold, the total weight of H_f^* is larger than :

$$T = \sum_{i=t_1}^{t_2} w_{v(i)}.$$

By Lemma 11

$$\mathbb{E}[T] = (t_2 - t_1) + o(t_2 - t_1).$$

By Theorem 21, there exist positive constants A'', A''' such that :

$$\mathbb{P}\left[T \leq \mathbb{E}(T) - \varepsilon'(t_2 - t_1)\right] \leq A'' \exp\left(\frac{-\varepsilon' f \ell_n^{1/3}}{A''}\right),$$

hence by the union bound the total weight of H_f^* is less than $t_2 - t_1 - \varepsilon'(t_2 - t_1)$ with probability at most :

$$\mathbb{P}\left[T \leq (t_2 - t_1) - \varepsilon'(t_2 - t_1)\right] + \mathbb{P}(E) \leq A' \exp\left(\frac{-f^\varepsilon}{A'}\right),$$

where $A > 0$ is a large constant. Moreover when E does not hold the total weight of H_f^* is less than :

$$T' = \sum_{i=0}^{t_3} w_{v(i)}.$$

By the same arguments H_f^* is more than $t_3 + \varepsilon' t_3$ with probability at most :

$$\mathbb{P}\left[T' \geq t_3 + \varepsilon' t_3\right] + \mathbb{P}(E) \leq A' \exp\left(\frac{-f^\varepsilon}{A'}\right).$$

□

2.4.2 The excess of the giant component.

The previous theorems give us information about the size of H_f^*. We now turn to its surplus. Recall that the surplus (or excess) is the number of edges we need to remove from a connected graph in order to make it a tree. The excess of a general graph is the sum of excesses of its connected components.

Theorem 30. *Suppose that Conditions 1 hold. Let* Exc *be the excess of* H_f^*, *there exists a positive constant* $A > 0$ *such that :*

$$\mathbb{P}(Exc \geq Af^3) \leq A \exp\left(\frac{-f}{A}\right).$$

Proof. By construction, if a component is discovered between times t_1 and t_2 of the process, then its excess is precisely

$$\sum_{i=t_1}^{t_2} \sum_{j=i+1}^{L_i+i-1} Y(v(i), v(j)).$$

Let $m = \frac{3f\ell_n^{2/3}}{2C}$. By Theorem 26 :

$$\mathbb{P}\left(\sup_{1\leq i\leq m}(L_i) \geq \frac{10f^2\ell_n^{1/3}}{C}\right) \leq A'' \exp\left(\frac{-f}{A''}\right). \tag{2.57}$$

By Theorem 28, there exists a constant $A' > 0$ such that the probability that H_f^* has size more than m is at most :

$$A' \exp\left(\frac{-f}{A'}\right). \tag{2.58}$$

Let E be the event that H_f^* has size less than m and $L_i \leq \frac{10f^2\ell_n^{1/3}}{C}$ for all $1 \leq i \leq m$. By the union bound between Inequalities (2.57) and (2.58) we get :

$$\mathbb{P}(\bar{E}) \leq A'' \exp\left(\frac{-f}{A''}\right), \tag{2.59}$$

for some large constants $A'' > 0$. Let $R = \frac{10f^2\ell_n^{1/3}}{C}$ and :

$$U(R,i) = \sum_{j=i+1}^{L_{i-1}+i-1} Y(v(i), v(j)) + \sum_{j=L_{i-1}+i}^{R+i} Y'(v(i), v(j)),$$

with $Y'(i,j)$ being a Bernoulli random variable independent of everything else and having the same distribution as $Y(i,j)$ for $i \neq j$. We have thus by the union bound for any $l \geq 0$:

$$\begin{aligned}\mathbb{P}\left(\text{Exc} \geq l\right) &\leq \mathbb{P}\left(\left(\sum_{i=1}^{|H_f^*|} \sum_{j=i+1}^{L_{i-1}+i-1} Y(v(i), v(j)) \geq l\right), E\right) + \mathbb{P}\left(\bar{E}\right) \\ &\leq \mathbb{P}\left(\sum_{i=1}^{m} U(R,i) \geq l\right) + \mathbb{P}\left(\bar{E}\right)\end{aligned} \tag{2.60}$$

Conditionally on $\mathcal{F}_n$ the $U(R,i)$'s are sums of independent Bernoulli random variables. This is true because the first sum in the definition of $U(R,i)$ consists on independent Bernoulli random variables as stated in Fact 23. Moreover, for any $(i,j)_{1\leq i,\, 1+i\leq j\leq L_i+i-1}$ by Equation 2.6 :

$$\mathbb{E}[Y(v(i), v(j))|\mathcal{F}_n] \leq w_{v(i)} w_{v(j)} p_f,$$

and

$$\mathbb{E}[Y(v(i), v(j))^2|\mathcal{F}_n] \leq w_{v(i)} w_{v(j)} p_f.$$

The first inequality yields :

$$\mathbb{E}\left[\sum_{i=1}^{m} U(R,i) \middle| \mathcal{F}_n\right] \leq \sum_{i=1}^{m} \sum_{j=i+1}^{(R+i)} w_{v(i)} w_{v(j)} p_f.$$

Hence, by Bernstein's inequality :

$$\mathbb{P}\left(\sum_{i=1}^{m} U(R,i) \geq l + \sum_{i=1}^{m} \sum_{j=i+1}^{(R+i)} w_{v(i)} w_{v(j)} p_f \middle| \mathcal{F}_n\right) \leq \exp\left(\frac{-l^2}{2l + 2\sum_{i=1}^{m} \sum_{j=i+1}^{(R+i)} w_{v(i)} w_{v(j)} p_f}\right). \tag{2.61}$$

Denote by $J_1, J_2, .., J_n$ i.i.d. copies of $v(1)$. From Lemma 8, there exists a constant $A' > 0$ such that :

$$\mathbb{E}\left[p_f \sum_{k=0}^{\lceil \frac{m}{R} \rceil} 2R \left(\sum_{j=kR+1}^{(k+2)R} {w_{J_i}}^2 \right)\right] \leq A' m R p_f. \tag{2.62}$$

Moreover, by Cauchy-Schwarz's inequality :

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{m}\sum_{j=i+1}^{(R+i)} w_{v(i)}w_{v(j)}p_f \geq (A'+1)mRp_f\right) &\leq \mathbb{P}\left(p_f \sum_{k=0}^{\lceil \frac{m}{R} \rceil} \left(\sum_{i=kR+1}^{(k+2)R} w_{v(i)} \right)^2 \geq (A'+1)mRp_f\right) \\
&\leq \mathbb{P}\left(p_f \sum_{k=0}^{\lceil \frac{m}{R} \rceil} 2R \left(\sum_{i=kR+1}^{(k+2)R} w_{v(i)}^2 \right) \geq (A'+1)mRp_f\right).
\end{aligned}$$

Hence, by Theorem 6 applied on the $(w_{v(i)}^2)_{1\leq i\leq m}$'s and Inequality (2.62) we have the following Chernoff bound which yields a Bernstein's inequality (Bernstein [1924], Boucheron et al. [2013]) :

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{m}\sum_{j=i+1}^{(R+i)} w_{v(i)}w_{v(j)}p_f \geq (A'+1)mRp_f\right) &\leq \mathbb{P}\left(\sum_{k=0}^{\lceil \frac{m}{R} \rceil} 2 \left(\sum_{i=kR+1}^{(k+2)R} w_{v(i)}^2 \right) \geq (A'+1)m\right) \\
&\leq \mathbb{E}\left[\exp\left(\sum_{k=0}^{\lceil \frac{m}{R} \rceil} 2 \left(\sum_{i=kR+1}^{(k+2)R} w_{v(i)}^2 \right)\right) \exp\left(-(A'+1)m\right)\right] \\
&\leq \mathbb{E}\left[\exp\left(\sum_{k=0}^{\lceil \frac{m}{R} \rceil} 2 \left(\sum_{i=kR+1}^{(k+2)R} w_{J_i}^2 \right)\right) \exp\left(-(A'+1)m\right)\right] \\
&\leq A \exp\left(\frac{-m^2}{An^{2/3}m}\right) \\
&\leq A \exp\left(\frac{-f}{A}\right).
\end{aligned} \tag{2.63}$$

Here the penultimate inequality uses the fact that $\mathbb{E}[w_{v(1)}^4] \leq n^{2/3}\mathbb{E}[w_{v(1)}^2]$ and Lemma 8. We have that $mRp_f = Af^3$ for some $A > 0$. By Equations 2.59, 2.60, 2.61 and 2.63, the union bound yields :

$$\begin{aligned}
\mathbb{P}\left(Exc \geq (A'+2)mRp_f\right) &\leq \mathbb{P}\left(\sum_{i=1}^{m} U(R,i) \geq (A'+2)mRp_f\right) + \mathbb{P}[\bar{E}] \\
&\leq \exp\left(\frac{-(mRp_f)^2}{2mRp_f + 2(A'+1)mRp_f}\right) + A'' \exp\left(\frac{-f}{A''}\right) \\
&+ \mathbb{P}\left(\sum_{i=1}^{m}\sum_{j=i+1}^{(R+i)} w_{v(i)}w_{v(j)}p_f \geq (A'+1)mRp_f\right) \\
&\leq A''' \exp\left(\frac{-(mRp_f)^2}{A'''(mRp_f)}\right) + A''' \exp\left(\frac{-f}{A'''}\right) \\
&\leq A \exp\left(\frac{-f}{A}\right),
\end{aligned}$$

where $A > 0$ is a large enough constant. □

2.4.3 The excess of the components discovered before the largest connected component.

Theorem 31. *Suppose that Conditions 1 hold. Let* Exc_0 *be the total excess of the components discovered before the largest component. There exists* $A > 0$ *such that for any* $0 < \varepsilon \leq 1$ *:*

$$\mathbb{P}\left(\mathrm{Exc}_0 \geq A f^{\varepsilon}\right) \leq A \exp\left(\frac{-f^{\varepsilon/2}}{A}\right).$$

Proof. We know from Theorem 28 that for any $0 < \bar{\varepsilon} \leq 1$ the exploration of the largest component starts before time $m = \frac{\ell_n^{2/3}}{f^{1-\bar{\varepsilon}}C}$ with probability at least :

$$1 - A \exp\left(\frac{-f^{\bar{\varepsilon}}}{A}\right). \tag{2.64}$$

In that case the total excess of components discovered before the largest one is at most :

$$\sum_{i=0}^{m} \sum_{j=i+1}^{L_{i-1}+i-1} Y(v(i), v(j)).$$

By Corollary 23.1 and Conditions 1, for any $0 \leq i \leq j \leq m$:

$$\mathbb{E}(L^0(j) - L^0(i)) \leq \frac{f^{\bar{\varepsilon}} \ell_n^{1/3}}{C}.$$

By this fact and Theorem 25 applied on $(m, 0, 0, y)$, there exists an $A > 0$ such that :

$$\mathbb{P}\left(\sup_{0 \leq i \leq j \leq m} (L^0(j) - L^0(i)) \geq \frac{2 f^{\bar{\varepsilon}} \ell_n^{1/3}}{C}\right) \leq A \exp\left(\frac{-f^{\bar{\varepsilon}}}{A}\right),$$

Remark that, deterministically,

$$\sup_{0 \leq k \leq m} L(k) \leq \sup_{0 \leq i \leq j \leq m} (L'(j) - L'(i)) \leq \sup_{0 \leq i \leq j \leq m} (L^0(j) - L^0(i)),$$

hence :

$$\begin{aligned} \mathbb{P}\left(\sup_{0 \leq i \leq m} L_i \geq \frac{2 f^{\bar{\varepsilon}} \ell_n^{1/3}}{C}\right) &\leq \mathbb{P}\left(\sup_{0 \leq i \leq j \leq m} (L^0(j) - L^0(i)) \geq \frac{2 f^{\bar{\varepsilon}} \ell_n^{1/3}}{C}\right), \\ &\leq A \exp\left(\frac{-f^{\bar{\varepsilon}}}{A}\right). \end{aligned} \tag{2.65}$$

Let $R = \frac{2 f^{\bar{\varepsilon}} \ell_n^{1/3}}{C}$. Let E be the event $\{\max_{0 \leq i \leq m} L_i \leq R\}$ and the exploration of the largest component starts before time m. Recall the definition of $U(R, i)$ from Theorem 30. We have for any $l \geq 0$ by the union bound :

$$\mathbb{P}\left(\mathrm{Exc}_0 \geq l\right) \leq \mathbb{P}\left(\sum_{i=0}^{m} U(R, i) \geq l\right) + \mathbb{P}[\bar{E}]. \tag{2.66}$$

We use the same idea as in Theorem 30. By Bernstein's inequality (Bernstein [1924]) :

$$\mathbb{P}\left(\sum_{i=1}^{m} U(R, i) \geq l + \sum_{i=1}^{m} \sum_{j=i+1}^{(R+i)} w_{v(i)} w_{v(j)} p_f \,\middle|\, \mathcal{F}_n\right) \leq \exp\left(\frac{-l^2}{2l + 2 \sum_{i=1}^{m} \sum_{j=i+1}^{(R+i)} w_{v(i)} w_{v(j)} p_f}\right). \tag{2.67}$$

Denote by $J_1, J_2, .., J_n$ i.i.d. copies of $v(1)$. Similarly to Equation (2.62), there exists a constant $A' > 0$ such that :

$$\mathbb{E}\left[p_f \sum_{k=0}^{\lceil \frac{m}{R} \rceil} 2R \left(\sum_{j=kR+1}^{(k+2)R} {w_{J_i}}^2 \right) \right] \leq A' mRp_f. \tag{2.68}$$

And similarly to Equation (2.63) we have for any $\lambda \geq 0$:

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=1}^{m} \sum_{j=i+1}^{(R+i)} w_{v(i)} w_{v(j)} p_f \geq (A'+1) mRf^{\lambda\bar{\varepsilon}} p_f \right) \\
\leq &\mathbb{E}\left[\exp\left(\sum_{k=0}^{\lceil \frac{m}{R} \rceil} 2 \left(\sum_{i=kR+1}^{(k+2)R} w_{J_i}^2 \right) \right) \exp\left(-(A'+1) mf^{\lambda\bar{\varepsilon}} \right) \right] \\
\leq &A \exp\left(\frac{-m^2 f^{2\lambda\bar{\varepsilon}}}{A(m\ell_n^{2/3} f^{\lambda\bar{\varepsilon}} + m\ell_n^{2/3})} \right) \\
\leq &A'' \exp\left(\frac{-f^{(\lambda+1)\bar{\varepsilon}-1}}{A''} \right).
\end{aligned} \tag{2.69}$$

And also, Equations (2.64) and (2.65) yield :

$$\mathbb{P}(\bar{E}) \leq A' \exp\left(\frac{-f^{\bar{\varepsilon}}}{A'} \right). \tag{2.70}$$

By Equations 2.66, 2.67, 2.69, and 2.70 the union bound yields for $A'' > 0$ large enough :

$$\begin{aligned}
\mathbb{P}\left(Exc_0 \geq (A'+2) mRf^{\lambda\bar{\varepsilon}} p_f \right) &\leq \mathbb{P}\left(\sum_{i=1}^{m} U(R,i) \geq (A'+2) mRf^{\lambda\bar{\varepsilon}} p_f \right) + \mathbb{P}[\bar{E}] \\
&\leq \exp\left(\frac{-(mRf^{\lambda\bar{\varepsilon}} p_f)^2}{A'''(mRf^{\lambda\bar{\varepsilon}} p_f} \right) + A'' \exp\left(\frac{-f^{\bar{\varepsilon}}}{A''} \right) \\
&\quad + \mathbb{P}\left(\sum_{i=1}^{m} \sum_{j=i+1}^{(R+i)} w_{v(i)} w_{v(j)} p_f \geq (A'+1) mRf^{\lambda\bar{\varepsilon}} p_f \right) \\
&\leq A''' \exp\left(\frac{-(mRf^{\lambda\bar{\varepsilon}} p_f)^2}{A'''(mRf^{\lambda\bar{\varepsilon}} p_f)} \right) + A'' \exp\left(\frac{-f^{\bar{\varepsilon}}}{A''} \right) \\
&\quad + A''' \exp\left(\frac{-f^{(\lambda+1)\bar{\varepsilon}-1}}{A'''} \right),
\end{aligned}$$

where $A > 0$ is a large enough constant. Moreover, we have for n large enough :

$$mRf^{\lambda\bar{\varepsilon}} p_f \geq \frac{1}{C^2} f^{(2+\lambda)\bar{\varepsilon}-1}$$

for some large constant $A > 0$. Hence, if we take :

$$\lambda = \frac{2}{\varepsilon},$$

and :

$$\varepsilon = (2+\lambda)\bar{\varepsilon} - 1.$$

We obtain $\bar{\varepsilon} = \varepsilon/2$ and :

$$(1+\lambda)\bar{\varepsilon} - 1 = \frac{\varepsilon}{2}.$$

This proves the inequality of the theorem. □

2.5 The structure of the tail's components

2.5.1 Preliminaries

We call tail of the exploration process the part of it that starts after H_f^* is fully explored and ends at n. In order to get bounds on the size, weight and excess of the tail, we will use two main ideas. Firstly we use an appropriate division of the interval that start after the exploration of H_f^*, and ends in n. Secondly we make use of the fact that the further we go in the exploration the smaller the weights we discover. These two ideas are formalized below. The rest of the proofs uses similar techniques to the ones presented in Section 4, but with the added complexity of incorporating these two ideas.

For $i \geq 1$, write :

$$\bar{k}_i = i^2 f((i+1)^2 - i^2).$$

For $\bar{k}_i > k \geq 0$, and as long as $t_k^i < \ell_n^{11/12}$, write :

$$t_k^i = t + \frac{(i^2-1)f\ell_n^{2/3}}{C} + \frac{k\ell_n^{2/3}}{Ci^2 f},$$

with $t = \frac{2(1-\varepsilon')f\ell_n^{2/3}}{C}$ and where $1/2 > \varepsilon' > 0$ is fixed from here on. Moreover, let $(\tilde{i}, \tilde{k})$ be the first time when $t_{\tilde{k}}^{\tilde{i}} \geq \ell_n^{11/12}$. For any $k > \tilde{k}$ let :

$$t_k^{\tilde{i}} = t + \frac{(\tilde{i}^2-1)f\ell_n^{2/3}}{C} + \frac{k\ell_n^{2/3}}{C\tilde{i}^2 f}.$$

$(\tilde{i}, \tilde{k})$ depends implicitly on ε'. Moreover, by construction $\tilde{i}^2 f = o(n^{1/3})$. We are only interested in $t_k^i \leq n$, and for simplicity, since there is no real difficulty in dealing with the boundaries, we assume everything is well truncated.

This construction gives a division of the interval between t and n in the following way : Take intervals of the form $\left[t_0^i, t_0^{i+1}\right)$. Such intervals get larger and larger. Divide each one of them into small intervals of the form $\left[t_k^i, t_{k+1}^i\right)$ that get smaller with i. The main idea here is that the large intervals, those where i changes, represent phases of the exploration where we will find connected components that are of size at most the size of small intervals $\left[t_k^i, t_{k+1}^i\right)$. Moreover Conditions 3 will be verified inside the small intervals for good enough deviation values, which will allow us to use all our concentration theorems. We start by showing that the maximum weight gets smaller the further we explore the tail.

Lemma 32. *Suppose that Conditions 1 hold. There exists a constant $A > 0$ such that :*

For any $1 \leq i \leq \tilde{i}$, the probability of discovering a weight larger than $\frac{\ell_n^{1/3}}{i\sqrt{f}}$ in the BFW after time t_0^i is less than :

$$A \exp\left(\frac{-i\sqrt{f}}{A}\right).$$

Proof. Recall that $(T_i)_{i\leq n}$ is a sequence of independent exponential variables with rates $(w_i/\ell_n)_{i\leq n}$. And that for any $x > 0$:

$$N(x) = \sum_{k=1}^{n} \mathbb{1}(T_k \leq x),$$

Moreover, recall that by the properties of exponential random variables, the order statistic indices $(\tilde{v}(1), \tilde{v}(2), ...\tilde{v}(n))$ of the $(T_k)_{k\leq n}$ have the same distribution as $(v(1), v(2), ...v(n))$.

Let $x = t_0^i/2$, then by Lemma 12, Conditions 1 and obvious bounds :

$$\mathbb{P}(N(x) \geq t_0^i) \leq A \exp\left(\frac{-t_0^i}{A}\right). \tag{2.71}$$

This equation shows that at time x, the weights with indices $(\tilde{v}(t_0^i), \tilde{v}(t_0^i+1), ...\tilde{v}(n))$ will not be picked yet with high probability. Denote the event $\{N(x) \geq t_0^i\}$ by E. For any k such that $w_k \geq \frac{\ell_n^{1/3}}{i\sqrt{f}}$, we have :

$$\begin{aligned}\mathbb{P}(T_k \geq x,\, \bar{E}) &\leq \mathbb{P}(T_k \geq x) \\ &\leq A \exp\left(\frac{-i\sqrt{f}}{A}\right),\end{aligned}$$

this equation shows that a large weight has a large probability of being picked before time x.

Recall that by Conditions 1 :

$$\sum_{k=1}^{n} w_k^3 = (\mathbb{E}[W^3] + o(1))n.$$

Hence, the total number of weights larger than $\frac{\ell_n^{1/3}}{i\sqrt{f}}$ is less than $A' i^3 f^{3/2}$, where $A' > 0$ is a large enough constant.

This yields :

$$\begin{aligned}\mathbb{P}\left(\sup_{k \geq t_0^i}(w_{v(k)}) \geq \frac{\ell_n^{1/3}}{i\sqrt{f}}\right) &\leq \mathbb{P}(E) + \sum_{k=1}^{n} \mathbb{P}(T_k \geq x,\, \bar{E}) \mathbb{1}\left(w_k \geq \frac{\ell_n^{1/3}}{i\sqrt{f}}\right) \\ &\leq \exp\left(\frac{-t_0^i}{A}\right) + A' A i^3 f^{3/2} \exp\left(\frac{-i\sqrt{f}}{A}\right) \\ &\leq A'' \exp\left(\frac{-i\sqrt{f}}{A''}\right),\end{aligned} \tag{2.72}$$

whith $A'' > 0$ a large constant and f large enough. □

We now use the same notations as in the proof above. For $0 \leq i \leq \tilde{i}$. Let B be the event that no weight larger than $\frac{\ell_n^{1/3}}{i\sqrt{f}}$ is present after time t_0^i. Then for any $t_0^i \leq x$, with the notation of Section 2 when B holds we have :

$$X(x) - X(u) = \sum_{k=1}^{n} w_k \mathbb{1}(u \leq T_k \leq x) \mathbb{1}\left(w_k \leq \frac{\ell_n^{1/3}}{i\sqrt{f}}\right),$$

And :

$$N(x) - N(u) = \sum_{k=1}^{n} \mathbb{1}(u \leq T_k \leq x) \mathbb{1}\left(w_k \leq \frac{\ell_n^{1/3}}{i\sqrt{f}}\right).$$

Moreover, clearly :

$$\mathbb{E}\left[\sum_{k=1}^{n} w_k \mathbb{1}(u \leq T_k \leq x) \mathbb{1}\left(w_k \leq \frac{\ell_n^{1/3}}{i\sqrt{f}}\right)\right] \leq \mathbb{E}[X(x) - X(u)],$$

and

$$\mathbb{E}\left[\sum_{k=1}^{n} \mathbb{1}(u \leq T_k \leq x) \mathbb{1}\left(w_k \leq \frac{\ell_n^{1/3}}{i\sqrt{f}}\right)\right] \leq \mathbb{E}[N(x) - N(u)].$$

By those remarks, when B holds one can redo the proofs of Theorems 20 by only taking nodes with weights smaller than $\frac{\ell_n^{1/3}}{i\sqrt{f}}$. Then use the union bound with Lemma 32 to obtain the following theorem which is in the spirit of Theorem 20.

Theorem 33. *Suppose that Conditions 1 hold. There exists a constant $A > 0$ such that the following holds :*
If $(m, l, 0, y)$ verify Conditions 3, and there exists $i \leq \tilde{i}$ such that $l \geq t_0^i$, and $m \leq t_0^{\tilde{i}}$ then :

$$\mathbb{P}\left[\sup_{l\leq u\leq w\leq m} \sum_{k=u}^{w} w_{v(k)} - \mathbb{E}\left[\sum_{k=u}^{w} w_{v(k)}\right] \geq y\right] \leq A\exp\left(\frac{-y^2}{A\left(y\frac{\ell_n^{1/3}}{i\sqrt{f}} + m - l\right)}\right) + A\exp\left(\frac{-i\sqrt{f}}{A}\right).$$

Moreover we have by Bernstein's inequality :

$$\begin{aligned}
\mathbb{P}\left(\exists(h,j),\, j \geq t_0^i,\, X_j^0 \geq \frac{2\ell_n^{1/3}}{i\sqrt{f}}\right) &\leq \sum_{k=t_0^i}^{n} \mathbb{P}\left(d_{v(k)} \geq \frac{2\ell_n^{1/3}}{i\sqrt{f}}\right) \\
&\leq \mathbb{P}(\bar{B}) + \sum_{k=0}^{n} \mathbb{1}\left(w_k \leq \frac{\ell_n^{1/3}}{i\sqrt{f}}\right) \mathbb{P}\left(d_k \geq \frac{2\ell_n^{1/3}}{i\sqrt{f}}\right) \\
&\leq \mathbb{P}(\bar{B}) + A'\exp\left(\frac{-\ell_n^{1/3}}{A' i\sqrt{f}}\right) \\
&\leq A\exp\left(\frac{-i\sqrt{f}}{A}\right)
\end{aligned} \tag{2.73}$$

where A is a large constant. This shows that, similarly to what we did in Section 2, one can assume that L^0 and L have increments of size at most $\frac{2\ell_n^{1/3}}{i\sqrt{f}}$ after time t_0^i. Using this fact, one can redo the proofs of Theorems 25 after time t_0^i. Then use the union bound with Lemma 32 and Equation (2.73) to obtain the following theorem which is in the spirit of Theorem 25.

Theorem 34. *Suppose that Conditions 1 hold. There exists a constant $A > 0$ such that the following holds :*
Let (m, l, y) be such that $(m, l, 0, y)$ verifies Conditions 3, and there exists $i \leq \tilde{i}$ such that $l \geq t_0^i$, and $m \leq t_0^i$. We have :

$$\mathbb{P}\left(\sup_{l\leq u\leq w\leq m} L_w^0 - L_u^0 - \mathbb{E}[L_w^0 - L_u^0] \geq y\right) \leq A\exp\left(\frac{-y^2}{A\left(y\frac{\ell_n^{1/3}}{i\sqrt{f}} + m - l\right)}\right) + A\exp\left(\frac{-i\sqrt{f}}{A}\right).$$

We will also need the following lemma. It states that the weights get smaller in probability the further we go in the exploration. For $1 \leq k \leq n$ let $w_k^i = w_k$ if $w_k \leq \frac{\ell_n^{1/3}}{i\sqrt{f}}$ and $w_k^i = \frac{\ell_n^{1/3}}{i\sqrt{f}}$ otherwise.

Lemma 35. *Let $1 \leq u \leq w \leq n$, then for any $x \geq 0$ and $1 \leq i \leq \tilde{i}$:*

$$\mathbb{P}(w_{v(u)}^i \geq x) \leq \mathbb{P}(w_{v(w)}^i \geq x).$$

Proof. Recall that $\mathcal{V}_u = (v(1), v(2), ..., v(u))$ for any $n \geq i \geq 1$. It is sufficient to prove the lemma for $w = u + 1$. In that case we have :

$$\mathbb{P}(w_{v(u)}^i \geq x | \mathcal{V}_{u-1}) = \frac{\sum_{k\notin\mathcal{V}_{u-1}} w_k \mathbb{1}(w_k^i \geq x)}{\sum_{k'\notin\mathcal{V}_{u-1}} w_{k'}}.$$

Let :

$$U = \sum_{k\notin\mathcal{V}_{i-1}} w_k \mathbb{1}(w_k^i \geq x),$$

and

$$V = \sum_{k\notin\mathcal{V}_{i-1}} w_k.$$

Since $V \geq U$ we have :

$$\begin{aligned}
\mathbb{P}(w^i_{v(u+1)} \geq x|\mathcal{V}_{u-1}) &= \sum_{k\notin\mathcal{V}_{u-1}} \mathbb{P}(v(u)=k|\mathcal{V}_{u-1})\mathbb{P}(w^i_{v(u+1)} \geq x|\mathcal{V}_{i-1},\, v(i)=k) \\
&= \sum_{k\notin\mathcal{V}_{u-1}} \frac{w_k}{V}\left(\frac{U-w_k\mathbb{1}(w^i_k\geq x)}{V-w_k}\right) \\
&= \sum_{k\notin\mathcal{V}_{u-1},\, w^i_k\geq x} \frac{w_k}{V}\left(\frac{U-w_k}{V-w_k}\right) + \sum_{k\notin\mathcal{V}_{u-1},\, w^i_k<x} \frac{w_k}{V}\left(\frac{U}{V-w_k}\right) \\
&\leq \sum_{k\notin\mathcal{V}_{u-1},\, w^i_k\geq x} \frac{w_k}{V}\left(\frac{U-x}{V-x}\right) + \sum_{k\notin\mathcal{V}_{u-1},\, w^i_k<x} \frac{w_k}{V}\left(\frac{U}{V-x}\right) \\
&= \frac{U}{V}\left(\frac{U-x}{V-x}\right) + \left(\frac{V-U}{V}\right)\left(\frac{U}{V-x}\right) \\
&= \frac{U}{V} \\
&= \mathbb{P}(w_{v(u)} \geq x|\mathcal{V}_{i-1}).
\end{aligned}$$

□

With this lemma in hand we can deal with the case when $m > \tilde{t^i_0}$.

Theorem 36. *Suppose that Conditions 1 hold. There exists a constant $A > 0$ such that the following holds :*
For $\tilde{t^i_0} < u \leq w$ and for any $y \geq 0$:

$$\mathbb{P}\left[\sum_{k=u}^{w}(w_{v(k)}-1) \geq y\right] \leq A\exp\left(\frac{-y^2}{A\left(y\frac{\ell_n^{1/3}}{\tilde{i}\sqrt{f}}+w-u\right)}\right) + A\exp\left(\frac{-\tilde{i}\sqrt{f}}{A}\right)$$

Proof. Let $\mathcal{A}$ be the even that no weight discovered after time $\tilde{t^i_0}$ is larger than $\frac{\ell_n^{1/3}}{\tilde{i}\sqrt{f}}$. Let $(J(i))_{i\geq u}$ be i.i.d with the distribution of $v(u)$. Theorem 1 from Ben-Hamou et al. [2018] still applies for the $(\tilde{w^i_v})_{v\geq 1}$'s and we get similarly to Theorem 6 :

$$\begin{aligned}
\mathbb{P}\left(\sum_{k=u}^{w}(w_{v(k)}-1) \geq y,\, \mathcal{A}\right) &\leq \mathbb{P}\left(\sum_{k=u}^{w}(\tilde{w^i_{v(k)}}-1) \geq y\right) \\
&\leq \mathbb{E}\left[\exp\left(\sum_{k=u}^{w}(\tilde{w^i_{v(k)}}-1)\right)\exp(-y)\right] \qquad (2.74) \\
&\leq \mathbb{E}\left[\exp\left(\sum_{k=u}^{w}(\tilde{w^i_{J(k)}}-1)\right)\exp(-y)\right].
\end{aligned}$$

By Lemma 35 we can apply an ordered coupling argument (Theorem 7.1 of den Hollander [2012]) in order to obtain :

$$\mathbb{E}\left[\exp\left(\sum_{k=u}^{w}(\tilde{w^i_{J(k)}}-1)\right)\exp(-y)\right] \leq \mathbb{E}\left[\exp\left(\sum_{k=u}^{w}(\tilde{w^i_{J'(k)}}-1)\right)\exp(-y)\right] \qquad (2.75)$$

where the $J'(k)$'s are i.i.d random variables with the distribution of $v(\tilde{t^i_0}+1)$. Moreover by Lemma 11 we have for any $k \geq u$:

$$\mathbb{E}[\tilde{w^i_{J'(k)}}] \leq 1$$

and by Lemma 8 :

$$\mathbb{E}[w_{J'(k)}^{\tilde{i}}{}^2] \leq C(1+o(1))$$

Hence, by Equation 2.74 and the Chernoff bound in Equation 2.75 we obtain the following Bernstein's inequality :

$$\mathbb{P}\left(\sum_{k=u}^{w}(w_{v(k)}-1)\geq y,\,\mathcal{A}\right) \leq 2\exp\left(\frac{-y^2}{A\left(y\frac{\ell_n^{1/3}}{C\tilde{i}\sqrt{f}}+\sum_{k=u}^{w}\mathbb{E}\left[(w_{J'(k)}^{\tilde{i}})^2\right]\right)}\right)$$
$$\leq A\exp\left(\frac{-y^2}{A\left(y\frac{\ell_n^{1/3}}{\tilde{i}\sqrt{f}}+w-u\right)}\right).$$

Moreover, by Theorem 32 :

$$\mathbb{P}(\mathcal{A}) \leq A\exp\left(\frac{-\tilde{i}\sqrt{f}}{A}\right).$$

We finish the proof by union bound between these last two inequalities. □

By the same method, we obtain the following theorem which deals with the $w_{v(i)}^2$'s.

Theorem 37. *Suppose that Conditions 1 hold. There exists a constant $A>0$ such that the following holds :*
For $i<\tilde{i}$ and $t_0^i \leq u \leq w \leq t_0^{\tilde{i}}$ and for any $y\geq 0$:

$$\mathbb{P}\left[\sum_{k=u}^{w}(w_{v(k)}^2-\mathbb{E}[w_{v(u)}^2])\geq y\right] \leq A\exp\left(\frac{-y^2}{A\frac{\ell_n^{2/3}}{i^2 f}(y+w-u)}\right)+A\exp\left(\frac{-\tilde{i}\sqrt{f}}{A}\right).$$

And for $t_0^{\tilde{i}}<u\leq w$ and for any $y\geq 0$:

$$\mathbb{P}\left[\sum_{k=u}^{w}(w_{v(k)}^2-\mathbb{E}[w_{v(t_0^{\tilde{i}})}^2])\geq y\right] \leq A\exp\left(\frac{-y^2}{A\frac{\ell_n^{2/3}}{i^2 f}(y+w-u)}\right)+A\exp\left(\frac{-\tilde{i}\sqrt{f}}{A}\right).$$

2.5.2 The size of connected components discovered after the largest connected component

We can now prove the main theorem on the concentration of the sizes of the components discovered after H_f^*. In order to do that we will once again study the event that L visits 0 in some intervals.

Theorem 38. *Suppose that Conditions 1 are verified. Let $i^*\in\mathbb{N}$ be the time at which the exploration of H_f^* ends.There exists a constant $A>0$ such that the following is true :*
The probability that there exists an $\tilde{i}\geq i\geq 1$ and $\bar{k}_i>k\geq 0$, such that L does not visit 0 between times $t_k^i-t+i^*$ and $t_{k+1}^i-t+i^*$, or times $t_{\bar{k}_i}^i-t+i^*$ and $t_0^{i+1}-t+i^*$ is at most :

$$A\exp\left(\frac{-\sqrt{f}}{A}\right)+A\exp\left(\frac{-n^{1/8}}{A}\right).$$

Proof. By Theorem 28 :

$$\mathbb{P}\left(\frac{2(1+\varepsilon')f\ell_n^{2/3}}{C}\geq i^*\geq\frac{2(1-\varepsilon')f\ell_n^{2/3}}{C}\right)\geq 1-A\exp\left(\frac{-\sqrt{f}}{A}\right). \tag{2.76}$$

Define E_k^i as the event that L does not visit 0 between times $t_k^i - t + i^*$ and time $t_{k+1}^i - t + i^*$, or $t_{\bar{k}_i}^i - t + i^*$ and $t_0^{i+1} - t + i^*$ if $k = \bar{k}_i$.

Deterministically, for any $0 \leq u \leq w \leq n$:

$$\mathbb{P}\left(L'_w - L'_u \geq 0\right) \leq \mathbb{P}\left(L_w^0 - L_u^0 \geq 0\right), \tag{2.77}$$

so it is sufficient to focus on L^0.

We start by dealing with $(i,k) = (1,0)$, then the rest of the proof consists in repeating the arguments we will give for $(i,k) = (1,0)$ with an induction.

In order to show that L visits 0 between i^* and $i^* + \frac{\ell_n^{2/3}}{Cf}$, recall that $t = \frac{2(1-\varepsilon')f\ell_n^{2/3}}{C}$ and let E be the event $t + \frac{2\varepsilon' f\ell_n^{2/3}}{C} \geq i^* \geq t$. Then :

$$\begin{aligned}\mathbb{P}\left(L^0_{i^*+\frac{\ell_n^{2/3}}{Cf}} - L^0_{i^*} \geq 0\right) &= \mathbb{P}\left(E, \left\{L^0_{i^*+\frac{\ell_n^{2/3}}{Cf}} - L^0_{i^*} \geq 0\right\}\right) + \mathbb{P}(\bar{E}) \\ &\leq \mathbb{P}\left(\sup_{t\leq u\leq t+\frac{2\varepsilon' f\ell_n^{2/3}}{C}} L^0_{u+\frac{\ell_n^{2/3}}{Cf}} - L_u^0 \geq 0\right) + \mathbb{P}(\bar{E}).\end{aligned} \tag{2.78}$$

Divide the interval between t and $t + \frac{2\varepsilon' f\ell_n^{2/3}}{C}$ by introducing intermediate terms of the form : $t'_j = t + \frac{j\ell_n^{2/3}}{fC}$. Let $\bar{j}$ be the largest integer such that $t'_{\bar{j}} \leq t + \frac{2\varepsilon' f\ell_n^{2/3}}{C}$, and suppose everything is well truncated i.e $t'_{\bar{j}} = t + \frac{2\varepsilon' f\ell_n^{2/3}}{C}$. Equation (2.78) then yields :

$$\mathbb{P}\left(\sup_{t\leq u\leq t+\frac{2\varepsilon' f\ell_n^{2/3}}{C}} L^0_{u+\frac{\ell_n^{2/3}}{Cf}} - L_u^0 \geq 0\right) \leq \sum_{j=1}^{\bar{j}} \mathbb{P}\left(\sup_{t'_{j-1}\leq u\leq t'_j} L^0_{u+\frac{\ell_n^{2/3}}{Cf}} - L_u^0 \geq 0\right). \tag{2.79}$$

For $\bar{j} \geq j \geq 1$ let :

$$y_j = \frac{\ell_n^{1/3}(1-2\varepsilon')}{2C} + \frac{\ell_n^{1/3}(j-1)}{2f^2C}.$$

By Corollary 23.1 and straightforward calculations :

$$\begin{aligned}\sup_{t'_{j-1}\leq k\leq t'_j} \mathbb{E}\left[L^0_{k+\frac{\ell_n^{2/3}}{fC}} - L_k^0\right] &\leq \frac{3}{4}\mathbb{E}\left[L^0_{t'_j} - L^0_{t'_{j-1}}\right] \\ &\leq \frac{-3y_j}{2}.\end{aligned}$$

Moreover, for any $\bar{j} \geq j \geq 1$, $(t'_j, t'_{j-1}, 0, y_j)$ verify Conditions 3. Hence, by Theorem 34 and the fact that, by definition, $\bar{j} \leq 2f^2$:

$$\begin{aligned}\sum_{j=1}^{\bar{j}} \mathbb{P}\left(\sup_{t'_{j-1}\leq u\leq t'_j} L^0_{u+\frac{\ell_n^{2/3}}{Cf}} - L_k^0 \geq 0\right) &\leq \sum_{j=1}^{\bar{j}} A\exp\left(\frac{-y_j^2}{A\left(y_j\frac{\ell_n^{1/3}}{\sqrt{f}} + f^{-1}\ell_n^{2/3}\right)}\right) + A\exp\left(\frac{-\sqrt{f}}{A}\right). \\ &\leq A' f^2 \exp\left(\frac{-\sqrt{f}}{A'}\right) \\ &\leq A'' \exp\left(\frac{-\sqrt{f}}{A''}\right),\end{aligned} \tag{2.80}$$

we finish the initialization by injecting Inequalities (2.76) and (2.80) in (2.78).

We now move to the heredity property. Write

$$\mathcal{E}_{i,k} := \cup_{(u,v)\leq(i,k)} E_v^u \cup \bar{E}.$$

Suppose that the following inequality holds for (i,k) :

$$\mathbb{P}\left(\mathcal{E}_{i,k}\right) \leq A\exp\left(\frac{-\sqrt{f}}{A}\right) + A\sum_{j=0}^{i}(i+1)^2\exp\left(\frac{-i\sqrt{f}}{A}\right) + Ak\exp\left(\frac{-i\sqrt{f}}{A}\right), \tag{2.81}$$

where $A>0$ is a large enough constant that does not depend on (i,k).

For now suppose that $(i,k) \leq (\tilde{i},\tilde{k})$. we want to prove a similar inequality for $(i,k+1)$ if $k+1<\bar{k}_i$, or $(i+1,0)$ if not. Suppose we are in the case $k+1<\bar{k}_i$, the other case is similar. Write $t_0 = t_k^i$, $t_1 = t_{k+1}^i + \frac{2\varepsilon' f\ell_n^{2/3}}{C}$. By definition of $\mathcal{E}_{(i,k)}$:

$$\mathbb{P}\left(\mathcal{E}_{(i,k+1)}\right) \leq \mathbb{P}\left(\sup_{t_0\leq u\leq t_1}\left(L^0_{u+\frac{\ell_n^{2/3}}{Ci^2f}} - L_u^0\right)\geq 0\right) + \mathbb{P}\left(\mathcal{E}_{(i,k)}\right). \tag{2.82}$$

By using a similar division to the one used in Inequality (2.80) we get again :

$$\mathbb{P}\left(\sup_{t_0\leq u\leq t_1}\left(L^0_{u+\frac{\ell_n^{2/3}}{Ci^2f}} - L_u^0\right)\geq 0\right) \leq A\exp\left(\frac{-i\sqrt{f}}{A}\right).$$

This finishes the induction in the case where $(i,k)\leq(\tilde{i},\tilde{k})$.

Now suppose that $(i,k)>(\tilde{i},\tilde{k})$, we cannot directly use Theorem 34 because t_k^i might be of order n. Thus, we use will use the coupling argument of Theorem 36. Similarly to Equation (2.79) we need to bound :

$$\sum_{j=1}^{\bar{j}}\mathbb{P}\left(\sup_{t'_{j-1}\leq u\leq t'_j} L^0_{u+\frac{\ell_n^{2/3}}{Cf}} - L_u^0\geq 0\right),$$

with $t'_j = t_k^{\tilde{i}} + \frac{j\ell_n^{2/3}}{\tilde{i}^2fC}$ and $\bar{j}$ the largest integer such that $t'_{\bar{j}} \leq t_k^{\tilde{i}} + \frac{2\varepsilon' f\ell_n^{2/3}}{C}$. Let

$$y = \frac{\ell_n^{1/3}(\tilde{i}^2-1)}{8\tilde{i}^2C^2}.$$

We have for any $u > t_{\tilde{k}}^{\tilde{i}}$:

$$\begin{aligned}
\tilde{L}^0_{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2f}} - \tilde{L}_u^0 &= \sum_{r=u+1}^{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2f}}\sum_{r'>u}\left(1-\exp\left(-w_{v(r)}w_{v(r')}p_f\right)\right) - \frac{\ell_n^{2/3}}{C\tilde{i}^2f} \\
&\leq \left(\sum_{r=u}^{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2f}} w_{v(r)}\right)\left(1+\frac{f}{\ell_n^{1/3}} - \sum_{r'=1}^{u} w_{v(r')}p_f\right) - \frac{\ell_n^{2/3}}{C\tilde{i}^2f} \\
&\leq \left(\sum_{r=u}^{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2f}} w_{v(r)}\right)\left(1+\frac{f}{\ell_n^{1/3}} - \sum_{r'=1}^{t_{\tilde{k}}^{\tilde{i}}} w_{v(r')}p_f\right) - \frac{\ell_n^{2/3}}{C\tilde{i}^2f}
\end{aligned} \tag{2.83}$$

for $u\geq 0$ let $\mathcal{A}_1(u)$ be the event :

$$\left\{\sum_{r=u+1}^{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2f}} w_{v(r)} < \frac{\ell_n^{2/3}}{C\tilde{i}^2f} + y/2\right\} \cap \left\{\left(\frac{\sum_{r'=1}^{t_{\tilde{k}}^{\tilde{i}}} w_{v(r')}}{\ell_n}\right) > \frac{t_{\tilde{k}}^{\tilde{i}}}{2}\right\}.$$

Then, if $\mathcal{A}_1(u)$ holds, then Equation (2.83) yields :

$$\tilde{L}^0_{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2 f}} - \tilde{L}^0_u \leq -\frac{y}{2}.$$

Let also $\mathcal{A}_2(u)$ be the event :

$$\left\{ \sum_{r=u}^{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2 f}} (w^2_{v(r)}) \leq 8y\frac{\ell_n^{1/3}}{\tilde{i}\sqrt{f}} + 2\frac{\ell_n^{2/3}}{C\tilde{i}^2 f} \right\}.$$

Then by Bernstein's inequality for martingales (Freedman [1975]) :

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t'_{j-1}\leq u\leq t'_j} \left(L^0_{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2 f}} - L^0_u - (\tilde{L}^0_{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2 f}} - \tilde{L}^0_u) \right) \geq \frac{y}{2}, \cap_{t'_{j-1}\leq u\leq t'_j} \mathcal{A}_2(u) \right) \\
\leq &\mathbb{P}\left(\sup_{t'_{j-1}\leq u\leq t'_j} \left(L^0_{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2 f}} - L^0_{t'_{j-1}} - (\tilde{L}^0_{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2 f}} - \tilde{L}^0_{t'_{j-1}}) \right) \geq \frac{y}{4}, \cap_{t'_{j-1}\leq u\leq t'_j} \mathcal{A}_2(u) \right) \\
&+ \mathbb{P}\left(\sup_{t'_{j-1}\leq u\leq t'_j} \left(L^0_u - L^0_{t'_{j-1}} - (\tilde{L}^0_{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2 f}} - \tilde{L}^0_{t'_{j-1}}) \right) \leq \frac{-y}{4}, \cap_{t'_{j-1}\leq u\leq t'_j} \mathcal{A}_2(u) \right) \\
&\leq \exp\left(\frac{-\tilde{i}\sqrt{f}}{A} \right),
\end{aligned} \tag{2.84}$$

where the last inequality uses the fact that $y^2 = \Theta(\ell_n^{2/3})$.

By Theorem 36, for any $u > t^{\tilde{i}}_{\tilde{k}}$:

$$\begin{aligned}
\mathbb{P}\left(\sum_{r=u}^{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2 f}} (w_{v(r)} - 1) \geq y/2 \right) &\leq A\exp\left(\frac{-y^2}{A\left(y\frac{\ell_n^{1/3}}{\tilde{i}\sqrt{f}} + \frac{\ell_n^{2/3}}{C\tilde{i}^2 f} \right)} \right) + A\exp\left(\frac{-\tilde{i}\sqrt{f}}{A} \right) \\
&\leq A'\exp\left(\frac{-\tilde{i}\sqrt{f}}{A'} \right),
\end{aligned} \tag{2.85}$$

with $A' > 0$ a large constant. By Theorem 37 we also get :

$$\mathbb{P}\left(\bar{\mathcal{A}}_2(u) \right) \leq A'\exp\left(\frac{-\tilde{i}\sqrt{f}}{A'} \right). \tag{2.86}$$

By Theorem 21 and straightforward computations we obtain :

$$\mathbb{P}\left(\sum_{r'=1}^{t^{\tilde{i}}_{\tilde{k}}} w_{v(r')} \leq \frac{t^{\tilde{i}}_{\tilde{k}}}{2} \right) \leq A'\exp\left(\frac{-\tilde{i}\sqrt{f}}{A'} \right). \tag{2.87}$$

By the union bound between inequalities (2.84), (2.85), (2.86) and (2.87) we obtain

$$\begin{aligned}
\mathbb{P}\left(\sup_{t'_{j-1}\leq u\leq t'_j} \left(L^0_{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2 f}} - L^0_u \right) \geq 0 \right) &\leq \exp\left(\frac{-\tilde{i}\sqrt{f}}{A} \right) + \sum_{u=t'_{j-1}}^{t'_j} \mathbb{P}(\bar{\mathcal{A}}_1) + \sum_{u=t'_{j-1}}^{t'_j} \mathbb{P}(\bar{\mathcal{A}}_2) \\
&\leq A''(t'_j - t'_{j-1})\exp\left(\frac{-\tilde{i}\sqrt{f}}{A''} \right),
\end{aligned} \tag{2.88}$$

where $A'' > 0$ is a large constant. Since $t^{\tilde{i}}_{\tilde{k}} > \ell_n^{11/12}$:

$$\begin{aligned}\ell_n^{11/12} &\leq t + \frac{(\tilde{i}^2-1)f\ell_n^{2/3}}{C} + \frac{k\ell_n^{2/3}}{C\tilde{i}^2 f} \\ &\leq \frac{3\tilde{i}^2 f\ell_n^{2/3}}{C},\end{aligned}$$

equation 2.88 yields for n large enough :

$$\begin{aligned}\sum_{j=1}^{\bar{j}} \mathbb{P}\left(\sup_{t'_{j-1}\leq u\leq t'_j}\left(L^0_{u+\frac{\ell_n^{2/3}}{C\tilde{i}^2 f}} - L^0_u\right) \geq 0\right) &\leq A''(t'_{\bar{j}} - t'_0)\exp\left(\frac{-\tilde{i}\sqrt{f}}{A''}\right), \\ &\leq A\exp\left(\frac{-\tilde{i}\sqrt{f}}{A}\right),\end{aligned}$$

where $A > 0$ is a large constant. This finishes the proof of the induction of Equation (2.81). By that same equation we obtain for n and f large enough :

$$\begin{aligned}\mathbb{P}\left(\cup_{(u,v)\leq(\tilde{i},n)} E^u_v \cup \bar{E}\right) &\leq A\exp\left(\frac{-\sqrt{f}}{A}\right) + A\sum_{i=1}^{\tilde{i}}(i+1)^2\exp\left(\frac{-i\sqrt{f}}{A}\right) + An\exp\left(\frac{-\tilde{i}\sqrt{f}}{A}\right) \\ &\leq A\exp\left(\frac{-\sqrt{f}}{A}\right) + A\sum_{i=1}^{\infty}(i+1)^2\exp\left(\frac{-i\sqrt{f}}{A}\right) + A'n\exp\left(\frac{-n^{1/8}}{A'}\right) \\ &\leq A''\exp\left(\frac{-\sqrt{f}}{A''}\right) + A''\exp\left(\frac{-n^{1/8}}{A''}\right).\end{aligned}$$

□

This theorem shows that, after exploring the largest connected component, we discover small connected components that become smaller and smaller the further the exploration process goes. From that, one can get multiple corollaries. A first one is that the total weights of the components also gets smaller and smaller. The proof is the same as that of Theorem 29 and is omitted.

Corollary 38.1. *Suppose that Conditions 1 hold. There exists a constant $A > 0$ such that the following holds :*

For any $\varepsilon > 0$, the probability that there exists an $i \geq 0$ and $\bar{k}_i \geq k \geq 0$, such that a connected component discovered between times $t^i_k - t + i^*$ and $t^i_{k+1} - t + i^*$ (or times $t^i_{\bar{k}_i} - t + i^*$ and $t^{i+1}_0 - t + i^*$) in the exploration process has total weight larger than $(1+\varepsilon)(t^i_{k+1} - t^i_k)$ (or $(1+\varepsilon)(t^{i+1}_0 - t^i_{\bar{k}_i})$), where $i^* \in \mathbb{N}$ is the time when the exploration of H^*_f ends, is at most :

$$A\exp\left(\frac{-\sqrt{f}}{A}\right) + A\exp\left(\frac{-n^{1/8}}{A}\right).$$

Another fact we can deduce from Theorem 38 is the following convergence in probability. Its proof is straightforward from Theorems 28 and 38.

Corollary 38.2. *Recall that $f = f(n)$ is such that $f(n) = o(n^{1/3})$. Suppose that $\lim_{n\to\infty} f(n) = +\infty$. Let $(|C_1|, |C_2|, |C_3|, ...)$ denote the sequence of sizes of the connected components of $G(n, \boldsymbol{W}, p_{f(n)})$ taken in decreasing order, with the convention $|C_i| = 0$ if there is no i-th largest component. We have the following convergence in probability for any $p > 7/3$ as $n \to \infty$:*

$$\left(\frac{|C_1|}{2f(n)\ell_n^{2/3}}, \frac{|C_2|}{\ell_n^{2/3}}, \frac{|C_3|}{\ell_n^{2/3}}, \frac{|C_4|}{\ell_n^{2/3}}, ...\right) \xrightarrow{\mathbb{P}} (C, 0, 0, ..),$$

in ℓ^p, the usual p norm.

Proof. By Theorem 28, for any $1 > \varepsilon' > 0$ there exists a constant $A > 0$ such that for n large enough :

$$\mathbb{P}\left(\left|\left(\frac{|C_1|}{2f(n)\ell_n^{2/3}} - C\right)^p\right| \geq (3\varepsilon')^p\right) \leq A\exp\left(\frac{-f(n)^{1/2}}{A}\right). \tag{2.89}$$

Recall the definition of $\tilde{i}$ and let

$$\varepsilon(f(n)) = \frac{1}{\sqrt{f(n)}^p} + \frac{1}{(Cf(n))^{p-1}}\sum_{i=1}^{\tilde{i}-1}\frac{1}{i^{2p-3}} + \frac{\ell_n^{1/3}}{(\tilde{i}^2 f)^{p-1}}.$$

We know that for any $(x_1, x_2, ...x_k)$ which are positive numbers :

$$\sum_{u=1}^{k} x_u^p \leq \left(\sum_{u=1}^{k} x_u\right)^p.$$

We showed in the end of the proof of Theorem 38 that $\ell_n^{1/4} = O(\tilde{i}^2 f)$, this yields $\lim_n \varepsilon(f(n)) = 0$. Using those remarks alongside Theorems 38 and 28, there exists a constant $A > 0$ such that :

$$\mathbb{P}\left(\sum_{k\geq 2}\left|\left(\frac{|C_k|}{\ell_n^{2/3}}\right)^p\right| \geq A\varepsilon(f(n))\right) \leq A\exp\left(\frac{-\sqrt{f(n)}}{A}\right) + A\exp\left(\frac{-n^{1/8}}{A}\right). \tag{2.90}$$

The corollary follows by the union bound Inequalities (2.89) and (2.90) □

Note : If we change $\ell_n^{11/12}$ to $\ell_n^{1-\varepsilon''}$ for $1/3 > \varepsilon'' > 0$ arbitrarily small in the definition of $t_{\tilde{k}}^{\tilde{i}}$ then Theorem 38 will hold with the term $n^{1/8}$ being replaced by $n^{\frac{-1+3\varepsilon''}{6}}$. And this shows that Corollary 38.2 holds in fact for any $p > 2$. Moreover, with the same technique one can also obtain the same convergence for the sequence of weights of the connected components of $G(\mathbf{W}, p_{f(n)})$. It is also easy to show that if $f(n)$ is of order n^{ε} for some $\varepsilon > 0$ then this convergence will hold in expectation for any moment larger than 1.

2.5.3 The excess of the tail

We showed that after discovering the giant component all the other components have size less than $\ell_n^{2/3}/f$ with high probability. We call excess of a discrete interval between 1 and n, the number of excess edges discovered in that interval of time during the exploration process, regardless of which connected component they belong to. In the following theorem we will first focus on getting bounds on the excess of small intervals, then getting bounds on the excess of the tail will be straightforward by using Theorem 38.

Theorem 39. *Suppose that Conditions 1 hold. There exists a constant $A > 0$ such that the following is true :*

For $\tilde{i} \geq i \geq 1$, for $\bar{k}_i \geq k \geq 0$ let Exc_k^i be the excess of the interval $[t_k^i, t_{k+1}^i)$. For any $\varepsilon > 0$:

$$\mathbb{P}\left(\sup_{k_i > k \geq 0}(\mathrm{Exc}_k^i) \geq f^{\varepsilon}\right) \leq A\exp\left(\frac{-f^{\varepsilon}\ln(i\sqrt{f})}{A}\right) + A\exp\left(\frac{-i\sqrt{f}}{A}\right) + A\exp\left(\frac{-\sqrt{f}}{A}\right) + A\exp\left(\frac{-n^{1/8}}{A}\right).$$

Proof. Let $k < k_i$. If $t_k^i \leq \ell_n^{11/12}$, by Theorem 34 :

$$\mathbb{P}\left(\sup_{t_{k-1}^i \leq u \leq w \leq t_{k+1}^i}(L_w^0 - L_u^0 - \mathbb{E}[L_w^0 - L_u^0]) \geq \ell_n^{1/3}\right) \leq A\exp\left(\frac{-i\sqrt{f}}{A'}\right). \tag{2.91}$$

By Corollary 23.1, for any $t^i_{k-1} \leq u \leq w \leq t^i_{k+1}$:

$$\mathbb{E}[L^0_w - L^0_u] \leq 0.$$

With the above inequality, Equation 2.91 yields :

$$\mathbb{P}\left(\sup_{t^i_{k-1}\leq u\leq w\leq t^i_{k+1}} (L^0_w - L^0_u) \geq \ell_n^{1/3}\right) \leq A\exp\left(\frac{-i\sqrt{f}}{A}\right). \tag{2.92}$$

And in fact, notice that this inequality also holds for $t^i_k > \ell_n^{11/12}$ by the method used to obtain Inequality (2.88). Denote the event "no connected component discovered after time t^i_0 has size larger $\frac{\ell_n^{2/3}}{i^2 fC}$" by $\mathcal{G}$. When $\mathcal{G}$ holds, L visits 0 in any interval of size $\frac{\ell_n^{2/3}}{i^2 fC}$ after t^i_0. In that case :

$$\sup_{t^i_k\leq r\leq t^i_{k+1}} L(r) \leq \sup_{t^i_{k-1}\leq u\leq w\leq t^i_{k+1}} (L^0_w - L^0_u).$$

This fact and Equation (2.92) yield :

$$\mathbb{P}\left(\sup_{t^i_k\leq r\leq t^i_{k+1}} L_r \geq \ell_n^{1/3}\right) \leq A'\exp\left(\frac{-i\sqrt{f}}{A'}\right) + \mathbb{P}(\bar{\mathcal{G}}). \tag{2.93}$$

Let $\mathcal{M} = \left\{\sup_{t^i_k\leq r\leq t^i_{k+1}} L_r \leq \ell_n^{1/3}\right\}$. By Equation (2.93) and Theorem 38 we obtain :

$$\mathbb{P}\left(\bar{\mathcal{M}}\right) \leq A\exp\left(\frac{-\sqrt{f}}{A}\right) + A\exp\left(\frac{-n^{1/8}}{A}\right) + A'\exp\left(\frac{-i\sqrt{f}}{A'}\right). \tag{2.94}$$

By the union bound :

$$\mathbb{P}\left(\mathrm{Exc}^i_k \geq l + \mathbb{E}[\mathrm{Exc}^i_k]\right) \leq \mathbb{P}\left(\mathrm{Exc}^i_k \geq l + \mathbb{E}[\mathrm{Exc}^i_k],\, \mathcal{M}\right) + \mathbb{P}(\bar{\mathcal{M}}). \tag{2.95}$$

Now we use the same method we used in Lemma 30. Let $R = \ell_n^{1/3}$ and define $\tilde{t} = t^i_{k+1} - t^i_k$. By Lemma 8 :

$$\mathbb{E}\left[p_f \sum_{r=\frac{t^i_{k-1}}{R}}^{\frac{t^i_k}{R}} 2R\left(\sum_{u=rR+1}^{(r+2)R} {w_{v(t^i_{k-1})}}^2\right)\right] \leq A\tilde{t}Rp_f, \tag{2.96}$$

Hence, by Equation (2.96) and Theorem 37 :

$$\begin{aligned}
&\mathbb{P}\left(\sum_{r=t^i_{k-1}}^{t^i_k}\sum_{u=r+1}^{(R+r)} w_{v(u)}w_{v(r)}p_f \geq 2A\tilde{t}Rp_f + \frac{1}{i\sqrt{f}}\right)\\
&\leq \mathbb{P}\left(p_f \sum_{r=\frac{t^i_{k-1}}{R}}^{\frac{t^i_k}{R}}\left(\sum_{u=rR+1}^{(r+2)R} w_{v(u)}\right)^2 \geq 2A\tilde{t}Rp_f + \frac{1}{i\sqrt{f}}\right)\\
&\leq \mathbb{P}\left(\sum_{r=\frac{t^i_{k-1}}{R}}^{\frac{t^i_k}{R}}\left(\sum_{u=rR+1}^{(r+2)R} w^2_{v(u)}\right) \geq A\tilde{t} + \frac{1}{2i\sqrt{f}Rp_f}\right)\\
&\leq A''\exp\left(\frac{-i\sqrt{f}}{A''}\right).
\end{aligned} \tag{2.97}$$

By the union bound between Equation (2.96) and Equation (2.65) :

$$\begin{aligned}
&\mathbb{P}\left(\sum_{r=t_{k-1}^{i}}^{t_k^i}\sum_{u=r+1}^{(L_r+i)} w_{v(r)}w_{v(u)}p_f \geq 2A\tilde{t}Rp_f + \frac{1}{i\sqrt{f}}\right)\\
&\leq \mathbb{P}\left(\sum_{r=t_{k-1}^{i}}^{t_k^i}\sum_{u=r+1}^{(R+i)} w_{v(r)}w_{v(u)}p_f \geq 2A\tilde{t}Rp_f + \frac{1}{i\sqrt{f}}\right) + \mathbb{P}(\bar{\mathcal{M}})\\
&\leq A\exp\left(\frac{-\sqrt{f}}{A}\right) + A\exp\left(\frac{-n^{1/8}}{A}\right) + A'\exp\left(\frac{-i\sqrt{f}}{A'}\right).
\end{aligned} \tag{2.98}$$

We know that for any $\varepsilon > 0$:

$$\mathbb{P}(\mathrm{Exc}_k^i \geq f^\varepsilon | \mathcal{F}_n) \leq \mathbb{P}\left(\sum_{r=t_{k-1}^{i}}^{t_k^i}\sum_{u=r+1}^{(L_r+r-1)} Y(v(r),v(u)) \geq f^\varepsilon \middle| \mathcal{F}_n\right)\mathbb{1}(\mathcal{M}) + \mathbb{1}(\bar{\mathcal{M}}). \tag{2.99}$$

Since we are dealing with a sum of Bernoulli random variables, this sum is larger than f^ε if and only if there are more than f^ε Bernoulli variables equal to 1. Let S be the random set of subsets of size f^ε (suppose that f^ε is an integer for simplicity) composed of couples (r,u) that appear as indices in the sum in the right-hand side of Equation (2.99), and let S' be the deterministic set of subsets of size f^ε composed of couples (r,u) that appear as indices in the sum in the right-hand side of Equation (2.99) when we replace L_u by R for all $t_{k-1}^i \leq r \leq t_k^i$. Then for f large enough :

$$\begin{aligned}
\mathbb{P}\left(\sum_{r=t_{k-1}^{i}}^{t_k^i}\sum_{u=r+1}^{(L_u+r-1)} Y(v(r),v(u)) \geq f^\varepsilon \middle| \mathcal{F}_n\right)\mathbb{1}(\mathcal{M}) &= \mathbb{P}\left(\bigcup_{M\in S}\bigcap_{(r,u)\in M}\{Y(v(r),v(u))=1\}\middle|\mathcal{F}_n\right)\mathbb{1}(\mathcal{M})\\
&\leq \sum_{M\in S'}\prod_{(r,u)\in M}\left(1-e^{-w_{v(r)}w_{v(u)}p_f}\right)\\
&\leq \sum_{M\in S'}\prod_{(r,u)\in M}\left(w_{v(r)}w_{v(u)}p_f\right)\\
&\leq \left(\sum_{r=t_{k-1}}^{t_k}\sum_{u=r+1}^{(R+r)} w_{v(r)}w_{v(u)}p_f\right)^{f^\varepsilon+1}.
\end{aligned}$$

By this fact and Equation (2.98) :

$$\begin{aligned}
\mathbb{P}(\mathrm{Exc}_k^i \geq f^\varepsilon) &\leq \left(A\tilde{t}Rp_f + \frac{1}{i\sqrt{f}}\right)^{f^\varepsilon+1} + A'\exp\left(\frac{-i\sqrt{f}}{A'}\right) + A\exp\left(\frac{-n^{1/8}}{A}\right) + A'\exp\left(\frac{-\sqrt{f}}{A'}\right)\\
&\leq \left(\frac{A''}{i^2 f} + \frac{1}{i\sqrt{f}}\right)^{f^\varepsilon+1} + A'\exp\left(\frac{-i\sqrt{f}}{A'}\right) + A\exp\left(\frac{-n^{1/8}}{A}\right) + A'\exp\left(\frac{-\sqrt{f}}{A'}\right)\\
&\leq \exp\left((f^\varepsilon+1)\left(\ln\left(\frac{A''}{i^2 f}\right) + \ln\left(1+\frac{i\sqrt{f}}{A''}\right)\right)\right) + A'\exp\left(\frac{-i\sqrt{f}}{A'}\right)\\
&\quad + A\exp\left(\frac{-n^{1/8}}{A}\right) + A'\exp\left(\frac{-\sqrt{f}}{A'}\right)\\
&\leq \exp\left(\frac{-f^\varepsilon \ln(i\sqrt{f})}{A'''}\right) + A'\exp\left(\frac{-i\sqrt{f}}{A'}\right) + A\exp\left(\frac{-n^{1/8}}{A}\right) + A'\exp\left(\frac{-\sqrt{f}}{A'}\right).
\end{aligned} \tag{2.100}$$

If $t_k^i \geq \ell_n^{11/12}$, then by definition $i = \tilde{i}$. And we obtain similarly :

$$\mathbb{P}(\mathrm{Exc}_k^{\tilde{i}} \geq f^\varepsilon) \leq A \exp\left(\frac{-f^\varepsilon \ln(\tilde{i}\sqrt{f})}{A}\right) + A \exp\left(\frac{-\tilde{i}\sqrt{f}}{A}\right) + A \exp\left(\frac{-\sqrt{f}}{A}\right) + A \exp\left(\frac{-n^{1/8}}{A}\right).$$

This finishes the proof. □

In Theorem 39 the term $A \exp\left(\frac{-\sqrt{f}}{A}\right)$ comes from applying Theorem 38, and that theorem gives a bound for all the connected components discovered after the giant connected component. Using this remark, we can sum over i. And using simple computations, we obtain the concentration of the total surplus of the tail.

Theorem 40. *Suppose that Conditions 1 hold. There exists $A > 0$, such that for any $\varepsilon > 0$, for f and n large enough, the probability that a connected component discovered after H_f^* has excess more than f^ε is at most :*

$$A \exp\left(\frac{-f^\varepsilon \ln(\sqrt{f})}{A}\right) + A \exp\left(\frac{-\sqrt{f}}{A}\right) + A \exp\left(\frac{-n^{1/8}}{A}\right).$$

As a Corollary of the work done here we obtain a natural global upper bound on L.

Corollary 40.1. *Suppose that Conditions 1 hold. There exists a constant $A > 0$ large enough, such that :*

$$\mathbb{P}\left(\sup_{t_0^1 \leq l \leq n} (L_l) \geq \ell_n^{1/3}\right) \leq A \exp\left(\frac{-\sqrt{f}}{A}\right) + A \exp\left(\frac{-n^{1/8}}{A}\right).$$

Proof. Let $1 \leq i \leq \tilde{i}$, and denote the event "no connected component discovered after time t_0^i has size larger $\frac{\ell_n^{2/3}}{i^2 fC}$" by G_i. when G_i holds, L visits 0 in any interval of size $\frac{\ell_n^{2/3}}{i^2 fC}$ after t_0^i. In that case :

$$\sup_{t_k^i \leq r \leq t_{k+1}^i} L_r \leq \sup_{t_{k-1}^i \leq u \leq w \leq t_{k+1}^i} (L_w^0 - L_u^0).$$

Moreover, by Equation (2.92) :

$$\mathbb{P}\left(\sup_{t_{k-1}^i \leq u \leq w \leq t_{k+1}^i} (L_w^0 - L_u^0) \geq \ell_n^{1/3}\right) \leq A \exp\left(\frac{-i\sqrt{f}}{A}\right),$$

with $A > 0$ a large constant independent of i. By summing this equation over $1 \leq k < \bar{k}_i - 1$ for every i, and then over $1 \leq i \leq \tilde{i}$ we obtain directly :

$$\mathbb{P}\left(\sup_{t_0^1 \leq r \leq n} (L_r) \geq \ell_n^{1/3}, \cap_{i \leq \tilde{i}} G_i\right) \leq A' \exp\left(\frac{-\sqrt{f}}{A'}\right) + A \exp\left(\frac{-n^{1/8}}{A}\right). \tag{2.101}$$

With $A' > 0$ a large constant. By Theorem 38 there exists a large constant $A > 0$ such that :

$$\mathbb{P}(\cup_{i \leq \tilde{i}} \bar{G}_i) \leq A \exp\left(\frac{-\sqrt{f}}{A}\right) + A \exp\left(\frac{-n^{1/8}}{A}\right). \tag{2.102}$$

By Equations (2.101) and (2.102) there exists a large constant $A > 0$ such that :

$$\begin{aligned}\mathbb{P}\left(\sup_{t_0^1 \leq r \leq n} (L_r) \geq \ell_n^{1/3}\right) &\leq \mathbb{P}(\cup_{i \leq \tilde{i}} G_i) + \mathbb{P}\left(\sup_{t_0^1 \leq r \leq n} (L_r) \geq \ell_n^{1/3}, \cap_{i \leq \tilde{i}} G_i\right) \\ &\leq A \exp\left(\frac{-\sqrt{f}}{A}\right) + A \exp\left(\frac{-n^{1/8}}{A}\right),\end{aligned}$$

which finishes the proof. □

This upper bound alongside Theorem 26 gives an upper bound for the whole process L. However, it can be refined, and it is not hard to show that L gets smaller the further we advance in the exploration. We elect to stop here and as a last result we use this upper bound on L and the theorems we showed in this chapter to give an upper bound on the number of connected components discovered in parts of the exploration of the graph.

Corollary 40.2. *Suppose that Conditions 1 hold. Recall that $i^* \in \mathbb{N}$ is the time at which the exploration of H_f^* ends. There exists a constant $A > 0$ such that the following is true :*
The probability that there exists an $\tilde{i} > i \geq 0$ and $\bar{k}_i > k \geq 0$, such that the number of connected components discovered between times $t_0^i - t + i^*$ and time $t_{\bar{k}}^i - t + i^*$, is more than $100 i^3 f^2 \ell_n^{1/3}$, is at most :

$$A \exp\left(\frac{-\sqrt{f}}{A}\right) + A \exp\left(\frac{-n^{1/8}}{A}\right).$$

Proof. Let $r = \frac{2f\ell_n^{2/3}}{C}$, $t_1 = t_0^i$, and $t_2 = t_{\bar{k}_i}^i + r$.
In order to prove this theorem we need to bound the number of times a new minima of L' is reached in the interval $[t_1, t_2]$. Since L' can only go down by 1, the number of new minimums created in the interval $[t_1, t_2]$ is smaller than

$$\inf_{t_1 \leq l \leq m \leq t_2} L'_m - L'_l.$$

Choose $x = -50 i^3 f^2 \ell_n^{1/3}$ and. Then $(t_2, t_1, 0, x)$ verifies Conditions 3. Hence, by Theorem 27 we have :

$$\begin{aligned} \mathbb{P}\left(\sup_{t_1 \leq u \leq w \leq t_2} \left|L'_w - L'_u - \tilde{L}_w - \tilde{L}_u\right| \geq -x\right) \leq & A \exp\left(\frac{-x^2}{A(xn^{1/3} + (t_2 - t_1))}\right) \\ \leq & A' \exp\left(\frac{-i^3 f^2}{A'}\right). \end{aligned} \tag{2.103}$$

For any $h > 0$, if $L_k < h$ for any $k \leq t_2$ then deterministically $\tilde{L}_m - \tilde{L}_l \geq \tilde{L}_m^h - \tilde{L}_l^h$ for any $1 \leq l \leq m \leq t_2$.

Hence, if $\tilde{L}_m - \tilde{L}_l \leq x$ for some $t_1 \leq l \leq m \leq t_2$ then one of the following events happens :

— There exists $0 \leq j \leq t_2$ such that $L_j \geq h$.

— There exists $t_1 \leq l \leq m \leq t_2$ such that $\tilde{L}_m^h - \tilde{L}_l^h \leq L'_m - L'_l \leq x$.

Let $h = \frac{10 f^2 \ell_n^{1/3}}{C}$. Then for the first event, by Theorem 26 and Corollary 40.1 :

$$\mathbb{P}\left(\sup_{1 \leq j \leq t_2} L_j \geq h\right) \leq A \exp\left(\frac{-\sqrt{f}}{A}\right) + A \exp\left(\frac{-n^{1/8}}{A}\right). \tag{2.104}$$

For the second event, Conditions 3 are verified for $(t_2, t_1, h, -x)$. By Corollary 23.1 and a quick computation, for any $t_2 \geq w \geq u \geq t_1$, we have

$$-\mathbb{E}[\tilde{L}_w^h - \tilde{L}_u^h] \leq -\mathbb{E}[\tilde{L}_{t_2}^h - \tilde{L}_{t_1}^h] \leq x/2.$$

We can thus apply Lemma 24 to obtain :

$$\begin{aligned} \mathbb{P}\left(\inf_{t_1 \leq u \leq w \leq t_2} \tilde{L}_w^h - \tilde{L}_u^h \leq x\right) \leq & \mathbb{P}\left(\inf_{t_1 \leq u \leq w \leq t_2} \tilde{L}_w^h - \tilde{L}_u^h - \mathbb{E}[\tilde{L}_w^h - \tilde{L}_u^h] \leq \frac{x}{2}\right) \\ \leq & A \exp\left(\frac{-x^2}{A(xn^{1/3} + (t_2 - t_1))}\right) \\ \leq & A' \exp\left(\frac{-i^3 f^2}{A'}\right), \end{aligned} \tag{2.105}$$

with $A' > 0$ a large constant that does not depend on i.

Recall that $i^* \in \mathbb{N}$ is the time at which the exploration of H_f^* ends. By Theorem 28 :

$$\mathbb{P}\left(\frac{3f\ell_n^{2/3}}{C} \geq i^* \geq \frac{f\ell_n^{2/3}}{C}\right) \geq 1 - A\exp\left(\frac{-\sqrt{f}}{A}\right). \tag{2.106}$$

When this event holds, we have $[t_0^i - t + i^*, t_{k_i}^i - t + i^*] \subset [t_1, t_2]$. Hence, summing Equations (2.105) and (2.103) for $\tilde{i} > i \geq 1$, and using the union bound with Equation (2.104) and (2.106) finishes the proof. □

Chapitre 3

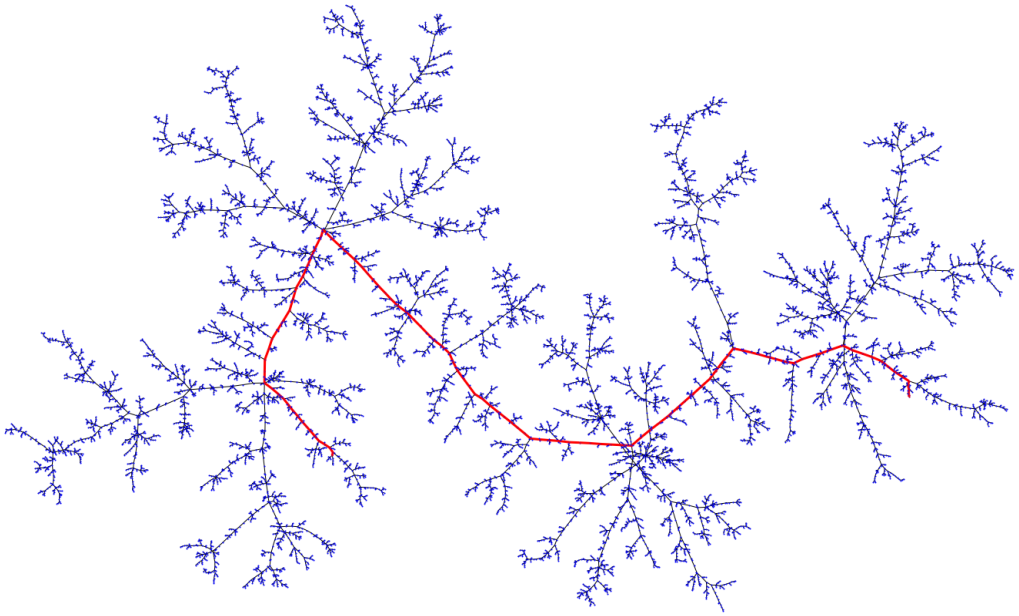
Diameter of inhomogeneous minimum spanning tree

Contents

3.1 Introduction . . . 111
3.1.1 The model . . . 111
3.1.2 Notations and definition of the exploration process . . . 112
3.1.3 Further discussion and related work in statistical physics . . . 115
3.2 First ingredients of the proof . . . 118
3.2.1 The phase transition . . . 118
3.2.2 The supercritical phase . . . 118
3.2.3 Known results . . . 120
3.3 Dealing with the small components . . . 121
3.3.1 Construction of the growth of small components . . . 122
3.3.2 Coupling with Galton-Watson trees . . . 126
3.4 Bounding the length of the longest path of the giant component . . . 132
3.4.1 A new pruning procedure for the giant component . . . 132
3.4.2 Bounding the longest path . . . 133
3.5 Proofs of Theorems 43, 45, 48 and 49 . . . 137

Diameter and typical distances of inhomogeneous minimum spanning tree

We study a new type of random minimum spanning trees. It is built on the complete graph where each vertex is given a weight, which is a positive real number. Then, each edge is given a capacity which is a random variable that only depends on the product of the weights of its endpoints. We then study the minimum spanning tree corresponding to the edge capacities. Under a condition of finite moments on the node weights, we show that the expected diameter and typical distances of this minimum spanning tree are of order $n^{1/3}$. This is a generalization of the results of Addario-Berry et al. [2009]. We then use our result to answer a conjecture in statistical physics about typical distances on a closely related object. This work also sets the ground for proving the existence of a non-trivial scaling limit of this spanning tree (a generalization of the result in Addario-Berry et al. [2017b]). Our proof is based on a detailed study of rank-1 critical inhomogeneous random graphs, done in Chapter 2, and novel couplings between exploration trees related to those graphs and Galton-Watson trees.



3.1 Introduction

3.1.1 The model

Let $G = (V, E)$ be a connected graph with n nodes and m edges. Let $e_1, e_2, ..., e_m$ be positive real numbers that represent capacities on the edges of G. Assume that these capacities are all distinct, then there exists a unique spanning tree $T = (V, E_T)$ that minimizes the sum of the capacities on the edges :

$$\sum_{e_i \in E_T} e_i.$$

We call this tree the minimum spanning tree (MST). If the graph is not connected the set of minimum spanning trees of its connected components is called the minimum spanning forest. If the weights are independent random variables with atomless distributions, the MST will be almost surely unique. The minimum spanning tree is an important object in combinatorial optimization, it can be easily computed (see the introduction of this thesis), and it can be used to construct approximations to more difficult problems such as the traveling salesman (Vazirani [2001]). The study of random minimum spanning trees is also of independent interest in statistical physics (Chen et al. [2006], Wu et al. [2006], Braunstein et al. [2007]).

From a probabilistic perspective, the minimum spanning tree has been studied extensively and on various graphs and models of randomness. Let $\mathcal{T}_n$ denote the minimum spanning tree of a complete graph of size n with i.i.d. $[0, 1]$-uniform random capacities on its edges. Frieze [1985] proved that the total capacity of $\mathcal{T}_n$ converges to $\zeta(3) = \sum_{k=1}^{+\infty} \frac{1}{k^3}$. Aldous [1990] found the limit distribution of the degree of node 1 in $\mathcal{T}_n$. These results are based on a local study of the minimum spanning tree.

Another type of questions concerns global properties of the random MST. For instance, Addario-Berry, Broutin, and Reed [2009] showed that the diameter (maximum number of edges of a shortest path between two nodes) of $\mathcal{T}_n$ is of order $n^{1/3}$. This estimate was used crucially by Addario-Berry et al. [2017b] to show that there exists a measured metric space which is a scaling limit of $\mathcal{T}_n$ with distances scaled by $n^{-1/3}$. In this chapter, we extend the results of Addario-Berry et al. [2009] to a more general class of random minimum spanning trees. Let $n \in \mathbb{N}$ and $\mathbf{W} = (w_1, w_2, ..w_n)$, with $0 < w_n \leq w_{n-1} \leq ... \leq w_1$, be a vector of positive weights, and consider the complete graph of size n. $\mathbf{W}$ will always depend implicitly on n. To each non-oriented edge $\{i, j\}$, $i \neq j$, we associate the random capacity $E_{\{i,j\}}$, which is an exponential random variable of rate $w_i w_j$. The capacities are then used to create a sequence of graphs. For each $p \in [0, +\infty]$ let $G(\mathbf{W}, p)$ be the graph on $\{1, 2...., n\}$ containing the edges of capacity at most p, so the edge set of $G(\mathbf{W}, p)$ is :

$$\left\{\{i, j\} | E_{\{i,j\}} \leq p\right\}.$$

Then $(G(\mathbf{W}, p))_{p\in[0,+\infty]}$ is an increasing sequence of graphs (for inclusion), and for each fixed value of p, $G(\mathbf{W}, p)$ is called a rank-1 inhomogeneous random graph. Now for each p, consider the forest $\mathcal{T}(\mathbf{W}, p)$ constructed by deleting edges from $G(\mathbf{W}, p)$ as follows :

We construct a sequence of graphs $(\mathcal{G}(\mathbf{W}, p, i))_{\binom{n}{2}\geq i\geq 1}$ such that $\mathcal{G}(\mathbf{W}, p, \binom{n}{2}) = \mathcal{T}(\mathbf{W}, p)$. First, sort the edges $(\{i, j\})_{i\neq j}$ of $G(\mathbf{W}, p)$ by decreasing order of their capacities $(E_{\{i,j\}})_{i\neq j}$. At step 1, $\mathcal{G}(\mathbf{W}, p, 1) = G(\mathbf{W}, p)$. If deleting the first edge in the order disconnects a connected component of $\mathcal{G}(\mathbf{W}, p, 1)$ then keep it. Otherwise, delete it. This gives $\mathcal{G}(\mathbf{W}, p, 2)$. Then, move on to the next edge and do the same. Continue like this by either keeping or deleting each edge consecutively. This procedure ensures that all the graphs $\mathcal{G}(\mathbf{W}, p, i)$ have the same number of connected components and only their number of cycles decreases. We call this the edges deletion algorithm, also known as bombing optimization (Braunstein et al. [2007]). It is easy to see that this procedure yields a forest in the end since we are removing all the cycles from the graph. Moreover, by construction, $\mathcal{T}(\mathbf{W}, p)$ is the minimum spanning forest of $G(\mathbf{W}, p)$ with respect to the capacities $(E_{\{i,j\}})_{i\neq j}$ because we delete the largest capacities first. Hence, $\mathcal{T}(\mathbf{W}, +\infty)$ is the

minimum spanning tree of the complete graph with respect to the capacities $(E_{\{i,j\}})_{i\neq j}$. The diameter of a graph is the largest graph distance (number of edges) between two of its nodes. If the graph is not connected its diameter is the maximum of the diameters of its connected components. The purpose of this chapter is the study of the diameter of $\mathcal{T}(\mathbf{W}, +\infty)$. Before going further in the discussions, we introduce some notations.

3.1.2 Notations and definition of the exploration process

We start by giving another definition of the graphs and defining the exploration processes we will work with. We use exactly the same definition we already used in Chapter 2. These definitions where already presented in Part 2.1.2, we only repeat them here to make this Chapter self sufficient. Consider $n \in \mathbb{N}$ vertices labeled $1, 2.., n$. For a vector of weights $\mathbf{W} = (w_1, w_2, ..w_n)$, where $0 < w_n \leq w_{n-1} \leq ... \leq w_1$, we create the inhomogeneous random graph associated to $\mathbf{W}$ and to $p \leq +\infty$ in the following way :

Each potential edge $\{i, j\}$ is in the graph with probability $1 - e^{-w_i w_j p}$ independently from everything else. This gives a random graph that we call the rank-1 inhomogeneous random graph associated to $\mathbf{W}$ and $p \leq +\infty$. This construction yields graphs with the marginal distribution of the sequence $(G(\mathbf{W}, p))_{p\in[0,+\infty]}$ presented previously. We will keep both those two graphs construction methods in mind in this chapter, and switch between the two depending on our needs.

Before stating the main theorem, we define an exploration process for $G(\mathbf{W}, p)$ seen as a graph from the sequence $(G(\mathbf{W}, p))_{p\in[0,+\infty]}$ for a fixed p. This process is based on an "horizontal" exploration of the graph, called the breadth-first walk (BFW). Write :

$$\ell_n = \sum_{i=1}^{n} w_i,$$

and recall that the weights depend implicitly on n. We say that a tree is spanning a graph if it has the same set of nodes and a subset of its edges. The BFW also naturally yields a spanning forest of $G(\mathbf{W}, p)$. That is a sequence of spanning trees of the connected component of $G(\mathbf{W}, p)$.

For each potential edge $\{i, j\}$ recall the definition of $E_{\{i,j\}}$ from the previous subsection. The BFW operates by steps, define the following sets of vertices. A vertex is always in exactly one of those sets :

— $(\mathcal{U}(i))_{1\leq i\leq n}$ is the sequence of sets of unexplored vertices at each step.
— $(\mathcal{D}(i))_{1\leq i\leq n}$ is the sequence of sets of discovered but not yet explored vertices at each step.
— $(\mathcal{F}(i))_{1\leq i\leq n}$ is the sequence of sets of explored vertices at each step.

First, choose a vertex i with probability :

$$\mathbb{P}(v(1) = i) = \frac{w_i}{\ell_n},$$

and call it $v(1)$. Let $\mathcal{V}$ be the set of all vertices labels, and $\mathcal{U}(1) = \mathcal{V} \setminus \{v(1)\}$, $\mathcal{D}(1) = \{v(1)\}$. At step 2, $v(1)$ is explored. The vertices j that are unexplored and such that $E_{\{j,v(1)\}} \leq p$ become discovered but not yet explored. We call them children of $v(1)$. Let $c(1)$ be the number of children of $v(1)$. Denote their labels by $(v(2), v(3), ..., v(c(1) + 1)$ in increasing order of their $E_{\{j,v(1)\}}$'s. For $i \geq 1$, denote the set $\{v(1), v(2), ..., v(i)\}$ by $\mathcal{V}_i$. Hence, at step 2 we have :

— $\mathcal{U}(2) = \mathcal{V} \setminus \mathcal{V}_{c(1)+1}$.
— $\mathcal{D}(2) = \mathcal{V}_{c(1)+1} \setminus \mathcal{V}_1$.
— $\mathcal{F}(2) = \mathcal{V}_1$.

Now, at the step 3, $v(2)$ becomes explored. The vertices j that are unexplored and such that $E_{\{j,v(2)\}} \leq p$ become discovered but not yet explored. We call them children of $v(2)$, and we denote their labels by $\{v(c(1) + 2), v(c(1) + 3)..., v(c(1) + c(2) + 3)\}$. The BFW continues like

this. At step $i+1$ node $v(i)$ becomes explored. If the set of discovered but not yet explored nodes become empty at some step i, this means that the exploration of a connected component is finished. In that case, we move on to the next step by choosing a vertex j with probability proportional to its weight w_j among the unexplored vertices (like we did for $v(1)$). We call this a size-biased sampling. The spanning forest associated to the the BFW is created by considering each node without a parent as a root, and adding only the edges between parent nodes and their children. We call the trees in that forest the exploration trees. By construction, exploration trees are spanning trees of the connected components of $G(\mathbf{W}, p)$. We say that a connected component is discovered at step i if its first node discovered by the BFW is $v(i)$. Similarly, we say that a connected component is explored at step i if its last node that becomes explored in the BFW is $v(i)$.

Generally, let $c(i)$ be the number of children of the node labeled $v(i)$. The exploration process associated to the BFW above is defined as follow, for $n-1 \geq i \geq 0$:

$$\begin{aligned} L'_0 &= 1, \\ L'(i+1) &= L'_i + c(i+1) - 1. \end{aligned}$$

The reflected exploration process is defined by

$$\begin{aligned} L_0 &= 1, \\ L(i+1) &= \max(L_i + c(i+1) - 1, 1). \end{aligned}$$

The increment of the process L' at step i is the number of nodes added to the set of discovered nodes in the BFW at step i. This number is at least -1 if the node being explored has no children. The process L' contains a lot of the information about $G(\mathbf{W}, p)$. For a more in-depth discussion on those process we refer the reader to the related part in Chapter 2.

The order of appearance of the nodes in the exploration process corresponds to a size-biased sampling. Formally, we have :

Lemma 41. *With the notations presented above, for :*

$$i \in \{0, 1, ..., n-1\},$$

and

$$j \in \{0, 1, ..., n\},$$

we have :

$$\mathbb{P}\left(v(1) = j\right) = \frac{w_j}{\ell_n}.$$

$$\mathbb{P}\left(v(i+1) = j \quad |\mathcal{V}_i\right) = \frac{w_j \mathbb{1}(j \notin \mathcal{V}_i)}{\ell_n - \sum_{k=1}^{i} w_{v(k)}}.$$

The proof of this fact is present at the end of part 2.1.2. We will assume the following conditions on $\mathbf{W}$ in the entire chapter.

Conditions 4. *There exists some positive random variable W such that :*

(i) *The distribution of a uniformly chosen weight w_X converges weakly to W.*

(ii) $\mathbb{E}[W^3] < \infty$.

(iii) $\mathbb{E}[W^2] = \mathbb{E}[W]$.

(iv) $\ell_n = \mathbb{E}[W]n + o(n^{2/3})$.

(v) $\sum_{k=1}^{n} w_k^2 = \mathbb{E}[W^2]n + o(n^{2/3})$.

(vi) $\sum_{k=1}^{n} w_k^3 = \mathbb{E}[W^3]n + o(1)$.

(vii) $\max_{i\leq n} w_i = o(n^{1/3})$.

We refer the reader to Chapter 2 for a more in-depth discussion of these conditions. Here, we only draw attention to the fact that an important case to keep in mind is when $(w_1, w_2, ..., w_n)$ are realizations of random variables $(W_1, W_2, ..., W_n)$ which are i.i.d. with distribution W. In that case the conditions can be seen as consequences of convergence theorems and hold almost surely.

We also define the size of a connected component $\mathcal{C}$, with vertices set $V(\mathcal{C})$, as the number of vertices in $\mathcal{C}$. The distance between two vertices of $\mathcal{C}$ is the number of edges in the smallest (in number of edges) path between them. We also define the weight of $\mathcal{C}$ as :

$$\sum_{i\in V(\mathcal{C})} w_i.$$

We call surplus (or excess) of $\mathcal{C}$ the number of edges that have to be removed from it in order to make it a tree. For instance, the surplus of a tree is 0, and the surplus of a cycle is 1. We can now state our two main theorems.

Theorem 42. *Suppose that Conditions 4 are verified. For any* $\ell_n^{4/3} - \ell_n^{1/3} \geq f_n \geq 0$ *the diameter of* $\mathcal{T}(\boldsymbol{W}, \frac{1}{\ell_n} + \frac{f_n}{\ell_n^{4/3}})$*, denoted by* $\mathrm{diam}(\mathcal{T}(\boldsymbol{W}, \frac{1}{\ell_n} + \frac{f_n}{\ell_n^{4/3}}))$*, verifies*[1] *:*

$$\mathbb{E}\left[\mathrm{diam}\left(\mathcal{T}\left(\boldsymbol{W}, \frac{1}{\ell_n} + \frac{f_n}{\ell_n^{4/3}}\right)\right)\right] = \Theta(n^{\frac{1}{3}}).$$

We will further discuss our choice of parameter f_n in this theorem in the next sections. The second theorem is related to typical distances in those minimum spanning trees.

Theorem 43. *Suppose that Conditions 4 are verified. For* $c > 1$ *and a sequence of parameters* $\tilde{p}_n \geq \frac{c}{\ell_n}$*, let* (U_{f_n}, V_{f_n}) *be two uniformly randomly drawn nodes from the largest tree in* $\mathcal{T}(\boldsymbol{W}, \tilde{p}_n)$*, and let* $d(U_{f_n}, V_{f_n})$ *be the distance between them in* $\mathcal{T}(\boldsymbol{W}, \tilde{p}_n)$*. We have :*

$$\mathbb{E}[d(U_{f_n}, V_{f_n})] = \Theta(n^{1/3}).$$

One important thing that will be clear from our proofs is that those theorems hold if we replace the inhomogeneous minimum spanning trees by other types of trees verifying an inclusion property. Let $a > 0$ and $(G_p)_{p\geq a}$ be an increasing process of graphs for inclusion. We say that the sequence of forests $(T_p)_{p\geq a}$ is an increasing process of spanning forests of $(G_p)_{p\geq a}$ if, for any $p \geq a$, T_p is a spanning forest of G_p, and the process $(T_p)_{p\geq a}$ is increasing for inclusion.

By definition $(\mathcal{T}(\mathbf{W}, p))_{p\geq 0}$ is an increasing process of spanning forests of $(G(\mathbf{W}, p))_{p\geq 0}$. A direct corollary of our proofs is the following :

Corollary 43.1. *Theorems 42 and 43 still hold if we replace the trees* $(\mathcal{T}(\boldsymbol{W}, p))_{p\geq \frac{1}{\ell_n}}$ *by any increasing process of spanning forests of* $(G(\boldsymbol{W}, p))_{p\geq \frac{1}{\ell_n}}$*.*

This corollary is true because our proofs only use the fact that $(\mathcal{T}(\mathbf{W}, p))_{p\geq 0}$ in an increasing process of spanning forests of $(G(\mathbf{W}, p))_{p\geq 0}$. And we never use any other property proper to $(\mathcal{T}(\mathbf{W}, p))_{p\geq 0}$.

1. The notation $\Theta(n^{\frac{1}{3}})$ means that there exists two constants $\varepsilon > 0$ and $A > 0$ and an $N > 0$, such that for any $n \geq N$ we have the diameter is larger than $\varepsilon n^{1/3}$ and smaller than $An^{1/3}$.

This is thus a generalization of Theorems 42 and 43. However, we emphasise the fact that given an increasing process of graphs for inclusion. An increasing process of spanning trees is uniquely determined by the spanning forest of the first graph in the process. The condition of being an increasing process of spanning trees is thus very strong and does not allow for much freedom in the choice of the process. Still, this is a nice addition that will be used in this chapter to tackle a conjecture from statistical physics.

3.1.3 Further discussion and related work in statistical physics

Addario-Berry et al. [2009] already showed a similar result to Theorem 42 with $\mathbf{W} = (1, 1, ..., 1)$. In this case the inhomogeneous random graph is in fact an Erdős-Rényi random graph. Our theorem is thus a generalization of theirs. It is also another hint at the fact that when the weights of rank-1 inhomogeneous random graphs verify a third moment condition, which corresponds to Conditions 4, then their asymptotic behavior is similar to Erdős-Rényi random graphs. Meaning that the inhomogeneity disappears asymptotically. Such behavior was already remarked for other asymptotic properties (see for instance Bhamidi et al. [2010]).

Another interesting result is in Addario-Berry et al. [2017b] where the authors showed the existence of a measured metric space that is a scaling limit of the minimum spanning tree associated to Erdős-Rényi random graph with distances rescaled by $n^{-1/3}$. In that article the concentration results that gave the order of the diameter of the minimum spanning tree corresponding to $\mathbf{W} = (1, 1, ..., 1)$ were crucially used. A similar result was obtained for 3-regular graphs with i.i.d. capacities on their edges in Addario-Berry and Sen [2019]. We do the same for general weights verifying Conditions 4 in the next chapter.

In an even more general setting, one can ease the Conditions 4. A large body of work (Chen et al. [2006], Braunstein et al. [2007], Stegehuis et al. [2017], van der Hofstad et al. [2018], Broutin et al. [2020]) is interested in weights related to power law distributions, the so called scale-free model. In that case the diameter is not expected to be of order $n^{1/3}$. If we suppose that the weights are drawn from a distribution with power law tail with parameter $4 \geq \alpha > 3$, then intuitive arguments suggest that the diameter of the minimum spanning tree should be of order $n^{\frac{\alpha-3}{\alpha-1}}$ (see Broutin et al. [2020]). The scaling limit of such trees should also be mutually singular from one another for different values of α. Some of the tools presented here could be useful to prove results for weights with power law distributions. However, more work is required to prove those conjectures.

The model presented here is closely related and generally called the Norros-Reitu model (Norros and Reittu [2006]) altough it is slightly different from the first model proposed by Norros and Reitu, since their model allows for multigraphs. This has no incidence on our results. In fact, our model is more closely related to the multiplicative coalescent introduced by Aldous in Aldous [1997], and further studied in Aldous and Limic [1998] by Aldous and Limic. Moreover, the results we present for this model also hold for closely related models. For instance, Chung-Lu (Chung and Lu [2006]) proposed a random graph with prescribed expected degrees. In that model two nodes $\{i, j\}$ are connected with probability :

$$p_{i,j} = \frac{w_i w_j}{\ell_n}.$$

It supposes that $\max_{i,j}(w_i w_j) \leq \ell_n$. The following classical theorem (proved in Bollobás et al. [2007]) holds for all those graph models and is crucial in understanding the evolution of the graph process $(G(\mathbf{W}, p))_{p\geq 0}$.

Theorem 44. *Recall that $\tilde{p}_n = \frac{c}{\ell_n}$. Take $G(\boldsymbol{W}, \tilde{p}_n)$ and suppose that Conditions 4 are verified, then the following results hold with high probability*[2] *:*

2. We say that a sequence of events E_n holds with high probability if $\lim_{n\to\infty} \mathbb{P}(E_n) = 1$.

— ***Subcritical regime*** *If $c < 1$ then the largest connected component is of size $o(n)$.*
— ***Supercritical regime*** *If $c > 1$ then the largest connected component is of size $\Theta(n)$ and for any $i > 1$ the i-th largest connected component is of size $o(n)$.*
— ***Critical regime*** *If $c = 1$ then for any $i \geq 1$ the i-th largest connected component is of size $\Theta(n^{2/3})$.*

A simulation-based conjecture from statistical physics (Chen et al. [2006], Wu et al. [2006], Braunstein et al. [2007]) is the following : Consider the largest component of $G(\mathbf{W}, \tilde{p}_n)$ for $c > 1$. Put i.i.d. capacities derived from some continuous distribution on the edges of that component. Then typical distances in the minimal spanning tree constructed on the largest component with those i.i.d. capacities scale like $n^{1/3}$. Here typical distances mean distance between two uniformly drawn nodes from the tree. First, notice that as long as the distribution used is atomless, it has no incidence on the distribution of the geometry of the tree obtained. Indeed, only the relative order of the edge capacities is relevant for the geometry of the minimum spanning tree. Hence, putting i.i.d. capacities on the edges amounts to taking a uniform random order on them for our purpose. In order to prove this conjecture, it is sufficient to prove a result similar to Theorem 43 but for the type of minimum spanning trees present in the statistical physics literature. In fact, this minimum spanning tree is not that much different from the inhomogeneous minimum spanning trees that we study in this chapter. One can see that, with a little more work, it can be shown that if Theorem 43 is true, then a similar theorem also holds for the minimum spanning trees constructed from i.i.d capacities on the giant component of supercritical inhomogeneuous random graphs. For clarity, after proving Theorem 43, we will prove a modified version of it to show the conjecture from statistical physics using our results. We thus prove the following Theorem :

Theorem 45. *Suppose that Conditions 4 are verified. For $c > 1$, and a sequence of parameters $\tilde{p}_n \geq \frac{c}{\ell_n}$, put i.i.d. capacities with an atomless distribution on the edges of the graphs $(\mathcal{G}(\boldsymbol{W}, \tilde{p}_n))_{n\geq 1}$. Let $\mathcal{T}'(\boldsymbol{W}, \tilde{p}_n)$ be the minimum spanning forest of $\mathcal{G}(\boldsymbol{W}, \tilde{p}_n)$ corresponding to those edge capacities. Let (U_{f_n}, V_{f_n}) be two uniformly randomly drawn nodes from the largest tree in $\mathcal{T}'(\boldsymbol{W}, \tilde{p}_n)$, and let $d(U_{f_n}, V_{f_n})$ be the distance between them in $\mathcal{T}'(\boldsymbol{W}, \tilde{p}_n)$. We have :*

$$\mathbb{E}[d(U_{f_n}, V_{f_n})] = \Theta(n^{1/3}).$$

Another closely related conjecture concerns the power law distribution case. It is generally also conjectured that typical distances in the minimum spanning tree with i.i.d. capacities on the edges of largest components of graphs with nodes weights following a distribution with a power law tail of parameter $4 > \alpha > 3$, either in inhomogeneous random graphs as studied here, or in the configuration model (Braunstein et al. [2007]), scale like $n^{\frac{\alpha-3}{\alpha-1}}$. Proving the existence of a non-trivial scaling limits for $\mathcal{T}(\mathbf{W}, +\infty)$ when we change Conditions 4 to conditions pertaining to power law distribution would also prove such a conjecture. A first step in that direction was done very recently in Bhamidi and Sen [2020]. The authors proved that the diameter of such trees is in fact of order $n^{\frac{\alpha-3}{\alpha-1}}$ in the inhomogeneous case and they even showed the existence of a non-trivial scaling limit for those trees. However, their work supposes restrictive conditions on the distribution other than just having power law tail. Thus, the general case remains open.

Lastly in this introduction, we remind the reader of Bernstein's inequality (Bernstein [1924], Boucheron et al. [2013]). This inequality will be used through this chapter in order to prove some concentration bounds.

Lemma 46. *Let $X_1, X_2, ..., X_n$ be i.i.d. random variables with finite second moment. Suppose that $|X_i| \leq M$, for all i. Then for any positive number x :*

$$\mathbb{P}\left(\left|\sum_{i\geq 1}^{n}(X_i - \mathbb{E}[X_i])\right| \geq x\right) \leq 2\exp\left(\frac{-\frac{1}{2}x^2}{\sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right] + \frac{1}{3}Mt}\right).$$

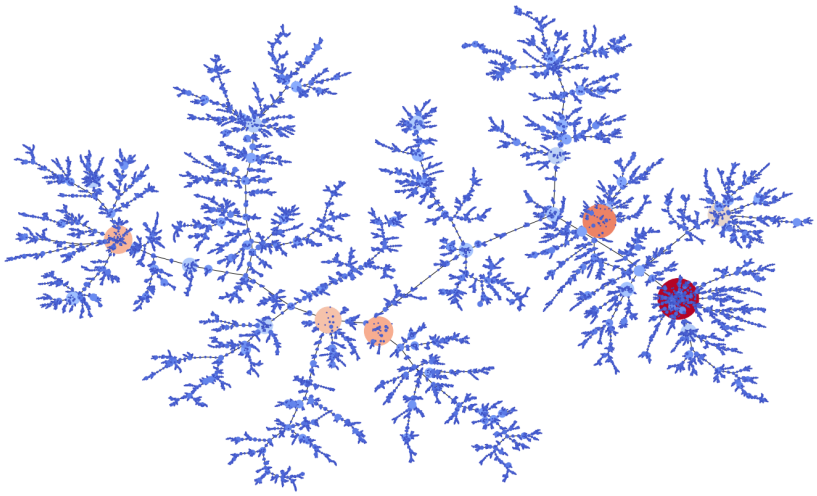


FIGURE 3.1 – A minimum spanning tree built on the giant component of a scale-free supercritical inhomogeneous random graph with parameter $\alpha = 3.5$. The edge capacities are i.i.d.. Nodes are coloured from largest in red to smallest in dark-blue. Node sizes are proportionnal to the square of their degrees.

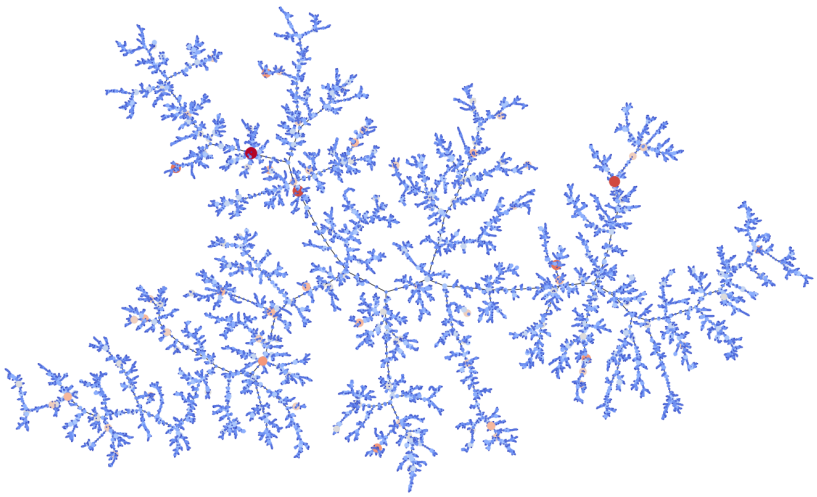


FIGURE 3.2 – A minimum spanning tree built on the giant component of a supercritical inhomogeneous random graph with node weights having finite third moments. The edge capacities are i.i.d.. Nodes are coloured from largest in red to smallest in dark-blue. Node sizes are proportionnal to the square of their degrees.

3.2 First ingredients of the proof

For $p \geq 0$. Any component of $G(\mathbf{W}, p)$ that is a tree is also a component of $\mathcal{T}(\mathbf{W}, p)$. Since $\mathcal{T}(\mathbf{W}, p) \subseteq \mathcal{T}(\mathbf{W}, +\infty)$, the diameter of any tree component of $G(\mathbf{W}, p)$ is a lower bound on $\mathrm{diam}(\mathcal{T}(\mathbf{W}, +\infty))$. The following theorem is a simple corollary of the much more general convergence result in Theorem 2.4 of Broutin et al. [2020].

Theorem 47. *There exists $\varepsilon > 0$ and $\varepsilon' > 0$ such that for all n large enough the largest connected component of $G(\boldsymbol{W}, \frac{1}{\ell_n})$ is a tree of diameter at least $\varepsilon' n^{1/3}$ with probability at least ε.*

Theorem 47 proves that there exist a constant $\varepsilon > 0$ such that $\mathbb{E}[\mathrm{diam}(\mathcal{T}(\mathbf{W}, +\infty))] \geq \varepsilon n^{1/3}$. The rest of this chapter is dedicated to proving the upper bound of Theorem 42.

3.2.1 The phase transition

The edge deletion algorithm provides a coupling between $\mathcal{T}(\mathbf{W}, +\infty)$ and the graph $G(\mathbf{W}, +\infty)$. Hence, by studying the evolution of the longest paths of the graphs in $(G(\mathbf{W}, p))_{p\in[0,+\infty]}$ we will be able to prove Theorem 42. As a first step, we want to know what are the important values of p in this process. That question was partly answered in Theorem 44. We see that there is a phase transition at $p = \frac{1}{\ell_n}$. It turns out that $p = p_f = \frac{1}{\ell_n} + \frac{f}{\ell_n^{4/3}}$, with $f \in \mathbb{R}$, constitutes a phase transition "window" for the size of the connected components sizes. In that window, the size and weight of the largest component increase smoothly. This is the so called critical regime. It is thus natural to investigate this critical regime in order to prove Theorem 42.

We will show that the diameter of $\mathcal{T}(\mathbf{W}, p_f)$ is a $O(n^{1/3})$, when $p = p_f = \frac{1}{n} + \frac{f}{n^{4/3}}$ with $f > 0$ a large constant. Then it only grows by $O(n^{1/3})$ when p becomes strictly larger than $\frac{1}{n}$. The increase of the diameter in the supercritical phase is easy to study, the main problem is to show that when we cross the critical regime, $p_f = \frac{1}{\ell_n} + \frac{f_n}{\ell_n^{4/3}}$ with f_n going to infinity, the diameter remains of order $n^{1/3}$. The value $f'_n = \frac{\ell_n^{1/3}}{\log(n)}$ which gives $p_{f'_n} = \frac{1}{\ell_n} + \frac{1}{\log(n)\ell_n}$ constitutes the pivotal moment where we will switch arguments : $p_{f'_n}$ is chosen like this because it is neither a critical nor a real supercritical value. We say that it is in the barely supercritical regime. In the rest of this section we will take Theorems 48 and 49 for granted and prove Theorem 42. The proofs of these theorems will be the focus of the rest of the chapter.

Notation : In the remainder of the chapter we drop the n from f_n for clarity, f will always be the critical parameter, moreover we will always assume $f = o(n^{1/3})$ and $f \geq F$, where $F > 0$ is a constant independent of n which is large enough for all our theorems to be true. Similarly $A, A', A'', ... \in \mathbb{R}^+$ will always be large constants independent of n. And $\varepsilon, \varepsilon', ... \in \mathbb{R}^+$ will always be small constants independent of n.

Theorem 48. *Assume Conditions 4 are verified, there exists a positive constant $A > 0$ such that : With probability at least $(1 - \frac{1}{n})$, the largest component, $H(\boldsymbol{W}, p_{f'_n})$, of $G(\boldsymbol{W}, p_{f'_n})$ has total weight at least $\frac{n}{A\log(n)}$ and every other connected component has total weight at most $A\log(n)^{1/2}n^{1/2}$ and longest path of length less than $A\log(n)^{1/4}n^{1/4}$.*

Theorem 49. *Assume Conditions 4 are verified, then :*

$$\mathbb{E}[diam(\mathcal{T}(\boldsymbol{W}, p_{f'_n}))] = O(n^{1/3})$$

3.2.2 The supercritical phase

Here we take Theorems 48 and 49 for granted. Let $\mathcal{C}$ be a connected component, with vertices set $V(\mathcal{C})$ and edge set $D(\mathcal{C})$, and suppose that $\mathcal{C}$ is not the largest connected component

of $G(\mathbf{W}, p_{f'_n})$. Let $\{I, J\}$ be the random edge with exactly one endpoint (say I) in $\mathcal{C}$ and with minimal capacities $E_{\{I,J\}}$ among edges with exactly one endpoint in $\mathcal{C}$. By construction, $\{I, J\}$ will necessarily be in $\mathcal{T}(\mathbf{W}; +\infty)$. We want to calculate $\mathbb{P}(J \in V(H(\mathbf{W}, p_{f'_n})))$, the probability that the other endpoint of the edge $\{I, J\}$ lies in the largest component $H(\mathbf{W}, p_{f'_n})$ with vertices set $V(H(\mathbf{W}, p_{f'_n}))$ and edge set $D(H(\mathbf{W}, p_{f'_n}))$. Recall that for any couple $\{i, j\}$, $E_{\{i,j\}}$ is an exponential random variable of parameter $w_i w_j$. So if we fix $\{i, j\}$, by classical results on exponential random variables :

$$\mathbb{P}\left(\forall l \neq i, l \neq j \quad E_{\{i,j\}} < E_{\{i,l\}}\right) = \frac{w_j}{\sum_{k=1}^{n} w_k}.$$

Similarly, write :

$$\mathbf{U} = \left\{ I \in V(\mathcal{C}), \exists j \in V(H(\mathbf{W}, p_{f'_n})), \forall l \in (\mathcal{V} \setminus (V(\mathcal{C}) \cup \{j\})\ E_{\{I,j\}} < E_{\{I,l\}} \right\}.$$

Then :

$$\mathbb{P}\left(\mathbf{U} | (V(H(\mathbf{W}, p_{f'_n})), V(\mathcal{C}))\right) = \frac{\sum_{k \in V(H(\mathbf{W}, p_{f'_n}))} w_k}{\sum_{k'=1}^{n} w_{k'} - \sum_{k'' \in V(\mathcal{C})} w_{k''}}.$$

Which implies that :

$$\mathbb{P}\left(J \in V(H(\mathbf{W}, p_{f'_n})) | (V(H(\mathbf{W}, p_{f'_n})), V(\mathcal{C}))\right) = \frac{\sum_{k \in H(\mathbf{W}, p_{f'_n})_V} w_k}{\sum_{k'=1}^{n} w_{k'} - \sum_{k'' \in V(\mathcal{C})} w_{k''}}. \tag{3.1}$$

By Theorem 48 and Equation (3.1), we get for n large enough :

$$\begin{aligned} 1 - \mathbb{P}\left(J \in V(H(\mathbf{W}, p_{f'_n}))\right) &\leq 1 - \frac{n}{A \log(n)(n - A \log(n)^{1/2} n^{1/2})} - \frac{1}{n} \\ &\leq 1 - \frac{2}{A \log(n)}, \end{aligned}$$

If J is not in $V(H(\mathbf{W}, p_{f'_n}))$, then J lies in another connected component $\mathcal{C}'$, with vertix set $V(\mathcal{C}')$ and edge set $D(\mathcal{C}')$. So the longest path of the newly created component will be at most $2A \log(n)^{1/4} n^{1/4}$ with probability at least $1 - \frac{2}{A \log(n)}$. Let $\{I', J'\}$ be the edge with the smallest capacities with exactly one endpoint I' in the newly created connected component $\mathcal{C}''$ which is the concatenation of $\mathcal{C}$ and $\mathcal{C}'$ through the edge $\{I, J\}$. By the same argument as before, conditionally on the event F that J is not in $V(H(\mathbf{W}, p_{f'_n}))$, we get for n large enough :

$$\begin{aligned} 1 - \mathbb{P}\left(J' \notin V(H(\mathbf{W}, p_{f'_n}))) | F\right) &\leq 1 - \frac{n}{\log(n)(n - 2A \log(n)^{1/2} n^{1/2})} - \frac{1}{n} \\ &\leq 1 - \frac{2}{A \log(n)}. \end{aligned}$$

By an immediate induction, for any r_n small enough, the probability that $\mathcal{C}$ is in a component of longest path larger than $r_n A \log(n)^{1/4} n^{1/4}$ when it connects to the largest connected component is at most $(1 - \frac{2}{A \log(n)})^{r_n}$. By taking $r_n = A(\log(n))^2$, it follows, using the inequality

$$\left(1 - \frac{2}{A \log(n)}\right) \leq \exp\left(\frac{-2}{A \log(n)}\right),$$

that the probability that the longest path of $\mathcal{C}$ reaches $A^2 \log(n)^{9/4} n^{1/4}$ before connecting to the largest component is less than $\frac{1}{n^2}$. Since $\mathcal{C}$ was arbitrary, and since there are at most n such components. We get by the union bound that the probability that there is a component that has a path longer than $A^2 \log(n)^{9/4} n^{1/4} = o(n^{1/3})$ when it connects to the largest connected component is at most $\frac{1}{n}$. Finally recall that we are building trees, and the diameter of the minimum spanning tree can at most go through two of those components.

In order to link the diameters of nested graphs we use the following lemma (Also used in Addario-Berry et al. [2009]).

Lemma 50. *Let H' and H be two connected graphs such that $H' \subset H$. Then*

$$diam(H) \leq diam(H') + 2lp(H(V - V(H'))) + 2.$$

Here lp stands for longest path and $H(V - V(H'))$ is the graph H from which we deleted the nodes of H' and any edge that had at least one endpoint in H'.

Proof. In order to show the lemma it is sufficient to show that between any two vertices H there exist a path of length at most than $diam(H') + 2lp(H(V - V(H'))) + 2$. Let x, y be two vertices of H, one can create a path from x to y in the following way.

— There exists a path π_1 between x and a vertex of H' that has length at most $lp(H(V - V(H'))) + 1$, let x' be the arrival vertex of this path.
— There exists a path π_2 between y and a vertex of H' that has length at most $lp(H(V - V(H'))) + 1$, let y' be the arrival vertex of this path.
— There exists a path π_3 between x' and y' inside H' that has length at most $diam(H')$.

Concatenating the three paths π_1, π_2 and π_3 provides the desired path.

□

Lemma 50 and the discussion above imply that :

Theorem 51. *Under conditions 4 :*

$$\mathbb{E}\left[\text{diam}(\mathcal{T}(\boldsymbol{W}; +\infty)) - \text{diam}(\mathcal{T}(\boldsymbol{W}; f'_n))\right] = O(\log(n)^{9/4} n^{1/4}).$$

Theorem 1 follows from this theorem and Theorem 49. Now what is left is proving Theorems 48 and 49. This is what we will do in the rest of the chapter.

3.2.3 Known results

Here we remind the reader of the main theorems in Chapter 2 that will be used in this chapter. Let :

$$C = \frac{\mathbb{E}[W^3]}{\mathbb{E}[W]} \geq 1.$$

Theorem 52. *Suppose that Conditions 4 hold. There exists a constant $A > 0$, such that for any $1 > \varepsilon' > 0$, $1 \geq \varepsilon > 0$, and $f = o(n^{1/3})$ large enough with probability at least :*

$$1 - A\exp\left(\frac{-f^{\varepsilon/2}}{A}\right) - A\exp\left(\frac{-n^{1/12}}{A}\right),$$

all the following events occur in $G(\boldsymbol{W}, p_f)$:

— *The size of the largest component is in the interval*

$$\left[\frac{2(1-\varepsilon'/2)f\ell_n^{2/3}}{C}, \frac{2(1+\varepsilon'/2)f\ell_n^{2/3}}{C}\right].$$

— *Every other connected component has size less than $\frac{A\ell_n^{2/3}}{f^{1-\varepsilon}}$.*
— *The weight of the largest component is in the interval*

$$\left[\frac{2(1-\varepsilon')f\ell_n^{2/3}}{C}, \frac{2(1+\varepsilon')f\ell_n^{2/3}}{C}\right].$$

— *Every other connected component has weight less than* $\frac{(1+\varepsilon')A\ell_n^{1/3}}{f^{1-\varepsilon}}$.
— *The surplus of the largest connected component is less than* Af^3.
— *The surplus of any other connected component is less than* Af^ε.

The Theorems of 2 hold with sharper bound containing $n^{1/8}$, but this bound is sufficient for us and we will use it in this chapter. This theorem contains many important statements for our proof. It gives bounds on the size, weight and surplus of the largest connected component of rank-1 inhomogeneous random graphs in the critical window, and also of its small components. For instance, by taking $f = f'_n$ in the above theorem we get the statement about the weights in Theorem 48. The Chapter 2 also provides bounds on the time the giant component is discovered during the exploration process. These results show a large part of Theorems 48 and 49. However they are not enough to bound the length of the longest paths of the graph. Information about the length of paths in the graph is still lacking. Although Theorem 52 contains a bound on the surplus of each connected component of $G(\mathbf{W}, p_{f'_n})$, it is not sufficient to prove upper bounds on the height of any spanning trees of the connected components in order to finish the proof. This is due to the fact that the surplus of the connected components of $G(\mathbf{W}, p_{f'_n})$ is too large. In the remainder of the chapter we will use a "snapshot" trick to fix this problem.

3.3 Dealing with the small components

In order to bound the longest path of a graph using its excess and its height we will use the following lemma.

Lemma 53. *Let G be a connected graph of excess q and such that there exist a spanning tree T of this graph of height h, then the longest path of the graph is of length at most*

$$2h(q+1)+q.$$

Proof. Let $\{D_1, D_2, D_3, ...\}$ be the set of edge-disjoint paths of maximum lengths in T taken in decreasing order of their lengths, if two paths of the same length exist at some point we choose one of them arbitrarily. We can create G by adding the excess edges consecutively to T.

Recursively, the first edge added will create a new longest path of length at most $|D_1| + |D_2| + 1 \leq 4h + 1$. The second edge added will at most create a new longest path containing D_1,D_2 and D_3 and of length $|D_1| + |D_2| + |D_3| + 2 \leq 6h + 2$. Hence, a direct induction shows that after adding q excess edges the longest path we can have is of length $2h(q+1)+q$. □

If we want to use this lemma, then bounding directly the height of the exploration tree of the largest component of $G(\mathbf{W}, p_{f'_n})$ will not be enough to prove Theorem 49. Indeed, Theorem 52 states that the excess of such a component is upper bounded by $f_n'^3 = \frac{\ell_n}{\log(n)^3}$. This is already much larger than the $n^{1/3}$ of Theorem 49. In order to circumvent this problem we will use the following steps.

1. Define a sequence of "snapshots", $(p_{f(i)})_{i\geq 0}$ such that :

   $$f(0) = F$$

   is a large constant, and

   $$f(i+1) = \frac{3}{2}f(i),$$

   for any $i \geq 0$.

2. For every $i \geq 0$, let $\tilde{\mathcal{V}}_i$ be the set of nodes of the largest connected component of $G(\mathbf{W}, p_{f(i+1)})$, and let $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$ be the sub-graph of $G(\mathbf{W}, p_{f(i+1)})$ from which we have taken out the nodes of the largest component of $G(\mathbf{W}, p_{f(i)})$ alongside any edge that has one of those nodes as endpoint. Show that the length of the longest path of this graph is bounded by $\frac{A\ell_n^{1/3}}{f(i+1)^{1/4}}$ w.h.p. This will also prove Theorem 48.
3. Show that the length of the longest path of the largest component of $G(\mathbf{W}, p_{f(i)})$ is smaller than $Af^5\ell_n^{1/3}$ w.h.p.
4. Use the precedent steps to show that when moving from $G(\mathbf{W}, p_{f(i)})$ to $G(\mathbf{W}, p_{f(i+1)})$, the diameter of the minimum spanning tree does not increase significantly. This will prove Theorem 49.

We provide the details for step 2 in the present section, step 3 in section 3.4, and step 4 in the last section.

3.3.1 Construction of the growth of small components

For $i \geq 1$, we want to bound the size, excess and length of longest path of the graph comprised of $G(\mathbf{W}, p_{f(i+1)})$ from which we have taken out the vertices of the largest component in $G(\mathbf{W}, p_{f(i)})$. Recall that we denote the set of those vertices by $\tilde{\mathcal{V}}_i$, and that graph by $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$. Bounding this length will allow us to bound the growth of the giant component when moving from $G(\mathbf{W}, p_{f(i)})$ to $G(\mathbf{W}, p_{f(i+1)})$. Instead of studying $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$ from scratch, we emphasis the fact that this graph is very similiar to $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i)})$ which is a critical inhomogeneous random graph from which we have taken the largest component, and which has already been studied extensively in Chapter 2. Let $K(i, \varepsilon')$ be the union of following events :

1. The size of the largest component of $G(\mathbf{W}, p_{f(i)})$ is in the interval

$$\left[\frac{2(1-\varepsilon'/2)f(i)\ell_n^{2/3}}{C} - \frac{\ell_n^{2/3}}{\sqrt{f(i)}C}, \frac{2(1+\varepsilon'/2)f(i)\ell_n^{2/3}}{C}\right].$$

2. The total weight of the largest connected component of $G(\mathbf{W}, p_{f(i)})$ is in the interval

$$\left[\frac{2(1-\varepsilon')f(i)\ell_n^{2/3}}{C}, \frac{2(1+\varepsilon')f(i)\ell_n^{2/3}}{C}\right].$$

By Theorems 28 and 29 from Chapter 2, there exists a large constant $A > 0$ for any $1 > \varepsilon' > 0$ and f large enough, $K(i, \varepsilon')$ happens with probability at least :

$$1 - A\exp\left(\frac{-\sqrt{f}}{A}\right). \tag{3.2}$$

Let $(\tilde{v}(1), \tilde{v}(2), ..)$ be the ordered labels in the exploration process of the graph $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$. Conditionally on the nodes of the largest component $\tilde{\mathcal{V}}_i$, $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i)})$ is an inhomogeneous random graph with vertices label set $\mathcal{V} \setminus \tilde{\mathcal{V}}_i$ and probability transition $p_{f(i)}$. And so $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$ is an inhomogeneous random graph with vertices label set $\mathcal{V} \setminus \tilde{\mathcal{V}}_i$ and with the slightly larger probability transition $p_{f(i+1)}$. Moereover, since the event $K(i, \varepsilon')$ is measurable with regard to $\tilde{\mathcal{V}}_i$. The same proofs of Lemmas 8, 9, 10 and 35 from Chapter 2 also yields similar results if we replace the $v(j)$'s with $\tilde{v}(j)$'s. We prove one of those results as an example here, and direct the reader to Chapter 2 for the rest of the proofs.

Lemma 54. *Let $l = o(n)$, we have :*

$$\mathbb{E}(w_{\tilde{v}(l)}\mathbb{1}(K(i,\varepsilon'))) \leq 1 + o(1).$$

Proof. First. Denote $\{\tilde{v}(1), \tilde{v}(2), ..., \tilde{v}(l)\}$ by $\mathbf{V}_{l-1}$. Then by definition :

$$\begin{aligned}\mathbb{E}[w_{\tilde{v}(l)}\mathbb{1}(K(i,\varepsilon'))] &= \mathbb{E}\left[\mathbb{E}[w_{\tilde{v}(l)}\mathbb{1}(K(i,\varepsilon'))|(\mathbf{V}_{l-1}, \tilde{\mathcal{V}}_i)]\right] \\ &= \mathbb{E}\left[\mathbb{1}(K(i,\varepsilon')) \sum_{j \notin \mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} \frac{w_j^2}{\ell_n - \sum\limits_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} w_j}\right].\end{aligned}$$

By Lemma 7 when $K(i,\varepsilon')$ holds :

$$\left\{\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i}^{l} w_j = o(n)\right\},$$

and :

$$\left\{\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i}^{l} w_j^2 = o(n)\right\}.$$

Hence :

$$\begin{aligned}\mathbb{E}\left[w_{\tilde{v}(l)}\mathbb{1}(K(i,\varepsilon'))\right] &= \mathbb{E}\left[\mathbb{1}(K(i,\varepsilon')) \sum_{j\notin\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} \frac{w_j^2}{\ell_n\left(1 - \frac{\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} w_j}{\ell_n}\right)}\right] \\ &= \mathbb{E}\left[\mathbb{1}(K(i,\varepsilon')) \sum_{j\notin\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} \frac{w_j^2}{\ell_n}\right](1+o(1)) \\ &= \mathbb{E}\left[\mathbb{1}(K(i,\varepsilon'))\sum_{j=1}^{n}\frac{w_j^2}{\ell_n}\right](1+o(1)) + o(1) \\ &\leq 1 + o(1)\end{aligned}$$

□

Similarly we also have :

Lemma 55. *Let $l = o(n)$. For $f(i)$ large enough and any integer $l \geq u > 0$ we have :*

$$\mathbb{E}(w_{\tilde{v}(l)}w_{\tilde{v}(u)}) \geq \frac{1}{2}.$$

Lemma 56. *Let $l = o(n)$, we have :*

$$\mathbb{E}(w_{\tilde{v}(l)}^2) \leq 1 + o(1).$$

Lemma 57. *Let $u' \geq u \geq 0$, then for any $x \geq 0$:*

$$\mathbb{P}(w_{\tilde{v}(u')}\mathbb{1}(K(i,\varepsilon')) \geq x) \leq \mathbb{P}(w_{\tilde{v}(u)}\mathbb{1}(K(i,\varepsilon')) \geq x)$$

Lemma 58. *Let :*

$$r = 2(1-\varepsilon')C^{-1}f(i)\ell_n^{2/3}.$$

For any $1 \geq \varepsilon > 0$. For $f(i)$ large enough and for any $l = o(n)$ we have :

$$\mathbb{E}[w_{\tilde{v}(l)}\mathbb{1}(K(i,\varepsilon'))] \leq 1 + \frac{l+r}{\ell_n}(1 - C + \varepsilon) + \frac{r(2\varepsilon'(1+\varepsilon')^{-1} + \varepsilon)}{\ell_n} + o\left(\frac{l + f(i)n^{2/3}}{n}\right).$$

Proof. Recall that $\{\tilde{v}(1), \tilde{v}(2), ..., \tilde{v}(l)\}$ is denoted by $\mathbf{V}_{l-1}$. Then by definition :

$$\begin{aligned}\mathbb{E}[w_{\tilde{v}(l)}\mathbb{1}(K(i,\varepsilon'))] &= \mathbb{E}\left[\mathbb{E}[w_{\tilde{v}(l)}\mathbb{1}(K(i,\varepsilon'))|(\mathbf{V}_{l-1}, \tilde{\mathcal{V}}_i)]\right] \\ &= \mathbb{E}\left[\mathbb{1}(K(i,\varepsilon')) \sum_{j\notin \mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} \frac{w_j^2}{\ell_n - \sum\limits_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} w_j}\right].\end{aligned}$$

By Lemma 7 when $K(i,\varepsilon')$ holds :

$$\left\{\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i}^{l} w_j = o(n)\right\},$$

Hence :

$$\begin{aligned}\mathbb{E}\left[w_{\tilde{v}(l)}\mathbb{1}(K(i,\varepsilon'))\right] &= \mathbb{E}\left[\mathbb{1}(K(i,\varepsilon')) \sum_{j\notin \mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} \frac{w_j^2}{\ell_n\left(1 - \frac{\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} w_j}{\ell_n}\right)}\right] \\ &= \mathbb{E}\left[\mathbb{1}(K(i,\varepsilon')) \sum_{j\notin \mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} \frac{w_j^2}{\ell_n}\left(1 + \frac{\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} w_j}{\ell_n}\right)\right] + o\left(\frac{l}{n}\right).\end{aligned}$$

Let :

$$A = \mathbb{E}\left[\frac{\mathbb{1}(K(i,\varepsilon'))(\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} w_j)(\sum_{i=1}^n w_j^2)}{\ell_n^2}\right],$$

and

$$B = \mathbb{E}\left[\frac{\mathbb{1}(K(i,\varepsilon'))\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i} w_j^2}{\ell_n}\right].$$

By definition $\mathbb{1}(K(i,\varepsilon'))$, Lemma 54, and Conditions 4 :

$$A \leq \frac{l+r}{\ell_n}(1+o(1)),$$

and, by the Cauchy-Schwarz inequality :

$$\begin{aligned}-B &\leq -\frac{l\mathbb{P}(K(i,\varepsilon'))}{\ell_n}(1+o(1))C - \mathbb{E}\left[\frac{\mathbb{1}(K(i,\varepsilon'))(\sum_{j\in\tilde{\mathcal{V}}_i} w_j^2)}{\ell_n}\right] \\ &\leq -\frac{l\mathbb{P}(K(i,\varepsilon'))}{\ell_n}(1+o(1))C - \mathbb{E}\left[\mathbb{1}(K(i,\varepsilon')\frac{(\sum_{j\in\tilde{\mathcal{V}}_i} w_j)^2}{|\tilde{\mathcal{V}}_i|\ell_n}\right] \\ &= -\mathbb{P}(K(i,\varepsilon'))\frac{l+t}{\ell_n}C(1+o(1)),\end{aligned}$$

where :

$$t = \frac{(\frac{2(1-\varepsilon')f(i)\ell_n^{2/3}}{C} - \frac{\ell_n^{2/3}}{\sqrt{f(i)C}})^2}{\frac{2(1+\varepsilon')f(i)\ell_n^{2/3}}{C}}.$$

If we let $f(i)$ tend to ∞ then t will be equivalent to $r(1-\varepsilon')(1+\varepsilon')^{-1}$. Hence, when $f(i)$ is large enough, and by definition of $K(i,\varepsilon')$ we obtain :

$$\begin{aligned}-B &\leq -\frac{l\mathbb{P}(K(i,\varepsilon'))}{\ell_n}(1+o(1))C - \mathbb{E}\left[\frac{\mathbb{1}(K(i,\varepsilon'))(\sum_{j\in\tilde{\mathcal{V}}_i} w_j^2)}{\ell_n}\right] \\ &= -\frac{l+r}{\ell_n}(C-\varepsilon)(1+o(1)) + \frac{r(2\varepsilon'(1+\varepsilon')^{-1}+\varepsilon)}{\ell_n}.\end{aligned}$$

It follows that :

$$\begin{aligned}
\mathbb{E}\left[w_{\tilde{v}(l)}\mathbb{1}(K(i,\varepsilon'))\right] &= \mathbb{E}\left[\mathbb{1}(K(i,\varepsilon'))\sum_{j\notin \mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i}\frac{w_j^2}{\ell_n}\left(1+\frac{\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i}w_j}{\ell_n}\right)\right]+o\left(\frac{l}{n}\right)\\
&\leq \frac{\sum_{i=1}^n w_j^2}{\ell_n}+A-B\\
&-\mathbb{E}\left[\mathbb{1}(K(i,\varepsilon'))\frac{\left(\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i}w_j^2\right)\left(\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i}w_j\right)}{\ell_n^2}\right]+o\left(\frac{l}{n}\right)\\
&= \frac{\sum_{i=1}^n w_j^2}{\ell_n}+A-B+o\left(\mathbb{E}\left[\mathbb{1}(K(i,\varepsilon'))\frac{\sum_{j\in\mathbf{V}_{l-1}\cup\tilde{\mathcal{V}}_i}w_j^2}{\ell_n}\right]\right)\\
&+o\left(\frac{l+f(i)n^{2/3}}{n}\right)\\
&\leq 1+\frac{l+r}{\ell_n}(1-C+\varepsilon)+\frac{r(2\varepsilon'(1+\varepsilon')^{-1}+\varepsilon)}{\ell_n}+o\left(\frac{l+f(i)n^{2/3}}{n}\right).
\end{aligned}$$

□

Conditionally on $\tilde{\mathcal{V}}_i$, the graph $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$ is an inhomogeneous random graph. When $K(i,\varepsilon')$ holds with ε' small enough, then the same computation of Corollary 23.1 from Chapter 2 shows that exploration process associated to $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$ has negative increments in expectation. Since $K(i,\varepsilon')$ is measurable with regards to $\tilde{\mathcal{V}}_i$. The same analysis done in Chapter 2 to study the graph $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i)})$ can also be done to study $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$. The upper bounds that where proved for the former graph, can thus be proved similarly with slightly larger constants for the latter graph under the event $K(i,\varepsilon')$. We can then use a union bound by adding the probability of $K(i,\varepsilon')$ not holding. By this argument the following two theorems are true similarly to their counterparts for the graph $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i)})$. Theorem 59 is related to Theorem 3 from Chapter 2, and Theorem 60 is related to Theorem 4 from Chapter 2.

Theorem 59. *Suppose that Conditions 4 hold. Let $\varepsilon'>0$. For any $1\geq\varepsilon>0$ and for f large enough, consider the following event :*
All the connected components of $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$ have size smaller than

$$\frac{\ell_n^{2/3}}{Cf(i)^{1-\varepsilon}},$$

and weight smaller than

$$\frac{(1+\varepsilon')\ell_n^{2/3}}{Cf(i)^{1-\varepsilon}}.$$

There exists a positive constant $A>0$ that only depend on the distribution of W such that the probability of this event not happening is at most :

$$A\left(\exp\left(\frac{-f(i+1)^{\varepsilon}}{A}\right)+\exp\left(\frac{-\sqrt{f(i+1)}}{A}\right)+\exp\left(\frac{-n^{1/12}}{A}\right)\right).$$

Theorem 60. *Suppose that Conditions 1 hold. Let* Exc *be the maximal excess of the connected components of $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$. There exists a positive constant $A>0$ such that, for any $1\geq\varepsilon>0$ the probability of*

$$\text{Exc}\geq Af(i+1)^{\varepsilon},$$

is at most :

$$A\left(\exp\left(\frac{-f(i+1)^{\varepsilon/2}}{A}\right)+\exp\left(\frac{-n^{1/12}}{A}\right)\right).$$

Those theorems deal with the sizes, weights and excesses of the connected components of $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$. In order to prove bounds on the longest paths of those components, we still need to bound the height of some well-chosen spanning trees and use Lemma 53. This is why we bound the height of their exploration trees.

3.3.2 Coupling with Galton-Watson trees

In this part, $i \geq 1$ is fixed. The construction described bellow is done conditionally on $\tilde{\mathcal{V}}_i$, the vertices set of the largest connected component of $G(\mathbf{W}, p_{f(i)})$.

Let $\tilde{L}$ be the exploration process of $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$. And for $k \geq 1$, let $\tilde{c}(k)$ be the number of children of node $\tilde{v}(k)$ in $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$. If a connected component of $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$ is discovered at time k then its exploration process will be :

$$\tilde{L}'_k(0) = 1.$$

And for $i \geq 0$:

$$\tilde{L}'_k(i+1) = \tilde{L}'_k(i) + \tilde{c}(k+i-1) - 1.$$

This exploration process stops when it hits 0. For $(k,l) \in (\mathcal{V} \setminus \tilde{\mathcal{V}}_i)^2$, let $X(k,l,i+1)$ the random variable equal to 1 if there is an edge between nodes (k,l) in $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$ and 0 otherwise. Then for any $k \geq 1$, by definition of the exploration process, if a connected component is discovered at time k the number of children of node $\tilde{v}(k)$ has the the same distribution as :

$$\sum_{l \geq 1+k} X(\tilde{v}(k), \tilde{v}(l), i+1).$$

Generally, if node $\tilde{v}(k+i)$ is in the same connected component as $\tilde{v}(k)$, then the distribution of $\tilde{c}(k+i)$ will be the same as the distribution of :

$$\sum_{l \geq \tilde{L}'_k(i)+k+i} X(\tilde{v}(k+i), \tilde{v}(l), i+1).$$

For each $k \geq 0$, let $\tilde{\mathbb{T}}(k)$ be the tree with exploration process $\tilde{L}'_k$. If we can bound the height of $\tilde{\mathbb{T}}(k)$ for every $k \geq 1$, then we will have a bound on the height of the largest exploration tree in $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$. In order to do so, we create a Galton-Watson tree $\mathbb{T}(k)$ that contains a sub-tree with the same distribution as $\tilde{\mathbb{T}}(k)$ in the following way :

Conditionally on $\mathbf{V}_{k-1} = (\tilde{v}(1), \tilde{v}(2), ..., \tilde{v}(k-1))$ and $\tilde{\mathcal{V}}_i$, we construct a Galton-Watson tree $\mathbb{T}(k)$. Start by one node, and for each new node created follow these two steps. First give the node a label which is a random variable with the distribution of $\tilde{v}(k)$ independently of everything else. And then conditionally on the node label being r give it a random number of children which is a Poisson random variable of parameter :

$$w_r \left(\ell_n - \sum_{l \in \mathbf{V}_{k-1} \cup \tilde{\mathcal{V}}_i} w_l \right) p_{f(i+1)},$$

independently of everything else [3]. Remark that this process may not end, in that case we get an infinite tree. $\mathbb{T}(k)$ is thus a Galton-Watson tree with reproduction distribution $\mu(k, i+1)$, which is, conditionally on $(\mathbf{V}_k, \tilde{\mathcal{V}}_i)$, a Poisson distribution of parameter

$$w_{\tilde{v}(k)} \left(\ell_n - \sum_{l \in \mathbf{V}_{k-1} \cup \tilde{\mathcal{V}}_i} w_l \right) p_{f(i+1)}.$$

3. We abused notations by using the $\mathbf{V}_{k-1}$ as a set instead of introducing a new notation.

In order to have a clear order on the nodes of $\mathbb{T}(k)$, we sort them in a breadth-first search fashion while creating the tree. We start from one node, which is the root of $\mathbb{T}(k)$. We give it order 1, label $r(1)$. The root then has $b(1)$ children with the distributions described above. The children of the root are given orders in $(2, 3..., b(1)+1)$ randomly uniformly. Now do same for the the node with order 2 and order its children $(b(1)+2, b(1)+2, ..., b(1)+b(2)+1)$ randomly uniformly too. Continue in this fashion until there are no new nodes or the construction never ends. We call this the breadth-first order of $\mathbb{T}(k)$.

We construct a subtree $\mathbb{T}'(k)$ by pruning $\mathbb{T}(k)$ with the following algorithm. $\mathbb{L}$ is the list of nodes currently investigated by the algorithm. Initially $\mathbb{L}$ contains just the root :

1. At step 1 color the root in red, then add the children of the root to $\mathbb{L}$.
2. At step i consider the i-th node added to $\mathbb{L}$ following the breadth-first order. If there are no other red nodes with the same label yet, then color it in red and add its childer to $\mathbb{L}$.
3. Continue like this as long as $\mathbb{L}$ is not empty.

Since there is only a finite number of labels, this process will eventually end. The nodes are always removed from $\mathbb{L}$ following the breadth-first order. After finishing this procedure, we obtain a subtree of $\mathbb{T}(k)$ composed only of red nodes that we denote by $\mathbb{T}'(k)$. This fact is easily seen from the procedure since a node can only be red if its parent is also red. Moreover, we have the following lemma. This result was already shown in Norros and Reittu [2006] (proposition 3.1) and proven again in Bhamidi et al. [2010] for $\mathbb{T}'(1)$. The proof for general $\mathbb{T}'(k)$ is the same conditionally on $\mathbf{V}_{k-1}$, and so we refer to those articles for a proof.

Lemma 61. *For $k \geq 1$, conditionally on $(\tilde{v}(1), \tilde{v}(2), ..., \tilde{v}(k-1))$ and $\tilde{\mathcal{V}}_i$, the tree $\mathbb{T}'(k)$ has the same distribution as $\tilde{\mathbb{T}}(k)$.*

By Lemma 61, in order to bound the heights of the exploration trees of $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$ it is sufficient to bound the heights of the trees $\mathbb{T}'(k)$ for $k \geq 1$. Since the $\mathbb{T}'(k)$'s are substrees of the $\mathbb{T}(k)$'s, we will bound the heights of the later. The following lemma revolves around classical results on Galton-Watson trees.

Lemma 62. *Suppose that Conditions 4 hold. For $k \geq 1$. Let $H(\mathbb{T}(k))$ be the height of tree $\mathbb{T}(k)$. For $m \geq 0$, denote the event $\{H(\mathbb{T}(k)) \geq m\} \cap K(i, 1/100)$ by $\boldsymbol{Q}(k, m)$. If $k \leq \ell_n^{11/12}$ then :*

$$\mathbb{P}\left(\boldsymbol{Q}\left(k, \frac{2\ell_n^{1/3}}{\sqrt{f(i+1)}}\right)\right) \leq \frac{A\sqrt{f(i+1)}}{\ell_n^{1/3}} \exp\left(\frac{-Ck}{2\sqrt{f(i+1)}\ell_n^{2/3}} - \frac{\sqrt{f(i+1)}}{8}\right).$$

And if $k \geq \ell_n^{11/12}$ then :

$$\mathbb{P}\left(\boldsymbol{Q}\left(k, \frac{2\ell_n^{1/3}}{\sqrt{f(i+1)}}\right)\right) \leq \frac{A\sqrt{f(i+1)}}{\ell_n^{1/3}} \exp\left(\frac{-C\ell_n^{11/12}}{2\sqrt{f(i+1)}\ell_n^{2/3}} - \frac{\sqrt{f(i+1)}}{8}\right).$$

Proof. For simplicity suppose that m is even. Clearly :

$$\begin{aligned} \mathbb{P}(\mathbf{Q}(k,m)) &\leq \mathbb{E}[\mu(k, i+1)\mathbb{1}(K(i, 1/100))]\mathbb{P}(\mathbf{Q}(k, m-1))) \\ &\leq \mathbb{E}[\mu(k, i+1)\mathbb{1}(K(i, 1/100))]^{m/2}\mathbb{P}(\mathbf{Q}(k, m/2)) \end{aligned} \tag{3.3}$$

Moreover, $\mathbb{1}(K(i, 1/100))$ is measurable with regard to $\tilde{\mathcal{V}}_i$ by definition. This remark and Lemma

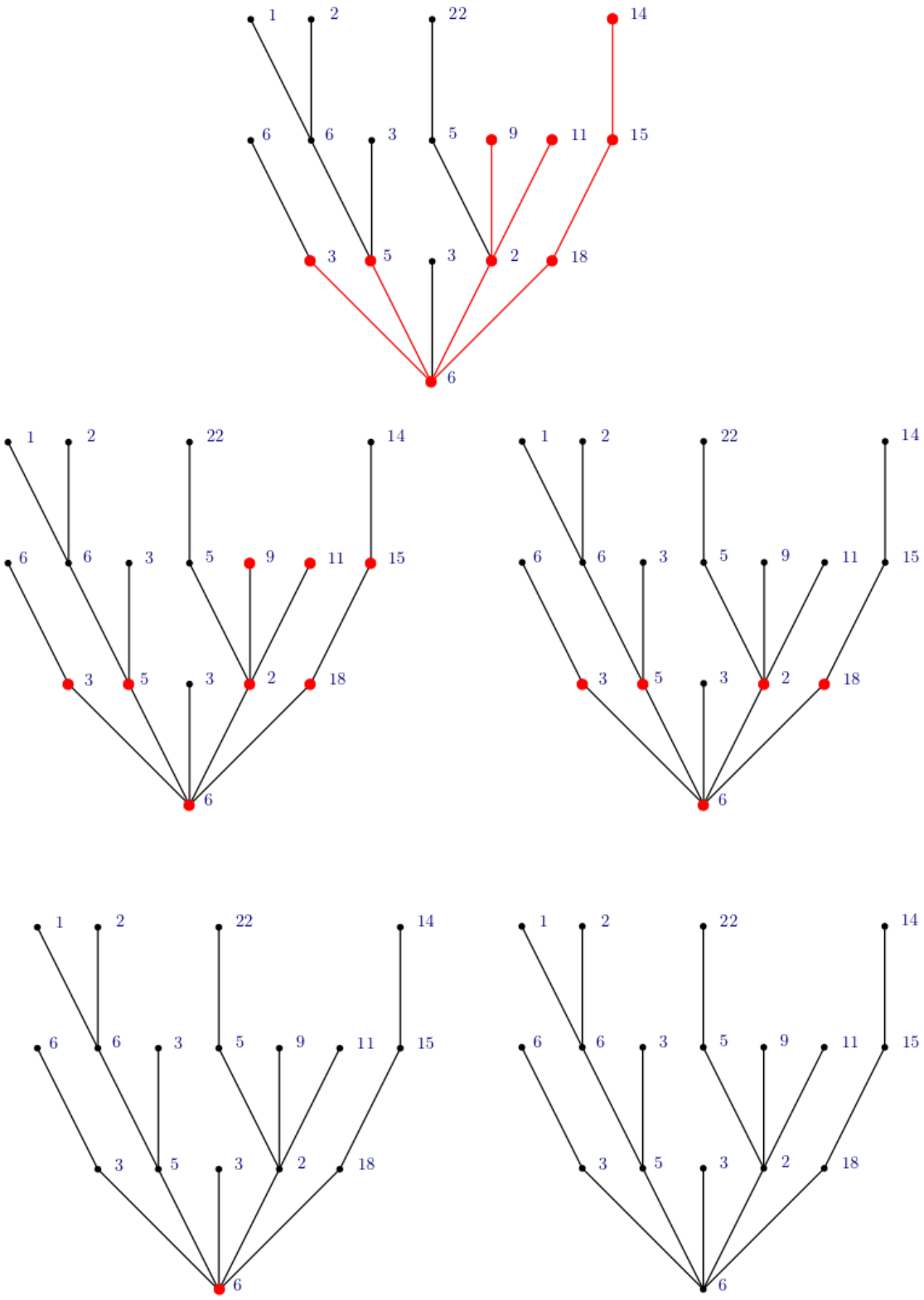


FIGURE 3.3 – An example of the colouring procedure. The breath-first order is implicitly induced by the drawing, nodes are ordered from bottom to top and from left to right. Each node has its label directly to its right. We only show the steps where a whole generation has been considered.

55 alongside the fact that for $x \in \mathbb{R}$, $1 - e^{-x} \leq x$ yield :

$$\begin{aligned}
&\mathbb{E}[\mu(v(k), i+1)\mathbb{1}(K(i,1/100))] \\
=&\mathbb{E}\left[\mathbb{E}\left[\mu(v(k), i+1)\mathbb{1}(K(i,1/100))\middle|(\tilde{\mathcal{V}}_i, \mathbf{V}_k)\right]\right] \\
\leq&\mathbb{E}\left[w_{\tilde{v}(k)}\left(\ell_n - \sum_{l\in\tilde{\mathcal{V}}_i} w_l\right) p_{f(i+1)}\mathbb{1}(K(i,1/100))\right] - \mathbb{E}\left[w_{\tilde{v}(k)}\left(\sum_{l\in\mathbf{V}_{k-1}} w_l\right) p_{f(i+1)}\mathbb{1}(K(i,1/100))\right] \\
\leq&\mathbb{E}\left[w_{\tilde{v}(k)}\left(\ell_n - \frac{99f(i)\ell_n^{2/3}}{50C}\right) p_{f(i+1)}\mathbb{1}(K(i,1/100))\right] - \frac{k}{2\ell_n}(1+o(1)).
\end{aligned}$$

For $k \leq \ell_n^{11/12}$ we have, if $C = 1$ then by Lemma :

$$\begin{aligned}
&\mathbb{E}[\mu(k, i+1)\mathbb{1}(K(i,1/100))] \\
&\leq (1+o(1))\left(1 + \frac{f(i+1)}{\ell_n^{1/3}} - \frac{99f(i)}{50\ell_n^{1/3}}\right) - \frac{k}{2\ell_n} + o\left(\frac{k + f(i)n^{2/3}}{n}\right) \\
&\leq 1 - \frac{k}{2\ell_n} - \frac{f(i)}{4\ell_n^{1/3}} + o\left(\frac{k + f(i)n^{2/3}}{n}\right).
\end{aligned} \tag{3.4}$$

If $C > 1$ and if we let $r = 99(50C)^{-1}f(i)\ell_n^{2/3}$ and let $\varepsilon = \min(1/100, (C-1)/2)$. Then by Lemma 58 applied with $\varepsilon > 0$ we have :

$$\begin{aligned}
&\mathbb{E}[\mu(k, i+1)\mathbb{1}(K(i,1/100))] \\
&\leq \left(1 + \frac{l+r}{\ell_n}(1 - C + \varepsilon) + \frac{r(2\varepsilon' + \varepsilon)}{\ell_n}\right)\left(1 + \frac{f(i+1)}{\ell_n^{1/3}} - \frac{99f(i)}{50C\ell_n^{1/3}}\right) - \frac{k}{2\ell_n} + o\left(\frac{k + f(i)n^{2/3}}{n}\right) \\
&\leq 1 + \frac{f(i)}{\ell_n^{1/3}}\left(\frac{99 - 99C + 4}{50C} + \frac{3}{2} - \frac{99}{50C}\right) - \frac{k}{2\ell_n} + o\left(\frac{k + f(i)n^{2/3}}{n}\right) \\
&\leq 1 - \frac{k}{2\ell_n} - \frac{f(i)}{4\ell_n^{1/3}} + o\left(\frac{k + f(i)n^{2/3}}{n}\right)
\end{aligned} \tag{3.5}$$

It is also well known for critical and thus also for sub-critical Galton-Watson trees (see for example Lemma 6.7 of Bhamidi et al. [2018]) that there exists a constant $A > 0$ such that for any $m \geq 0$:

$$\mathbb{P}(\mathbf{Q}(k, m/2)) \leq \frac{A}{m}. \tag{3.6}$$

Hence, by Equation (3.3), (3.4), (3.5), and (3.6), we have :

$$\begin{aligned}
\mathbb{P}\left(\mathbf{Q}\left(k, \frac{2\ell_n^{1/3}}{\sqrt{f(i+1)}}\right)\right) &\leq \frac{A\sqrt{f(i+1)}}{\ell_n^{1/3}}\mathbb{E}[\mu(k, i+1)\mathbb{1}(K(i,1/100))]^{\frac{\ell_n^{1/3}}{\sqrt{f(i+1)}}} \\
&= \frac{A\sqrt{f(i+1)}}{\ell_n^{1/3}}\exp\left(\frac{-k}{2\sqrt{f(i+1)}\ell_n^{2/3}} - \frac{\sqrt{f(i+1)}}{4} + o\left(\frac{k + f(i)n^{2/3}}{\sqrt{f(i+1)}n^{2/3}}\right)\right) \\
&\leq \frac{A\sqrt{f(i+1)}}{\ell_n^{1/3}}\exp\left(\frac{-k}{4\sqrt{f(i+1)}\ell_n^{2/3}} - \frac{\sqrt{f(i+1)}}{8}\right).
\end{aligned} \tag{3.7}$$

Moreover, for $k > \ell_n^{11/12}$, it is sufficient to show that for any $u \geq 1$:

$$\mathbb{E}[\mu(\tilde{v}(u), i+1)\mathbb{1}(K(i,1/100))] \leq \mathbb{E}[\mu(\tilde{v}(u+1), i+1)\mathbb{1}(K(i,1/100))],$$

This would prove that :

$$\mathbb{E}[\mu(\tilde{v}(k), i+1)\mathbb{1}(K(i,1/100))] \leq \mathbb{E}[\mu(\tilde{v}(\ell_n^{11/12}), i+1)\mathbb{1}(K(i,1/100))],$$

and then we can finish by using Equations (3.7). Notice that :

$$\begin{aligned}
&\mathbb{E}[\mu(\tilde{v}(u+1), i+1)\mathbb{1}(K(i,1/100))] \\
=&\mathbb{E}\left[\mathbb{E}\left[\mu(\tilde{v}(u+1), i+1)\mathbb{1}(K(i,1/100))\middle|(\tilde{\mathcal{V}}_i, \mathbf{V}_u)\right]\right] \\
\leq&\mathbb{E}\left[\mathbb{E}[w_{\tilde{v}(u+1)}|(\tilde{\mathcal{V}}_i, \mathbf{V}_u)]\left(\ell_n - \sum_{l\in\tilde{\mathcal{V}}_i\cup\mathbf{V}_u} w_l\right) p_{f(i+1)}\mathbb{1}(K(i,1/100))\right] \\
\leq&\mathbb{E}\left[\mathbb{E}[w_{\tilde{v}(u+1)}|(\tilde{\mathcal{V}}_i, \mathbf{V}_u)]\left(\ell_n - \sum_{l\in\tilde{\mathcal{V}}_i\cup\mathbf{V}_{u-1}} w_l\right) p_{f(i+1)}\mathbb{1}(K(i,1/100))\right] \\
=&\mathbb{E}\left[\mathbb{E}[w_{\tilde{v}(u+1)}|(\tilde{\mathcal{V}}_i, \mathbf{V}_{u-1})]\left(\ell_n - \sum_{l\in\tilde{\mathcal{V}}_i\cup\mathbf{V}_{u-1}} w_l\right) p_{f(i+1)}\mathbb{1}(K(i,1/100))\right].
\end{aligned}$$

It is thus sufficient to prove that :

$$\mathbb{E}[w_{\tilde{v}(u+1)}|(\tilde{\mathcal{V}}_i, \mathbf{V}_{u-1})] \leq \mathbb{E}[w_{\tilde{v}(u)}|(\tilde{\mathcal{V}}_i, \mathbf{V}_{u-1})],$$

and this is was already showed in the proof of Lemma 35 of Chapter 2 for similar weights. This finishes the proof. □

In order to bound the height of all the exploration trees at once, we need a bound on the number of connected components in $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$. The order of nodes in the exploration process of $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$ has the same distribution as that of $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i)})$. The exploration process of $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i)})$ is thus stochastically upper bounded by the exploration process of $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$. Since, by definiton, the number of connected components discovered in an interval of time of the BFW correspond to the number of new minimums attained by the exploration process in that same interval. It is clear by this argument that for any $m\in\mathbb{N}$ and $(t_1,t_2)\in\mathbb{N}^2$ such that $t_1\leq t_2$, the probability of discovering more than m connected components in the BFW of $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$ between times t_1 and t_2 is smaller than the probability of discovering more than m connected components in the BFW of $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i)})$.

Lemma 63. *Suppose that Conditions 4 hold. Let* $r_1 = \frac{\ell_n^{2/3}}{\sqrt{f}C}$. *Let* D_0 *be the event "The number of connected component discovered before time* r_1 *in the exploration process of* $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$ *is larger than* $\frac{\sqrt{f(i)}\ell_n^{1/3}}{2C}$."

For $j\geq 1$, *consider the interval*

$$I_j = \left[r_1 + \frac{(j^2-1)f(i+1)\ell_n^{2/3}}{C}, r_1 + \frac{((j+1)^2-1)f(i+1)\ell_n^{2/3}}{C}\right)$$

and let $\tilde{j}$ *be the greatest integer such that :*

$$\frac{(\tilde{j}^2-1)f(i+1)\ell_n^{2/3}}{C} \leq \ell_n^{11/12}.$$

For some $\tilde{j} > j \geq 1$, *let* D_j *be the event "There are more than* $100j^3 f(i)^2\ell_n^{1/3}$ *connected components discovered in the interval* I_j *in the exploration process of* $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$."

There exists a constant $A>0$ *such that :*

$$\mathbb{P}\left(\bigcup_{j=0}^{\tilde{j}}(D_j)\right) \leq A\exp\left(\frac{-\sqrt{f}}{A}\right) + A\exp\left(\frac{-n^{1/12}}{A}\right).$$

Proof. By the arguments given before, it is sufficient to show this result for the exploration process of $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i)})$. Recall that the exploration process can only go down by 1, so the number of connected components discovered in an interval of time is equal the number of times a new minimum was attained by the exploration process in that same interval of time. Hence, by Equation 73 in Chapter 2, the probability of discovering more than $\frac{\sqrt{f}\ell_n^{1/3}}{2C}$ connected components before time r_1 in the exploration process of $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i)})$ is at most :

$$A \exp\left(\frac{-\sqrt{f}}{A}\right). \tag{3.8}$$

And by Corollary 40.2 in Chapter 2 we obtain :

$$\mathbb{P}\left(\bigcup_{j=1}^{\tilde{j}-1}(D_j)\right) \leq A \exp\left(\frac{-\sqrt{f}}{A}\right) + A \exp\left(\frac{-n^{1/12}}{A}\right). \tag{3.9}$$

□

With Lemmas 62 and 63 we get bounds on the height of the exploration trees of $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$. On the other hand, Theorem 60 gives bounds on its excess. With this in hand, the following theorem combines those results in order to give an upper bound for the length of the longest path of $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$.

Theorem 64. *Suppose that Conditions 4 hold. There exists a constant $A > 0$ such that :*

$$\mathbb{P}\left(\mathrm{lp}(G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})) \geq \frac{A\ell_n^{1/3}}{f(i+1)^{1/4}}\right) \leq A\left(\exp\left(\frac{-f(i+1)^{1/8}}{A}\right)\right).$$

Proof. By Equation (3.2), there exists a constant $A > 0$ such that, the probability of the event $K(i, 1/100)$ not holding is at most :

$$A \exp\left(\frac{-f(i)}{A}\right). \tag{3.10}$$

When that event holds, it is sufficient to study the event $\mathbf{Q}\left(k, \frac{2\ell_n^{1/3}}{\sqrt{f(i+1)}}\right)$ and bound the surplus and then use the union bound. Recall the definition of the events $(D_j)_{j<\tilde{j}}$ and the intervals $(I_j)_{j<\tilde{j}}$ from Lemma 63. If all those events are verified, then by Lemmas 62 and 63, in each interval I_j, the probability that there exists an exploration tree with height larger than $\frac{2\ell_n^{1/3}}{\sqrt{f(i+1)}}$ is at most :

$$B_j = 100j^3 f(i)^2 \ell_n^{1/3} \frac{A\sqrt{f(i+1)}}{\ell_n^{1/3}} \exp\left(\frac{-C(j^2-1)f(i+1)\ell_n^{2/3}}{2C\sqrt{f(i+1)}\ell_n^{2/3}} - \frac{\sqrt{f(i+1)}}{8}\right).$$

Recall that $r_1 = \frac{\ell_n^{2/3}}{\sqrt{f}C}$. The probability that there exists an exploration tree with height larger than $\frac{2\ell_n^{1/3}}{\sqrt{f(i+1)}}$ before time r_1 is at most :

$$B_0 = \frac{\sqrt{f(i)}\ell_n^{1/3}}{2C} \frac{A\sqrt{f(i+1)}}{\ell_n^{1/3}} \exp\left(-\frac{\sqrt{f(i+1)}}{8}\right).$$

Hence, for it to be an exploration tree with height larger than $\frac{2\ell_n^{1/3}}{\sqrt{f(i+1)}}$, either one of the events $(D_j)_{0\leq j\leq\tilde{j}}$ does not hold or $K(i, 1/100)$ does not hold, or one of the event $\mathbf{Q}\left(k, \frac{2\ell_n^{1/3}}{\sqrt{f(i+1)}}\right)$ does

not hold or there is a connected component discovered after time $\ell_n^{11/12}$ with an exploration tree with height larger than $\frac{2\ell_n^{1/3}}{\sqrt{f(i+1)}}$. There are less than n connected component discovered after time $\ell_n^{11/12}$. Hence, by Equation (3.10) and Lemmas 62 and 63, there exists constants $A > 0$ and $A' > 0$ such that, the probability that there exists a connected component with an exploration tree of height larger than $\frac{2\ell_n^{1/3}}{\sqrt{f(i+1)}}$ is at most :

$$A\exp\left(\frac{-f(i)}{A}\right) + \sum_{j=0}^{\tilde{j}} B_j + n\frac{A\sqrt{f(i+1)}}{\ell_n^{1/3}}\exp\left(\frac{-C\ell_n^{11/12}}{2\sqrt{f(i+1)}\ell_n^{2/3}} - \frac{\sqrt{f(i+1)}}{8}\right) + \mathbb{P}\left(\bigcup_{j=1}^{\tilde{j}}(D_j)\right)$$
$$\leq A'\exp\left(\frac{-\sqrt{f(i)}}{A'}\right) + A'\exp\left(\frac{-n^{1/12}}{A'}\right). \tag{3.11}$$

Recall that Exc is the maximal excess of the connected components of $G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})$. By Theorem 60, there exists a positive constant $A'' > 0$ such that, for any $1 > \varepsilon > 0$:

$$\mathbb{P}\left(\text{Exc} \geq A''f(i+1)^{\varepsilon}\right) \leq A''\left(\exp\left(\frac{-f(i+1)^{\varepsilon/2}}{A''}\right) + \exp\left(\frac{-n^{1/12}}{A''}\right)\right). \tag{3.12}$$

Taking $\varepsilon = 1/4$ in Equation (3.12) and using Lemma 53 with Equations (3.11) and (3.12) then yields :

$$\mathbb{P}\left(\text{lp}(G(\mathcal{V}\setminus\tilde{\mathcal{V}}_i, p_{f(i+1)})) \geq \frac{A\ell_n^{1/3}}{f(i+1)^{1/4}}\right) \leq A\left(\exp\left(\frac{-f(i+1)^{1/8}}{A}\right) + \exp\left(\frac{-n^{1/12}}{A}\right)\right)$$
$$\leq A'\left(\exp\left(\frac{-f(i+1)^{1/8}}{A'}\right)\right).$$

□

3.4 Bounding the length of the longest path of the giant component

3.4.1 A new pruning procedure for the giant component

Take $f \geq f(0)$, we consider the exploration process until time $t = \frac{5f\ell_n^{2/3}}{2C}$. We know by Chapter 2 that the giant component is fully discovered before t with some high probability. In order to bound the length of the longest path of the giant component, a natural idea would be to couple the trees discovered before times t with a sequence of i.i.d. Galton-Watson trees with the distribution of $\tilde{\mathbb{T}}(1)$ by using the same colouring of the nodes described in Sub-section 3.3.2. However, $\tilde{\mathbb{T}}(1)$ is slightly supercritical. Thus, some of the Galton-Watson trees we couple with will have infinite size. To solve this problem we first add i.i.d. cuts with probability $\frac{2f}{\ell_n^{1/3}}$ to the edges of $G(\mathbf{W}, p_f)$. Then we couple the new smaller exploration trees obtained with i.i.d. Galton-Watson trees with the the distribution of :

$$\sum_{l=1}^{n}\tilde{X}(v(1), l, i),$$

where $\tilde{X}(k, l, i)$ is the product of two i.i.d. Bernoulli random variables. The first one corresponds to "having an edge" and has parameter :

$$\left(1 - \exp(-w_k w_l p_{f(i+1)})\right),$$

the second one corresponds to "not having a cut" and has parameter :

$$\left(1-\frac{2f}{\ell_n^{1/3}}\right).$$

We then apply the procedure of Sub-section 3.3.2 on those Galton-Watson trees. This yields a new two steps procedure, first pruning then coloring, which gives a forest of i.i.d. slightly sub-critical Galton-Watson trees. Moreover, the sum of heights of those Galton-Watson trees is an upper bound of the height of any exploration tree discovered before time t.

We start by bounding the number of Galton-Watson trees obtained by this construction. A new tree is created at each cut. This means that the total number of Galton-Watson trees after pruning is equal to the number of cuts plus the number of initial trees. However, if we are only interested in the height of one tree. Only the number of cuts is of interest. Hence, even though we have a bound on the initial number of trees before time t by Chapter 2. We, in fact only need a bound on the number of cuts. This is given in the following lemma.

Lemma 65. *The number of cuts in the pruning procedure is less than* $\frac{6f^2\ell_n^{1/3}}{C}$ *with probability at least :*

$$1-\exp\left(-2Cf^3\right).$$

Proof. Because a tree of size m has $m-1$ edges, there are less than t edges in the exploration trees before time t. Let J be the number of edge cuts, then :

$$\mathbb{E}[J]\leq\frac{5f^2\ell_n^{1/3}}{C}.$$

Hence, by Hoeffding's inequality :

$$\begin{aligned}\mathbb{P}\left(J\geq\frac{6f^2\ell_n^{1/3}}{C}\right)&\leq\exp\left(\frac{-(f^2\ell_n^{1/3})^2}{C^2t}\right)\\&\leq\exp\left(\frac{-f^3}{5C}\right).\end{aligned}$$

□

Since $C=\frac{\mathbb{E}[W^3]}{\mathbb{E}[W]}\geq 1$, this lemma shows that w.h.p the number of Galton-Watson trees obtained by the pruning then coupling procedure is less than $6f^2\ell_n^{1/3}+1\leq 7f^2\ell_n^{1/3}$.

3.4.2 Bounding the longest path

With this in hand, we can prove a bound on the height of the exploration tree of the largest component. We denote the Galton-Watson trees constructed by pruning by $(\mathbb{T}^{\top}(k))_{k\geq 1}$. We use a simple union bound. Either there is some tree discovered before time t with large height, or the largest connected component is discovered after time t. We already have a bound on the latter event, and we bound the former event to get the following theorem. Bounding the height of an exploration tree discovered before time t by the sum of heights of all the $(\mathbb{T}^{\top}(k))_{k\geq 1}$'s will not yield a good enough upper bound. But this first step will allow us to use a recursive argument which yields the tight bound presented here.

Theorem 66. *Suppose that Conditions 4 hold. There exist constants* $A'>0$ *and* $A>0$ *such that the exploration tree of the largest component of* $G(\boldsymbol{W},p_f)$*, denoted by* $H(\boldsymbol{W},p_f)$*, has height smaller than* $A'f^2\ell_n^{1/3}$ *with probability at least :*

$$1-A\exp\left(\frac{-f}{A}\right).$$

Proof. Let $\mathcal{B}_1$ be the event : the exploration of the largest component of $G(\mathbf{W}, p_f)$ is not already done at time t. By Theorem 1 in Chapter 2 :

$$\mathbb{P}(\mathcal{B}_1) \leq A \exp\left(\frac{-f}{A}\right), \tag{3.13}$$

with $A > 0$ a large constant.

Let R be the random number of cuts obtained after pruning the exploration trees up to time t. And let $\mathcal{B}_2$ be the event : R is larger than $7f^2\ell_n^{1/3}$. By Lemma 65 :

$$\mathbb{P}(\mathcal{B}_2) \leq A \exp\left(\frac{-f^3}{A}\right). \tag{3.14}$$

If the largest component is discovered before time t, the height of its exploration tree $\mathrm{ht}(H(\mathbf{W}, p_f))$ will be smaller than the sum of heights of all the trees obtained after pruning. By exactly the same method we used in the proof of Lemma 62, there exists $A > 0$ such that for any $r > 0$:

$$\mathbb{P}(\mathrm{ht}(\mathbb{T}^\top(1))) \geq r) \leq \frac{A}{r} \exp\left(\frac{-fr}{2\ell_n^{1/3}}\right), \tag{3.15}$$

and by the union bound :

$$\mathbb{P}\left(\sup_{k \leq 7f^2\ell_n^{1/3}} \mathrm{ht}(\mathbb{T}^\top(k))) \geq \ell_n^{1/3}\right) \leq A' \exp\left(\frac{-f}{A'}\right), \tag{3.16}$$

with $A' > 0$ a large constant. Denote the event

$$\left\{\sup_{k \leq 7f^2\ell_n^{1/3}} \mathrm{ht}(\mathbb{T}^\top(k))) \geq \ell_n^{1/3}\right\}$$

by $\mathcal{B}_3$. Using Equations (3.13) and (3.14) alongside Equation (3.16) shows that :

$$\mathbb{P}(\mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3) \leq A \exp\left(\frac{-f}{A}\right). \tag{3.17}$$

Let $\tilde{\mathcal{B}} = \bar{\mathcal{B}}_1 \cap \bar{\mathcal{B}}_2 \cap \bar{\mathcal{B}}_3$. If $\tilde{\mathcal{B}}$ holds, then the height of the exploration tree of the largest component is smaller than the sum of heights of the trees $(\mathbb{T}^\top(k))_{k \leq 7f^2\ell_n^{1/3}}$. By a small computation, since $f = o(n^{1/3})$, we have by Equation (3.15) :

$$\begin{aligned}
\mathbb{E}[\mathrm{ht}(\mathbb{T}^\top(1))] &= \sum_{r=0}^{\infty} \mathbb{P}(\mathrm{ht}(\mathbb{T}^\top(1)) \geq r) \\
&\leq 1 + \sum_{r=1}^{\infty} \frac{A}{r} \exp\left(\frac{-fr}{2\ell_n^{1/3}}\right) \\
&\leq 1 + \sum_{r=1}^{\frac{\ell_n^{1/3}}{f}} \frac{A}{r} + \sum_{r=\frac{\ell_n^{1/3}}{f}+1}^{\infty} \frac{Af}{\ell_n^{1/3}} \exp\left(\frac{-fr}{2\ell_n^{1/3}}\right) \\
&\leq A' \log\left(\frac{\ell_n^{1/3}}{f}\right) + \frac{A'f}{\ell_n^{1/3}\left(1 - \exp\left(\frac{-f}{2\ell_n^{1/3}}\right)\right)} \\
&= O\left(\log\left(\frac{n^{1/3}}{f}\right)\right)
\end{aligned} \tag{3.18}$$

and similarily

$$\begin{aligned}
\mathbb{E}[\mathrm{ht}(\mathbb{T}^{\top}(1))^2] &= \sum_{r=0}^{\infty} r^2 \mathbb{P}(\mathrm{ht}(\mathbb{T}^{\top}(1)) = r) \\
&\leq \sum_{r=0}^{\infty} r(r+1)\mathbb{P}(\mathrm{ht}(\mathbb{T}^{\top}(1)) = r) \\
&= \sum_{r=0}^{\infty} 2\mathbb{P}(\mathrm{ht}(\mathbb{T}^{\top}(1)) = r) \left(\sum_{j=0}^{r} j \right) \\
&= 2\sum_{j=0}^{\infty} j \left(\sum_{r=j}^{\infty} \mathbb{P}(\mathrm{ht}(\mathbb{T}^{\top}(1)) = r) \right) \\
&= 2\sum_{j=0}^{\infty} j\mathbb{P}(\mathrm{ht}(\mathbb{T}^{\top}(1)) \geq j) \\
&= O\left(\frac{n^{1/3}}{f} \right).
\end{aligned} \tag{3.19}$$

Hence, by Bernstein's inequality (Bernstein [1924]) for $A' > 0$ large enough, there exists a large constant $A > 0$ such that :

$$\begin{aligned}
&\mathbb{P}\left(\sum_{k=1}^{R} \mathrm{ht}(\mathbb{T}^{\top}(k)) \mathbb{1}(\tilde{\mathcal{B}}) \geq A' f^2 \ell_n^{1/3} \log\left(\frac{n^{1/3}}{f} \right) \right) \\
&\leq \mathbb{P}\left(\sum_{k=1}^{7f^2\ell_n^{1/3}} \mathrm{ht}(\mathbb{T}^{\top}(k))) \mathbb{1}(\mathrm{ht}(\mathbb{T}^{\top}(k)) \leq \ell_n^{1/3}) \geq A' f^2 \ell_n^{1/3} \log\left(\frac{n^{1/3}}{f} \right) \right) \\
&\leq A \exp\left(\frac{-\log\left(\frac{n^{1/3}}{f} \right) f^2}{A} \right).
\end{aligned} \tag{3.20}$$

Suppose now that the height of the exploration tree of any connected component discovered before time t is at most $A' f^2 \ell_n^{1/3} \log\left(\frac{n^{1/3}}{f} \right)$, with A' the constant of Equation (3.20). Denote that event by $\mathcal{S}$. In that case, the number of cuts (denoted by $\mathcal{P}$) in one of the longuest paths in the exploration tree of $H(\mathbf{W}, p_f)$ verifies by Bernstein's inequality (Bernstein [1924]) :

$$\mathbb{P}\left(\mathcal{P}\mathbb{1}(\mathcal{S}) \geq 5A' f^3 \log\left(\frac{n^{1/3}}{f} \right) \right) \leq A'' \exp\left(\frac{-\log\left(\frac{n^{1/3}}{f} \right) f^2}{A''} \right), \tag{3.21}$$

It is well known that for n large enough, for any $7f^2\ell_n^{1/3} \geq u \geq 1$:

$$\binom{7f^2\ell_n^{1/3}}{u} \leq \frac{(7f^2\ell_n^{1/3})^u}{u!} \leq \left(\frac{7f^2\ell_n^{1/3} e}{u} \right)^u$$

With this we can bound the height of the u'th highest tree, denoted by $\mathbb{T}_u^{\top}$. By Equation (3.15)

there exist some large constants $A > 0$ and $A' > 0$ such that :

$$\begin{aligned}\mathbb{P}\left(\mathrm{ht}(\mathbb{T}_u^{\top})\mathbb{1}(\tilde{\mathcal{B}}) \geq \frac{\ell_n^{1/3}}{\sqrt{u}}\right) &\leq \left(\frac{A\sqrt{u}}{\ell_n^{1/3}}\exp\left(\frac{-f}{2\sqrt{u}}\right)\right)^u \binom{7f^2\ell_n^{1/3}}{u} \\ &\leq \left(\frac{A\sqrt{u}}{\ell_n^{1/3}}\exp\left(\frac{-f}{2\sqrt{u}}\right)\right)^u \left(\frac{7ef^2\ell_n^{1/3}}{u}\right)^u \\ &\leq \left(\frac{7Aef^2}{\sqrt{u}}\exp\left(\frac{-f}{2\sqrt{u}}\right)\right)^u \\ &\leq A'\exp\left(\frac{-\sqrt{u}f - u\log\left(\frac{u}{f}\right)}{A'}\right).\end{aligned} \tag{3.22}$$

Let $m = 5A'f^3\log\left(\frac{n^{1/3}}{f}\right)$. If $\mathcal{P} \leq m$, then the height of the largest component is smaller than the sum of heights of the m highest Galton-Watson trees $(\mathbb{T}_u^{\top})_{u\leq m}$. By the union bound, using Equations (3.17),(3.20), (3.21), and (3.22), there exist $A > 0$ and $A' > 0$ large enough such that :

$$\begin{aligned}&\mathbb{P}\left(\sum_{u=1}^{m}\mathrm{ht}(\mathbb{T}_u^{\top}) \geq \ell_n^{1/3}\left(\sum_{u=1}^{m}\frac{1}{\sqrt{u}}\right)\right) \\ &\leq A\exp\left(\frac{-f}{A}\right) + A\exp\left(\frac{-\log\left(\frac{n^{1/3}}{f}\right)f^2}{A}\right) + \sum_{u=1}^{m}A\exp\left(\frac{-\sqrt{u}f - u\log\left(\frac{u}{f}\right)}{A}\right) \\ &\leq A'\exp\left(\frac{-f}{A'}\right) + A\exp\left(\frac{-\log\left(\frac{n^{1/3}}{f}\right)f^2}{A}\right).\end{aligned} \tag{3.23}$$

Moreover, by comparison with an integral :

$$\sum_{u=1}^{m}\frac{1}{\sqrt{u}} \leq 2\sqrt{m}.$$

This shows that the height of the largest component is in fact smaller than $2\sqrt{m}\ell_n^{1/3}$ w.h.p. We can repeat the same argument as before to show that the number of cuts $\mathcal{P}$ in the largest path verifies :

$$\mathbb{P}\left(\mathcal{P} \geq 9f\sqrt{m}\right) \leq A'\exp\left(\frac{-f}{A'}\right) + A\exp\left(\frac{-\log\left(\frac{n^{1/3}}{f}\right)f^2}{A}\right) + A\exp\left(\frac{-\sqrt{m}}{A}\right).$$

By repeating the same arguments recursively, we get for any $l \geq 0$:

$$\mathbb{P}\left(\mathrm{ht}(H(\mathbf{W}, p_f)) \geq \ell_n^{1/3}U_{l+1}\right) \leq A\exp\left(\frac{-f}{A}\right) + \sum_{u=0}^{t}A\exp\left(\frac{-U_l}{A}\right), \tag{3.24}$$

where :

$$U_0 = m,$$

and :

$$U_{l+1} = 9f\sqrt{U_l}.$$

The sequence $(U_l)_{l\geq 0}$ converges to $81f^2$. If we define $R_l = U_l/(81f^2)$, then :

$$R_{l+1} = \sqrt{R_l}.$$

It is clear that for $l_0 = \lceil \log(\ln(R_0)) \rceil$ we have $1 < R_{l_0} \leq e$. Hence :

$$\begin{aligned}
\mathbb{P}\left(\mathrm{ht}(H(\mathbf{W}, p_f)) \geq \ell_n^{1/3} 81 e f^2\right) &\leq A \exp\left(\frac{-f}{A}\right) + \sum_{l=0}^{l_0} A \exp\left(\frac{-U_l}{A}\right) \\
&\leq A \exp\left(\frac{-f}{A}\right) + \sum_{l=0}^{l_0} A \exp\left(\frac{-81 f^2 R_{l_0}^{2^l}}{A}\right) \\
&\leq A \exp\left(\frac{-f}{A}\right) + \sum_{l=0}^{l_0} A' \exp\left(\frac{-81 f^2 (l+1) R_{l_0}}{A'}\right) \\
&\leq A \exp\left(\frac{-f}{A}\right) + A'' \exp\left(\frac{-f^2}{A''}\right),
\end{aligned} \tag{3.25}$$

which finishes the proof. □

With a little more work, we could show that the bound given above is in fact tight, in the sense that for any $\varepsilon < 1$ the height of the largest component will be larger than $\frac{f^{1+\varepsilon} \ell_n^{1/3}}{A}$ w.h.p. Finally by Theorem 52, Lemma 53 and Theorem 66 we get :

Theorem 67. *Suppose that Conditions 4 hold. There exist constants $A' > 0$ and $A > 0$ such that :*

$$\mathbb{P}(\mathrm{lp}(H(\boldsymbol{W}, p_f)) \geq A' f^5 \ell_n^{1/3}) \leq A \exp\left(\frac{-f^{1/8}}{A}\right).$$

3.5 Proofs of Theorems 43, 45, 48 and 49

The different statements of Theorem 48 were all proved in the previous sections. The statements about the weights are shown in Theorem 52, and the statement about the longest path of small components is a direct corollary of Theorem 64.

Now we finish the proof of Theorem 49. recall that $(p_{f(i)})_{i \geq 0}$ is a sequence such that $f(0) = F$ the large constant defined at the end of Subsection 2.1, and $f(i+1) = \frac{3}{2} f(i)$ for any $i \geq 0$. We stop at $f(t_n)$ the smallest element larger than $f'_n = \frac{\ell_n^{1/3}}{\log(n)}$. For $0 \leq i \leq t_n$ we define the following events, $A > 0$ is a large constant :

1. $E_1(i)$ is the event where $H(\mathbf{W}, p_{f(i)})$ has size between $\frac{3 f(i) \ell_n^{2/3}}{2C}$ and $\frac{5 f(i) \ell_n^{2/3}}{2C}$.
2. $E_2(i)$ is the event where the longest path of $H(\mathbf{W}, p_{f(i)})$ has length at most $A f(i)^5 \ell_n^{1/3}$.
3. $E_3(i)$ is the event where every connected component of $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$ has size at most $\frac{A \ell_n^{2/3}}{f(i)^{3/4}}$, and longest path at most $\frac{A \ell_n^{1/3}}{f(i)^{1/4}}$.

Notice that $j \leq 2 \log(n)$. Let r be the smallest value such that $E_1(i)$, $E_2(i)$, and $E_3(i)$ hold for every $r \leq l \leq j$. When those events hold it is clear that $\mathrm{diam}(\mathcal{T}(\mathbf{W}, p_{f(j)}))$ is reached on the minimum spanning tree of $H(\mathbf{W}, p_{f(j)})$. By Lemma 50 and a simple computation :

$$\begin{aligned}
\mathrm{diam}(\mathcal{T}(\mathbf{W}, p_{f(j)})) &\leq \mathrm{diam}(\mathcal{T}(\mathbf{W}, p_{f(r)})) + \sum_{k=r}^{j} \frac{3 A \ell_n^{1/3}}{f(k)^{1/4}} \\
&\leq \mathrm{diam}(\mathcal{T}(\mathbf{W}, p_{f(r)})) + \frac{A' \ell_n^{1/3}}{f(r)^{1/4}} \\
&\leq A'' f(r)^5 \ell_n^{1/3}.
\end{aligned}$$

If $r = i+1$ then one of the events $E_1(i)$, $E_2(i)$, or $E_3(i)$ does not hold, because if not we would have $r \leq i$. By Theorems 52, 59, 60, 64 and Theorem 67, the probability of one of those events

not happening is at most :

$$\mathbb{P}(r = i+1) \leq A \exp\left(\frac{-f(i)^{1/8}}{A}\right).$$

Combining the two inequalities above yields :

$$\begin{aligned}\mathbb{E}\left[\mathrm{diam}(\mathcal{T}(\mathbf{W}, p_{f(j)}))\right] &\leq \sum_{i=1}^{j-1} A'' f(i+1)^5 \ell_n^{1/3} \mathbb{P}(r = i+1) \\ &\leq \sum_{i=1}^{j-1} A'' f(i+1)^5 \ell_n^{1/3} A \exp\left(\frac{-f(i)^{1/8}}{A}\right) \\ &= O(n^{1/3}).\end{aligned}$$

The lower bound was provided in Theorem 47. We move now to Theorem 43.

Proof of Theorem 43. Since the diameter is an upper bound of typical distances, by Theorem 42 we already have :

$$\mathbb{E}[d(U_{f_n}, V_{f_n})] = O(n^{1/3}).$$

We only need to show now that there exist some constants $\varepsilon > 0$ and $\varepsilon' > 0$ such that typical distances in the largest tree in $\mathcal{T}(\mathbf{W}, \tilde{p}_n)$ are larger than $\varepsilon n^{1/3}$ with probability at least ε'. Denote that tree by $B(\tilde{p}_n)$. And more generally for any $p > 0$ denote the largest tree in $\mathcal{T}(\mathbf{W}, p)$ by $B(p)$. Let $f > 0$ be a large constant, and $p_f = \frac{1}{\ell_n} + \frac{f}{\ell_n^{4/3}}$. For n large enough $p_f < \tilde{p}_n$. For a constant $\varepsilon > 0$, consider the following events :

- $\mathcal{A}_1$ is the event where $B(p_f)$ is a sub-tree of $B(\tilde{p}_n)$, and the size and weight of $B(p_f)$ are smaller than $3f\ell_n^{2/3}$ for every n large enough.
- $\mathcal{A}_2$ is the event where the size and weight of $B(\tilde{p}_n)$ is larger than εn for every n large enough.
- $\mathcal{A}_3$ is the event where there exist two sub-trees of $B(p_f)$, that we denote by $B_1(p_f)$ and $B_2(p_f)$, such that the minimal distance in $B(p_f)$ between a node in $B_1(p_f)$ and a node in $B_2(p_f)$ is larger than $\varepsilon n^{1/3}$. Moreover, the sizes and weights of $B_1(p_f)$ and $B_2(p_f)$ are larger than $\varepsilon n^{2/3}$ for every n large enough.

Then by Theorems 52 and 44, for any $\varepsilon_1 > 0$, if f is large enough then $\mathcal{A}_1$ holds with probability larger than $1 - \varepsilon_1$. Moreover, by Theorem 44, there exists $\varepsilon > 0$ such that A_2 holds with high probability. And also, Broutin et al. [2020] show the convergence of the largest component of $G(\mathbf{W}, p_f)$ with graph distances scaled by $n^{1/3}$ and sizes scaled by $n^{2/3}$ in the Gromov-Hausdorff-Prokhorov topology to a continuous compact connected graph. We refer to their article for a detailed analysis of this general result. Here we just emphasize the fact that their limiting object is a non-trivial compact metric space which containse two subspaces of mass $\zeta > 0$ and at distance at least ζ with positive probability $\zeta' > 0$. This result directly implies the existence of $\varepsilon_3 > 0$ such that A_3 holds with probability at least ε_3.

Let $\mathcal{A}(\varepsilon) = \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3$. By the discussion above, there exists two constants $\varepsilon > 0$ and $\varepsilon_4 > 0$ such that, for a constant $f > 0$ large enough, $\mathcal{A}(\varepsilon)$ holds with probability at least ε_4. This shows that there exists some $\varepsilon_5 > 0$ such that, by denoting the total weight of a graph G by $W(G)$, we have [4] :

$$\mathbb{E}\left[W\left(B_1(p_f)\right)\right] \geq \varepsilon_5 \mathbb{E}\left[W\left(B(p_f)\right)\right]. \tag{3.26}$$

For $\tilde{p}_n \geq p \geq p_f$, let $B_1(\tilde{p}_n)$ be the sub-tree of $B(\tilde{p}_n)$ obtained by applying Kruskal's algorithm to $B_1(p_f)$ between p_f and p. And define $B_2(\tilde{p}_n)$ similarly. Clearly the distance between $B_1(\tilde{p}_n)$ and $B_2(\tilde{p}_n)$ is larger than $\varepsilon_3 n^{1/3}$. It is thus sufficient to show that the sizes of $B_1(\tilde{p}_n)$ and $B_2(\tilde{p}_n)$ are on expectation proportional to n. By symmetry it is sufficient to show this for $B_1(\tilde{p}_n)$.

4. Here we just set the weights and sizes to 0 if $B_1(p_f)$ or $B_2(p_f)$ do not exist

Let $(r_1, r_2, r_3, ...)$ be the sequence of increasing values between p_f and $\tilde{p}_n$ at which a new edge is added by Kruskal's algorithm to $(B(r_i^-))_{i\geq 1}$. At each value r_i, an edge is created between $B(r_i^-)$ and the rest of $\mathcal{T}(\mathbf{W}, r_i^-)$. And by the properties of exponential random variables, the probability that such an edge gets connected to $B_1(r_i^-)$ is equal to

$$\frac{W(B_1(r_i^-))}{W(B(r_i^-))},$$

Hence the process of weights

$$(W(B_1(r_i^-)), W(B(r_i)) - W(B_1(r_i^-)))_{i\geq 1},$$

corresponds to a Polya Urn where a ball with of random weight is added at each step, it is easy to check by induction that if at step i, there exists $\varepsilon > 0$ such that :

$$\mathbb{E}[W(B_1(r_i^-))] = \varepsilon \mathbb{E}[W(B(r_i^-))],$$

Then at step $i+1$, a connected component $C(i)$ is added and we get :

$$\begin{aligned}\mathbb{E}[W(B_1(r_{i+1}^-))] &= \varepsilon \mathbb{E}[W(B(r_i^-))] + \varepsilon \mathbb{E}[W(C(i))] \\ &= \varepsilon \mathbb{E}[W(B(r_{i+1}^-)],\end{aligned}$$

with the same ε of step i. We know that (for instance by Theorem 44), there exists $\varepsilon_8 > 0$ such that for n large enough

$$\mathbb{E}\left[W\left(B\left(\tilde{p}_n\right)\right)\right] \geq \varepsilon_8 n.$$

Given that the event $\mathcal{A}(\varepsilon)$ holds with positive probability, this, alongside Equation 3.26, shows that there exists $\varepsilon_6 > 0$ such that :

$$\mathbb{E}\left[W\left(B_1\left(\tilde{p}_n\right)\right)\right] \geq \varepsilon_6 n. \tag{3.27}$$

And the same also holds for $B_2\left(\tilde{p}_n\right)$. Finally, we use this lower bound on the weight in order to get a lower bound on what we actually need, the size.

Recall that the node weights depend implicitly on n. Suppose that there exists an $N > 0$ and subsets $(\mathcal{S}_n)_{n\geq N}$ of the nodes such that :

$$\sum_{k\in\mathcal{S}_n} w_k \geq \varepsilon_6 n,$$

for every $n \geq N$, but the sizes of those sets verify $|\mathcal{S}_n| = o(n)$. Then by Lemma 7 :

$$\sum_{k\in\mathcal{S}_n} w_k \leq \sum_{k=1}^{|\mathcal{S}_n|} w_k = o(n).$$

Hence no such sets exist, and Equation 3.27 implies the existence of ε_7 such that :

$$\mathbb{E}\left[\left|B_1\left(\tilde{p}_n\right)\right|\right] \geq \varepsilon_7 n,$$

for any n large enough. We have thus shown that there exist two sub-trees, $B_1\left(\tilde{p}_n\right)$ and $B_2\left(\tilde{p}_n\right)$, of $B\left(\tilde{p}_n\right)$ of sizes proportional to n and such that the distance between the two sub-trees is larger than $\varepsilon_7 n$ on-expectation. This finishes the proof. □

Finally we explain how one can obtain the same result for the trees related to statistical physics. Meaning we take the minimum spanning tree of the largest component of $G(\mathbf{W}, \tilde{p}_n)$ with i.i.d capacities on its edges.

Proof of Theorem 45. Under Conditions 4, and for n large enough, we can also construct that minimal spanning tree by using an edge deletion algorithm with exponentially distributed edge capacities. Fix $c > 1$ and to each non-oriented edge $\{i, j\}$, $i \neq j$, associate the random capacity $E'_{\{i,j\}}$, which is an exponential random variable of rate 1. The capacities are then used to create a sequence of graphs. First keep each edge independently with probability $p_{\{i,j\}} = \frac{w_i w_j c}{\ell_n} = o(n^{-1/3})$. Once this operation is done we obtain a graph $G'(\mathbf{W}, +\infty)$. We can then construct an increasing sequence of graphs $(G'(\mathbf{W}, p))_{p\geq 0}$ for inclusion by letting $G'(\mathbf{W}, p)$ be the subgraph of $G'(\mathbf{W}, +\infty)$ with edge capacities verifying :

$$\left\{\{i, j\} | E'_{\{i,j\}} \leq p\right\}.$$

Under Conditions 4 :

$$1 - \exp\left(\frac{-w_i w_j c}{\ell_n}\right) = p_{\{i,j\}}\left(1 - \frac{p_{\{i,j\}}}{2} + o(p_{\{i,j\}})\right). \tag{3.28}$$

Since by conditions 4 :

$$\sum_{i<j} p^3_{\{i,j\}} = o(1),$$

Corollary 2.13 from Janson [2010] shows that $G'(\mathbf{W}, +\infty)$ is equivalent to $G(\mathbf{W}, \frac{c}{\ell_n})$ in the sense defined by Janson [2010]. Moreover, if $\{i, j\} \neq \{i', j'\}$, then the edge capacities $E'_{\{i,j\}}$ and $E'_{\{i',j'\}}$ are independent. Equation (3.28) then implies that $G'(\mathbf{W}, +\infty)$ can be coupled edge by edge with some graph $G(\mathbf{W}', \frac{c}{\ell_n})$:

$$G'(\mathbf{W}, +\infty) = G\left(\mathbf{W}', \frac{c}{\ell_n}\right),$$

and $\mathbf{W}'$ verifies Conditions 4 and can be obtained from $\mathbf{W}$ using Equation (3.28). We say that $G'(\mathbf{W}, +\infty)$ is asymptotically equivalent to $G(\mathbf{W}, \tilde{p}_n)$. Generally, for every $p \geq 0$, $G'(\mathbf{W}, p)$ is asymptotically equivalent to $G(\mathbf{W}, (1 - e^{-p})\tilde{p}_n)$ and they are also equivalent in the sense of Janson [2010].

Now for each p, consider the forest $\mathcal{T}'(\mathbf{W}, p)$ constructed by the edge deletion algorithm : Sort the edges $(\{i, j\})_{i\leq n, j\leq n}$ of $G'(\mathbf{W}, p)$ by decreasing order of their capacities $(E'_{\{i,j\}})_{i\leq n, j\leq n}$. Then starting from the first edge, only keep an edge if removing it would disconnect a connected component. By construction, for any $p \leq \infty$, $\mathcal{T}'(\mathbf{W}, p)$ is the minimum spanning forest of $G'(\mathbf{W}, p)$ with respect to the capacities $(E'_{\{i,j\}})_{i\leq n, j\leq n}$. Hence, $(\mathcal{T}'(\mathbf{W}, p))_{p\geq 0}$ is an increasing process of spanning trees for $(G'(\mathbf{W}, p)_{p\geq 0}$. So in order to prove the lower bound of typical distances, if we define $\mathcal{A}'(\varepsilon)$ similarly to $\mathcal{A}(\varepsilon)$ but for the forests $(\mathcal{T}'(\mathbf{W}, p))_{p\geq 0}$, it is sufficient to prove that there exist $\varepsilon > 0$ and $\varepsilon' > 0$ such that $\mathcal{A}'(\varepsilon)$ holds with probability at least ε'. This is true because $\mathcal{A}(\varepsilon)$ holds with probability ε'' for some $\varepsilon'' > 0$, and then $\mathcal{A}'(\varepsilon)$ holds with probability at least ε' because of the asymptotic equivalence relation between $G(\mathbf{W}, (1 - e^{-p})\tilde{p}_n)$ and $G'(\mathbf{W}, p)$ for every $p \geq 0$. For the upper bound, it is clear with the asymptotic equivalence relation that Theorem 48 also holds for $G'(\mathbf{W}, p)$, similarly with the asymptotic equivalence relation the proof of Theorem 49 transposes to $\mathcal{T}'(\mathbf{W}, p)$ directly. And finally the same discussion done after Theorem 49 can be done to deal with the supercritical phase of the graphs $(G'(\mathbf{W}, p))_{p\geq 0}$ using again the asymptotic equivalence relation. □

Chapitre 4

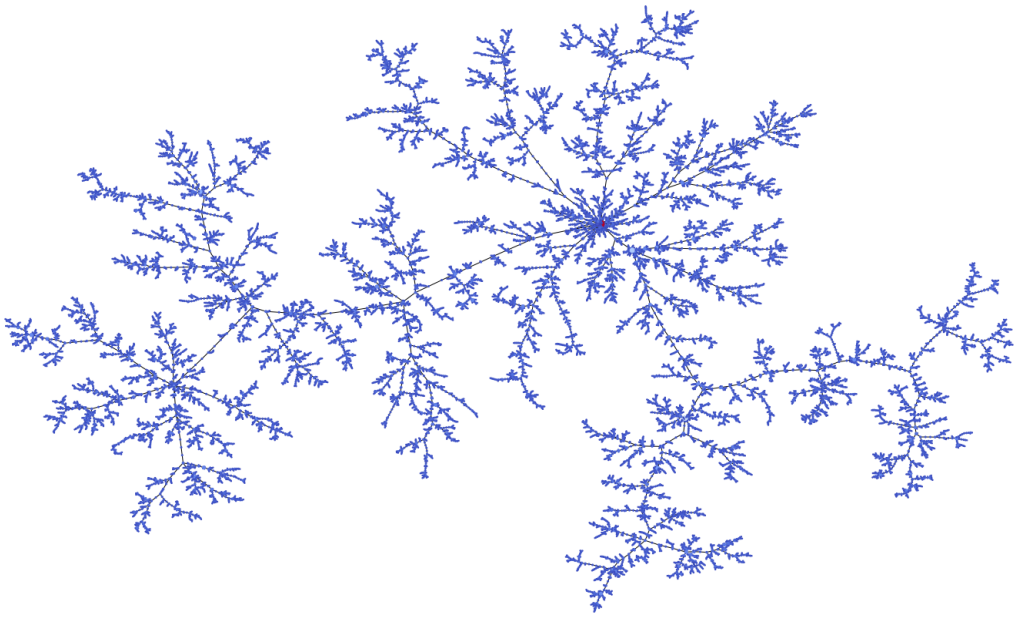
Scaling limit of inhomogeneous minimum spanning tree

Contents

4.1 introduction 143
4.1.1 Defnitions and main results 143
4.1.2 Related work 144
4.1.3 Organization of the chapter 145
4.2 Metric space notions and convergence 146
4.2.1 Gromov-Haussdorf distance 146
4.2.2 General definitions 147
4.3 Cycle breaking algorithm in the continuous and discrete settings . . 149
4.3.1 Definition of discrete and continuous cycle-breaking 149
4.3.2 Relation between discrete and continuous cycle breaking 150
4.4 Convergence of the discrete minimum spanning tree to its scaling limit 151
4.4.1 The scaling limit of inhomogeneous random graphs 151
4.4.2 Preliminary results and convergence of minimum spanning trees in the critical window 154
4.4.3 Convergence of the largest minimum spanning tree in the supercritical regime 155
4.4.4 Properties of the scaling limit 157

Scaling limit of inhomogeneous minimum spanning tree

In this chapter we show that the largest inhomogeneous minimum spanning tree in the forest $\mathcal{T}'(\mathbf{W}, \infty)$ defined in Chapter 3, when endowed with the graph distance renormalized by $n^{1/3}$, converges in distribution as n tends to infinity in the Gromov-Hausdorff topology to a limit compact metric space. We show that the limiting space has the same distribution as the random binary $\mathbb{R}$-tree of Addario-Berry, Broutin, Goldschmidt, and Miermont [2017b] up to a scale factor. We also show, for any $\lambda > 0$ large enough, the convergence of the minimum spanning trees in the forest $\mathcal{T}'(\mathbf{W}, p(\lambda))$ obtained by the partial Kruskal algorithm studied in Chapter 3. We use the same overall method as Addario-Berry et al. [2017b] and address all the new difficulties that arise in generalizing their result. We use the concentration bounds of Chapter 3 in order to prove the main technical tightness lemma that was shown for Erdős-Rényi random graph in Addario-Berry et al. [2017b] for inhomogeneous random graphs. We also use the results from Bhamidi et al. [2010], Bhamidi et al. [2017] and more recently Broutin et al. [2020] that show the convergence of the inhomogeneous random graphs of Chapter 2 to the same scaling limit as that of Erdős-Rényi random graphs up to a scale factor. This chapter is a nice wrap up of all the results of Chapters 2 and 3 alongside many other articles (Addario-Berry et al. [2017b], Broutin et al. [2020], Addario-Berry and Sen [2019] ...) that formally proves that the inhomogeneity in minimum spanning trees with weights having a finite third moment does not have a strong influence in the limit.



4.1 introduction

4.1.1 Defnitions and main results

In all this chapter $c > 1$ is a constant and $\tilde{p}_n = \frac{c}{\ell_n}$ is a supercritical probability. We recall here the construction of the inhomogeneous minimum spanning forest $\mathcal{T}'(\mathbf{W}_n, \infty)$. Let $n \in \mathbb{N}$ and $\mathbf{W}_n = (w_1, w_2, ..w_n)$, with $0 < w_n \leq w_{n-1} \leq ... \leq w_1$, first keep each edge independently with probability $p_{\{i,j\}} = \frac{w_i w_j c}{\ell_n} = o(n^{-1/3})$. Once this operation is done we obtain the graph $G'(\mathbf{W}, +\infty)$. Then to each non-oriented edge $\{i, j\}$, $i \neq j$, associate the random capacity $E_{\{i,j\}}$, which is an exponential random variable of rate 1. $\mathcal{T}'(\mathbf{W}_n, \infty)$ is the minimum spanning forest of $G'(\mathbf{W}_n, \infty)$ associated to the capacities $E'_{\{i,j\}}$.

On the other hand those same capacities can also be used to create a sequence of graphs. We construct an increasing sequence (for inclusion) of graphs $(G'(\mathbf{W}_n, p))_{p\geq 0}$ by letting $G'(\mathbf{W}_n, p)$ be the subgraph of $G'(\mathbf{W}_n, \infty)$ with edge capacities verifying :

$$\left\{\{i, j\} | E'_{\{i,j\}} \leq p\right\}.$$

The minimum spanning forests $(\mathcal{T}'(\mathbf{W}_n, p))_{p\geq 0}$ correspond to the graphs $(G'(\mathbf{W}_n, p))_{p\geq 0}$ with edge capacities $E'_{\{i,j\}}$. Recall from the proof of Theorem 45 that for every $\infty \geq p \geq 0$, $G'(\mathbf{W}_n, p)$ is asymptotically equivalent to $G(\mathbf{W}_n, (1-e^{-p})\tilde{p}_n)$. Meaning that $G'(\mathbf{W}, p)$ can be coupled edge by edge with some graph $G(\mathbf{W}', (1 - e^{-p})\tilde{p}_n)$:

$$G'(\mathbf{W}, p) = G\left(\mathbf{W}', (1 - e^{-p})\tilde{p}_n\right),$$

and $\mathbf{W}'$ verifies Conditions 4 and can be obtained from $\mathbf{W}$ using Equation (3.28). For $\lambda > 0$, we define

$$p'(n, \lambda) = -\ln\left(1 - \frac{1}{c} - \frac{\lambda}{c\ell_n^{1/3}}\right).$$

We drop the n and write $p'(\lambda)$ when the context is clear. By a simple computation $G'(\mathbf{W}_n, p'(\lambda))$ is asymptotically equivalent to $G(\mathbf{W}_n, p(\lambda))$ for all finite λ and $p(\lambda) = \frac{1+\lambda\ell_n^{-1/3}}{\ell_n}$. Some of the results that we use in this chapter where proved for $G(\mathbf{W}_n, p(\lambda))$, all those proofs also hold for $G'(\mathbf{W}_n, p'(\lambda))$, either by our asymptotic equivalence relation or by the one defined in Janson [2010]. Let $\mathcal{T}'(\mathbf{W}_n)$ be the largest tree in $\mathcal{T}'(\mathbf{W}_n, \infty)$.

Before stating our main theorem, we give a brief description of the space we are working with. This space will be rigorously introduced in the next section. Let $\mathcal{M}$ be the set of isometry-equivalence classes of compact metric spaces, and let d_{GH} denote the Gromov-Hausdorff distance on $\mathcal{M}$. $(\mathcal{M}, d_{\mathrm{GH}})$ is a connected polish space (complete and separable metric space). $\mathcal{T}'(\mathbf{W}_n)$ can be seen as an element of $(\mathcal{M}, d_{\mathrm{GH}})$, in order to do so we construct $\mathscr{T}'(\mathbf{W}_n)$ from $\mathcal{T}'(\mathbf{W}_n)$ by rescaling distances by $n^{1/3}$. The main theorem of this chapter is the following.

Theorem 68. *If $\boldsymbol{W}_n$ verifies Conditions 4. Then there exists a random $\mathbb{R}$-tree $\mathscr{M}(W)$ such that, as $n \to \infty$*

$$\mathscr{T}'(\boldsymbol{W}_n) \xrightarrow{\mathrm{d}} \mathscr{M}(W),$$

in the space $(\mathcal{M}, d_{\mathrm{GH}})$. Moreover, $\mathscr{M}(W)$ is almost surely binary, and its Minkowsky dimension exist almost surely and is 3.

We will also show that if $(\mathbf{W'}_n)_{n\geq 1}$ also verifies conditions 4, then there exists a simple relation between the distribution of $\mathscr{M}(W)$ and that of $\mathscr{M}(W')$. A consequence of this last statement is that under conditions 4, all the trees $\mathscr{M}(W)$ are "pretty similar" to the tree $\mathscr{M}$ in Addario-Berry et al. [2017b] from which we have dropped the mass measure. This further affirms the idea that if the weights are in the basin of attraction of a Gaussian distribution, then the inhomogeneity does not change the distribution significantly in the limit.

In order to prove this theorem we will use the detailed description of $G(\mathbf{W}_n, p)$ in the critical and barely supercritical regime done in Chapters 2 and 3. Recall that

$$\ell_n = \sum_{k=1}^{n} w_k.$$

For $\lambda \in \mathbb{R}$ write

$$(G(\mathbf{W}_n, p(\lambda), i))_{i\geq 1}$$

for the components of $G(\mathbf{W}_n, p(\lambda))$, with $p(\lambda) = 1/\ell_n + \lambda/\ell_n^{4/3}$, listed in decreasing order of their sizes, if two components have the same sizes the one containing the smallest label goes first. For each $i \geq 1$, let $\mathscr{G}(\mathbf{W}_n, p(\lambda), i)$ be the metric spaces obtained from $G(\mathbf{W}_n, p(\lambda), i)$ by rescaling distances by $n^{1/3}$, and let

$$\mathscr{G}(\mathbf{W}_n, p(\lambda)) = (\mathscr{G}(\mathbf{W}_n, p(\lambda), i))_{i\geq 1}.$$

Similarly, we define a sequence $(\mathcal{T}(\mathbf{W}_n, p(\lambda), i))_{i\geq 1}$ of graphs, and a sequence $\mathscr{T}(W, p(\lambda)) = (\mathscr{T}(W, p(\lambda), i))_{i\geq 1}$ of metric spaces by starting from $\mathcal{T}(\mathbf{W}, p(\lambda))$ instead of $G(\mathbf{W}, p(\lambda))$. We use the same notation for the forest $\mathcal{T}'(\mathbf{W}_n, p'(\lambda))$ and the graphs $G'(\mathbf{W}_n, p'(\lambda))$.

The second main theorem of this chapter is the following

Theorem 69. *Fix $\lambda \in \mathbb{R}$. If $\boldsymbol{W}_n$ verifies Conditions 4, then there exists a sequence $\mathscr{M}(W, \lambda) = (\mathscr{M}(W, \lambda, i), i \geq 1)$ of random compact metric spaces, such that for any $i \geq 1$ as $n \to \infty$*

$$\mathscr{T}'(\boldsymbol{W}_n, p'(\lambda), i) \xrightarrow{\mathrm{d}} \mathscr{M}(W, \lambda, i)$$

in the space $\mathcal{M}, d_{\mathrm{GH}}$.

This is equivalent to convergence of the whole sequence in the product topology. Moreover, a direct Corollary of our work is the following.

Corollary 69.1. *As $\lambda \to \infty$, $\mathscr{M}(W, \lambda, 1)$ converges in distribution to $\mathscr{M}(W)$ in $(\mathcal{M}, d_{\mathrm{GH}})$.*

4.1.2 Related work

For an overview on minimum spanning trees, we refer the reader to the introduction this thesis and of Chapter 3. Here, we only discuss the work related to the scaling limit.

This chapter is a direct extension of the work of Addario-Berry et al. [2017b]. We follow the same overall steps as them, and use some of their results on convergence of metric spaces. In Addario-Berry et al. [2017b], similar versions of Theorems 68 and 69 are proven for the random minimum spanning tree related to Erdős-Rényi random graphs. This is equivalent to taking $\mathbf{W} = (1, 1, ..., 1)$ in our work. However, their result is stronger for two reasons. First they prove convergence in the Gromov-Hausdorff-Prokhorov topology. This means that they add a mass measure to their spaces, distributed evenly among all vertices. This allows them for instance to show that the $\mathbb{R}$-tree $\mathscr{M}(\mathbf{W})$ seen as a compact measured metric space has all its mass on its leaves. We could do the same here, at the cost of more technical computations, by giving each vertex of our graphs a mass proportional to its weight. Since we intend on keeping this chapter simple we decided to drop the mass measure and only show convergence in the Gromov-Hausdorff topology.

The second difference is that the version of Theorem 69 proven in Addario-Berry et al. [2017b] is done for a different metric. Their metric induces a finer topology, called $(\mathbb{L}_4, \mathrm{dist}^4_{\mathrm{GHP}})$ (see Section 2.1 of their article for a definition of this metric). In order to do so they use the result of Addario-Berry, Broutin, and Goldschmidt [2012] that shows the convergence of rescaled critical Erdős-Rényi random graphs to a sequence of connected compact metric spaces with respect to

the topology $(\mathbb{L}_4, \mathrm{dist}^4_{\mathrm{GHP}})$. Such a result is for now not available for the inhomogeneous random graphs $G(\mathbf{W}, p(\lambda))$. Bhamidi et al. [2010] only show the convergence of the components sizes, the convergence of the metric spaces is done in Bhamidi et al. [2017] under more restrictive conditions, and in Broutin, Duquesne, and Wang [2020] for more general conditions, but only in the product topology. If we drop the mass, and are only interested in convergence in $(\mathbb{L}_4, \mathrm{dist}^4_{\mathrm{GH}})$, by carefully redoing the proof of Addario-Berry et al. [2017b], one can see that in fact we do not need the convergence of the whole graphs with the metric $(\mathbb{L}_4, \mathrm{dist}^4_{\mathrm{GH}})$. We only need the following convergence for any $\lambda > 0$ large enough :

$$\lim_{N\to\infty} \limsup_{n\to\infty} \sum_{k\geq N} \mathrm{diam}(G(\mathbf{W}, \lambda, k))^4 = 0,$$

where $\mathrm{diam}(G(\mathbf{W}, \lambda, k))$ is the diameter of the graph $G(\mathbf{W}, \lambda, k)$.

Another closely related work was done in Addario-Berry and Sen [2019]. There the authors show that the minimum spanning tree obtained by assigning i.i.d. weights to the random 3-regular graph also converges when rescaled by $n^{1/3}$ to the same scaling limit as that of Addario-Berry et al. [2017b] up to a scaling constant with respect to the Gromov-Hausdorff-Prokhorov topology. They further determine the scaling constant to be $6^{1/3}$. Finally, Bhamidi and Sen [2020] show the existence of a scaling limit $\mathscr{T}'(\mathbf{W}_n)$ but with $\mathbf{W}_n$ verifying some restrictive "scale free" conditions instead of Conditions 4. It seems likely that the limit discovered in Addario-Berry et al. [2017b] is universal, and should appear in other various settings. Interestingly, a lot of questions remain unanswered about this object. For instance, is there a more "direct" construction of this object ? What is the distribution of its typical distances and its diameter ? The other famous $\mathbb{R}$-tree is the continuum random tree, which is encoded by a Brownian excursion (Aldous [1993]), what type of random process encodes the random $\mathbb{R}$-tree of Addario-Berry et al. [2017b] ? There are other discrete graphs that have a mean-field behavior such as the hypercube. Is it true that the renormalised minimum spanning tree obtained by assigning i.i.d. weights to the edges of such graphs also converges to a rescaled version of that continuous tree ? This chapter alongside the work of Addario-Berry and Sen [2019] seem to assert this idea.

4.1.3 Organization of the chapter

We start in Section 4.2 by defining the necessary metric space notions, and providing some properties of those spaces that are needed for our proofs. This section is a briefer version of a similar section in Addario-Berry et al. [2017b] in which the authors also define how one can add a mass measure to metric spaces. We do not need this since we work in the Gromov-Hausdorff topology where only the metric distances between isometry classes are relevant. We define two important special cases of metric spaces, the so-called $\mathbb{R}$-trees, and the $\mathbb{R}$-graphs. In doing so we also define paths, cycles, the skeleton and other relevant notions for our use. We end the section with Theorem 73 which characterizes the cores of $\mathbb{R}$-graphs.

Next, in Section 4.3, we explain two algorithms that yield respectively minimum spanning trees in the discrete and continuous setting. These algorithms formalize an intuitive way of cutting down edges in order to go from a discrete graph or an $\mathbb{R}$-graph respectively to a discrete or continuous minimum spanning tree. In doing so we also formalize the idea of continuous minimum spanning trees by relating it directly to the discrete setting, and we show that one can switch between $\mathbb{R}$-graphs and discrete graphs freely provided the $\mathbb{R}$-graph verify some conditions. This work was also done in Addario-Berry et al. [2017b].

Finally in Section 4.4 we show our main theorems. We start the section by recalling general theorems about the convergence of inhomogeneous random graphs. In the setting we are working with, which is that of Conditions 4, the convergence of such graphs for the Gromov-Hausdorff distance was proved incrementally. First there is the work by Aldous in Aldous [1997] then Bhamidi et al. [2010] that shows the convergence of the list of component weights and sizes respectively,

rescaled by $n^{2/3}$ and taken in decreasing order. Bhamidi et al. [2017] showed the convergence in distribution of the graphs $(G(\mathbf{W}_n, p(\lambda)))_{n\geq 1}$ with distances in the connected components rescaled by $n^{1/3}$, seen as a sequence of compact metric spaces with the Gromov-Hausdorff-Prokhorov distance for the product topology to a limiting sequence of $\mathbb{R}$-graph $\mathscr{G}(W, p(\lambda))$. However their work used more restrictive conditions than the one we have here. They also showed the same convergence for the topology $(\mathbb{L}_4, \mathrm{dist}^4_{\mathrm{GHP}})$ with even more restrictive conditions. The work we use here is that of Broutin et al. [2020], they show the same convergence as Bhamidi et al. [2017] with the Gromov-Hausdorff distance under our conditions and for the product topology. In fact their work covers a much larger set of conditions, and they also show the convergence of other functionals of the graphs, such as the surplus. After presenting this convergence and the appropriate objects and notions, such as how to code $\mathbb{R}$-graphs with excursions and point processes, we then show Theorem 69 which is a direct consequence of the work done up to that point.

Before moving to the proof of Theorem 68 we tackle the leader problem. Informally, we want to show that the largest connected component of the graphs $(G(\mathbf{W}_n, p(\lambda)))_{n\geq 1}$ does not change indefinitely as λ and n go to infinity. Instead there exists an almost sure "time" λ after which the largest connected component stays the same. This was shown formally for Erdős-Rényi random graphs in Łuczak [1990], and it was only recently in Addario-Berry, Bhamidi, and Sen [2017a] that the same result was shown for the graphs $(G(\mathbf{W}_n, p(\lambda)))_{n\geq 1}$. We extend slightly the result of Addario-Berry et al. [2017a] by using some of the concentration results of Chapter 2 in the second limit result of Theorem 82. This is done in order to get the right tightness result tailored for our needs. The final crucial result that we prove is related to tightness of the graphs when $\lambda \to \infty$, this is the lemma that allows us to get Theorem 68 from 69. This is done in Lemma 83, and in there we crucially use results from Chapter 3 that depend on results from Chapter 2. There is a similar lemma shown in Addario-Berry et al. [2017b] for Erdős-Rényi random graphs, however it contained a minor error that we correct in this work. Finally, we prove Theorem 69 and give some geometric properties of the limiting object in that theorem.

4.2 Metric space notions and convergence

4.2.1 Gromov-Haussdorf distance

Given a metric space (X, d), we write $[X, d]$ for the isometry class of (X, d). We simply write X for (X, d) or $[X, d]$ when no ambiguity is possible. We define the diameter of (X, d) as $\mathrm{diam}((X, d)) = \sup_{x,y\in X} d(x, y)$. The diameter may be infinite.

The Gromov-Hausdorff distance measures how far two metric spaces are from being isometric. Before defining the Gromov-Hausdorff distance. We start by defining the Hausdorff distance. Let X, X' be two subspaces of a same compact metric space (M, d). The **Hausdorff distance** between X and X' is defined as follow :

$$d_{\mathrm{H}}(X, X') = \max\left\{ \sup_{x\in X} \inf_{x'\in Y} d(x, y),\ \sup_{x'\in Y} \inf_{x\in X} d(x, x') \right\}.$$

Now, suppose that X and X' do not lie in the same overall space. We would still like to compare them, and we do so by mapping them to the same spaces. Formally, if X and X' are two compact metric spaces, the **Gromov-Haussdorf distance** $d_{\mathrm{GH}}(X, X')$ is defined as the infimum of all numbers $d_{\mathrm{H}}(f(X), g(X'))$ for all metric spaces M and isometric embeddings $f : X \to M$ and $g : X' \to M$.

Remark that this defines a distance between the isometry classes of X and X'. This is the original definition of the Gromov-Hausdorff distance, however this definition can be difficult to use because of the infimum over all embeddings. We give here a more suitable definition for our use. Let (X, d) and (X', d') be two metric spaces. Let C be a subspace of $X \times X'$, the **distortion**

$\mathrm{dis}(C)$ is defined as follow :

$$\mathrm{dis}(C) = \sup\left\{|d(x,y) - d'(x',y)| : (x,x') \in C, (y,y') \in C\right\}.$$

A **correspondence** C between X and X' is a measurable subset of $X \times X'$ such that for every $x \in X$ there exists $x' \in X'$ such that $(x,x') \in C$ and vice versa. Let $\mathscr{C}(X,X')$ be the set of correspondences between X and X', then the **Gromov-Haussdorff distance** between the isometry classes of X and X' can also be defined as follow :

$$d_{\mathrm{GH}}(X,Y) = \frac{1}{2}\inf\{\mathrm{dis}(C) \,:\, C \in \mathscr{C}(X,X')\}.$$

Moreover, there exists a correspondence whose distortion achieves this infimum. A proof of those facts can be found in Kalton and Ostrovskii [1999]. Writing $\bar{\mathcal{M}}$ for the set of isometry classes of compact metric spaces, it can be verified that $(\bar{\mathcal{M}}, d_{\mathrm{GH}})$ is itself a polish space.

Let $(X,d,(x_1,x_2,...,x_k))$ and $(X',d',(x'_1,x'_2,...,x'_k))$ be metric spaces, each with an ordered set of k distinguished points. Such spaces are called k-pointed metric spaces. We say that those spaces are isometrically equivalent if there exists an isometry $\phi : X \to X'$ such that $\phi(x_i) = x'_i$ for every $i \in \{1,2,...,k\}$. As before, we write $[X,d,(x_1,x_2,...,x_k)]$ for the isometry equivalence class of $(X,d,(x_1,x_2,...,x_k))$. Similarly, we define the k-pointed Gromov-Hausdorff distance as

$$d^k_{\mathrm{GH}} = \frac{1}{2}\inf\{\mathrm{dis}(C) \,:\, C \in \mathscr{C}(X,X') \text{ such that } (x_i,x'_i) \in C, 1 \le i \le k\}.$$

Like before, $(\bar{\mathcal{M}}^k, d^k_{\mathrm{GH}})$ is itself a polish space.

4.2.2 General definitions

Let (X,d) be a metric space. For $x \in X$ and $r \ge 0$, we define the open ball of radius r and center x as $B_r(x) = \{y \in X \,:\, d(x,y) < r\}$, and the closed ball of radius r and center x as $\bar{B}_r(x) = \{y \in X \,:\, d(x,y) \le r\}$.

Let $\mathcal{C}([a,b],X)$ be the set of continuous function from $[a,b]$ to X, we call them paths from a to b. If $f \in \mathcal{C}([a,b],X)$, the length of f is defined as follows

$$\mathrm{len}(f) = \sup\left\{\sum_{i=1}^{k} d(f(t_{i-1},f(t_i)) \,:\, k \ge 1\ t_0,t_1,...,t_k \in [a,b]\ t_0 \le t_1 \le ... \le t_k\right\}.$$

An arc is the image of a path, it is simple if the path is injective. A geodesic path between $x,y \in X$ is an isometric path $f : [a,b] \to X$ such that $f(a) = x$, $f(b) = y$ and $\mathrm{len}(f) = b - a$. Informally a geodesic path is a shortest path between x and y, its image is called a geodesic arc. A metric space is called a geodesic space if there exists a geodesic path between any two points. Given a compact geodesic space X, given $x \in X$, a cycle in X is a path $f : [a,b] \to X$, with $a < b$, such that $f(a) = x$, $f(b) = x$ and the restriction of f to $[a,b)$ is injective.

A metric space (X,d) is an **$\mathbb{R}$-tree** if for any two points $x,y \in X$ we have the following two properties :

— There exists a unique isometry $c_{u,v} : [0,d(u,v)] \to X$ such that $c_{u,v}(0) = u$ and $c_{u,v}(d(u,v)) = v$.

— If $f : [0,1] \to X$ is an injective path such that $f(0) = u$ and $f(1) = v$ then $f([0,1]) = c_{u,v}([0,d(u,v)])$.

In other terms, a metric space is an $\mathbb{R}$-tree if and only if it is acyclic and geodesic. If (X,d) is an $\mathbb{R}$-tree and $x \in X$, then the degree of x, $\deg_X(x)$ is the number of connected component of $X \setminus \{x\}$. A leaf is a point of degree 1.

A compact metric space (X,d) is an **$\mathbb{R}$-graph** if for every $x \in X$, there exists an $\varepsilon > 0$, such that the open ball of center x and radius ε in X with the distance induced by d in it is an $\mathbb{R}$-tree.

If (X, d) is an $\mathbb{R}$-graph and $x \in X$, the degree of x in X corresponds to its degree in a ball of radius $\varepsilon > 0$ small enough for it to be an $\mathbb{R}$-tree. This definition is independent of the choice of $\varepsilon > 0$ as long as it is small enoughu. We define likewise

$$\mathcal{L}(X) = \{x \in X \,:\, \deg_X(x) = 1\},$$

and

$$\mathrm{skel}(X) = \{x \in X \,:\, \deg_X(x) \geq 2\}.$$

$\mathcal{L}(X)$ is called the set of leaves of X, and $\mathrm{skel}(X)$ is called the skeleton of X. We also have :

$$\mathrm{skel}(X) = \bigcup_{x,y \in X; c \in \Gamma(x,y)} c \setminus \{x, y\},$$

where $\Gamma(x, y)$ is the set of all geodesic arcs between x and y. If X is separable, then by this latter definition $\mathrm{skel}(X)$ can be written as a countable union countable. Hence, there is a unique σ-finite Borel measure ℓ on X such that $\ell(f) = \mathrm{len}(f)$ for every injective path f on X, and $\ell(X \setminus \mathrm{skel}(X)) = 0$. This measure is called the Hausdorff measure of dimension 1, or length measure on X.

Suppose that X is an $\mathbb{R}$-graph. We call branch-point of X, a point of degree at least 3. We denote by $K(X)$ the set of branch-points of X. The core of X, $\mathrm{core}(X)$ is the maximal closed subset of X having only points of degree 2 or more. We define $\mathrm{conn}(X)$ as the set of points x of $\mathrm{core}(X)$ such that $X \setminus \{x\}$ is connected. These two definitions can be extended to discrete graphs. We give now some properties of $\mathbb{R}$-graphs, all these properties are proved in Addario-Berry et al. [2017b].

Proposition 70. *Let (X, d) be an $\mathbb{R}$-graph. Its core is the union of all its cycles. If it is not empty then $(\mathrm{cor}(X), d)$ is an $\mathbb{R}$-graph with no leaves. If it is empty then (X, d) is an $\mathbb{R}$-tree.*

Let X be an $\mathbb{R}$-graph, let $x \in X \setminus \mathrm{cor}(X)$, and let f be the geodesic arc between x and a point $y \in \mathrm{cor}(X)$ with minimal length ℓ. f is unique, or else there would be a cycle, composed of the two different paths of X not in $\mathrm{cor}(X)$. Let $\alpha(x)$ be the closest point in $\mathrm{cor}(X)$ to x. We call $\alpha(x)$ the point of attachment of x and set by convention $\alpha(x) = x$ if $x \in \mathrm{cor}(X)$.

Proposition 71. *The relation $x \sim y \iff \alpha(x) = \alpha(y)$ on X is an equivalence relation. Denote by $[x]$ the equivalence class of x, then $([x], d)$ is a compact $\mathbb{R}$-tree. Moreover $[x]$ is a singleton if and only if $x \in cor(X)$ and $\deg_X(x) = \deg_{\mathrm{cor}(X)}(x)$.*

We also have the following sufficient condition for a point x of X to be in $\mathrm{conn}(X)$.

Proposition 72. *Suppose that $x \in \mathrm{cor}(X)$ has degree 2 and is in the image of a cycle of X. Then $x \in \mathrm{conn}(X)$.*

One can see a discrete multigraph as an $\mathbb{R}$-graph if we associate edge lengths which are positive real numbers. Moreover, multiple discrete graphs with edge lengths can be glued on specific points to obtain an $\mathbb{R}$-graph. This is called the metric gluing (see Burago, Burago, and Ivanov [2001]). The following theorem is also proved in Addario-Berry et al. [2017b]. It gives a precise characterization of graphs with no leaves. Such graphs are important because all the cores of $\mathbb{R}$-graphs are of this form.

Theorem 73. *Suppose that X is an $\mathbb{R}$-graph with no leaves. Then X is either the image of a cycle, or it is the metric gluing of a finite multigraph with edge lengths and nodes of degree at least 3. The Kernel of X, $\ker(X) = (k(X), e(X))$ is the discrete multigraph associated to X, without the edge lengths, if such a multigraph exists. Or it is a multigraph composed of a single node of degree 2 if X is the image of a cycle.*

The surplus, or excess, of a connected multigraph $G = (V, E)$ is $s(G) = |E| - |V| + 1$. For an $\mathbb{R}$-graph X, we set the surplus $s(X) = s(\ker(X))$ in case the kernel is not empty. Otherwise, X is a compact $\mathbb{R}$-tree, and we set $s(X) = 0$. If $s(X) \geq 2$ then the degree of every vertex in the kernel of X is at least 3. Hence $2|e(X)| \geq 3k(X)$, which gives :

$$2s(X) - 2 \geq |k(X)|. \tag{4.1}$$

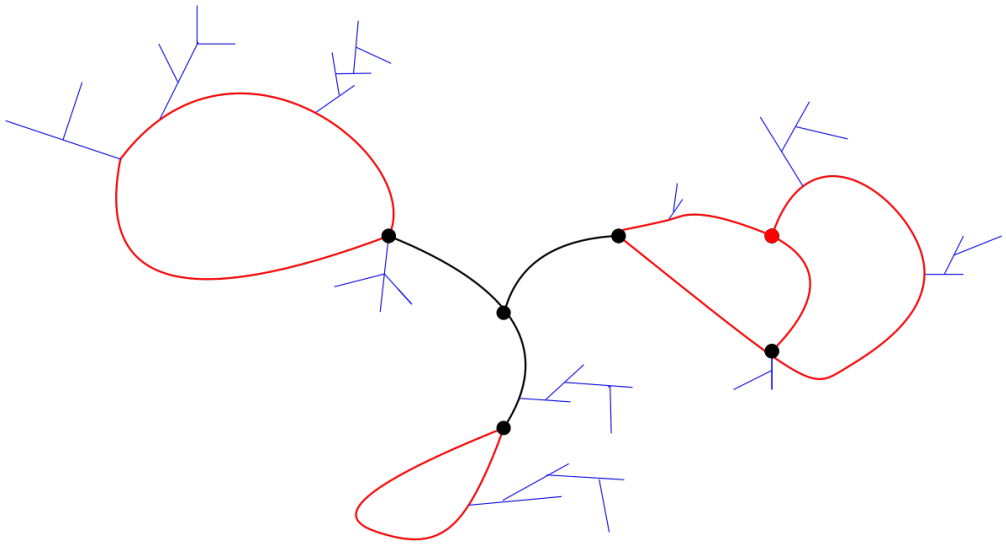


FIGURE 4.1 – An example of an $\mathbb{R}$-graph. Here, $\text{core}(X)$ is drawn with thick lines in black and red, and $\text{conn}(X)$ is drawn in thick lines in red. The large dots represent the vertices of $\ker(X)$. For this graph we have $s(X) = 4$.

4.3 Cycle breaking algorithm in the continuous and discrete settings

In this section we present the different classical cycle-breaking algorithms. We then state the relation between the discrete and continuous procedure, and with the minimum spanning tree. The work here is similar to that of Addario-Berry et al. [2017b] section 3, or Addario-Berry and Sen [2019] section 4.2. The proofs of the results stated here can be found in the former article for a stronger version that considers graphs with a mass measure.

4.3.1 Definition of discrete and continuous cycle-breaking

Let $G = (V, E)$ be a finite discrete multigraph, similarly to the continuous case, we define $\text{conn}(G)$ as the set of edges e such that $G \setminus e = (V, E \setminus e)$ is not empty. If $s(G) > 0$ then $\text{conn}(G)$ is not empty. Let $K(G, .)$ be the law of the multigraph $G \setminus U$, where U is a uniform edge in $\text{conn}(G)$. If $s(G) = 0$, let $K(G, .)$ be the dirac mass at G. Then K is a Markov kernel (Reiss [1993]) from the set of graphs of surplus s to the set of graphs of surplus $s - 1 \vee 0$. For $n \in \mathbb{N}$, let K^n be the application n times of K, then clearly $K^n(G, .)$ does not depend on n for $n \geq s(G)$. Let $K^\infty(G, .)$ be the common value of the $K^n(G, .)$'s for $n \geq s(G)$. We have the following proposition.

Proposition 74. *The probability distribution $K^\infty(G,.)$ is equal to the law of the minimum spanning tree of G when its edges capacities are random, exchangeable, and distinct.*

Similarly, there is a continuous cutting procedure for $\mathbb{R}$-graphs that yield the minimum spanning tree. Recall the definition of $\mathrm{conn}(X)$ from the previous section. For $x \in \mathrm{conn}(X)$ we define the space X cut at x, (X_x, d_x) as follows. (X_x, d_x) is the metric completion of $(X \setminus \{x\}, d_{X\setminus\{x\}})$, where for $y, z \in X$ we have $d_{X\setminus\{x\}}(y,z)$ is the minimal length of a path from y to z that does not have x in its image. Informally (X_x, d_x) corresponds to removing x from X and replacing it by two leaves in place of x, then taking the length distance induced by d on that new space.

We say that $x \in X$ is a regular point if $deg_X(x) = 2$. If X is an $\mathbb{R}$-graph, then the marked space $(X, d, x) \in \bar{\mathcal{M}}^1$ is said to be safely pointed if x is regular. If $s(X) > 0$ and $\mathcal{L} = \ell(. \cap \mathrm{conn}(X))$ is the length measure restricted to $\mathrm{conn}(X)$, then $\mathcal{L}$-almost every point is regular. Since $\mathcal{L}$ is finite by Theorem 73, we can define $\mathcal{K}(X,.)$ as the distributzion of X_x, where x is sampled according to the probability distribution $L/L(\mathrm{conn}(X))$. And $\mathcal{K}(X,.)$ is the Dirac mass on X if $s(X) = 0$. Again, $\mathcal{K}(X,.)$ is a Markov kernel from the set of $\mathbb{R}$-graphs with surplus s to the set of $\mathbb{R}$-graphs with surplus $s - 1 \vee 0$. Similarly to the discrete case we write $\mathcal{K}^\infty(G,.) = \mathcal{K}^{s(X)}(G,.)$.

For $r \in (0,1)$ let $\mathcal{A}_r$ be the set of $\mathbb{R}$-graphs with $s(X) \leq 1/r$, and such that when we see their core as a graph with edge lengths it verifies

$$\min_{e \in e(X)} \ell(e) \geq r, \quad \sum_{e \in e(X)} \ell(e) \leq 1/r,$$

if $s(X) = 1$ this condition amounts to the cycle composing the core of X being of length between r and $1/r$. We have the following important theorem

Theorem 75. *Fix $r \in (0,1)$. Let $(X^n, d^n)_{n\geq 1}$ be a sequence of $\mathbb{R}$-graphs in $\mathcal{A}_r$. If that sequence converges to $(X,d) \in \mathcal{A}_r$ in $(\mathcal{M}, d_{\mathrm{GH}})$, then $(\mathcal{K}^\infty(X^n,.))_{n\geq 1}$ converges weakly to $\mathcal{K}^\infty(X,.)$.*

4.3.2 Relation between discrete and continuous cycle breaking

If given a finite connected multigraph $G = (V, E)$, one can construct two continuous version of it. First, let (V, d) be a metric space with d being the normal graph distance in the discrete sense. Or consider the metric space $(m(G), d_{m(G)})$ created from G by letting the edges be segments of length 1. Clearly $(m(G), d_{m(G)})$ is an $\mathbb{R}$-graph, and it contains a subspace isometric to (V, d). Moreover, if H is the core of G, meaning the maximal subgraph of G of minimal degree two, then $\mathrm{core}(m(G))$ is isometric to $(m(H), d_{m(H)})$.

On the other hand, one can construct discrete versions of some $\mathbb{R}$-graphs. We say that a path $f : [a,b] \to X$ is a local geodesic path between x and y, if $f(a) = x$, $f(b) = y$, and for any $t \in [a,b]$, there exists a neighborhood V of t such that the restriction of f to V is a geodesic path. Let (X,d) be an $\mathbb{R}$-graph and let S_X be the set of points of X of degree at least 3. We say that (X,d) has **integer lengths** if all local geodesic paths between points in S_X have lengths in $\mathbb{N}$. If X is compact and has integer lengths, then necessarily $|S_X| < \infty$. Let

$$v(X) = \{x \in X : d(x, S_X) \in \mathbb{N}\},$$

remark that $S_X \subset v(X)$, and if X is compact and has integer lengths then $|v(x)| < \infty$. If we remove all the points in $v(X)$ from X, it will be separated into a finite collection of paths. Those paths fall into two categories :

— Open paths of length 1 between two points of $v(X)$.
— Half-open paths of length strictly smaller than 1 between a point of $v(X)$ and a leaf of X.

We construct the discrete multigraph associated to X, $g(X)$, by creating an edge between the endpoints of each open path. Let $e(X)$ be the set of such edges and $g(X) = (v(X), e(X))$. Let (X, d) be an $\mathbb{R}$-graph with integer lengths and surplus $s(X)$. Let $x_1, x_2, ..., x_{s(X)}$ be the points chosen by the applications of $\mathcal{K}$ to X. For $i \geq 1$; let $X_{x_1,x_2,...,x_i}$ be the space cut successively at the points $x_1, x_2, ..., x_i$. If $i < s(X)$, recall that x_{i+1} is chosen according to $L/L(X)$ on $\mathrm{conn}(X_{x_1,x_2,...,x_i})$. Since $v(X)$ is finite, almost surely $x_i \notin v(X)$. Hence, x_i fall necessarily in an edge $e_i \in e(X)$. By construction, the edges $(e_i)_{s(X)\geq i\geq 1}$ are distinct. Define :

$$g_i(X) = (v(X), e(X) \setminus \{e_1, e_2, ..., e_i\})$$

for $s(X) \geq i \geq 1$, and $g_0(X) = X$. Clearly, $g_i(X)$ is connected and has surplus $s(X) - i$. Let $\mathrm{cut}(X)$ be the random $\mathbb{R}$-graph obtained by applying $\mathcal{K}^\infty$ to (X, d). We the two following intuitive propositions

Proposition 76. *We have* $d_{\mathrm{GH}}(\mathrm{cut}(X), g_{s(X)}(X)) < 1$.

Proposition 77. *The graph* $g(\mathrm{cut}(X))$ *has the same distribution as the MST of* $g(X)$ *when the edges* $e \in e(X)$ *are given exchangeable, distinct random capacities.*

We end this section by a quick note on metric gluing. For a more in-depth discussion, we refer the reader to Burago et al. [2001]. Let (X, d) be an $\mathbb{R}$-graph and x, y be two distinct point of X. Let $\sim$ be smallest equivalence relation on (X, d) for which $x \sim y$. We define the space $X^{x,y}$ as the quotient metric space of (X, d) by $\sim$. This space should be understood informally as X with x and y identified togethe, it is again an $\mathbb{R}$-graph. Moreover $[x]$ the equivalence class of x in $X^{x,y}$ verifies $\deg_{X^{x,y}}([x]) = \deg_X(x) + \deg_X(y)$. For a finite set $R = \{\{x_i, y_j\},\ i \leq k\}$ for some $k \in \mathbb{N}$ and $x, y \in \mathbb{R}$, we define similarly X^R by identifying each x_i and y_i, and this also yields an $\mathbb{R}$-graph.

4.4 Convergence of the discrete minimum spanning tree to its scaling limit

4.4.1 The scaling limit of inhomogeneous random graphs

In this part we define the scaling limit of inhomogeneous random graphs with node weights verifying the Conditions 4. Then we state the relevant theorems about the convergence of inhomoneous random graphs to their scaling limits.

Take $0 < a < b$, an excursion in $[a, b]$ is a continuous real function h on that interval such that $h(a) = h(b) = 0$. The length of the excursion is $b - a$. Given an excursion h in $[a, b]$, on can construct an $\mathbb{R}$-tree as follow. Let d_h be the pseudo-metric on $[a, b]$ defined by :

$$d_h(s, t) = h(s) + h(t) - 2 \inf_{x\in[s,t]}(h(x)), \text{ for any } s, t \in [a, b].$$

The relation $s \sim t \iff d_h(s, t) = 0$ is an equivalence relation. Let $\mathscr{T}_h$ be the quotient space of $[a, b]$ by $\sim$, and $\bar{d}_h$ be the distance induced by d_h on that quotient space. Then $(\mathscr{T}_h, \bar{d}_h)$ is a compact $\mathbb{R}$-tree.

In order to obtain $\mathbb{R}$-graphs from this construction, wee need to define the points where the space is "glued". For $a > 0$, an excursion h in $[0, a]$, and a countable set $\mathscr{P} \subset \mathbb{R}^+ \times \mathbb{R}^+$, let

$$h \cap \mathscr{P} = \{(x, y) \in \mathscr{P};\ 0 \leq x \leq a,\ 0 \leq y < h(x)\}.$$

Suppose that $|h \cap \mathscr{P}| < \infty$, there exists $k \in \mathbb{N}$ such that $h \cap \mathscr{P} = \{(x_i, y_i);\ 1 \leq i \leq k\}$. For $1 \leq i \leq k$ let

$$r(x_i, y_i) = \inf\{x \geq x_i;\ h(x) \leq y_i\},$$

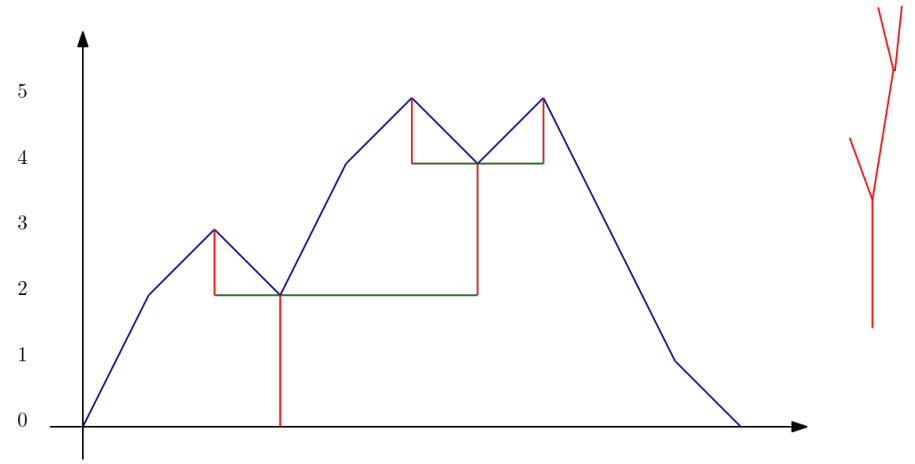


FIGURE 4.2 – An excursion (in blue) and an embedding in the plan of the associated $\mathbb{R}$-tree (in red).

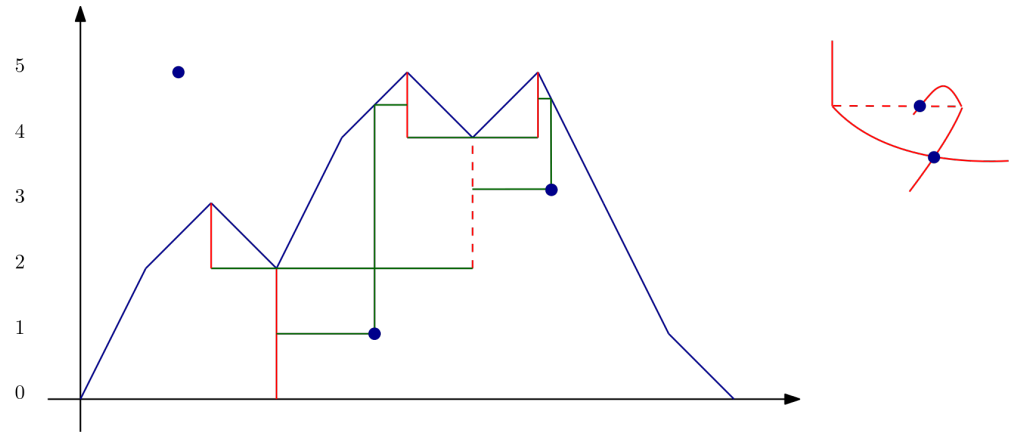


FIGURE 4.3 – An excursion and a set of points in the plan (in blue). An embedding in the plan of the associated $\mathbb{R}$-graph (in red). We drew with dashed lines a particular segment in order to make the embedding clearer.

and identify the equivalence class of x_i and that of $r(x_i, y_i)$ in $\mathscr{T}_h$. The resulting metric space $(\mathscr{G}(h, \mathscr{P}), \tilde{d}_h)$, where $\tilde{d}_h$ is the distance naturally induced by $\bar{d}_h$ on $\mathscr{G}(h, \mathscr{P})$, is an $\mathbb{R}$-graph. It is also clear that $s(\mathscr{G}(h, \mathscr{P})) = k$.

Take $\lambda > 0$ and recall that $p(\lambda) = \frac{1}{n} + \frac{\lambda}{n^{4/3}}$. Let $(G(\mathbf{W}_n, p(\lambda)))_{n\geq 1}$ be a sequence of graphs with node weights verifying Conditions 4, recall the definition of W from those same conditions. Let B be a standard Brownian motion and define :

$$F(W)^\lambda(s) = \sqrt{\frac{\mathbb{E}[W^3]}{\mathbb{E}[W]}} B(s) + \lambda s - \frac{s^2 \mathbb{E}[W^3]}{2\mathbb{E}[W]^2}.$$

Aldous (Aldous [1997]) showed that the excursions of the reflected process $(F(W)^\lambda(s) - \min_{u\in[0,s]} F(W)^\lambda(u))_{s\geq 0}$ can be ordered in a decreasing manner and that the sequence of decreasing excursion lays in $\ell_2^\downarrow$, the space of decreasing sequence with finite ℓ^2 norm. Moreover, in that same article for the Erdős-Rényi case then (Bhamidi et al. [2010]), for inhomogeneous random graphs with finite third moments it is showed that, if we write these excursion sizes as

$\mathbf{Z}(W) = (Z_1(W), Z_2(W), ...)$ and recall $\mathscr{G}(\mathbf{W}_n, p(\lambda)) = (\mathscr{G}(\mathbf{W}_n, p(\lambda), i))_{i\geq 1}$ [1] are the connected components of $G(\mathbf{W}_n, p(\lambda))$ taken in decreasing order, then we have as n goes to infinity :

$$\frac{|\mathscr{G}(\mathbf{W}_n, p(\lambda))|}{n^{2/3}} \xrightarrow{\mathrm{d}} \mathbf{Z}(W),$$

where the convergence takes place in distribution with respect to the the ℓ^2 topology.

Let $\delta(1)$ be a random variable following the Dirac distribution at 1, if $W = \delta(1)$ the $(G(\mathbf{W}_n, p(\lambda)))_{n\geq 1}$ are Erdős-Rényi random graphs. Consider the excursions above 0 of the reflected process $(F(W)^\lambda(s) - \min_{u\in[0,s]} F(W)^\lambda(u))_{s\geq 0}$. Shift those excursion to start from $(0,0)$ and list the shifted excursions in decreasing order $(\zeta^1, \zeta^2, ...,)$. Let $(\mathcal{P}_i)_{i\geq 1}$ be a sequence of i.i.d. Poisson point processes on $\mathbb{R}^+ \times \mathbb{R}^+$. Consider the sequence of $\mathbb{R}$-graphs $\mathcal{G}_\lambda(W) = \big(\mathcal{G}(\zeta^1, \mathcal{P}_1), \mathcal{G}(\zeta^2, \mathcal{P}_2), ...\big) = (\mathcal{G}^1_\lambda(W), \mathcal{G}^2_\lambda(W), ...)$ constructed by the procedure we defined before. Addario-Berry et al. [2012] showed that :

Theorem 78. *Let $\lambda > 0$, and $(G(\boldsymbol{W}_n, p(\lambda))_{n\geq 1}$ be a sequence of i.i.d. Erdős-Rényi random graphs. This means that we take $W = \delta(1)$. Recall that $\mathscr{G}(\boldsymbol{W}_n, p(\lambda)) = (\mathscr{G}(\boldsymbol{W}_n, p(\lambda), i))_{i\geq 1}$ is the sequence the connected components of $G(\boldsymbol{W}_n, p(\lambda))$ taken in decreasing order of their sizes and seen as $\mathbb{R}$-graphs with edge length rescaled by $n^{-1/3}$. Then we have the following convergence as $n \to \infty$*

$$\mathscr{G}(\boldsymbol{W}_n, p(\lambda)) \xrightarrow{\mathrm{d}} \mathcal{G}_\lambda(\delta(1))$$

in $(\mathbb{L}^4, \mathrm{dist}_{GH})$.

This result was extended by Broutin et al. [2020] to the inhomogeneous random graphs we are interested in

Theorem 79. *Let $\lambda > 0$, and let $(G(\boldsymbol{W}_n, p_\lambda))_{n\geq 1}$ be a sequence of graphs with node weights verifying Conditions 4. Recall that $\mathscr{G}(\boldsymbol{W}_n, p(\lambda)) = (\mathscr{G}(\boldsymbol{W}_n, p(\lambda), i))_{i\geq 1}$ is the sequence the connected components of $G(\boldsymbol{W}_n, p(\lambda))$ taken in decreasing order of their sizes and seen as $\mathbb{R}$-graphs with edge length rescaled by $n^{-1/3}$. Let*

$$\kappa = \frac{\mathbb{E}[W]}{\mathbb{E}[W^3]^{2/3}}.$$

Then we have the following convergence as $n \to \infty$

$$\mathscr{G}(\boldsymbol{W}_n, p(\lambda)) \xrightarrow{\mathrm{d}} \kappa\mathcal{G}_\lambda(\kappa\delta(1))$$

in the product topology. Here the factor κ means that the distance of $\mathcal{G}_\lambda(\kappa\delta(1))$ are scaled by κ.

In the rest of the chapter, we will always use κ for $\frac{\mathbb{E}[W]}{\mathbb{E}[W^3]^{2/3}}$. For an $\mathbb{R}$-graph (X, d), let $r(X)$ be the minimal length of a core edge in X. Recall that $s(X)$ is the surplus of X and recall the definition of $k(X)$ from the previous section. Clearly :

$$r(X) = \inf\{d(u,v) \ : u, v \in k(X)\},$$

if $\ker(X)$ is not empty. We set $r(X) = \infty$ if $\mathrm{cor}(X)$ is empty, and $r(X) = \ell(c)$ if X is the image of a cycle c. We have the following Theorem

Theorem 80. *Let $\lambda > 0$, and let $(G(\boldsymbol{W}_n, p'(\lambda)))_{n\geq 1}$ be a sequence of graphs with node weights verifying Conditions 4. Write $(\mathcal{G}^1_\lambda(\kappa\delta(1)), \mathcal{G}^2_\lambda(\kappa\delta(1)), ...)$ for the connected components of the graph $\mathcal{G}_\lambda(\kappa\delta(1))$ taken in decreasing order of their sizes. We have the following joint convergences in distribution for the product topology :*

$$\mathscr{G}(\boldsymbol{W}_n, p(\lambda)) \xrightarrow{\mathrm{d}} \kappa\mathcal{G}_\lambda(\kappa\delta(1))$$

$$(s(\mathscr{G}(\boldsymbol{W}_n, p(\lambda), i)), i \geq 1) \xrightarrow{\mathrm{d}} (s(\kappa\mathcal{G}^i_\lambda(\kappa\delta(1)), i \geq 1)$$

$$(r(\mathscr{G}(\boldsymbol{W}_n, p(\lambda), i)), i \geq 1) \xrightarrow{\mathrm{d}} (r(\kappa\mathcal{G}^i_\lambda(\kappa\delta(1)), i \geq 1).$$

1. We consider $\mathscr{G}(\mathbf{W}_n, p(\lambda))$ to be an infinite sequence of spaces by adding spaces that consist of a single point at the end of the finite sequence of non-empty connected components

This theorem is a consequence of Theorem 2.4 in Broutin et al. [2020]. The last convergence in Theorem 80 is justified by the last argument in the proof of Theorem 4.1 from Addario-Berry et al. [2017b].

Of course all the results of this Section also hold for the graphs $(G'(\mathbf{W}_n, p'(\lambda)))_{n\geq 1}$ by Janson [2010].

4.4.2 Preliminary results and convergence of minimum spanning trees in the critical window

From now on we always suppose that the sequence $(\mathbf{W}_n)_{n\geq 1}$ verifies Conditions 4. Recall the definitions of $G'(\mathbf{W}_n, \infty)$ and $(G'(\mathbf{W}_n, p'(\lambda)))_{n\geq 1}$ and the corresponding minimum spanning trees. We start by showing the convergence of the trees in $(\mathcal{T}'(\mathbf{W}_n, p(\lambda)))_{n\geq 1}$ for $\lambda \in \mathbb{R}$ with distance rescaled by a factor $n^{1/3}$ to a limiting collection of $\mathbb{R}$-trees, then in the following subsection we will use this convergence and a tightness argument to show the convergence of $\mathcal{T}'(\mathbf{W}_n)$ with distances rescaled by $n^{1/3}$.

Recall the cutting procedure from Section 4.3.2, and recall the definition of $\mathcal{G}_\lambda(\delta(1))$ from the previous section. We extend the cutting procedure to $\mathcal{G}_\lambda(\delta(1))$ by cutting independently each $\mathbb{R}$-graph in it. Let $\mathscr{M}(W, \lambda, i) = \mathrm{cut}(\kappa\mathcal{G}^i_\lambda(\kappa\delta(1)))$ for every $i \geq 1$. We have the following more precise version of Theorem 69.

Theorem 81. *Suppose the weights $(\boldsymbol{W}_n)_{i\geq 1}$ verify Conditions 4. Then fix $\lambda \in \mathbb{R}$. We have for any $i \geq 1$, as $n \to \infty$*

$$\mathscr{T}'(\boldsymbol{W}_n, p'(\lambda), i) \xrightarrow{\mathrm{d}} \mathscr{M}(W, \lambda, i)$$

in the topology of d_{GH}.

Proof. By Skorokhod's representation Theorem (Billingsley [1968] Theorem 6.7), we work in probability space in a which the convergence in Theorem 80 holds almost surely. For $r \in (0, 1)$, recall the definition of $\mathcal{A}_r$ from Theorem 75. It is stated in the proof of Theorem 4.1 from Addario-Berry et al. [2017b] that there exists an $r \in (0, 1)$ such that, almost surely, $\mathcal{G}^i_\lambda(\delta(1)) \in \mathcal{A}_r$. Clearly then, there exists $r' \in (0, 1)$ such that $\kappa\mathcal{G}^i_\lambda(\kappa\delta(1)) \in \mathcal{A}_{r'}$. Thus, by the convergence in Theorem 80 we also have $\mathscr{G}'(\mathbf{W}_n, p'(\lambda), i) \in \mathcal{A}_{r''}$ almost surely for n large enough and some $r'' > 0$. Hence, by Theorem 75, almost surely

$$d_{\mathrm{GH}}\left(\mathrm{cut}\left(\mathscr{G}'(\mathbf{W}_n, p'(\lambda), i)\right), \mathrm{cut}(\kappa\mathcal{G}^i_\lambda(\kappa\delta(1)))\right) \to 0.$$

By Proposition 76 and 77, $\mathrm{cut}\left(\mathscr{G}'(\mathbf{W}_n, p'(\lambda), i)\right)$ has the same distribution as $\mathscr{T}'(\mathbf{W}_n, p'(\lambda), i)$, and this ends the proof. □

This proves Theorem 69. We now move to the leader problem. Before showing our main Theorem, we need to ensure that the largest tree of $(G'(\mathbf{W}_n, p'_\lambda))_{n\geq 1}$ stops changing at some point when λ gets large. Let Λ^n be the random time such that $G'^1(\mathbf{W}_n, p'_\lambda) \subset G'^1(\mathbf{W}_n, p'_{\lambda'})$ for any $\lambda' > \lambda \geq \Lambda^n$. Λ^n is the last time when the leader component in term of size changes. Similarly define Λ'^n as the random time after which the component with the largest total weight does not change. We have the following Theorem.

Theorem 82. *Suppose the weights $(\boldsymbol{W}_n)_{i\geq 1}$ verify Conditions 4. The random variable Λ^n is tight in the following sense :*

$$\lim_{\lambda\to\infty} \limsup_{n\to\infty} \mathbb{P}(\Lambda'^n > \lambda) = 0,$$

and also

$$\lim_{\lambda\to\infty} \limsup_{n\to\infty} \mathbb{P}(\Lambda^n > \lambda) = 0.$$

Proof. The first statement of this theorem is a direct consequence of Theorem 2.1 from Addario-Berry et al. [2017a], with the following parameters, which verify Conditions 1 from that article, for any $n \geq 1$ and $i \leq n$:

$$x_i^{(n)} = \frac{w_i \mathbb{E}[W^3]^{1/3}}{n^{2/3}\mathbb{E}[W]}.$$

Theorems 1 and 3 from Chapter 2 show that with high probability as $\lambda \to \infty$ the largest component in term of number of vertices is also the largest component in term of weight. This result and the first statement imply the second statement. □

4.4.3 Convergence of the largest minimum spanning tree in the super-critical regime

In this subsection, we prove Theorem 68. But before that, we start by showing the main tightness result that will allow us to deduce Theorem 68 and Corollary 69.1 from Theorem 81.

Lemma 83. *Suppose the weights $(\boldsymbol{W}_n)_{i\geq 1}$ verify Conditions 4. There exists constants $A > 0$ and $A' > 0$ such that the following holds. Let $\varepsilon \in (0,1)$, and let $\lambda_0 > 0$ be a large enough constant. There exists some $\varepsilon' > 0$ such that for any $\lambda \geq \lambda_0$ we have :*

$$\limsup_{n\to\infty} \mathbb{P}\left(d_{\mathrm{H}}\left(\mathscr{T}'(\boldsymbol{W}_n, p'(\lambda), 1), \mathscr{T}'(\boldsymbol{W}_n)\right) \geq \frac{A'}{\lambda^{1-\varepsilon}} \middle| \Lambda^n \leq \lambda_0 \right) \leq A \exp\left(\frac{-\lambda^{\varepsilon'}}{A}\right).$$

Proof. We work with $G(\mathbf{W}_n, p(\lambda))$, by our asymptotic equivalence relation, all the results in this proof are true for $G'(\mathbf{W}_n, p'(\lambda))$. We proceed by using snapshots, as was already done in Chapter 2. Take $f(0) > 0$ large enough, such that there exists $i \geq 0$ such that $\frac{3^i}{2^i} f(0) = \lambda_0$. For $i \geq 0$ let $f(i+1) = \frac{3}{2} f(i)$. We stop at $f(t_n)$ the smallest element larger than $f'_n = \frac{\ell_n^{1/3}}{\log(n)}$. Recall that $\mathcal{V}$ is the set of all nodes, and $\tilde{\mathcal{V}}_i$ is the set of nodes of the largest component of $G(\mathbf{W}_n, p_{f(i)})$. Similarly to the proof of Theorem 49, for $0 \leq i \leq t_n$ define the following event, let $\varepsilon_0 = 1 - \varepsilon$, and take $A > 0$ to be a large constant, then

— $E(i)$ is the event that every connected component of $G(\mathcal{V} \setminus \tilde{\mathcal{V}}_i, p_{f(i+1)})$ has longest path of length at most $\frac{A\ell_n^{1/3}}{f(i)^{\varepsilon_0}}$.

Let r be the smallest value such that $E(l)$ holds for every $t_n \geq l \geq r$. If $r = i+1$ then $E(i)$ does not hold, because if not we would have $r \leq i$. As was already shown in the proof of Theorem 49, we have for $f(0)$ and n large enough :

$$\mathbb{P}(r = i+1) \leq A \exp\left(\frac{-f(i)^{1/12}}{A}\right),$$

when $\varepsilon_0 = 1/4$. It is straightforward to prove, by changing the constants in that proof, that for any $1 > \varepsilon_0 > 0$, there exists $1 > \varepsilon' > 0$ and a large constant $A > 0$, such that for any $t_n \geq i \geq 0$:

$$\mathbb{P}(r = i+1) \leq A \exp\left(\frac{-f(i)^{\varepsilon'}}{A}\right). \tag{4.2}$$

Hence :

$$\begin{aligned}\mathbb{P}(r > i) &\leq \sum_{k=i}^{t_n - 1} A \exp\left(\frac{-f(i)^{\varepsilon'}}{A}\right) \\ &\leq A' \exp\left(\frac{-f(i)^{\varepsilon'}}{A'}\right),\end{aligned}$$

where $A' > A$ is a large enough constant. Recall also that in Section 2.2 of Chapter 2 we proved that

$$\mathbb{E}[\mathrm{diam}(\mathcal{T}(\mathbf{W}_n, p(f(t_n)), 1))] - \mathbb{E}[\mathrm{diam}(\mathcal{T}(\mathbf{W}_n))] = O(\log(n)^{9/4} n^{1/4}).$$

The same proof also shows that :

$$\mathbb{E}[\mathrm{diam}(\mathcal{T}'(\mathbf{W}_n, p'(f(t_n)), 1))] - \mathbb{E}[\mathrm{diam}(\mathcal{T}'(\mathbf{W}_n))] = O(\log(n)^{9/4} n^{1/4}).$$

And in that same proof we also incidentally showed that :

$$\mathbb{P}\left(d_{\mathrm{H}}(\mathscr{T}'(\mathbf{W}_n, p'(f(t_n)), 1), \mathscr{T}'(\mathbf{W}_n)) \geq \log(n)^{9/4} n^{-1/12}\right) \leq \frac{1}{n}. \tag{4.3}$$

For $i < t_n$, if $\Lambda^n \leq f(i)$, then for any $\lambda \in [f(i), f(i+1)]$ we have :

$$d_{\mathrm{H}}(\mathcal{T}'(\mathbf{W}_n, p'(\lambda), 1), \mathcal{T}'(\mathbf{W}_n, p'(f(t_n)), 1)) \leq d_{\mathrm{H}}(\mathcal{T}'(\mathbf{W}_n, p'(f(i)), 1), \mathcal{T}'(\mathbf{W}_n, p'(f(t_n)), 1)).$$

Moreover, if $r \leq i$, we have for n large enough :

$$d_{\mathrm{H}}(\mathscr{T}'(\mathbf{W}_n, p'(f(i)), 1), \mathscr{T}'(\mathbf{W}_n, p'(f(t_n)), 1)) \leq \sum_{k=i+1}^{t_n} A' f(k)^{\varepsilon - 1} \leq A'' f(i+1)^{\varepsilon - 1}, \tag{4.4}$$

where $A' > 0$ and $A'' > 0$ are some large constants.

Let i_0 be such that $\lambda \in [f(i_0 - 1), f(i_0))$, and let A'' is the constant that appears in Equation (4.4). Since $f(t_n) \to \infty$ when $n \to \infty$, we have $i_0 < t_n$ for n large enough. We then have for any n large enough, using Equation (4.3) :

$$\begin{aligned}
&\mathbb{P}\left(d_{\mathrm{H}}\left(\mathscr{T}'(\mathbf{W}_n, p'(\lambda), 1), \mathscr{T}'(\mathbf{W}_n)\right) > \frac{2A''}{\lambda^{1-\varepsilon}} \middle| \Lambda^n \leq \lambda_0\right) \\
\leq &\mathbb{P}\left(d_{\mathrm{H}}\left(\mathscr{T}'(\mathbf{W}_n, p'(\lambda), 1), \mathscr{T}'(\mathbf{W}_n, p'(f(t_n)), 1)\right) > \frac{A''}{\lambda^{1-\varepsilon}} \middle| \Lambda^n \leq \lambda_0\right) \\
&+ \mathbb{P}\left(d_{\mathrm{H}}\left(\mathscr{T}'(\mathbf{W}_n, p'(f(t_n)), 1), \mathscr{T}'(\mathbf{W}_n)\right) > \frac{A''}{\lambda^{1-\varepsilon}} \middle| \Lambda^n \leq \lambda_0\right) \\
\leq &\frac{1}{\mathbb{P}(\Lambda^n \leq \lambda_0)} \left(\mathbb{P}\left(d_{\mathrm{H}}\left(\mathscr{T}'(\mathbf{W}_n, p'(f(i_0 - 1)), 1), \mathscr{T}'(\mathbf{W}_n, p'(f(t_n)), 1)\right) > A'' f(i_0)^{\varepsilon - 1}\right) + \frac{1}{n}\right).
\end{aligned}$$

By Theorem 82, there exists a constant $A > 0$ such that $\frac{1}{\mathbb{P}(\Lambda^n \leq \lambda_0)} \leq B$ for λ_0 large enough. This fact alongside Equations (4.2) and (4.4) then yield :

$$\begin{aligned}
&\mathbb{P}\left(d_{\mathrm{H}}\left(\mathscr{T}'(\mathbf{W}_n, p'(\lambda), 1), \mathscr{T}'(\mathbf{W}_n)\right) > \frac{2A''}{\lambda^{1-\varepsilon}} \middle| \Lambda^n \leq \lambda_0\right) \\
\leq &B\left(\mathbb{P}\left(d_{\mathrm{H}}\left(\mathscr{T}'(\mathbf{W}_n, p'(f(i_0 - 1)), 1), \mathscr{T}'(\mathbf{W}_n, p'(f(t_n)), 1)\right)\right) > A'' f(i)^{\varepsilon - 1}\right) + \frac{1}{n}\right) \\
\leq &B\left(\mathbb{P}\left(r > i_0 - 1\right) + \frac{1}{n}\right) \\
\leq &B\left(A' \exp\left(\frac{-f(i_0)^{\varepsilon'}}{A'}\right) + \frac{1}{n}\right).
\end{aligned}$$

We finish the proof by letting $n \to \infty$. □

Finally, we can prove Theorem 68 and the related Corollary 69.1.

Theorem 84. *Suppose the weights* $(\boldsymbol{W}_n)_{i\geq 1}$ *verify Conditions 4. There exists a random compact metric space* $\mathscr{M}(W)$ *such that as* n *goes to infinity :*

$$\mathscr{T}'(\boldsymbol{W}_n) \xrightarrow{\mathrm{d}} \mathscr{M}(W).$$

Moreover, as λ *goes to infinity :*

$$\mathscr{M}(W, \lambda, 1) \xrightarrow{\mathrm{d}} \mathscr{M}(W)$$

Proof. Recall that the space $(\bar{\mathscr{M}}, d_{\mathrm{GH}})$ is a polish space. By Theorem 81 we have for any $\lambda \in \mathbb{R}$:

$$\mathscr{T}'(\mathbf{W}_n, p'(\lambda), i) \xrightarrow{\mathrm{d}} \mathscr{M}(W, \lambda, i)$$

in $(\bar{\mathscr{M}}, d_{\mathrm{GH}})$. The Theorem follows from this alongside Lemma 83 and the so-called principle of accompanying laws (Stroock [1993] Theorem 3.1.14 or Theorem 9.1.13 in the second edition). □

4.4.4 Properties of the scaling limit

We end this Chapter by giving some properties of the limit $\mathscr{M}(W)$. These properties where proved in Addario-Berry et al. [2017b] for $\mathscr{M}(\delta(1))$. The scaling parameter κ has no incidence on the properties we give here, the proofs remain exactly similar and are thus omitted.

A natural object $\mathscr{M}(W)$ can be compared to is the continuum random tree. Recall that this latter tree is defined by $\mathscr{T} = \mathscr{T}_{2e}$, where $e = (e(t), 0 \leq t \leq 1)$ is a standard normalized Brownian excursion. The following Theorem gives some properties that are shared between the continuum random tree and $\mathscr{M}(W)$.

Theorem 85. $\mathscr{M}(W)$ *is a compact* $\mathbb{R}$*-tree which is almost surely binary.*

In Addario-Berry et al. [2017b], a mass measure is added to the metric spaces. The mass measure added to $\mathscr{M}(W)$ is the natural limit (for the the Gromov-Hausdorff-Prokorhov distance) of the mass measure on the trees $(\mathscr{T}'(\mathbf{W}_n))_{n\geq 1}$ where each node is given mass $1/n$. The aforementioned mass measure on $\mathscr{M}(W)$ is concentrated on its leaves. This is also the case for $\mathscr{T}$. However, although those two $\mathbb{R}$-trees share these properties, they are drastically different. Let us define the Minkowsky dimension, also called box-counting dimension. Let (X, d) be a compact metric space, for $r > 0$ define $N(X, r)$ as the minimal number of open balls of radius r necessary to cover X. Then the lower **Minkowsky dimension** of X is

$$\underline{\dim}_{\mathrm{M}}(X) = \liminf_{r\downarrow 0} \frac{\log(N(X, r))}{\log(1/r)}.$$

The upper Minkowsky dimension is defined similarly by

$$\overline{\dim}_{\mathrm{M}}(X) = \limsup_{r\downarrow 0} \frac{\log(N(X, r))}{\log(1/r)}.$$

If $\underline{\dim}_{\mathrm{M}}(X) = \overline{\dim}_{\mathrm{M}}(X)$, then we say that X has Minkowsky dimension $\dim_{\mathrm{M}}(X) = \underline{\dim}_{\mathrm{M}}(X)$. We have the following Theorem from Addario-Berry et al. [2017b]

Theorem 86. *The Minkowsky dimension of* $\mathscr{M}(W)$ *exists and is equal to* 3 *almost surely.*

We also know from previous work (Aldous [1991a], Aldous [1991b], Aldous [1993]), that the Minkowsky dimension of $\mathscr{T}$ exists and is equal to 2 almost surely, this shows that $\mathscr{T}$ is not $\mathscr{M}(W)$ in the following sense.

Corollary 86.1. *For any* $a > 0$*, let* $a\mathscr{T}$ *be* $\mathscr{T}$ *with distances rescaled by* a*. Then the laws of* $a\mathscr{T}$ *and* $\mathscr{M}(W)$ *are mutually singular.*

This Corollary is a direct consequence of the Theorem above it. Apart from those results, not much is know about the tree $\mathscr{M}(\delta(1))$. However, there are some results that are strongly believed to be true. For instance, for a metric space (X, d), for $u \geq 0$, define the u-dimensional Hausdorff outer measure as follow :

$$\mathcal{H}^u(X) = \liminf_{r \downarrow 0} \left\{ \sum_i r_i^u \ : \ \text{There exists a covering of X by open balls of radius } 0 < r_i < r \right\}.$$

The **Hausdorff dimension** of X is defined by :

$$\dim_{\mathrm{H}}(X) = \inf\{u \geq 0 \ : \ \mathcal{H}^u(X) = 0\}.$$

It is easily checked that the Minkowsky dimension is always at least equal to the Hausdorff dimension. In many cases those two notions of dimension are equal, for instance, we know that the Hausdorff dimension of $\mathscr{T}$ is also almost surely 2. We conjecture that the Hausdorff dimension of the $\mathbb{R}$-trees $\mathscr{M}(W)$ is almost surely equal to their Minkowsky dimension, that is 3.

Bibliography

R. Abraham and J.-F. Delmas. Local limits of conditioned Galton-Watson trees : the condensation case. *Electron. J. Probab.*, 19 :no. 56, 29, 2014. doi : 10.1214/ejp.v19-3164. URL https://doi.org/10.1214/ejp.v19-3164. (Cité en page 28.)

R. Abraham, J.-F. Delmas, and P. Hoscheit. A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.*, 18 :no. 14, 21, 2013. doi : 10.1214/EJP.v18-2116. URL https://doi.org/10.1214/EJP.v18-2116. (Cité en page 40.)

L. Addario-Berry. Most trees are short and fat. *Probability Theory and Related Fields*, 173(1-2) : 1–26, 2019. (Cité en pages 22 et 32.)

L. Addario-Berry and S. Sen. Geometry of the minimal spanning tree of a random 3-regular graph. *arXiv preprint arXiv :1810.03802*, 2019. (Cité en pages 47, 115, 142, 145 et 149.)

L. Addario-Berry, N. Broutin, and B. Reed. The diameter of the minimum spanning tree of a complete graph. In *Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities*, Discrete Math. Theor. Comput. Sci. Proc., AG, pages 237–248. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2006. (Cité en pages 12, 37, 38, 46 et 60.)

L. Addario-Berry, N. Broutin, and B. Reed. Critical random graphs and the structure of a minimum spanning tree. *Random Structures Algorithms*, 35(3) :323–347, 2009. ISSN 1042-9832. doi : 10.1002/rsa.20241. URL https://doi.org/10.1002/rsa.20241. (Cité en pages 61, 75, 110, 111, 115 et 120.)

L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probability Theory and Related Fields*, 152(3-4) :367–406, 2012. (Cité en pages 30, 33, 43, 61, 144 et 153.)

L. Addario-Berry, L. Devroye, and S. Janson. Sub-Gaussian tail bounds for the width and height of conditioned Galton-Watson trees. *Ann. Probab.*, 41(2) :1072–1087, 2013. ISSN 0091-1798. doi : 10.1214/12-AOP758. URL https://doi.org/10.1214/12-AOP758. (Cité en pages 32 et 38.)

L. Addario-Berry, S. Bhamidi, and S. Sen. A probabilistic approach to the leader problem in random graphs. *arXiv preprint arXiv :1703.09908*, 2017a. (Cité en pages 49, 146 et 155.)

L. Addario-Berry, N. Broutin, C. Goldschmidt, and G. Miermont. The scaling limit of the minimum spanning tree of the complete graph. *Ann. Probab.*, 45(5) :3075–3144, 2017b. ISSN 0091-1798. doi : 10.1214/16-AOP1132. URL https://doi.org/10.1214/16-AOP1132. (Cité en pages 12, 21, 30, 33, 38, 47, 50, 61, 110, 111, 115, 142, 143, 144, 145, 146, 148, 149, 154 et 157.)

D. Aldous. A random tree model associated with random graphs. *Random Structures & Algorithms*, 1(4) :383–402, 1990. (Cité en pages 11 et 111.)

D. Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1) :1–28, 1991a. ISSN 0091-1798. URL http://links.jstor.org/sici?sici=0091-1798(199101)19:1<1:TCRTI>2.0.CO;2-B&origin=MSN. (Cité en pages 28 et 157.)

D. Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge Univ. Press, Cambridge, 1991b. doi : 10.1017/CBO9780511662980.003. URL https://doi.org/10.1017/CBO9780511662980.003. (Cité en pages 28 et 157.)

D. Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1) :248–289, 1993. ISSN 0091-1798. URL http://links.jstor.org/sici?sici=0091-1798(199301)21:1<248:TCRTI>2.0.CO;2-1&origin=MSN. (Cité en pages 21, 28, 29, 145 et 157.)

D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25(2) :812–854, 1997. ISSN 0091-1798. doi : 10.1214/aop/1024404421. URL https://doi.org/10.1214/aop/1024404421. (Cité en pages 21, 33, 42, 45, 57, 60, 61, 115, 145 et 152.)

D. Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. *Electron. J. Probab.*, 3 :No. 3, 59 pp. 1998. ISSN 1083-6489. doi : 10.1214/EJP.v3-25. URL https://doi.org/10.1214/EJP.v3-25. (Cité en pages 21, 33, 60 et 115.)

D. Aldous and J. Pitman. Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent. *Probab. Theory Related Fields*, 118(4) :455–482, 2000. ISSN 0178-8051. doi : 10.1007/PL00008751. URL https://doi.org/10.1007/PL00008751. (Cité en page 36.)

K. B. Athreya and P. E. Ney. *Branching Processes*. Springer-Verlag, New York-Heidelberg, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196. (Cité en page 21.)

K. Azuma. Weighted sums of certain dependent random variables. *Tohoku Math. J. (2)*, 19 : 357–367, 1967. ISSN 0040-8735. doi : 10.2748/tmj/1178243286. URL https://doi.org/10.2748/tmj/1178243286. (Cité en page 17.)

J. Beardwood, J. H. Halton, and J. M. Hammersley. The shortest path through many points. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 55, pages 299–327. Cambridge University Press, 1959. (Cité en page 10.)

A. Ben-Hamou, Y. Peres, and J. Salez. Weighted sampling without replacement. *Brazilian Journal of Probability and Statistics*, 32(3) :657–669, 2018. (Cité en pages 19, 45, 61, 63 et 96.)

S. Bernstein. On a modification of Chebyshev's inequality and of the error formula of laplace. *Ann. Sci. Inst. Sav. Ukraine, Sect. Math*, 1(4) :38–49, 1924. (Cité en pages 17, 32, 44, 63, 67, 78, 83, 90, 91, 116 et 135.)

A. Beveridge, A. Frieze, and C. McDiarmid. Random minimum length spanning trees in regular graphs. *Combinatorica*, 18(3) :311–333, 1998. (Cité en page 11.)

S. Bhamidi and S. Sen. Geometry of the minimal spanning tree in the heavy-tailed regime : new universality classes. *arXiv preprint arXiv :2009.10696*, 2020. (Cité en pages 116 et 145.)

S. Bhamidi, R. van der Hofstad, and J. S. H. van Leeuwaarden. Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electron. J. Probab.*, 15 :no. 54, 1682–1703, 2010. doi : 10.1214/EJP.v15-817. URL https://doi.org/10.1214/EJP.v15-817. (Cité en pages 33, 36, 42, 57, 58, 61, 63, 115, 127, 142, 145 et 152.)

S. Bhamidi, S. Sen, and X. Wang. Continuum limit of critical inhomogeneous random graphs. *Probab. Theory Related Fields*, 169(1-2) :565–641, 2017. ISSN 0178-8051. doi : 10.1007/s00440-016-0737-x. URL https://doi.org/10.1007/s00440-016-0737-x. (Cité en pages 61, 142, 145 et 146.)

S. Bhamidi, R. van der Hofstad, and S. Sen. The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs. *Probability Theory and Related Fields*, 170(1-2) :387–474, 2018. (Cité en pages 30, 35, 36, 39, 50 et 129.)

P. Billingsley. *Convergence of probability measures*. John Wiley & Sons, Inc., New York-London-Sydney, 1968. (Cité en pages 28 et 154.)

B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms*, 31 :3–122, 2007. (Cité en pages 35, 60 et 115.)

O. Borůvka. About a certain minimal problem. *Práce mor. přírodově d. spol. Brně*, 3 :37–58, 1926. (Cité en page 10.)

S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. ISBN 978-0-19-953525-5. doi : 10.1093/acprof:oso/9780199535255.001.0001. URL https://doi.org/10.1093/acprof:oso/9780199535255.001.0001. A nonasymptotic theory of independence, With a foreword by Michel Ledoux. (Cité en pages 17, 45, 70, 90 et 116.)

L. A. Braunstein, S. V. Buldyrev, R. Cohen, S. Havlin, and H. E. Stanley. Optimal paths in disordered complex networks. *Physical review letters*, 91(16) :168701, 2003. (Cité en page 39.)

L. A. Braunstein, Z. Wu, Y. Chen, S. V. Buldyrev, T. Kalisky, S. Sreenivasan, R. Cohen, E. Lopez, S. Havlin, and H. E. Stanley. Optimal path and minimal spanning trees in random weighted networks. *International Journal of Bifurcation and Chaos*, 17(07) :2215–2255, 2007. (Cité en pages 39, 111, 115 et 116.)

T. Britton, M. Deijfen, and A. Martin-Löf. Generating simple random graphs with prescribed degree distribution. *Journal of Statistical Physics*, 124(6) :1377–1397, Sep 2006. (Cité en pages 33 et 39.)

N. Broutin, T. Duquesne, and M. Wang. Limits of multiplicative inhomogeneous random graphs and Lèvy trees : The continuum graphs. *arXiv preprint arXiv :1804.05871*, 2018. (Cité en page 61.)

N. Broutin, T. Duquesne, and M. Wang. Limits of multiplicative inhomogeneous random graphs and Lèvy trees : Limit theorems. *arXiv preprint arXiv :2002.02769*, 2020. (Cité en pages 36, 39, 43, 46, 57, 61, 115, 118, 138, 142, 145, 146, 153 et 154.)

D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. ISBN 0-8218-2129-6. doi : 10.1090/gsm/033. URL https://doi.org/10.1090/gsm/033. (Cité en pages 148 et 151.)

M. Camarri and J. Pitman. Limit distributions and random trees derived from the birthday problem with unequal probabilities. *Electronic Journal of Probability*, 5, 2000. (Cité en page 21.)

Y. Chen, E. López, S. Havlin, and H. E. Stanley. Universal behavior of optimal paths in weighted networks with general disorder. *Physical review letters*, 96(6) :068702, 2006. (Cité en pages 39, 111, 115 et 116.)

H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statistics*, 23 :493–507, 1952. ISSN 0003-4851. doi : 10.1214/aoms/1177729330. URL https://doi.org/10.1214/aoms/1177729330. (Cité en page 17.)

F. Chung and L. Lu. *Complex graphs and networks*, volume 107 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC ; by the American Mathematical Society, Providence, RI, 2006. ISBN 978-0-8218-3657-6 ; 0-8218-3657-9. doi : 10.1090/cbms/107. URL https://doi.org/10.1090/cbms/107. (Cité en pages 33, 60 et 115.)

B. Davis and McDonald. An elementary proof of the local central limit theorem. *Journal of Theoretical Probability*, 8 :693–701, 07 1995. doi : 10.1007/BF02218051. (Cité en page 27.)

F. den Hollander. Probability theory : The coupling method. 2012. (Cité en page 96.)

S. Dhara, R. van der Hofstad, J. S. van Leeuwaarden, and S. Sen. Critical window for the configuration model : finite third moment degrees. *Electronic Journal of Probability*, 22, 2017. (Cité en page 33.)

R. M. Dudley. Sample functions of the Gaussian process. *Ann. Probability*, 1(1) :66–103, 1973. ISSN 0091-1798. doi : 10.1214/aop/1176997026. URL https://doi.org/10.1214/aop/1176997026. (Cité en page 70.)

T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281) :vi+147, 2002. ISSN 0303-1179. (Cité en page 21.)

T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lèvy trees. *Probability Theory and Related Fields*, 131(4) :553–603, 2005. (Cité en page 30.)

P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1) :17–60, 1960. (Cité en page 32.)

P. Flajolet and A. Odlyzko. The average height of binary trees and other simple trees. *J. Comput. System Sci.*, 25(2) :171–213, 1982. ISSN 0022-0000. doi : 10.1016/0022-0000(82)90004-6. URL https://doi.org/10.1016/0022-0000(82)90004-6. (Cité en page 38.)

D. A. Freedman. On tail probabilities for martingales. *Ann. Probability*, 3 :100–118, 1975. ISSN 0091-1798. doi : 10.1214/aop/1176996452. URL https://doi.org/10.1214/aop/1176996452. (Cité en pages 77, 83 et 100.)

A. Frieze, M. Ruszinkó, and L. Thoma. A note on random minimum length spanning trees. *the Electronic Journal of Combinatorics*, 7(1) :R41, 2000. (Cité en page 11.)

A. M. Frieze. On the value of a random minimum spanning tree problem. *Discrete Applied Mathematics*, 10(1) :47–56, 1985. (Cité en pages 11 et 111.)

A. M. Frieze and C. J. McDiarmid. On random minimum length spanning trees. *Combinatorica*, 9(4) :363–374, 1989. (Cité en page 11.)

E. N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4) :1141–1144, 1959. (Cité en page 32.)

R. Graham and P. Hell. On the history of the minimum spanning tree problem. *Annals of the History of Computing*, 7 :43–57, 02 1985. doi : 10.1109/MAHC.1985.10011. (Cité en page 10.)

W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58 :13–30, 1963a. ISSN 0162-1459. URL http://links.jstor.org/sici?sici=0162-1459(196303)58:301<13:PIFSOB>2.0.CO;2-D&origin=MSN. (Cité en page 17.)

W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58 :13–30, 1963b. ISSN 0162-1459. URL http://links.jstor.org/sici?sici=0162-1459(196303)58:301<13:PIFSOB>2.0.CO;2-D&origin=MSN. (Cité en page 19.)

S. Janson. The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph. *Random Structures & Algorithms*, 7(4) :337–355, 1995. (Cité en page 11.)

S. Janson. Asymptotic equivalence and contiguity of some random graphs. *Random Structures & Algorithms*, 36(1) :26–45, 2010. (Cité en pages 140, 143 et 154.)

S. Janson, T. Łuczak, and A. Rucinski. *Random Graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000. ISBN 0-471-17541-2. doi : 10.1002/9781118032718. URL https://doi.org/10.1002/9781118032718. (Cité en page 38.)

N. J. Kalton and M. I. Ostrovskii. Distances between banach spaces. In *Forum Mathematicum*, volume 11, pages 17–48. De Gruyter, 1999. (Cité en pages 41 et 147.)

I. Kortchemski. Sub-exponential tail bounds for conditioned stable Bienaymé-Galton-Watson trees. *Probab. Theory Related Fields*, 168(1-2) :1–40, 2017. ISSN 0178-8051. doi : 10.1007/s00440-016-0704-6. URL https://doi.org/10.1007/s00440-016-0704-6. (Cité en page 32.)

J. B. Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical society*, 7(1) :48–50, 1956. (Cité en page 13.)

R. Laroche, O. Safsafi, N. Broutin, and R. Féraud. Batched multi-armed bandits with crowd externalities. In *Proceedings of the 33rd Advances in Neural Information Processing Systems (NeurIPS, submitted)*, 2020. (Cité en page 8.)

J.-F. Le Gall and Y. Le Jan. Branching processes in Lèvy processes : the exploration process. *The Annals of Probability*, 26(1) :213–252, 1998. (Cité en page 24.)

T. D.-J.-F. Le Gall and T. Duquesne. Random trees, Lèvy processes and spatial branching processes. *Astérisque*, 281, 2002. (Cité en page 30.)

T. Łuczak. Component behavior near the critical point of the random graph process. *Random Structures & Algorithms*, 1(3) :287–310, 1990. (Cité en pages 33, 47, 49, 61 et 146.)

T. Luczak. Cycles in a random graph near the critical point. *Random Structures & Algorithms*, 2(4) :421–439, 1991. (Cité en page 47.)

T. Łuczak, B. Pittel, and J. C. Wierman. The structure of a random graph at the point of the phase transition. *Transactions of the American Mathematical Society*, 341(2) :721–748, 1994. (Cité en page 33.)

J.-F. Marckert and A. Mokkadem. The depth first processes of galton–watson trees converge to the same brownian excursion. *The Annals of Probability*, 31(3) :1655–1678, 2003. (Cité en page 25.)

J. Neveu. Arbres et processus de galton-watson. *Annales de l'I.H.P. Probabilités et statistiques*, 22(2) :199–207, 1986. URL http://www.numdam.org/item/AIHPB_1986__22_2_199_0. (Cité en page 19.)

I. Norros and H. Reittu. On a conditionally Poissonian graph process. *Adv. in Appl. Probab.*, 38(1) :59–75, 2006. ISSN 0001-8678. doi : 10.1239/aap/1143936140. URL https://doi.org/10.1239/aap/1143936140. (Cité en pages 33, 39, 60, 115 et 127.)

R. Otter. The multiplicative process. *The Annals of Mathematical Statistics*, pages 206–224, 1949. (Cité en page 27.)

J. Pitman. Enumerations of trees and forests related to branching processes and random walks. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, volume 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 163–180. Amer. Math. Soc., Providence, RI, 1998. (Cité en page 27.)

R. C. Prim. Shortest connection networks and some generalizations. *The Bell System Technical Journal*, 36(6) :1389–1401, 1957. (Cité en page 13.)

R.-D. Reiss. *A course on point processes*. Springer Series in Statistics. Springer-Verlag, New York, 1993. ISBN 0-387-97924-7. doi : 10.1007/978-1-4613-9308-5. URL https://doi.org/10.1007/978-1-4613-9308-5. (Cité en page 149.)

D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999. ISBN 3-540-64325-7. doi : 10.1007/978-3-662-06400-9. URL https://doi.org/10.1007/978-3-662-06400-9. (Cité en pages 18 et 30.)

R. J. Serfling. Probability inequalities for the sum in sampling without replacement. *Ann. Statist.*, 2 :39–48, 1974. ISSN 0090-5364. URL http://links.jstor.org/sici?sici=0090-5364(197401)2:1<39:PIFTSI>2.0.CO;2-6&origin=MSN. (Cité en pages 19, 50 et 62.)

C. Stegehuis, R. van der Hofstad, A. Janssen, and J. van Leeuwaarden. Clustering spectrum of scale-free networks. *Physical Review E*, 96(4), 10 2017. ISSN 2470-0045. doi : 10.1103/PhysRevE.96.042309. (Cité en pages 21 et 115.)

D. W. Stroock. *Probability theory, an analytic view*. Cambridge University Press, Cambridge, 1993. ISBN 0-521-43123-9. (Cité en page 157.)

M. Talagrand. *The generic chaining*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005. ISBN 3-540-24518-9. Upper and lower bounds of stochastic processes. (Cité en page 70.)

R. van der Hofstad, S. Kliem, and J. van Leeuwaarden. Cluster tails for critical power-law inhomogeneous random graphs. *Journal of Statistical Physics*, 171(1) :38–95, 4 2018. ISSN 0022-4715. doi : 10.1007/s10955-018-1978-0. (Cité en pages 61 et 115.)

V. V. Vazirani. *Approximation algorithms*. Springer-Verlag, Berlin, 2001. ISBN 3-540-65367-8. (Cité en page 111.)

H. W. Watson and F. Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4 :138–144, 1875. (Cité en page 21.)

S. Willard. *General topology*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1970. (Cité en page 41.)

Z. Wu, L. A. Braunstein, S. Havlin, and H. E. Stanley. Transport in weighted networks : partition into superhighways and roads. *Physical Review Letters*, 96(14) :148702, 2006. (Cité en pages 111 et 116.)

J. E. Yukich. *Probability theory of classical Euclidean optimization problems*, volume 1675 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998. ISBN 3-540-63666-8. doi : 10.1007/BFb0093472. URL https://doi.org/10.1007/BFb0093472. (Cité en page 11.)

Résumé : Cette thèse porte sur un type particulier d'arbres couvrants minimums aléatoires inhomogènes, ainsi que sur les graphes sous-jacents à ces arbres. Pour $n \in \mathbb{N}$ grand, nous posons des poids $(w_i)_{i\leq n}$ strictement positifs sur les noeuds du graphe complet de taille n. De plus, les poids $(w_i)_{i\leq n}$ que nous considérerons dans cette thèse vérifient une condition particulière dite de troisième moment fini. Cette condition est naturelle au vu de l'histoire de ces arbres. Puis, à chaque arête $\{i, j\}$ du graphe complet nous attribuons une capacité qui est une variable aléatoire exponentielle de paramètre $w_i w_j$ indépendamment du reste. Nous construisons ensuite l'arbre couvrant minimum du graphe complet avec ces capacités.

Dans le Chapitre 2 nous obtenons des propriétés asymptotiques des graphes de rang-1 inhomogènes dans la fenêtre dite à peine sur-critique. La nouveauté apportée dans ce chapitre consiste en l'étude détaillée de ces graphes inhomogènes, ainsi qu'en l'obtention de nouvelles inégalités de concentration pour les tirages sans remises.

Dans le Chapitre 3, nous utilisons les résultats du Chapitre 2 afin de démontrer que l'espérence des distances typiques et du diamètre de nos arbres couvrants minimums aléatoires inhomogènes sont de l'ordre de $n^{1/3}$ quand n est grand. Comme corollaire de notre travail, nous répondons par l'affirmative à une conjecture issue de la physique statistique concernant les distances typiques d'arbres couvrants minimums proches de ceux qu'on étudie.

Finalement, dans le Chapitre 4, nous démontrons que nos arbres couvrants minimums, vus comme des espaces métriques et avec des distances renormalisées par $n^{1/3}$ convergent en loi vers un espace métrique compact non-trivial pour la topologie de Gromov-Hausdorff.

Abstract : In this thesis we study a specific type of inhomogeneous random minimum spanning trees and their related graphs. For large $n \in \mathbb{N}$, we put positive weights $(w_i)_{i\leq n}$ on the nodes of the complete graph of size n. Moreover, the weights $(w_i)_{i\leq n}$ that we consider in this thesis verify a finite third moment condition. This condition is natural given the history of those trees. Then, we give to each edge $\{i, j\}$ of the complete graph an edge capacity which is an exponential random variable of parameter $w_i w_j$ independently of everything else. We then build the minimum spanning tree of the complete graph with these edge capacities.

In Chapter 2 we prove asymptotic properties for rank-1 inhomogeneous random graphs in the so-called barely super-critical regime. The novelty of this chapter lies in the detailed study of those graphs, and on proofs of original concentration inequalities for sampling without replacement.

In Chapter 3, we use the results of Chapter 2 in order to prove that the expectation of the typical distances and of the diameter of our minimum spanning trees is of order $n^{1/3}$ when n is large. As a corollary of our work, we answer a conjecture from statistical physics regarding typical distances in a model of minimum spanning trees closely related to our model.

Finally, in Chapter 4, we prove that our minimum spanning trees, seen as metric spaces and with distances rescaled by $n^{1/3}$ converge in distribution to a non-trivial compact metric space for the Gromov-Hausdorff topology.