



Scaling limits of inhomogeneous continuum random trees and graphs

Arthur Blanc-Renaudie

► To cite this version:

Arthur Blanc-Renaudie. Scaling limits of inhomogeneous continuum random trees and graphs. Probability [math.PR]. Sorbonne Université, 2022. English. NNT : 2022SORUS320 . tel-03879046v2

HAL Id: tel-03879046

<https://hal.science/tel-03879046v2>

Submitted on 1 Feb 2023

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

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

THÈSE

Spécialité : Mathématiques

Présentée pour obtenir le grade de :
Docteur de Sorbonne Université.

par

Arthur BLANC-RENAUDIE

Limites d'échelles d'arbres et de graphes inhomogènes

Soutenue le 24 novembre 2022 devant la commission d'examen:

M.	Nicolas BROUTIN	Directeur
M.	Nicolas CURIEN	Rapporteur
M.	Thomas DUQUESNE	Examineur
M ^{me}	Christina GOLDSCHMIDT	Examinatrice
M ^{me}	Bénédicte HAAS	Examinatrice
M.	Jean-François MARCKERT	Examineur
M.	Grégory MIERMONT	Examineur

et s'appuyant sur le rapport de M. Asaf NACHMIAS (Rapporteur).

Contents

1	Introduction	9
1.1	Main motivations, background, and algorithms	10
1.2	Overview of the chapters	28
2	Compactness and dimensions of ICRT	47
2.1	Introduction	48
2.2	Model and definition of the fractal dimensions	48
2.3	Main results	51
2.4	Preliminaries	53
2.5	The mass measure	56
2.6	Compactness	62
2.7	Fractal dimensions : proof of theorem 2.3.4	66
2.8	Appendix: Concentration inequalities	68
3	Limit of trees with fixed degree sequence	71
3.1	Introduction	71
3.2	Models and algorithms	75
3.3	Notions of convergence	78
3.4	Main results	81
3.5	Convergence of the first branches of \mathcal{D} -trees	85
3.6	Gromov–Prokhorov convergence of \mathcal{D} -trees	90
3.7	GHP convergence and height of \mathcal{D} -trees	92
3.8	Additional remarks	102
3.9	Appendix	106
4	Limit of connected multigraph	109
4.1	Introduction	110
4.2	Notions of convergence	112
4.3	Constructions of \mathcal{D} -trees, \mathcal{P} -trees and Θ -ICRT	113
4.4	Constructions of (\mathcal{D}, k) -graphs, (\mathcal{P}, k) -graphs and (Θ, k) -ICRG	115
4.5	Main results	118
4.6	Study of the bias	120
4.7	Proof of the main theorems	128
4.8	Configuration model and multiplicative graphs	129
4.9	Appendix A: \mathbb{R} -tree reconstruction problem	133
4.10	Appendix B: \square_c is continuous	135
5	Looptree, Fennec, and Snake of ICRT.	137
5.1	Introduction	138
5.2	Model and main results.	140
5.3	Recalls from Section 2.4	145
5.4	Basic properties on \mathcal{L}	146

5.5	Preliminary results on left, front, right	149
5.6	Construction and Hölder continuity of the contour path.	154
5.7	Fractal dimensions of the looptree	158
5.8	Proof of Theorem 5.2.4	163
5.9	Proof and extension of Theorem 5.2.5	165
5.10	Appendix A: Proof of Lemma 5.2.7, and measurability	167
5.11	Appendix B: Holder continuity of a Gaussian free field.	170
5.12	Appendix C: Supremum of empirical process.	171
Bibliography		175

Remerciements

Attention!!! Ames sensibles s'abstenir! Je dédie ces lignes à ma défunte mère qui m'a appris dans ces derniers instants que l'on ne pouvait jamais assez remercier les gens de leur vivant.

J'écris ces remerciements pour montrer à quel point certaines personnes comptent pour moi. Sans elles, je n'aurais très probablement pas pu passer certains moments difficiles de ma vie. Mais ces moments me reviennent violemment alors que j'écris péniblement ces lignes. J'espère juste que la lecture de ces remerciements ne vous sera pas aussi pénible.

Tout d'abord, j'aimerais te remercier Nicolas de m'avoir aidé pendant tout le long de la thèse. J'ai probablement dû te décevoir plusieurs fois à cause de mes différentes périodes de stagnation. Mais tu as toujours été très gentil et patient envers moi. Grâce à ça, même après la mort de ma mère, j'ai toujours eu un endroit où je pouvais me sentir chez moi et quelqu'un avec qui me confier. (J'aurais juste aimé que le COVID ne foute pas la merde dans ce nouveau foyer...) Aussi, merci infiniment de m'avoir lu et conseillé un nombre indénombrable de fois sur la rédaction, jusqu'à ce que j'apprenne finalement comment bien écrire. Pour les maths, même si je les faisais principalement seul, je savais que si j'avais la moindre question, je pouvais toujours compter sur toi. J'avais juste à venir te voir, et j'en ressortais toujours avec une réponse raisonnable. (J'espère que cette méthode miracle nous servira pour nos futures collaborations!)

Merci, à toi, Nicolas C. d'avoir rapporté ma thèse en si peu de temps. Merci, aussi de m'avoir guidé dans le domaine pendant mon une bonne partie du Master, et pour m'avoir redonné goût aux maths après une longue période de dégoût. Tu es probablement le meilleur prof que j'aurais pu avoir à l'époque, et clairement, sans toi, je n'aurais pas pu faire de la recherche aujourd'hui.

Thanks Asaf Nachmias for reporting my PhD on such a short notice. Thanks also for accepting me as a postdoc in TAU. I hope our collaborations will be fruitful!

Je remercie maintenant les autres membres du Jury: Thomas Duquesne, Christina Goldshmidt, Bénédicte Haas, Jean-François Marckert, Grégory Miermont. Aussi loin que des étoiles, aussi brillant que des soleils, si je vous ai invité à ma soutenance de thèse c'est que je vous tiens en haute estime. Et je suis énormément honoré que vous ayez accepté mon invitation.

Ensuite j'aimerais remercier les différents profs de maths qui m'ont fourni ma dose d'opium. Merci à mon prof de 5e qui m'a permis d'apprendre quasiment tout le programme de lycée. Merci à Mr Matthieu Schavsinski qui m'a suivi de la 4e à la terminale. Merci à Animath de m'avoir occupé pendant le lycée avec les olympiades. Merci à Villani d'avoir pris le temps de voir le mioche que j'étais en terminal. Le mini entretien m'a vraiment motivé à me spécialiser en analyse. Merci à Nicolas Tosel de m'avoir fourni en livre autant que je le souhaitais. J'ai du manger des livres et des livres de licence et masters à ne plus savoir compter. Mais trop manger amène au dégoût, et j'ai donc quasiment arrêté les maths jusqu'au moment où par chance j'ai eu Nicolas Curien.

Il est finalement l'heure de remercier ma famille, ou du moins ce qu'il en reste... Merci, maman, de m'avoir protégé de mon père et des vautours qui tournaient autour. Merci d'avoir tout fait en sorte que j'ai mon propre monde autour des maths. J'espère qu'il y a de la musique au paradis. Merci frerot d'avoir pris soin de ton autiste de frère. Je sais que l'entraide c'est vraiment ce qui nous a permis de survivre dans la famille, mais pour l'instant j'arrive pas trop à t'aider. Merci Tatie, Tonton de vous être énormément occupés de moi depuis la mort de maman. Vous m'avez vraiment aidé à me sortir de la dépression à de nombreuses reprises et à éviter de retomber dedans. Clairement sans votre aide ma thèse ne serait pas aussi avancée.

Résumé

Français

Dans cette thèse on étudie certains modèles d'arbres (\mathcal{D} -arbre, \mathcal{P} -arbre, ICRT) et de graphes (modèle de configuration, graphe multiplicatif) à suite de degrés fixés. Pour cela, on développe de nouveaux algorithmes qui construisent ces modèles en collant des branches les unes sur les autres. En analysant ces constructions, on obtient des résultats sur la géométrie de nos modèles.

Pour les analyser, on utilise principalement deux méthodes. Tout d'abord, on modifie nos algorithmes pour étudier les tailles des premières branches et là où elles sont collées. Ensuite, pour prouver que nos modèles sont proches de leurs premières branches, on utilise la méthode de chaînage. Plus précisément, on divise nos algorithmes en grandes étapes, et on prouve qu'entre deux grandes étapes les objets que l'on construit ne changent pas beaucoup.

Dans le chapitre 2, on étudie les ICRT et notamment leur compacité et dimensions fractales. Dans le chapitre 3, on prouve des limites d'échelles des arbres à suite de degrés fixés, et on majore leur hauteur. Dans le chapitre 4, on prouve des limites d'échelles pour les multigraphes à suites de degrés fixés et surplus fixés, et on précise des connexions entre le modèle de configuration et les graphes multiplicatifs. Dans le chapitre 5, on invente une théorie d' \mathbb{R} -arbre plan ce qui nous permet de définir et d'étudier les ICRT plans, leurs "arbres-boucles", et des champs sur ces objets. Ce chapitre a pour but d'être appliqué à l'étude des cartes aléatoires à suite de face-degrés fixés.

English

In this thesis, we study several models of trees (\mathcal{D} -trees, \mathcal{P} -trees, ICRT) and graphs (configuration model, multiplicative graph) with fixed degree sequence. To this end, we introduce several new stick-breaking constructions for those models, which glue branches on one another. By studying those constructions, we prove several results on the geometry of our models.

To study those constructions, we mainly use two methods. First, we develop several modifications of our algorithms to study precisely the size of the first branches and where they are glued. Then, to prove that the trees or graphs are close from the first branches, we use the chaining method. More precisely, we split our algorithms in main steps, and prove that between two main steps the objects we construct do not change much.

In Chapter 2, we study ICRT and notably their compactness and fractal dimensions. In Chapter 3, we prove scaling limits for trees with fixed degree sequence, and we upper bound their height. In Chapter 4, we prove scaling limits for multitrees with fixed degree sequence and fixed surplus, and we explain several connections between the configuration model and multiplicative graphs. In Chapter 5, we develop a new theory of plane \mathbb{R} -tree, to define and study plane ICRT, the ICRT's looptrees, and some fields on those spaces. This chapter has been developed for the study of random maps with fixed face-degree sequence.

Chapter 1

Introduction

"Judge an artist not by the quality of what is framed and hanging on the walls, but by the quality of what's in the wastebasket."

Anon., quoted by Leslie Lamport

In this PhD thesis, we study uniform rooted trees with fixed degree sequence, and their limits, \mathcal{P} -trees and inhomogeneous continuum random trees (ICRT). This thesis contains four articles:

- [29] Compactness and fractal dimensions of inhomogeneous continuum random trees. Probability Theory and Related Fields. <https://doi.org/10.1007/s00440-022-01138-9>
- [28] Limit of trees with fixed degree sequence. arXiv:2110.03378
- [27] Limit of connected multigraph with fixed degree sequence. arXiv:2112.07725
- [30] Looptree, Fennec, and Snake of ICRT. arXiv:2203.10891

In this introduction we explain the backgrounds, results, methods, and applications of those articles. For applications, we will not detail all implications of our results to the old ones. Instead, we will explain how they may or can be used to prove new ones. Notably, some of our results were actually partly written to be used in two forthcoming articles, which should have been in this thesis:

- [26] Limit of looptree, fennec, and snake of trees with fixed degree sequence. In preparation.
- [25] A few notes on ICRT excursions. In preparation.

Since they are closely related to the others, we will explain some of their background and goals. Along the way, the framed algorithms and results are new.

Contents

1.1	Main motivations, background, and algorithms	10
1.1.1	The choice of the models	10
1.1.2	Geometric background	15
1.1.3	Main algorithms	19
1.1.4	The chaining method	24
1.1.5	Some new martingale inequalities	27
1.2	Overview of the chapters	28
1.2.1	Compactness and fractal dimensions of ICRT	28
1.2.2	Limit of trees with fixed degree sequence	31
1.2.3	Limit of connected multigraph with fixed degree sequence	35
1.2.4	Looptree, Fennec, and Snake of ICRT	38

1.1 Main motivations, background, and algorithms

1.1.1 The choice of the models

The configuration model

Our goal is to study graphs, also called networks. A graph is a set of vertices V connected by some edges $E \subset \{\{v, w\}, v \neq w \in V\}$. Let $(V_i)_{i \in \mathbb{N}}$ denote some vertices. Graphs are used in many sciences to encode and study a large amount of information in a concise way. And usually, having only partial information on real life networks, one use probabilistic models for their studies. The main idea is to infer some properties of real life network by proving that they hold with high probabilities for the theoretic graphs.

By doing so, we can use the properties of real life networks that we expect to hold, together with some randomness to "fill our gap of knowledge". To choose a model, we must make a tradeoff: On the one hand, with more information, the theoretic graphs usually get closer to real life networks. On the other hand, with more information, the graphs get harder to study.

In this thesis, our only information are the degrees. Given a graph $G = (V, E)$, we call the degree of $v \in V$, the number of edges incident to v , that is $\deg_G(v) := \#\{w \in V, \{v, w\} \in E\}$. Thus, given a degree sequence $(d_i)_{1 \leq i \leq n}$, we consider a random graph G uniform among all those with vertices $(V_i)_{1 \leq i \leq n}$ and such that for every $1 \leq i \leq n$, $\deg_G(v_i) = d_i$.

For technical reasons, we actually consider the configuration model introduced by Bollabas [31]. In this model, we allow multiple edges between two vertices, and "self-loops", that is edges between a vertex and itself. Since there are few multiple edges and few self-loops, and since they appear with small probability, the configuration model is close to uniform graphs with fixed degree sequence. Moreover, this model is much simpler to construct.

The configuration model is widely used as a reference model for social network: Vertices are peoples, edges represent a friendship relation, and the degrees are the numbers of friends. This model is often criticized as in social networks people tends to have connections with others that are similar in some way (ideas, jobs, locations...), while in this model the vertices are "indifferent". Despite those critics, this "indifference" leads to many useful combinatorial properties.

Multiplicative graphs

Another model of reference in both mathematics and computer science is the Erdős–Rényi graph introduced in [51]. In this thesis we are interested in one of its generalization: multiplicative graphs. In this model, given some weights $(w_i)_{1 \leq i \leq n}$, we sample a graph $G = (\{V_i\}_{1 \leq i \leq n}, E)$ such that independently for every $1 \leq i \neq j \leq n$, the edge $\{V_i, V_j\}$ is in E with probability $1 - e^{-w_i w_j}$.

This model is introduced, by Chung and Lu [41], Norros and Reitu [73], Britton and Deijfun [34], as a modification of the configuration model. Their idea is to replace the degrees by mean degrees, to keep the independency of the edges of the Erdős Renyi model, to simplify the study of the graphs. This proximity between those two models appear in several of our algorithms. And, we formalize this connection by proving that the configuration model can converge toward multiplicative graphs.

Multiplicative graphs may also be generalized as follows: Given some weights $(w_{i,j})_{1 \leq i, j \leq n}$, we may sample $G = (\{V_i\}_{1 \leq i \leq n}, E)$ such that independently for every $1 \leq i \neq j \leq n$, the edge $\{V_i, V_j\}$ is in E with probability $1 - e^{-w_{i,j}}$. The name "multiplicative graphs" then comes from $w_{i,j} = w_i w_j$. Although this generalization can modelize some "preferences" between the vertices, it loses many key combinatorial properties of the configuration model and multiplicative graphs.

From graphs to \mathcal{D} -trees and \mathcal{P} -trees

We first recall some background. Consider $G = (V, E)$ a graph. A tuple (v_1, v_2, \dots, v_k) is a path from v_1 to v_k in G if $\{v_1, v_2\}, \dots, \{v_{k-1}, v_k\} \in E$. We call k its length. Two vertices $v, v' \in V$ are connected if there is a path from v to v' . G is connected if every $v, v' \in V$ are connected. The connected components of G are the largest subsets $C \subset V$ such that every $v, v' \in C$ are connected. A path $(v_1, v_2, v_3, \dots, v_k, v_1)$ is a cycle if $k \geq 3$ and $(v_i)_{1 \leq i \leq k}$ are distinct. G is acyclic if it does not have any cycle. A tree is a connected acyclic graph.

The configuration model and multiplicative graphs can morally be divided in three regimes. When the degrees are small, the graphs only have small connected components. When the degrees are large, there is one large component that contains most of the vertices, and the others are small. In the critical regime, there are many connected components of "intermediate size". (See e.g. Dhara [46] Chapter 1 and reference therein for a detailed account of the literature on the subject.)

In this thesis we focus on the critical regime. In this regime the components are "tree like": By deleting a few edges, those components become trees. Reciprocally, the component can be obtained by adding some edges to some "tree versions" of the configuration model and multiplicative graphs, " \mathcal{D} -trees" and " \mathcal{P} -trees". We will detail this connection later, which is the heart of Chapter 4.

Given a degree sequence $\mathcal{D} = (d_i)_{1 \leq i \leq n}$, we call a \mathcal{D} -tree a random tree T uniform among all those with vertices $\{V_i\}_{1 \leq i \leq n}$ and such that for $1 \leq i \leq n$, $\deg_T(v_i) = d_i$. Given some weights $\mathcal{P} = (p_i)_{1 \leq i \leq n}$, we call a \mathcal{P} -tree a random tree T such that for every tree $t = (\{V_i\}_{1 \leq i \leq n}, E)$,

$$\mathbb{P}(T = t) \propto \prod_{e=\{V_i, V_j\} \in E} (p_i p_j), \quad (1.1)$$

writing \propto for proportional. For convenience, we use some tiny modifications of those models.

Some brief words on a generalization of \mathcal{P} -trees.

Similarly to multiplicative graphs one can generalize \mathcal{P} -trees as follows: Given some weights $\mathcal{W} = (w_{i,j})_{1 \leq i, j \leq n}$, we call a \mathcal{W} -tree a random tree T such that for every tree $t = (\{V_i\}_{1 \leq i \leq n}, E)$,

$$\mathbb{P}(T = t) \propto \prod_{e=\{V_i, V_j\} \in E} w_{i,j}.$$

Again \mathcal{P} -trees can be seen as "multiplicative-trees" with $w_{i,j} = p_i p_j$. Another important example of \mathcal{W} -tree are the uniform spanning trees: Consider $G = (\{V_i\}_{1 \leq i \leq n}, E_G)$ a connected graph. A spanning tree of G is a tree $T = (V, E)$ such that $E \subset E_G$. The adjacency matrix of G is $M_G = (\mathbf{1}_{\{V_i, V_j\} \in E_G})_{1 \leq i, j \leq n}$. It is easy to check that a M_G -tree is a uniform spanning tree of G .

\mathcal{W} -trees can be sampled with Aldous–Broder’s algorithm [9, 35] and Wilson’s algorithm [81]. Those algorithms are currently widely used for the study of uniform spanning trees. We refer for instance to [18, 19, 56, 59, 74, 76] for the study of uniform spanning trees in finite dimensions. Since our algorithms are also derived from Aldous–Broder’s algorithm, we expect that some of the methods developed in this thesis should be useful to study \mathcal{W} -trees.

Connection with Galton–Watson trees

Independently of the studies around the configuration model, Galton–Watson trees have received a tremendous amount of attention, and notably in combinatoric. This model was introduced by Galton and Watson in 1875, as a model of genealogical tree, to study the extinction of surnames. In this model, vertices represent peoples, and edges represent a parent-child relation. For simplicity, we consider only the people who give their surname, so everybody has one parent.

Galton–Watson trees can be sampled as follows: Start with an ancestor. It gives birth to a certain number of children according to a probability distribution μ on $\{0\} \cup \mathbb{N}$. Then each of those children gives birth independently of the others to a certain number of children according to μ . Then the same holds for their children, and so on. . .

To define Galton–Watson trees rigorously, we need a way to encode them. To be coherent with the rest of the thesis, we do not use Neveu’s notations, and instead use labeled rooted trees. A rooted graph (resp. tree) (V, E, ρ) is a graph (V, E) with a distinguished vertex ρ called the root. Think of the root as the ancestor. Since it has 0 parent, its number of children is its degree. Since the other vertices have one parent, their number of children is their degrees - 1. Hence given a probability distribution μ on $\{0\} \cup \mathbb{N}$, a Galton–Watson tree with vertices $\{V_i\}_{1 \leq i \leq n}$ is a random tree T such that for every rooted tree $t = (\{V_i\}_{1 \leq i \leq n}, E, \rho)$,

$$\mathbb{P}(T = t) \propto \prod_{i=1}^n (\deg_t(V_i) - 1 + \mathbf{1}_{V_i=\rho})! \prod_{i=1}^n \mu(\deg_t(V_i) - 1 + \mathbf{1}_{V_i=\rho}), \quad (1.2)$$

(where the factorials come from a counting argument.)

Before we explain how Galton–Watson and \mathcal{D} -trees are related, we need to clarify something: \mathcal{D} -trees stand for either of two closely related models. On the one hand, we consider a uniform tree T with true degree sequence $(d_i)_{1 \leq i \leq n}$, that is such that for every $1 \leq i \leq n$, V_i have degree d_i . On the other hand, we consider a uniform rooted tree T with fixed degree sequence $(d_i)_{1 \leq i \leq n}$, that is such that for every $1 \leq i \leq n$, V_i have d_i children. There is actually a simple bijection between those two models. From a tree T with true degree sequence $(d_i)_{1 \leq i \leq n}$ with $d_n = 1$, one can construct a rooted tree with fixed degree sequence $(d_i - 1)_{1 \leq i \leq n-1}$ by rooting T at the only vertex connected to V_n and by removing V_n . Since this bijection only removes a leaf, that is a vertex of degree 1, it does not change our results.

Then, by (1.2), a Galton–Watson tree conditioned on its degree sequence (the number of children of $(V_i)_{i \in \mathbb{N}}$) is a \mathcal{D} -tree. Hence, Galton–Watson trees can be studied in two parts: the degrees, and \mathcal{D} -trees. This thesis focus on \mathcal{D} -trees, while the degrees can be studied completely apart as a sequence of independent random variables conditioned by their sum.

Some brief words on the connections with multiplicative and additive coalescent

Coalescent processes arise naturally in many sciences (see e.g. Bertoin [22] or Berestycki [21]). The main principle is the following: we have some particles with some masses and those particles tends to fuse at some rate to form new particles. To study coalescent we must make some assumptions: First two disjoint fusions cannot hold at the same time. Then there is no fusion of three or more particles. Also we often assume the law of conservation of mass that is (a) when two particles fuse, the mass of the fusion is the sum of their mass, and (b) outside of fusions masses do not change. Finally, the rate at which two particles fuse is a function $\kappa : \mathbb{R}^{+2} \rightarrow \mathbb{R}^+$ of their masses.

This thesis is related to two coalescent. Indeed, the multiplicative graphs, and by extension the configuration model, are related to the multiplicative coalescent, where $\kappa : (x, y) \mapsto xy$. Also \mathcal{P} -trees, and by extension \mathcal{D} -trees, are related to the additive coalescent where $\kappa : (x, y) \mapsto x + y$. Since those relations helped us find our algorithms, we will detail them. Moreover, they showcase two crucial points of view on our models. On the one hand, the configuration model and \mathcal{D} -trees may be seen as multiplicative graphs and \mathcal{P} -trees with "mass loss". On the other hand, multiplicative graphs and \mathcal{P} -trees are configuration models and \mathcal{D} -trees with "infinite degrees and leaves".

Connection with multiplicative coalescent

Informally the configuration model is constructed as follows. Consider $\mathcal{D} = (d_i)_{1 \leq i \leq n}$ a degree sequence. Consider around each V_i , $1 \leq i \leq n$, d_i half edges. The edges of the configuration model are obtained by uniformly fusing those half-edges pair by pair.

Formally, we encode a multigraph G on $\{V_i\}_{i \in \mathbb{N}}$ as a matrix $(\#_{i,j}(G))_{i,j \in \mathbb{N}}$, where for every $i, j \in \mathbb{N}$, $\#_{i,j}$ represents the number of edges between V_i and V_j . We call a function $f : I \mapsto I$ a matching if $f \circ f = Id$ and for every $x \in I$, $f(x) \neq x$.

Algorithm 1.1. Construction of the configuration model from $\mathcal{D} = (d_1, \dots, d_s)$:

- Let $f = (f_1, f_2)$ be a uniform matching of $\{(i, j)\}_{1 \leq i \leq s, 1 \leq j \leq d_i}$.
- The configuration model is the random multigraph $\text{CM}^{\mathcal{D}}$ with vertices $(V_i)_{1 \leq i \leq s}$ and such that for $1 \leq i \leq s$, $\#_{i,i}(\text{CM}^{\mathcal{D}}) := \frac{1}{2} \sum_{a=1}^{d_i} \mathbf{1}_{f_1(i,a)=i}$ and for $1 \leq i \neq j \leq s$,

$$\#_{i,j}(\text{CM}^{\mathcal{D}}) := \sum_{a=1}^{d_i} \mathbf{1}_{f_1(i,a)=j} = \sum_{a=1}^{d_j} \mathbf{1}_{f_2(j,a)=i}.$$

Informally, the configuration model can be constructed by recursively forming a full edge with a uniform pair of half edges. By doing so, we have at each steps a collection of "half graphs" which have some full edges and some half edges that still needs to be completed (see Figure 1.1). Since when two half edges fuse, their half graphs fuse, this construction may be seen as a coalescent. To add some independency in our construction, we consider that two half edges fuse at rate one. Two half graphs with n_1 and n_2 half edges then fuse together at rate $n_1 n_2$. And when they fuse, since two of their half edges fuse, their fusion have $n_1 + n_2 - 2$ half edges. In addition a half graph with n_1 half edges, can have two half edges that fuse together at rate $\binom{n_1}{2}$.

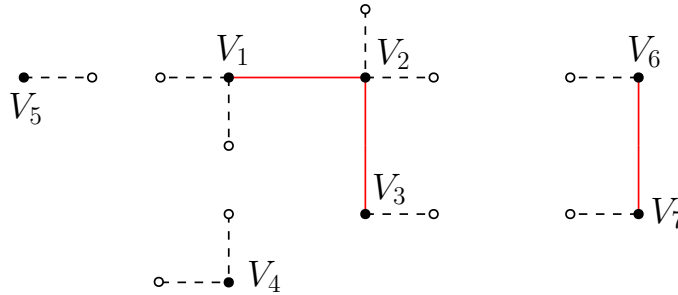


Figure 1.1: A partial construction of a configuration model with $\mathcal{D} = (3, 4, 2, 2, 1, 2, 2)$. Half edges are in black dashed. The full edges of the configuration model already constructed are in red.

This coalescent is thus a multiplicative coalescent with some "mass loss" when half edges fuse. This mass loss appears repeatedly in this thesis, and is the source of the new technical difficulties of dealing with the configuration model and \mathcal{D} -trees.

We can actually modify the configuration model to preserve the mass. Indeed, by multiplying all the degrees by a constant $c \rightarrow \infty$, we proportionally preserve the rates at which the half graphs fuse, and the mass loss is negligible. However, we obtain at the limit infinitely many edges between each pair of vertices. To avoid this issue, we temporary add some leaves that we remove on the limit. Moreover, we prove in Chapter 4 that with the right amount of leaves the configuration model converges toward multiplicative multigraph:

Algorithm 1: The multiplicative multigraph from $\mathcal{P} = (p_1, \dots, p_n) \in \mathbb{R}^{+*n}$:

- Let $(N_{i,j}^{\mathcal{P}})_{1 \leq i, j \leq s}$ be independent Poisson random variables, such that for $1 \leq i \leq s$, $N_{i,i}^{\mathcal{P}}$ have mean $p_i^2/2$ and for $1 \leq i \neq j \leq s$, $N_{i,j}^{\mathcal{P}}$ have mean $p_i p_j$.
- The multiplicative multigraph is the random multigraph $\text{MG}^{\mathcal{P}+}$ with vertices $(V_i)_{1 \leq i \leq s}$ and such that for every $1 \leq i, j \leq s$, $\#_{i,j}(\text{MG}^{\mathcal{P}+}) := N_{i,j}^{\mathcal{P}}$.

Then by deleting the multiple edges and loops we recover multiplicative graphs. This is not surprising since multiplicative graphs were introduced [41, 73, 34] as modifications of the configuration model which preserves the independency of the edges, and because the "mass loss" is due to the lack of independence in the configuration model. Furthermore, the connection between Erdős–Rényi graph and multiplicative coalescent was already discovered by Aldous [12]. Nonetheless, understanding that multiplicative graphs are "configuration models with infinite degree and leaves" helped us find many of our algorithms and results.

Connection with additive coalescent

Recall that \mathcal{D} -trees and \mathcal{P} -trees are "tree versions" of the configuration model and multiplicative graphs. We discuss here "tree versions" of the coalescents presented before.

First let us introduce some notions. A plane tree is a rooted tree with for each vertex an order for its children. We encode a plane tree T as a matrix $(V_{i,j}(T))_{1 \leq i \leq n, 1 \leq j \leq d_i}$, where for $1 \leq i \leq n$, $1 \leq j \leq d_i$, $V_{i,j}(T)$ is the j^{th} child of V_i . A matrix $(V_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d_i}$ is a rooted tree if and only if $((V_i)_{1 \leq i \leq n}, (\{V_i, V_{i,j}\})_{1 \leq i \leq n, 1 \leq j \leq d_i})$ is a tree, all vertices but the root have 1 parent, and the root has 0 parent. As a result, we can construct plane \mathcal{D} -trees as follows:

Algorithm 1.2. Matching construction of a plane \mathcal{D} -tree from $\mathcal{D} = (d_1, \dots, d_n)$:

- Let $(A_i)_{1 \leq i \leq n}$ be a uniform permutation of $(V_i)_{1 \leq i \leq n}$.
- Let $((B_i, j_i))_{1 \leq i \leq n-1}$ be a uniform permutation of $((V_i, j))_{1 \leq i \leq n, 1 \leq j \leq d_i}$.
- Let $T = (V_{i,j}(T))_{1 \leq i \leq n, 1 \leq j \leq d_i}$ be such that for every $1 \leq i \leq n-1$, $V_{B_i, c_i}(T) = A_i$.
- Let $T^{\mathcal{D}}$ be the matrix T conditioned by being a rooted tree.

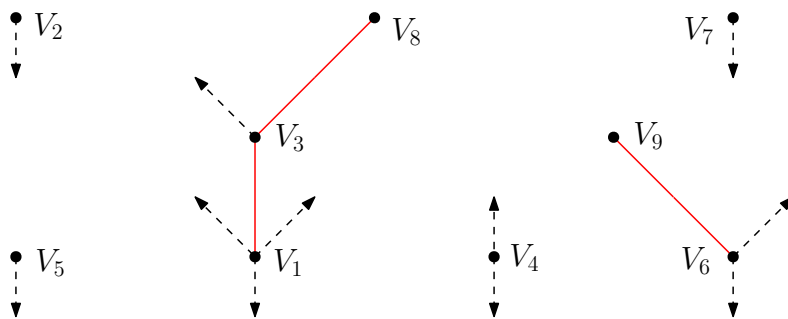


Figure 1.2: A partial coalescent construction of a \mathcal{D} -tree with $\mathcal{D} = (3, 0, 2, 1, 0, 2, 0, 0, 0)$. Half edges are in black dashed. The full edges already constructed are in red. The arrows provide the orientation of the half edges (downward/upward).

Algorithm 1.2 can morally be interpreted as follows. First consider for each $1 \leq i \leq n$, one half downward edge and d_i half upward edges around V_i . The half downward edges are encoded by $(V_i)_{1 \leq i \leq n}$, and the half upward edges by $((V_i, j))_{1 \leq i \leq n, 1 \leq j \leq d_i}$. By permuting those half edges we are actually forming pair between those half edges. And the final edges are the fusions of an half downward edge with an half upward edge.

As for the configuration model we can construct those pairs recursively. We must however take care of not forming any cycle by pairing an half upward edge with its wrong half downward edge. By doing so, we have at each step a collection of "half trees" which have some half upward edges and one half downward edge (see Figure 1.2). And when we pair two half edges two half trees fuse. Moreover, by a counting argument (see Addario-Berry, Barrett [3]), at each step the choice of the pair of a downward edge and an upward edge is uniform among those that does not form a cycle. Then we may consider that each such pair fuse a rate 1. By doing so, two half tree with a_1 and a_2 upward edges fuse at rate $a_1 + a_2$ and their fusion have $a_1 + a_2 - 1$ upward edges.

Hence, this coalescent is an additive coalescent with "mass loss" when two half edges fuse. Here again we can modify \mathcal{D} -trees to preserve the mass by multiplying the degrees by $c \rightarrow \infty$ and adding leaves. Moreover, we prove in Chapter 3, that in this case \mathcal{D} -trees converge toward \mathcal{P} -trees. Finally we recover the link between additive coalescent and \mathcal{P} -trees (see Evans, Pitman [52]).

1.1.2 Geometric background

"Scaling limit" as a main objective.

Now that we have introduced our models, let us explain what kind of properties we are studying. We want to study the "geometry" of our graphs. More precisely, let $G = (V, E)$ denote a graph. One can put a metric d_G on G such that for every $x, y \in V$ the graph distance between x and y , $d_G(x, y)$ is the minimum length of a path in G between x and y . We are interested in the distance between some random points, in the diameter of G , in the height of a tree $T = (V, E, \rho)$ that is in $\max_{x \in V} d_T(x, \rho)$, in more complex object such as a random walk...

We study such properties for "large" random graphs, that is as the number of vertices diverges. To this end we follow a method proposed by Aldous in [10, 8, 11]: Instead of studying each property individually, we study all of them at once by proving and studying some "scaling limits".

The "scaling limits" of graphs are then metric spaces. And for trees, the limits are \mathbb{R} -trees (geodesic loopless metric spaces, see Le Gall [61]). Since we want to study the geometry of the limits, we naturally look at their compactness and fractal dimensions.

"Scaling limits" are also interesting to provide universal point of views on several models. Indeed, if several models share the same limits, then they all asymptotically share similar properties. As a result, each model may be a way to understand or even study many others. Then, the properties that hold for many models tend to also hold in real life networks.

Brief overview of the past results on the geometry of our spaces

In the pioneer papers [10, 8, 11], Aldous introduce two main approaches to study random trees: stick-breaking construction, and tree encodings by real process. Le Gall [61] throughly develop tree encodings. This allows Le Gall and Le Jan [64, 63] to construct Lévy trees from Lévy processes. From there Le Gall and Duquesne [50, 49] study in detail the geometry of Lévy trees, and notably prove that Galton–Watson trees converge toward Lévy trees.

At the same period, using stick-breaking construction, Aldous, Camarri, and Pitman [16, 40] prove that \mathcal{P} -trees converge toward ICRT. Then Aldous Miermont, and Pitman [14, 15] introduce encodings of \mathcal{P} -trees and ICRT. They however fail to obtain precise results on the geometry of \mathcal{P} -trees and ICRT, and so leave some conjectures by comparison with Lévy trees [50, 49].

The objective of Chapter 2, 3 is to solve those conjectures using stick-breaking constructions. Notably, we prove that \mathcal{D} -trees converge toward ICRT. This implies, since Galton–Watson trees are mix of \mathcal{D} -trees with random degree sequence, that Lévy trees are mix of ICRT. We thus extend several results from [50, 49].

Due to the many development around Galton–Watson trees, tree encodings have earned an interest of their own. For this reasons we come back in Chapter 5 on the encodings of ICRT. Although those encodings are real processes, our approach is entirely based on stick-breaking construction, and we only see those processes as geometric objects on \mathbb{R}^2 constructed from ICRT.

In parallel, scaling limits have been used to study the configuration model (see e.g. [42, 54, 47]) and multiplicative graphs (see e.g. [37, 36]). However, for technical reasons, the previous works consider random degrees sequence. This weaken the main strength of our models which is to incorporate arbitrary degree sequence. We remove this technical assumption in Chapter 4.

Fractal dimensions

This part can be skimmed on a first reading. In the entire part X is a metric space, and for every $x \in X$, $\varepsilon > 0$, $B(x, \varepsilon)$ denotes the closed ball centered at x with radius ε .

Definition. (Minkowski dimensions) For every $\varepsilon > 0$ let N_ε be the minimal number of closed balls of radius ε to cover X . Define the Minkowski lower box and upper box dimensions respectively by

$$\underline{\dim}(X) := \liminf_{l \rightarrow \infty} \frac{\log N_{1/l}}{\log l} \quad \text{and} \quad \overline{\dim}(X) := \limsup_{l \rightarrow \infty} \frac{\log N_{1/l}}{\log l}.$$

Definition. (Packing dimension) For every $s \geq 0$ and $A \subset X$ let

$$P_0^s(A) := \limsup_{\delta \rightarrow 0} \left\{ \sum_{i \in I} \text{diam}(B_i)^s \mid \{B_i\}_{i \in I} \text{ are disjoint balls } B(x, r) \text{ with } x \in A \text{ and } r \leq \delta \right\}.$$

and

$$P^s(X) := \inf \left\{ \sum_{i=1}^{\infty} P_0^s(A_i) \mid X \subset \bigcup_{i=1}^{\infty} A_i \right\}.$$

Then P^s is a decreasing function of s , and we define the packing dimension of X as

$$\dim_P(X) := \sup\{s, P^s(X) < \infty\}.$$

Definition. (Hausdorff dimension) For every $s, r \geq 0$ write

$$H_r^s(X) := \inf_{\text{diam}(A_i) \leq r} \left\{ \sum_{i=1}^{\infty} \text{diam}(A_i)^s \mid X \subseteq \bigcup_{i=1}^{\infty} A_i \right\}.$$

The Hausdorff dimension of X is defined by

$$\dim_H(X) := \sup \left\{ s, \sup_{r \in \mathbb{R}^+} H_r^s(X) < \infty \right\}.$$

We have the well-known inequalities (see e.g. Falconer [53]). As a result, we only need to lower bound the Hausdorff and Packing dimensions, and to upper bound the Minkowski dimensions.

Lemma 1.1.1. For every metric space X we have

$$\dim_H(X) \leq \underline{\dim}(X) \leq \overline{\dim}(X) \quad \text{and} \quad \dim_H(X) \leq \dim_P(X) \leq \overline{\dim}(X).$$

To lower bound the Hausdorff and Packing dimensions, we do not use their definitions, and we use instead the following extension of Theorem 6.9, and Theorem 6.11 from Matilla [70].

Lemma 1.1.2. Let p be a Borel probability measure on X and $s \in \mathbb{R}^+$.

- a) If p -almost everywhere $\liminf p(B(x, \varepsilon))\varepsilon^{-s} < +\infty$ as $\varepsilon \rightarrow 0$, then $\dim_P(X) \geq s$.
- b) If p -almost everywhere $p(B(x, \varepsilon)) = O(\varepsilon^s)$ as $\varepsilon \rightarrow 0$, then $\dim_H(X) \geq s$.

Some brief general words on transport theory

To formally define "scaling limit", we need to explain the meaning of "having a similar geometry". To this end, we use some distances from transport theory (see e.g. Villani [80] Chapter 27). Those distances can be decomposed in two main parts: First a way to "move" a metric space with eventually some additional structure, and a way to "deform" them. The distance between two metric space then consists in the minimal cost of the deformations.

To study the graph distances, we must conserve the metrics when we "move" the metric spaces. For this reason, isometries (functions from a metric space to another which preserves the metrics) are our references to move the metric spaces. In addition, our metrics space may be "pointed", to force that while we move our spaces, some specific vertices reach some specific locations. For instance, to study the height, we must point the trees at their root in order to preserve the height. Similarly, when we consider the distance between random vertices in a graph, we must preserve the law of the distance and so the probability measure on the metric spaces.

On the other hand, when we choose a way to "deform" the spaces we must estimate by how much some quantities differ (height, diameter, distances between some vertices...). So we must be able to change those quantities, and the cost of the deformations must be relevant for them.

In the rest of the subsection we formally define the main metrics we consider. The reader may wish to simply skim this part on a first reading, and refer back to it if needed.

Gromov–Prokhorov (GP) topology

A measured metric space is a triple (X, d, μ) such that (X, d) is a Polish (separable complete metric) space and μ is a Borel probability measure on X . Two such spaces $(X, d, \mu), (X', d', \mu')$ are called isometry-equivalent iff there exists an isometry $f : X \rightarrow X'$ such that if $f_*\mu$ is the image of μ by f then $f_*\mu = \mu'$. Let \mathbb{K}_{GP} be the set of isometry-equivalent classes of measured metric space. Given a measured metric space (X, d, μ) , we write $[X, d, \mu]$ for the isometry-equivalence class of (X, d, μ) and frequently use the notation X for either (X, d, μ) or $[X, d, \mu]$.

We now recall the definition of the Prokhorov's distance. Consider a metric space (X, d) . For every $A \subset X$ and $\varepsilon > 0$ let $A^\varepsilon := \{x \in X, d(x, A) < \varepsilon\}$. Then given two (Borel) probability measures μ, ν on X , the Prokhorov distance between μ and ν is defined by

$$d_P(\mu, \nu) := \inf\{\varepsilon > 0: \mu\{A\} \leq \nu\{A^\varepsilon\} \text{ and } \nu\{A\} \leq \mu\{A^\varepsilon\}, \text{ for all Borel set } A \subset X\}.$$

The Gromov–Prokhorov (GP) distance is an extension of the Prokhorov's distance: For every $(X, d, \mu), (X', d', \mu') \in \mathbb{K}_{GP}$ the Gromov–Prokhorov distance between X and X' is defined by

$$d_{GP}((X, d, \mu), (X', d', \mu')) := \inf_{S, \phi, \phi'} d_P(\phi_*\mu, \phi'_*\mu'),$$

where the infimum is taken over all metric spaces S and isometric embeddings $\phi : X \rightarrow S, \phi' : X' \rightarrow S$. d_{GP} is indeed a distance on \mathbb{K}_{GP} and $(\mathbb{K}_{GP}, d_{GP})$ is a Polish space (see e.g. [1]).

We use another convenient characterization of the GP topology which relies on convergence of distance matrices: For every measured metric space (X, d^X, μ^X) let $(x_i^X)_{i \in \mathbb{N}}$ be a sequence of i.i.d. random variables of common distribution μ^X and let $M^X := (d^X(x_i^X, x_j^X))_{(i,j) \in \mathbb{N}^2}$. We have the following result from [66],

Lemma 1.1.3. *Let $(X^n)_{n \in \mathbb{N}} \in \mathbb{K}_{GP}^{\mathbb{N}}$ and let $X \in \mathbb{K}_{GP}$ then $X^n \xrightarrow{GP} X$ as $n \rightarrow \infty$ if and only if M^{X^n} converges in distribution toward M^X .*

With the next extension, we morally need less information on the limit to prove convergence. Its proof, in Chapter 3, is short and use only basic topology.

Lemma 1

Let $(X^n)_{n \in \mathbb{N}} \in \mathbb{K}_{\text{GP}}^{\mathbb{N}}$ and let $X \in \mathbb{K}_{\text{GP}}$. Let $(y_i^X)_{i \in \mathbb{N}}$ be a sequence of random variables on X and let $N^X := (d^X(y_i^X, y_j^X))_{(i,j) \in \mathbb{N}^2}$. If

$$M^{X^n} \xrightarrow{(d)} N^X \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^n \delta_{y_i^X} \xrightarrow{(d)} \mu^X,$$

then $X^n \xrightarrow{\text{GP}} X$ and thus M^X and N^X have the same distribution.

Gromov–Hausdorff (GH) topology

Let \mathbb{K}_{GH} be the set of isometry-equivalent classes of compact metric space. For every metric space (X, d) , we write $[X, d]$ for the isometry-equivalent class of (X, d) , and frequently use the notation X for either (X, d) or $[X, d]$.

For every metric space (X, d) , the Hausdorff distance between $A, B \subset X$ is given by

$$d_H(A, B) := \inf\{\varepsilon > 0, A \subset B^\varepsilon, B \subset A^\varepsilon\}.$$

The Gromov–Hausdorff distance between $(X, d), (X', d') \in \mathbb{K}_{\text{GH}}$ is given by

$$d_{\text{GH}}((X, d), (X', d')) := \inf_{S, \phi, \phi'} (d_H(\phi(X), \phi'(X'))),$$

where the infimum is taken over all metric spaces S and isometric embeddings $\phi : X \rightarrow S$, $\phi' : X' \rightarrow S$. d_{GH} is indeed a distance on \mathbb{K}_{GH} and $(\mathbb{K}_{\text{GH}}, d_{\text{GH}})$ is a Polish space. (see e.g. [1])

Gromov–Hausdorff–Prokhorov (GHP) topology

Let $\mathbb{K}_{\text{GHP}} \subset \mathbb{K}_{\text{GP}}$ be the set of isometry-equivalent classes of compact measured metric space. The Gromov–Hausdorff–Prokhorov distance between $(X, d, \mu), (X', d', \mu') \in \mathbb{K}_{\text{GHP}}$ is given by

$$d_{\text{GHP}}((X, d, \mu), (X', d', \mu')) := \inf_{S, \phi, \phi'} (d_P(\phi_*\mu, \phi'_*\mu') + d_H(\phi(X), \phi'(X'))),$$

where the infimum is taken over all metric spaces S and isometric embeddings $\phi : X \rightarrow S$, $\phi' : X' \rightarrow S$. d_{GHP} is indeed a distance on \mathbb{K}_{GHP} and $(\mathbb{K}_{\text{GHP}}, d_{\text{GHP}})$ is a Polish space. (see [1])

Note that GHP convergence implies GP convergence and that random variables GHP measurable are also GH measurable. For all $[X, d, p] \in \mathbb{K}_{\text{GHP}}$, let $[X, d]$ denote its natural projection on \mathbb{K}_{GH} . Note that GHP convergence implies GH convergence of the projections on \mathbb{K}_{GH} , then that the projection on \mathbb{K}_{GH} is a measurable function. Moreover, we prove the following result in Chapter 3:

Lemma 2

Let $([X^n, d^n, p^n])_{n \in \mathbb{N}}$ and $[X, d, p]$ be GHP measurable random variables in \mathbb{K}_{GHP} . Assume that almost surely $[X, d, p]$ have full support. Assume that $([X^n, d^n, p^n])_{n \in \mathbb{N}}$ converges weakly toward $[X, d, p]$ in a GP sens, and that $([X^n, d^n])_{n \in \mathbb{N}}$ converges weakly toward $[X, d]$ in a GH sens. Then $([X, d, p])_{n \in \mathbb{N}}$ converges weakly toward $[X, d, p]$ in a GHP sens.

The proof morally goes as follows: We use the GP and GH convergence to prove GHP tightness. We then consider a GHP limit $[X', d', p']$, which is thus also a GP and a GH limit. So we may assume that almost surely $[X, d, p] =^{GP} [X', d', p']$. On the one hand, the GH convergences imply that X and X' have the same size. On the other hand, the GP convergences imply that the support of p and p' have the same size. As a result, since p has full support, p' must also has full support. We deduce with basic topology that almost surely $[X, d, p] =^{GHP} [X', d', p']$.

1.1.3 Main algorithms

Notations: Throughout this thesis, similar variables for our models share similar notations. To avoid any ambiguity, the models that we are using and their parameters are indicated by superscripts $\mathcal{D}_n, \mathcal{P}_n, \Theta_n \dots$. We often drop those superscripts when the context is clear.

Time complexity and geometry

Up to this point, we defined our models, we explained some of their combinatorial interpretations, and we decided to study their scaling limits. We now present our main tools to study their geometry. First, since we research some complex properties, we want "simple" constructions for our models. The "simplicity" of a construction can be measured with its time complexity, that is the number of steps needed to run it. In research, having low complexity, means that we need to study less steps, and that we may access more directly to some specific informations.

Then, constructing a tree of size n , requires at least n steps to add each vertex. We call a construction linear if it takes $O(n)$ steps. And a construction may be seen as "simple" if it is linear. In fact, many linear algorithms, notably coalescents, are hard to use, and we can somehow be faster. Indeed, to measure the distance between a few vertices, we do not need to construct the whole tree as we only need the paths between those vertices. Moreover, studying the distance between random vertices is by definition directly equivalent to study the GP geometry of our trees.

More generally, we use the notion of leaf tightness introduced by Aldous [11]. First, a subgraph (resp. subtree) (V', E') of a graph (resp. tree) (V, E) , is a graph such that $V' \subset V$ and $E' \subset E$. Given a tree $T = (V, E)$ and some vertices $V' \subset V$, the subtree $T' = (W, E')$ of T spanned by V' , is the subtree of T such that for every $v \in V$, $v \in W$ iff v is on a path between two vertices in V' . Morally, a family of tree is leaf tight if the subtrees spanned by a few leaves approach well the trees.

Formally, consider some random trees $(T^n)_{n \in \mathbb{N}}$, with eventually some additional structures. Let $(d^n)_{n \in \mathbb{N}}$ be the graph distances on $(T^n)_{n \in \mathbb{N}}$, and let $(\lambda^n)_{n \in \mathbb{N}}$ be some scaling factors in \mathbb{R}^{+*} . For every $n \in \mathbb{N}$ let $(L_i^n)_{i \in \mathbb{N}}$ be independent uniform leaves in T^n . And let for every $n \in \mathbb{N}$, $k \in \mathbb{N}$, T_k^n be the subtree of T^n spanned by L_1^n, \dots, L_k^n . Finally let d_* be a distance on metric spaces, with eventually some additional structures. We say that $(T^n)_{n \in \mathbb{N}}$ is leaf tight for d_* if

$$\forall \varepsilon > 0, \limsup_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(d_*((T_k^n, \lambda^n d^n), (T^n, \lambda^n d^n)) > \varepsilon) = 0.$$

Aggregations of paths/Stick-breaking constructions

Given a tree T , a subtree T' of T , a vertex $v \in T$, the projection of v on T' is the closer vertex from v in T' . Any tree T with root ρ and leaves L_1, \dots, L_k is an union of some paths (see Figure 1.3): the path between L_1 and ρ , the path between L_2 and its projection on the subtree spanned by ρ, L_1 , the path between L_3 and its projection on the subtree spanned by ρ, L_1, L_2 , and so on. . .

Based on this description, we use Aldous–Bröder's algorithm [9, 35] to construct our trees. Let $(A_i)_i$ be a sequence of vertices. Start with a root vertex A_1 then recursively add the edge $\{A_{i-1}, A_i\}$ when A_i is "new", that is when $A_i \notin \{A_j\}_{j < i}$. In this thesis, we add an edge between A_{i-1} and a new leaf when A_i is a "repetition", that is when $A_i \in \{A_j\}_{j < i}$. By doing so, the previous paths corresponds exactly to the vertices between two repetitions.

A better way to understand this aggregation of paths is through stick-breaking constructions. Morally, first consider $(A_i)_i$ as a line of distinct vertices. Then recall that there is no edges $\{A_{i-1}, A_i\}$ when A_i is not new, so cut those edges. This gives you a collection of paths or sticks. Finally identify the vertices which are equal, or equivalently glue the paths at the repeated vertices. So our trees can be seen as a collection of sticks glued on one another. And we can study their geometry through the positions of the cuts and the positions of the glue-points.

Stick breaking construction of \mathcal{D} -trees

A sequence $(d_i)_{1 \leq i \leq s}$ is a degree sequence of a rooted tree if and only if $\sum_{i=1}^s d_i = s - 1$, and by convention $d_1 \geq d_2 \geq \dots \geq d_s$. Let $\Omega_{\mathcal{D}}$ be the set of such sequences. Recall that we encode a plane tree T as a matrix $(V_{i,j})_{1 \leq i \leq n, 1 \leq j \leq d_i}$, where $V_{i,j}$ denote the j^{th} child of V_i . Also, for every $\mathcal{D} \in \Omega_{\mathcal{D}}$ we let $L_1^{\mathcal{D}}, L_2^{\mathcal{D}}, \dots$ be the leaves (that is the vertices V_{a_1}, V_{a_2}, \dots with $a_1 \leq a_2 \leq \dots$ and $d_{a_1} = d_{a_2} = \dots = 0$). For every graph $G = (V, E)$ and edge $e = \{v_1, v_2\}$, $G \cup e$ denotes the graph $(V \cup \{v_1, v_2\}, E \cup \{e\})$. We say that a vertex $v \in G$ if $v \in V$. The following algorithm is a bijection between some permutations and plane trees with degree sequence $\mathcal{D} = (d_1, d_2, \dots, d_s)$.

Algorithm 2: Stick-breaking construction of a plane \mathcal{D} -tree (see Figure 1.3).

- Let $(A_i^{\mathcal{D}}, U_i^{\mathcal{D}})_{1 \leq i \leq s-1}$ be a uniform permutation of $\{V_i, j\}_{1 \leq i \leq s, 1 \leq j \leq d_i}$.
- For every $1 \leq i \leq s - 1$ let

$$W_{A_i, U_i} = \begin{cases} A_{i+1} & \text{if } A_{i+1} \notin \{A_j\}_{j \leq i}. \\ L_{\inf\{k, L_k \notin T_{i-1}\}} & \text{if } A_i \in \{A_j\}_{j \leq i} \text{ or } i = s - 1. \end{cases}$$

- Let $T^{\mathcal{D}}$ be the plane tree $(W_{i,j})_{1 \leq i \leq s, 1 \leq j \leq d_i}$.

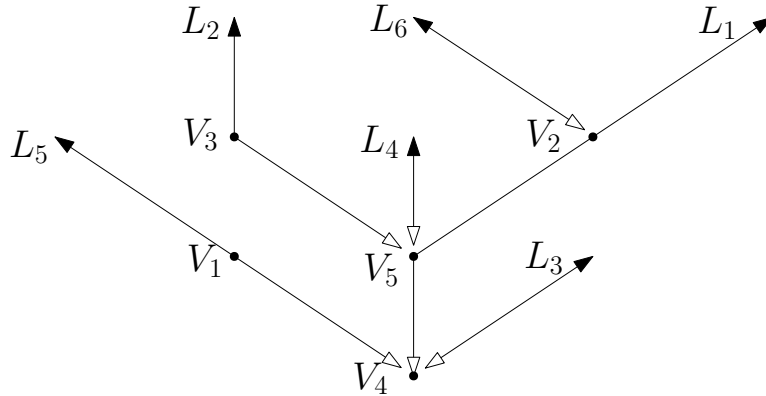


Figure 1.3: Stick breaking construction of a plane \mathcal{D} -tree with $((A_i^{\mathcal{D}}, U_i^{\mathcal{D}}))_{1 \leq i \leq s-1} = ((V_4, 2), (V_5, 3), (V_2, 2), (V_5, 1), (V_3, 1), (V_4, 3), (V_5, 2), (V_4, 1), (V_1, 1), (V_2, 1))$. The exploration starts at V_4 then follows the white-black arrow toward L_1 , then jumps at V_5 toward L_2 and so on...

Let us give a combinatorial interpretation of Algorithm 2. First, recall the connection between plane \mathcal{D} -trees and additive coalescent. In particular, recall that for every $1 \leq i \leq s$, we consider one downward and d_i upward edges around V_i . Also, we can construct plane \mathcal{D} -trees by recursively fusing a uniform pair of a downward edge and an upward edge that does not form any cycle. We now instead sample only the upward edges, and choose the downward edges. More precisely, we first sample a permutation of the upward edges. Then we may fuse at each step the upward edge with the downward edge corresponding to the next upward edge. Those fusions have several issues. First, we try to fuse upward edges with a downward already fused. Also, we are forgetting the leaves. To avoid those problems, instead of fusing some upward edges with a downward edge already fused, we instead fuse them with a downward edge of a new leaf.

This construction may be seen in a different order: Start at L_1 . Repeat until you hit the root: look at your downward edge, look at the upward edge fused with this downward edge, go to the corresponding vertex... Next restart at L_2 and go down as previously to the subtree spanned by the root and L_1 . Repeat for L_3, L_4, \dots . In a wider setting, this order is used in Wilson's algorithm,

which is also a modification of Aldous–Broder algorithm to recursively look at the subtrees spanned by specific vertices. We changed this order thanks to the many symmetries of \mathcal{D} -trees.

In Chapters 3–4, we use the following simplification of Algorithm 2, which was independently found by Addario-Berry, Donderwinkel, Maazoun, and Martin in [6] while I was writing this thesis.

Algorithm 3: Stick-breaking construction of a rooted \mathcal{D} -tree.

- Let $(A_i^{\mathcal{D}})_{1 \leq i \leq s-1}$ be a uniform \mathcal{D} -tuple (tuple such that $\forall i \in \mathbb{N}$, V_i appears d_i times).
- Let $T_1^{\mathcal{D}} := (\{A_1\}, \emptyset)$ then for every $2 \leq i \leq s$ let

$$T_i^{\mathcal{D}} := \begin{cases} T_{i-1} \cup \{A_{i-1}, A_i\} & \text{if } A_i \notin T_{i-1}. \\ T_{i-1} \cup \{A_{i-1}, L_{\inf\{k, L_k \notin T_{i-1}\}}\} & \text{if } A_i \in T_{i-1} \text{ or } i = s. \end{cases}$$

- Let $T^{\mathcal{D}}$ denotes the rooted tree (T_s, A_1) .

Stick-breaking construction of \mathcal{P} -trees

Historically, \mathcal{P} -trees was first introduced by Pitman and Evans [52] to study additive coalescent. They also noticed that \mathcal{P} -trees can be constructed by Aldous–Bröder algorithm. This construction was then used by Aldous, Camarri, Pitman [16, 40] to study the geometry of \mathcal{P} -trees.

In this thesis, we slightly extends this model to consider all the possible limits of \mathcal{D} -trees. Morally, in terms of coalescent, we are adding some "dust", which consists in infinitely many particles of mass null with total mass no null. From a tree point of view, many vertices of small degrees may contribute together to the total degrees.

Let $\Omega_{\mathcal{P}}$ be the set of sequence $(p_i)_{i \in \mathbb{N} \cup \{\infty\}}$ in \mathbb{R}^+ such that $\sum_{i=1}^{\infty} p_i + p_{\infty} = 1$, $p_1 > 0$ and $p_1 \geq p_2 \geq \dots$. Let $(V_{i,\infty})_{i \in \mathbb{N}}$ and $(L_i)_{i \in \mathbb{N}}$ be some vertices different from $(V_i)_{i \in \mathbb{N}}$. For every $\mathcal{P} \in \Omega_{\mathcal{P}}$, the \mathcal{P} -tree is the tree constructed as follows:

Algorithm 4: Definition of the \mathcal{P} -tree for $\mathcal{P} = (p_i)_{i \in \mathbb{N} \cup \{\infty\}} \in \Omega_{\mathcal{P}}$.

- Let $(A_i^{\mathcal{P}})_{i \in \mathbb{N}}$ be i.i.d. random variables such that for all $i \in \mathbb{N}$, $\mathbb{P}(A_1^{\mathcal{P}} = V_i) = p_i$.
- For every $i \in \mathbb{N}$, let $B_i^{\mathcal{P}} = A_i$ if $A_i \in \mathbb{N}$, and let $B_i^{\mathcal{P}} = V_{\infty,i}$ otherwise.
- Let $T_1^{\mathcal{P}} := (\{B_1\}, \emptyset)$ then for every $i \geq 2$ let

$$T_i^{\mathcal{P}} := \begin{cases} T_{i-1} \cup \{B_{i-1}, B_i\} & \text{if } B_i \notin T_{i-1}. \\ T_{i-1} \cup \{B_{i-1}, L_{\inf\{k \in \mathbb{N}, L_k \notin T_{i-1}\}}\} & \text{if } B_i \in T_{i-1}. \end{cases}$$

- Let $T^{\mathcal{P}}$ denote the rooted tree $(\bigcup_{n \in \mathbb{N}} T_n, B_1)$.

Using that \mathcal{P} -trees are \mathcal{D} -trees with "infinite degrees and leaves", we deduced Algorithm 4 from the weak limit of $(A_i^{\mathcal{D}})_{i \in \mathbb{N}}$ in Algorithm 3. In particular, we introduced the vertices $(V_{\infty,i})_{i \in \mathbb{N}}$ to fill the eventual "gap" produced by the vertices with small degrees. Moreover, we introduced the leaves $(L_i)_{i \in \mathbb{N}}$. We think that they are crucial to fully understand several objects related to \mathcal{P} -trees. However, \mathcal{P} -trees are no longer separable with the leaves, which leads to several topological issues. For this reason we sometimes omit them.

From \mathcal{D} -trees to their scaling limits ICRT

First let us introduce a generic stick breaking construction. It takes for input two sequences in \mathbb{R}^+ called cuts $\mathbf{y} = (y_i)_{i \in \mathbb{N}}$ and glue points $\mathbf{z} = (z_i)_{i \in \mathbb{N}}$, which satisfy

$$\forall i < j, \quad y_i < y_j \quad ; \quad y_i \longrightarrow \infty \quad ; \quad \forall i \in \mathbb{N}, \quad z_i \leq y_i,$$

and creates an \mathbb{R} -tree (loopless geodesic metric space) by recursively "gluing" segment $(y_i, y_{i+1}]$ at z_i , or rigorously, by constructing a consistent sequence of distances $(d_n)_{n \in \mathbb{N}}$ on $([0, y_n])_{n \in \mathbb{N}}$.

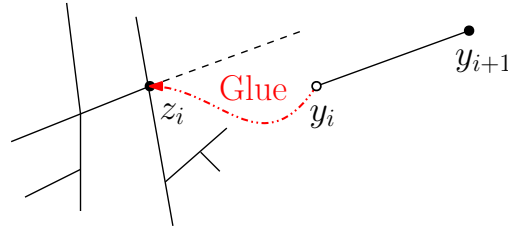


Figure 1.4: A typical step of the stick-breaking construction: the "gluing" of $(y_i, y_{i+1}]$ at z_i .

Algorithm 5: Generic stick-breaking construction of \mathbb{R} -tree (see Figure 1.4).

- Let d_0 be the trivial metric on $[0, 0]$.
- For each $i \geq 0$ define the metric d_{i+1} on $[0, y_{i+1}]$ such that for each $x \leq y$:

$$d_{i+1}(x, y) := \begin{cases} d_i(x, y) & \text{if } x, y \in [0, y_i] \\ d_i(x, z_i) + |y - y_i| & \text{if } x \in [0, y_i], y \in (y_i, y_{i+1}] \\ |x - y| & \text{if } x, y \in (y_i, y_{i+1}], \end{cases}$$

where by convention $y_0 := 0$ and $z_0 := 0$.

- Let d be the unique metric on \mathbb{R}^+ which agrees with d_i on $[0, y_i]$ for each $i \in \mathbb{N}$.
- Let $\text{SB}(\mathbf{y}, \mathbf{z})$ be the completion of (\mathbb{R}^+, d) .

Recall Algorithms 2, 3. To study the geometry of \mathcal{D} -trees (and similarly for \mathcal{P} -trees) we are interested in the cuts, that is the indexes $Y_1^{\mathcal{D}}, Y_2^{\mathcal{D}}, \dots$ such that $A_i^{\mathcal{D}} \notin \{A_j^{\mathcal{D}}\}_{j < i}$, and in the glue points that is the indexes $Z_1^{\mathcal{D}}, Z_2^{\mathcal{D}}, \dots$ such that for $i \geq 1$, $Z_i^{\mathcal{D}} := \inf\{z, A_z = A_{Y_i^{\mathcal{D}}}\}$. To this end, we modify our algorithms to have direct access to those variables.

First, we want to accelerate our constructions. Generally, there are two basic ways to accelerate an algorithm: approximation, and parallel processing. We can do both by adding independency. To this end, we use a classic trick: We generate the uniform permutation of Algorithm 2 by ordering independent exponential random variables. We then rewrite an approximation of Algorithm 2 using those variables. This idea was already used by Aldous, Camarri, Pitman [16, 40] to study \mathcal{P} -trees.

We also introduce another modification which is central in the whole thesis. In algorithm 3 we look at each step at a vertex A_i and act differently depending of whether it is new or not. The main idea is to invert the order: we first see if it is new or not, and then look at its precise value. To this end, we need at each step the probability of taking an old vertex. Using Algorithm 2, this probability is the probability of choosing an half upward edges in the tree we already constructed. For this reason, we keep track of the number of half upward edges in the tree using a measure μ . To do so, we look at the instants $(X_i)_{i \in \mathbb{N}}$ when we see the vertices $(V_i)_{i \in \mathbb{N}}$ for the first time.

This measure also indicates where those upward edges are, which is necessary to know where will be the glue points. We often call μ the "uniform measure", since at each step, conditionally on taking an old vertex, we take an uniform upward edge in the tree already constructed.

Finally, recall that we lose half upward edges when we fuse them with an half downward edge. To add independency and simplify μ , we forget this mass loss. Precisely, we do not remove the upward edges when we get a repetition. As a result, for each new vertex V_i in the tree we get $d_i - 1$ upward edges, and the number of upward edge does not decrease. Then we rectify the bad fusions.

Also to have some "scaling limits" we need to rescale the distances in the tree. Some basic estimations shows that the typical distance in the tree is of order $\sigma^D := \sqrt{\sum_{i=1}^s d_i(d_i - 1)}$.

Combining those ideas we get the next algorithm. Although, it may seem more complex, since we are only interested in the cuts and glue points it repels the non essential informations to the end.

Algorithm 6: Definition of the continuum \mathcal{D} -tree.

- Let $(X_i)_{i \in \mathbb{N}}$ be independent exponential random variables of parameter $(d_i/\sigma)_{i \in \mathbb{N}}$.
- Let μ be the measure on \mathbb{R}^+ defined by $\mu = \sum_{i=1}^s \delta_{X_i} (d_i - 1) / \sigma$.
- Let $(\bar{Y}_i, \bar{Z}_i)_{i \in \mathbb{N}}$ be a Poisson point process on \mathbb{R}^{+2} of intensity $\mathbf{1}_{y \geq z} dy \times d\mu(z)$.
- For every $i \in \mathbb{N}$ with $d_i \geq 1$, let $U_{X,i}$ be uniform in $\{1, \dots, d_i\}$.
- For $j \in \mathbb{N}$, let U_j be uniform in $\{1, \dots, d_j\} \setminus \{U_{X,i}\}$ where i is such that $X_i = \bar{Z}_j$.
- Let $k_1 < \dots < k_N$ be the indexes such that $(Z_k, U_k) \notin \{(Z_j, U_j), j < k\}$.
- For all $1 \leq i \leq N$, let $Y_i = \bar{Y}_{k_i}$ and let $Z_i = \bar{Z}_{k_i}$. For all $i > N$, let $Y_i = Z_i = +\infty$.
- The continuum \mathcal{D} -tree is the random \mathbb{R} -tree $\text{SB}(Y_i, Z_i)$ (see Algorithm 5).

Finally, by making each line of Algorithm 6 converge, and by removing the non essential informations, we recover the ICRT introduced by Aldous, Camarri, and Pitman in [16, 40]. Let Ω_Θ be the space of sequences $(\theta_i)_{i \in \{0\} \cup \mathbb{N}}$ in \mathbb{R}^+ such that $\sum_{i=0}^\infty \theta_i^2 = 1$ and such that $\theta_1 \geq \theta_2 \geq \dots$.

Algorithm 7: New construction of the Θ -ICRT for $\Theta \in \Omega_\Theta$.

- Let $(X_i)_{i \in \mathbb{N}}$ be independent exponential random variables of parameter $(\theta_i)_{i \in \mathbb{N}}$.
- Let μ be the measure on \mathbb{R}^+ defined by $\mu = \theta_0^2 dx + \sum_{i=1}^\infty \delta_{X_i} \theta_i$.
- Let $(Y_i, Z_i)_{i \in \mathbb{N}}$ be a Poisson point process on \mathbb{R}^{+2} of intensity $\mathbf{1}_{y \geq z} dy \times d\mu(z)$.
- The Θ -ICRT is defined as $(\mathcal{T}, d) = \text{SB}((Y_i)_{i \in \mathbb{N}}, (Z_i)_{i \in \mathbb{N}})$ (see Algorithm 5).

In addition, in Chapter 5, to define plane ICRT, we keep the lines 4 and 5 of algorithm 6.

The original definition of Aldous, Camarri, Pitman [16, 40] of the ICRT is slightly different. Indeed, at each step we first look at if the vertices are new or not, and then look at their values. Since they chose otherwise they found the following algorithm, which is harder to use:

Algorithm 1.3. Classical construction of the Θ -ICRT from [16, 40].

- Let $(A_i, B_i)_{i \in \mathbb{N}}$ be a Poisson point process of intensity θ_0^2 on $\{(a, b) \in \mathbb{R}^{+2} : b \leq a\}$.
- Let $((A_{i,j})_{j \in \{0\} \cup \mathbb{N}})_{i \in \mathbb{N}}$ be a family of independent Poisson point processes of intensity $(\theta_i)_{i \in \mathbb{N}}$ on \mathbb{R}^+ and independent of $(A_i, B_i)_{i \in \mathbb{N}}$.
- Sort the elements of the (almost surely) locally finite set $\bigcup_{i=1}^\infty \{A_i\} \cup \bigcup_{i=1}^\infty \bigcup_{j=1}^\infty \{A_{i,j}\}$ as $U = (U_i)_{i \geq 1}$ with $U_1 < U_2 < \dots$.
- For $i \geq 1$, let $V_i = \begin{cases} B_j & \text{if } U_i \text{ is of the form } A_j \\ A_{i,0} & \text{otherwise} \end{cases}$ and let $V = (V_i)_{i \geq 1}$.
- The (old) Θ -ICRT is defined as $(\mathcal{T}^*, d^*) = \text{SB}(U, V)$.

We prove in Chapter 2, using a simple computation, that the two constructions are equivalent.

Proposition 3

Let (\mathcal{T}, d) be the tree constructed by Algorithm 7. Let (\mathcal{T}^*, d^*) be the tree constructed by Algorithm 1.3. Then (\mathcal{T}, d) and (\mathcal{T}^*, d^*) have the same distribution.

1.1.4 The chaining method

The chaining method is one of the most important basic technique in concentration theory. It is directly related to trees, and has found many applications to study random metric spaces and stick-breaking constructions (see e.g. Aldous [10], Amini, Devroye, Griffiths, Olver [17], Curien, Haas [43], Sénizergues [77]). In this thesis, we use this method in most of our main results' proof.

On the other hand, by understanding the metric structure of trees spanning a metric space, one can show precise bounds for suprema of empirical process¹. Following this principle, we proved general concentration inequalities that have their own interest.

Main principle

The goal of this method is to estimate the max of a given function f on a space S . To this end, consider a sequence of increasing subspaces² $(S_i)_{i \geq 0}$ of S "approximating" S , and for every $i \geq 0$, a projection $p_i : S_{i+1} \rightarrow S_i$. The main idea is that if $(S_i)_{i \in \mathbb{N}}$ are properly chosen then

$$\max_{x \in S} f(x) \leq \sum_{i=0}^{\infty} \max_{x \in S_{i+1}} (f(x) - f(p_i(x))). \quad (1.3)$$

As a result, one can decompose a complex estimate into many simpler ones.

Interpretation with trees

First let us introduce some definitions. Consider a rooted tree $T = (V, E, \rho)$, and consider for each edge $e \in E$ a length $l(e) \in \mathbb{R}^+$. The length of a path v_1, v_2, \dots, v_k is $\sum_{i=1}^{k-1} l(v_i, v_{i+1})$. As before, the graph distances between two vertices is the minimal length of a path between them. And we write $H_l(T)$ as the height of tree T . Also, given for each edge $e \in E$ a cost $c(e)$, we define the total cost of the tree as being $c(T) := \sum_{e \in E} c(e)$.

Now, for simplicity, we assume that $S_0 = \{\rho\}$ with $f(\rho) = 0$. We consider the edges $E = \bigcup_{i=1}^{\infty} \{x \in S_{i+1}, \{x, p_i(x)\}\}$, and the rooted tree, $T := (S, E, \rho)$. And for every $i \in \mathbb{N}$, $x \in S_{i+1}$ we give to the edge $(x, p_i(x))$ the length $|f(x) - f(p_i(x))|$. So (1.3) can be rewritten as

$$\max_{x \in S} f(x) \leq H_l(T).$$

We now add probabilities. Consider some arbitrary length $l : E \mapsto \mathbb{R}^+$. We associate to l some costs $c_l : \{x, y\} \in E \mapsto \mathbb{P}(|f(x) - f(y)| > l(e))$. By (1.3) and an union bound,

$$\mathbb{P} \left(\max_{x \in S} f(x) > H_l(T) \right) \leq c_L(T).$$

Hence, to optimize the chaining method we minimize the height of a tree given a maximum cost. To this end, we need some a priori on the "geometry" of (S, f) to properly choose $(S_i)_{i \in \mathbb{N}}$ and l . However, there is no proper way to study the "geometry" of (S, f) since (S, f) is not a metric space. For this reason, outside some precise setting, optimizing the chaining method require some a priori informal optimization and geometric guesses.

¹Talagrand [79] p.148 "Many of the results of this book were first discovered by following the idea that T is "large" if and only if it contains a large tree".

²For more complex algorithms one may want to consider general $(X_i)_{i \in \mathbb{N}}$ and a family of functions f_i on X_i . (See e.g. Sénizergues thesis [78] for several such applications of this method).

Basic examples, and contributions to empirical process

First consider f a Brownian motion on $[0, 1]$. $[0, 1]$ can be approximated via dyadic numbers. For every $i \geq 0$ let $S_i := \{k/2^i, 0 \leq k < 2^i\}$. And for $i \geq 0$ let $p_i : x \in S_{i+1} \mapsto \max\{y \in S_i, y \leq x\}$. We thus approximate $[0, 1]$ with a dyadic tree (see Figure 1.5). And since for every $x \leq y \in [0, 1]$, $f(x) - f(y)$ have order $\sqrt{y - x}$ we give a length of order $\sqrt{y - x}$ to an edge (x, y) in the tree. By being more precise we can obtain sharp upper bound on the maximum of a Brownian motion. This simple example, called the classic or generic chaining, is extended in many settings where we have some "dimension-type conditions" on S , and some "Hölder-type conditions" on f .

This kind of chaining immediately fails when f has a discontinuity. For instance consider $f : x \in [0, 1] \mapsto \sum_{i=1}^n \mathbf{1}_{x \leq U_i} - nx$, where $(U_i)_{1 \leq i \leq n}$ are independent and uniform on $[0, 1]$. In this case, we technically count the jumps an infinite number of times in the tree. To avoid this issue, one usually approximate $[0, 1]$ as $\{i/2^k\}_{0 \leq i < 2^k}$ where 2^k have approximatively order n . Then, in each of the intervals of the form $[i/2^k, (i+1)/2^k]$, f slightly vary since there are very few jumps.

This method however fails when the sizes of the jumps are inhomogeneous. To deal with this case we need to understand better the "geometry" of $([0, 1], f)$. The main idea is that the instants of the jumps are important to understand its geometry of the graph of f . Hence, to approximate better the "geometry" of $([0, 1], f)$ we need to use those jumps. By adding those jumps to the previous dyadic decomposition we prove the following results:

Lemma 4

Let $(\lambda_i)_{i \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. Let $(X_i)_{i \in \mathbb{N}}$ be independent real random variables with density $(f_i)_{i \in \mathbb{N}}$. Let $a \leq b$. Let $S : x \mapsto \sum_{i=1}^{\infty} \lambda_i (\mathbf{1}_{X_i \leq x} - \mathbb{P}(X_i \leq x))$. Assume that the previous series converges uniformly on $[a, b]$.

(a) Assume that $(\lambda_i)_{i \in \mathbb{N}}$ is decreasing. Let,

$$C := \sum_{q=1}^{\infty} |\lambda_{2^q}| \quad ; \quad D := 2 \log(2) \sum_{q=1}^{\infty} q |\lambda_{2^q}|$$

$$W := \sum_{i=1}^{\infty} \lambda_i^2 (b-a) \max_{a \leq x \leq b} f_i(x) \quad ; \quad E := \sum_{i=1}^{\infty} \frac{|\lambda_i|}{i} (b-a) \max_{a \leq x \leq b} f_i(x).$$

For every $t > 0$,

$$\mathbb{P} \left(\sup_{a \leq x \leq b} |S(x) - S(a)| > 6\sqrt{Wt} + (t+1)C + D + 2E \right) \leq e^{3-t}.$$

(b) Let

$$V := \sum_{i=1}^{\infty} \lambda_i^2 \quad ; \quad M := (b-a) \max_{i \in \mathbb{N}} \max_{a \leq x \leq b} f_i(x).$$

For every $t > 0$,

$$\mathbb{P} \left(\sup_{a \leq x \leq b} |S(x) - S(a)| > 7\sqrt{Vt} + (t+1)M \right) \leq e^{3-t}.$$

Although those results are already precise enough to be used in a large variety of settings, we think that there are many rooms for improvement and notably to generalize them. Also, the bounds are bad to deal with the i such that $|\lambda_i|$ is large and $\mathbb{P}(X_i \in [a, b])$ is small. To avoid this issue, we advise to condition by $(\mathbf{1}_{X_i \in [a, b]})_{i \in \mathbb{N}}$, and replace, up to reordering, $(\lambda_i)_{i \in \mathbb{N}}$ by $(\lambda_i \mathbf{1}_{X_i \in [a, b]})_{i \in \mathbb{N}}$.

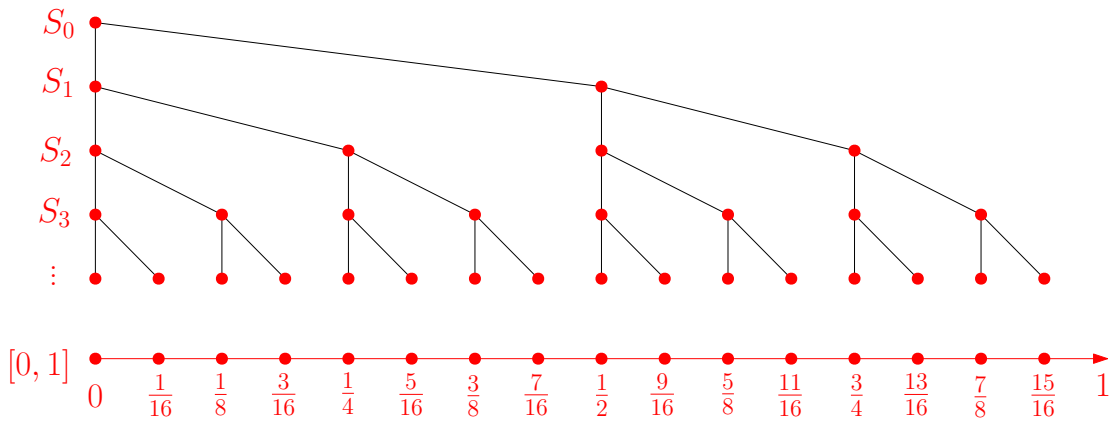


Figure 1.5: Approximating $[0, 1]$ with a dyadic tree. Points in a same column have the same value.

We then derive from Proposition 4, and a slicing argument (see [32] Section 13.7) the next bound, which will be used in [26] to consider non centered process on trees. It can also be used to study several other process related to trees (the mass $x \mapsto \mu[0, x]$, process with exchangeable increments, the exploration process of the configuration model/multiplicative graphs...)

Lemma 5

We keep the notations of Proposition 4. For every $k \in \mathbb{N}$, let

$$S_k : x \mapsto \sum_{i=1}^k \lambda_i (\mathbf{1}_{X_i \leq x} - \mathbb{P}(X_i \leq x)).$$

For every $t > 0$,

$$\mathbb{P} \left(\sup_{k \in \mathbb{N}} \sup_{a \leq x \leq b} |S_k(x) - S_k(a)| > 42\sqrt{V}(t + M) \right) \leq e^{6-t}.$$

Chaining method to study random geometry

Our algorithms can be resumed as follows: we construct a tree T (with eventually some additional structures) by recursively adding some edges and vertices to some subtrees. We want to estimate the distance between those subtrees and the whole tree. To this end, in most of our main results' proof, we use the chaining method. Notably, we use it to study compactness, compute fractal dimensions, prove leaf tightness, estimate height...

Although we detail those applications in the next section, let us explain how we use the chaining method in general. The main idea is to split the construction of the trees in different main steps, and to prove that between two main steps the trees $(T_i)_{i \in \mathbb{N}}$ that we construct does not change much. For simplicity let us consider a reference example. We want to estimate the Hausdorff distance d_H between our trees. Let d be the graph distance on T . For $x \in T$, $j \in \mathbb{N}$, let $p_j(x)$ be the projection of x on T_j , that is the vertex y on T_j which minimizes $d(x, y)$. By definition of the Hausdorff distance, for $i \in \mathbb{N}$, $d_H(T_i, T) = \max_{x \in T} d(x, p_i(x))$. And the chaining argument (1.3) with the function $x \in T \mapsto d(x, p_i(x))$ can be rewritten into,

$$d_H(T_i, T) \leq \sum_{j \geq i} \max_{x \in T_{j+1}} d(x, p_j(x)). \quad (1.4)$$

Here using the chaining method might seem problematic. Indeed, to optimize this method we need to understand the "geometry" of (T, d) . But, we use the chaining method to study (T, d) . Actually, the main strength of this method is that it only needs a few geometric results/guesses to prove precise bounds. In our case, we already have the necessary information: The typical length between any new leaf and the previous subtree is given by the length of a single branch.

Moreover in (1.4) the maximum is reach on the leaves of T_{j+1} . So we only need to estimate and optimize the number of leaves in $(T_i)_{i \in \mathbb{N}}$ and the distance between a leaf and the previous subtree. Alternatively, for technical reasons, we approximate the maximum by not considering leaves but instead sufficiently many vertices on $(T_j)_{j \in \mathbb{N}}$.

1.1.5 Some new martingale inequalities

We regroup here some martingale concentration inequalities proved in this thesis. Their proofs are both disconnected from the rest of the thesis, and classic (moments computation+Doob inequality). Nonetheless, those results are strong and precise enough to have their own interest. The first result concerns the speed of convergence of general Pólya urns.

Lemma 6

Let $\{m_n\}_{n \geq 0}$ be a positive real-valued sequence. Let $(A_n)_{n \geq 0}$ be a sequence of positive real-valued random variables such that $A_0 \leq m_0$ and such that for every $n \geq 0$,

$$\mathbb{P}(A_{n+1} = A_n + m_{n+1} | A_n) = \frac{A_n}{M_n} \quad ; \quad \mathbb{P}(A_{n+1} = A_n | A_n) = \frac{M_n - A_n}{M_n},$$

where for every $n \geq 0$, $M_n = \sum_{i=0}^n m_i$.

a) If $\sum_{n=0}^{\infty} m_n^2 / M_n^2 < \infty$, then almost surely for every $a \geq 0$ and $t \in \mathbb{R}^+$,

$$\begin{aligned} \mathbb{P}\left(\sup_{i \geq a} \left| \frac{A_i}{M_i} - \frac{A_a}{M_a} \right| > t \frac{A_a}{M_a} \middle| A_a\right) \\ \leq 2 \exp\left(-\frac{(t^2/4)(A_a/M_a)}{\sum_{n>a} \frac{m_n^2}{M_n^2} + t \max\left(\sum_{n>a} \frac{m_n^2}{M_n^2}, \max_{n>a} \frac{m_n}{M_n}\right)}\right). \end{aligned}$$

b) If $\{m_n\}_{n \in \mathbb{N}}$ is bounded, then almost surely for every $a \geq 0$ and $t \in \mathbb{R}^+$,

$$\mathbb{P}\left(\sup_{i \geq a} \left| \frac{A_i}{M_i} - \frac{A_a}{M_a} \right| > t \frac{A_a}{M_a} \middle| A_a\right) \leq 2 \exp\left(-\frac{t^2}{4(1+t)} \frac{A_a}{\max_{n>a} m_n}\right).$$

The second result concerns empirical process, and is derived from Marcinkiewick's inequality.

Lemma 7

Let $\kappa \geq 2$, and let $(x_i)_{i \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. Let $(X_i)_{i \in \mathbb{N}}$ be independent centered, random variables. For every $k \leq n$ let $S_k := x \in \mathbb{R} \mapsto \sum_{i \leq k} \mathbf{1}_{x \leq x_i} X_i$. Assume that for every $i \in \mathbb{N}$, $v_i := \mathbb{E}[X_i^\kappa]^{2/\kappa} < \infty$, and let $V := \sum_{i \in \mathbb{N}} v_i$. Then for every $t > 0$,

$$\mathbb{P}\left(\sup_{i \in \mathbb{N}} \|S_i\|_{\infty} > t\right) \leq C_{\kappa} \left(\sqrt{V}/t\right)^{\kappa},$$

where C_{κ} is a constant which depends only on κ .

1.2 Overview of the chapters

In Section 1.1 we presented this thesis as a whole, our models, algorithms, and their connections. We now focus on each chapter appart, and particularly on our main results and proofs. Since we now follow the publication order, the logical order of this section and Section 1.1 differ.

1.2.1 Compactness and fractal dimensions of ICRT

We discuss here Chapter 2, whose aim is twofold. We introduce several tools to study ICRT, notably a stick-breaking construction (recall Algorithm 7), and a probability measure on ICRT. We also prove refined results on ICRT, in particular concerning compactness and fractal dimensions.

We first prove with basic concentration inequalities several bounds on μ and the cuts $(Y_i)_{i \in \mathbb{N}}$. To simplify the notations, for $n \in \mathbb{N}$ let $l_n := Y_n - Y_{n-1}$ denote the length of the n th branch, and let $m_n := \mu(Y_{n-1}, Y_n]$ denote its weight. Then let $M_n := \mu[0, Y_n] = m_1 + \dots + m_n$.

Lemma 8

The following assertion hold a.s. (almost surely):

- (a) The map $l \mapsto \mathbb{E}[\mu[0, l]]$ is concave.
- (b) As $l \rightarrow \infty$, $\mu[0, l] \sim \mathbb{E}[\mu[0, l]] = \theta_0^2 l + o(l)$.
- (c) For every l large enough, there are at most $2l\mu[0, l] \leq 2l^2$ cuts on $[0, l]$.
- (d) For every i large enough, $l_{i+1} \leq 5 \log(Y_i)/M_i$.
- (e) For every i large enough, $m_{i+1} \leq \log(Y_i)^2/Y_i$.

The uniform probability measure on ICRT

As explained in Section 1.1, we prove that the first branches approximate well the whole tree \mathcal{T} . We consider two main metrics: GH and GP. We deal here with GP approximation, by introducing the "uniform probability measure" on ICRT.

To this end, we need the ICRT to have infinite mass, that is $\theta_0 > 0$ or $\sum_{i=1}^{\infty} \theta_i = \infty$. In this case, we write $\mu = \infty$ since a.s. $\mu(\mathbb{R}^+) = \infty$. Otherwise the ICRT has finite mass and $\mu < \infty$. For technical reasons, there is no proper way to define the uniform probability measure on ICRT with finite mass. So we assume throughout this section that $\mu = \infty$.

Let for every $l \in \mathbb{R}^+$, $\mathcal{T}_l = ([0, l], d)$, let μ_l be the restriction of μ on \mathcal{T}_l , and let $p_l := \mu_l / \mu[0, l]$.

Theorem 1

Almost surely $(p_l)_{l \in \mathbb{R}^+}$ converges weakly toward a probability measure p on \mathcal{T} . Furthermore, p has support \mathcal{T} , has no atoms and gives measure 1 to the set of leaves, (that is the set of $x \in \mathcal{T}$ such that $\mathcal{T} \setminus \{x\}$ is connected).

The next equivalent definitions may be interpreted with the discrete setting as the uniform probability measure on the edges/leaves. So we call p the "uniform probability measure" on \mathcal{T} .

Proposition 9

Let μ^{\sim} be the Lebesgue measure on \mathbb{R}^+ and $\mu^{\bullet} = \sum_{i=1}^{\infty} \delta_{Y_i}$. For every $l \in \mathbb{R}^+$ let μ_l^{\sim} (resp. μ_l^{\bullet}) be the restriction of μ^{\sim} (resp. μ^{\bullet}) to $\mathcal{T}_l = ([0, l], d)$. Also let for every $l \in \mathbb{R}^+$, $p_l^{\sim} = \mu_l^{\sim} / \mu_l^{\sim}[0, l]$ and $p_l^{\bullet} = \mu_l^{\bullet} / \mu_l^{\sim}[0, l]$. Then

$$p_l^{\sim} \xrightarrow[l \rightarrow \infty]{\text{weakly}} p \quad \text{and} \quad p_l^{\bullet} \xrightarrow[l \rightarrow \infty]{\text{weakly}} p.$$

To prove Theorem 1, and Proposition 9 we estimate precisely the evolution of the mass μ as we add branches to the tree. The main idea is that the mass evolves as a Pólya Urn. More precisely, for all $x \in \mathcal{T}$, $l \in \mathbb{R}^+$, let the projection of x in \mathcal{T}_l be the nearest point from x in \mathcal{T}_l . Also for every $l \in \mathbb{R}^+$, $S \subset \mathcal{T}$, let $S^{\uparrow l}$ be the set of $x \in \mathcal{T}$ such that the projection of x in \mathcal{T}_l is in S . For every $i \in \mathbb{N}$, $S \subset \mathcal{T}_{Y_i}$, conditionally on $\mu, \{Y_i\}_{i \in \mathbb{N}}, \{\mu_{Y_j}(S^{\uparrow})\}_{j \geq i}$ is a generalized Pólya urn. So, with Lemmas 6 and 8 (e), we can estimate the evolution of the mass glued on a set.

Then to estimate globally μ , we cover $(\mathcal{T}_l)_{l>0}$ with many small sets and estimate the mass glued on them. This idea is not new and goes back to the very study of the CRT by Aldous [10]. But in our case, many topological argument are much more complex for non compact ICRT.

Comparisons with Lévy trees

Recall that Lévy trees are the limits of Galton–Watson trees which are mix of \mathcal{D} -trees. And recall that we will prove in Chapter 3 that \mathcal{D} -trees converge toward ICRT. So Lévy trees are ICRT with random parameter. We now precise this connection to compare our results with some previous ones, and notably those of Le Gall and Le Jan [64, 63], and of Le Gall and Duquesne [50, 49].

For short, Lévy trees are constructed from real random processes modified from Lévy processes. Morally, the large degrees are given by the jumps of the Lévy process, and for ICRT to $(\theta_i)_{i \in \mathbb{N}}$. The small degrees contribute proportionally to the brownian part of the process, and for ICRT to θ_0 . For both models, the geometry is dictated by the tail of the degrees. Equivalently, for Lévy trees, since this tail is morally concentrated, it is dictated by the Laplace exponent ψ of the Lévy process. An analog of ψ for ICRT is the function $l \mapsto l\mathbb{E}[\mu[0, l]]$.

Compactness of ICRT

The next result was conjectured by Aldous, Miermont, Pitman [14] by comparison with Lévy trees .

Theorem 2

The ICRT is almost surely compact if and only if

$$\int_0^\infty \frac{dl}{l\mathbb{E}[\mu[0, l]]} < \infty.$$

To prove Theorem 2 we estimate the Hausdorff distance for $l \in \mathbb{R}^+$ between \mathcal{T}_l and \mathcal{T} . More precisely, \mathcal{T} is compact if and only if $(\mathcal{T}_l)_{l \in \mathbb{R}^+}$ is a Cauchy sequence for the Hausdorff distance. We prove it with the chaining method (recall Section 1.1.4).

Our starting point is to estimate the distance for every $y, z \in \mathbb{R}^+$ between z and \mathcal{T}_y . To this end, we adapt an approach of Curien and Haas [43] (see also Sénizergues [77] for an extensive use): We follow the path from z to \mathcal{T}_y . In [43, 77], they use deterministic branches, and show that each branch $[Y_i, Y_{i+1})$ intersect the path independently of each other with probability m_i/M_i . When it does the length of the intersection is at most l_i . As a result, conditionally on the branches, the distance between z and \mathcal{T}_y is bounded by a sum of independent random variables. Here we shortcut this argument, by sampling the cuts and glue points as we go down in the tree (or equivalently \mathbb{R}^+). By estimating the length of this path, we obtain the next exponential bound:

Lemma 10

For every $y, z \in \mathbb{R}^+$, $t \geq 6$, conditionally on μ , $\mathbb{P}(d(\mathcal{T}_y, z) > 4t/\mu[0, x]) \leq e^{-t}$.

We then apply the chaining method. Here, the generic chaining seems optimal. More precisely, for every $l \in \mathbb{R}^+$, by Lemma 8 (a), there exists a unique real $\mathcal{X}_l \in \mathbb{R}^+$ such that $\mathbb{E}[\mu[0, \mathcal{X}_l]] = l$.

With a topological argument, and Lemma 10 we deduce that almost surely for every k large enough,

$$d_H(\mathcal{T}_{\mathcal{X}_{2^{k-1}}}, \mathcal{T}_{\mathcal{X}_{2^k}}) \leq 21 \log \mathcal{X}_{2^k} / 2^k. \quad (1.5)$$

Hence, the ICRT is a.s. compact when $\sum \log \mathcal{X}_{2^k} / 2^k < \infty$, which appear to be equivalent to the condition of Theorem 2.

Reciprocally, it is easy to prove a reverse bound for (1.5). However, this does not yield a reverse bound for $d_H(\mathcal{T}_{\mathcal{X}_{2^k}}, \mathcal{T})$. To do so, we adapt an argument of Amini, Devroye, Griffiths, Olver [17]. The main idea is to prove that a.s. for every k large enough there are many segments of length of order $\mathcal{X}_{2^k} / 2^k$ between $\mathcal{X}_{2^{k-1}}$ and \mathcal{X}_{2^k} , and that moreover those segments form long chains by being glued on one another. In conclusion we show that morally a.s. for every k large enough,

$$d_H(\mathcal{T}_{\mathcal{X}_{2^k}}, \mathcal{T}) \approx \sum_{n=k}^{\infty} \frac{\log \mathcal{X}_{2^n}}{\mathcal{X}_{2^n}} \approx \int_{\mathcal{X}_{2^k}}^{\infty} \frac{dl}{l \mu[0, l]}. \quad (1.6)$$

Fractal dimensions

Theorem 3

Recall the definitions of the fractal dimensions in Section 1.1.2. Almost surely

$$\dim_P(\mathcal{T}) = \overline{\dim}(\mathcal{T}) = 1 + \limsup_{l \rightarrow \infty} \frac{\log l}{\log \mathbb{E}[\mu[0, l]]}.$$

Furthermore if $\log l = \mathbb{E}[\mu[0, l]]^{o(1)}$ then

$$\dim_H(\mathcal{T}) = \underline{\dim}(\mathcal{T}) = 1 + \liminf_{l \rightarrow \infty} \frac{\log l}{\log \mathbb{E}[\mu[0, l]]}.$$

To prove Theorem 3, by Lemma 1.1.1, it is enough to upperbound the Minkowski dimensions and to lowerbound the packing and Hausdorff dimensions. To upperbound the Minkowski dimensions we first cover for $l \in \mathbb{R}^+$, $\mathcal{T}_{\mathcal{X}_l}$ with some balls then increase their radius to cover \mathcal{T} . The computations are straightforward. First $\mathcal{T}_{\mathcal{X}_l}$ have total length \mathcal{X}_l and so can be covered with $l\mathcal{X}_l$ balls of radius $3/l$. Then by (1.5) and the condition of Theorem 3, $d(\mathcal{T}_{\mathcal{X}_l}, \mathcal{T}) = O(1/l^{1+o(1)})$. (For the upper Minkowski dimension the upperbound is otherwise trivial.)

To lowerbound the packing and Hausdorff dimensions, the main idea is to estimate p of the open balls around a uniform random vertex on \mathcal{T} . To do so we proceed in two steps. First we lower bound the mass glued on a branch. This is done as explained previously with a Polya urn argument. Then, we lower bound for $l \in \mathbb{R}^+$ the distance between a uniform vertex and \mathcal{T}_l .

Using discrete trees as an alternative.

Many arguments of Chapter 2 can be shortcut using that \mathcal{P} -trees converge toward ICRT. Indeed this convergence yield many equalities in law. For instance, for every $i < j \in \mathbb{N}$, conditionally on \mathcal{T}_{Y_i} , $d(\mathcal{T}_{Y_i}, Y_j) \stackrel{(d)}{=} d(\mathcal{T}_{Y_i}, Y_{i+1})$ which completely replace Lemma 10. Another example, concerns the fractal dimensions. Since the root, Y_1, Y_2, \dots are uniform vertex we can estimate the mass around a uniform vertex by looking at the neighborhood of 0, or even at the probability that Y_1 (and $Y_2, Y_3 \dots$ when needed) are close from 0. Also, we repeat many computations of Chapter 2 in Chapter 3. We could have, as Aldous [10], study at the same time our discrete models and their limits.

In this thesis, we decided to avoid those shortcuts as most as possible for several reasons. First, they complicate our setting. Then, they are less generalizable to small perturbations of our algorithms. Also, they actually require many technical arguments to be fully rigorous.

1.2.2 Limit of trees with fixed degree sequence

Regimes of convergence

Recall Section 1.1.3. We discuss here Chapter 3 where we study the limits of our trees. The methods are the same as in the last section. But some proofs are more technical since we need to be more precise and deal with the mass loss of \mathcal{D} -trees. On the other hand, we shortcut some proofs.

First let us precise the different regimes we work on. Recall $\Omega_{\mathcal{D}}, \Omega_{\mathcal{P}}, \Omega_{\Theta}$ from Section 1.1.3. Let $(\mathcal{D}_n)_{n \in \mathbb{N}}, (\mathcal{P}_n)_{n \in \mathbb{N}}, (\Theta_n)_{n \in \mathbb{N}}$ denote some sequences in $\Omega_{\mathcal{D}}, \Omega_{\mathcal{P}}, \Omega_{\Theta}$ respectively.

Assumption 1.1 ($\mathcal{D}_n \Rightarrow \mathcal{P}$). For all $i \geq 1$, $d_i^{\mathcal{D}_n}/s^{\mathcal{D}_n} \rightarrow p_i^{\mathcal{P}}$ and $s^{\mathcal{D}_n} \rightarrow \infty$.

Assumption 1.2 ($\mathcal{D}_n \Rightarrow \Theta$). For all $i \geq 1$, $d_i^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n} \rightarrow \theta_i^{\Theta}$ and $d_1^{\mathcal{D}_n}/s^{\mathcal{D}_n} \rightarrow 0$.

Assumption 1.3 ($\mathcal{P}_n \Rightarrow \Theta$). For all $i \geq 1$, $p_i^{\mathcal{P}_n}/\sigma^{\mathcal{P}_n} \rightarrow \theta_i^{\Theta}$ and $p_1^{\mathcal{P}_n} \rightarrow 0$.

Assumption 1.4 ($\Theta_n \Rightarrow \Theta$). For all $i \geq 1$, $\theta_i^{\Theta_n} \rightarrow \theta_i^{\Theta}$.

In addition, one can put a topology on $\Omega := \Omega_{\mathcal{D}} \cup \Omega_{\mathcal{P}} \cup \Omega_{\Theta}$ such that \Rightarrow coincide with the notion of convergence on Ω . The subset $\Omega_{\mathcal{D}}$ is dense, so our results on \mathcal{D} -trees imply the others. Moreover, (Ω, \Rightarrow) is Polish, so our results can be used to study trees with random degree sequence.

Convergence of the first branches

Recall that to prove that our trees converge, we first prove that the first branches converge, and finish by proving the leaf tightness. To prove the convergence of the first branches, we transform our algorithms and prove that each lines of the algorithms converge as explained in Section 1.1.3.

Our first main result in Chapter 3 says that under the above regimes the first branches converge. I wrote it not only for Chapter 3, but also to study several other objects in the forthcoming [26, 25]. We detail those objects (looptree, contour process, exploration process, snake...) in Section 1.2.4. Since this result is technical to formally state, let us only explain what we mean by "convergence of the first branches".

First, to prove GH convergence, we prove that the cuts and glue points converge. Also, to prove the GP convergence, we may look at the measure μ which dictates where the mass is glued. Actually, by Lemma 1, to study the GP convergence, we only need to prove that the distances between random vertices converge. And this can be done by using only the cuts and glue points. On the other hand, we still prove that μ converge to study several other objects. Also, in some of those objects the vertices are expanded proportionally to their degrees. So we prove the convergence of the indicator functions that a branch is glued on a fixed vertex of large degree.

Lastly, the convergence of the first branches presented in Chapter 3 is slightly incomplete. Indeed, this chapter omit the plane order of the trees. We will add this order in [26, 25] by sampling it with uniform random variables.

Gromov-Prokhorov convergence

Beforehand let us introduce some notations. Let $\Omega_{\mathbf{p}}$ be the set of measures on $\{V_i\}_{i \in \mathbb{N}}$. We say that a sequence $(\mathbf{p}_n)_{n \in \mathbb{N}} \in \Omega_{\mathbf{p}}^{\mathbb{N}}$ converges toward $\mathbf{p} \in \Omega_{\mathbf{p}}$ when $\max_{i \in \mathbb{N}} |\mathbf{p}_n(V_i) - \mathbf{p}(V_i)| \rightarrow 0$. For every $\mathcal{D} \in \Omega_{\mathcal{D}}$ let $\mathbf{p}^{\mathcal{D}}$ denote a probability measure with support on $\mathcal{V}^{\mathcal{D}} := \{V_i, 1 \leq i \leq s^{\mathcal{D}}\}$. Similarly, for all $\mathcal{P} \in \Omega_{\mathcal{P}}$ let $\mathbf{p}^{\mathcal{P}}$ denote a probability measure with support on $\mathcal{V}^{\mathcal{P}} := \{V_i^{\mathcal{P}}\}_{i: p_i > 0}$. Also, we sometimes let 0 denote the null measure on $\{V_i\}_{i \in \mathbb{N}}$. Recall that " $\mu^{\Theta} = \infty$ " means $\theta_0^{\Theta} > 0$ or $\sum_{i=1}^{\infty} \theta_i^{\Theta} = \infty$. Also recall the definition of \mathbf{p}^{Θ} from Theorem 1.

Next for every $\mathcal{D} \in \Omega$ let $s^{\mathcal{D}} := \max\{i, d_i^{\mathcal{D}} \neq 0\}$, and recall that $(\sigma^{\mathcal{D}})^2 = \sum_{i=1}^{\infty} d_i(d_i - 1)$. Also for every $\mathcal{P} \in \Omega_{\mathcal{P}}$, let $(\sigma^{\mathcal{P}})^2 := \sum_{i=1}^{\infty} (p_i^{\mathcal{P}})^2$. Finally, let $d^{\mathcal{D}}$ denote the graph distance on $T^{\mathcal{D}}$, and let $d^{\mathcal{P}}$ denote the graph distance on $T^{\mathcal{P}}$.

Theorem 4

The following convergences hold weakly for the GP topology (see Section 1.1.2).

(a) If $\mathcal{D}_n \Rightarrow \mathcal{P}$ and $\mathfrak{p}^{\mathcal{D}_n} \rightarrow \mathfrak{p}^{\mathcal{P}}$ then

$$(\mathcal{V}^{\mathcal{D}_n}, d^{\mathcal{D}_n}, \mathfrak{p}^{\mathcal{D}_n}) \xrightarrow{\text{WGP}} (\mathcal{V}^{\mathcal{P}}, d^{\mathcal{P}}, \mathfrak{p}^{\mathcal{P}}).$$

(b) If $\mathcal{D}_n \Rightarrow \Theta$, $\mathfrak{p}^{\mathcal{D}_n} \rightarrow 0$, and $\mu^{\Theta} = \infty$ then

$$(\mathcal{V}^{\mathcal{D}_n}, (\sigma^{\mathcal{D}_n}/s^{\mathcal{D}_n})d^{\mathcal{D}_n}, \mathfrak{p}^{\mathcal{D}_n}) \xrightarrow{\text{WGP}} (\mathcal{T}^{\Theta}, d^{\Theta}, \mathfrak{p}^{\Theta}).$$

(c) If $\mathcal{P}_n \Rightarrow \Theta$, $\mathfrak{p}^{\mathcal{P}_n} \rightarrow 0$, and $\mu^{\Theta} = \infty$ then

$$(\mathcal{V}^{\mathcal{P}_n}, \sigma^{\mathcal{P}_n}d^{\mathcal{P}_n}, \mathfrak{p}^{\mathcal{P}_n}) \xrightarrow{\text{WGP}} (\mathcal{T}^{\Theta}, d^{\Theta}, \mathfrak{p}^{\Theta}).$$

(d) If $\Theta_n \Rightarrow \Theta$, and $\mu^{\Theta} = \infty$ then

$$(\mathcal{T}^{\Theta_n}, d^{\Theta_n}, \mathfrak{p}^{\Theta_n}) \xrightarrow{\text{WGP}} (\mathcal{T}^{\Theta}, d^{\Theta}, \mathfrak{p}^{\Theta}).$$

When $\mu^{\Theta} < \infty$, the distances between random vertices still converge, but p^{Θ} cannot be defined. (c) and (d) was already mostly proved by Aldous, Camarri, and Pitman [16, 40]. Our only contribution to those results is the small extension of the \mathcal{P} -tree model. (a) directly follows from the convergence of the first branches. For (b), although the convergence of the first branches imply that the distances between random leaves converge, we need to consider arbitrary random vertices that are not necessary leaves.

To this end, we modify Algorithm 3 to also construct the subtrees spanned by specific vertices. With this algorithm we prove that the subtree spanned by fixed vertices of small degree is close in distribution to the subtree spanned by the first leaves. Since random vertices tend to have small degree this shows the GP convergence. The algorithm goes as follows:

Algorithm 8: General stick-breaking construction of a \mathcal{D} -tree from a permutation $(W_i)_{1 \leq i \leq s^{\mathcal{D}}}$ of $\mathcal{V}^{\mathcal{D}}$.

- Let $A^{\mathcal{D}} = (A_1, \dots, A_{s-1})$ be a uniform \mathcal{D} -tuple.
- Let $T_1^{\mathcal{D}, W} := (\{A_1\}, \emptyset)$ then for every $2 \leq i \leq s$ let

$$T_i^{\mathcal{D}, W} := \begin{cases} T_{i-1} \cup \{A_{i-1}, A_i\} & \text{if } A_i \notin T_{i-1}, \\ T_{i-1} \cup \{A_{i-1}, W_{\inf\{k, W_k \notin T_{i-1}\}}\} & \text{if } A_i \in T_{i-1} \text{ or } i = s. \end{cases}$$

- Let $T^{\mathcal{D}, W}$ denote the rooted tree (T_s, A_1) .

Gromov–Hausdorff–Prokhorov convergence

By Lemma 2 and Theorem 1, the GHP convergence follows from the GP and GH convergence. So we only have to consider the GH convergence. To this end, we use the same approach we used to prove the compactness of ICRT. With this approach we compute precisely the Hausdorff distance

between the first leaves and the whole tree. In particular, recall from (1.6) the next approximation:

$$d_H(\mathcal{T}_y^\Theta, \mathcal{T}^\Theta) \approx \int_y^\infty \frac{dl}{l\mathbb{E}[\mu^\Theta[0, l]]}.$$

Hence, to have leaf tightness, we assume for ICRT that:

Assumption 1.5.

$$\lim_{y \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \int_y^{+\infty} \frac{dl}{l\mathbb{E}[\mu^{\Theta_n}[0, l]]} = 0.$$

We do similarly for \mathcal{P} -trees. Recall Algorithm 4. For $i \in \mathbb{N}$ let $X_i := \inf\{j \in \mathbb{N}, B_j^\mathcal{P} = V_i\}$. Then let $\mu^\mathcal{P} := \sum_{i=1}^\infty p_i/\sigma^\mathcal{P} \delta_{\sigma^\mathcal{P} X_i}$. Here, the previous integral diverges since $\mu^\mathcal{P}[0, \infty] < \infty$. To solve this issue, we truncate the integral and rely on

$$d_H(\mathcal{T}_y^\Theta, \mathcal{T}_{y'}^\Theta) \approx \int_y^{y'} \frac{dl}{l\mathbb{E}[\mu^\Theta[0, l]]}.$$

The main idea is then to take y' as large as possible such that the approximation stays "optimal" and then to estimate $d_H(\mathcal{T}_{y'}, \mathcal{T})$. For this reason we assume for \mathcal{P} -trees that:

Assumption 1.6. The two following assumptions hold:

- (i) For every $\mathcal{P} \in \Omega_\mathcal{P}$, let $t^\mathcal{P} := \inf\{l \in \mathbb{N}, \mathbb{E}[\mu^\mathcal{P}[0, l]] \geq 1/(2\sigma^\mathcal{P})\}$. We have,

$$\lim_{y \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \int_y^{\sigma^{\mathcal{P}_n} t^{\mathcal{P}_n}} \frac{dl}{l\mathbb{E}[\mu^{\mathcal{P}_n}[0, l]]} = 0.$$

- (ii) $\ln s^{\mathcal{P}_n} = o(1/\sigma^{\mathcal{P}_n})$, where for $\mathcal{P} \in \Omega_\mathcal{P}$, $s^\mathcal{P} := \#\{i \in \mathbb{N}, p_i^\mathcal{P} > 0\}$.

For \mathcal{D} -trees, we add an extra assumption to deal with the vertices of degree 1. Those vertices are particular to deal with since they cannot result in repetitions, but still add length to $T^\mathcal{D}$. Recall Algorithm 3. For $\mathcal{D} \in \Omega$, let $N^\mathcal{D}$ be the number of leaves in $\mathcal{T}^\mathcal{D}$. For $i \in \mathbb{N}$, let $X_i := \inf\{j \in \mathbb{N}, A_j^\mathcal{D} = V_i\}$. Then let $\mu^\mathcal{D} := \sum_{i=1}^\infty (d_i - 1)/\sigma^\mathcal{D} \delta_{\sigma^\mathcal{D}/s^\mathcal{D} X_i}$.

Assumption 1.7. The three following assumptions hold :

- (i) Let for every $\mathcal{D} \in \Omega_\mathcal{D}$, $t^\mathcal{D} := \inf\{l \in \mathbb{N}, \mathbb{E}[\mu^\mathcal{D}[0, l]] \geq N^\mathcal{D}/(2\sigma^\mathcal{D})\}$. We have,

$$\lim_{y \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \int_y^{\sigma^{\mathcal{D}_n} t^{\mathcal{D}_n}/s^{\mathcal{D}_n}} \frac{dl}{l\mathbb{E}[\mu^{\mathcal{D}_n}[0, l]]} = 0.$$

- (ii) $\ln(s_{\geq 2}^{\mathcal{D}_n})s^{\mathcal{D}_n}/N^{\mathcal{D}_n} = o(s^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n})$ where for all $\mathcal{D} \in \Omega_\mathcal{D}$, $s_{\geq 2}^\mathcal{D} := \#\{i \in \mathbb{N}, d_i^\mathcal{D} \geq 2\}$.
 (iii) $\ln(N^{\mathcal{D}_n})/\ln(s^{\mathcal{D}_n}/s_1^{\mathcal{D}_n}) = o(s^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n})$ where for all $\mathcal{D} \in \Omega_\mathcal{D}$, $s_1^\mathcal{D} := \#\{i \in \mathbb{N}, d_i^\mathcal{D} = 1\}$.

Those assumptions are close from optimal, and only the assumptions (ii) can be improved. We refer to Section 3.4.2 for details. Instead, we would like to insist on how to apply those results. First it is not necessary to compute $t^\mathcal{P}$ or $t^\mathcal{D}$. Indeed, since $\int dl/(l\mu[0, l])$ vary slowly, $t^\mathcal{D}$ can morally be replaced by an upper bound at small cost. In particular, note that $t^{\mathcal{P}_n} = O(s^{\mathcal{P}_n})$ and $t^{\mathcal{D}_n} = O(s_2^{\mathcal{D}_n})$. Moreover, under the assumptions (ii) it is equivalent to use $t^\mathcal{P}$, $t^\mathcal{D}$ or those upper bounds.

Also, one do not need to compute exactly $\mathbb{E}[\mu[0, l]]$. Moreover, μ can be approximated by a sum of independent random variables (see Algorithm 6). Then, we do not need to estimate $\mathbb{E}[\mu[0, l]]$ for l small. And for l large one morally only needs the tail of the degrees to approximate $\mathbb{E}[\mu[0, l]]$. Hence, for random degree sequence, since this tail tends to be morally concentrated, one can prove and use large deviation inequalities for $\mathbb{E}[\mu[0, l]]$.

Theorem 5

The following convergences hold weakly for the GHP-topology.

(a) If $\mathcal{D}_n \Rightarrow \Theta$, $\mathcal{M}^{\mathcal{D}_n} \rightarrow 0$, and Assumption 1.7 is satisfied then

$$(\mathcal{V}^{\mathcal{D}_n}, (\sigma^{\mathcal{D}_n}/s^{\mathcal{D}_n})d^{\mathcal{D}_n}, \mathcal{M}^{\mathcal{D}_n}) \xrightarrow{\text{WGHP}} (\mathcal{T}^\Theta, d^\Theta, \mathfrak{p}^\Theta).$$

(b) If $\mathcal{P}_n \Rightarrow \Theta$, $\mathcal{M}^{\mathcal{P}_n} \rightarrow 0$, and Assumption 1.6 is satisfied then

$$(\mathcal{V}^{\mathcal{P}_n}, \sigma^{\mathcal{P}_n}d^{\mathcal{P}_n}, \mathcal{M}^{\mathcal{P}_n}) \xrightarrow{\text{WGHP}} (\mathcal{T}^\Theta, d^\Theta, \mathfrak{p}^\Theta).$$

(c) If $\Theta_n \Rightarrow \Theta$, and Assumption 1.5 is satisfied then

$$(\mathcal{T}^{\Theta_n}, d^{\Theta_n}, \mathfrak{p}^{\Theta_n}) \xrightarrow{\text{WGHP}} (\mathcal{T}^\Theta, d^\Theta, \mathfrak{p}^\Theta).$$

Height of \mathcal{D} -trees, \mathcal{P} -trees and ICRT

With the same method, we obtain some upper bounds for the height of \mathcal{D} -trees, \mathcal{P} -trees and ICRT. Those bounds are similar to those proved by Addario–Berry, Devroye, Janson [7] and Kortchemski [58] for Galton–Watson trees and Lévy trees.

Theorem 6

There exists some constants $c, C > 0$ such that:

(a) For every $\mathcal{D} \in \Omega_{\mathcal{D}}$ and $x \in \mathbb{R}^+$:

$$\mathbb{P} \left(c \frac{\sigma^{\mathcal{D}}}{s^{\mathcal{D}}} H(T^{\mathcal{D}}) > x + \int_1^{\frac{\sigma^{\mathcal{D}}}{s^{\mathcal{D}}} t^{\mathcal{D}}} \frac{dl}{l \mathbb{E}[\mu^{\mathcal{D}}[0, l]]} + \frac{\ln(s_{\geq 2}^{\mathcal{D}}) \sigma^{\mathcal{D}}}{N^{\mathcal{D}}} + \frac{\sigma^{\mathcal{D}}}{s^{\mathcal{D}}} \frac{\ln(N^{\mathcal{D}})}{\ln\left(\frac{s^{\mathcal{D}}}{s_1^{\mathcal{D}}}\right)} \right) \leq C e^{-cx \mathbb{E}[\mu^{\mathcal{D}}[0, x]]}.$$

(b) For every $\mathcal{P} \in \Omega_{\mathcal{P}}$ and $x \in \mathbb{R}^+$:

$$\mathbb{P} \left(c \sigma^{\mathcal{P}} H(T^{\mathcal{P}}) > x + \int_1^{\sigma^{\mathcal{P}} t^{\mathcal{P}}} \frac{dl}{l \mathbb{E}[\mu^{\mathcal{P}}[0, l]]} + \sigma^{\mathcal{P}} \ln(s^{\mathcal{P}}) \right) \leq C e^{-cx \mathbb{E}[\mu^{\mathcal{P}}[0, x]]}.$$

(c) For every $\Theta \in \Omega_{\Theta}$ and $x \in \mathbb{R}^+$:

$$\mathbb{P} \left(c H(\mathcal{T}^\Theta) > x + \int_x^{+\infty} \frac{dl}{l \mathbb{E}[\mu^\Theta[0, l]]} \right) \leq C e^{-cx \mathbb{E}[\mu^\Theta[0, x]]}.$$

Rerooting and rebranching principle

Since \mathcal{D} -trees converge toward ICRT, and since \mathcal{D} -trees satisfies some equality in distribution, then so must ICRT. From those equalities, we derive several main technical principles. First for $i \in \mathbb{N}$, one can reroot the ICRT at X_i by simply changing in Algorithm 7, X_i by 0. It is useful to estimate what happen around a vertex of large degree. Then $\{Y_i\}_{i \geq 0}$ behaves as independent random leaves. So we can use the first cuts to estimate the mass of some sets in the ICRT.

As explained earlier at the end of Section 1.2.1, we tried to avoid using those principles. However, we needed them to compute the fractal dimensions of the ICRT's looptree in Chapter 5. And we will expansively use them in [14] to prove the dual local time property for ICRT.

1.2.3 Limit of connected multigraph with fixed degree sequence

Main approach to study the connected components

We now discuss Chapter 4, which concerns the configuration model and multiplicative graphs. To study the geometry of their connected components, Addario-Berry, Broutin, and Goldschmidt [4], developed a general approach which is divided in two main steps:

- (a) First one encodes the random graphs into stochastic processes, and study those processes to deduce several limits for relevant quantities of the largest connected components such as the size, surplus, degrees. . . This has been noticed in the ground-breaking work of Aldous [12].
- (b) Then, one use those convergences to reduce the problem to a study of a single connected component conditioned on those quantities.

This approach was first used in [4] for Erdős–Rényi graphs. It has then been further developed for multiplicative graphs and the configuration model in many different regimes. We refer to [4, 24, 23] for the homogeneous case, [42, 47, 54] for the power law case, and [36, 37] for a unified regime for multiplicative graphs. In Chapter 4 we solve (b), under what we believe the weakest assumptions. So we reduce the study of the largest connected components to (a), which tends to be simpler.

The other main contribution of Chapter 4 is that multiplicative (multi)graphs are limits of the configuration model. We have further developed this point of view all along this introduction. Notably, we will come back at the end of this section to (a), to provide a new combinatorial point of view on the exploration process of Broutin, Duquesne, Wang [36, 37].

Cycle breaking algorithm.

The surplus of a connected multigraph is the number of edges that one needs to delete to transform it into a tree. For $\mathcal{D} = (d_1, d_2, \dots, d_s)$, and $k \in \mathbb{N}$, we call a (\mathcal{D}, k) -graph a uniform connected multigraph among those with surplus k and such that for every $1 \leq i \leq k$, V_i has degree $d_i + 1$. In Chapter 4 we generalize the results of Chapter 3 to (\mathcal{D}, k) -graphs and their limits (\mathcal{P}, k) -graphs and (Θ, k) -ICRG (inhomogeneous continuum random graphs).

To this end, we actually focus on the connections between (\mathcal{D}, k) -graphs and \mathcal{D} -trees. The main idea is to revert the cycle breaking algorithm of Addario-Berry, Broutin, Goldschmidt, Miermont [5] to construct (\mathcal{D}, k) -graphs from a biased \mathcal{D} -trees. We use this algorithm as follows (see Figure 1.6). Take a connected multigraph with surplus k , and repeat k times: choose an edge uniformly among all the edges that can be removed without disconnecting the graph, then cut this edge in the middle. By doing so, we add $2k$ named leaves $(\star_i)_{1 \leq i \leq 2k}$ and keep the degrees. Then to reconstruct the original multigraph it suffices to repair the broken edges by fusing the pairs in $(\star_i)_{1 \leq i \leq 2k}$.

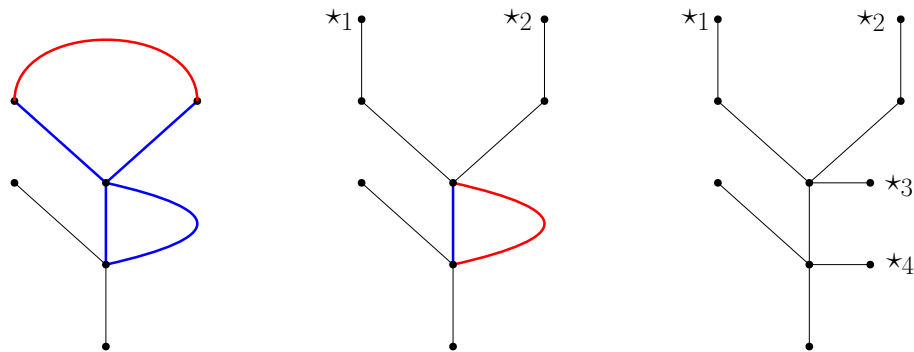


Figure 1.6: The steps of a cycle breaking of a multigraph with surplus 2. At each step we chose an edge (in red) among those that does not disconnect the graph (in red/blue) and "break it". Those edges can be repaired by fusing \star_1 with \star_2 and \star_3 with \star_4 . We omit the labels of the other vertices.

Although the cycle breaking preserves the degrees, it is not a bijection between (\mathcal{D}, k) -graphs and \mathcal{D} -trees. Indeed, for a given graphs there are many corresponding trees. To avoid this issue we bias the trees by the probability that they was obtained at the end of the cycle breaking. This bias is then a function of $(d(\star_i, \star_j))_{1 \leq i, j \leq 2k}$. And by proving that it converges, we deduce from Chapter 3, that (\mathcal{D}, k) -graphs also converges.

Finally, although the GP convergence of \mathcal{D} -trees implies the convergence of $(d(\star_i, \star_j))_{1 \leq i, j \leq 2k}$, some extra care is needed since the bias explodes when $(\star_i)_{1 \leq i \leq 2k}$ are close. For this reason, the main part of the proofs consists in lowerbounding $(d(\star_i, \star_j))_{1 \leq i, j \leq 2k}$, which we do with the first branches of our stick-breaking constructions.

The exploration processes of the configuration model

We now discuss the exploration processes used to study the degrees and surplus of the configuration model and multiplicative graphs. We first discuss the configuration model. Recall that in this model, we construct a multigraph by uniformly fusing some pair of half edges. Actually, we can construct the multigraph by recursively fusing an half edge that we chose together with a uniform one.

The main idea is then to construct the connected components one by one. Think of this construction as an exploration (see Figure 1.7). We have at each step a connected component with some full edges and some half edges. We pick at each step one half edge of our connected component then look at the half edges it is fused. (The order we pick the edges is not relevant here.) We then add the new edge, the eventual corresponding new vertex and its half edges. When our component no longer have half edges we explore a new one starting from a single vertex.

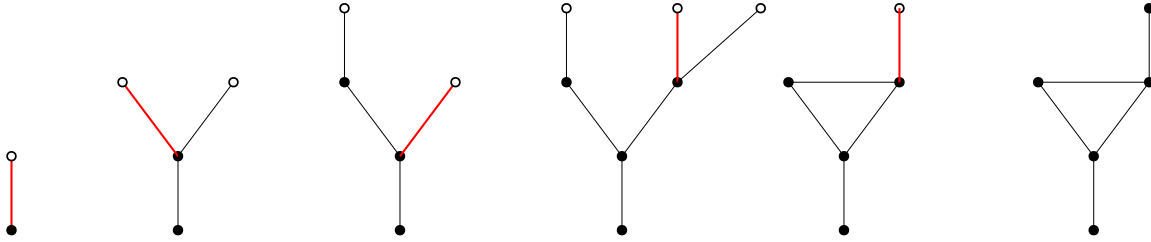


Figure 1.7: An exploration of a connected component. At each step the edge we will explore is red. Half edges are represented as segments between a known black vertex and a unknown white vertex.

The exploration process is then the real process which associate to each step the number of half edges of our current connected component minus twice the number of component we have explored. This process thus encodes the sizes of the connected components, the degrees of the vertices in the components, and so the surplus of the components.

It can be approximated as follows. Let $\mathcal{D} = (d_1, \dots, d_k)$ be our degree sequence. Let $S = \sum_{j=1}^k d_j$. Let $(A_i, U_i)_{1 \leq i \leq S}$ be a permutation of our half edges $(i, j)_{1 \leq i \leq k, 1 \leq j \leq d_i}$. Each time we need to sample a new edge, or pick a new half edge to start a new connected components we pick the first half edge in $(A_i^{\mathcal{D}}, U_i^{\mathcal{D}})_{1 \leq i \leq S}$ that is not yet fused. By doing so, an approximation of the exploration process is

$$\mathfrak{X}^{\mathcal{D}} : i \mapsto \sum_{j=1}^i (d_{A_j} \mathbf{1}_{A_j \notin \{A_1, \dots, A_{j-1}\}} - 2). \quad (1.7)$$

Indeed, when we see a new vertex we add its half edges, and at each step we lose the two half edges we fuse. It is though only an approximation as we need to morally skip some steps in $(A_i, U_i)_{1 \leq i \leq S}$.

A new point of view on the exploration processes of multiplicative graph.

In Chapter 4 we prove that the configuration model can converge toward multiplicative multigraphs. Recall Algorithms 1.1 and 1. Recall also that multiplicative graphs are obtained by removing the multiple edges and loops of multiplicative multigraphs.

Lemma 11

Let $\mathcal{P} = (p_1, \dots, p_s)$ a tuple in \mathbb{R}^{+*} . For $n \in \mathbb{N}$, let $\mathcal{D}_n = (d_i^n)_{1 \leq i \leq s^n}$ a tuple of integer with even sum. If $s^n \rightarrow \infty$, and for $1 \leq i \leq s$, $d_i^n \sim \sqrt{s^n} p_i$, and for $s < i \leq s^n$, $d_i^n = 1$. Then,

$$(\#_{i,j}(\text{CM}^{\mathcal{D}_n}))_{1 \leq i,j \leq s} \xrightarrow{(d)} (\#_{i,j}(\text{MG}^{\mathcal{P}+}))_{1 \leq i,j \leq s}.$$

Based on this convergence, we present a new point of view on the exploration process of multiplicative graphs. We refer to Aldous, Limic [13] who introduce a similar process to study the sizes of the connected components. We also refer to Broutin, Duquesne, Wang [36, 37] who recently introduced a way to reconstruct multiplicative graphs from their exploration processes.

First, we can rewrite (1.7), by writing for $i \in \mathbb{N}$, $X_i^{\mathcal{D}} := \inf\{t \in \mathbb{N}, A_t^{\mathcal{D}} = V_i\}$, as

$$\mathfrak{X}^{\mathcal{D}_n} : i \mapsto \sum_{1 \leq k \leq s^n} d_k^n \mathbf{1}_{X_k^n \leq i} - 2i.$$

Then under the assumptions of Lemma 11 it is easy to check that $(X_i^n / \sqrt{s^n})_{1 \leq i \leq s}$ converges toward some independent exponential random variables $(X_i^{\mathcal{P}})_{1 \leq i \leq s}$ of rate $(p_i)_{1 \leq i \leq s}$. One then deduce that the rescaled process $t \in \mathbb{R}^+ \mapsto \mathfrak{X}_i^{\mathcal{D}_n}(\lfloor t / \sqrt{s^n} \rfloor) / \sqrt{s^n}$ converges as $n \rightarrow \infty$ toward

$$\mathfrak{X}^{\mathcal{P}} : t \mapsto \sum_{1 \leq k \leq s} p_k \mathbf{1}_{X_k^{\mathcal{P}} \leq t} - t.$$

This is exactly the exploration process used by Broutin, Duquesne, Wang [36, 37].

To morally compare with the configuration model, first recall that multiplicative multigraphs can be seen as a configuration model with "infinite degrees and leaves". Then $(X_k)_{k \leq s}$ are the first instants we see $(V_k)_{k \leq s}$. The jumps then corresponds to the new vertices we encounter in the exploration, and to the fact we add their half edges. And the linear drift corresponds to the fusions. Here since the largest degrees are negligible in front of the total degree, we actually spent most of the time fusing an half edge of a vertex of large degree with an half edge of a new leaf.

Once we have the exploration process the main question is to reconstruct the connected components. To this end, we know at each step which edge of which vertex we decided to explore. (Here different ways to explore lead to different constructions.) And we know when we encounter a new vertex. When we do, we add an edge between the vertex we explore and the new one. We do not know however when we pick an half edge of a vertex we already explored in the exploration.

We can still add the edges in this cases with extra randomness. Indeed, \mathfrak{X} directly gives how many half edges we currently have in our connected components, and depending of our exploration we also know where they are. So we know at each step the probability of picking an half edge in our connected component and where. This idea was first introduced by Addario-Berry, Broutin, and Goldschmidt [4], and then extended for multiplicative graphs by Broutin, Duquesne, Wang [36, 37].

The major point of Chapter 4 is that we actually do not need to consider those edges. Indeed from the exploration process we have the degrees of the vertices in each connected components, and the surplus which corresponds to the number of edges between between a new vertex and an old one. From those degrees and surplus, we have results on the geometry of the connected components.

1.2.4 Looptree, Fennec, and Snake of ICRT

Plane trees and plane \mathbb{R} -trees

We now discuss Chapter 4 and the closely related forthcoming papers [26, 25]. In those papers we will prove several limits of objects constructed from \mathcal{D} -trees. In Chapter 4 we define and study the limits. The methods of Chapter 4 are close from those we used to study ICRT in Chapter 2, and the main difficulty was to formally define the limits.

The object we study was usually defined for Lévy trees using their associated Lévy processes. We might have adapted those definitions using the processes of Aldous, Miermont, Pitman [14, 15]. However, as they explained those processes are hard to study. So we decided to use instead our stick-breaking constructions. To do so we introduced a new theory of plane \mathbb{R} -tree.

Recall that a plane tree is a rooted tree with for each of its vertex an order for its children. Similarly one can give to a \mathbb{R} -tree \mathcal{T} , for each vertex $x \in \mathcal{T}$, an order on the connected components of $\mathcal{T} \setminus \{x\}$. This was done by Aldous [11] for binary continuum trees, then by Duquesne [48] for general \mathbb{R} -trees. However, using general order is not enough to define several objects. And it leads to measurability issues which prevents us from studying limits. For those reasons, we introduce the notion of angle which can be seen as a numbering of the children of each vertex.

For every \mathbb{R} -tree \mathcal{T} , and $x, y \in \mathcal{T}$, let $\llbracket x, y \rrbracket$ denote the geodesic path between x and y . The closest common ancestor of $x, y \in \mathcal{T}$ is the vertex $x \wedge y \in \llbracket \rho, x \rrbracket \cap \llbracket \rho, y \rrbracket$ which maximizes $d(\rho, z)$. For every $x \in \mathcal{T}$, the degree $\deg(x)$ of x in \mathcal{T} is the number of connected components of $\mathcal{T} \setminus \{x\}$. An angle function on a rooted \mathbb{R} -tree (\mathcal{T}, d, ρ) is a function $u : \mathcal{T}^2 \rightarrow [0, 1]$ such that:

- For all $x \in \mathcal{T}$, $u_{x,\rho} = u_{x,x} = 0$.
- For all $x \in \mathcal{T}$, $y, z \in \mathcal{T} \setminus \{x\}$, $u_{x,y} = u_{x,z}$ iff y and z are connected in $\mathcal{T} \setminus \{x\}$.

A plane \mathbb{R} -tree is a rooted \mathbb{R} -tree equipped with an angle function. For technical reasons, we further assume that a plane \mathbb{R} -tree is balanced: for every $x, y \in \mathcal{T}$ if $\deg(x) = 2$ then $u_{x,y} \in \{0, 1/2\}$.

Plane ICRT

Recall Algorithm 7. We root ICRT at 0. To define an angle function, it suffices to define for each $x \in \mathcal{T}$, for each connected component C of $\mathcal{T} \setminus \{x\}$ with $0 \notin C$, the value of $U_{x,y} \in [0, 1]$ for a unique $y \in C$. So, the following algorithm a.s. does well define an angle function on ICRT.

Algorithm 9: Construction of the uniform angle function U on the ICRT.

- Let $(U_{X,i})_{i \in \mathbb{N}}, (U_{Z,i})_{i \in \mathbb{N}}$ be independent uniform random variables in $[0, 1]$.
- Let U be the unique angle function on \mathcal{T} such that:
 - For every $i \in \mathbb{N}$, $U(X_i, Y_{\inf\{a \in \mathbb{N} : Y_a > X_i\}}) = U_{X,i}$
 - For every $i \in \mathbb{N}$, $U(Z_i, Y_{i+1}) = U_{Z,i}$.
 - For every $x \in \mathbb{R}^+ \setminus \bigcup_{i \in \mathbb{N}} \{X_i, Y_i\}$, we have $U(x, Y_{\inf\{a \in \mathbb{N} : Y_a > x\}}) = 1/2$.

This angle function was already in the lines 4 and 5 of Algorithm 6 the construction of the continuum \mathcal{D} -trees. And Algorithm 9 is just the limits of those lines. This will be useful in [26, 25] to study scaling limits of objects constructed from plane \mathcal{D} -trees.

Before we explain our other constructions we have to speak about \mathcal{P} -trees. Those trees can be seen as ICRT with finite mass and a modified distance (see Section 3.5.2). Since we do not use here the tree distances, we simply see \mathcal{P} -trees as ICRT with finite mass. This also corresponds with the fact that \mathcal{P} -trees can be seen as having infinite degrees and leaves.

The contour order

To encode a tree with real processes, we need to define an order on the trees. We informally define this order as follows (see Figure 1.8). Embed the tree in the plan. Start at the left of the root. Then continuously turn clockwise around the tree. Stop when you reach the right of the root. We use the order we visit the vertices.

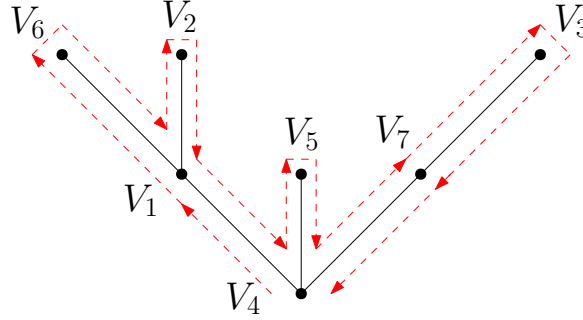


Figure 1.8: The plane tree is black, the order we explore the tree is given by the red arrows. We visit in order $V_4, V_1, V_6, V_1, V_2, V_1, V_4, V_5, V_4, V_7, V_3, V_7, V_4$. The discrete contour order is $(V_4, 1) \prec (V_1, 1) \prec (V_6, 1) \prec (V_1, 2) \prec (V_2, 1) \prec (V_1, 3) \prec (V_4, 2) \prec (V_5, 1) \prec (V_4, 3) \prec (V_7, 1) \dots$

This order is however badly defined as we visit each vertex several times. To avoid this issue we might use the first instants we visit each vertex. But this order "jump" in the tree which raises several issues. Instead, we consider each vertex several times and consider an order on $\{V_i\} \times \mathbb{N}$. The limit analog is then an order on the couples of a vertex and an angle.

We now formally define the order. Let $(\mathcal{T}, d, \rho, u)$ a plane \mathbb{R} -tree. For $x \in \mathcal{T}$, $(y, w) \in \mathcal{T} \times [0, 1]$ let $u_{x,y,w} := u_{x,y}$ if $x \neq y$ and let $u_{x,y,w} := w$ otherwise. For $\alpha = (x, v), \beta = (y, w) \in \mathcal{T} \times [0, 1]$, we say that α is at the left of β (or β is at the right of α) and write $\alpha \curvearrowright \beta$ if $u_{x \wedge y, x, v} < u_{x \wedge y, y, w}$. We say that β is in front of α and write $\alpha \triangleleft \beta$ if $x \in \llbracket 0, y \rrbracket$ and $u_{x,y,w} = v$. (See Figure 1.9.)

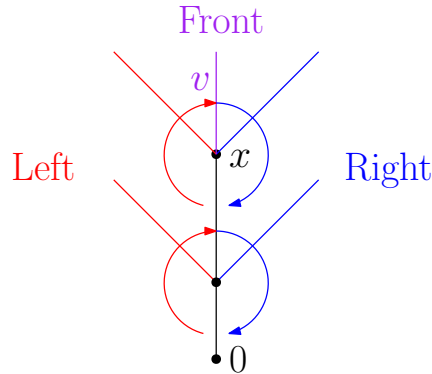


Figure 1.9: A vertex $(x, v) \in \mathcal{T} \times [0, 1]$ with its left (red), front (purple), right (blue) are represented.

Lemma 12: Definition of the contour order.

Let \prec be the binary relation defined on $\mathcal{T} \times [0, 1]$ such that for all $\alpha, \beta \in \mathcal{T} \times [0, 1]$:

$$\alpha \prec \beta \iff (\alpha \curvearrowright \beta) \text{ or } (\alpha \triangleleft \beta).$$

Then \prec is a total order relation on $\mathcal{T} \times [0, 1]$, and is called the contour order.

The contour path on ICRT

Several real processes are constructed as follows. Let f denote a function defined on a plane tree. Consider the previous exploration and write the vertices we visit in order as x_0, x_1, x_2, \dots . We consider the process $\mathfrak{F} : i \mapsto f(x_i)$. The contour path is a way to formalise the function $i \mapsto x_i$. The advantage of using this path is that we can split and study our constructions in two parts: the contour path, and some functions defined on our trees.

We formally define this path as follows. Recall the probability measure \mathbf{p} on the ICRT \mathcal{T} . Let $\mathbf{p}_{\mathcal{L}} := \mathbf{p} \times \mathbf{1}_{l \in [0,1]} dl$. The mass on the left of $\alpha \in \mathcal{T} \times [0, 1]$ is $\mathbf{p}_{\mathcal{L}}(\alpha) := \mathbf{p}_{\mathcal{L}}\{\beta : \beta \curvearrowright \alpha\}$. Think of the mass of the left as morally the number of steps before we see $x \in \mathcal{T} \times [0, 1]$.

Proposition 13

Assume that the ICRT is almost surely compact. Then almost surely there exists a unique continuous function $\mathfrak{C}_{\mathcal{T}} : [0, 1] \mapsto \mathcal{T}$ such that for every $(x, u) \in \mathcal{T} \times [0, 1]$, $\mathfrak{C}_{\mathcal{T}}(\mathbf{p}_{\mathcal{L}}(x, u)) = x$. We call $\mathfrak{C}_{\mathcal{T}}$ the contour path on \mathcal{T} .

Looptrees

For several processes it is more convenient to use instead a contour path on $\mathcal{T} \times [0, 1]$. To this end, we use the looptrees introduced by Curien and Kortchemski [44]. Informally, the looptree of a tree is obtained by replacing each vertex by a loop of size proportional to its degree.

Formally for every plane tree $T = (V_{i,j}(T))_{1 \leq i \leq s, 1 \leq j \leq d_i}$ and $1 \leq i \leq s$ let $V_{1,0}(T) = V_{1,d_i+1}(T) = V_i$. We define the looptree of T as the graph $\mathcal{L}(T)$ on $\{V_i\}_{1 \leq i \leq s}$ with edges $\bigcup_{1 \leq i \leq s: d_i \geq 1} \{V_{i,j}(T), V_{i,j+1}(T)\}_{0 \leq j \leq d_i}$. In [26] we will prove that the looptrees of plane \mathcal{D} -trees with a distance close from the graph distance converge toward the ICRT looptrees.

To define the looptree \mathcal{L} of an ICRT \mathcal{T} we want to replace each vertex by a loop. So we let $\mathcal{L} = \mathcal{T} \times [0, 1]$ with a proper pseudo-distance $d_{\mathcal{L}}$ corresponding to the cycles. Morally, the size of those cycles are proportional to the degrees. Also recall that morally the vertices of large degrees are $(X_i)_{i \in \mathbb{N}}$, and their degrees are proportional to $(\theta_i)_{i \in \mathbb{N}}$. Also recall that the parameter θ_0 corresponds to the contribution of the vertices of small degree.

So, we formally define the ICRT looptree as follows: Let \mathfrak{c} be the distance in the torus $[0, 1]$. For every $x, y \in \mathcal{T}, u \in [0, 1]$, let $U_{x,y,u} = U_{x,y}$ if $x \neq y$ and let $U_{x,y,u} = u$ otherwise. We define a pseudo-distance $d_{\mathcal{L}}$ on $\mathcal{T} \times [0, 1]$ such that for every $(x, u), (y, v) \in \mathcal{T} \times [0, 1]$,

$$d_{\mathcal{L}}((x, u), (y, v)) := \theta_0^2 / 4 d_{\mathcal{T}}(x, y) + \sum_{i \in \mathbb{N}} \theta_i \mathfrak{c}(U_{X_i, x, u}, U_{X_i, y, v}).$$

Finally let $(\mathcal{L}, d_{\mathcal{L}})$ be the completion of the pseudo-metric space $(\mathcal{T} \times [0, 1], d_{\mathcal{L}})$.

We prove the next results in Chapter 5. The proofs are similar to those we used to study ICRT.

Theorem 7

Almost surely $d_{\mathcal{L}}$ is finite on $\mathcal{T} \times [0, 1]$, and $(\mathcal{L}, d_{\mathcal{L}})$ is compact.

Theorem 8

Let,

$$\underline{\mathfrak{d}} := 1 + \liminf_{l \rightarrow \infty} \frac{\log \mathbb{E}[\mu[0, l]]}{\log l} \quad ; \quad \bar{\mathfrak{d}} := 1 + \limsup_{l \rightarrow \infty} \frac{\log \mathbb{E}[\mu[0, l]]}{\log l}. \quad (1.8)$$

Almost surely the upper Minkowski dimension, and Packing dimension of \mathcal{L} are $\bar{\mathfrak{d}}$. Almost surely the lower Minkowski dimension, and Hausdorff dimension of \mathcal{L} are $\underline{\mathfrak{d}}$.

The contour path on the looptree of the ICRT.

Let $\sim_{\mathcal{L}}$ denote the equivalent relation on \mathcal{L} such that for every $\alpha, \beta \in \mathcal{L}$, $\alpha \sim_{\mathcal{L}} \beta$ iff $d_{\mathcal{L}}(\alpha, \beta) = 0$. We prove in Section 5.5.3 that a.s. \mathfrak{p}_{\wedge} extends on \mathcal{L} to a function continuous at $\mathcal{L} \setminus (\mathbb{R}^+ \times [0, 1])$.

Theorem 9

Almost surely there exists a continuous function $\mathfrak{C} : [0, 1] \mapsto \mathcal{L}$ such that for every $\alpha \in \mathcal{L}$, $\mathfrak{C}(\mathfrak{p}_{\wedge}(\alpha)) \sim_{\mathcal{L}} \alpha$. We call \mathfrak{C} the contour path on \mathcal{L} (see Figure 1.10).

To prove Theorem 9, we first construct the contour path on the looptrees of the first branches. We then prove that those paths converge. Our proof is very technical and use many types of argument : the chaining method for the looptrees, our estimates on μ and several topological arguments to estimate the evolution of \mathfrak{p}_{\wedge} , some logical arguments to deal with the plane relations.

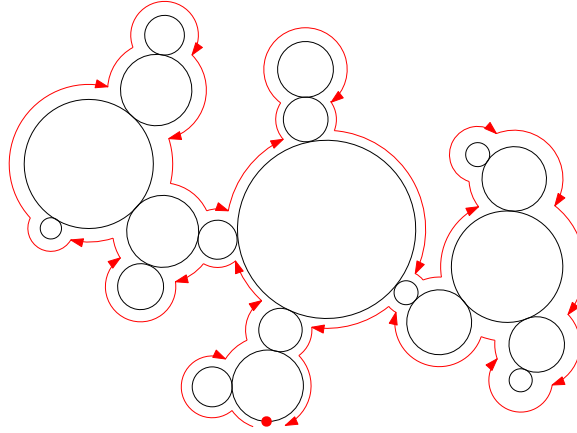


Figure 1.10: A looptree \mathcal{L} , with its contour path $\mathfrak{C} : [0, 1] \mapsto \mathcal{L}$ in red. The path start at the root $(0, 0)$ then "turn around" each cycle clockwise. It is continuous, "surjective", not "injective".

The density of $\{\mathfrak{p}_{\wedge}(\alpha), \alpha \in \mathcal{L}\}$ (see Lemma 5.5.3) implies that \mathfrak{C} is unique up to $\sim_{\mathcal{L}}$. Moreover it also implies with the existence of \mathfrak{C} the next result which in turns implies Proposition 13. We will use it to construct several classical processes from ICRT.

Proposition 14

Almost surely for every continuous function $F : \mathcal{L} \mapsto \mathbb{R}$, $F \circ \mathfrak{C}$ is the unique continuous function f such that for every $\alpha \in \mathcal{L}$, $f(\mathfrak{p}_{\wedge}(\alpha)) = F(\alpha)$.

Alongside Theorem 9 we also prove the next result. Recall that a function $f : (X, d) \mapsto (Y, d')$ is α -holder continuous if there exists $C \in \mathbb{R}^+$ such that $\forall x, y \in X$, $d'(f(x), f(y)) \leq C d(x, y)^\alpha$.

Theorem 10

Recall (1.8). Almost surely \mathfrak{C} is Hölder continuous with any exponent smaller than $1/\bar{d}$.

The height process

The height process is the simplest example of process that can be constructed with the contour path. For discrete trees it is informally constructed as follows. Follow the contour of the tree. Associate to each step i the height of your position (see Figure 1.11). So we define the Height process of the ICRT as $\mathfrak{H} : x \in [0, 1] \mapsto d_{\mathcal{T}}(0, \mathfrak{C}_{\mathcal{T}}(x))$.

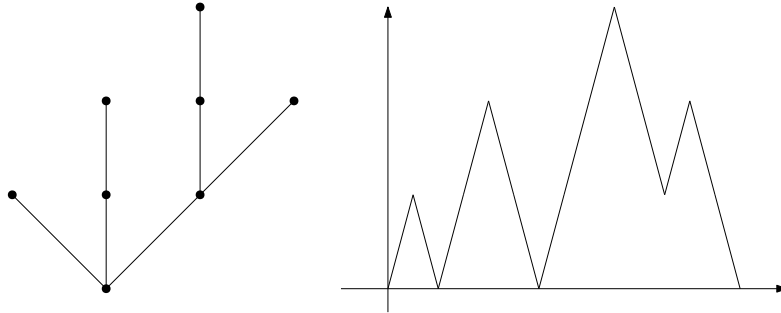


Figure 1.11: A plane tree with the graph of its height process.

Actually Duquesne [48] already gave, with its own formalism, a similar definition for general \mathbb{R} -trees. He furthermore prove that writing for every $x, y \in [0, 1]$,

$$d(x, y) = \mathfrak{H}(x) + \mathfrak{H}(y) - \inf_{z \in [x, y]} \mathfrak{H}(z),$$

$([0, 1], d)$, after taking its quotient by identifying points of $[0, 1]$ that are at d -pseudo distance 0, is isometric to original tree \mathcal{T} . This relation is well known for the discrete trees. And the height process is used precisely because this relation can be used to prove scaling limits of discrete trees. We refer to Le Gall [61] for more details on the subject.

Lukasiewicz walk and Exploration process

The exploration process is in a weird position. Indeed, it is one of the easiest process to study when one constructs the tree with real processes, but is hard to study with stick-breaking constructions. For discrete trees this process is called Lukasiewicz walk and is constructed as follows. Consider the contour of the tree. We associate at each new instant we see a new vertex the number of vertex that we have not visited that are connected to a vertex visited previously (see Figure 1.12).

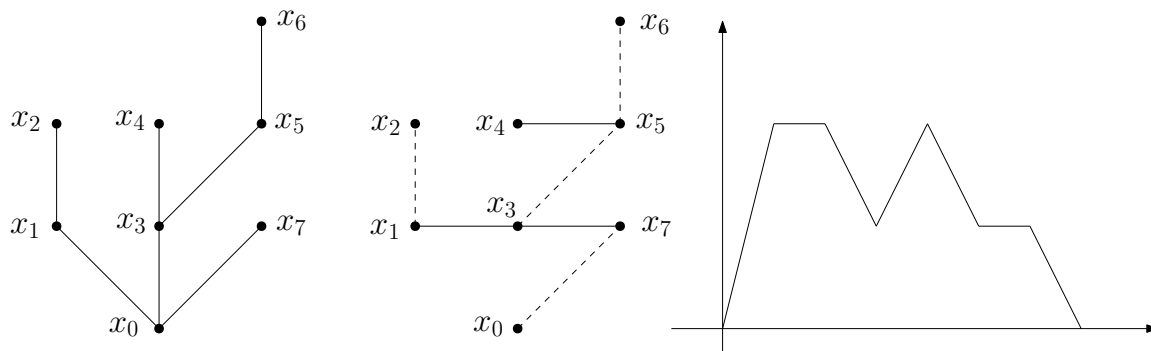


Figure 1.12: A plane tree, its dual, and its Lukasiewicz walk. We indicate the order we visit each vertex for the first time by $x_0, x_1, x_2, \dots, x_7$. The dashed edges of the dual have length 0. Note that the Lukasiewicz walk associate to each $i \in \mathbb{N}$ the distance between x_i and the root in the dual.

The Lukasiewicz walks of Galton–Watson trees are random walk with independent increments. Those random walks converge toward the exploration processes of Lévy trees, the Lévy processes. This is the starting point of the works on Galton–Watson trees and Lévy trees (see e.g. [50, 49]).

By analogy Aldous, Miermont, Pitman [14, 15] defined the exploration processes of ICRT as processes with exchangeable increments. They also defined the exploration process of \mathcal{P} -trees. This process is similar to the exploration of multiplicative graphs presented in Section 1.2.3. Also, similarly, the exploration process of \mathcal{P} -trees appear as the limit of the Lukasiewicz walk of \mathcal{D} -trees. And by seeing \mathcal{P} -trees as having infinite degrees and leaves, their exploration processes can be interpreted as a Lukasiewicz walk. We refer to Section 1.2.3 for a detailed analog point of view.

We now give an other definition. To distinguish from processes with exchangeable increments, we call our processes the Lukasiewicz walk. And we will prove in [25], using the unicity of the limits of the exploration process of \mathcal{P} -trees, that the Lukasiewicz walk and the exploration process of ICRT coincide. First let

$$\mathfrak{L} : (x, u) \in \mathcal{T} \times [0, 1] \mapsto \frac{\theta_0^2}{2} d_{\mathcal{T}}(0, x) + \sum_{i \in \mathbb{N} : X_i \in \llbracket 0, x \rrbracket} \theta_i (1 - U_{X_i, x, u}). \quad (1.9)$$

Then the Lukasiewicz walk of the ICRT is the unique càdlàg process such that for every $x \in \mathcal{T}$, $\mathfrak{X}(\mathfrak{p}_{\wedge}(x, 0)) = \mathfrak{L}(x, 0)$ (see [25]).

A geometric way to think (1.9) is through a dual tree. We construct the dual of a plane tree as follows (see Figure 1.12). For each vertex, for each of its children we add an edge between the children and the next one for the plane order. For the last child we add instead an edge of length 0 between this child and the father. We remove all the other edges. A simple counting argument then shows that the Lukasiewicz walk associates to a vertex its distance to the root in the dual tree.

This dual tree is very close from the looptree. Indeed, it can be obtained from the looptree by removing for each vertex of the original tree its edge connecting it to its left most child. They are also similar at the limit, and many partial results of Chapter 5 can be used to study the dual of ICRT. In particular, a.s. the dual is also compact and have the same fractal dimensions as the looptree. However because some edges are missing in the dual, the contour path is discontinuous on this tree. This makes many topological arguments around the Lukasiewicz walk very complex.

One of the most important relations between the tree and its dual concerns their local times: Assume that the ICRT is almost surely compact. Then a.s. for every $x > 0$, writing dl for the Lebesgue measure,

$$\mathfrak{H}(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} dl \left\{ 0 \leq y \leq x \leq 1 : \mathfrak{X}(y) \leq \inf_{z \in [0, x]} \mathfrak{X}(z) + \varepsilon \right\}, \quad (1.10)$$

and

$$\mathfrak{X}(x) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} dl \left\{ x \leq y \leq 1 : \mathfrak{H}(y) \leq \inf_{z \in [x, y]} \mathfrak{H}(z) + \varepsilon \right\}. \quad (1.11)$$

We will prove those dual relations in [25]. The first relation have been used repeatedly to prove and study scaling limits of Galton–Watson trees. And we found the second one while proving the first. The main idea behind our proof is to estimate the mass around a branch on the ICRT and its dual.

More precisely we have to estimate for each vertex x the \mathfrak{p} -mass of the part that is at the left (resp. right) of x and at distance at most ε in the dual (resp. ICRT). To this end, we use the fact that $\{Y_i\}_{i \in \mathbb{N}}$ behaves like independent uniform random leaves. Hence, by estimating the probability that Y_2, Y_3, \dots are at distance ε of the first branch and at the left (resp. right) of Y_1 we estimate the mass around a branch between 0 and a uniform leaf. By proving sufficiently precise concentration inequalities we can then conclude with some topological arguments.

This proof differ a lot from the rest of the literature. Indeed, to prove local times one would usually use some Markov property, but we use here our stick-breaking constructions. This proof is thus unique to the ICRT, and it is still open to prove local times for general non Markov process.

Field on discrete trees

We consider functions on trees constructed as follows. Let $\mathcal{D} = (d_1, \dots, d_s)$ be a degree sequence. Let $T = (V_{i,j}(T))_{1 \leq i \leq s, 1 \leq j \leq d_i}$ be a plane tree. Let $M_f = (f_i(j))_{1 \leq i \leq s, 1 \leq j \leq d_i}$ be a matrix in \mathbb{R}^+ . We consider the function \mathfrak{F}^{T, M_f} on T such that if ρ denotes the root of T then,

$$\mathfrak{F}(\rho) = 0 \quad \text{and} \quad \forall 1 \leq i \leq s, \forall 1 \leq j \leq d_i, \quad \mathfrak{F}(V_{i,j}(T)) = \mathfrak{F}(V_i) + f_i(j).$$

In [26] we will prove the convergence of \mathfrak{F} when \mathcal{D} converge, taking T a \mathcal{D} -tree, under some regime of convergence of M_f . To prove the convergence of F we proceed as usual by first proving the convergence of F along the first branches then proving leaf tightness.

The function F is easy to study along the first branches. Indeed, recall Algorithm 2. Studying F on the first branches is mostly equivalent to study (with the first branches themselves) the process $k \mapsto \sum_{i=1}^k f_{A_i, U_i}$, which can be done with some classical methods.

To prove the leaf tightness of F one can adapt the proof of the leaf tightness of \mathcal{D} -trees in Chapter 3. However, there is no general way to optimize this method for every matrix M_f , so one must think case by case. To this end, one must first upperbound by how much F can vary along a single branch. We proved Lemmas 4, 5, and 7 for this reason.

Some brief words on an application to random maps.

The below remarks on random maps are written for the specialist readers. For the others we refer to the courses of Miermont [71] and of Le Gall and Miermont [65] for an introduction to the subject.

Due to the relations with uniform planar maps with fixed even face degree sequences, we are interested in the case where for every $1 \leq i \leq s$, $\{f_i(j)\}_{0 \leq j \leq d_i}$ is a random walk with independent increments ± 1 conditioned by $f_i(0) = f_i(d_i) = 0$. We refer to Bouttier, Di Francesco, Guitter [33] and to Janson, Stefánsson [55] for the bijections involved. We also refer to Marzouk [68, 67, 69] for elaborate discussions on the subject.

More generally, we think that studying \mathfrak{F} for more M_f should be useful to study maps with arbitrary degrees. Indeed the above bijections send those maps to some trees with two types of edges and a field. By counting for each vertex its number of edges of a given type, we can define two kind of degrees. When those degrees are fixed, we can extend our stick-breaking constructions to the involved trees. (For three or more types our algorithms are biased.) However, because we could not estimate those degrees we could not get any relevant results on those maps.

Fennec and Snake of ICRT

Under the above described regimes, we will prove in [26] that F converges toward the "fennec" (for field+snake) of the ICRT (see also Marzouk [69] for a similar result). We now recall the definition of the fennec from Chapter 5.

Let \mathcal{T} denote the ICRT, let \mathcal{L} denote its looptree. Let $\mathfrak{B} : \mathbb{R}^+ \rightarrow \mathbb{R}$ be a Brownian motion. Define inductively \mathfrak{G} on \mathbb{R}^+ such that for every $i \in \mathbb{N}$ and $Y_i < x \leq Y_{i+1}$, we have $\mathfrak{G}(x) := \mathfrak{G}(Z_i) + \mathfrak{B}(x) - \mathfrak{B}(Y_i)$. To construct the fennec \mathfrak{F} , we need to show that \mathfrak{G} extends to a continuous function on \mathcal{T} if $\theta_0 > 0$. We actually prove the much stronger result with the chaining method.

Theorem 11

Almost surely \mathfrak{G} extends to a continuous function on \mathcal{T} if

$$\int_0^\infty \frac{dl}{l \sqrt{\mathbb{E}[\mu[0, l]]}} < \infty. \quad (1.12)$$

In Chapter 5, we prove that when the ICRT is a.s. compact, in particular when $\theta_0 > 0$, a.s. $\mathcal{L} = \mathcal{T} \times [0, 1]$, and a.s. $p_{\mathcal{L}, \mathcal{T}} : (x, u) \in \mathcal{L} \rightarrow x$ is continuous. Then the fennec is

$$\mathfrak{F} : \alpha \in \mathcal{L} \mapsto \frac{\theta_0}{\sqrt{6}} \mathfrak{G} \circ p_{\mathcal{L}, \mathcal{T}}(\alpha) + \sum_{i=1}^{\infty} \sqrt{\theta_i} \mathfrak{B}_i(U_{X_i, \alpha}). \quad (1.13)$$

With another application of the chaining method we prove the next result.

Theorem 12

Almost surely the sum in (1.13) converges uniformly on \mathcal{L} , so \mathfrak{F} is continuous.

We actually prove a stronger statement where the functions $(\mathfrak{B}_i)_{i \in \mathbb{N}}$ are replaced by more general random functions (see Section 5.9). We proved this extension in order to eventually study other regime of convergence for the random field described in the previous part.

A direct corollary of Theorem 12 is that \mathfrak{F} is a Gaussian free field on \mathcal{L} :

Proposition 15

Almost surely, conditionally on $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, U$, for every $\alpha, \beta \in \mathcal{L}$, $\mathfrak{F}(\alpha) - \mathfrak{F}(\beta)$ is Gaussian with variance

$$d'_{\mathcal{L}}(\alpha, \beta) := \frac{\theta_0^2}{6} d_{\mathcal{T}}(\alpha, \beta) + \sum_{i \in \mathbb{N}} \theta_i |U_{X_i, \alpha} - U_{X_i, \beta}| (1 - |U_{X_i, \alpha} - U_{X_i, \beta}|).$$

Finally we define the snake of the ICRT as $\mathcal{Z} := \mathfrak{F} \circ \mathfrak{C}$. This process has historically been used mostly to study the fennec. We think that the fennec should be used when one work with stick-breaking constructions, while the snake should be used when one work with tree encodings.

Due to the connections with the fractal properties of random maps we prove the next result.

Theorem 13

Almost surely \mathfrak{F} is Hölder continuous with any exponent smaller than $1/2$, and \mathcal{Z} is Hölder continuous with any exponent smaller than $1/2\bar{d}$.

We deduce the first part from the fact that \mathfrak{F} is a Gaussian free field on \mathcal{L} , which has finite upper Minkowski dimension. Our proof is quite general and is again based on the chaining method. We then deduce the second part using the Hölder continuity of \mathfrak{C} (Theorem 10).

A brief remark on random maps of ICRT

Since we defined for ICRT the looptrees, the contour order, and the fennecs, one can potentially use the bijection of Bouttier, Di Francesco, Guitter [33] to define the ICRT maps. However, they may be defined in two ways. On the one hand, one can consider the map constructed from the whole looptree. On the other hand, one can consider the maps constructed from the first branches then take the eventual limit as the number of branches grows. To prove that those definitions are coherent, one would then have to prove leaf tightness. Proving leaf tightness would also be a way to prove scaling limits. However, as we add branches to the tree the maps evolve in a very complex way.

Chapter 2

Compactness and fractal dimensions of inhomogeneous continuum random trees.

"Easy reading is damned hard writing."

Nathaniel Hawthorne

This chapter is adapted from [29], which is published in Probability Theory and Related Fields. Reproduced with permission from Springer Nature.

We introduce a new stick-breaking construction for inhomogeneous continuum random trees (ICRT). This new construction allows us to prove the necessary and sufficient condition for compactness conjectured by Aldous, Miermont and Pitman [14] by comparison with Lévy trees. We also compute the fractal dimensions (Minkowski, Packing, Hausdorff).

Contents

2.1	Introduction	48
2.2	Model and definition of the fractal dimensions	48
2.2.1	The ICRT and its construction	48
2.2.2	Fractal dimension	50
2.3	Main results	51
2.4	Preliminaries	53
2.4.1	Fundamental properties of μ	53
2.4.2	Key results on cuts and sticks	54
2.4.3	An estimate of distances in \mathcal{T}	55
2.5	The mass measure	56
2.5.1	The mass conservation lemma	57
2.5.2	Weak convergence of μ_n : proof of Theorem 2.3.1	58
2.5.3	Other convergences toward p : proof of Proposition 2.3.2	60
2.6	Compactness	62
2.6.1	Equivalent condition	62
2.6.2	The condition of Theorem 2.3.3 is sufficient for compactness	62
2.6.3	The condition of Theorem 2.3.3 is necessary for compactness	63
2.7	Fractal dimensions : proof of theorem 2.3.4	66
2.7.1	Upper bound for the Minkowski dimensions	66
2.7.2	Lower bound for the Packing dimension and the Hausdorff dimension	67
2.8	Appendix: Concentration inequalities	68

2.1 Introduction

Since the pioneer work of Aldous in [10], the study of continuum random trees (CRT) is considered as a powerful tool to study properties of large random discrete trees. In particular, it has been conjectured in [10], that the Brownian CRT is a universal limit for numerous models of trees with large height. This has been verified over and over. Furthermore the Brownian CRT model has been extended, for discrete trees with smaller height, toward two main distinct directions. On the one hand, Lévy trees are introduced, in Le Gall Le Jan [64] (see also Le Gall Duquesne [50, 49] for an extensive treatment), as limits of Galton-Watson trees. On the other hand, inhomogeneous continuum random trees (ICRT) are introduced by Aldous, Camarri and Pitman, in [16, 40], as limits of \mathcal{P} -trees. Those two distinct but similar models leave the following main problem: Finding a universal model for limits of random discrete trees (with no restriction on the height).

To solve this problem, we prove in a forthcoming paper (Chapter 3), that ICRT appears as limits of uniform random trees with fixed degree sequence. Since many models of interest can be studied under the spectrum of those trees, this proves that ICRT are universal. In particular Lévy trees are ICRT with random parameters. The aim of the present paper is twofold: obtain refined information about the ICRT, the universal limit object in particular concerning compactness and fractal dimensions, and introduce some tools for convergence that will be used in Chapter 3.

Our main results are derived from a new version of the stick-breaking construction of the ICRT from Aldous, Pitman [16]. Stick-breaking constructions generate a \mathbb{R} -tree (a loopless geodesic space see Le Gall [61] for an extensive treatment) and are separated in two steps:

- the line \mathbb{R}^+ is first cut into the segments ("sticks") $[0, Y_1], (Y_1, Y_2], (Y_2, Y_3] \dots$
- the segments are then re-arranged sequentially in a tree-like fashion by gluing $(Y_i, Y_{i+1}]$ at a point $Z_i \leq Y_i$. (see Figure 2.1)

Such a construction has been introduced by Aldous [10] for the Brownian CRT. Recently Amini, Devroye, Griffiths, Olver in [17] studied a case where cuts are fixed with $(Y_{i+1} - Y_i)_{i \in \mathbb{N}}$ decreasing. The condition of monotonicity has been removed by Curien and Haas in [43] where they construct a probability measure on \mathcal{T} , give a sufficient criterion for compactness of \mathcal{T} and compute the Hausdorff dimension of \mathcal{T} . We use similar methods in a setting where cuts and glue points are generated according to a random measure μ on \mathbb{R}^+ .

2.2 Model and definition of the fractal dimensions

2.2.1 The ICRT and its construction

Let us first present a generic deterministic stick-breaking construction. It takes for input two sequences in \mathbb{R}^+ called cuts $\mathbf{y} = (y_i)_{i \in \mathbb{N}}$ and glue points $\mathbf{z} = (z_i)_{i \in \mathbb{N}}$, which satisfy

$$\forall i < j, \quad y_i < y_j \quad ; \quad y_i \longrightarrow \infty \quad ; \quad \forall i \in \mathbb{N}, \quad z_i \leq y_i,$$

and creates an \mathbb{R} -tree by recursively "gluing" segment $(y_i, y_{i+1}]$ at position z_i (see Figure 2.1), or rigorously, by constructing recursively a consistent sequence of distances $(d_n)_{n \in \mathbb{N}}$ on $([0, y_n])_{n \in \mathbb{N}}$.

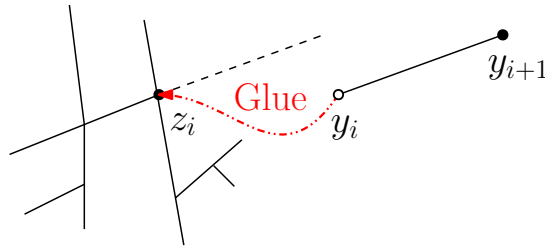


Figure 2.1: A typical step of the stick-breaking construction: the "gluing" of $(y_i, y_{i+1}]$ at z_i .

Algorithm 2.1. Generic stick-breaking construction.

- Let d_0 be the trivial distance on $\{0\}$.
- For each $n \geq 1$ define d_n on $[0, y_n]$ such that for each $x \leq y$:

$$d_n(x, y) := \begin{cases} d_{n-1}(x, y) & \text{if } x, y \in [0, y_{n-1}] \\ d_{n-1}(x, z_{n-1}) + |y - y_{n-1}| & \text{if } x \in [0, y_{n-1}], y \in (y_{n-1}, y_n] \\ |x - y| & \text{if } x, y \in (y_{n-1}, y_n] \end{cases}$$

where by convention $y_0 := 0$ and $z_0 := 0$.

- Let d be the unique metric on \mathbb{R}^+ which agrees with d_n on $[0, y_n]$ for each $n \in \mathbb{N}$.
- Let $\text{SB}(\mathbf{y}, \mathbf{z})$ be the completion of (\mathbb{R}^+, d) .

Remark. There is a more general way of gluing metric space. (see [39] for definition or [77] for similar work in this context). We prefer to work directly on \mathbb{R}^+ for practical reasons.

Now, let Ω be the space of sequences $\{\theta_i\}_{i \in \mathbb{N}}$ in \mathbb{R}^+ such that:

$$\sum_{i=0}^{\infty} \theta_i^2 = 1 \quad ; \quad \theta_1 \geq \theta_2 \geq \dots \quad ; \quad \theta_0 \neq 0 \text{ or } \sum_{i=1}^{\infty} \theta_i = \infty.$$

The ICRT of parameter $\Theta \in \Omega$ is the random \mathbb{R} -tree constructed via the following algorithm.

Algorithm 2.2. Classical construction of the Θ -ICRT from [16, 40]

- Let $(A_i, B_i)_{i \in \mathbb{N}}$ be a Poisson point process of intensity θ_0^2 on $\{(a, b) \in \mathbb{R}^{+2} : b \leq a\}$.
- Let $((A_{i,j})_{j \in \{0\} \cup \mathbb{N}})_{i \in \mathbb{N}}$ be a family of independent Poisson point processes of intensity $(\theta_i)_{i \in \mathbb{N}}$ on \mathbb{R}^+ and independent of $(A_i, B_i)_{i \in \mathbb{N}}$.
- Sort the elements of the (almost surely) locally finite set $\bigcup_{i=1}^{\infty} \{A_i\} \cup \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} \{A_{i,j}\}$ as $U = (U_i)_{i \geq 1}$ with $U_1 < U_2 < \dots$.
- For $i \geq 1$, let $V_i = \begin{cases} B_j & \text{if } U_i \text{ is of the form } A_j \\ A_{i,0} & \text{otherwise} \end{cases}$ and let $V = (V_i)_{i \geq 1}$.
- The (old) Θ -ICRT is defined as $(\mathcal{T}^*, d^*) = \text{SB}(U, V)$.

For technical reasons, it is convenient to deal with the following alternative construction.

Algorithm 2.3. New construction of the Θ -ICRT

- Let $(X_i)_{i \in \mathbb{N}}$ be a family of independent exponential random variables of parameter $(\theta_i)_{i \in \mathbb{N}}$.
- Let μ be the measure on \mathbb{R}^+ defined by $\mu = \theta_0^2 dx + \sum_{i=1}^{\infty} \delta_{X_i} \theta_i$.
- For each $l \in \mathbb{R}^+$ let μ_l be the restriction of μ to $[0, l]$.
- Let $(Y_i)_{i \in \mathbb{N}}$ be a Poisson point process on \mathbb{R}^+ of rate $\mu[0, l]dl$.
- Let $(Z_i)_{i \in \mathbb{N}}$ be a family of independent random variables with respective laws $\frac{\mu_{Y_i}}{\mu[0, Y_i]}$, $i \in \mathbb{N}$.
- The (new) Θ -ICRT is defined as $(\mathcal{T}, d) = \text{SB}(Y, Z)$.

Remarks. • The construction may fail because $\mu[0, l]$ may be infinite for some l . However, since $\mathbb{E}[\mu[0, l]] < \infty$, this almost surely never happens. (See Lemma 2.4.1)

• When $\theta_0 = 1$, the ICRT is the Brownian CRT.

• When $\theta_0 = 0$ and $\sum_{i=0}^{\infty} \theta_i < \infty$, the ICRT is still well defined and can be seen as a \mathcal{P} -tree with a modified distance (see Section 3.5.2, or Camarri Pitman [40]). Although we exclude this case to avoid any technical issues starting Section 2.5, it is easy to check the following properties. First the ICRT is almost surely not compact nor even separable. So the dimensions we compute are infinite. Also $\mu(\mathbb{R}^+) < \infty$ so Theorem 2.3.1 still holds. But Proposition 2.3.2 does not hold in this case.

The constructions in Algorithms 2.2 and 2.3 are equivalent that is:

Lemma 2.2.1. (\mathcal{T}^*, d^*) and (\mathcal{T}, d) have the same distribution.

Proof. First conditionally on $\{A_{i,0}\}_{i \in \mathbb{N}}$, $\{U_i, V_i\}_{i \in \mathbb{N}}$ is a Poisson point process on $\Delta := \{(a, b) \in \mathbb{R}^{+2} : b \leq a\}$ of intensity

$$\theta_0^2 dx dy + \sum_{i=1}^{\infty} \theta_i \mathbf{1}_{A_{i,0} \leq x} dx \times \delta_{A_{i,0}}.$$

Also, conditionally on $(X_i)_{i \in \mathbb{N}}$, $(Y_i, Z_i)_{i \in \mathbb{N}}$ is a Poisson point process on Δ of intensity

$$\theta_0^2 dx dy + \sum_{i=1}^{\infty} \theta_i \mathbf{1}_{X_i \leq x} dx \times \delta_{X_i}.$$

So since $(X_i)_{i \in \mathbb{N}}$ and $\{A_{i,0}\}_{i \in \mathbb{N}}$ have the same distribution, (U, V) and (Y, Z) also have the same distribution. Finally $(\mathcal{T}^*, d^*) = \text{SB}(U, V)$ and $(\mathcal{T}, d) = \text{SB}(Y, Z)$ have the same distribution. \square

Finally let us introduce some notation that will simplify many expressions later.

Definition. For $n \in \mathbb{N}$ let $l_n := Y_n - Y_{n-1}$ denote the length of the n th segment, and let $m_n := \mu(Y_{n-1}, Y_n]$ denote its weight. Then let $M_n := \mu[0, Y_n] = m_1 + \dots + m_n$.

2.2.2 Fractal dimension

In the entire section X is a metric space and for every $x \in X$, $\varepsilon > 0$, $B(x, \varepsilon)$ denotes the closed ball centered at x with radius ε . We recall the definitions of the fractal dimensions we compute in this paper.

Definition. (Minkowski dimensions) For every $\varepsilon > 0$ let N_ε be the minimal number of closed balls of radius ε to cover X . Define the Minkowski lower box and upper box dimensions respectively by

$$\underline{\dim}(X) := \liminf_{l \rightarrow \infty} \frac{\log N_{1/l}}{\log l} \quad \text{and} \quad \overline{\dim}(X) := \limsup_{l \rightarrow \infty} \frac{\log N_{1/l}}{\log l}.$$

Definition. (Packing dimension) For every $s \geq 0$ and $A \subset X$ let

$$P_0^s(A) := \limsup_{\delta \rightarrow 0} \left\{ \sum_{i \in I} \text{diam}(B_i)^s \mid \{B_i\}_{i \in I} \text{ are disjoint balls } B(x, r) \text{ with } x \in A \text{ and } r \leq \delta \right\}.$$

and

$$P^s(X) := \inf \left\{ \sum_{i=1}^{\infty} P_0^s(A_i) \mid X \subset \bigcup_{i=1}^{\infty} A_i \right\}.$$

Then P^s is a decreasing function of s , and we define the packing dimension of X as

$$\dim_P(X) := \sup\{s, P^s(X) < \infty\}.$$

Definition. (*Hausdorff dimension*) For every $s, r \geq 0$ write

$$H_r^s(X) := \inf_{\text{diam}(A_i) \leq r} \left\{ \sum_{i=1}^{\infty} \text{diam}(A_i)^s \mid X \subseteq \bigcup_{i=1}^{\infty} A_i \right\}.$$

The Hausdorff dimension of X is defined by

$$\dim_H(X) := \sup \left\{ s, \sup_{r \in \mathbb{R}^+} H_r^s(X) < \infty \right\}.$$

To compute the Packing dimension and Hausdorff dimension of the ICRT we will use the following extension of Theorem 6.9, and Theorem 6.11 from [70]. ([70] deals with subsets of Euclidian space, but the same arguments hold for every metric space.)

Lemma 2.2.2. Let p be a Borel probability measure on X and $s \in \mathbb{R}^+$.

- a) If p -almost everywhere $\liminf p(B(x, \varepsilon))\varepsilon^{-s} < +\infty$ as $\varepsilon \rightarrow 0$, then $\dim_P(X) \geq s$.
- b) If p -almost everywhere $p(B(x, \varepsilon)) = O(\varepsilon^s)$ as $\varepsilon \rightarrow 0$, then $\dim_H(X) \geq s$.

We have the well-known inequalities (see e.g. Chapter 3 of Falconer [53]):

Lemma 2.2.3. For every metric space X we have

$$\dim_H(X) \leq \underline{\dim}(X) \leq \overline{\dim}(X) \quad \text{and} \quad \dim_H(X) \leq \dim_P(X) \leq \overline{\dim}(X).$$

2.3 Main results

The first theorem defines a probability measure on ICRT.

Theorem 2.3.1. Almost surely there is a probability measure p on the tree \mathcal{T} such that

$$p_l := \frac{\mu_l}{\mu[0, l]} \xrightarrow[l \rightarrow \infty]{\text{weakly}} p.$$

Furthermore p has support \mathcal{T} , has no atoms and gives measure 1 to the set of leaves (the set of $x \in \mathcal{T}$ such that $\mathcal{T} \setminus \{x\}$ is connected).

This probability is also the limit of other natural empirical measures on \mathcal{T} :

Proposition 2.3.2. Let μ^\sim be the Lebesgue measure on \mathbb{R}^+ and $\mu^\bullet = \sum_{i=1}^{\infty} \delta_{Y_i}$. For every $l \in \mathbb{R}^+$ let μ_l^\sim (resp. μ_l^\bullet) be the restriction of μ^\sim (resp. μ^\bullet) to $\mathcal{T}_l = ([0, l], d)$. Also let for every $l \in \mathbb{R}^+$, $p_l^\sim = \frac{\mu_l^\sim}{\mu_l^\sim[0, l]}$ and $p_l^\bullet = \frac{\mu_l^\bullet}{\mu_l^\bullet[0, l]}$. Then

$$p_l^\bullet \xrightarrow[l \rightarrow \infty]{\text{weakly}} p \quad \text{and} \quad p_l^\sim \xrightarrow[l \rightarrow \infty]{\text{weakly}} p.$$

Intuitively speaking this comes from the fact that μ "dictates" how segments are glued together so the convergence of p_l implies the convergence of many others quantities.

Remark. Proposition 2.3.2 shows that p corresponds to the probability measure introduced in Aldous Pitman [16]. In particular independent leafs sampled by p "behave" like $(Y_i)_{i \in \mathbb{N}}$. (see [16] Corollary 8)

Then we prove the conjecture of Aldous, Miermont, Pitman in [14] about compactness.

Theorem 2.3.3. *The ICRT is almost surely compact if and only if*

$$\int_0^\infty \frac{dl}{l\mathbb{E}[\mu[0, l]]} < \infty. \quad (2.1)$$

Remark. The conjecture in [14] is based on a comparison between the ICRT and Levy trees introduced by Le Gall Le Jan [64]. Levy trees are characterized by their Laplace exponent ψ and are compact if and only if $\int_0^\infty \frac{dl}{\psi(l)} < \infty$ (see [49]). The formulation of the conjecture in [14] is based on an analog of the Laplace exponent in the setting of ICRT, which behaves like $\mathbb{E}[\mu[0, l]]$ (see Lemma 2.6.1) which turns out to be equivalent to (2.1).

For the proof of Theorem 2.3.3, we first translate the condition in (2.1) into a more convenient one: it turns out (Lemma 2.6.1) that

$$\int_0^\infty \frac{dl}{l\mathbb{E}[\mu[0, l]]} < \infty \quad \text{if and only if} \quad \sum_{n=1}^\infty \frac{\log \mathcal{X}_{2^n}}{\mathcal{X}_{2^n}} < \infty,$$

where for every $l \in \mathbb{R}^+$, \mathcal{X}_l is the real number such that $\mathbb{E}[\mu[0, \mathcal{X}_l]] = l$ (see Lemma 2.4.1 for existence and uniqueness).

To prove that the condition is sufficient, we will upper bound the law of the distance between a random point in $\mathcal{T}_{\mathcal{X}_{2^n}}$ and its projection on $\mathcal{T}_{\mathcal{X}_{2^{n-1}}}$. We then use this bound to prove that

$$d_H(\mathcal{T}_{\mathcal{X}_{2^n}}, \mathcal{T}_{\mathcal{X}_{2^{n-1}}}) \leq C \frac{\log \mathcal{X}_{2^n}}{\mathcal{X}_{2^n}},$$

where d_H denotes the Hausdorff distance on subsets of \mathcal{T} . For the Hausdorff topology, Cauchy sequences of compact sets converge toward a compact set so this proves that $\sum_{n=1}^\infty \frac{\log \mathcal{X}_{2^n}}{\mathcal{X}_{2^n}} < \infty$ implies that \mathcal{T} is compact.

The fact that the condition is necessary follows from an adaptation of an argument of Amini, Devroye, Griffiths, Olver in [17]. We show that, for some fixed constants $c, C \in (0, \infty)$ and for all k large enough:

$$c \sum_{n=k+1}^\infty \frac{\log \mathcal{X}_{2^n}}{\mathcal{X}_{2^n}} \leq d_H(\mathcal{T}, \mathcal{T}_{\mathcal{X}_{2^k}}) \leq C \sum_{n=k+1}^\infty \frac{\log \mathcal{X}_{2^n}}{\mathcal{X}_{2^n}}.$$

We then proceed to the computation of some fractal dimensions.

Theorem 2.3.4. *Almost surely*

$$\dim_P(\mathcal{T}) = \overline{\dim}(\mathcal{T}) = 1 + \limsup_{l \rightarrow \infty} \frac{\log l}{\log \mathbb{E}[\mu[0, l]]}.$$

Furthermore if $\log l = \mathbb{E}[\mu[0, l]]^{o(1)}$ then

$$\dim_H(\mathcal{T}) = \underline{\dim}(\mathcal{T}) = 1 + \liminf_{l \rightarrow \infty} \frac{\log l}{\log \mathbb{E}[\mu[0, l]]}.$$

Remark. If one replaces $\mathbb{E}[\mu[0, l]]$ by the Laplace exponent ψ then one recovers the formulas for the fractal dimensions of Levy trees obtained by Duquesne and Le Gall [50].

To prove Theorem 2.3.4, it suffices by Lemma 2.2.3 to upper bound the Minkowski dimensions and to lower bound the Packing and Hausdorff dimension. To upper bound $\overline{\dim}(\mathcal{T})$ and $\underline{\dim}(\mathcal{T})$ we use some cover of \mathcal{T} which relies on $\log l = \mathbb{E}[\mu[0, l]]^{o(1)}$. Then we derive the lower bound on $\dim_P(\mathcal{T})$ and $\dim_H(\mathcal{T})$ from Lemma 2.2.2.

2.4 Preliminaries

This section should be seen as a tool box: we gather here a collection of lemmas that will be used repeatedly throughout the paper. Most of them are straightforward.

2.4.1 Fundamental properties of μ

Lemma 2.4.1. *The map $l \rightarrow \mathbb{E}[\mu[0, l]]$ is differentiable and its derivative decreases to θ_0^2 as $l \rightarrow \infty$ we thus have as $l \rightarrow \infty$:*

$$\mathbb{E}[\mu[0, l]] = \theta_0^2 l + o(l).$$

Proof. By Fubini's theorem,

$$\mathbb{E}[\mu[0, l] - \theta_0^2 l] = \mathbb{E}\left[\sum_{i=1}^{\infty} \theta_i \mathbf{1}_{X_i \leq l}\right] = \sum_{i=1}^{\infty} \theta_i \mathbb{P}(X_i \leq l) = \sum_{i=1}^{\infty} \theta_i (1 - e^{-\theta_i l}).$$

Thus,

$$\mathbb{E}[\mu[0, l]] = \theta_0^2 l + \sum_{i=1}^{\infty} \theta_i (1 - e^{-\theta_i l}). \quad (2.2)$$

Each term of the sum is positive and increasing so we can differentiate term by term:

$$\frac{d}{dl} \mathbb{E}[\mu[0, l]] = \theta_0^2 + \sum_{i=1}^{\infty} \theta_i^2 e^{-\theta_i l}.$$

Since $\sum_{i=1}^{\infty} \theta_i^2 < \infty$, by bounded convergence the last term decreases to θ_0^2 as $l \rightarrow \infty$. \square

Lemma 2.4.1 implies that the map $l \mapsto \mathbb{E}[\mu[0, l]]$ is strictly increasing, continuous, and diverges, so is invertible. Thus for every $l \in \mathbb{R}^+$, there is a well-defined real number \mathcal{X}_l with $\mathbb{E}[\mu[0, \mathcal{X}_l]] = l$.

Lemma 2.4.2. *We have almost surely*

$$\mu[0, l] \underset{l \rightarrow \infty}{\sim} \mathbb{E}[\mu[0, l]].$$

Proof. For every $l \in \mathbb{R}^+$ the variance of μ_l is given by:

$$\mathbb{V}[\mu[0, l]] = \mathbb{V}\left[\theta_0^2 l + \sum_{i=1}^{\infty} \theta_i \mathbf{1}_{X_i \leq l}\right] = \sum_{i=1}^{\infty} \mathbb{V}[\theta_i \mathbf{1}_{X_i \leq l}] \leq \sum_{i=1}^{\infty} \theta_i^2 \leq 1.$$

Therefore for every $n \in \mathbb{N}$,

$$\mathbb{P}(|\mu[0, \mathcal{X}_{n^2}] - \mathbb{E}[\mu[0, \mathcal{X}_{n^2}]]| > n) \leq \frac{1}{n^2}.$$

By definition of \mathcal{X}_n we deduce by the Borel–Cantelli lemma that for every n large enough

$$n^2 - n \leq \mu[0, \mathcal{X}_{n^2}] \leq n^2 + n.$$

We thus have almost surely

$$\mu[0, \mathcal{X}_{n^2}] \sim \mathbb{E}[\mu[0, \mathcal{X}_{n^2}]] = n^2.$$

This result is then extended to every $l \in \mathbb{R}^+$ by monotonicity of $l \mapsto \mu[0, l]$. \square

Note that Lemmas 2.4.1 and 2.4.2 implies that for every l large enough $\mu[0, l] \leq l$.

The following lemma should be seen as an estimate for the "density" and "jump" of $l \mapsto \mu[0, l]$.

Lemma 2.4.3. *Almost surely there exists $L_0 \in \mathbb{R}$ such that for every $l \geq L_0$ and $0 \leq \delta \leq l$,*

$$\mu[l, l + \delta] \leq 2\delta \frac{\mathbb{E}[\mu[0, l]]}{l} + \frac{13 \log(l)}{l}.$$

Proof. First let us prove a concentration inequality for $\mu[l, l + \delta]$. We have by Fubini's Theorem,

$$\mathbb{E} \left[e^{\frac{l}{2} \mu[l, l + \delta]} \right] = \mathbb{E} \left[e^{\frac{l}{2} \theta_0^2 \delta} \prod_{i=1}^{\infty} e^{\frac{l}{2} \theta_i \mathbf{1}_{l \leq X_i \leq l + \delta}} \right] = e^{\frac{l}{2} \theta_0^2 \delta} \prod_{i=1}^{\infty} \left(1 + (e^{\frac{l}{2} \theta_i} - 1) \mathbb{P}(l \leq X_i \leq l + \delta) \right). \quad (2.3)$$

Furthermore we have for every $i \in \mathbb{N}$, since X_i is an exponential random variable of parameter θ_i ,

$$(e^{\frac{l}{2} \theta_i} - 1) \mathbb{P}(l \leq X_i \leq l + \delta) = (1 - e^{-\frac{l}{2} \theta_i}) \mathbb{P}\left(\frac{l}{2} \leq X_i \leq \frac{l}{2} + \delta\right) \leq \frac{l}{2} \theta_i \mathbb{P}\left(\frac{l}{2} \leq X_i \leq \frac{l}{2} + \delta\right),$$

Therefore by (2.3) and (2.2),

$$\begin{aligned} \mathbb{E} \left[e^{\frac{l}{2} \mu[l, l + \delta]} \right] &\leq \exp \left(\frac{l}{2} \theta_0^2 \delta + \sum_{i=1}^{\infty} \frac{l}{2} \theta_i \mathbb{P}\left(\frac{l}{2} \leq X_i \leq \frac{l}{2} + \delta\right) \right) \\ &= \exp \left(\frac{l}{2} \mathbb{E} \left[\mu \left[\frac{l}{2}, \frac{l}{2} + \delta \right] \right] \right). \end{aligned} \quad (2.4)$$

Moreover by Lemma 2.4.1, $t \mapsto \mathbb{E}[\mu[0, t]]$ is concave and increasing, hence,

$$\mathbb{E} \left[\mu \left[\frac{l}{2}, \frac{l}{2} + \delta \right] \right] \leq \frac{2\delta}{l} \mathbb{E} \left[\mu \left[0, \frac{l}{2} \right] \right] \leq \frac{2\delta}{l} \mathbb{E}[\mu[0, l]]. \quad (2.5)$$

Finally it follows from Markov's inequality, (2.4), and (2.5) that for every $l, l', t \in \mathbb{R}^+$,

$$\mathbb{P} \left(\mu[l, l + \delta] \geq \frac{2\delta}{l} \mathbb{E}[\mu[0, l]] + \frac{2t}{l} \right) \leq e^{-t}. \quad (2.6)$$

We now derive the desired result from (2.6). First by the Borel–Cantelli Lemma, there exists almost surely an $N \in \mathbb{N}$, such that for every $n \geq N$ and $n \leq m \leq 8n$,

$$\mu[\sqrt{n}, \sqrt{m}] \leq 2 \frac{(\sqrt{m} - \sqrt{n})}{\sqrt{n}} \mathbb{E}[\mu[0, \sqrt{n}]] + \frac{6 \log(n)}{\sqrt{n}}.$$

Now fix $l \geq N + 10$, $0 \leq \delta \leq l$ then let $n := \max\{i, \sqrt{i} \leq l\}$ and let $m := \min\{i, l + \delta \leq \sqrt{i}\}$. Since $l \mapsto \mu[0, l]$ is non decreasing, we have,

$$\begin{aligned} \mu[l, l + \delta] &\leq \mu[\sqrt{n}, \sqrt{m}] \leq 2 \frac{\sqrt{m} - \sqrt{n}}{\sqrt{n}} \mathbb{E}[\mu[0, \sqrt{n}]] + 12 \frac{\log(\sqrt{n})}{\sqrt{n}} \\ &\leq 2 \frac{\delta + 2/l}{l - 1/l} \mathbb{E}[\mu[0, l]] + 12 \frac{\log(l)}{l}. \end{aligned} \quad (2.7)$$

Finally by Lemma 2.4.1, $\mathbb{E}[\mu[0, l]] = O(l)$ as $l \rightarrow \infty$ and the desired result follows from (2.7). \square

2.4.2 Key results on cuts and sticks

For every random variables A, B on \mathbb{R} we recall that A is stochastically dominated by B , and write $A \leq_{\text{st}} B$ iff for every $t \in \mathbb{R}^+$, $\mathbb{P}(A \geq t) \leq \mathbb{P}(B \geq t)$.

Lemma 2.4.4. *Almost surely there exists $L_0 \in \mathbb{R}^+$ such that for every $l \geq L_0$ there are at most $2l\mu[0, l] \leq 2l^2$ cuts on $[0, l]$.*

Proof. Conditionally on μ , $\{Y_i\}_{i \in \mathbb{N}}$ is a Poisson point process with rate $\mu[0, l]dl$ so the number of cuts in $[0, l]$ is stochastically dominated by a Poisson random variable α with mean $l\mu[0, l]$ and for l large enough

$$\mathbb{P}\left(\alpha \geq \frac{3}{2}l\mu[0, l]\right) \leq \frac{1}{l^2}.$$

Thus by the Borel–Cantelli lemma almost surely for every $l \in \mathbb{N}$ large enough, there are at most $\frac{3}{2}l\mu[0, l]$ cuts on $[0, l]$. This can be easily extended to all $l \in \mathbb{R}^+$ large enough using Lemmas 2.4.1 and 2.4.2. We omit the straightforward details. \square

Lemma 2.4.5. *Almost surely there exists $i_0 \in \mathbb{N}$ such that for every $i \geq i_0$:*

$$l_{i+1} \leq \frac{5 \log(Y_i)}{M_i}.$$

Proof. Because the cuts are made at rate $\mu[0, l]dl$, for every $i \in \mathbb{N}$, $(Y_{i+1} - Y_i)\mu[0, Y_i]$ is stochastically dominated by an exponential random variable with mean one. Therefore

$$\mathbb{P}((Y_{i+1} - Y_i)\mu[0, Y_i] \geq 2 \log(i)) \leq 1/i^2.$$

So by the Borel–Cantelli lemma and Lemma 2.4.4, for every i large enough,

$$Y_{i+1} - Y_i \leq \frac{2 \log(i)}{\mu[0, Y_i]} \leq \frac{2 \log(2Y_i^2)}{\mu[0, Y_i]} \leq \frac{5 \log(Y_i)}{\mu[0, Y_i]}. \quad \square$$

Lemma 2.4.6. *Almost surely there exists $L_0 \in \mathbb{R}^+$ such that for all $l \geq L_0$ and $i \in \mathbb{N}$ with $Y_i \geq l$,*

$$m_{i+1} \leq \frac{\log^2 l}{l}.$$

Proof. We have by Lemmas 2.4.5, 2.4.2, and 2.4.3, as $i \rightarrow \infty$,

$$m_{i+1} \leq \mu\left[Y_i, Y_i + \frac{5 \log Y_i}{\mu[0, Y_i]}\right] \leq O\left(\frac{\log Y_i}{\mu[0, Y_i]} \frac{\mu[0, Y_i]}{Y_i} + \frac{\log Y_i}{Y_i}\right) = o\left(\frac{\log^2 Y_i}{Y_i}\right). \quad \square$$

2.4.3 An estimate of distances in \mathcal{T}

For every $l \in \mathbb{R}^+$, let $\text{Exp}(l)$ denote an exponential random variable of mean l .

Lemma 2.4.7. *For every $x, y \in \mathbb{R}^+$, $t \geq 6$, conditionally on μ , $\mathbb{P}(d(\mathcal{T}_x, y) > 4t/\mu[0, x]) \leq e^{-t}$.*

Remark. Proving an equivalent of Lemma 2.4.7 is crucial for each studies on stick-breaking constructions, notably for compactness [43, 77] and convergence [10], Chapter 3. Although the proof below uses strong properties on μ , more general methods can be found in [10, 43, 77], Chapter 3. Finally, we believe that such methods can be useful to study several other classes of algorithms.

Proof. To simplify the notation let for every $l \in \mathbb{R}^+$, $\mathcal{F}_l := \sigma(\mu, \{(Y_i, Z_i)\}_{i \in \mathbb{N}} \cap [l, +\infty) \times \mathbb{R}^+)$. We first prove that if $\mu[0, x] \geq 2\mu[0, y]$ then conditionally on \mathcal{F}_y , $d(\mathcal{T}_x, y) \leq_{\text{st}} \text{Exp}(\frac{2}{\mu[0, x]})$. If $y \leq x$ then $d(\mathcal{T}_x, y) = 0$. We assume henceforth that it is not the case. Let us "follow" the geodesic path from y to \mathcal{T}_x . More precisely we define the next sequence by induction (see Figure 2.2). Let

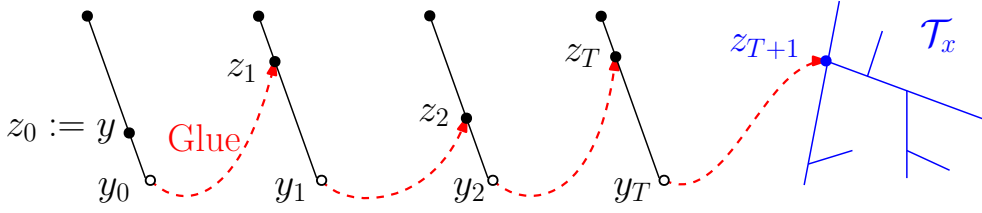


Figure 2.2: A typical construction of $(y_i, z_i)_{i \in \mathbb{N}}$. Note that in general we do not know if $y_T \in \mathcal{T}_x$.

$z_0 := y$, then for every $i \geq 0$, let $k_i := \max\{k \in \mathbb{N} : Y_k < z_i\}$ and let $y_i := Y_{k_i}$, and $z_{i+1} := Z_{k_i}$. Additionally let T denote the smallest integer such that $z_{T+1} \leq x$. Note that

$$d(y, \mathcal{T}_x) = \sum_{i=0}^T (z_i - \max(y_i, x)). \quad (2.8)$$

Now recall that conditionally on μ , $\{(Y_i, Z_i), i \in \mathbb{N}\}$ is a Poisson point process, so $\{(y_i, z_i), \mathcal{F}_{y_i}\}_{i \geq 0}$ is a Markov chain. Also note that $T+1 = \inf\{n : z_n < x\}$ is a stopping time for $\{(y_i, z_i), \mathcal{F}_{y_i}\}_{i \geq 0}$. Moreover, for every $i \in \mathbb{N}$ conditionally on (μ, y_i, z_i) , z_{i+1} has law p_{y_i} . Hence if $y_i \geq x$,

$$\mathbb{P}(z_{i+1} \leq x | \mu, y_i, z_i) = p_{y_i}[0, x] \geq \frac{\mu[0, x]}{\mu[0, y]} \geq \frac{1}{2}.$$

So T is stochastically dominated by a geometric random variable of parameter $1/2$. Furthermore, if $i \leq T$, conditionally on (μ, y_i) , $\{Y_i\}_{i \in \mathbb{N} \cap [x, y_i]}$ is a Poisson point process of rate $\mu[0, l] \geq \mu[0, x]$ so $z_i - \max(y_i, x) \leq_{\text{st}} \text{Exp}(\frac{1}{\mu[0, x]})$. Finally it follows from (2.8) that $d(\mathcal{T}_x, y) \leq_{\text{st}} \text{Exp}(\frac{2}{\mu[0, x]})$.

Let us now treat the general case. As previously, we bound $d(\mathcal{T}_x, y)$ by following the geodesic path between \mathcal{T}_x and y . More precisely, let for every $i \geq 0$, $x_i := \inf\{a \in \mathbb{R}^+, \mu[0, a] \geq 2^i \mu[0, x]\}$ and let y_i be the nearest point from y on $[0, x_i]$. Note that

$$d(x, y) = \sum_{i=0}^{+\infty} d(y_i, y_{i+1}). \quad (2.9)$$

Then for every $i \geq 0$, since $2\mu[0, x_i] \geq \mu[0, x_{i+1}]$, the first case yields, conditionally on $\mathcal{F}_{y_{i+1}}$,

$$d(y_i, y_{i+1}) = d(\mathcal{T}_{x_i}, y_{i+1}) \leq_{\text{st}} \text{Exp}\left(\frac{2}{\mu[0, x_i]}\right) \leq_{\text{st}} \text{Exp}\left(\frac{2^{1-i}}{\mu[0, x]}\right). \quad (2.10)$$

Finally since for every $j > i$, $d(y_j, y_{j+1})$ is $\mathcal{F}_{y_{i+1}}$ measurable, the desired result follows from (2.9), (2.10) and Lemma 2.8.3. \square

2.5 The mass measure

First we prove Lemma 2.5.1 that describes precisely the evolution of the mass μ as we add branches to the tree. Then we prove that $(p_l)_{l \geq 0}$ is tight and use Lemma 2.5.1 to prove that for every bounded Lipschitz function $(p_l(f))_{l \geq 0}$ converges. It proves, by the Portmanteau Theorem, that $(p_l)_{l \geq 0}$ converges weakly toward a probability measure p (Theorem 2.3.1). Then we adapt the argument to prove Proposition 2.3.2.

2.5.1 The mass conservation lemma

Definition. For every $l \in \mathbb{R}^+$ let the projection of x in \mathcal{T}_l be the nearest point from x in \mathcal{T}_l . Also for every $S \subset \mathcal{T}$, let $S^{\uparrow l}$ be the set of $x \in \mathcal{T}$ such that the projection of x in \mathcal{T}_l is in S .

Lemma 2.5.1. Almost surely $(\mu, (Y_i)_{i \in \mathbb{N}})$ satisfy the following property. For every a large enough, conditionally on \mathcal{T}_{Y_a} , for every measurable set $S \subset \mathcal{T}_{Y_a}$, the following assertions hold.

- (i) Almost surely $\{p_l(S^{\uparrow Y_a})\}_{l \in \mathbb{R}^+}$ converges toward a real number $p(S^{\uparrow Y_a})$.
- (ii) If $\mu(S) \geq \frac{\log^6 Y_a}{Y_a}$ with probability at least $1 - \frac{1}{Y_a^5}$, for every $l \geq Y_a$

$$\left(1 - \frac{1}{\log Y_a}\right) p_{Y_a}(S) \leq p_l(S^{\uparrow Y_a}) \leq \left(1 + \frac{1}{\log Y_a}\right) p_{Y_a}(S).$$

- (iii) If $\mu(S) \leq \frac{\log^6 Y_a}{Y_a}$ with probability at least $1 - \frac{1}{Y_a^5}$, for every $l \geq Y_a$

$$p_l(S^{\uparrow Y_a}) \leq \frac{(\log Y_a)^6}{Y_a M_a}.$$

Proof. First for every $i \geq a$, let $A_i := \mu_{Y_i}(S^{\uparrow Y_a})$ and $\mathcal{F}_i := \sigma(\mu, \{Y_n\}_{n \in \mathbb{N}}, \{Z_n\}_{1 \leq n < i})$. Note that for every $i \geq a$, since Z_i has law $\frac{\mu_{Y_i}}{M_i}$, we have $(Y_i, Y_{i+1}] \subset S^{\uparrow Y_a}$ with probability $\frac{A_i}{M_i}$ so

$$\mathbb{P}(A_{i+1} = A_i + m_{i+1}) = \frac{A_i}{M_i} \quad ; \quad \mathbb{P}(A_{i+1} = A_i) = \frac{M_i - A_i}{M_i}.$$

Thus $(A_i, \mathcal{F}_i)_{i \geq a}$ can be seen as a Pólya urn in the sense of Lemma 2.8.1. Furthermore by Lemma 2.4.6, we have almost surely for every a large enough, $\max_{n > a} m_n \leq \frac{\log^2 Y_a}{Y_a}$, hence by Lemma 2.8.1 (b), for every $t \in [0, 1]$,

$$\mathbb{P}\left(\sup_{i \geq a} \left| \frac{A_i}{M_i} - \frac{A_a}{M_a} \right| > t \frac{A_a}{M_a} \mid A_a\right) \leq 2 \exp\left(-\frac{t^2}{8} \frac{A_a Y_a}{\log^2 Y_a}\right). \quad (2.11)$$

Also still by Lemma 2.4.6 we have for every $a \in \mathbb{N}$ large enough, $i \geq a$, and $Y_i \leq l \leq Y_{i+1}$,

$$p_l(S^{\uparrow Y_a}) = \frac{\mu_l(S^{\uparrow Y_a})}{\mu[0, l]} \leq \frac{\mu_{Y_i}(S^{\uparrow Y_a}) + m_{i+1}}{\mu[0, Y_i]} = \frac{A_i}{M_i} + \frac{m_{i+1}}{M_i} \leq \frac{A_i}{M_i} + \frac{\log^2 Y_a}{Y_a M_a},$$

and similarly

$$p_l(S^{\uparrow Y_a}) = \frac{\mu_l(S^{\uparrow Y_a})}{\mu[0, l]} \geq \frac{\mu_{Y_{i+1}}(S^{\uparrow Y_a}) - m_{i+1}}{\mu[0, Y_{i+1}]} = \frac{A_{i+1}}{M_{i+1}} - \frac{m_{i+1}}{M_{i+1}} \geq \frac{A_{i+1}}{M_{i+1}} - \frac{\log^2 Y_a}{Y_a M_a}.$$

Therefore,

$$\sup_{l \geq Y_a} \left| p_l(S^{\uparrow Y_a}) - \frac{A_a}{M_a} \right| \leq \sup_{i \geq a} \left| \frac{A_i}{M_i} - \frac{A_a}{M_a} \right| + \frac{\log^2 Y_a}{Y_a M_a}. \quad (2.12)$$

The claims in (i) (ii) (iii) are applications of the inequalities in (2.11) and (2.12).

Consider first (i). Note that (2.11) implies that $\{\frac{A_i}{M_i}\}_{i \in \mathbb{N}}$ is almost surely Cauchy, and hence converges. Furthermore $\{\frac{\log^2 Y_a}{Y_a M_a}\}_{a \in \mathbb{N}}$ almost surely converges to 0. (i) then follows from (2.12).

Towards (ii), we have by assumption $A_a \geq \frac{\log^6 Y_a}{Y_a}$ so if $a \geq 10$, $\frac{\log^2(Y_a)}{Y_a M_a} \leq \frac{1}{2 \log Y_a} \frac{A_a}{M_a}$. Therefore by (2.12) it suffices to estimate the right-hand side of (2.11) with $t = \frac{1}{2 \log Y_a}$:

$$2 \exp\left(-\frac{t^2}{8} \frac{A_a Y_a}{\log^2 Y_a}\right) \leq 2 \exp\left(-\frac{1}{32 \log^2 Y_a} \frac{\log^6 Y_a}{\log^2 Y_a}\right) = o\left(\frac{1}{Y_a^5}\right)$$

and (ii) follows. (iii) can be treated similarly using (2.11) with $t = \frac{\log^6 Y_a}{2 Y_a A_a}$. We leave the details to the reader. This concludes the proof. \square

2.5.2 Weak convergence of μ_n : proof of Theorem 2.3.1

In this section we prove Theorem 2.3.1. Let us start with the tightness of $(p_l)_{l \in \mathbb{R}^+}$ which follows from the following lemma.

Lemma 2.5.2. *For every $n \in \mathbb{N}$ let A_n be the set of $x \in \mathcal{T}$ such that, $d(x, [0, \mathcal{X}_{2^n}]) \leq 8n/2^n$ and $B_n := \bigcap_{m \geq n} A_m$. The following assertions hold:*

- (i) *Almost surely for every n large enough, for every $l \geq 0$, $p_l(B_n) \geq 1 - 2^{-2n}$.*
- (ii) *For every n large enough B_n is compact.*

Proof. First for every $n \leq m \in \mathbb{N}$ large enough, conditionally on μ we have by Fubini's theorem, Lemma 2.4.7, and Lemma 2.4.2,

$$\mathbb{E}[p_{\mathcal{X}_{2^m}}(\mathcal{T} \setminus A_n) | \mu] = \int_0^{\mathcal{X}_{2^m}} \mathbb{P}(x \notin A_n) \frac{d\mu(x)}{\mu[0, \mathcal{X}_{2^m}]} \leq e^{-\frac{8n}{2^n} \frac{\mu[0, \mathcal{X}_{2^n}]}{4}} = e^{-2n(1+o(1))}.$$

It directly follows by Markov's inequality and the Borel–Cantelli lemma that almost surely for every n large enough and $m \geq n$,

$$p_{\mathcal{X}_{2^m}}(\mathcal{T} \setminus A_n) \leq 2^{-2n-3}.$$

Therefore for every $n \in \mathbb{N}$ and $l \geq \mathcal{X}_{2^n}$, writing k for the smallest integer such that $l \leq \mathcal{X}_{2^k}$ we have by Lemma 2.4.2,

$$p_l(\mathcal{T} \setminus A_n) \leq \frac{\mu[0, \mathcal{X}_{2^k}]}{\mu[0, l]} p_{\mathcal{X}_{2^k}}(\mathcal{T} \setminus A_n) \leq \frac{\mu[0, \mathcal{X}_{2^k}]}{\mu[0, \mathcal{X}_{2^{k-1}}]} p_{\mathcal{X}_{2^k}}(\mathcal{T} \setminus A_n) \leq 2^{-2n-1}.$$

Note that the latter is also true for $l \leq \mathcal{X}_{2^n}$ since in this case $\mathcal{T}_l \subset \mathcal{T}_{\mathcal{X}_{2^n}} \subset A_n$. (i) then follows from a union bound on n .

Toward (ii), note that A_m is a closed set for $m \geq n$, so B_n is a closed set as well. Therefore it suffices to show that any sequence $(x_i)_{i \in \mathbb{N}}$ in B_n has an accumulation point. Fix $(x_i)_{i \in \mathbb{N}}$ then for every $m \in \mathbb{N}$ let x_i^m be the projection of x_i on $[0, \mathcal{X}_{2^m}]$. Since for every $m \in \mathbb{N}$, $\mathcal{T}_{\mathcal{X}_{2^m}}$ is compact, by a diagonal extraction procedure there exists an increasing function $\phi : \mathbb{N} \mapsto \mathbb{N}$ such that for every $m \in \mathbb{N}$, $(x_{\phi(i)}^m)_{i \in \mathbb{N}}$ converges. Hence, for every $m \geq n$ there exists $N \in \mathbb{N}$ such that for every $a, b \geq N$, $d(x_{\phi(a)}^m, x_{\phi(b)}^m) \leq 1/m$ and so

$$d(x_{\phi(a)}, x_{\phi(b)}) \leq d(x_{\phi(a)}, x_{\phi(a)}^m) + d(x_{\phi(a)}^m, x_{\phi(b)}^m) + d(x_{\phi(b)}^m, x_{\phi(b)}) \leq \frac{8m}{2^m} + \frac{1}{m} + \frac{8m}{2^m}.$$

Therefore $(x_{\phi(i)})_{i \in \mathbb{N}}$ is Cauchy and thus converges since \mathcal{T} is complete by definition. Since $(x_i)_{i \in \mathbb{N}}$ is arbitrary, B_n is compact. \square

Definition. Let \mathbb{F} be the set of positive, 1-Lipschitz functions that are bounded by 1 on \mathcal{T} . For every finite measure ν on \mathcal{T} and measurable function $f : \mathcal{T} \rightarrow \mathbb{R}$ let $\nu(f) := \int_{\mathcal{T}} f(x) d\nu(x)$.

Lemma 2.5.3. *Almost surely, for every $f \in \mathbb{F}$, $p_l(f)$ converges as $l \rightarrow \infty$.*

Proof. First for every $a \in \mathbb{N}$ let $\{I_i^a\}_{1 \leq i \leq N_a}$ be a partition of $\mathcal{T}_{Y_a} = ([0, Y_a], d)$ into intervals of diameter at most $1/a$. Then for every $a \in \mathbb{N}$ and $1 \leq i \leq N_a$ let $J_i^a := (I_i^a)^{\uparrow Y_a}$ and let $x_i^a \in I_i^a$. Note that for every $a \in \mathbb{N}$, $\{J_i^a\}_{1 \leq i \leq N_a}$ is a partition of \mathcal{T} . So for every $l \geq Y_a$ and $f \in \mathbb{F}$,

$$p_l(f) = \sum_{i=1}^{N_a} p_l(\mathbf{1}_{J_i^a} f) = \sum_{i=1}^{N_a} p_l(J_i^a f(x_i^a)) + \sum_{i=1}^{N_a} p_l(\mathbf{1}_{J_i^a} (f - f(x_i^a))) . \quad (2.13)$$

By Lemma 2.5.1 (i), almost surely for every $f \in \mathbb{F}$ the first sum converges toward $\sum_{i=1}^{N_a} p(J_i^a) f(x_i^a)$ as l goes to infinity. Let us bound the second sum in order to prove that $(p_l(f))_{l \in \mathbb{R}^+}$ is Cauchy. For every $a \in \mathbb{N}$ let k_a be the largest integer such that $\mathcal{X}_{2^{k_a}} \leq Y_a$. We have for every $f \in \mathbb{F}$:

$$\sum_{i=1}^{N_a} p_l(\mathbf{1}_{J_i^a}(f - f(x_i^a))) = \sum_{i=1}^{N_a} p_l(\mathbf{1}_{J_i^a \cap B_{k_a}}(f - f(x_i^a))) + p_l(\mathbf{1}_{\mathcal{T} \setminus B_{k_a}}(f - f(x_i^a)))$$

and

$$\left| \sum_{i=1}^{N_a} p_l(\mathbf{1}_{J_i^a}(f - f(x_i^a))) \right| \leq \sum_{i=1}^{N_a} p_l(\mathbf{1}_{J_i^a \cap B_{k_a}} |f - f(x_i^a)|) + p_l(\mathcal{T} \setminus B_{k_a}).$$

Furthermore for every $a \in \mathbb{N}$ and $1 \leq i \leq N_a$, recall that by definition J_i^a has diameter at most $\frac{1}{a}$ and that $d_H([0, \mathcal{X}_{2^{k_a}}], B_{k_a}) \leq 8k_a 2^{-k_a}$. Therefore $J_i^a \cap B_{k_a} = (J_i^a)^{\uparrow Y_a} \cap B_{k_a}$ has diameter at most $\delta_a := \frac{1}{a} + \frac{16k_a}{2^{k_a}}$. Hence for every $f \in \mathbb{F}$,

$$\left| \sum_{i=1}^{N_a} p_l(\mathbf{1}_{J_i^a}(f - f(x_i^a))) \right| \leq \sum_{i=1}^{N_a} p_l(J_i^a \cap B_{k_a}) \delta_a + p_l(\mathcal{T} \setminus B_{k_a}) \leq \delta_a + p_l(\mathcal{T} \setminus B_{k_a}).$$

Moreover by Lemma 2.5.2 for every a large enough $p_l(\mathcal{T} \setminus B_{k_a}) \leq 2^{-2k_a}$. Finally for every $f \in \mathbb{F}$,

$$\limsup_{a \rightarrow \infty} \limsup_{l \rightarrow \infty} \left| \sum_{i=1}^{N_a} p_l(\mathbf{1}_{J_i^a}(f - f(x_i^a))) \right| = 0, \quad (2.14)$$

which implies together with (2.13) that $(p_l(f))_{l \in \mathbb{R}^+}$ is Cauchy and thus converges. \square

Proof of Theorem 2.3.1. First by lemma 2.5.2, $(p_l)_{l \in \mathbb{R}^+}$ is tight. The convergence of $(p_l)_{l \in \mathbb{R}^+}$ then directly follows from Lemma 2.5.3 and the Portmanteau theorem.

Towards proving that p has full support, we first prove that μ has almost surely full support. Note that it suffices to prove that for every $a < b \in \mathbb{R}^+$, almost surely $\mu[a, b] > 0$. If $\theta_0 > 0$ then $\mu[a, b] \geq (b - a)\theta_0^2 > 0$. So we assume henceforth that $\theta_0 = 0$. Note that in this case, $\sum_{i=1}^{\infty} \theta_i = \infty$. Moreover, recall that $\{X_i\}_{i \in \mathbb{N}}$ is a family of independent exponential random variables of parameter $\{\theta_i\}_{i \in \mathbb{N}}$ so that,

$$\sum_{i=1}^{\infty} \mathbb{P}(X_i \in [a, b]) = \sum_{i=1}^{\infty} e^{-\theta_i a} (1 - e^{-\theta_i(b-a)}) = \infty.$$

Therefore by the Borel–Cantelli lemma, for every $a, b \in \mathbb{R}^+$ almost surely there exists an $i \in \mathbb{N}$ such that $X_i \in [a, b]$ and so $\mu[a, b] \geq \theta_i > 0$. Thus, μ has almost surely full support.

Next we prove that p also has full support. Fix $x \in \mathbb{R}^+$ and $\varepsilon > 0$. Additionally for every $a \in \mathbb{N}$ let k_a be the largest integer such that $\mathcal{X}_{2^{k_a}} \leq Y_a$. Note that for every $a \in \mathbb{N}$ large enough, by definition of B_{k_a} , $B(x, \varepsilon)^{\uparrow Y_a} \cap B_{k_a}$ has diameter at most $\varepsilon + 16k_a 2^{-k_a} \leq 2\varepsilon$. It follows that,

$$p(B(x, 2\varepsilon)) \geq p(B(x, \varepsilon)^{\uparrow Y_a} \cap B_{k_a}) \geq p(B(x, \varepsilon)^{\uparrow Y_a}) - p(\mathcal{T} \setminus B_{k_a}). \quad (2.15)$$

On the one hand, recall that almost surely $\mu(B(x, \varepsilon)) > 0$. Thus by Lemma 2.5.1 (ii), for every a large enough, with probability at least $1 - 1/Y_a^5$,

$$p(B(x, \varepsilon)^{\uparrow Y_a}) \geq \frac{1}{2} p_{Y_a}(B(x, \varepsilon)) = \frac{\mu_{Y_a}(B(x, \varepsilon))}{2M_a}.$$

On the other hand, by Lemmas 2.5.2 (i), 2.4.2, and the definition of k_a , for every a large enough,

$$p(\mathcal{T} \setminus B_{k_a}) \leq 2^{-2k_a} \leq 2\mu[0, \mathcal{X}_{2k_a}]^{-2} \leq 2\mu[0, Y_a]^{-2} = o(1/M_a).$$

Therefore by (2.15), almost surely $p(B(x, 2\varepsilon)) > 0$. Since x, ε were arbitrary and since rational numbers are dense on \mathcal{T} , it follows that p has full support.

Finally, we prove that almost surely p gives measure 1 to the set of leaves and is non-atomic. For every $\varepsilon > 0$ and $S \subset \mathcal{T}$, let $B(S, \varepsilon) = \{x \in \mathcal{T} : d(x, S) < \varepsilon\}$. Then let $(\varepsilon_a)_{a \in \mathbb{N}}$ be a sequence of positive real numbers decreasing sufficiently fast so that for every $a > 0$ and $0 \leq i < a$ we have $\mu_{Y_a}(B((Y_i, Y_{i+1}], \varepsilon_a)) \leq 2\mu_{Y_a}(Y_a, Y_{a+1}]$. By Lemma 2.5.1 (ii) (iii), for every a large enough and $0 \leq i < a$, with probability at least $1 - 1/Y_a^5$, for every $l \geq Y_a$,

$$p_l \left(B((Y_i, Y_{i+1}], \varepsilon_a)^{\uparrow Y_a} \right) \leq \max \left\{ 2p_{Y_a} \left(B((Y_i, Y_{i+1}], \varepsilon_a) \right); \frac{(\log Y_a)^6}{Y_a M_a} \right\}. \quad (2.16)$$

Since by Lemma 2.4.4 $a = O(Y_a^2)$, the Borel–Cantelli lemma implies that almost surely (2.16) is true for every a large enough, $0 \leq i < a$ and $l \geq Y_a$. Furthermore by Lemma 2.4.6 $M := \max_{i \in \mathbb{N}} \mu(Y_i, Y_{i+1}] < \infty$. Also note that $\frac{(\log Y_a)^6}{Y_a} \rightarrow 0$ as $a \rightarrow +\infty$. Therefore for every a large enough, $0 \leq i < a$, and $l \geq a$:

$$p_l \left(B((Y_i, Y_{i+1}], \varepsilon_a)^{\uparrow Y_a} \right) \leq \frac{4M}{M_a}. \quad (2.17)$$

Moreover since for every $a \in \mathbb{N}$ the projection on \mathcal{T}_{Y_a} (see 2.5.1 for definition) is a continuous function, for every $0 \leq i < a$, $B((Y_i, Y_{i+1}], \varepsilon_a)^{\uparrow Y_a}$ is open. Thus by letting $l \rightarrow \infty$ in (2.17), the Portmanteau theorem yields for a large enough:

$$p \left(B((Y_i, Y_{i+1}], \varepsilon_a)^{\uparrow Y_a} \right) \leq \frac{4M}{M_a}, \quad (2.18)$$

which tends to 0 as $a \rightarrow \infty$. So for every $i \in \mathbb{N}$, $p[Y_i, Y_{i+1}] = 0$. Summing over all $i \in \mathbb{N}$ we get $p(\mathbb{R}^+) = 0$ and so p gives measure 1 to the set of leaves. Note that (2.18) also yield for every $a \in \mathbb{N}$,

$$\sup_{x \in \mathcal{T}} p\{x\} = \max_{0 \leq i < a} \sup_{x \in [Y_i, Y_{i+1}]^{\uparrow Y_a}} p\{x\} \leq \frac{4M}{M_a},$$

which implies, taking $a \rightarrow \infty$, that p is non-atomic. □

2.5.3 Other convergences toward p : proof of Proposition 2.3.2

In this section we prove Proposition 2.3.2, and the following stronger result.

Lemma 2.5.4. *Let μ^α be a random Borel measure on \mathbb{R}^+ which is $\sigma(\mu, \{Y_i\}_{i \in \mathbb{N}})$ measurable. Let for every $l \in \mathbb{R}^+$, μ_l^α be the restriction of μ^α to $\mathcal{T}_l = ([0, l], d)$ and $p_l^\alpha := \frac{\mu_l^\alpha}{\mu^\alpha[0, l]}$. Suppose that almost surely the following assertions hold:*

- (i) *For every $l > 0$ $\mu^\alpha[0, l] < \infty$, and $\mu^\alpha(\mathbb{R}^+) = +\infty$.*
- (ii) *There exists $\varepsilon > 0$ such that $\mu^\alpha(Y_{i-1}, Y_i] = o(\mu^\alpha[0, Y_i]^{1-\varepsilon})$.*
- (iii) *For all $\varepsilon > 0$, $\sum_{i=1}^n \mu^\alpha(Y_{i-1}, Y_i] \mathbf{1}_{Y_i - Y_{i-1} > \varepsilon} = o(\mu^\alpha[0, Y_n])$.*

Then almost surely $\{p_l^\alpha\}_{l \in \mathbb{R}^+}$ converges weakly toward p .

To prove Lemma 2.5.4, we first show the following strong law of large number.

Lemma 2.5.5. *Let μ^α be such as in Lemma 2.5.4 and $S \subset \mathcal{T}$ be a random measurable set such that for every n large enough, $p_{Y_n}(S)$ is $\sigma(\mu, \{Y_i\}_{i \in \mathbb{N}}, \{Z_i\}_{1 \leq i < n})$ measurable. Almost surely,*

$$\limsup_{n \rightarrow \infty} \sum_{i=1}^n p^\alpha(Y_{i-1}, Y_i) \mathbf{1}_{Z_{i-1} \in S} \leq \limsup_{l \rightarrow \infty} p_l(S).$$

Proof. Let $\{U_n\}_{n \in \mathbb{N}}$ be a family of independent uniform random variables on $[0, 1]$. Since for every $n \in \mathbb{N}$, conditionally on $(\mu, \{Y_j\}_{j \in \mathbb{N}}, \{Z_j\}_{1 \leq j < n})$, Z_n has law p_{Y_n} , we may couple $\{U_n\}_{n \in \mathbb{N}}$ and $\{Z_n\}_{n \in \mathbb{N}}$ such that for every n large enough, $Z_n \in S$ if and only if $U_n \leq p_{Y_n}(S)$. Therefore, by Lemma 2.8.2 and assumptions (i) and (ii), almost surely for every $t > \limsup_{l \rightarrow \infty} p_l(S)$,

$$\limsup_{n \rightarrow \infty} \frac{\sum_{i=1}^n \mu^\alpha(Y_{i-1}, Y_i) \mathbf{1}_{Z_{i-1} \in S}}{\mu^\alpha[0, Y_n]} \leq \limsup_{n \rightarrow \infty} \frac{\sum_{i=1}^n \mu^\alpha(Y_{i-1}, Y_i) \mathbf{1}_{U_{i-1} \leq t}}{\mu^\alpha[0, Y_n]} = t. \quad (2.19)$$

Taking $t \rightarrow \limsup_{l \rightarrow \infty} p_l(S)$ in (2.19) yields the desired inequality. \square

Proof of Lemma 2.5.4. First by the Portmanteau's theorem it suffices to prove that for every $f \in \mathbb{F}$, $p_l^\alpha(f) \rightarrow p(f)$ where \mathbb{F} is the set of positive, 1-Lipschitz functions that are bounded by 1 on \mathcal{T} . Moreover since we work with probability measures and since for every $f \in \mathbb{F}$, $(1 - f) \in \mathbb{F}$, it suffices to prove instead that for every $f \in \mathbb{F}$, $\limsup p_l^\alpha(f) \leq p(f)$. To this end, we proceed as in the proof of Lemma 2.5.3 and will hence use the same notations. In addition, for $\varepsilon > 0$ let $\Gamma_\varepsilon := \bigcup_{i, Y_{i+1} - Y_i \leq \varepsilon} (Y_i, Y_{i+1})$ and for every $x \in \mathcal{T}$, let $\zeta(x) := Z_{\max\{i, Y_i \leq x\}}$.

Now fix $\varepsilon > 0$ and recall from the proof of Lemma 2.5.3 that for every $a \in \mathbb{N}$, $\{J_i^a\}_{1 \leq i \leq N_a}$, is a partition of \mathcal{T} , so for every $f \in \mathbb{F}$ and $l \geq Y_a$,

$$p_l^\alpha(f) \leq \sum_{i=1}^{N_a} p_l^\alpha(f \mathbf{1}_{z(\cdot) \in J_i^a \cap B_{k_a}} \mathbf{1}_{\Gamma_\varepsilon}) + p_l^\alpha(f \mathbf{1}_{z(\cdot) \notin B_{k_a}}) + p_l^\alpha(f \mathbf{1}_{\mathcal{T} \setminus \Gamma_\varepsilon}). \quad (2.20)$$

We now upper bound each term of (2.20) separately. First, recall that $J_i^a \cap B_{k_a}$ have diameter at most δ_a , thus for every $1 \leq i \leq N_a$ and $s \in S_i^a := \{x, z(x) \in J_i^a \cap B_{k_a}\} \cap \Gamma_\varepsilon$ we have $d(s, x_i^a) \leq \varepsilon + \delta_a$. Therefore for every $f \in \mathbb{F}$,

$$\sum_{i=1}^{N_a} p_l^\alpha(f \mathbf{1}_{S_i^a}) \leq \sum_{i=1}^{N_a} p_l^\alpha((f(x_i^a) + \varepsilon + \delta_a) \mathbf{1}_{S_i^a}) \leq \sum_{i=1}^{N_a} f(x_i^a) p_l^\alpha(\mathbf{1}_{z(\cdot) \in J_i^a}) + \varepsilon + \delta_a. \quad (2.21)$$

Furthermore by Lemma 2.5.1 (i) almost surely for every $a \in \mathbb{N}$ and $1 \leq i \leq N_a$, $p_l(J_i^a) \rightarrow p(J_i^a)$ as $l \rightarrow \infty$, hence by Lemma 2.5.5 almost surely

$$\limsup_{l \rightarrow \infty} p_l^\alpha(\mathbf{1}_{z(\cdot) \in J_i^a}) \leq p(J_i^a).$$

Therefore since $\delta_a \rightarrow 0$ as $a \rightarrow \infty$, we have by (2.21) and (2.14) for every $f \in \mathbb{F}$,

$$\limsup_{a \rightarrow \infty} \limsup_{l \rightarrow \infty} \sum_{i=1}^{N_a} p_l^\alpha(f \mathbf{1}_{S_i^a}) \leq \limsup_{a \rightarrow \infty} \sum_{i=1}^{N_a} f(x_i^a) p(J_i^a) + \varepsilon \leq p(f) + \varepsilon.$$

Next we have by Lemma 2.5.2 (i) and Lemma 2.5.5, almost surely for every a large enough,

$$\limsup_{l \rightarrow \infty} p_l^\alpha(\mathbf{1}_{z(\cdot) \notin B_{k_a}}) \leq 2^{-2k_a}.$$

Futhermore by assumption (iii), $p_l^\alpha(\mathbf{1}_{\mathcal{T} \setminus \Gamma_\varepsilon}) \rightarrow 0$ as $l \rightarrow 0$. Finally (2.20) yields, for every $f \in \mathbb{F}$,

$$\limsup_{l \rightarrow \infty} p_l^\alpha(f) \leq \limsup_{a \rightarrow \infty} p(f) + \varepsilon + 2^{-2k_a} = p(f) + \varepsilon.$$

Taking $\varepsilon \rightarrow 0$ in the previous inequality concludes the proof. \square

Proof of Proposition 2.3.2. We now justify that $\mu^{\sim}, \mu^{\bullet}$ satisfy the assumptions of Lemma 2.5.4. First (i) and the $\sigma(\mu, \{Y_i\}_{i \in \mathbb{N}})$ measurability for $\mu^{\sim}, \mu^{\bullet}$ are straightforward from their definitions. (ii) for μ^{\sim} is an immediate consequence of Lemma 2.4.5. (ii) for μ^{\bullet} comes from $\mu^{\bullet}(Y_n, Y_{n+1}] = 1$. (iii) is a little tedious to prove and follows directly from the fact that conditionally on μ , $\{Y_i\}_{i \in \mathbb{N}}$ is a Poisson point process with rate $\mu[0, l]dl$ and that by Lemma 2.4.2 almost surely $\mu[0, l] \rightarrow \infty$ as $l \rightarrow \infty$. We omit the details. This concludes the proof of Proposition 2.3.2. \square

2.6 Compactness

2.6.1 Equivalent condition

In this section, we obtain a condition equivalent to that of Theorem 2.3.3 which is more convenient to study the compactness of the ICRT from the bounds provided by Lemmas 2.6.2 and 2.6.5. Additionally we also prove that the condition conjectured in [14] is also equivalent to that of Theorem 2.3.3. For $l \geq 0$, recall that \mathcal{X}_l is defined by $\mathbb{E}[\mu[0, \mathcal{X}_l]] = l$ and let

$$\psi(l) := \frac{\theta_0^2}{2} l^2 + \sum_{i=1}^{\infty} (e^{-l\theta_i} - 1 + l\theta_i).$$

Lemma 2.6.1. *The following conditions are equivalent:*

$$(i) \quad \int^{+\infty} \frac{dl}{\mathbb{E}[\mu[0, l]]} < +\infty \quad , \quad (ii) \quad \int^{+\infty} \frac{dl}{\psi(l)} < +\infty \quad , \quad (iii) \quad \sum_{n=1}^{+\infty} \frac{\log \mathcal{X}_{2^n}}{2^n} < \infty.$$

Proof. Since for every $x \in \mathbb{R}^+$, $e^{-x} - 1 + x \leq x(1 - e^{-x}) \leq 2(e^{-x} - 1 + x)$, for every $l \geq 0$:

$$\frac{\theta_0^2}{2} l^2 + \sum_{i=1}^{\infty} (e^{-l\theta_i} - 1 + l\theta_i) \leq \theta_0^2 l^2 + \sum_{i=1}^{\infty} l\theta_i (1 - e^{-\theta_i l}) \leq \theta_0^2 l^2 + \sum_{i=1}^{\infty} 2(e^{-l\theta_i} - 1 + l\theta_i).$$

So by (2.2) for every $l \geq 0$, $\psi(l) \leq \mathbb{E}[\mu[0, l]] \leq 2\psi(l)$. It follows readily that (i) and (ii) are equivalent. Furthermore

$$\int_{\mathcal{X}_1}^{\infty} \frac{dl}{\mathbb{E}[\mu[0, l]]} = \sum_{k=0}^{\infty} \int_{\mathcal{X}_{2^k}}^{\mathcal{X}_{2^{k+1}}} \frac{dl}{\mathbb{E}[\mu[0, l]]} \leq \sum_{k=0}^{\infty} \int_{\mathcal{X}_{2^k}}^{\mathcal{X}_{2^{k+1}}} \frac{dl}{l2^k} = \sum_{k=1}^{\infty} \frac{\log \mathcal{X}_{2^k}}{2^k} - \log \mathcal{X}_1,$$

and similarly

$$\int_{\mathcal{X}_1}^{\infty} \frac{dl}{\mathbb{E}[\mu[0, l]]} = \sum_{k=0}^{\infty} \int_{\mathcal{X}_{2^k}}^{\mathcal{X}_{2^{k+1}}} \frac{dl}{\mathbb{E}[\mu[0, l]]} \geq \sum_{k=0}^{\infty} \int_{\mathcal{X}_{2^k}}^{\mathcal{X}_{2^{k+1}}} \frac{dl}{l2^{k+1}} = \sum_{k=1}^{\infty} \frac{\log \mathcal{X}_{2^k}}{2^{k+1}} - \frac{\log \mathcal{X}_1}{2}.$$

So (i) and (iii) are equivalent. \square

2.6.2 The condition of Theorem 2.3.3 is sufficient for compactness

The aim of this section is to prove Lemma 2.6.2 below. This Lemma implies that under condition (iii) of Lemma 2.6.1, $(\mathcal{T}_{\mathcal{X}_{2^k}})_{k \in \mathbb{N}}$ is a Cauchy sequence of compact sets for the Hausdorff topology and thus converges toward a compact set. Since $(\mathcal{T}_{\mathcal{X}_{2^k}})_{k \in \mathbb{N}}$ is increasing (for \subset) toward \mathcal{T} , \mathcal{T} is the only possible limit, and hence is compact.

Lemma 2.6.2. *Almost surely, for every k large enough:*

$$d_H(\mathcal{T}_{\mathcal{X}_{2^{k-1}}}, \mathcal{T}_{\mathcal{X}_{2^k}}) \leq 21 \frac{\log \mathcal{X}_{2^k}}{2^k}.$$

Proof. For every $k \in \mathbb{N}$ and $x \in \mathcal{T}$, let $E_k(x)$ denote the event $d(x, [0, \mathcal{X}_{2^{k-1}}]) > 20 \log \mathcal{X}_{2^k} 2^{-k}$. First by Fubini's theorem and Lemma 2.4.7, we have conditionally on μ , for every k large enough,

$$\mathbb{E} \left[\int_0^{\mathcal{X}_{2^k}} \mathbf{1}_{E_k(x)} dx \middle| \mu \right] = \int_0^{\mathcal{X}_{2^k}} \mathbb{P}(E_k(x) | \mu) dx \leq \mathcal{X}_{2^k} \exp \left(-5 \frac{\log \mathcal{X}_{2^k}}{2^k} \mu[0, \mathcal{X}_{2^{k-1}}] \right).$$

Then by Lemma 2.4.2 as k goes to infinity $\mu[0, \mathcal{X}_{2^k}] \sim 2^k$. So for every k large enough:

$$\mathbb{E} \left[\int_0^{\mathcal{X}_{2^k}} \mathbf{1}_{E_k(x)} dx \middle| \mu \right] \leq \mathcal{X}_{2^k}^{-4/3}.$$

Furthermore by Lemma 2.4.1, $2^k = O(\mathcal{X}_{2^k})$ so $\sum \mathcal{X}_{2^k}^{-1/3} < \infty$. Hence by Markov's inequality and the Borel–Cantelli lemma, for every k large enough:

$$\int_0^{\mathcal{X}_{2^k}} \mathbf{1}_{E_k(x)} dx < \mathcal{X}_{2^k}^{-1}.$$

Note that it implies that, for every k large enough and $x \in [0, \mathcal{X}_{2^k}]$,

$$d(x, [0, \mathcal{X}_{2^{k-1}}]) \leq 20 \log \mathcal{X}_{2^k} 2^{-k} + \mathcal{X}_{2^k}^{-1},$$

since otherwise the geodesic path from x to $[0, \mathcal{X}_{2^{k-1}}]$ would contain a segment S of length at least $\frac{1}{\mathcal{X}_{2^k}}$ such that $d(S, [0, \mathcal{X}_{2^{k-1}}]) > 20 \log \mathcal{X}_{2^k} 2^{-k}$. Finally by Lemma 2.4.1, for every k large enough $\mathcal{X}_{2^k} \geq 2^k$, hence $\mathcal{X}_{2^k}^{-1} \leq \log \mathcal{X}_{2^k} 2^{-k}$. This concludes the proof. \square

2.6.3 The condition of Theorem 2.3.3 is necessary for compactness

The following section is organized as follows: Lemma 2.6.3 defines and proves the existence of "long" segments, Lemma 2.6.4 proves that they tend to "aggregate". Lemma 2.6.5 deduces a lower bound on $d_H(\mathcal{T}_{\mathcal{X}_{2^k}}, \mathcal{T})$ from the two previous lemmas, thus proving that the condition is necessary. Finally Lemma 2.6.6 gives a more precise view of the geometry of the tree in the non-compact case: "the tree is infinite in every direction".

Lemma 2.6.3. *For every $n \in \mathbb{N}$ let $L_n := \frac{\log \mathcal{X}_{2^n}}{2^{n+2}}$ and let \mathcal{I}_n be the set of segments $[Y_a + L_n, Y_{a+1}]$ with*

$$Y_a \in [\mathcal{X}_{2^n}, \mathcal{X}_{2^{n+1}}] \quad ; \quad Y_a + L_n \leq Y_{a+1} \quad ; \quad \mu[Y_a + L_n, Y_{a+1}] \geq 1/\mathcal{X}_{2^{n+1}}^2.$$

Almost surely for every n large enough we have $\#\mathcal{I}_n \geq 2^{n+2} \mathcal{X}_{2^n}^{1/3}$.

Proof. Write \mathcal{I}'_n for the set of segments $[Y_a + L_n, Y_{a+1}]$ with

$$Y_a \in [\mathcal{X}_{2^n}, \mathcal{X}_{2^{n+1}}] \quad ; \quad Y_a + L_n \leq Y_{a+1} \quad ; \quad \mu[Y_a + L_n, Y_{a+1}] < 1/\mathcal{X}_{2^{n+1}}^2.$$

First by Lemmas 2.4.4 and 2.4.2, for every n large enough, there are at most $2^{n+2} \mathcal{X}_{2^{n+1}}$ cuts on $[0, \mathcal{X}_{2^{n+1}}]$, hence $\#\mathcal{I}'_n \leq 2^{n+2} \mathcal{X}_{2^{n+1}}$. Furthermore, by Lemma 2.4.6, for every n large enough and $I \in \mathcal{I}_n$, we have $\mu(I) \leq \log^2 \mathcal{X}_{2^n} / \mathcal{X}_{2^n}$ so

$$\sum_{I \in \mathcal{I}_n \cup \mathcal{I}'_n} \mu(I) = \sum_{I \in \mathcal{I}_n} \mu(I) + \sum_{I \in \mathcal{I}'_n} \mu(I) \leq \#\mathcal{I}_n \frac{\log^2 \mathcal{X}_{2^n}}{\mathcal{X}_{2^n}} + \frac{2^{n+2}}{\mathcal{X}_{2^{n+1}}}.$$

Therefore, since $\mathcal{X}_{2^{n+1}} \geq \mathcal{X}_{2^n}$, it suffices to prove that, writing $S_n := \bigcup_{I \in \mathcal{I}_n \cup \mathcal{I}'_n} I$,

$$\mu(S_n) > 2^{n+2} \mathcal{X}_{2^n}^{-2/3} \log^2 \mathcal{X}_{2^n} + 2^{n+2} / \mathcal{X}_{2^n}. \quad (2.22)$$

Note that for every $x \in [\mathcal{X}_{2^n}, \mathcal{X}_{2^{n+1}}]$, $x \in S_n$ if and only if there is a cut in $[\mathcal{X}_{2^n}, x]$ and no cut in $[x - L_n, x]$. So if there is a cut in $[\mathcal{X}_{2^n}, \mathcal{X}_{2^n} + 1]$,

$$\mu(S_n) \geq \int_{\mathcal{X}_{2^n} + L_n + 1}^{\mathcal{X}_{2^n} + 1} \emptyset_{x-L_n, x} d\mu(x),$$

where for every $x \leq y$, $\emptyset_{x,y} := \mathbf{1}_{\forall i \in \mathbb{N}, Y_i \notin [x,y]}$. Let A_n denote the right-hand side above. Since, conditionally on μ , $(Y_i)_{i \in \mathbb{N}}$ is a Poisson point process of rate $\mu[0, l]dl$, for every n large enough

$$\mathbb{P}(A_n > \mu(S_n) | \mu) \leq \mathbb{P}(\emptyset_{\mathcal{X}_{2^n}, \mathcal{X}_{2^n} + 1} = 0 | \mu) \leq e^{-\mu[0, \mathcal{X}_{2^n}]} \leq e^{-2^{n-1}}.$$

Therefore, by the Borel–Cantelli lemma, almost surely for every n large enough $\mu(S_n) \geq A_n$.

We now lower bound A_n via a second moment method. We have, still by the properties of $(Y_i)_{i \in \mathbb{N}}$,

$$\mathbb{E}[A_n | \mu] \geq \int_{\mathcal{X}_{2^n} + L_n + 1}^{\mathcal{X}_{2^n} + 1} e^{-\mu[0, \mathcal{X}_{2^n} + 1]} L_n d\mu(x) = e^{-\mu[0, \mathcal{X}_{2^n} + 1]} L_n \mu[\mathcal{X}_{2^n} + L_n + 1, \mathcal{X}_{2^n} + 1]. \quad (2.23)$$

Furthermore note that $\frac{1+L_n}{\mathcal{X}_{2^n}} \rightarrow 0$ as $n \rightarrow \infty$, hence by Lemmas 2.4.2 and 2.4.1 a.s. as $n \rightarrow \infty$,

$$\mu[\mathcal{X}_{2^n} + L_n + 1, \mathcal{X}_{2^n} + 1] = \mathbb{E}[\mu[0, \mathcal{X}_{2^n} + 1]](1 + o(1)) - \mathbb{E}[\mu[0, \mathcal{X}_{2^n} + L_n + 1]](1 + o(1)) \sim 2^n.$$

It follows from (2.23), Lemma 2.4.2, and the definition of L_n that, as $n \rightarrow \infty$,

$$\mathbb{E}[A_n | \mu] \geq \mathcal{X}_{2^n}^{-1/2+o(1)} 2^n. \quad (2.24)$$

Moreover we have by Fubini's theorem,

$$\mathbb{V}[A_n | \mu] = \int_{\mathcal{X}_{2^n} + L_n + 1}^{\mathcal{X}_{2^n} + 1} \int_{\mathcal{X}_{2^n} + L_n + 1}^{\mathcal{X}_{2^n} + 1} \text{Cov}[\emptyset_{x-L_n, x} \emptyset_{y-L_n, y} | \mu] d\mu(x) d\mu(y).$$

Note that for every $x, y \in \mathbb{R}^+$, $\text{Cov}[\emptyset_{x-L_n, x} \emptyset_{y-L_n, y} | \mu] \leq \mathbb{E}[\emptyset_{y-L_n, y} | \mu]$, and that conditionally on μ , $\emptyset_{x-L_n, x}$ and $\emptyset_{y-L_n, y}$ are independent when $|y - x| > L_n$. It follows that,

$$\begin{aligned} \mathbb{V}[A_n | \mu] &\leq \int_{\mathcal{X}_{2^n} + L_n + 1}^{\mathcal{X}_{2^n} + 1} \mathbb{E}[\emptyset_{y-L_n, y} | \mu] \int_{y-L_n}^{y+L_n} d\mu(x) d\mu(y) \\ &\leq \mathbb{E}[A_n | \mu] \max_{\mathcal{X}_{2^n} + L_n + 1 \leq y \leq \mathcal{X}_{2^n} + 1} \mu[y - L_n, y + L_n]. \end{aligned} \quad (2.25)$$

Furthermore by Lemma 2.4.3, for every n large enough and $y \in [\mathcal{X}_{2^n} + 1, \mathcal{X}_{2^{n+1}} - L_n]$,

$$\mu[y, y + 2L_n] \leq 4L_n \frac{\mathbb{E}[\mu[0, y]]}{y} + 13 \frac{\log y}{y} \leq 4 \frac{\log \mathcal{X}_{2^n}}{2^{n+2}} \frac{\mathbb{E}[\mu[0, \mathcal{X}_{2^{n+1}}]]}{\mathcal{X}_{2^n}} + 13 \frac{\log \mathcal{X}_{2^n}}{\mathcal{X}_{2^n}}. \quad (2.26)$$

Put together (2.25) and (2.26) yield as $n \rightarrow \infty$,

$$\mathbb{V}[A_n | \mu] \leq \mathbb{E}[A_n | \mu] \frac{17 \log \mathcal{X}_{2^n}}{\mathcal{X}_{2^n}}. \quad (2.27)$$

Therefore, by Chebyshev's inequality, (2.27), and (2.24), we have as $n \rightarrow \infty$,

$$\mathbb{P}\left(A_n \leq \frac{\mathbb{E}[A_n | \mu]}{2} \mid \mu\right) \leq 4 \frac{\mathbb{V}[A_n | \mu]}{\mathbb{E}[A_n | \mu]^2} \leq \frac{O(1) \log \mathcal{X}_{2^n}}{\mathcal{X}_{2^n} \mathbb{E}[A_n | \mu]} \leq 2^{-n} \mathcal{X}_{2^n}^{-1/2+o(1)}.$$

So by the Borel–Cantelli lemma, almost surely for every n large enough $A_n \geq \mathbb{E}[A_n | \mu]/2$. Finally the inequality in (2.22) follows from (2.24) and the fact that for every n large enough $\mu(S_n) \geq A_n$. This concludes the proof. \square

Formally we call the segments in $\bigcup_{n \in \mathbb{N}} \mathcal{I}_n$ "long". The following lemma proves that those long segments tend to "glue" to one another.

Lemma 2.6.4. *For every $I \in \bigcup_{n \in \mathbb{N}} \mathcal{I}_n$ let a_I denote the only integer such that $I \subset (Y_{a_I}, Y_{a_I+1}]$. Almost surely for every $n, m \in \mathbb{N}$ large enough with $n < m$ and $\mathcal{X}_{2^m} \geq \mathcal{X}_{2^{n+1}}^8$, for every $I \in \mathcal{I}_n$ there exists $I' \in \mathcal{I}_m$ such that $Z_{a_{I'}} \in I$. In this case we say that I' is glued on I .*

Proof. Conditionally on $\mathcal{F} := \sigma(\mu, (Y_i)_{i \geq 1}, (Z_i)_{i \geq 1})$ are independent random variables with law $(p_{Y_i})_{i \in \mathbb{N}}$ so for every $i \in \mathbb{N}$ and $I \in \mathcal{I}_n$

$$\mathbb{P}(\forall I' \in \mathcal{I}_m, Z_{a_{I'}} \notin I \mid \mathcal{F}) = \prod_{I' \in \mathcal{I}_m} \left(1 - \frac{\mu(I)}{\mu[0, Y_{a_{I'}}]}\right) \leq \exp\left(-\#\mathcal{I}_m \frac{\mu(I)}{\mu[0, \mathcal{X}_{2^{m+1}}]}\right).$$

Furthermore we have by definition of I_n , $\mu(I) \leq \mathcal{X}_{2^{n+1}}^{-2} \leq \mathcal{X}_{2^m}^{-1/4}$. It follows from Lemmas 2.6.3 and 2.4.2 that for every m large enough,

$$\mathbb{P}(\forall I' \in \mathcal{I}_m, Z_{a_{I'}} \notin I \mid \mathcal{F}) \leq \exp\left(-\mathcal{X}_{2^m}^{-1/4} \frac{2^{m+2} \mathcal{X}_{2^m}^{1/3}}{2^{m+2}}\right) = \exp\left(-\mathcal{X}_{2^m}^{1/12}\right).$$

Moreover, by Lemma 2.4.4, for every i large enough $\#\mathcal{I}_n \leq 2\mathcal{X}_{2^{n+1}}^2$, and by Lemma 2.4.1 for every $m \in \mathbb{N}$, $\mathcal{X}_{2^m} \geq 2^m$. So for every m large enough,

$$\mathbb{P}(\exists I \in \mathcal{I}_n, \forall I' \in \mathcal{I}_m, Z_{a_{I'}} \notin I \mid \mathcal{F}) \leq 2\mathcal{X}_{2^n}^2 e^{-\mathcal{X}_{2^m}^{1/12}} \leq f(\mathcal{X}_{2^m}) \leq f(2^m),$$

where $f : x \mapsto 2x^2 e^{-x^{1/12}}$. Since $\sum_{n=0}^{\infty} \sum_{m=n}^{\infty} f(2^m) < \infty$ the Borel–Cantelli lemma yields the desired result. \square

Lemma 2.6.5. *Almost surely for every k large enough:*

$$d_H(\mathcal{T}_{\mathcal{X}_{2^k}}, \mathcal{T}) \geq \frac{1}{128} \sum_{n=k}^{\infty} \frac{\log \mathcal{X}_{2^n}}{2^n}.$$

Proof. First define by induction $(n_i)_{i \in \mathbb{N}}$ such that $n_0 = k$ and such that for every $i \geq 0$, $n_{i+1} = \inf\{n \in \mathbb{N} : n > n_i, \mathcal{X}_{2^n} \geq \mathcal{X}_{2^{n_i}}^8\}$. Note that for $i \in \mathbb{N}$, $\mathcal{X}_{2^{n_{2i+2}}} \geq \mathcal{X}_{2^{n_{2i+1}}}^8 \geq \mathcal{X}_{2^{n_i+1}}^8$, so by Lemma 2.6.4, there exists a sequence $(I_i)_{i \in \mathbb{N}}$ such that for every $i \in \mathbb{N}$, $I_i \in \mathcal{I}_{n_{2i}}$ and I_{i+1} is glued on I_i . On this event, note that for every $j \in \mathbb{N}$ and $x \in I_j$, x is at distance at least $\sum_{i=0}^{j-1} L_{n_{2i}}$ of \mathcal{X}_{2^k} so $d_H(\mathcal{T}_{\mathcal{X}_{2^k}}, \mathcal{T}) \geq \sum_{i=0}^{\infty} L_{n_{2i}}$. Similarly we have $d_H(\mathcal{T}_{\mathcal{X}_{2^k}}, \mathcal{T}) \geq \sum_{i=0}^{\infty} L_{n_{2i+1}}$, hence

$$d_H(\mathcal{T}_{\mathcal{X}_{2^k}}, \mathcal{T}) \geq \frac{1}{2} \sum_{i=0}^{\infty} L_{n_i}.$$

Finally we compare $\sum_{i=0}^{\infty} L_{n_i}$ with $\sum_{n=k+1}^{\infty} \frac{\log \mathcal{X}_{2^n}}{2^n}$. By definition of $\{n_i\}_{i \in \mathbb{N}}$ we have:

$$\sum_{n=k}^{\infty} \frac{\log \mathcal{X}_{2^n}}{2^n} = \sum_{i=0}^{\infty} \sum_{n=n_i}^{n_{i+1}-1} \frac{\log \mathcal{X}_{2^n}}{2^n} \leq \sum_{i=0}^{\infty} \sum_{n=n_i}^{n_{i+1}-1} \frac{8 \log \mathcal{X}_{2^{n_i}}}{2^n} \leq 64 \sum_{i=0}^{\infty} \frac{\log \mathcal{X}_{2^{n_i}}}{2^{n_i+2}} = 64 \sum_{i=0}^{\infty} L_{n_i}.$$

This concludes the proof. \square

The previous lemma proves that when $\sum \frac{\log \mathcal{X}_{2^n}}{2^n} = \infty$ the tree is not compact, thus finishing the proof of Theorem 2.3.1. The next lemma gives a more precise description of the geometry of the tree in the non-compact case: "the tree is infinite in every direction".

Lemma 2.6.6. *Suppose that $\sum \log \mathcal{X}_{2^n} 2^{-n} = \infty$ then almost surely for every $a < b < c$, $[a, b]^{\uparrow c}$ has infinite diameter.*

Remark. A similar result is proved in Le Gall and Le Jan [16] for non-compact Lévy trees: the set of values taken by the height process on any non-trivial open interval contains a half line $[a, \infty)$.

Proof. First one may adapt the argument of the proof of Lemma 2.6.5 to prove that for every $k \in \mathbb{N}$ large enough and $I \in \mathcal{I}_k$,

$$\text{diam}((Y_{a_I}, Y_{a_I+1}]^{\uparrow Y_{a_I+1}}) \geq L_k + \frac{1}{2} \left(\sum_{i=1}^{\infty} L_{n_{2i+1}} + \sum_{i=1}^{\infty} L_{n_{2i}} \right) \geq \frac{1}{64} \sum_{n=n_2}^{\infty} \frac{\log \mathcal{X}_{2^n}}{2^n} = \infty, \quad (2.28)$$

where $\{n_i\}_{i \in \mathbb{N}}$ is defined in the proof of Lemma 2.6.5 and a_I in Lemma 2.6.4. We leave the details to the reader.

We now fix $a < b < c \in \mathbb{R}^+$. Since conditionally on $\mathcal{F} := \sigma(\mu, (Y_i)_{i \geq 1})$, $(Z_i)_{i \geq 1}$ are independent random variables with law $(p_{Y_i})_{i \in \mathbb{N}}$, we have for every $m \in \mathbb{N}$,

$$\mathbb{P}(\forall I \in \mathcal{I}_m, Z_{a_I} \notin [a, b] | \mathcal{F}) = \prod_{I' \in \mathcal{I}_m} \left(1 - \frac{\mu[a, b]}{\mu[0, Y_{a_{I'}}]} \right) \leq \exp \left(- \frac{\mu[a, b] \# \mathcal{I}_m}{\mu[0, \mathcal{X}_{2^{m+1}}]} \right).$$

Since μ has full support it follows from Lemmas 2.4.2 and 2.6.3 that the right-hand side above converges to 0 as $m \rightarrow \infty$. Therefore for every $n \in \mathbb{N}$, there exists almost surely $m \geq n$ and $I \in \mathcal{I}_m$ such that $Z_{a_I} \in [a, b]$. It follows from (2.28) that if m is large enough $[a, b]^{\uparrow Y_{a_I}}$ has infinite diameter, hence $[a, b]^{\uparrow c}$ also has infinite diameter. Since $a < b < c$ are arbitrary and since rational numbers are dense on \mathbb{R}^+ , the desired claim follows. \square

2.7 Fractal dimensions : proof of theorem 2.3.4

In this section we prove Theorem 2.3.4. By Lemma 2.2.3, it suffices to upper bound the Minkowski dimensions and to lower bound the Packing and Hausdorff dimension. We obtain the upper bounds from some simple cover of \mathcal{T} and we derive the lower bounds from Lemma 2.2.2.

2.7.1 Upper bound for the Minkowski dimensions

First from the change of variables $u = \mathcal{X}_l$, note that the upper bound for the Minkowski dimensions given by Theorem 2.3.4 are equivalent to

$$(a) \quad \overline{\dim}(\mathcal{T}) \leq \limsup_{l \rightarrow \infty} \frac{\log(l\mathcal{X}_l)}{\log l} \quad \text{and} \quad (b) \quad \underline{\dim}(\mathcal{T}) \leq \liminf_{l \rightarrow \infty} \frac{\log(l\mathcal{X}_l)}{\log l} \quad \text{when} \quad \log \mathcal{X}_l = l^{o(1)}.$$

Then for every $l \in \mathbb{R}$, $\mathcal{T}_{\mathcal{X}_l}$ has total length \mathcal{X}_l , hence one can construct a cover of $\mathcal{T}_{\mathcal{X}_l}$ using $l\mathcal{X}_l$ balls of radius $2/l$. By increasing the radius of those balls by $d_H(\mathcal{T}_{\mathcal{X}_l}, \mathcal{T})$ one obtains a cover of \mathcal{T} . So for every $l \in \mathbb{R}^+$,

$$N_{2/l+d_H(\mathcal{X}_l, \mathcal{T})} \leq l\mathcal{X}_l. \quad (2.29)$$

The claims (a) and (b) are applications of the inequality in (2.29).

Toward proving (a), we may assume that $\log \mathcal{X}_l = O(\log l)$ since otherwise the bound is trivial. It follows from Lemma 2.6.2 that $d_H(\mathcal{T}_{\mathcal{X}_{2^{k-1}}}, \mathcal{T}_{\mathcal{X}_{2^k}}) = O(k/2^k)$ and summing over all $k \geq \log_2(l)$, we obtain $d_H(\mathcal{X}_l, \mathcal{T}) = O(\log(l)/l)$. Therefore by (2.29),

$$\overline{\dim}(\mathcal{T}) = \limsup_{l \rightarrow \infty} \frac{\log N_{1/l}}{\log l} = \limsup_{l \rightarrow \infty} \frac{\log N_{2/l+d_H(\mathcal{T}_{\mathcal{X}_l}, \mathcal{T})}}{-\log(2/l + d_H(\mathcal{T}_{\mathcal{X}_l}, \mathcal{T}))} \leq \limsup_{l \rightarrow \infty} \frac{\log(l\mathcal{X}_l)}{\log l}.$$

and (a) follows. (b) can be treated similarly by observing that Lemma 2.6.2 and $\log \mathcal{X}_l = l^{o(1)}$ implies that $d_H(\mathcal{X}_l, \mathcal{T}) = l^{-1+o(1)}$. We leave the details to the reader. This concludes the proof.

2.7.2 Lower bound for the Packing dimension and the Hausdorff dimension

In this section we show that almost surely,

$$\dim_P(\mathcal{T}) \geq \alpha := 1 + \limsup_{l \rightarrow \infty} \frac{\log l}{\log \mathbb{E}[\mu[0, l]]} \quad \text{and} \quad \dim_H(\mathcal{T}) \geq \beta := 1 + \liminf_{l \rightarrow \infty} \frac{\log l}{\log \mathbb{E}[\mu[0, l]]}.$$

To this end, by lemma 2.2.2 it suffices to prove that if A is a random variable with law p then almost surely for every $\delta > 0$, $\liminf p(B(A, \varepsilon))\varepsilon^{-\alpha-\delta} < \infty$ and $p(B(A, \varepsilon)) = O(\varepsilon^{\beta+\delta})$ as $\varepsilon \rightarrow 0$. The two previous inequalities can be proved via an elementary computation using $\mu[0, l] \sim \mathbb{E}[\mu[0, l]]$ (Lemma 2.4.2),

$$(a) \quad p(B(A, d(A, \mathcal{T}_{Y_i}))) \leq \frac{1}{Y_i^{1+o(1)}\mu[0, Y_i]} \quad \text{and} \quad (b) \quad d(A, \mathcal{T}_{Y_{i+1}}) \geq \mu[0, Y_i]^{-1+o(1)}.$$

We omit the details and focus on the proof of (a) and (b).

Toward (a), let γ be the geodesic path from 0 to A and let for every $i \in \mathbb{N}$, $j_i := \min\{j \geq i : (Y_j, Y_{j+1}] \cap \gamma \neq \emptyset\}$. Note that since by Theorem 2.3.1 almost surely $A \notin \mathbb{R}^+$,

$$B(A, d(A, \mathcal{T}_{Y_i})) \subset \{Z_{j_i}\} \cup (Y_{j_i}, Y_{j_i+1}]^{\uparrow Y_{j_i+1}}.$$

Furthermore we have by Lemma 2.4.6, $\mu(Y_{j_i}, Y_{j_i+1}] \leq \frac{\log^2 Y_i}{Y_i}$. Therefore by Lemma 2.5.1 (iii), conditionally on $(\mu, \{Y_j\}_{j \in \mathbb{N}})$, for every i large enough with probability at least $1 - Y_{j_i}^{-5} \geq 1 - Y_i^{-5}$,

$$p(B(A, d(A, \mathcal{T}_{Y_i}))) \leq p\left(\uparrow Y_{j_i}, Y_{j_i+1}]^{\uparrow Y_{j_i+1}}\right) \leq 2 \frac{\log^6 Y_{j_i}}{Y_{j_i} \mu[0, Y_{j_i}]} \leq 2 \frac{\log^6 Y_i}{Y_i \mu[0, Y_i]}. \quad (2.30)$$

Moreover by Lemma 2.4.4, we have $i = O(Y_i^2)$, hence $\sum_{i=1}^{\infty} Y_i^{-5} < \infty$. The Borel–Cantelli lemma then yields that almost surely (2.30) holds for every i large enough, hence (a) holds.

Toward (b), let us first upper bound $\{p(S_n)\}_{n \in \mathbb{N}}$ where for $n \in \mathbb{N}$, S_n denotes the set of $x \in \mathcal{T}$ such that $d(x, [0, \mathcal{X}_{2^n}]) \leq \delta_n := \frac{1}{2^{n/6}}$. Let for every $n \in \mathbb{N}$,

$$a_n := \max\{a : Y_a \leq \mathcal{X}_{n^2 2^n}\}, \quad S'_n := \{x \in (\mathcal{X}_{2^n} + \delta_n, Y_{a_n}] : [x - \delta_n, x] \cap \{Y_i\}_{i \in \mathbb{N}} \neq \emptyset\}.$$

Note that for every $n \in \mathbb{N}$, $S_n \subset ([0, \mathcal{X}_{2^n} + \delta_n] \cup S'_n)^{\uparrow Y_{a_n}}$. Therefore by Lemma 2.5.1 (ii), 2.4.2 and 2.4.6, almost surely for every n large enough:

$$p(S_n) \leq 2p_{Y_{a_n}}([0, \mathcal{X}_{2^n} + \delta_n] \cup S'_n) = 2 \frac{\mu[0, \mathcal{X}_{2^n} + \delta_n] + \mu(S'_n)}{\mu[0, \mathcal{X}_{n^2 2^n}] - \mu(Y_{a_n}, \mathcal{X}_{n^2 2^n}]} \leq 4 \frac{2^n + \mu(S'_n)}{n^2 2^n}. \quad (2.31)$$

Furthermore since conditionally on μ , $(Y_i)_{i \in \mathbb{N}}$ is a Poisson point process of rate $\mu[0, l]dl$, we have by Fubini's theorem, for every $n \in \mathbb{N}$:

$$\begin{aligned} \mathbb{E}[\mu(S'_n)|\mu] &\leq \int_0^{\mathcal{X}_{n^2 2^n}} \mathbb{P}([x - \delta_n, x] \cap \{Y_i\}_{i \in \mathbb{N}} \neq \emptyset | \mu) d\mu(x) \\ &\leq \int_0^{\mathcal{X}_{n^2 2^n}} 1 - e^{-\delta_n \mu[0, \mathcal{X}_{n^2 2^n}]} d\mu(x) \\ &\leq \delta_n \mu[0, \mathcal{X}_{n^2 2^n}]^2. \end{aligned}$$

It directly follows from Lemma 2.4.2 that almost surely $\mathbb{E}[\mu(S'_n)|\mu] = O(2^n/n^2)$. Thus by Markov's inequality and the Borel–Cantelli lemma a.s. $\mu(S'_n) = O(2^n)$. Therefore by (2.31), $p(S_n) = O(1/n^2)$, hence by the Borel–Cantelli lemma a.s. for every n large enough, $A \notin S_n$.

Finally let for every $i \in \mathbb{N}$, $n_i := \inf\{n \in \mathbb{N}, Y_{i+1} \leq \mathcal{X}_{2^n}\}$. We have by Lemmas 2.4.2 and 2.4.6 a.s. $\mu[0, Y_i] \geq 2^{n_i+O(1)}$. Hence, since for every i large enough $A \notin S_{n_i}$, we have,

$$d(A, Y_{i+1}) \geq d(A, \mathcal{X}_{2^{n_i}}) \geq \frac{1}{2^{n_i n_i^6}} \geq \mu[0, Y_i]^{-1+o(1)}.$$

This concludes the proof of (b) and therefore of Theorem 2.3.4.

Acknowledgment Thanks are due to Nicolas Broutin for interesting conversations and numerous advice on earlier versions of this paper.

2.8 Appendix: Concentration inequalities

We first prove an exponential concentration inequality for general Pólya urns.

Lemma 2.8.1. *Let $\{m_n\}_{n \geq 0}$ be a positive real-valued sequence. Let $(A_n)_{n \geq 0}$ be a sequence of positive real-valued random variables such that $A_0 \leq m_0$ and such that for every $n \geq 0$,*

$$\mathbb{P}(A_{n+1} = A_n + m_{n+1} | A_n) = \frac{A_n}{M_n} \quad ; \quad \mathbb{P}(A_{n+1} = A_n | A_n) = \frac{M_n - A_n}{M_n},$$

where for every $n \geq 0$, $M_n = \sum_{i=0}^n m_i$. We say that in this case $(A_n)_{n \geq 0}$ is a $(A_0, \{m_i\}_{i \geq 0})$ Pólya urn.

a) *If $\sum_{n=0}^{\infty} \frac{m_n^2}{M_n^2} < \infty$, then almost surely for every $a \geq 0$ and $t \in \mathbb{R}^+$,*

$$\mathbb{P}\left(\sup_{i \geq a} \left| \frac{A_i}{M_i} - \frac{A_a}{M_a} \right| > t \frac{A_a}{M_a} \middle| A_a\right) \leq 2 \exp\left(-\frac{\frac{t^2}{4} \frac{A_a}{M_a}}{\sum_{n>a} \frac{m_n^2}{M_n^2} + t \max\left(\sum_{n>a} \frac{m_n^2}{M_n^2}, \max_{n>a} \frac{m_n}{M_n}\right)}\right).$$

b) *If $\{m_n\}_{n \in \mathbb{N}}$ is bounded, then almost surely for every $a \geq 0$ and $t \in \mathbb{R}^+$,*

$$\mathbb{P}\left(\sup_{i \geq a} \left| \frac{A_i}{M_i} - \frac{A_a}{M_a} \right| > t \frac{A_a}{M_a} \middle| A_a\right) \leq 2 \exp\left(-\frac{t^2}{4(1+t)} \frac{A_a}{\max_{n>a} m_n}\right).$$

Remark. Note that Lemma 2.8.1 implies that almost surely $(\frac{A_i}{M_i})_{i \in \mathbb{N}}$ is a Cauchy sequence and so converges. The statement should then be seen as an estimate on the speed of convergence.

Proof. First let us explain why (b) follows from (a). We have for every $a \in \mathbb{N}$,

$$\sum_{n>a} \frac{m_n^2}{M_n^2} \leq \max_{n>a} m_n \sum_{n>a} \frac{m_n}{M_n^2} \leq \max_{n>a} m_n \int_{M_a}^{+\infty} \frac{dx}{x^2} = \frac{\max_{n>a} m_n}{M_a},$$

and (b) follows by replacing $\max_{n>a} \frac{m_n}{M_n}$ and $\sum_{n>a} \frac{m_n^2}{M_n^2}$ by the upper bound $\frac{\max_{n>a} m_n}{M_a}$ in (a).

We focus henceforth on (a). To simplify the notation set for every $n \in \mathbb{N}$, $X_n := A_n/M_n$ and $\delta_n := m_n/M_n$. Also we write for every $a \in \mathbb{N}$, $\mathbb{E}^{(a)}[\dots] = \mathbb{E}[\dots | X_a]$. We first prove by induction that for every $a, b, c \in \mathbb{N}$ and $\lambda \in \mathbb{R}$ satisfying

$$a \leq b \leq c \quad ; \quad |\lambda| \leq \Lambda_a := \frac{1}{4 \max\{\sum_{n>a} \delta_n^2, \max_{n>a} \delta_n\}} \quad (2.32)$$

we have

$$\mathbb{E}^{(a)}[e^{\lambda X_c}] \leq f(a, b, c, \lambda) := \mathbb{E}^{(a)}[e^{\lambda(1+\lambda \sum_{n=b+1}^c \delta_n^2) X_b}]. \quad (P(a, b, c, \lambda))$$

Note that when $b = c$, $P(a, b, c, \lambda)$ is trivial. Therefore it suffices to prove that for every a, b, c, λ such that $a \leq b < c$ and $\lambda \leq \Lambda_a$ that $f(a, b+1, c, \lambda) \leq f(a, b, c, \lambda)$. Fix $a \leq b < c$, $\lambda \leq \Lambda_a$ and let $\gamma := \lambda(1 + \lambda \sum_{n=b+2}^c \delta_n^2)$. We have,

$$\begin{aligned} f(a, b+1, c, \lambda) &= \mathbb{E}^{(a)}[e^{\gamma X_{b+1}}] \\ &= \mathbb{E}^{(a)}\left[e^{\gamma X_b} \mathbb{E}^{(b)}[e^{\gamma(X_{b+1}-X_b)}]\right] \\ &= \mathbb{E}^{(a)}\left[e^{\gamma X_b} \left(X_b e^{\gamma\left(\frac{A_b+m_{b+1}}{M_{b+1}}-X_b\right)} + (1-X_b) e^{\gamma\left(\frac{A_b}{M_{b+1}}-X_b\right)}\right)\right] \\ &= \mathbb{E}^{(a)}\left[e^{\gamma X_b} \left(X_b e^{\gamma \delta_{b+1}(1-X_b)} + (1-X_b) e^{-\gamma \delta_{b+1} X_b}\right)\right]. \end{aligned} \quad (2.33)$$

Furthermore by (2.32), $|\gamma| \leq \frac{5}{4}|\lambda| \leq \frac{5}{16\delta_{b+1}}$, hence since for every $0 \leq x \leq 1$ and $|c| \leq \frac{5}{16}$, $xe^{c(1-x)} + (1-x)e^{-cx} \leq e^{\frac{16}{25}c^2x}$, we have

$$X_b e^{\gamma\delta_{b+1}(1-X_b)} + (1-X_b) e^{-\gamma\delta_{b+1}X_b} \leq e^{\frac{16}{25}\gamma^2\delta_{b+1}^2 X_b}. \quad (2.34)$$

Finally by (2.33), (2.34), and $|\gamma| \leq \frac{5}{4}|\lambda|$,

$$f(a, b+1, c, \lambda) \leq \mathbb{E}^{(a)} \left[e^{(\gamma + \frac{16}{25}\gamma^2\delta_{b+1}^2)X_b} \right] \leq \mathbb{E}^{(a)} \left[e^{(\gamma + \lambda^2\delta_{b+1}^2)X_b} \right] = f(a, b, c, \lambda).$$

This concludes our proof by induction of $P(a, b, c, \lambda)$.

We now fix $a \in \mathbb{N}$. For every $n \in \mathbb{N}$ and $0 \leq \lambda \leq \Lambda_a$, by $P(a, a, n, \lambda)$ and $P(a, a, n, -\lambda)$, we have the sub-Gaussian bound, $\mathbb{E}^{(a)} [e^{\lambda|X_n - X_a|}] \leq 2e^{\lambda^2 V_a/2}$ where $V_a := 2X_a \sum_{i>a} \delta_i^2$. Furthermore note that $\{X_n\}_{n \geq a}$ is a martingale, and hence that for every $\lambda \in \mathbb{R}^+$, $\{e^{\lambda|X_n - X_a|}\}_{n \geq a}$ is a sub-martingale. It follows by Doob's inequality that for every $t \in \mathbb{R}^+$ and $0 \leq \lambda \leq \Lambda_a$,

$$\mathbb{P}^{(a)} \left(\sup_{n \geq a} |X_n - X_a| \geq t \right) = \mathbb{P}^{(a)} \left(\sup_{n \geq a} e^{\lambda|X_n - X_a|} \geq e^{\lambda t} \right) \leq \sup_{n \geq a} \frac{\mathbb{E}^{(a)} [e^{\lambda|X_n - X_a|}]}{e^{\lambda t}} \leq 2e^{\lambda^2 \frac{V_a}{2} - \lambda t} \quad (2.35)$$

On the one hand, for every $0 \leq t \leq V_a \Lambda_a$, taking $\lambda := t/V_a$ in (2.35) gives,

$$\mathbb{P}^{(a)} \left(\sup_{n \geq a} |X_n - X_a| \geq t \right) \leq 2e^{-\frac{t^2}{2V_a}}. \quad (2.36)$$

On the other hand, for every $t > V_a \Lambda_a$, taking $\lambda := \Lambda_a$ in (2.35) gives,

$$\mathbb{P}^{(a)} \left(\sup_{n \geq a} |X_n - X_a| \geq t \right) \leq 2e^{\Lambda_a^2 \frac{V_a}{2} - \Lambda_a t} = 2e^{\frac{t^2}{V_a} \left(\frac{1}{2} \left(\frac{t}{\Lambda_a V_a} \right)^{-2} - \left(\frac{t}{\Lambda_a V_a} \right)^{-1} \right)},$$

hence since for every $x \geq 1$, $\frac{1}{2x^2} - \frac{1}{x} \leq -\frac{1}{2+x}$,

$$\mathbb{P}^{(a)} \left(\sup_{n \geq a} |X_n - X_a| \geq t \right) \leq 2e^{-\frac{t^2}{2V_a + t/\Lambda_a}}. \quad (2.37)$$

By (2.36) the last inequality is also true for $0 \leq t \leq V_a \Lambda_a$. The desired inequality then directly follows from a reorganization of the different terms in (2.37). We omit the straightforward details. \square

Lemma 2.8.2. *Let $p > 0$, $(X_i)_{i \in \mathbb{N}}$ be a family of independent Bernoulli random variables with mean p . Let $(a_i)_{i \in \mathbb{N}}$ be a positive real-valued sequence and let for every $n \in \mathbb{N}$, $A_n := \sum_{i=1}^n a_i$. Suppose that $A_n \rightarrow \infty$ and $\sum_{n=1}^{\infty} \frac{a_n^2}{A_n^2} < \infty$, then almost surely $\sum_{i=1}^n a_i X_i \sim p A_n$.*

Proof. Since $\sum_{n=1}^{\infty} \frac{a_n^2}{A_n^2} < \infty$ and $\{X_i\}_{i \in \mathbb{N}}$ are independent random variables, the classical three series theorem implies that $S_n := \sum_{i=1}^n \frac{a_i(X_i - p)}{A_i}$ almost surely converges as $n \rightarrow \infty$. Therefore,

$$\sum_{i=1}^n \frac{a_i(X_i - p)}{A_n} = \sum_{i=1}^n \frac{A_i}{A_n} (S_i - S_{i-1}) = S_n - \sum_{i=1}^{n-1} S_i \frac{A_{i+1} - A_i}{A_n} \xrightarrow{n \rightarrow \infty} 0. \quad \square$$

Remark. If there exists $\varepsilon > 0$ such that $a_i = O(A_i^{1-\varepsilon})$, then as n goes to infinity,

$$\sum_{n=1}^{\infty} \frac{a_n^2}{A_n^2} = O \left(\sum_{i=1}^{\infty} \frac{a_i}{A_i^{1+\varepsilon}} \right) = O \left(\int_1^{A_n} \frac{dx}{x^{1+\varepsilon}} \right) = O(1).$$

Lemma 2.8.3. *Let $\lambda > 0$. Let $(X_i)_{i \geq 0}$ be a family of independent exponential random variables of mean $(m/2^i)_{i \in \mathbb{N}}$. Let $S = \sum_{i \geq 0} X_i$. Then for every $t \geq 6$, $\mathbb{P}(S \geq 2mt) \leq e^{-t}$.*

Proof. Without loss of generality, we may assume that $m = 1$. Then, with $\lambda = 2/3$,

$$\mathbb{E}[e^{\lambda S}] = \prod_{i \geq 0} \mathbb{E}[e^{\lambda X_i}] = \prod_{i \geq 0} \frac{1}{1 - \lambda 2^{-i}} \leq e^2.$$

Hence, by Markov's inequality, for every $t \geq 6$,

$$\mathbb{P}(S \geq 2t) \leq \mathbb{E}[e^{\lambda S}] e^{-\lambda 2t} \leq e^{2-2\lambda t} \leq e^{-t}. \quad \square$$

Chapter 3

Limit of trees with fixed degree sequence

"Never spend more than a year on anything."

Jeff Ullman

This chapter is adapted from [28].

We use a new stick-breaking construction, also introduced independently by Addario-Berry, Donderwinkel, Maazoun, and Martin in [6], to study uniform rooted trees with fixed degree sequence \mathcal{D} (\mathcal{D} -trees). With this construction we prove, under natural conditions of convergence of the degree sequence, that \mathcal{D} -trees converge toward either \mathcal{P} -trees or ICRT. Alongside we upper-bound the height of \mathcal{D} -trees. We deduce similar results for \mathcal{P} -trees and ICRT. We also confirm a conjecture of Aldous, Miermont, and Pitman [14] stating that Lévy trees are ICRT with random parameters.

3.1 Introduction

Contents

3.1	Introduction	71
3.1.1	Overview of the results	72
3.1.2	Overview of the proof	72
3.1.3	Historical motivations.	73
3.2	Models and algorithms	75
3.2.1	\mathcal{D} -trees: the new stick breaking construction	75
3.2.2	\mathcal{P} -trees: construction and connection with \mathcal{D} -trees	77
3.2.3	ICRT: construction and connection with \mathcal{D} -trees	77
3.3	Notions of convergence	78
3.3.1	Stick-breaking topology	78
3.3.2	Gromov–Prokhorov (GP) topology	79
3.3.3	Gromov–Hausdorff (GH) topology	80
3.3.4	Gromov–Hausdorff–Prokhorov (GHP) topology	80
3.3.5	Measurability	81
3.4	Main results	81
3.4.1	Convergence of the first branches	81
3.4.2	Gromov–Prokhorov convergence	81
3.4.3	Gromov–Hausdorff–Prokhorov convergence	82
3.4.4	Height of \mathcal{D} -trees, \mathcal{P} -trees and ICRT	84

3.5	Convergence of the first branches of \mathcal{D}-trees	85
3.5.1	Convergence of \mathcal{D} -trees toward \mathcal{P} -trees: proof of Theorem 3.4.1 (a)	85
3.5.2	Preliminaries: \mathcal{D} -trees and \mathcal{P} -trees can be seen as modifications of ICRT	85
3.5.3	Convergence of \mathcal{D} -trees toward ICRT: proof of Theorem 3.4.1 (b)	87
3.6	Gromov–Prokhorov convergence of \mathcal{D}-trees	90
3.6.1	Preliminary: vertices with small degree behaves like leaves	90
3.6.2	Convergence of \mathcal{D} -trees toward ICRT: proof of Theorem 3.4.2 (b)	91
3.7	GHP convergence and height of \mathcal{D}-trees	92
3.7.1	Preliminaries: purely technical results on μ and $\tilde{\mu}$	92
3.7.2	Height of \mathcal{D} -trees: Proof of Theorem 3.4.4 (a)	95
3.7.3	GHP convergence of \mathcal{D} -trees: Proof of Theorem 3.4.3 (a)	100
3.8	Additional remarks	102
3.8.1	The space $\Omega := \Omega_{\mathcal{D}} \cup \Omega_{\mathcal{P}} \cup \Omega_{\Theta}$, its topology and applications	102
3.8.2	Re-rooting principle for ICRT	104
3.9	Appendix	106
3.9.1	Proof that Algorithm 3.1 and Algorithm 3.6 sample \mathcal{D} -trees	106
3.9.2	GP and GH convergence imply GHP convergence: Proof of Lemma 3.3.3	106

3.1.1 Overview of the results

Let $\{V_i\}_{i \in \mathbb{N}}$ be a set of vertices. A rooted tree T have degree sequence $(d_i)_{1 \leq i \leq s}$ if T has vertices $\{V_i\}_{1 \leq i \leq s}$ and for every $1 \leq i \leq s$, V_i has d_i children.

The aim of this paper is to study \mathcal{D} -trees: uniform rooted trees with fixed degree sequence \mathcal{D} . Notably we show, under natural conditions of convergence of the degree sequence, that \mathcal{D} -trees converge either toward \mathcal{P} -trees, or after normalisation by $s/\sqrt{\sum_{i=1}^s d_i(d_i - 1)}$ toward ICRT, for the Gromov–Prokhorov (GP) topology. Furthermore we show, under a given tightness assumption, that \mathcal{D} -trees also converge toward ICRT for the Gromov–Hausdorff–Prokhorov (GHP) topology. Alongside we provide a near optimal upper bound for the height of \mathcal{D} -trees.

Those results are in direct continuity with the work of Aldous, Camarri, Pitman [16, 40], which proves that \mathcal{P} -trees converge toward ICRT for the GP topology. We also complete their results by considering the GHP topology, and by proving an upper bound for the height of \mathcal{P} -trees and ICRT.

When taken together, those results on \mathcal{D} -trees, \mathcal{P} -trees and ICRT can be used to study the different models of trees that are uniform when conditioned to their exact degree sequence (see the end of the introduction and Section 3.8.1 for more details).

Finally, this paper is simplified thanks to the results of a companion paper (Chapter 2), which introduces and studies a new construction for ICRT. Notably, the probability measure introduced in Chapter 2 and the compactness of ICRT avoid us many computations and topological arguments.

3.1.2 Overview of the proof

Our approach relies on a new construction for \mathcal{D} -trees, which can be seen in three different ways: as a modification of Aldous–Broder algorithm, as a recursive construction of subtrees spanned by specific leaves, and as a stick-breaking construction. Our proof is strongly based on those points of view so let us recall them: (see also Figure 3.1 for a construction)

- (a) *Aldous–Broder algorithm*: Fix an arbitrary random walk $(A_i)_i$. Start with a single vertex A_1 , then recursively add the edge $\{A_{i-1}, A_i\}$ when A_i is "new", that is when $A_i \notin \{A_j\}_{j < i}$. It is well known, since its introduction in [9, 35], that this algorithm yields a random tree.

Along the construction, we add an edge between A_{i-1} and a new leaf when A_i is not "new". By doing so, the number of children of a vertex equals the number of times it is visited. Moreover we show that if $(A_i)_i$ is uniform among all \mathcal{D} -tuple, that is tuple such that for every $i \in \mathbb{N}$ V_i appears d_i times, then this algorithm constructs a \mathcal{D} -tree.

- (b) *Subtrees spanned by specific leaves*: Note that the last algorithm, when stopped at the k^{th} repetition (k^{th} index such that $A_i \in \{A_j\}_{j < i}$) constructs a subtree spanned by the root and k leaves. In fact the labels of those k leaves can be fixed in advance: In other words, the stopped algorithm constructs the subtree of a \mathcal{D} -tree spanned by the root and k fixed leaves. This allows us to study the distance matrix between random leaves, and hence, by definition, the geometry of a \mathcal{D} -tree in a Gromov–Prokhorov sense.
- (c) *Aggregation of paths / Stick breaking construction*: Note that the previous random walk $(A_i)_i$ starts at the root A_1 then follows the minimal path toward the first leaf, then "jumps" to follow the minimal path between a "repeated" vertex and the second leaf and so on... Hence, one can see the previous spanning tree between the root and the k^{th} first leaves as an aggregation of k minimal paths. A better way of understanding this aggregation of paths is through a stick-breaking construction: First consider $(A_i)_i$ as a "line of distinct vertices". Then recall that there is no edges $\{A_{i-1}, A_i\}$ when A_i is not "new", so "cut" those edges. This gives you a collection of paths or "sticks". Finally identify the vertices which are equal, or equivalently "glue" the paths at the repeated vertices. So the previous aggregation is fully understood through the positions of the "cuts" and the positions of the "glue-points".

To sum up, we use a modification of Aldous–Broder algorithm to construct \mathcal{D} -trees from some random tuples. Through a study of repetitions in those tuples, we study this algorithm as a stick-breaking construction. In particular, we prove the convergence of the "cuts" and the "glue-points". We then use this convergence to prove the convergence of the subtrees spanned by specific leaves, which turns out to be exactly what is needed to prove the GP convergence of \mathcal{D} -trees.

The idea of using an Aldous–Broder algorithm to study the geometry of a random tree is not new and goes back to the very study of the CRT by Aldous [10]. It is also used by Aldous, Camarri, Pitman in [16, 40] to prove that \mathcal{P} -trees converge toward ICRT for the GP topology. Here such an approach is possible thanks to our new construction for \mathcal{D} -trees. Furthermore its strong similarity with \mathcal{P} -trees construction explains the similarity between \mathcal{D} -trees, \mathcal{P} -trees and ICRT.

Finally, we deduce the GHP convergence from the GP and GH convergence (see Lemma 3.3.3). To prove the GH convergence, we prove that, in addition of the convergence of the first branches, the whole tree is close from its first branches. To do so, we consider a sequence of trees $(T_l)_{l \geq 0}$, corresponding to the different steps of our construction, and we upper bound for a well-chosen sequence (l_n) , the sum $\sum_n d_H(T_{l_n}, T_{l_{n+1}})$.

We deduce our upper bound for the height of \mathcal{D} -trees from the same estimates.

3.1.3 Historical motivations.

Since the pioneer work of Aldous [10], scaling limits of random discrete trees are at the center of many studies. The aim of such a study is to find an algorithm to construct a given model, and to show that, while the number of vertices diverges, the algorithm "converges" (in a suitable sense). By doing so, one can prove that the corresponding discrete tree converges, after proper scaling, toward a limit tree constructed from the limit algorithm, in order to study geometric properties (height, diameter, ...) of the discrete tree from the limit tree.

The main strength of this approach is to give a universal point of view on several models at the same time. Indeed, if several models have the same limit, then they all share similar properties.

This approach becomes even stronger when one considers that there may be several constructions for each model and that each construction may be a way to study all the connected models.

Hence, \mathcal{D} -trees, as essential "building blocks", are prime tools for the study of many other models. Let us briefly present some of the main models connected with \mathcal{D} -trees (see also Broutin, Marckert [38] for some detailed applications of \mathcal{D} -trees in the finite variance case).

- *Galton–Watson trees and Lévy trees:* Since Aldous [11], scaling limit of Galton–Watson trees have generated a large amount of literature due to the many connections with random walks, Lévy processes, superprocesses (see Le Gall [61, 60] for an introduction). The most important results are due to Duquesne Le Gall, [50, 49], who prove that Galton–Watson trees converge toward Lévy trees and study some geometric properties of Lévy trees (see also Le Gall, Le Jan [64, 63]).

Galton–Watson trees can be seen as \mathcal{D} -trees with random degree sequence \mathcal{D} . Hence, provided some preliminary estimates for the degree distribution, our results implies the convergence of many conditioned Galton–Watson trees (see Section 3.8.1). We do not prove such estimates, since they are already proved in most studies. However given the results of [61, 60], we can confirm the conjecture of Aldous, Miermont, Pitman [14] stating that Lévy trees are ICRT with random parameters. This explains why Lévy trees [61, 60] and ICRT (Chapter 2) have similar geometries.

Finally our stick-breaking construction is a new tool for Galton–Watson trees. In particular, until now, there has been a strong interest in several random objects build from those trees. Such objects may often be built by following the same construction, hence providing a powerful tool for their studies. As an example, some forthcoming notes [25] will use our construction to study height process, Lukasiewicz walk and snake of \mathcal{D} -trees and ICRT. The study of those processes are motivated by random planar maps with fixed degree sequence (see Marzouk [68]).

- *Additive coalescent:* Coalescent processes arise naturally in many sciences (see Bertoin [22] or Berestycki [21]). A classical way of studying those processes is to embed them into genealogical trees. Sadly, in general, little is known on those trees and many questions remain open.

Still, there is one unique model, the additive coalescent, who is far more easier to study than every other coalescent due to his many properties. The corresponding trees, \mathcal{P} -trees (biased by the root), and their scaling limits, ICRT, are thoroughly studied in [16, 40], Chapter 2, and their geometries are already well understood thanks to their stick breaking constructions.

Despite \mathcal{P} -trees may be seen as \mathcal{D} -trees with random degree sequence \mathcal{D} , studying them as such makes little sense. Indeed \mathcal{P} -trees stick-breaking construction is easier to study than \mathcal{D} -trees. Instead, we provide a new description of \mathcal{P} -trees as "degenerate" \mathcal{D} -trees. This description gives a new and fresh point of view on many of the strong "combinatorial" properties of \mathcal{P} -trees and ICRT, and is a powerful tool to prove new ones (see for instance Section 3.8.2).

- *Multiplicative graphs, and the configuration model:* Recently, those models of graphs have interested many computer scientists as natural generalizations of Erdős Renyi graph, which seems closer to real life network thanks to the "inhomogeneity in their degree distribution" (see e.g. [72]).

Without entering in the details, those models are "graph versions" of \mathcal{P} -trees and \mathcal{D} -trees. Notably, one can construct and study those graphs from \mathcal{D} -trees and \mathcal{P} -trees (see e.g. [37, 36] on multiplicative graph, and [54, 47, 42] on configuration model). Hence, our results on \mathcal{P} -trees and \mathcal{D} -trees should help at improving the current understanding of those graphs, and in particular they should help at removing the arbitrary and omnipresent randomness in the degree sequence, which exists only to avoid some technical issues. (See e.g. Chapter 4.)

Moreover our upper bound for the height of \mathcal{D} -trees and \mathcal{P} -trees should help at proving upper bounds for the height of those graphs (see e.g. Safsafi thesis [75]). Such an upper bound has been one of the main missing ingredients for a precise study of Prim Algorithm in all generality (see e.g. Addario-Berry, Broutin, Goldschmidt, Miermont [5] for the "Erdős Renyi case").

Plan of the paper: In Section 3.2, we present some constructions for \mathcal{D} -trees, \mathcal{P} -trees and ICRT. In Section 3.3 we introduce the different topologies that we are using in this paper (SB, GP, GHP). Our main results are stated in Section 3.4. Their proof are carried in Section 3.5 (SB), Section 3.6 (GP), and Section 3.7 (GHP and height). Finally we discuss some applications of our main results in Section 3.8: we prove in Section 3.8.1 that Lévy trees are ICRT with random parameters, and we introduce in Section 3.8.2 a new computation tool for ICRT: the re-rooting principle.

Notations: Throughout the paper, similar variables for \mathcal{D} -trees, \mathcal{P} -trees, Θ -ICRT share similar notations. To avoid any ambiguity, the models that we are using and their parameters are indicated by superscripts $\mathcal{D}_n, \mathcal{P}_n, \Theta_n$. We often drop those superscripts when the context is clear. We use the letter d for both degrees and distances. From the context, it is always clear which one it refers to.

3.2 Models and algorithms

3.2.1 \mathcal{D} -trees: the new stick breaking construction

Recall that a sequence (d_1, \dots, d_s) is a degree sequence of a tree if and only if $\sum_{i=1}^s d_i = s - 1$, and by convention $d_1 \geq d_2 \geq \dots \geq d_s$. Let $\Omega_{\mathcal{D}}$ be the set of such sequences.

Let us present our construction for \mathcal{D} -tree. For simplicity, we use the next conventions: For every graph $G = (V, E)$ and edge $e = \{v_1, v_2\}$, $G \cup e$ denotes the graph $(V \cup \{v_1, v_2\}, E \cup \{e\})$. We say that a vertex $v \in G$ if $v \in V$. Also, for every $\mathcal{D} \in \Omega_{\mathcal{D}}$ we let $L_1^{\mathcal{D}}, L_2^{\mathcal{D}}, \dots$ be the leaves (that is the vertices V_{a_1}, V_{a_2}, \dots with $a_1 \leq a_2 \leq \dots$ and $d_{a_1} = d_{a_2} = \dots = 0$).

Algorithm 3.1. Stick-breaking construction of a \mathcal{D} -tree (see Figure 3.1).

- Let $A^{\mathcal{D}} = (A_i^{\mathcal{D}})_{1 \leq i \leq s-1}$ be a uniform \mathcal{D} -tuple (tuple such that $\forall i \in \mathbb{N}$, V_i appears d_i times).
- Let $T_1^{\mathcal{D}} := (\{A_1\}, \emptyset)$ then for every $2 \leq i \leq s$ let

$$T_i^{\mathcal{D}} := \begin{cases} T_{i-1} \cup \{A_{i-1}, A_i\} & \text{if } A_i \notin T_{i-1}. \\ T_{i-1} \cup \{A_{i-1}, L_{\inf\{k, L_k \notin T_{i-1}\}}\} & \text{if } A_i \in T_{i-1} \text{ or } i = s. \end{cases}$$

- Let $T^{\mathcal{D}}$ denotes the rooted tree (T_s, A_1) .

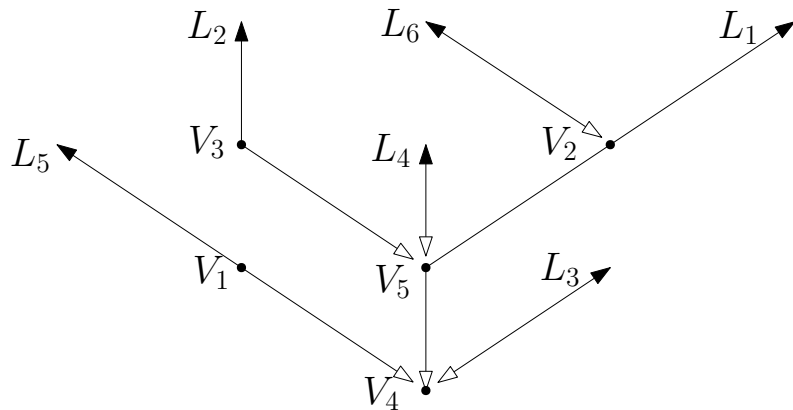


Figure 3.1: Stick breaking construction of a \mathcal{D} -tree with $\mathcal{D} = (1, 2, 1, 3, 3, 0, 0, \dots)$ and $(A_i^{\mathcal{D}})_{1 \leq i \leq s-1} = (V_4, V_5, V_2, V_5, V_3, V_4, V_5, V_4, V_1, V_2)$. The exploration starts at V_4 then follows the white-black arrow toward L_1 , then jumps at V_5 to follow the path toward L_2 and so on...

Proposition 3.2.1. *For every $\mathcal{D} \in \Omega_{\mathcal{D}}$, $T^{\mathcal{D}}$ is a \mathcal{D} -tree.*

Proof. See Appendix. □

Remarks. (a) Addario-Berry, Donderwinkel, Maazoun, and Martin [6] independently found the same construction while I was writing the article.

(b) We introduce later a generalisation of Algorithm 3.1, Algorithm 3.6, which will be used to simplify many of our computations.

(c) One might prefer to consider "pure degree sequence": A tree T have pure degree sequence $(d_i)_{1 \leq i \leq s}$ if T have vertices $(V_i)_{1 \leq i \leq s}$, and for every $1 \leq i \leq s$, V_i have degree d_i (not d_i children). From a tree T with pure degree sequence (d_1, \dots, d_s) with $d_s = 1$, one can construct a tree with degree sequence $(d_1 - 1, \dots, d_{s-1} - 1)$ by rooting T at the unique vertex adjacent to V_s and then removing V_s . This is a well known bijection between tree with pure degree sequence (d_1, \dots, d_s) and tree with degree sequence $(d_1 - 1, \dots, d_{s-1} - 1)$. Since this bijection only removes one leaf, it does not change much the geometry of the tree. So our results on \mathcal{D} -trees easily adapt to uniform trees with fixed pure degree sequence.

As claimed in the introduction 3.1.2 (d), we are interested in the stick-breaking point of view to prove the convergence of \mathcal{D} -trees. To this end, let us formally define the cuts and the glue points. Note that there is $N := \sum_{i=1}^s (d_i - 1) \mathbf{1}_{d_i \geq 1}$ indexes Y_1, \dots, Y_N such that $A_i \in \{A_1, \dots, A_{i-1}\}$. Also, for every $1 \leq i \leq N$, let Z_i denote the smallest integer z such that $A_z = A_{Y_i}$. We call $\{Y_i\}_{1 \leq i \leq N}$ the cuts, and $\{Z_i\}_{1 \leq i \leq N}$ the glue points.

Let us introduce some new tools to study those processes. First, note that the cuts and the glue points are encoded by "when" and "where" there are some "repetitions" in $A^{\mathcal{D}}$. Therefore, it is natural to look at the law of A_i conditioned on $(A_j)_{1 \leq j < i}$. In particular, since $A^{\mathcal{D}}$ is a uniform \mathcal{D} -tuple, note that for every $1 \leq i \leq s - 1$ and $1 \leq a \leq s$:

$$\mathbb{P}(A_i = V_a | (A_j)_{1 \leq j < i}) = \frac{d_a - \#\{1 \leq j < i : A_j = V_a\}}{s - i}. \quad (3.1)$$

Hence, we have for every $j < i \leq s - 1$,

$$\mathbb{P}((i, j) \in \{(Y_l, Z_l)\}_{l \in \mathbb{N}} | (A_k)_{1 \leq k < i}) = \mathbf{1}_{A_j \notin \{A_k\}_{1 \leq k < j}} \frac{d_{A_j} - \#\{1 \leq k < i : A_k = A_j\}}{s - i}. \quad (3.2)$$

To fully understand the behavior of the cuts and the glue points, let us rewrite (3.2). Since we are mostly interested in the first cuts, note that typically $\#\{1 \leq k < i : A_k = A_j\} = 1$ and so we have the following informal approximation

$$\mathbb{P}((i, j) \in \{(Y_l, Z_l)\}_{l \in \mathbb{N}} | (A_k)_{1 \leq k < i}) \approx \frac{\mu(j)}{s - i}, \quad (3.3)$$

where μ is the real measure defined by

$$\mu := \sum_{j=1}^{s-1} \delta_j \mathbf{1}_{A_j \notin \{A_k\}_{1 \leq k < j}} (d_{A_j} - 1). \quad (3.4)$$

As a result, the cuts and glue points are mostly encoded by this measure μ . Finally to understand the behavior of μ , let us rewrite (3.4). Let for every $i \in \mathbb{N}$,

$$X_i := \inf\{1 \leq j \leq s, A_j = V_i\}.$$

(by convention $X_i = \infty$ and $\delta_\infty = 0$ when there is no such index) and note that

$$\mu = \sum_{i=1}^s \delta_{X_i} (d_i - 1). \quad (3.5)$$

So μ can be directly studied through $(X_i)_{1 \leq i \leq s}$.

Finally we need to rescale \mathcal{D} -trees. So let for $\mathcal{D} \in \Omega_{\mathcal{D}}$, $(\sigma^{\mathcal{D}})^2 := \sum_{i=1}^s d_i (d_i - 1)$.

3.2.2 \mathcal{P} -trees: construction and connection with \mathcal{D} -trees

Fix $(\mathcal{D}_n)_{n \in \mathbb{N}}$ a sequence in $\Omega_{\mathcal{D}}$ and assume that for every $i \in \mathbb{N}$, $d_i^{\mathcal{D}_n}/s^{\mathcal{D}_n} \rightarrow p_i$ as $n \rightarrow \infty$. Note that $p_\infty := 1 - \sum_{i=1}^{\infty} p_i \geq 0$. We suppose further that $p_1 > 0$ and let $\mathcal{P} := (p_i)_{i \in \mathbb{N} \cup \{\infty\}}$.

In this regime $(\mathcal{D}_n)_{n \in \mathbb{N}}$ converges: Let V_∞ be an additional vertex and put a metric on $\{V_i\}_{i \in \mathbb{N} \cup \{\infty\}}$ such that $V_i \rightarrow V_\infty$ as $i \rightarrow \infty$. One can easily check that $(\mathcal{D}_n)_{n \in \mathbb{N}}$ converges weakly toward $(A_i^{\mathcal{P}})_{i \in \mathbb{N}}$ as $n \rightarrow \infty$ where $(A_i^{\mathcal{P}})_{i \in \mathbb{N}}$ is a family of i.i.d. random variables such that for every $i \in \mathbb{N} \cup \{\infty\}$, $\mathbb{P}(A_1^{\mathcal{P}} = V_i) = p_i$.

Hence, intuitively, the steps of Algorithm 3.1, which are encoded by those tuples, "converge". Therefore \mathcal{D}_n -trees should converge toward the tree $T^{\mathcal{P}}$ defined by the "limit algorithm":

Let $\Omega_{\mathcal{P}}$ be the set of sequence $(p_i)_{i \in \mathbb{N} \cup \{\infty\}}$ in \mathbb{R}^+ such that $\sum_{i=1}^{\infty} p_i + p_\infty = 1$, $p_1 > 0$ and $p_1 \geq p_2 \geq \dots$. For every $\mathcal{P} \in \Omega_{\mathcal{P}}$, the \mathcal{P} -tree is the tree constructed as follows:

Algorithm 3.2. Definition of the \mathcal{P} -tree for $\mathcal{P} \in \Omega_{\mathcal{P}}$.

- Let $(A_i^{\mathcal{P}})_{i \in \mathbb{N}}$ be a family of i.i.d. random variables such that for all $i \in \mathbb{N}$, $\mathbb{P}(A_1^{\mathcal{P}} = V_i) = p_i$.
- For every $i \in \mathbb{N}$, let $B_i^{\mathcal{P}} = A_i$ if $A_i \in \mathbb{N}$, and let $B_i^{\mathcal{P}} = V_{\infty, i}$ otherwise.
- Let $T_1^{\mathcal{P}} := (\{B_1\}, \emptyset)$ then for every $i \geq 2$ let

$$T_i^{\mathcal{P}} := \begin{cases} T_{i-1} \cup \{B_{i-1}, B_i\} & \text{if } B_i \notin T_{i-1}. \\ T_{i-1} & \text{if } B_i \in T_{i-1}. \end{cases} \quad (3.6)$$

- Let $T^{\mathcal{P}}$ denote the rooted tree $(\bigcup_{n \in \mathbb{N}} T_n, B_1)$.

Remarks. (a) We choose to delete the leaves $\{L_i\}_{i \in \mathbb{N}}$ of $T^{\mathcal{P}}$ in (3.6) to avoid measurability issue. (b) Note that the infinite vertex V_∞ is split into many others. Indeed this vertex "fills the gap" produced by vertices with small degree. The introduction of V_∞ allows us to consider a slightly more general definition than the one introduced in [40] which requires $\sum_{i=1}^{\infty} p_i = 1$.

Finally let us introduce a list of notations similar to those for \mathcal{D} -trees. For every $\mathcal{P} \in \Omega_{\mathcal{P}}$:

- Let $p_1^{\mathcal{P}}, \dots, p_\infty^{\mathcal{P}} \in \mathbb{R}^+$ be such that $\mathcal{P} = (p_1^{\mathcal{P}}, \dots, p_\infty^{\mathcal{P}})$.
- Let $s^{\mathcal{P}} := \max\{i \in \mathbb{N} \cup \{\infty\}, p_i > 0\}$ and let $(\sigma^{\mathcal{P}})^2 = \sum_{i=1}^{\infty} (p_i)^2$.
- Let $(Y_i^{\mathcal{P}})_{i \in \mathbb{N}}$ be the indexes of repetitions in $(B_i)_{i \in \mathbb{N}}$.
- For $i \in \mathbb{N}$, let $Z_i^{\mathcal{P}} := \inf\{z, B_z = B_{Y_i}\}$.
- For $i \in \mathbb{N}$, let $X_i^{\mathcal{P}} = \inf\{j \in \mathbb{N}, A_j = i\}$. Then let $\mu^{\mathcal{P}} = \sum_{i=1}^{\infty} p_i \delta_{X_i}$.

3.2.3 ICRT: construction and connection with \mathcal{D} -trees

Fix $(\mathcal{D}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{D}}^{\mathbb{N}}$ and assume that $d_1^{\mathcal{D}_n}/s^{\mathcal{D}_n} \rightarrow \infty$ and that $\forall i \in \mathbb{N}$, $d_i^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n} \rightarrow \theta_i$ as $n \rightarrow \infty$. Note that $\sum_{i=1}^{\infty} \theta_i^2 \leq 1$ and let $\theta_0 := 1 - \sum_{i=1}^{\infty} \theta_i^2$. Then let $\Theta := (\theta_i)_{i \in \{0\} \cup \mathbb{N}}$.

Let us briefly explain why \mathcal{D}_n -trees converge in this regime. We stay informal here, and details are done in Section 3.5.3. First some quick estimates show that $(X_i^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n})_{i \in \mathbb{N}}$ converge weakly toward some exponential random variables $(X_i^{\Theta})_{i \in \mathbb{N}}$. By (3.5), this implies that after proper rescaling $(\mu^{\mathcal{D}_n})_{n \in \mathbb{N}}$ should converges weakly toward some random measure μ^{Θ} . Finally, by (3.3) the cuts and the glue points are "encoded" by $\mu^{\mathcal{D}_n}$, so they should also converge after rescaling.

Those estimates suggest that \mathcal{D}_n -trees should converge after rescaling toward some limit tree. If one pursue the computation further, then one notice that this tree is in fact the Θ -ICRT.

We now construct the Θ -ICRT. First let us introduce a generic stick breaking construction. It takes for input two sequences in \mathbb{R}^+ called cuts $\mathbf{y} = (y_i)_{i \in \mathbb{N}}$ and glue points $\mathbf{z} = (z_i)_{i \in \mathbb{N}}$, which satisfy

$$\forall i < j, \quad y_i < y_j \quad ; \quad y_i \longrightarrow \infty \quad ; \quad \forall i \in \mathbb{N}, \quad z_i \leq y_i,$$

and creates an \mathbb{R} -tree (loopless geodesic metric space) by recursively "gluing" segment $(y_i, y_{i+1}]$ on position z_i (see Figure 3.2), or rigorously, by constructing a consistent sequence of distances $(d_n)_{n \in \mathbb{N}}$ on $([0, y_n])_{n \in \mathbb{N}}$.

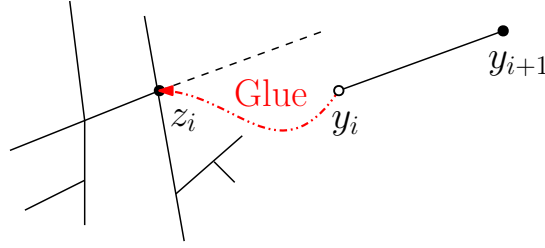


Figure 3.2: A typical step of the stick-breaking construction: the "gluing" of $(y_i, y_{i+1}]$ at z_i .

Algorithm 3.3. Generic stick-breaking construction of \mathbb{R} -tree.

- Let d_0 be the trivial metric on $[0, 0]$.
- For each $i \geq 0$ define the metric d_{i+1} on $[0, y_{i+1}]$ such that for each $x \leq y$:

$$d_{i+1}(x, y) := \begin{cases} d_i(x, y) & \text{if } x, y \in [0, y_i] \\ d_i(x, z_i) + |y - y_i| & \text{if } x \in [0, y_i], y \in (y_i, y_{i+1}] \\ |x - y| & \text{if } x, y \in (y_i, y_{i+1}], \end{cases}$$

where by convention $y_0 := 0$ and $z_0 := 0$.

- Let d be the unique metric on \mathbb{R}^+ which agrees with d_i on $[0, y_i]$ for each $i \in \mathbb{N}$.
- Let $\text{SB}(\mathbf{y}, \mathbf{z})$ be the completion of (\mathbb{R}^+, d) .

Now, let Ω_Θ be the space of sequences $(\theta_i)_{i \in \{0\} \cup \mathbb{N}}$ in \mathbb{R}^+ such that $\sum_{i=0}^\infty \theta_i^2 = 1$ and such that $\theta_1 \geq \theta_2 \geq \dots$. For every $\Theta \in \Omega_\Theta$, the Θ -ICRT is the random \mathbb{R} -tree constructed as follows:

Algorithm 3.4. Construction of Θ -ICRT (from Section 2.2)

- Let $(X_i)_{i \in \mathbb{N}}$ be a family of independent exponential random variables of parameter $(\theta_i)_{i \in \mathbb{N}}$.
- Let μ be the measure on \mathbb{R}^+ defined by $\mu = \theta_0^2 dx + \sum_{i=1}^\infty \delta_{X_i} \theta_i$.
- Let $(Y_i, Z_i)_{i \in \mathbb{N}}$ be a Poisson point process on $\{(y, z) \in \mathbb{R}^{+2} : y \geq z\}$ of intensity $dy \times d\mu$.
- The Θ -ICRT is defined as $(\mathcal{T}, d) = \text{SB}((Y_i)_{i \in \mathbb{N}}, (Z_i)_{i \in \mathbb{N}})$. (see Algorithm 3.3)

Remarks. (a) Although one usually excludes the ICRT with $\theta_0 = 0$ and $\sum_{i=1}^\infty \theta_i < \infty$, there are some \mathcal{D} -trees which converge toward such ICRT. Furthermore those ICRT may be seen as \mathcal{P} -trees with a modified distance (see section 3.5.2).

(b) The term $\theta_0^2 dx$ in μ^θ encompass the fact that some cuts may occur "because" of vertices with small degree. The case where $\theta_0 = 1$ corresponds to the Brownian CRT.

3.3 Notions of convergence

The reader may wish to simply skim this section on a first reading, referring back to it as needed.

3.3.1 Stick-breaking topology

Let \mathbb{M} denote the set of locally finite positive Borel measure on \mathbb{R}^+ . Let $\mathbb{K}_{\text{SB}} := (\mathbb{R}^+)^{3\mathbb{N}} \times \mathbb{M}$. For every $\Lambda \in \Omega := \Omega_{\mathcal{D}} \cup \Omega_{\mathcal{P}} \cup \Omega_\Theta$ let

$$\Upsilon^\Lambda := ((X_i^\Lambda, Y_i^\Lambda, Z_i^\Lambda)_{i \in \mathbb{N}}, \mu^\Lambda) \in \mathbb{K}_{\text{SB}},$$

where by convention for $\mathcal{D} \in \Omega_{\mathcal{D}}$, and $i > N^{\mathcal{D}}$, $Y_i = Z_i = +\infty$.

We say that a sequence $((X_i^n, Y_i^n, Z_i^n)_{i \in \mathbb{N}}, \mu^n)_{n \in \mathbb{N}}$ converges toward $((X_i, Y_i, Z_i)_{i \in \mathbb{N}}, \mu)$ for the stick-breaking (SB) topology if and only if

- a) For every $i \in \mathbb{N}$, $X_i^n \rightarrow X_i$, $Y_i^n \rightarrow Y_i$, $Z_i^n \rightarrow Z_i$.
- b) For every $i, j \in \mathbb{N}$, $\mathbf{1}_{X_i^n = Z_j^n} \rightarrow \mathbf{1}_{X_i = Z_j}$.
- c) $\mu^n \rightarrow \mu$ for the Skorohod topology on \mathbb{M} (that is $l \mapsto \mu^n[0, l]$ converges toward $l \mapsto \mu[0, l]$).

Note that the SB topology on \mathbb{K}_{SB} is well defined and that \mathbb{K}_{SB} equipped with the SB topology is a Polish space (separable complete metrisable space) as a product of Polish spaces.

Remark. b) is here to "avoid fusion" between two branching points. Although (b) is not used in this paper, it is crucial for our study in [25]. The presence of $(X_i)_{i \in \mathbb{N}}$ is also motivated by [25].

Finally let us introduce some notations to rescale our trees on \mathbb{K}_{SB} . For every $\lambda_1, \lambda_2 > 0$, and $\mu \in \mathbb{M}$ let $(\lambda_1, \lambda_2)\mu \in \mathbb{M}$ such that for every Borel set B ,

$$(\lambda_1, \lambda_2)\mu(B) = \lambda_2 \mu\{\lambda_1 x, x \in B\}.$$

For every $\lambda_1, \lambda_2 > 0$ and $\Upsilon := \{(X_i, Y_i, Z_i)_{i \in \mathbb{N}}, \mu\} \in \mathbb{K}_{\text{SB}}$ let

$$(\lambda_1, \lambda_2)\Upsilon := ((\lambda_1 X_i, \lambda_1 Y_i, \lambda_1 Z_i)_{i \in \mathbb{N}}, (\lambda_1, \lambda_2)\mu).$$

Note that for every $\lambda_1, \lambda_2 > 0$, the map $\Upsilon \mapsto (f, \lambda)\Upsilon$ is continuous and so measurable.

3.3.2 Gromov–Prokhorov (GP) topology

A measured metric space is a triple (X, d, μ) such that (X, d) is a Polish space and μ is a Borel probability measure on X . Two such spaces (X, d, μ) , (X', d', μ') are called isometry-equivalent if and only if there exists an isometry $f : X \rightarrow X'$ such that if $f_*\mu$ is the image of μ by f then $f_*\mu = \mu'$. Let \mathbb{K}_{GP} be the set of isometry-equivalent classes of measured metric space. Given a measured metric space (X, d, μ) , we write $[X, d, \mu]$ for the isometry-equivalence class of (X, d, μ) and frequently use the notation X for either (X, d, μ) or $[X, d, \mu]$.

We now recall the definition of the Prokhorov's distance. Consider a metric space (X, d) . For every $A \subset X$ and $\varepsilon > 0$ let $A^\varepsilon := \{x \in X, d(x, A) < \varepsilon\}$. Then given two (Borel) probability measures μ, ν on X , the Prokhorov distance between μ and ν is defined by

$$d_P(\mu, \nu) := \inf\{\varepsilon > 0: \mu\{A\} \leq \nu\{A^\varepsilon\} \text{ and } \nu\{A\} \leq \mu\{A^\varepsilon\}, \text{ for all Borel set } A \subset X\}.$$

The Gromov–Prokhorov (GP) distance is an extension of the Prokhorov's distance: For every $(X, d, \mu), (X', d', \mu') \in \mathbb{K}_{\text{GP}}$ the Gromov–Prokhorov distance between X and X' is defined by

$$d_{\text{GP}}((X, d, \mu), (X', d', \mu')) := \inf_{S, \phi, \phi'} d_P(\phi_*\mu, \phi'_*\mu'),$$

where the infimum is taken over all metric spaces S and isometric embeddings $\phi : X \rightarrow S$, $\phi' : X' \rightarrow S$. d_{GP} is indeed a distance on \mathbb{K}_{GP} and $(\mathbb{K}_{\text{GP}}, d_{\text{GP}})$ is a Polish space (see e.g. [1]).

Remark. For stick breaking construction a natural choice for S is the completion of $(\mathbb{R}^{+\mathbb{N}}, |||_1)$ as one may simply use one direction for each branch (see Aldous [10] for more details).

We use another convenient characterization of the GP topology which relies on convergence of distance matrices: For every measured metric space (X, d^X, μ^X) let $(x_i^X)_{i \in \mathbb{N}}$ be a sequence of i.i.d. random variables of common distribution μ^X and let $M^X := (d^X(x_i^X, x_j^X))_{(i,j) \in \mathbb{N}^2}$. We have the following result from [66],

Lemma 3.3.1. *Let $(X^n)_{n \in \mathbb{N}} \in \mathbb{K}_{\text{GP}}^{\mathbb{N}}$ and let $X \in \mathbb{K}_{\text{GP}}$ then $X^n \xrightarrow{GP} X$ as $n \rightarrow \infty$ if and only if M^{X^n} converges in distribution toward M^X .*

For convenience issue we use the following extension of Lemma 3.3.1.

Lemma 3.3.2. *Let $(X^n)_{n \in \mathbb{N}} \in \mathbb{K}_{GP}^{\mathbb{N}}$ and let $X \in \mathbb{K}_{GP}$. Let $(y_i^X)_{i \in \mathbb{N}}$ be a sequence of random variables on X and let $N^X := (d^X(y_i^X, y_j^X))_{(i,j) \in \mathbb{N}^2}$. If*

$$M^{X^n} \xrightarrow{(d)} N^X \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^n \delta_{y_i^X} \xrightarrow{(d)} \mu^X,$$

then $X^n \xrightarrow{GP} X$ and thus M^X and N^X have the same distribution.

Proof. Fix $k \leq m \in \mathbb{N}$. Let (A_1, \dots, A_k) be a uniform tuple of k different integers in $\{1, \dots, m\}$. Since $M^{X^n} \xrightarrow{(d)} N^X$, we have

$$(d^{X^n}(x_i^{X^n}, x_j^{X^n}))_{1 \leq i, j \leq k} \xrightarrow{(d)} (d^{X^n}(x_{A_i}^{X^n}, x_{A_j}^{X^n}))_{1 \leq i, j \leq k} \xrightarrow{(d)} (d^X(y_{A_i}^X, y_{A_j}^X))_{1 \leq i, j \leq k}.$$

Now since as $m \rightarrow \infty$, $\frac{1}{m} \sum_{i=1}^m \delta_{y_i^X} \rightarrow \mu^X$, taking $m \rightarrow +\infty$ in the above equation yields

$$(d^{X^n}(x_i^{X^n}, x_j^{X^n}))_{1 \leq i, j \leq k} \xrightarrow{(d)} (d^X(x_i^X, x_j^X))_{1 \leq i, j \leq k}.$$

Finally, since k is arbitrary, Lemma 3.3.1 concludes the proof. \square

3.3.3 Gromov–Hausdorff (GH) topology

Let \mathbb{K}_{GH} be the set of isometry-equivalent classes of compact metric space. For every metric space (X, d) , we write $[X, d]$ for the isometry-equivalent class of (X, d) , and frequently use the notation X for either (X, d) or $[X, d]$.

For every metric space (X, d) , the Hausdorff distance between $A, B \subset X$ is given by

$$d_H(A, B) := \inf\{\varepsilon > 0, A \subset B^\varepsilon, B \subset A^\varepsilon\}.$$

The Gromov–Hausdorff distance between $(X, d), (X', d') \in \mathbb{K}_{GH}$ is given by

$$d_{GH}((X, d), (X', d')) := \inf_{S, \phi, \phi'} (d_H(\phi(X), \phi'(X'))),$$

where the infimum is taken over all metric spaces S and isometric embeddings $\phi : X \rightarrow S$, $\phi' : X' \rightarrow S$. d_{GH} is indeed a distance on \mathbb{K}_{GH} and $(\mathbb{K}_{GH}, d_{GH})$ is a Polish space. (see e.g. [1])

3.3.4 Gromov–Hausdorff–Prokhorov (GHP) topology

Let $\mathbb{K}_{GHP} \subset \mathbb{K}_{GP}$ be the set of isometry-equivalent classes of compact measured metric space. The Gromov–Hausdorff–Prokhorov distance between $(X, d, \mu), (X', d', \mu') \in \mathbb{K}_{GHP}$ is given by

$$d_{GHP}((X, d, \mu), (X', d', \mu')) := \inf_{S, \phi, \phi'} (d_P(\phi_*\mu, \phi'_*\mu') + d_H(\phi(X), \phi'(X'))),$$

where the infimum is taken over all metric spaces S and isometric embeddings $\phi : X \rightarrow S$, $\phi' : X' \rightarrow S$. d_{GHP} is indeed a distance on \mathbb{K}_{GHP} and $(\mathbb{K}_{GHP}, d_{GHP})$ is a Polish space. (see [1])

Note that GHP convergence implies GP convergence, then that random variables GHP measurable are also GH measurable. For every $[X, d, p] \in \mathbb{K}_{GHP}$, let $[X, d]$ denote its natural projection on \mathbb{K}_{GH} . Note that GHP convergence implies GH convergence of the projections on \mathbb{K}_{GH} , then that the projection on \mathbb{K}_{GH} is a measurable function. We will need the following statement.

Lemma 3.3.3. *Let $([X^n, d^n, p^n])_{n \in \mathbb{N}}$ and $[X, d, p]$ be GHP measurable random variables in \mathbb{K}_{GHP} . Assume that almost surely $[X, d, p]$ have full support. Assume that $([X^n, d^n, p^n])_{n \in \mathbb{N}}$ converges weakly toward $[X, d, p]$ in a GP sens, and that $([X^n, d^n])_{n \in \mathbb{N}}$ converges weakly toward $[X, d]$ in a GH sens. Then $([X, d, p])_{n \in \mathbb{N}}$ converges weakly toward $[X, d, p]$ in a GHP sens.*

Proof. See Appendix \square

3.3.5 Measurability

We briefly explain why \mathcal{D} -trees, \mathcal{P} -trees, and Θ -ICRT are measurable for the SB, GP, and GHP topologies. First \mathcal{D} -trees are measurable as discrete random variables. Then, it is easy to prove from the definitions of X, Y, Z, μ , that \mathcal{P} -trees and ICRT are SB measurable. Also, it is proved in [16, 40] that \mathcal{P} -trees and ICRT are GP measurable. Furthermore, \mathcal{P} -trees are GHP measurable, since we restrict to a finite discrete case.

The GHP measurability of compact Θ -ICRT is harder to prove so let us give a sketch of proof. Let for every $l > 0$, μ_l^Θ be the restriction of μ^Θ to $[0, l]$ and let $\mathcal{T}_l^\Theta := ([0, l], d^\Theta, \mu_l^\Theta / \mu^\Theta[0, l])$. One can show that, for every $l > 0$, \mathcal{T}_l^Θ is GHP measurable by proving the continuity of SB (Algorithm 3.3) from \mathbb{K}_{SB} to those subtrees. Then, by Theorems 2.3.1 and 2.3.3, almost surely $(\mathcal{T}_l^\Theta)_{l \in \mathbb{R}_+}$ converges toward \mathcal{T}^Θ for the GHP topology so \mathcal{T}^Θ is GHP measurable.

3.4 Main results

In the whole section $(\mathcal{D}_n)_{n \in \mathbb{N}}, (\mathcal{P}_n)_{n \in \mathbb{N}}, (\Theta_n)_{n \in \mathbb{N}}$ are fixed sequences in $\Omega_{\mathcal{D}}, \Omega_{\mathcal{P}}, \Omega_{\Theta}$ respectively. Furthermore we always work under one of the following regimes:

Assumption 3.1 ($\mathcal{D}_n \Rightarrow \mathcal{P}$). For all $i \geq 1$, $d_i^{\mathcal{D}_n} / s^{\mathcal{D}_n} \rightarrow p_i^{\mathcal{P}}$ and $s^{\mathcal{D}_n} \rightarrow \infty$.

Assumption 3.2 ($\mathcal{D}_n \Rightarrow \Theta$). For all $i \geq 1$, $d_i^{\mathcal{D}_n} / \sigma^{\mathcal{D}_n} \rightarrow \theta_i^\Theta$ and $d_1^{\mathcal{D}_n} / s^{\mathcal{D}_n} \rightarrow 0$.

Assumption 3.3 ($\mathcal{P}_n \Rightarrow \Theta$). For all $i \geq 1$, $p_i^{\mathcal{P}_n} / \sigma^{\mathcal{P}_n} \rightarrow \theta_i^\Theta$ and $p_1^{\mathcal{P}_n} \rightarrow 0$.

Assumption 3.4 ($\Theta_n \Rightarrow \Theta$). For all $i \geq 1$, $\theta_i^{\Theta_n} \rightarrow \theta_i^\Theta$.

A few words on \Rightarrow . In most of the paper \Rightarrow is only used as a notation. However, one can put a topology on $\Omega := \Omega_{\mathcal{D}} \cup \Omega_{\mathcal{P}} \cup \Omega_{\Theta}$ such that \Rightarrow corresponds with the notion of convergence on Ω . Let us briefly explain the main advantages of this approach (see Section 3.8.1 for more details).

First, one can check that (Ω, \Rightarrow) is a Polish space, so one can do probability on this space. Moreover, one can see our results as continuity results for the function which associate to a set of parameters a tree. Hence, our results can be used to study trees with random degree distributions.

Furthermore one can check that $\Omega_{\mathcal{D}}$ is dense, so our results on \mathcal{D} -trees implies all the others.

3.4.1 Convergence of the first branches

Intuitively, the next technical result on Algorithms 3.1, 3.2, 3.4 tells that the first branches converge. It is absolutely central for this paper (recall Introduction 3.1.2) and also for the forthcoming [25].

Theorem 3.4.1. *The following convergences hold weakly for the SB topology (see Section 3.3.1).*

- (a) Suppose that $\mathcal{D}_n \Rightarrow \mathcal{P}$, then $(1, 1/s^{\mathcal{D}_n}) \Upsilon^{\mathcal{D}_n} \xrightarrow{\text{WSB}} \Upsilon^{\mathcal{P}}$.
- (b) Suppose that $\mathcal{D}_n \Rightarrow \Theta$, then $(\sigma^{\mathcal{D}_n} / s^{\mathcal{D}_n}, 1/\sigma^{\mathcal{D}_n}) \Upsilon^{\mathcal{D}_n} \xrightarrow{\text{WSB}} \Upsilon^\Theta$.
- (c) Suppose that $\mathcal{P}_n \Rightarrow \Theta$, then $(\sigma^{\mathcal{P}_n}, 1/\sigma^{\mathcal{P}_n}) \Upsilon^{\mathcal{P}_n} \xrightarrow{\text{WSB}} \Upsilon^\Theta$.
- (d) Suppose that $\Theta_n \Rightarrow \Theta$, then $\Upsilon^{\Theta_n} \xrightarrow{\text{WSB}} \Upsilon^\Theta$.

3.4.2 Gromov–Prokhorov convergence

First let us specify the measures that we consider. Let $\Omega_{\mathcal{M}}$ be the set of measures on $\{V_i\}_{i \in \mathbb{N}}$. We say that a sequence $(\mathcal{M}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{M}}^{\mathbb{N}}$ converges toward $\mathcal{M} \in \Omega_{\mathcal{M}}$ iff $\max_{i \in \mathbb{N}} |\mathcal{M}_n(V_i) - \mathcal{M}(V_i)| \rightarrow 0$. In the whole paper, for every $\mathcal{D} \in \Omega_{\mathcal{D}}$, $\mathcal{M}^{\mathcal{D}}$ is a fixed probability measure with support on $\mathcal{V}^{\mathcal{D}} := \{V_i, 1 \leq i \leq s^{\mathcal{D}}\}$. Similarly, for every $\mathcal{P} \in \Omega_{\mathcal{P}}$, $\mathcal{M}^{\mathcal{P}}$ is a fixed probability measure with support on $\mathcal{V}^{\mathcal{P}} := \{V_i^{\mathcal{P}}\}_{i: p_i > 0}$. Also, we sometimes let 0 denote the null measure on $\{V_i\}_{i \in \mathbb{N}}$.

Then we recall the probability measure on ICRT introduced in Chapter 2. For convenient issue, we write $\mu^\Theta = \infty$ when either $\theta_0^\Theta > 0$ or $\sum_{i=1}^\infty \theta_i^\Theta = \infty$, (since $\mu^\Theta = \infty$ iff a.s. $\mu^\Theta[0, \infty] = \infty$).

Definition (Theorem 2.3.1). *Let $\Theta \in \Omega_\Theta$ be such that $\mu^\Theta = \infty$. For every $l > 0$, let μ_l^Θ be the restriction of μ^Θ on $[0, l]$ and let $\mathbf{p}_l^\Theta := \mu_l^\Theta / \mu^\Theta[0, l]$. Almost surely there exists a probability measure \mathbf{p}^Θ on \mathcal{T}^Θ such that $(\mathbf{p}_l^\Theta)_{l>0}$ converges weakly toward \mathbf{p}^Θ .*

Let us now state the main result of this section. In what follows, $d^\mathcal{D}$ denotes the graph distance on $T^\mathcal{D}$ and similarly $d^\mathcal{P}$ denotes the graph distance on $T^\mathcal{P}$.

Theorem 3.4.2. *The following convergences hold weakly for the GP topology (see Section 3.3.2).*

(a) *If $\mathcal{D}_n \Rightarrow \mathcal{P}$ and $\mathcal{M}^{\mathcal{D}_n} \rightarrow \mathcal{M}^\mathcal{P}$ then*

$$(\mathcal{V}^{\mathcal{D}_n}, d^{\mathcal{D}_n}, \mathcal{M}^{\mathcal{D}_n}) \xrightarrow{WGP} (\mathcal{V}^\mathcal{P}, d^\mathcal{P}, \mathcal{M}^\mathcal{P}).$$

(b) *If $\mathcal{D}_n \Rightarrow \Theta$, $\mathcal{M}^{\mathcal{D}_n} \rightarrow 0$, and $\mu^\Theta = \infty$ then*

$$(\mathcal{V}^{\mathcal{D}_n}, (\sigma^{\mathcal{D}_n} / s^{\mathcal{D}_n}) d^{\mathcal{D}_n}, \mathcal{M}^{\mathcal{D}_n}) \xrightarrow{WGP} (\mathcal{T}^\Theta, d^\Theta, \mathbf{p}^\Theta).$$

(c) *If $\mathcal{P}_n \Rightarrow \Theta$, $\mathcal{M}^{\mathcal{P}_n} \rightarrow 0$, and $\mu^\Theta = \infty$ then*

$$(\mathcal{V}^{\mathcal{P}_n}, \sigma^{\mathcal{P}_n} d^{\mathcal{P}_n}, \mathcal{M}^{\mathcal{P}_n}) \xrightarrow{WGP} (\mathcal{T}^\Theta, d^\Theta, \mathbf{p}^\Theta).$$

(d) *If $\Theta_n \Rightarrow \Theta$, and $\mu^\Theta = \infty$ then*

$$(\mathcal{T}^{\Theta_n}, d^{\Theta_n}, \mathbf{p}^{\Theta_n}) \xrightarrow{WGP} (\mathcal{T}^\Theta, d^\Theta, \mathbf{p}^\Theta).$$

Remarks. • Theorem 3.4.2 (a) follows from Theorem 3.4.1 (a) since we consider discrete trees.

• There are several canonical measures on trees: uniform on leaves, uniform on vertices. . . We use generic measures $\mathcal{M}^{\mathcal{D}_n}$ so that Theorem 3.4.2 applies to such a variety of measures.

• If $\mathcal{M}^{\mathcal{D}_n}$ denotes the uniform measure on the leaves of $T^{\mathcal{D}_n}$, then Theorem 3.4.2 (b) follows from Theorem 3.4.1 (b), Lemma 3.3.2 and the weak convergence $\frac{1}{n} \sum_{i=1}^n \delta_{Y_i^\Theta} \rightarrow \mathbf{p}^\Theta$ (Proposition 2.3.2).

To prove the convergence for general $\mathcal{M}^{\mathcal{D}_n}$ we use a coupling between Algorithms 3.1 and 3.6.

• In passing, by Lemma 3.3.2, we reprove (see [16] Theorem 9) that the cuts "are" uniform vertices: Let $\Theta \in \Omega_\Theta$ such that $\mu^\Theta = \infty$. Let (A_i^Θ) be a family of i.i.d.random variables with law \mathbf{p}^Θ . Then

$$(d^\Theta(Y_i^\Theta, Y_j^\Theta))_{i,j \in \mathbb{N}} \stackrel{(d)}{=} (d^\Theta(A_i^\Theta, A_j^\Theta))_{i,j \in \mathbb{N}}.$$

• Results (c) and (d) are already mostly proved by Aldous, Camarri and Pitman in [16, 40]. Our only contribution to those results is the small extension of the \mathcal{P} -tree model.

3.4.3 Gromov–Hausdorff–Prokhorov convergence

We deduce (see Lemma 3.3.3) the GHP convergence, from the GP convergence given by Theorem 3.4.2, and the GH convergence. To prove the GH convergence, we use the convergence of the first branches, and prove that the whole trees are close from their first branches. We already estimated the distance between the ICRT and its first branches in Chapter 2, and we apply similar methods for \mathcal{D} -trees and \mathcal{P} -trees. From those estimates, we know when there is convergence, and so state assumptions to focus only on those cases.

Let us enter the details. We proved in Section 2.6 the following approximation for ICRT:

$$d_H(\mathcal{T}_y^\Theta, \mathcal{T}^\Theta) \approx \int_y^\infty \frac{dl}{l \mathbb{E}[\mu^\Theta[0, l]]}, \quad (3.7)$$

where for every $\Theta \in \Omega_\Theta$ and $y \in \mathbb{R}^+$, $\mathcal{T}_y^\Theta = ([0, y], d^\Theta, \mathbf{p}_y^\Theta)$. Hence, since we want the trees to be close from their first branches, we assume for ICRT that:

Assumption 3.5.

$$\lim_{y \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \int_y^{+\infty} \frac{dl}{l \mathbb{E}[\mu^{\Theta_n}[0, l]]} = 0.$$

Let us make a similar assumption for \mathcal{P} -trees. Here, the previous integral does not converge since $\mu^{\mathcal{P}}[0, \infty] = \sum_{i=1}^{\infty} p_i < \infty$. To solve this issue, we truncate the integral and rely on from the following generalisation of (3.7):

$$d_H(\mathcal{T}_y^{\Theta}, \mathcal{T}_{y'}^{\Theta}) \approx \int_y^{y'} \frac{dl}{l \mathbb{E}[\mu^{\Theta}[0, l]]}. \quad (3.8)$$

The main idea is then to take y' as large as possible such that the approximation (3.8) stays "optimal" and then to estimate $d_H(\mathcal{T}_{y'}, \mathcal{T})$. For this reason we assume for \mathcal{P} -trees that:

Assumption 3.6. The two following assumptions hold:

- (i) For every $\mathcal{P} \in \Omega_{\mathcal{P}}$, let $t^{\mathcal{P}} := \inf \{l \in \mathbb{N}, \mathbb{E}[\mu^{\mathcal{P}}[0, l]] \geq 1/2\}$ and let $\bar{\mu}^{\mathcal{P}} := (\sigma^{\mathcal{P}}, 1/\sigma^{\mathcal{P}})\mu^{\mathcal{P}}$ (see Section 3.3.1 for definition of rescaled measure). We have,

$$\lim_{y \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \int_y^{\sigma^{\mathcal{P}_n} t^{\mathcal{P}_n}} \frac{dl}{l \mathbb{E}[\bar{\mu}^{\mathcal{P}_n}[0, l]]} = 0.$$

- (ii) $\ln s^{\mathcal{P}_n} = o(1/\sigma^{\mathcal{P}_n})$.

For \mathcal{D} -trees, we make similar assumptions with another one for vertices of degree 1. Those vertices have the particularity of "not adding mass" to μ while "adding length" to $T^{\mathcal{D}}$. This is an issue, since intuitively, as the construction goes on, adding vertices normally increases μ which results into having smaller branches.

Assumption 3.7. The three following assumptions hold :

- (i) Let for every $\mathcal{D} \in \Omega_{\mathcal{D}}$, $t^{\mathcal{D}} := \inf \{l \in \mathbb{N}, \mathbb{E}[\mu^{\mathcal{D}}[0, l]] \geq N^{\mathcal{D}}/2\}$ and $\bar{\mu}^{\mathcal{D}} := (\sigma^{\mathcal{D}}/s^{\mathcal{D}}, 1/\sigma^{\mathcal{D}})\mu^{\mathcal{D}}$. We have,

$$\lim_{y \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \int_y^{\frac{\sigma^{\mathcal{D}_n}}{s^{\mathcal{D}_n}} t^{\mathcal{D}_n}} \frac{dl}{l \mathbb{E}[\bar{\mu}^{\mathcal{D}_n}[0, l]]} = 0.$$

- (ii) $\ln \left(\frac{s^{\mathcal{D}_n}}{s_{\geq 2}^{\mathcal{D}_n}} \right) \frac{s^{\mathcal{D}_n}}{N^{\mathcal{D}_n}} = o \left(\frac{s^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} \right)$ where for every $\mathcal{D} \in \Omega_{\mathcal{D}}$, $s_{\geq 2}^{\mathcal{D}} := \#\{i \in \mathbb{N}, d_i^{\mathcal{D}} \geq 2\}$.
 (iii) $\frac{\ln(N^{\mathcal{D}_n})}{\ln(s^{\mathcal{D}_n}/s_1^{\mathcal{D}_n})} = o \left(\frac{s^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} \right)$ where for every $\mathcal{D} \in \Omega_{\mathcal{D}}$, $s_1^{\mathcal{D}} := \#\{i \in \mathbb{N}, d_i^{\mathcal{D}} = 1\}$.

Remarks. We discuss here the necessity of the different assumptions for the GHP convergence.

- First (i) corresponds to the compactness of ICRT. Although we do not prove it, we are convinced that (i) is necessary as one can adapt the argument of Section 2.6.3 to prove that there is no possible convergence otherwise. In applications $t^{\mathcal{D}}$ and $t^{\mathcal{P}}$ are hard to compute, however since the integral "grows slowly" one can often choose much more convenient bounds for the integral.
- Although natural, Assumptions 3.6 (ii) and 3.7 (ii) are not necessary as Lemma 3.7.8 is sub-optimal.
- Assumption 3.7 (iii) is optimal: The right hand side is the typical distance between two vertices. The left hand side is the typical size of the longest paths that are composed of vertices of degree 1. Hence when those paths are too long \mathcal{D}_n -trees can not converge.
- Although (ii) and (iii) of Assumption 3.7 look similar, they are not equivalent. However, we have the following implications: If (ii) holds and either $s_1^{\mathcal{D}_n} = O(s_{\geq 2}^{\mathcal{D}_n})$ or $N^{\mathcal{D}_n} = O(s_{\geq 2}^{\mathcal{D}_n})$, then (iii) holds. If (iii) holds and $s_{\geq 2}^{\mathcal{D}_n} = O(s_1^{\mathcal{D}_n})$ then (ii) holds.

Theorem 3.4.3. *The following convergences hold weakly for the GHP-topology.*

(a) *If $\mathcal{D}_n \Rightarrow \Theta$, $\mathcal{M}^{\mathcal{D}_n} \rightarrow 0$, and Assumption 3.7 is satisfied then*

$$(\mathcal{V}^{\mathcal{D}_n}, (\sigma^{\mathcal{D}_n}/s^{\mathcal{D}_n})d^{\mathcal{D}_n}, \mathcal{M}^{\mathcal{D}_n}) \xrightarrow{WGHP} (\mathcal{T}^\Theta, d^\Theta, \mathfrak{p}^\Theta).$$

(b) *If $\mathcal{P}_n \Rightarrow \Theta$, $\mathcal{M}^{\mathcal{P}_n} \rightarrow 0$, and Assumption 3.6 is satisfied then*

$$(\mathcal{V}^{\mathcal{P}_n}, \sigma^{\mathcal{P}_n}d^{\mathcal{P}_n}, \mathcal{M}^{\mathcal{P}_n}) \xrightarrow{WGHP} (\mathcal{T}^\Theta, d^\Theta, \mathfrak{p}^\Theta).$$

(c) *If $\Theta_n \Rightarrow \Theta$, and Assumption 3.5 is satisfied then*

$$(\mathcal{T}^{\Theta_n}, d^{\Theta_n}, \mathfrak{p}^{\Theta_n}) \xrightarrow{WGHP} (\mathcal{T}^\Theta, d^\Theta, \mathfrak{p}^\Theta).$$

Our proof of Theorem 3.4.3 follows a method of Aldous [10]. The main idea is to upper bound for a suitable sequence (k_n) , the sum $\sum_n d_H(\mathcal{T}_{Y_{k_n}}, \mathcal{T}_{Y_{k_{n+1}}})$. To this end note that for every n ,

$$d_H(\mathcal{T}_{Y_{k_n}}, \mathcal{T}_{Y_{k_{n+1}}}) = \max_{k_n < i \leq k_{n+1}} d(Y_i, \mathcal{T}_{Y_{k_n}}),$$

and that for $i > k_n$,

$$d(Y_i, \mathcal{T}_{Y_{k_n}}) \stackrel{(d)}{=} d(Y_{k_{n+1}}, \mathcal{T}_{Y_{k_n}}) = Y_{k_{n+1}} - Y_{k_n}.$$

The rest of the proof is just a computation using large deviation inequalities and union bounds.

Remark. This method was used in [10] to prove that the Brownian CRT is compact, and by taking the limit of those bounds one can give an alternative proof for the compactness of ICRT.

3.4.4 Height of \mathcal{D} -trees, \mathcal{P} -trees and ICRT

With the same method, we obtain some upper bounds for the height of \mathcal{D} -trees, \mathcal{P} -trees and ICRT:

Theorem 3.4.4. *There exists some constants $c, C > 0$ such that:*

(a) *For every $\mathcal{D} \in \Omega_{\mathcal{D}}$ and $x \in \mathbb{R}^+$:*

$$\mathbb{P} \left(c \frac{\sigma^{\mathcal{D}}}{s^{\mathcal{D}}} H(T^{\mathcal{D}}) > x + \int_1^{\frac{\sigma^{\mathcal{D}}}{s^{\mathcal{D}}} t^{\mathcal{D}}} \frac{dl}{l \mathbb{E}[\bar{\mu}^{\mathcal{D}}[0, l]]} + \frac{\ln(s_{\geq 2}^{\mathcal{D}}) \sigma^{\mathcal{D}}}{N^{\mathcal{D}}} + \frac{\sigma^{\mathcal{D}}}{s^{\mathcal{D}}} \frac{\ln(N^{\mathcal{D}})}{\ln\left(\frac{s^{\mathcal{D}}}{s_1^{\mathcal{D}}}\right)} \right) \leq C e^{-cx \mathbb{E}[\bar{\mu}^{\mathcal{D}}[0, x]]}.$$

(b) *For every $\mathcal{P} \in \Omega_{\mathcal{P}}$ and $x \in \mathbb{R}^+$:*

$$\mathbb{P} \left(c \sigma^{\mathcal{P}} H(T^{\mathcal{P}}) > x + \int_1^{\sigma^{\mathcal{P}} t^{\mathcal{P}}} \frac{dl}{l \mathbb{E}[\bar{\mu}^{\mathcal{P}}[0, l]]} + \sigma^{\mathcal{P}} \ln(s^{\mathcal{P}}) \right) \leq C e^{-cx \mathbb{E}[\bar{\mu}^{\mathcal{P}}[0, x]]}.$$

(c) *For every $\Theta \in \Omega_{\Theta}$ and $x \in \mathbb{R}^+$:*

$$\mathbb{P} \left(c H(\mathcal{T}^\Theta) > x + \int_x^{+\infty} \frac{dl}{l \mathbb{E}[\mu^\Theta[0, l]]} \right) \leq C e^{-cx \mathbb{E}[\mu^\Theta[0, x]]}.$$

Remarks. • There is a (non matching) lower bound in the case of the ICRT. Indeed, note that $H(\mathcal{T}^\Theta) \geq Y_1^\Theta$ then that

$$\mathbb{P}(Y_1^\Theta \geq x | \mu^\Theta) = e^{-\int_0^x \mu^\Theta[0, l] dl} \geq e^{-x \mu^\Theta[0, x]},$$

and then by convexity of $l \mapsto e^l$,

$$\mathbb{P}(Y_1^\Theta \geq x) \geq e^{-x \mathbb{E}[\mu^\Theta[0, x]]}.$$

Although we do not pursue in this direction, one can show similar bounds for \mathcal{D} -trees and \mathcal{P} -trees.

- The other terms are almost optimal as they correspond to Assumptions 3.5, 3.6, 3.7.
- Addario–Berry, Devroye, Janson [7] and Kortchemski [58] prove similar upper bounds for Galton–Watson trees and Lévy trees. Also Addario–Berry [2] prove a weaker result on \mathcal{D} -trees, which is similar to our result in the "small"-degree case and simpler to compute.

3.5 Convergence of the first branches of \mathcal{D} -trees

3.5.1 Convergence of \mathcal{D} -trees toward \mathcal{P} -trees: proof of Theorem 3.4.1 (a)

Recall Algorithms 3.1 and 3.2. Let $(\mathcal{D}_n)_{n \in \mathbb{N}}$ be a sequence in $\Omega_{\mathcal{D}}$ and let $\mathcal{P} \in \Omega_{\mathcal{P}}$ such that $\mathcal{D}_n \Rightarrow \mathcal{P}$. We saw in Section 3.2.2 that in this case $(A^{\mathcal{D}_n})_{n \in \mathbb{N}}$ converges weakly toward $A^{\mathcal{P}}$ as $n \rightarrow \infty$. Hence, we may assume that a.s. for every $i \in \mathbb{N}$, $A_i^{\mathcal{D}_n} \rightarrow A_i^{\mathcal{P}}$.

Our goal is to prove that $(1, 1/s^{\mathcal{D}_n})\Upsilon^{\mathcal{D}_n} \rightarrow \Upsilon^{\mathcal{P}}$ for the SB topology defined in Section 3.3.1. Since we work with discrete random variables, note that it suffices to show that for every $i \in \mathbb{N}$, $X_i^{\mathcal{D}_n} \rightarrow X_i^{\mathcal{P}}$, $Y_i^{\mathcal{D}_n} \rightarrow Y_i^{\mathcal{P}}$, $Z_i^{\mathcal{D}_n} \rightarrow Z_i^{\mathcal{P}}$ and that $\mu^{\mathcal{D}_n}/s^{\mathcal{D}_n} \rightarrow \mu^{\mathcal{P}}$. First we have for every $i \in \mathbb{N}$,

$$X_i^{\mathcal{D}_n} = \inf\{j \in \mathbb{N}, A_j^{\mathcal{D}_n} = V_i\} \longrightarrow \inf\{j \in \mathbb{N}, A_j^{\mathcal{P}} = V_i\} = X_i^{\mathcal{P}}.$$

Then we have for the Skorohod topology: (There is no issue involved with the sum of diracs here since all atoms lie at integer point.)

$$\frac{1}{s^{\mathcal{D}_n}}\mu^{\mathcal{D}_n} = \sum_{i=1}^{s^{\mathcal{D}_n}} \delta_{X_i^{\mathcal{D}_n}} \frac{d_i^{\mathcal{D}_n} - 1}{s^{\mathcal{D}_n}} \xrightarrow{\text{sko}} \sum_{i=1}^{\infty} \delta_{X_i^{\mathcal{P}}} p_i^{\mathcal{P}} = \mu^{\mathcal{P}}.$$

We now study the cuts and the glue points which correspond to repetitions in $(A_i^{\mathcal{D}_n})_{1 \leq i \leq s^{\mathcal{D}_n}-1}$. Here, some extra care is required since the vertex V_{∞} "is split" into many others in Algorithm 3.2. For every $i, j, k, n \in \mathbb{N}$, we have

$$\mathbb{P}\left(A_i^{\mathcal{D}_n} = A_j^{\mathcal{D}_n} \in \{V_a\}_{a>k}\right) = \sum_{a=k+1}^{+\infty} \frac{d_a^{\mathcal{D}_n}}{(s^{\mathcal{D}_n} - 1)} \frac{(d_a^{\mathcal{D}_n} - 1)}{(s^{\mathcal{D}_n} - 2)} \leq \sum_{a=k+1}^{+\infty} \frac{1}{a^2} \xrightarrow{k \rightarrow \infty} 0.$$

So, we may also assume that for every $i, j \in \mathbb{N}$, there exists $k \in \mathbb{N}$ such that for every n large enough we do not have $A_i^{\mathcal{D}_n} = A_j^{\mathcal{D}_n} \in \{V_a\}_{a>k}$. Or equivalently, for every $i, j \in \mathbb{N}$ such that $A_i^{\mathcal{P}} = A_j^{\mathcal{P}} = V_{\infty}$, for every n large enough $A_i^{\mathcal{D}_n} \neq A_j^{\mathcal{D}_n}$. Also, for every $i, j \in \mathbb{N}$ such that $A_i^{\mathcal{P}} = A_j^{\mathcal{P}} \neq V_{\infty}$, for every n large enough $A_i^{\mathcal{D}_n} = A_j^{\mathcal{D}_n}$. It directly follows that for every $i \in \mathbb{N}$, $Y_i^{\mathcal{D}_n} \rightarrow Y_i^{\mathcal{P}}$ and $Z_i^{\mathcal{D}_n} \rightarrow Z_i^{\mathcal{P}}$. This concludes the proof.

3.5.2 Preliminaries: \mathcal{D} -trees and \mathcal{P} -trees can be seen as modifications of ICRT

In this section, we give a rigorous meaning (Lemma 3.5.1) to the following informal statement: If one gives to each edge e of a \mathcal{P} -tree a length l_e , such that $(l_e)_{e \in T^{\mathcal{P}}}$ is a family of i.i.d. exponential random variables of mean $\sigma^{\mathcal{P}}$, then this \mathcal{P} -tree is quasi isometric to an ICRT. Also, although the connection between \mathcal{D} -trees and ICRT is not as simple, we prove a similar statement for \mathcal{D} -trees (Lemma 3.5.2), which is key in our proof of Theorem 3.4.1 (b).

We first introduce "length change" on \mathbb{K}^{SB} (recall Section 3.3.1). Let \mathbb{F} be the set of increasing bijection from \mathbb{R}^+ to \mathbb{R}^+ . For every $f \in \mathbb{F}$, $\lambda > 0$ and $\Upsilon = ((X_i, Y_i, Z_i)_{i \in \mathbb{N}}, \mu) \in \mathbb{K}_{\text{SB}}$, let

$$(f, \lambda)\Upsilon := ((f(X_i), f(Y_i), f(Z_i))_{i \in \mathbb{N}}, \lambda f_{\star} \mu).$$

Note that for every $f \in \mathbb{F}$, $\lambda > 0$, the map $\Upsilon \mapsto (f, \lambda)\Upsilon$ is continuous and so measurable.

Lemma 3.5.1. *Let $\mathcal{P} \in \Omega_{\mathcal{P}}$, let $\Theta = (0, p_1/\sigma, p_2/\sigma, \dots)$. Note that $\Theta \in \Omega_{\Theta}$. Let $\{l_i\}_{i \in \mathbb{N}}$ be a family of i.i.d. exponential random variables of mean σ . Let $f \in \mathbb{F}$ be such that for every $i \geq 0$, $f(\sigma i) = \sum_{k=1}^i l_k$. Then $(f, 1/\sigma)\Upsilon^{\mathcal{P}}$ and Υ^{Θ} have the same distribution.*

Proof. We omit the proof of Lemma 3.5.1 since it would be too similar to the proof of Lemma 3.5.2. (See also Camarri, Pitman [40] Section 4.1 who introduce a similar connection.) \square

The connection between \mathcal{D} -trees and ICRT is not as simple. However we have the following construction which "replaces" ICRT.

Algorithm 3.5. Definition of "the continuum \mathcal{D} -tree": $\tilde{\Upsilon}^{\mathcal{D}}$.

- Let $(\tilde{X}_i)_{i \in \mathbb{N}}$ be a family of independent exponential random variables of parameter $(d_i/\sigma)_{i \in \mathbb{N}}$.
- Let $\tilde{\mu}$ be the measure on \mathbb{R}^+ defined by $\tilde{\mu} = \sum_{i=1}^s \delta_{\tilde{X}_i} (d_i - 1) / \sigma$.
- Let $(\tilde{Y}_i, \tilde{Z}_i)_{i \in \mathbb{N}}$ be a Poisson point process on $\{(y, z) \in \mathbb{R}^{+2} : y \geq z\}$ of intensity $dy \times d\tilde{\mu}$.
- For every $i \in \mathbb{N}$ with $d_i \geq 1$, let $U_{X,i}$ be uniform in $\{1, \dots, d_i\}$.
- For every $j \in \mathbb{N}$, let U_j be uniform in $\{1, \dots, d_i\} \setminus \{U_{X,i}\}$ where i is the unique index such that $\tilde{X}_i = \tilde{Z}_j$. (U_j is well defined since \tilde{Z}_j is in the support of $\tilde{\mu}$, which is $\{\tilde{X}_i\}_{i: d_i > 0}$.)
- Let $k_1 < \dots < k_N$ be the indexes such that $(Z_k, U_k) \notin \{(Z_j, U_j), j < k\}$. (There is a.s. N such indexes since $N = \sum_{i: d_i > 0} d_i - 1 = \#\{(i, u) \in \mathbb{N}^2 : d_i > 0, 1 \leq u \leq d_i, u \neq U_{X,i}\}$.)
- For every $1 \leq i \leq N$, let $\tilde{Y}_i = \tilde{Y}_{k_i}$ and let $\tilde{Z}_i = \tilde{Z}_{k_i}$. For every $i > N$, let $\tilde{Y}_i = \tilde{Z}_i = +\infty$.
- Let $\tilde{\Upsilon}^{\mathcal{D}} = ((\tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i)_{i \in \mathbb{N}}, \tilde{\mu})$.

The following result is the analog of Lemma 3.5.1 for \mathcal{D} -trees.

Lemma 3.5.2. *Let $\mathcal{D} \in \Omega_{\mathcal{D}}$. Let $(E_i^{\mathcal{D}})_{1 \leq i \leq s^{\mathcal{D}}-1}$ be a family of independent exponential random variables of mean $(\sigma^{\mathcal{D}}/(s^{\mathcal{D}} - i))_{1 \leq i \leq s^{\mathcal{D}}-1}$ and let $f^{\mathcal{D}} \in \mathbb{F}$ such that for every $1 \leq i \leq s^{\mathcal{D}} - 1$, $f^{\mathcal{D}}(i) = \sum_{k=1}^i E_k^{\mathcal{D}}$. Then $(f^{\mathcal{D}}, 1/\sigma^{\mathcal{D}})\Upsilon^{\mathcal{D}}$ and $\tilde{\Upsilon}^{\mathcal{D}}$ have the same distribution.*

Proof. Note that it suffices to construct a coupling between Algorithms 3.1 and 3.5 such that a.s. $(f(X_i), f(Y_i), f(Z_i))_{i \in \mathbb{N}} = (\tilde{X}_i, \tilde{Y}_i, \tilde{Z}_i)_{i \in \mathbb{N}}$. To this end we use the same "starting randomness" for both algorithms:

First let $I := \{(i, j) \in \mathbb{N}^2, 1 \leq i \leq s, 1 \leq j \leq d_i\}$. For every $(i, j) \in I$, let $P_{i,j}$ be a Poisson point process on \mathbb{R}^+ with rate $1/\sigma$ and let $E_{i,j} = \min P_{i,j}$. For every $i \in \mathbb{N}$ such that $d_i \geq 1$ let $k_i = \operatorname{argmin}\{E_{i,j}, 1 \leq j \leq d_i\}$. Finally let $I_2 = I \setminus \{(i, k_i)\}_{1 \leq i \leq s}$.

Toward Algorithm 3.1, sort $\{(E_{i,j}, V_i, j)\}_{(i,j) \in I}$ by the first coordinate as $(t_i, B_i, K_i)_{1 \leq i \leq s-1}$. One can easily check, that $(t_i)_{1 \leq i \leq s-1}$ is independent of $(B_i, k_i)_{1 \leq i \leq s-1}$, that $(t_i)_{i \in \mathbb{N}}$ have the same distribution as $(f(i))_{1 \leq i \leq s-1}$, and that $(B_i, k_i)_{1 \leq i \leq s-1}$ is a uniform permutation of I . We omit the trivial details. As a result, $(t_i, B_i)_{1 \leq i \leq s-1}$ have the same distribution as $(f(i), A_i)_{1 \leq i \leq s-1}$. Therefore we may assume that for every $1 \leq i \leq s - 1$, $(t_i, B_i) = (f(i), A_i)$. It directly follows that for every $i \in \mathbb{N}$ such that $d_i \geq 1$,

$$f(X_i) = f(\min\{j \in \mathbb{N}, A_j = V_i\}) = \min\{f(j), A_j = V_i\} = \min\{E_{i,k}, 1 \leq k \leq d_i\} = E_{i,k_i}. \quad (3.9)$$

Then by a similar argument,

$$\{f(Y_i), f(Z_i)\}_{1 \leq i \leq N} = \{E_{i,j}, E_{i,k_i}\}_{(i,j) \in I_2}. \quad (3.10)$$

Toward Algorithm 3.5, note that E_{i,k_i} is an exponential random variable of mean d_i/σ and that k_i is uniform in $\{1, \dots, d_i\}$, hence we may assume that

$$\tilde{X}_i = E_{i,k_i} \quad ; \quad U_{X,i} = k_i. \quad (3.11)$$

Then, note that conditionally on $(k_i)_{1 \leq i \leq s}$ and on $(E_{i,k_i})_{1 \leq i \leq s}$, $\bigcup_{(i,j) \in I_2} P_{i,j} \times \{E_{i,k_i}\} \times \{j\}$ is a Poisson point process on \mathbb{R}^{+3} of intensity

$$\sum_{(i,j) \in I_2} \mathbf{1}_{E_{i,k_i} \leq x} dx \times \delta_{E_{i,k_i}} \times \delta_j. \quad (3.12)$$

Also note that conditionally on $(\tilde{X}_i)_{1 \leq i \leq s}$ and on $(U_{X,i})_{1 \leq i \leq s}$, $\{\tilde{Y}_i, \tilde{Z}_i, U_{Z,i}\}_{i \in \mathbb{N}}$ is also a Poisson point process with the same intensity as in (3.12). So we may assume that

$$\{\tilde{Y}_i, \tilde{Z}_i, U_{Z,i}\}_{i \in \mathbb{N}} = \bigcup_{(i,j) \in I_2} P_{i,j} \times \{E_{i,k_i}\} \times \{j\}. \quad (3.13)$$

Therefore, by "keeping" the first point of each line $\{y, z, u\}_{y \in \mathbb{R}^+}$ in both side and by "deleting" the last coordinate, we have,

$$\{\tilde{Y}_i, \tilde{Z}_i\}_{1 \leq i \leq N} = \{E_{i,j}, E_{i,k_i}\}_{(i,j) \in I_2}. \quad (3.14)$$

Finally by (3.9) and (3.13) we have $(f(X_i))_{i \in \mathbb{N}} = (\tilde{X}_i)_{i \in \mathbb{N}}$. Also by (3.10), (3.14), and by monotony of $(Y_i)_{i \in \mathbb{N}}$ and $(\tilde{Y}_i)_{i \in \mathbb{N}}$ we have $(f(Y_i), f(Z_i))_{i \in \mathbb{N}} = (\tilde{Y}_i, \tilde{Z}_i)_{i \in \mathbb{N}}$. \square

3.5.3 Convergence of \mathcal{D} -trees toward ICRT: proof of Theorem 3.4.1 (b)

Fix $(\mathcal{D}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{D}}^{\mathbb{N}}$ and $\Theta \in \Omega_{\Theta}$. We assume that $\mathcal{D}_n \Rightarrow \Theta$, that is $d_1^{\mathcal{D}_n}/s^{\mathcal{D}_n} \rightarrow 0$ and $\forall i \in \mathbb{N}$, $d_i^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n} \rightarrow \theta_i^{\Theta}$, and we prove that $(\sigma^{\mathcal{D}_n}/s^{\mathcal{D}_n}, 1/\sigma^{\mathcal{D}_n})\Upsilon^{\mathcal{D}_n} \rightarrow \Upsilon^{\Theta}$ weakly for the SB topology. To this end, we follow the steps of Algorithm 3.5 and concludes by using Lemma 3.5.2.

Lemma 3.5.3. *If $\mathcal{D}_n \Rightarrow \Theta$ then jointly in distribution for every $i \in \mathbb{N}$, $\tilde{X}_i^{\mathcal{D}_n} \rightarrow X_i^{\Theta}$ and $\tilde{\mu}^{\mathcal{D}_n} \rightarrow \mu$.*

Proof. The convergence of the variables $(\tilde{X}_i^{\mathcal{D}_n})_{i \in \mathbb{N}}$ is immediate from their definitions. Hence by the Skorohod representation theorem we may assume that a.s. for every $i \in \mathbb{N}$, $\tilde{X}_i^{\mathcal{D}_n} \rightarrow X_i^{\Theta}$.

We now prove that jointly with the previous convergence, $\tilde{\mu}^{\mathcal{D}_n} \rightarrow \mu^{\Theta}$ weakly for the Skorohod topology. To this end, for every $\mathcal{D} \in \Omega_{\mathcal{D}}$ and $k \in \mathbb{N}$ we split $x \mapsto \tilde{\mu}^{\mathcal{D}}[0, x]$ into

$$F^{\mathcal{D}, \leq k} : x \mapsto \sum_{i=1}^k \mathbf{1}_{\tilde{X}_i^{\mathcal{D}} \leq x} (d_i^{\mathcal{D}} - 1)/\sigma^{\mathcal{D}} \quad \text{and} \quad F^{\mathcal{D}, > k} : x \mapsto \sum_{i=k+1}^{\infty} \mathbf{1}_{\tilde{X}_i^{\mathcal{D}} \leq x} (d_i^{\mathcal{D}} - 1)/\sigma^{\mathcal{D}}.$$

And we show that if $(k_n)_{n \in \mathbb{N}}$ is a sequence increasing sufficiently slowly to $+\infty$ then for every $x > 0$ (a) almost surely $F^{\mathcal{D}_n, \leq k_n}(x) \rightarrow \sum_{i=1}^{\infty} \mathbf{1}_{X_i^{\Theta} \leq x} \theta_i$ and (b) $F^{\mathcal{D}_n, > k_n}(x) \rightarrow \theta_0^2 x$ in probability. Note that provided (c) $(\tilde{\mu}^{\mathcal{D}_n})_{n \in \mathbb{N}}$ is tight for the Skorohod topology, a sum of (a) and (b) yields the desired result.

Toward (a), recall that for every $i \in \mathbb{N}$, a.s. $\tilde{X}_i^{\mathcal{D}_n} \rightarrow X_i^{\Theta}$, and that $\mathcal{D}_n \Rightarrow \Theta$. So if $(k_n)_{n \in \mathbb{N}}$ increases sufficiently slowly to $+\infty$ then by bounded convergence for every $x \in \mathbb{R}^+$ a.s.

$$F^{\mathcal{D}_n, \leq k_n}(x) = \sum_{i=1}^{k_n} \mathbf{1}_{\tilde{X}_i^{\mathcal{D}_n} \leq x} (d_i^{\mathcal{D}_n} - 1)/\sigma^{\mathcal{D}_n} \xrightarrow{n \rightarrow \infty} \sum_{i=1}^{\infty} \mathbf{1}_{X_i^{\Theta} \leq x} \theta_i^{\Theta}.$$

To prove (b) we use a second moment method. We have for every $x \geq 0$,

$$\mathbb{E} \left[F^{\mathcal{D}_n, > k_n}(x) \right] = \sum_{i=k_n+1}^{\infty} \frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \mathbb{P} \left(\tilde{X}_i^{\mathcal{D}_n} \leq x \right) = \sum_{i=k_n+1}^{\infty} \frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \left(1 - \exp \left(x \frac{d_i^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} \right) \right).$$

Then, since $k_n \rightarrow \infty$ and since $\mathcal{D}_n \Rightarrow \Theta$,

$$\mathbb{E} \left[F^{\mathcal{D}_n, > k_n}(x) \right] \sim \sum_{i=k_n+1}^{\infty} x \frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \frac{d_i^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} = x - x \sum_{i=1}^{k_n} \frac{d_i^{\mathcal{D}_n} (d_i^{\mathcal{D}_n} - 1)}{(\sigma^{\mathcal{D}_n})^2} \rightarrow x - x \sum_{i=1}^{+\infty} \theta_i^2 = x \theta_0^2,$$

where the last convergence holds by bounded convergence when $(k_n)_{n \in \mathbb{N}}$ increases sufficiently slowly to $+\infty$. Also similarly for every $x \geq 0$,

$$\mathbb{V} \left[F^{\mathcal{D}_n, > k_n}(x) \right] \leq \sum_{i=k_n+1}^{+\infty} \left(\frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \right)^2 \mathbb{P} \left(\tilde{X}_i^{\mathcal{D}_n} \leq x \right) \leq \frac{d_{k_n+1}^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} \mathbb{E} \left[F^{\mathcal{D}_n, > k_n}(x) \right] = o(1).$$

And (b) follows.

Toward (c), for every $0 \leq x \leq y \leq z$, and $n \in \mathbb{N}$ we have,

$$\begin{aligned}
\Delta^n(x, y, z) &:= \mathbb{E} [|\tilde{\mu}^{\mathcal{D}_n}(y) - \tilde{\mu}^{\mathcal{D}_n}(x)| |\tilde{\mu}^{\mathcal{D}_n}(z) - \tilde{\mu}^{\mathcal{D}_n}(y)|] \\
&= \mathbb{E} \left[\sum_{i=1}^{\infty} \mathbf{1}_{x < \tilde{X}_i^{\mathcal{D}_n} \leq y} \frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \sum_{i=1}^{\infty} \mathbf{1}_{y < \tilde{X}_i^{\mathcal{D}_n} \leq z} \frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \right] \\
&= \sum_{i \neq j} \mathbb{P} \left(x < \tilde{X}_i^{\mathcal{D}_n} \leq y \right) \frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \mathbb{P} \left(y < \tilde{X}_j^{\mathcal{D}_n} \leq z \right) \frac{d_j^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \\
&\leq \left(\sum_{i=1}^{\infty} \mathbb{P} \left(x < \tilde{X}_i^{\mathcal{D}_n} \leq z \right) \frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \right)^2. \tag{3.15}
\end{aligned}$$

Then by definition of $(\tilde{X}_i^{\mathcal{D}_n})_{i \in \mathbb{N}}$ and by convexity of $x \mapsto e^x$,

$$\mathbb{P}(x < \tilde{X}_i^{\mathcal{D}_n} \leq z) = e^{-d_i^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n}z} - e^{-d_i^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n}x} \leq d_i^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n}(z - x).$$

Hence, by (3.15) and definition of $\sigma^{\mathcal{D}_n}$,

$$\Delta^n(x, y, z) \leq \left(\sum_{i, \theta_i=1}^{\infty} \frac{d_i^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} (z - x) \frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \right)^2 = (z - x)^2.$$

Finally since x, y, z are arbitrary (c) follows. \square

It directly follows from Lemma 3.5.3, and the definition of $(\hat{Y}_i^{\mathcal{D}_n}, \hat{Z}_i^{\mathcal{D}_n})$ (see Algorithm 3.5), that

$$\hat{\Upsilon}^{\mathcal{D}_n} := ((\tilde{X}_i^{\mathcal{D}_n}, \hat{Y}_i^{\mathcal{D}_n}, \hat{Z}_i^{\mathcal{D}_n})_{i \in \mathbb{N}}, \tilde{\mu}^{\mathcal{D}}) \xrightarrow[n \rightarrow \infty]{\text{WSB}} \Upsilon^{\Theta}. \tag{3.16}$$

We omit the trivial details, and refer for instance to Lemma 4.24 (ii') of Kallenberg [57] for more precision on convergence of Poisson point process.

The following lemma implies that we can "replace" in (3.16) $(\hat{Y}_i^{\mathcal{D}_n}, \hat{Z}_i^{\mathcal{D}_n})_{i \in \mathbb{N}}$ by $(\tilde{Y}_i^{\mathcal{D}_n}, \tilde{Z}_i^{\mathcal{D}_n})_{i \in \mathbb{N}}$.

Lemma 3.5.4. *Assume that $\mathcal{D}_n \Rightarrow \Theta$ then for every $i \in \mathbb{N}$,*

$$\mathbb{P} \left(\hat{Y}_i^{\mathcal{D}_n} = \tilde{Y}_i^{\mathcal{D}_n} \text{ and } \hat{Z}_i^{\mathcal{D}_n} = \tilde{Z}_i^{\mathcal{D}_n} \right) \xrightarrow[n \rightarrow \infty]{} 1.$$

Proof. Let for $\mathcal{D} \in \Omega_{\mathcal{D}}$, $m^{\mathcal{D}} := \inf((\tilde{Y}_i^{\mathcal{D}})_{i \in \mathbb{N}} \setminus (\hat{Y}_i^{\mathcal{D}})_{1 \leq i \leq N})$. Note that it is enough to show that for every $l \in \mathbb{R}^+$, $\mathbb{P}(m^{\mathcal{D}_n} > l) \rightarrow 1$. To this end, we lower bound $\mathbb{P}(m^{\mathcal{D}} > l)$ for \mathcal{D} and l fixed.

First let us recall some notations introduced in the proof of Lemma 3.5.2. Let $I := \{(i, j) \in \mathbb{N}^2, 1 \leq i \leq s, 1 \leq j \leq d_i\}$. For every $(i, j) \in I$, let $P_{i,j}$ be a Poisson point process on \mathbb{R}^+ with rate $1/\sigma$ and let $E_{i,j} = \min P_{i,j}$. For every i such that $d_i \geq 1$ let $k_i = \operatorname{argmin}\{E_{i,j}, 1 \leq j \leq d_i\}$. Then let $I_2 = I \setminus \{i, k_i\}_{1 \leq i \leq s}$.

Now, recall that $\{\hat{Y}_i\}_{i \in \mathbb{N}} \setminus \{\hat{Y}_i\}_{i \in \mathbb{N}} = \bigcup_{(i,j) \in I_2} \{P_{i,j} \setminus \{E_{i,j}\}\}$. So that by the union bound,

$$\mathbb{P}(m > l) \leq \sum_{i=1}^s \mathbb{P} \left(\exists 1 \leq j \leq d_i : (i, j) \in I_2, \#(P_{i,j} \cap [0, l]) \geq 2 \right). \tag{3.17}$$

To simplify the notation, let $\sum_{1 \leq i \leq s} S_i$ denote the sum above. We split this sum in two according to "whether d_i is small or large", and we upper bound each part.

On the one hand, note that if there exists $1 \leq j \leq d_i$ such that $(i, j) \in I_2$ and $\#(P_{i,j} \cap [0, l]) \geq 2$ then $\#(\bigcup_{1 \leq j \leq d_i} P_{i,j} \cap [0, l]) \geq 3$ and $d_i \geq 2$. So that for every $1 \leq i \leq s$,

$$S_i \leq \mathbb{P} \left(\# \left(\bigcup_{1 \leq j \leq d_i} P_{i,j} \cap [0, l] \right) \geq 3 \right) = 1 - e^{-ld_i/\sigma} \left(1 + \frac{ld_i}{\sigma} + \frac{l^2 d_i^2}{2\sigma^2} \right) \leq \left(l \frac{d_i}{\sigma} \right)^3,$$

where the last inequality comes from the fact that for every $x \geq 0$, $1 - e^{-x}(1 + x + x^2/2) \leq x^3$. Then for every $\varepsilon > 0$,

$$\sum_{i, d_i \leq \varepsilon \sigma} S_i \leq \sum_{i, d_i \leq \varepsilon \sigma} \mathbf{1}_{d_i \geq 2} \left(l \frac{d_i}{\sigma} \right)^3 \leq 2l^3 \varepsilon \sum_{i, d_i \leq \varepsilon \sigma} \frac{d_i(d_i - 1)}{\sigma^2} \leq 2l^3 \varepsilon. \quad (3.18)$$

On the other hand, we have by a similar argument,

$$S_i \leq d_i \mathbb{P}(\#(P_{i,1} \cap [0, l]) \geq 2) \leq d_i l^2 / \sigma^2,$$

and when $\varepsilon \sigma > 1$,

$$\sum_{i, d_i > \varepsilon \sigma} P_i \leq \sum_{i, d_i > \varepsilon \sigma} l^2 \frac{d_i}{\sigma^2} \leq \sum_{i, d_i > \varepsilon \sigma} \frac{l^2}{\varepsilon \sigma - 1} \frac{d_i(d_i - 1)}{\sigma^2} \leq \frac{l^2}{\varepsilon \sigma - 1}. \quad (3.19)$$

Therefore, summing (3.18) and (3.19), we have for every $n \in \mathbb{N}$,

$$\mathbb{P}(m^{\mathcal{D}_n} > l) \leq 2l^3 \varepsilon + \frac{l^2}{\varepsilon \sigma^{\mathcal{D}_n} - 1}. \quad (3.20)$$

Then as $n \rightarrow +\infty$, $\mathcal{D}_n \Rightarrow \Theta$ so $\sigma^{\mathcal{D}_n} \rightarrow \infty$ and the right hand term in (3.20) converges to 2ε . Since ε is arbitrary, $\mathbb{P}(m^{\mathcal{D}_n} > l) \rightarrow 0$. Since l is arbitrary, this concludes the proof. \square

Proof of Theorem 3.4.1 (b). Recall for $\mathcal{D} \in \Omega_{\mathcal{D}}$ the definitions of $f^{\mathcal{D}}$ and $\{E_i^{\mathcal{D}}\}_{1 \leq i < s^{\mathcal{D}}}$ introduced in Lemma 3.5.2. First it directly follows from Lemma 3.5.4 and equation (3.16) that weakly for the SB topology $\tilde{\Upsilon}^{\mathcal{D}_n} \rightarrow \Upsilon^{\Theta}$ as $n \rightarrow \infty$. Furthermore by Lemma 3.5.2, $(f^{\mathcal{D}_n}, 1/\sigma^{\mathcal{D}_n})\Upsilon^{\mathcal{D}_n}$ and $\tilde{\Upsilon}^{\mathcal{D}_n}$ have the same distribution. So, by Skorohod representation theorem we may assume that both the convergence and the equality holds almost surely, hence almost surely

$$(f^{\mathcal{D}_n}, 1/\sigma^{\mathcal{D}_n})\Upsilon^{\mathcal{D}_n} \rightarrow \Upsilon^{\Theta}. \quad (3.21)$$

Let for $n \in \mathbb{N}$, $\lambda_n := s^{\mathcal{D}_n} / \sigma^{\mathcal{D}_n}$. For every $x \in \mathbb{R}^+$ and $n \in \mathbb{N}$, we have by definition of $f^{\mathcal{D}_n}$ and $\{E_i^{\mathcal{D}_n}\}_{1 \leq i < s^{\mathcal{D}_n}}$,

$$\mathbb{E}[f^{\mathcal{D}_n}(\lfloor \lambda_n x \rfloor)] = \sum_{i=1}^{\lfloor \lambda_n x \rfloor} \mathbb{E}[E_i^{\mathcal{D}_n}] = \sum_{i=1}^{\lfloor \lambda_n x \rfloor} \frac{\sigma^{\mathcal{D}_n}}{s^{\mathcal{D}_n} - i} = \lfloor \lambda_n x \rfloor \frac{\sigma^{\mathcal{D}_n}}{s^{\mathcal{D}_n} - O(\lambda_n)} \xrightarrow{n \rightarrow \infty} x.$$

Similarly,

$$\mathbb{V}[f^{\mathcal{D}_n}(\lfloor \lambda_n x \rfloor)] = \sum_{i=1}^{\lfloor \lambda_n x \rfloor} \mathbb{V}[E_i] = \sum_{i=1}^{\lfloor \lambda_n x \rfloor} \left(\frac{\sigma^{\mathcal{D}_n}}{s^{\mathcal{D}_n} - i} \right)^2 \rightarrow 0.$$

Therefore $x \mapsto f^{\mathcal{D}_n}(\lambda_n x)$ converges in distribution for $\|\cdot\|_{\infty}$ toward the identity on any interval. Then by Skorohod representation theorem we may assume that this convergence holds almost surely. The convergence $(1/\lambda_n, 1/\sigma^{\mathcal{D}_n})\Upsilon^{\mathcal{D}_n} \rightarrow^{\text{SB}} \Upsilon^{\Theta}$ then follows from (3.21) and a deterministic topological verification. \square

3.6 Gromov–Prokhorov convergence of \mathcal{D} -trees

3.6.1 Preliminary: vertices with small degree behaves like leaves

Recall that, by Lemma 3.3.2, to prove the GP convergence, we need to show that the distances between random vertices converge. To do so, we introduce a generalisation of Algorithm 3.1 which constructs sequentially the subtree spanned by the root and W_1, \dots, W_i where $W = (W_i)_{1 \leq i \leq s^{\mathcal{D}}}$ is any permutation of $\mathcal{V}^{\mathcal{D}}$. Then, we couple it with Algorithm 3.1 to show that whenever W_1, W_2, \dots have "small" degree, those subtrees are close from the subtrees spanned by the first leaves (see Lemma 3.6.2). This agrees with the intuition that vertices with small degree behave like leaves. Naturally, this is also true for random vertices, which typically are distinct and have small degree.

Algorithm 3.6. General stick-breaking construction of \mathcal{D} -tree.

- Let $A^{\mathcal{D}} = (A_1, \dots, A_{s-1})$ be a uniform \mathcal{D} -tuple.
- Let $T_1^{\mathcal{D}, W} := (\{A_1\}, \emptyset)$ then for every $2 \leq i \leq s$ let

$$T_i^{\mathcal{D}, W} := \begin{cases} T_{i-1} \cup \{A_{i-1}, A_i\} & \text{if } A_i \notin T_{i-1}, \\ T_{i-1} \cup \{A_{i-1}, W_{\inf\{k, W_k \notin T_{i-1}\}}\} & \text{if } A_i \in T_{i-1} \text{ or } i = s. \end{cases}$$

- Let $T^{\mathcal{D}, W}$ denote the rooted tree (T_s, A_1) .

Remark. If for every $1 \leq i \leq N+1$, $W_i = L_i$ then Algorithm 3.1 and Algorithm 3.6 follow the exact same steps, hence $T^{\mathcal{D}, W} = T^{\mathcal{D}}$.

Proposition 3.6.1. *For every $\mathcal{D} \in \Omega_{\mathcal{D}}$, and W permutation of $\mathcal{V}^{\mathcal{D}}$, $T^{\mathcal{D}, W}$ is a \mathcal{D} -tree.*

Proof. See Appendix. □

Before stating the main result of this section, let us define the relabelling operation. For every graph $G = (V, E)$ and bijection $f : V \rightarrow V'$, let $f(G) := (V', \{\{f(x), f(y)\}\}_{\{x, y\} \in E})$.

Lemma 3.6.2. *Let $\mathcal{D} \in \Omega_{\mathcal{D}}$ and $W = \{W_i\}_{1 \leq i \leq s^{\mathcal{D}}}$ be a permutation of $\mathcal{V}^{\mathcal{D}}$. Let $1 \leq k \leq N^{\mathcal{D}}$ and let $f_k : \mathcal{V}^{\mathcal{D}} \rightarrow \mathcal{V}^{\mathcal{D}}$ be a bijection such that:*

$$\forall 1 \leq i \leq k \quad f_k(W_i) = L_i \quad ; \quad \forall V \notin \{W_i\}_{1 \leq i \leq k} \cup \{L_i\}_{1 \leq i \leq k} \quad f_k(V) = V.$$

Then for every $1 \leq l \leq s^{\mathcal{D}}$, (where by convention for $i \in \mathbb{N}$, $d_{V_i}^{\mathcal{D}} := d_i^{\mathcal{D}})$

$$\mathbb{P}\left(f_k\left(T_{Y_k^{\mathcal{D}}}^{\mathcal{D}, W}\right) \neq T_{Y_k^{\mathcal{D}}}^{\mathcal{D}}\right) \leq \mathbb{P}(Y_k^{\mathcal{D}} > l) + l(d_{W_1}^{\mathcal{D}} + \dots d_{W_k}^{\mathcal{D}})/(s^{\mathcal{D}} - 1).$$

Proof. Note that $f_k(T_{Y_k}^W) = T_{Y_k}$ whenever $\{W_i\}_{1 \leq i \leq k} \cap \{A_1, \dots, A_{Y_k}\} = \emptyset$, since in this case, up to relabelling, Algorithm 3.1 and Algorithm 3.6 follow the exact same steps. Therefore:

$$\begin{aligned} \mathbb{P}\left(f_k(T_{Y_k}^W) \neq T_{Y_k}\right) &\leq \mathbb{P}(Y_k > l) + \mathbb{P}(\{W_i\}_{1 \leq i \leq k} \cap \{A_1, \dots, A_l\} \neq \emptyset), \\ &\leq \mathbb{P}(Y_k > l) + \sum_{i=1}^l \sum_{j=1}^k \mathbb{P}(A_i = W_j), \\ &= \mathbb{P}(Y_k > l) + l(d_{W_1} + \dots d_{W_k})/(s - 1). \end{aligned} \quad \square$$

3.6.2 Convergence of \mathcal{D} -trees toward ICRT: proof of Theorem 3.4.2 (b)

To simplify the notations we write for $\mathcal{D} \in \Omega_{\mathcal{D}}$, $\lambda^{\mathcal{D}} := s^{\mathcal{D}}/\sigma^{\mathcal{D}}$.

First recall the assumptions of Theorem 3.4.2 (b): $\mathcal{D}_n \Rightarrow \Theta$, $\mathcal{M}^{\mathcal{D}_n} \rightarrow 0$, and $\mu^{\Theta} = \infty$. In this case we have by Theorem 3.4.1 (b) the following joint convergence in distribution,

$$((Y_i^{\mathcal{D}_n}/\lambda^{\mathcal{D}_n})_{i \in \mathbb{N}}, (Z_i^{\mathcal{D}_n}/\lambda^{\mathcal{D}_n})_{i \in \mathbb{N}}) \xrightarrow{W} ((Y_i^{\Theta})_{i \in \mathbb{N}}, (Z_i^{\Theta})_{i \in \mathbb{N}}). \quad (3.22)$$

Moreover, note that the distance between the leaves of a tree constructed by stick breaking is "entirely determined" by the positions of the cuts and the glue points. More precisely, for every $i < j \in \mathbb{N}$, there exists a measurable function $g_{i,j} : \mathbb{R}^{+2j} \mapsto \mathbb{R}^+$ such that for every $n \in \mathbb{N}$ large enough (such that there is at least j leaves) and $\Theta \in \Omega_{\Theta}$:

$$(d^{\mathcal{D}_n}/\lambda^{\mathcal{D}_n})(L_i^{\mathcal{D}_n}, L_j^{\mathcal{D}_n}) = g_{i,j}((Y_k^{\mathcal{D}_n}/\lambda^{\mathcal{D}_n})_{1 \leq k \leq j}, (Z_k^{\mathcal{D}_n}/\lambda^{\mathcal{D}_n})_{1 \leq k \leq j}),$$

and

$$d^{\Theta}(Y_i^{\Theta}, Y_j^{\Theta}) = g_{i,j}((Y_k^{\Theta})_{1 \leq k \leq j}, (Z_k^{\Theta})_{1 \leq k \leq j}).$$

Therefore, by (3.22) we have the following joint convergence in distribution:

$$((d^{\mathcal{D}_n}/\lambda^{\mathcal{D}_n})(L_i^{\mathcal{D}_n}, L_j^{\mathcal{D}_n}))_{i,j \in \mathbb{N}} \xrightarrow{W} (d^{\Theta}(Y_i^{\Theta}, Y_j^{\Theta}))_{i,j \in \mathbb{N}}. \quad (3.23)$$

Now for every $n \in \mathbb{N}$ let $(W_i^{\mathcal{D}_n})_{i \in \mathbb{N}}$ be a family of i.i.d. random variables with law $\mathcal{M}^{\mathcal{D}_n}$. Recall by Lemma 3.3.1 and by Proposition 2.3.2 (see Section 3.4.2) that it suffices to prove that

$$((d^{\mathcal{D}_n}/\lambda^{\mathcal{D}_n})(W_i^{\mathcal{D}_n}, W_j^{\mathcal{D}_n}))_{i,j \in \mathbb{N}} \xrightarrow{W} (d^{\Theta}(Y_i^{\Theta}, Y_j^{\Theta}))_{i,j \in \mathbb{N}}. \quad (3.24)$$

To this end we use the coupling introduced in Section 3.6.1 to derive (3.24) from (3.23).

Beforehand, note that for every $n \in \mathbb{N}$, $(W_i^{\mathcal{D}_n})_{i \in \mathbb{N}}$ is not a permutation of $\mathcal{V}^{\mathcal{D}_n}$ so we can not directly apply Lemma 3.6.2. However, since $\mathcal{M}^{\mathcal{D}_n} \rightarrow 0$ we have for every $i, j \in \mathbb{N}$, $\mathbb{P}(W_i = W_j) \rightarrow 0$. Hence, there exists a family $(\tilde{W}_i^{\mathcal{D}_n})_{i \in \mathbb{N}} = ((\tilde{W}_i^{\mathcal{D}_n})_{1 \leq i \leq s^{\mathcal{D}_n}})_{n \in \mathbb{N}}$ of random permutations of $\mathcal{V}^{\mathcal{D}_n}$ such that for every $i \in \mathbb{N}$, $\mathbb{P}(W_i^{\mathcal{D}_n} = \tilde{W}_i^{\mathcal{D}_n}) \rightarrow 0$.

We now apply Lemma 3.6.2 to those permutations: We have for every $n, k \in \mathbb{N}$ and $l \in \mathbb{R}^+$,

$$\mathbb{P}\left(f_k^{\mathcal{D}_n}\left(T_{Y_k^{\mathcal{D}_n}}^{\mathcal{D}_n, \tilde{W}^{\mathcal{D}_n}}\right) \neq T_{Y_k^{\mathcal{D}_n}}^{\mathcal{D}_n}\right) \leq \mathbb{P}\left(Y_k^{\mathcal{D}_n} > \lambda^{\mathcal{D}_n} l\right) + \mathbb{E}\left[\frac{\lambda^{\mathcal{D}_n} l}{s^{\mathcal{D}_n} - 1} \left(d_{\tilde{W}_1^{\mathcal{D}_n}}^{\mathcal{D}_n} + \dots + d_{\tilde{W}_k^{\mathcal{D}_n}}^{\mathcal{D}_n}\right)\right],$$

where $f_k^{\mathcal{D}_n}$ is the relabelling function defined in Lemma 3.6.2. Moreover note that the last term converges to 0. Indeed since $\mathcal{M}^{\mathcal{D}_n} \rightarrow 0$, for every $i \in \mathbb{N}$, $d_{\tilde{W}_i^{\mathcal{D}_n}}^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n} \rightarrow 0$ in probability. And by bounded convergence the last also convergence holds in expectation since $d_1^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n} \rightarrow \theta_1 < \infty$. Therefore for every k, l fixed,

$$\limsup_{n \rightarrow \infty} \mathbb{P}\left(f_k^{\mathcal{D}_n}\left(T_{Y_k^{\mathcal{D}_n}}^{\mathcal{D}_n, \tilde{W}^{\mathcal{D}_n}}\right) \neq T_{Y_k^{\mathcal{D}_n}}^{\mathcal{D}_n}\right) \leq \limsup_{n \rightarrow \infty} \mathbb{P}\left(Y_k^{\mathcal{D}_n} > \lambda^{\mathcal{D}_n} l\right).$$

Therefore, since l is arbitrary and since by Theorem 3.4.1 (b) $\lambda^{\mathcal{D}_n} Y_k^{\mathcal{D}_n} \rightarrow Y_k^{\Theta}$ weakly as $n \rightarrow \infty$,

$$\limsup_{n \rightarrow \infty} \mathbb{P}\left(f_k^{\mathcal{D}_n}\left(T_{Y_k^{\mathcal{D}_n}}^{\mathcal{D}_n, \tilde{W}^{\mathcal{D}_n}}\right) \neq T_{Y_k^{\mathcal{D}_n}}^{\mathcal{D}_n}\right) = 0. \quad (3.25)$$

Finally, since relabelling does not change the distance in the tree,

$$d_{TV}\left((d^{\mathcal{D}_n}(L_i^{\mathcal{D}_n}, L_j^{\mathcal{D}_n}))_{i,j \leq k}, (d^{\mathcal{D}_n}(\tilde{W}_i^{\mathcal{D}_n}, \tilde{W}_j^{\mathcal{D}_n}))_{i,j \leq k}\right) \rightarrow 0,$$

where d_{TV} stands for the total variation distance. Therefore since $\forall i \in \mathbb{N}$, $\mathbb{P}(W_i^{\mathcal{D}_n} = \tilde{W}_i^{\mathcal{D}_n}) \rightarrow 0$,

$$d_{TV}\left((d^{\mathcal{D}_n}(L_i^{\mathcal{D}_n}, L_j^{\mathcal{D}_n}))_{i,j \leq k}, (d^{\mathcal{D}_n}(W_i^{\mathcal{D}_n}, W_j^{\mathcal{D}_n}))_{i,j \leq k}\right) \rightarrow 0. \quad (3.26)$$

And finally, (3.24) directly follows from (3.23) and (3.26). This concludes the proof.

3.7 Gromov–Hausdorff–Prokhorov convergence and height of \mathcal{D} -trees

3.7.1 Preliminaries: purely technical results on μ and $\tilde{\mu}$

The reader can safely skim this section on a first reading and refer back to it when needed.

We want to estimate μ since it "dictates" the stick breaking construction. However μ is hard to estimate as a sum of dependent random variables. For this reason, we estimate instead $\tilde{\mu}$ introduced in Algorithm 3.5, and we use in the next sections Lemma 3.5.2 to "replace" μ by $\tilde{\mu}$ when needed.

Lemma 3.7.1. *For every $\mathcal{D} \in \Omega_{\mathcal{D}}$ and $l \in \mathbb{R}^+$,*

$$\mathbb{P}(\tilde{\mu}^{\mathcal{D}}[0, l] \leq \mathbb{E}[\tilde{\mu}^{\mathcal{D}}[0, l]]/2) \leq e^{-l\mathbb{E}[\tilde{\mu}^{\mathcal{D}}[0, l]]/4}.$$

Proof. Recall that $(\tilde{X}_i)_{i \in \mathbb{N}}$ is a family of independent exponential random variables of parameter $(d_i/\sigma)_{i \in \mathbb{N}}$, and that $\tilde{\mu} = \sum_{i=1}^s \delta_{\tilde{X}_i} (d_i - 1)/\sigma$. So for every $1 \leq i \leq s$, $\mathbb{P}(\tilde{X}_i \leq l) = 1 - e^{-ld_i/\sigma}$, and hence $\tilde{\mu}[0, l] - \mathbb{E}[\tilde{\mu}[0, l]]$ can be written as a sum of independent centered random variables:

$$\tilde{\mu}[0, l] - \mathbb{E}[\tilde{\mu}[0, l]] = \sum_{i=1}^s M_i \quad \text{where for } 1 \leq i \leq s \quad M_i := \left(\mathbf{1}_{\tilde{X}_i \leq l} - 1 + e^{-ld_i/\sigma} \right) (d_i - 1)/\sigma.$$

We now compute some exponential moments. First for every i such that $d_i \geq 2$,

$$\begin{aligned} \mathbb{E} \left[\exp \left(-l \frac{d_i}{d_i - 1} M_i \right) \right] &= \exp \left(-(ld_i/\sigma) e^{-ld_i/\sigma} \right) \left(\mathbb{P}(\tilde{X}_i \leq l) + \mathbb{P}(\tilde{X}_i > l) e^{ld_i/\sigma} \right) \\ &= \exp \left(-(ld_i/\sigma) e^{-ld_i/\sigma} \right) \left(1 - e^{-ld_i/\sigma} + e^{-ld_i/\sigma} e^{ld_i/\sigma} \right) \\ &= \exp \left(-(ld_i/\sigma) e^{-ld_i/\sigma} + \log \left(2 - e^{-ld_i/\sigma} \right) \right). \end{aligned}$$

Then since for every $x \geq 0$, $-xe^{-x} + \log(2 - e^{-x}) \leq x(1 - e^{-x})/4$,

$$\mathbb{E} \left[\exp \left(-l \frac{d_i}{d_i - 1} M_i \right) \right] \leq \exp \left((ld_i/\sigma)(1 - e^{-ld_i/\sigma})/4 \right).$$

So by concavity of $t \rightarrow t^{(d_i-1)/d_i}$ on \mathbb{R}^+ ,

$$\mathbb{E}[\exp(-lM_i)] \leq \exp \left(l \frac{d_i - 1}{\sigma} (1 - e^{-ld_i/\sigma})/4 \right).$$

Therefore, since $M_i = 0$ when $d_i \leq 1$, multiplying over all i such that $d_i \geq 2$ we get,

$$\mathbb{E}[\exp(-l(\tilde{\mu}[0, l] - \mathbb{E}[\tilde{\mu}[0, l]]))] \leq \exp(l\mathbb{E}[\tilde{\mu}[0, l]]/4).$$

Finally the desired result follows from a simple application of Markov's inequality. \square

We now upper bound some "numbers of cuts". More precisely, for every $0 \leq m \leq N/\sigma$ let

$$\mathcal{X}_m := \inf\{l \in \mathbb{R}^+, \mathbb{E}[\tilde{\mu}[0, l]] \leq m\} \quad \text{and let} \quad \xi_m := \min\{i \in \mathbb{N}, \tilde{\mu}[0, \tilde{Y}_i] > m\}. \quad (3.27)$$

It is easy to check that $l \rightarrow \mathbb{E}[\tilde{\mu}[0, l]]$ is a continuous and strictly increasing function of l so that for every $0 \leq m \leq N/\sigma$, $\mathbb{E}[\tilde{\mu}[0, \mathcal{X}_m]] = m$. Moreover, we have the following upper bound on ξ_m :

Lemma 3.7.2. *For every $0 \leq m \leq N/(2\sigma)$,*

$$\mathbb{P}(\xi_m \geq 3\mathcal{X}_{2m}m + 1) \leq 2e^{-\mathcal{X}_{2m}m/2}.$$

Proof. First by Lemma 3.7.1 $\tilde{\mu}[0, \mathcal{X}_{2m}] \leq m$ with probability at most $e^{-\mathcal{X}_{2m}m/2}$. In the following we work conditionally on $\tilde{\mu}$ and assume that this event does not hold.

Note that $\xi_m - 1 = \max\{i, \tilde{\mu}[0, \hat{Y}_i] \leq m\} \leq \max\{i, \tilde{\mu}[0, \hat{Y}_i] \leq m\}$. Also since conditionally on $\tilde{\mu}$, $\{\hat{Y}_i\}_{i \in \mathbb{N}}$ is a Poisson point process on \mathbb{R}^+ with intensity $\tilde{\mu}[0, l]dl$, $\max\{i, \mu[0, \hat{Y}_i] \leq l\sigma\}$ is a Poisson random variable with mean $\int_0^\alpha \tilde{\mu}[0, t]dt$ where $\alpha := \max\{a \in \mathbb{R}^+, \tilde{\mu}[0, a] \leq m\}$. Therefore since by our assumption $\alpha \leq \mathcal{X}_{2m}$ and since $\tilde{\mu}[0, \alpha] \leq m$, $\xi_m - 1$ is bounded by a Poisson random variable of mean $\mathcal{X}_{2m}m$. Finally the result follows from some basic concentration inequalities for Poisson random variables (see e.g. [32] p.23). \square

Lemma 3.7.3. *The following assertions holds:*

- (a) For every $0 \leq m \leq N/(16\sigma)$, $\mathcal{X}_{2m} \leq \sigma/8$.
- (b) For every $x \geq 0$, $\mathbb{E}[\tilde{\mu}[0, x]] \leq x$. Hence, for every $0 \leq m < N/\sigma$, $m \leq \mathcal{X}_m$.
- (c) For every $x \geq 1/2$, $\mathbb{E}[\tilde{\mu}[0, x]] \geq 1/6$.

Proof. Toward (a), simply note that $x \mapsto \mathbb{E}[\tilde{\mu}[0, x]]$ is increasing and that

$$\mathbb{E}\left[\tilde{\mu}\left[0, \frac{\sigma}{8}\right]\right] = \sum_{i=1}^s \frac{d_i - 1}{\sigma} \left(1 - e^{-\frac{d_i}{8}}\right) \geq \left(1 - e^{-\frac{2}{8}}\right) \sum_{i=1}^s \frac{d_i - 1}{\sigma} \mathbf{1}_{d_i \geq 2} = \left(1 - e^{-\frac{2}{8}}\right) \frac{N}{\sigma} \geq 2m.$$

Similarly for (b), note that

$$\mathbb{E}[\tilde{\mu}[0, x]] = \sum_{i=1}^s \frac{d_i - 1}{\sigma} \left(1 - e^{-d_i \frac{x}{\sigma}}\right) \leq \sum_{i=1}^s \frac{(d_i - 1)(d_i)x}{\sigma^2} = x.$$

Finally for (c), since for every $0 \leq x \leq 1$, $1 - e^{-x} \geq x/3$, note that

$$\mathbb{E}[\tilde{\mu}[0, 1/2]] = \sum_{i=1}^s \frac{d_i - 1}{\sigma} \left(1 - e^{-\frac{d_i}{2\sigma}}\right) \geq \sum_{i=1}^s \frac{(d_i - 1)(d_i)}{6\sigma^2} = \frac{1}{6}. \quad \square$$

Lemma 3.7.4. *Recall that $\bar{\mu}$ is the real measure on \mathbb{R}^+ such that for every $l > 0$, $\frac{\mu^{\mathcal{D}}}{\sigma^{\mathcal{D}}} \left[0, \frac{s^{\mathcal{D}}}{\sigma^{\mathcal{D}}} l\right]$. For every $l \in \mathbb{R}^+$, we have $\mathbb{E}[\bar{\mu}[0, l]] \leq 2\mathbb{E}[\tilde{\mu}[0, l]]$.*

Proof. Fix $l \in \mathbb{R}^+$. Since $\mu = \sum_{i=1}^s \delta_{X_i}(d_i - 1) = \sum_{i=1}^s \delta_{\inf\{1 \leq j \leq s, A_j = V_i\}}(d_i - 1)$ and since $\tilde{\mu} = \sum_{i=1}^s \delta_{\tilde{X}_i}(d_i - 1)/\sigma$, we have by (3.1) and linearity of the expectation,

$$\mathbb{E}[\bar{\mu}[0, l]] = \sum_{i=1}^s \left(1 - \mathbb{P}\left(A_1, \dots, A_{\lfloor \frac{sl}{\sigma} \rfloor} \neq V_i\right)\right) \frac{d_i - 1}{\sigma} = \sum_{i=1}^s \left(1 - \prod_{j=1}^{\lfloor \frac{sl}{\sigma} \rfloor} \left(1 - \frac{d_i}{s - j}\right)\right) \frac{d_i - 1}{\sigma},$$

and

$$\mathbb{E}[\tilde{\mu}[0, l]] = \sum_{i=1}^s \left(1 - e^{-l \frac{d_i}{\sigma}}\right) \frac{d_i - 1}{\sigma}.$$

On the one hand, if $l > \sigma/2$ then:

$$\mathbb{E}[\tilde{\mu}[0, l]] = \sum_{i=1}^s \left(1 - e^{-l \frac{d_i}{\sigma}}\right) \frac{d_i - 1}{\sigma} \geq \sum_{i=1}^s \left(1 - e^{-1/2}\right) \frac{d_i - 1}{\sigma} \geq \frac{1}{2} \sum_{i=1}^s \frac{d_i - 1}{\sigma} \geq \frac{1}{2} \mathbb{E}[\bar{\mu}[0, l]].$$

On the other hand, if $l \leq \sigma/2$ then, by convexity of $x \mapsto e^x$, for every $1 \leq i \leq s$:

$$1 - \prod_{j=1}^{\lfloor \frac{sl}{\sigma} \rfloor} \left(1 - \frac{d_i}{s - j}\right) \leq 1 - \left(1 - 2\frac{d_i}{s}\right)^{\lfloor sl/\sigma \rfloor} \leq 1 - e^{2d_i l/\sigma} \leq 2(1 - e^{-d_i l/\sigma}).$$

Hence by summing over all $1 \leq i \leq \sigma$, we get $\mathbb{E}[\bar{\mu}[0, l]] \leq 2\mathbb{E}[\tilde{\mu}[0, l]]$. \square

We now prove several results that will be used to prove Theorem 3.4.3 (a).

Lemma 3.7.5. *Let $(\mathcal{D}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{D}}^{\mathbb{N}}$ and $\Theta \in \Omega_{\Theta}$. We have the following assertions:*

- (a) *If $\mathcal{D}_n \Rightarrow \Theta$, then for every $x \in \mathbb{R}^+$, $\mathbb{E}[\tilde{\mu}^{\mathcal{D}_n}[0, x]] \rightarrow \mathbb{E}[\mu^{\Theta}[0, x]]$.*
- (b) *If $\mathcal{D}_n \Rightarrow \Theta$ and Assumption 3.7 is satisfied, then $\mu^{\Theta} = \infty$.*
- (c) *If $\mathcal{D}_n \Rightarrow \Theta$ and Assumption 3.7 is satisfied, then the Θ -ICRT is almost surely compact.*

Proof. Toward (a), we have have by bounded convergence, provided some justifications,

$$\begin{aligned} \mathbb{E}[\tilde{\mu}^{\mathcal{D}_n}[0, x]] &= x + \sum_{i=1}^s \frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \left(1 - e^{-x d_i^{\mathcal{D}_n} / \sigma^{\mathcal{D}_n}} - x \frac{d_i^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} \right), \\ &\rightarrow x + \sum_{i=1}^{\infty} \theta_i (1 - e^{-x \theta_i} - x \theta_i) = \mathbb{E}[\mu^{\Theta}[0, x]]. \end{aligned} \quad (3.28)$$

So it remains to justify (3.28). We have for every $n \in \mathbb{N}$ and $i \in \mathbb{N}$ such that $d_i^{\mathcal{D}_n} \geq 2$,

$$\left(\frac{d_i^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} \right)^2 \leq 2 \frac{d_i^{\mathcal{D}_n} (d_i^{\mathcal{D}_n} - 1)}{(\sigma^{\mathcal{D}_n})^2} \leq 2 \frac{\frac{1}{i} \sum_{j=1}^i d_j^{\mathcal{D}_n} (d_j^{\mathcal{D}_n} - 1)}{(\sigma^{\mathcal{D}_n})^2} \leq \frac{2}{i}.$$

So for every $n \in \mathbb{N}$ and $i \in \mathbb{N}$, since for every $0 \leq y \leq \sqrt{2}$, $e^y - 1 - x \leq y^2$,

$$\frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \left(1 - e^{-x d_i^{\mathcal{D}_n} / \sigma^{\mathcal{D}_n}} - x \frac{d_i^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} \right) \leq \frac{d_i^{\mathcal{D}_n} - 1}{\sigma^{\mathcal{D}_n}} \left(\frac{d_i^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} \right)^2 \leq \frac{2^{3/2}}{i^{3/2}}.$$

And the convergence in (3.28) follows.

We now prove (b) and (c). First by Assumption 3.7 (ii), $\sigma^{\mathcal{D}_n} = o(N^{\mathcal{D}_n})$. Then, by Lemmas 3.7.4 and 3.7.3 (b), for every $n \in \mathbb{N}$, $x \geq 0$,

$$\mathbb{E}[\mu^{\mathcal{D}_n}[0, (s^{\mathcal{D}_n} / \sigma^{\mathcal{D}_n})x]] = \sigma^{\mathcal{D}_n} \mathbb{E}[\tilde{\mu}^{\mathcal{D}_n}[0, x]] \leq 2x\sigma^{\mathcal{D}_n} = o(N^{\mathcal{D}_n}).$$

So, writing for every $\mathcal{D} \in \Omega_{\mathcal{D}}$, $t^{\mathcal{D}} := \inf \{l \in \mathbb{N}, \mathbb{E}[\mu^{\mathcal{D}}[0, l]] \geq N^{\mathcal{D}}/2\}$, we have for every $x > 0$ and $n \in \mathbb{N}$ large enough,

$$\mathbb{E}[\mu^{\mathcal{D}_n}[0, t^{\mathcal{D}_n}]] > \mathbb{E}[\mu^{\mathcal{D}_n}[0, (s^{\mathcal{D}_n} / \sigma^{\mathcal{D}_n})x]].$$

Thus, $(\sigma^{\mathcal{D}_n} / s^{\mathcal{D}_n})t^{\mathcal{D}_n} \rightarrow \infty$.

Moreover, since for every $n \in \mathbb{N}$, $x \mapsto \mathbb{E}[\tilde{\mu}^{\mathcal{D}_n}[0, x]]$ is increasing, we have by (a) and by bounded convergence, for every $0 \leq x \leq y$,

$$\int_x^y \frac{dl}{l \mathbb{E}[\tilde{\mu}^{\mathcal{D}_n}[0, l]]} \rightarrow \int_x^y \frac{dl}{l \mathbb{E}[\mu^{\Theta}[0, l]]}.$$

Hence, by Lemma 3.7.4 and since $(\sigma^{\mathcal{D}_n} / s^{\mathcal{D}_n})t^{\mathcal{D}_n} \rightarrow \infty$, we have for every $x \geq 0$,

$$\liminf_{n \rightarrow \infty} \int_x^{\frac{\sigma^{\mathcal{D}_n}}{s^{\mathcal{D}_n}} t^{\mathcal{D}_n}} \frac{dl}{l \mathbb{E}[\tilde{\mu}^{\mathcal{D}_n}[0, l]]} \geq \liminf_{n \rightarrow \infty} \int_x^{\frac{\sigma^{\mathcal{D}_n}}{s^{\mathcal{D}_n}} t^{\mathcal{D}_n}} \frac{dl}{2l \mathbb{E}[\tilde{\mu}^{\mathcal{D}_n}[0, l]]} \geq \frac{1}{2} \int_x^{\infty} \frac{dl}{l \mathbb{E}[\mu^{\Theta}[0, l]]}.$$

It directly follows from Assumption 3.7 (i) that $\int_{\mathbb{R}^+} \frac{dl}{l \mathbb{E}[\mu^{\Theta}[0, l]]} < \infty$. Therefore, $\mu^{\Theta} = \infty$, and by Theorem 2.3.3 the Θ -ICRT is a.s. compact. \square

3.7.2 Height of \mathcal{D} -trees: Proof of Theorem 3.4.4 (a)

Fix $\mathcal{D} \in \Omega_{\mathcal{D}}$ and $x \in \mathbb{R}^+$. Our proof of Theorem 3.4.4 (a) is based on the following inequality

$$H(T) \leq H(T_{Y_{j_0}}) + \sum_{j=0}^{K-1} d_H(T_{Y_{j_i}}, T_{Y_{j_{i+1}}}) + d_H(T_{Y_{j_K}}, T), \quad (3.29)$$

where $0 \leq j_0 \leq \dots \leq j_K \leq N$ and $K \in \mathbb{N}$ are well chosen. We will detail this choice later.

Beforehand, let us introduce some notations that will simplify many expressions. First let for every $i \leq j \in \mathbb{N}$,

$$\Delta(i, j) := d_H(T_{Y_i}, T_{Y_j}). \quad (3.30)$$

Then for every $i \in \mathbb{N}$, let:

$$\Delta(0, i) := H(T_{Y_i}) ; \Delta(i, 2^{\geq}) := \max_{1 \leq a \leq s, d_i \geq 2} d(T_{Y_i}, V_a) ; \Delta(2^{\geq}, \infty) := \max_{1 \leq b \leq s} d(\{V_a\}_{1 \leq a \leq s, d_a \geq 2}, V_b).$$

So (3.29) can be rewritten with those notations, and by splitting the last term in two, as:

$$H(T) \leq \Delta(0, j_0) + \sum_{j=0}^{K-1} \Delta(j_i, j_{i+1}) + \Delta(j_K, 2^{\geq}) + \Delta(2^{\geq}, \infty). \quad (3.31)$$

Also recall from (3.27) the definitions of \mathcal{X}_m and ξ_m . Define $(E_i)_{1 \leq i \leq s-1}$ and f as in Lemma 3.5.2. By Lemma 3.5.2 we may assume and will for the rest of the section that a.s. $(f^{\mathcal{D}}, 1/\sigma^{\mathcal{D}})\Upsilon^{\mathcal{D}} = \tilde{\Upsilon}^{\mathcal{D}}$. In particular, for every $0 \leq m \leq N/\sigma$,

$$\xi_m = \min\{i \in \mathbb{N}, \mu[0, Y_i] > \sigma m\}. \quad (3.32)$$

The rest of the section is organized as follows: We upper bound each of the four terms of (3.31) then we sum the upper bounds to prove Theorem 3.4.4 (a).

First we have the following upper bound on $\Delta(0, i) \leq Y_i$:

Lemma 3.7.6. *For every $0 \leq m \leq N/(2\sigma)$ such that $\mathcal{X}_{2m} \leq \sigma/8$,*

$$\mathbb{P}\left(Y_{\xi_m} > 3\frac{s}{\sigma}\mathcal{X}_{2m} + 2\right) \leq 5e^{-m\mathcal{X}_{2m}/2}.$$

Proof. Let $\alpha := \inf\{i \in \mathbb{N}, \mu[0, i] > \sigma m\}$. Let us upper bound α and then $Y_{\xi_m} - \alpha$. First since $(f^{\mathcal{D}}, 1/\sigma^{\mathcal{D}})\Upsilon^{\mathcal{D}} = \tilde{\Upsilon}^{\mathcal{D}}$ we have $\alpha = \inf\{i \in \mathbb{N}, \tilde{\mu}[0, f(i)] > m\}$. So by monotony of $x \mapsto \tilde{\mu}[0, x]$ and by Lemma 3.7.1,

$$\mathbb{P}(\mathcal{X}_{2m} \leq f(\alpha_m - 1)) \leq \mathbb{P}(\tilde{\mu}[0, \mathcal{X}_{2m}] \leq m) \leq e^{-m\mathcal{X}_{2m}/2}. \quad (3.33)$$

Then, by definition f is increasing and for every $1 \leq i \leq s$, $f(i) = \sum_{k=1}^i E_k^{\mathcal{D}}$ so,

$$\begin{aligned} \mathbb{P}((f(\alpha - 1) < \mathcal{X}_{2m}) \cap (\alpha \geq 2(s/\sigma)\mathcal{X}_{2m} + 2)) &\leq \mathbb{P}(f(2(s/\sigma)\mathcal{X}_{2m} + 1) < \mathcal{X}_{2m}) \\ &\leq \mathbb{P}\left(\sum_{i=1}^{\lceil 2(s/\sigma)\mathcal{X}_{2m} \rceil} E_i < \mathcal{X}_{2m}\right). \end{aligned} \quad (3.34)$$

Furthermore, since $\{E_i\}$ is a family of independent exponential random variables of mean greater than σ/s , we have by classical results on the Gamma distributions (see e.g. [32] Section 2.4),

$$\mathbb{P}\left(\sum_{i=1}^{\lceil 2(s/\sigma)\mathcal{X}_{2m} \rceil} E_i < \mathcal{X}_{2m}\right) \leq e^{-(s/(2\sigma)\mathcal{X}_{2m})}.$$

Therefore, since $m \leq N/(8\sigma) \leq s/(8\sigma)$ we have by (3.33) and (3.34),

$$\mathbb{P}(\alpha \geq 2(s/\sigma)\mathcal{X}_{2m} + 2) \leq 2e^{-m\mathcal{X}_{2m}/2}. \quad (3.35)$$

Now let us upper bound $Y_{\xi_m} - \alpha$. Note from (3.32) that Y_{ξ_m} is the first index of repetition in $\{A_i\}_{1 \leq i \leq s-1}$ after α , and that α is a stopping time for $\{A_i\}_{1 \leq i \leq s-1}$. So by (3.1), conditionally on $\{A_i\}_{1 \leq i \leq \alpha}$, $Y_{\xi_m} - \alpha$ is bounded by a geometric random variable of parameter $(\mu[0, \alpha] - \xi_m + 1)/s$. Hence,

$$\mathbb{P}(Y_{\xi_m} - \alpha \geq (s/\sigma)\mathcal{X}_{2m} | \{A_i\}_{1 \leq i \leq \alpha}) \leq \exp(-\mathcal{X}_{2m}(\mu[0, \alpha] - \xi_m + 1)/\sigma). \quad (3.36)$$

Furthermore by definition of α , by $\mathcal{X}_{2m} \leq \sigma/8$, and by Lemma 3.7.2,

$$\mathbb{P}(\mu[0, \alpha] - \xi_m + 1 \leq \sigma m/2) \leq \mathbb{P}(\xi_m \geq \sigma m/2 + 1) \leq \mathbb{P}(\xi_m \geq 3m\mathcal{X}_m + 1) \leq 2e^{-m\mathcal{X}_{2m}/2}.$$

So by (3.36),

$$\mathbb{P}(Y_{\xi_m} - \alpha \geq (s/\sigma)\mathcal{X}_{2m}) \leq \mathbb{P}(\mu[0, \alpha] - \xi_m - 1 \leq \sigma m/2) + e^{-m\mathcal{X}_{2m}/2} \leq 3e^{-m\mathcal{X}_{2m}/2}.$$

Finally summing the above equation with (3.35) yields the desired inequality. \square

Lemma 3.7.7. *We have the following upper bounds on $\Delta(j, k)$:*

(a) *For every $1 \leq j < k \leq N$, and $t > 0$,*

$$\mathbb{P}(\Delta(j, k) \geq t | \{A_i\}_{1 \leq i \leq Y_j}) \leq (k - j)e^{-t(\mu[0, j] - j)/s}.$$

(b) *For every $0 \leq m \leq N/(2\sigma)$ such that $\mathcal{X}_{2m} \leq \sigma/8$ and $t > 0$,*

$$\mathbb{P}(\Delta(\xi_{m/2}, \xi_m) > t) \leq 2e^{-\mathcal{X}_{2m}m/2} + 3\mathcal{X}_{2m}me^{-t(\sigma m)/(8s)}.$$

Proof. Toward (a), note that $\Delta(j, k) = \max_{j < i \leq k} d(L_i, T_{Y_j})$ where $(L_i)_{1 \leq i \leq N}$ are the different leaves in \mathcal{V}^D that are used in Algorithm 3.1. Also by symmetry of the leaves, for every $j < i \leq k$

$$d(L_i, T_{Y_j}) \stackrel{(d)}{=} d(L_{i+1}, T_{Y_j}) = Y_{j+1} - Y_j. \quad (3.37)$$

Hence, for every $t \in \mathbb{R}^+$

$$\mathbb{P}(\Delta(j, k) \geq t | \{A_i\}_{1 \leq i \leq Y_j}) \leq (k - j)\mathbb{P}(Y_{j+1} - Y_j \geq t).$$

On the other hand, Y_{j+1} is by definition the first repetition in $\{A_i\}_{1 \leq i \leq s-1}$ after Y_j , so by (3.2) $Y_{j+1} - Y_j$ is bounded by a geometric random variable of parameter $(\mu[0, j] - j)/s$ and (a) follows.

Toward (b), note that $\Delta(\xi_{m/2}, \xi_m) = 0$ if $\xi_{m/2} = \xi_m$. So, by union bound, and by Lemma 3.7.2,

$$\begin{aligned} \mathbb{P}(\Delta(\xi_{m/2}, \xi_m) > t) &\leq \mathbb{P}(\xi_m \geq 3\mathcal{X}_{2m}m + 1) + \mathbb{P}((\Delta(\xi_{m/2}, 3\mathcal{X}_{2m}m + 1) > t) \cap (\xi_{m/2} \leq 3\mathcal{X}_{2m}m)) \\ &\leq 2e^{-\mathcal{X}_{2m}m/2} + \mathbb{P}((\Delta(\xi_{m/2}, 3\mathcal{X}_{2m}m + 1) > t) \cap (\xi_{m/2} \leq 3\mathcal{X}_{2m}m)). \end{aligned}$$

Furthermore by (a), since $\xi_{m/2}$ is a stopping time for $(A_i)_{1 \leq i \leq \sigma}$,

$$\mathbb{P}((\Delta(\xi_{m/2}, 3\mathcal{X}_{2m}m + 1) > t) \cap (\xi_{m/2} \leq 3\mathcal{X}_{2m}m)) \leq 3\mathcal{X}_{2m}m\mathbb{E} \left[\exp \left(-t \frac{\mu[0, Y_{\xi_{m/2}}] - 3\mathcal{X}_{2m}m}{s} \right) \right].$$

So, since by (3.32) a.s. $\mu[0, Y_{\xi_{m/2}}] \geq \sigma m/2$, and since $\mathcal{X}_{2m} \leq \sigma/8$,

$$\mathbb{P}((\Delta(\xi_{m/2}, 3\mathcal{X}_{2m}m + 1) > t) \cap (\xi_{m/2} \leq 3\mathcal{X}_{2m}m)) \leq 3\mathcal{X}_{2m}me^{-t(\sigma m)/(8s)}.$$

This concludes the proof. \square

Lemma 3.7.8. *We have the following upper bounds on $\Delta(j, 2^\geq)$:*

(a) *For every $1 \leq j \leq N$ and $t > 0$,*

$$\mathbb{P}(\Delta(j, 2^\geq) > t) \leq s_{\geq 2} e^{-t(\mu[0, Y_j] - j)/s}.$$

(b) *For every $0 \leq m \leq N/(2\sigma)$ such that $\mathcal{X}_{2m} \leq \sigma/8$ and $t > 0$,*

$$\mathbb{P}(\Delta(\xi_m, 2^\geq) > t) \leq 2e^{-\mathcal{X}_{2m}m/2} + s_{\geq 2} e^{-t(\sigma m)/(8s)}.$$

Proof. Toward (a), it is enough to prove that for every a such that $d_a \geq 2$ and $1 \leq j \leq N$,

$$\mathbb{P}(d(V_a, T_{Y_j}) > t \mid \{A_i\}_{1 \leq i \leq Y_j}) \leq e^{-t(\mu[0, Y_j] - j)/s}. \quad (3.38)$$

To this end, let us use Algorithm 3.6. Let $W = (W_i)_{1 \leq i \leq s}$ be any permutation of $\mathcal{V}^{\mathcal{D}}$ such that for every $1 \leq i \leq j$, $W_i = L_i$ and such that $W_{j+1} = V_a$. Note that Algorithm 3.1 and Algorithm 3.6 follow the exact same steps until Y_j so a.s. $T_{Y_j} = T_{Y_j}^W$. Thus since T and T^W are both \mathcal{D} -trees,

$$d(V_a, T_{Y_j}) \stackrel{(d)}{=} d^W(V_a, T_{Y_j}^W).$$

Also note that

$$d^W(V_a, T_{Y_j}^W) \leq (Y_{j+1} - Y_j) \quad (3.39)$$

since one of the two following cases must happen:

- Either $V_a \notin (A_i)_{1 \leq i \leq Y_{j+1}}$ and then Algorithm 3.1 and Algorithm 3.6 follow the exact same steps until $Y_{j+1} - 1$. And at the next step L_{j+1} is "relabelled" V_a . So $d^W(V_a, T_{Y_j}^W) = (Y_{j+1} - Y_j)$.
- Or $V_a \in (A_i)_{1 \leq i \leq Y_{j+1}}$, and then Algorithm 3.1 and Algorithm 3.6 follow the exact same steps until $i := \inf\{k \in \mathbb{N}, A_k = V_a\} \leq Y_{j+1}$. So $d^W(V_a, T_{Y_j}^W) = \max(i - Y_j, 0) \leq Y_{j+1} - Y_j$.

Therefore (3.39) holds, and (3.38) follows from Lemma 3.7.7.

The proof of (b) is similar to the proof of Lemma 3.7.7 (b). We omit the details. \square

Lemma 3.7.9. *If $s \geq 2$ then for every $t \geq 0$,*

$$\mathbb{P}\left(\Delta(2^\geq, \infty) > 3 + t \frac{s}{\sigma} + \frac{\ln(N)}{\ln((s-1)/s_1)}\right) \leq e^{-tN/\sigma}.$$

Proof. First, note that $\Delta(2^\geq, \infty)$ corresponds to the length of the largest path whose vertices have degree 0 or 1. Furthermore, since those paths cannot contain any glue point,

$$\Delta(2^\geq, \infty) \leq 2 + \max\{(j-i), 1 \leq i \leq j \leq s-1, (V_{A_{i+1}}, V_{A_{i+2}}, \dots, V_{A_j}) \subset \mathcal{V}_1\}.$$

where \mathcal{V}_1 is the set of vertices with $d_i = 1$. Then by a simple union bound, for every $t \in \mathbb{N}$,

$$\begin{aligned} \mathbb{P}(\Delta(2^\geq, \infty) \geq 2 + t) &\leq \sum_{i=1}^{s-t-1} \mathbb{P}(V_{A_i} \notin \mathcal{V}_1, V_{i+1}, \dots, V_{i+t} \in \mathcal{V}_1) \\ &= \sum_{i=1}^{s-t-1} \frac{s-1-s_1}{s-1} \frac{s_1}{s-2} \frac{s_1-1}{s-3} \dots \frac{s_1-t+1}{s-t-1} \\ &= (s-1-s_1) \frac{s_1}{s-1} \frac{s_1-1}{s-2} \dots \frac{s_1-t+1}{s-t} \leq 2N \left(\frac{s_1}{s-1}\right)^t, \end{aligned}$$

where we use for the last inequality the fact that $(s - 1 - s_1) = N + s_{\geq 2} \leq 2N$ and $s_1 \leq s - 1$. Thus, by monotony of $t \mapsto \mathbb{P}(\Delta(2^{\geq}, \infty) \geq t)$, we have for every $t > 0$,

$$\mathbb{P}(\Delta(2^{\geq}, \infty) > 3 + t) \leq 2N \left(\frac{s_1}{s - 1} \right)^t.$$

Therefore,

$$P := \mathbb{P} \left(\Delta(2^{\geq}, \infty) > 3 + t \frac{s}{\sigma} + \frac{\ln(N)}{\ln((s - 1)/s_1)} \right) \leq 2 \left(\frac{s_1}{s - 1} \right)^{t(s/\sigma)}.$$

Finally, since for every $x, y \geq 0$, $x^y \leq e^{-(1-x)y}$, and since $s = (N + 1) + s_1 + s_{\geq 2}$,

$$P \leq 2e^{-t(s/\sigma)(1-s_1/(s-1))} \leq 2e^{-t(s/\sigma)(N/(s-1))} \leq 2e^{-tN/\sigma}. \quad \square$$

Proof of Theorem 3.4.4 (a). Fix $x \geq 0$. Recall that we want to use (3.31) to upper bound $H(T)$.

Beforehand, let us make some assumptions to exclude some trivial but annoying cases: Since C can be chosen arbitrary large in Theorem 3.4.4 (a), we may assume, without loss of generality, that $x\mathbb{E}[\tilde{\mu}[0, x]] \geq 200$. This assumption together with Lemmas 3.7.3 (b), and 3.7.4 imply that $x \geq 10$. Similarly, we may assume $s \geq 10$. Also we may assume $N \geq 2$, since otherwise $\sigma = 0$ and the result is trivial in this case. Finally we may assume $x \leq \sigma/8$, since otherwise the result is trivial with $c = 1/8$ because $H(T) \leq s$.

We now treat the general case. Let $m := \frac{1}{2}\mathbb{E}[\tilde{\mu}[0, x]]$ then let $K := \inf\{k \in \mathbb{N}, 2^{k+1}m > \frac{N}{8\sigma}\}$. To simplify the forthcoming computations, let for every $i \in \mathbb{N}$, $\alpha_i := 2^i m \mathcal{X}_{2^{i+1}m}$. Note that $\mathcal{X}_{2m} = x \leq \sigma/8$ and that for every $1 \leq i \leq K$, $2^i m \leq \frac{N}{8\sigma}$ so by Lemma 3.7.3 (a) $\mathcal{X}_{2^{i+1}m} \leq \sigma/8$. Hence, we have by Lemmas 3.7.6, 3.7.7 (b), 3.7.8 (b), 3.7.9 respectively, for every $0 \leq i < K$:

$$\mathbb{P} \left(\Delta(0, \xi_m) > 3 \frac{s}{\sigma} x + 1 \right) \leq 5e^{-\alpha_0/2}, \quad (3.40)$$

$$\mathbb{P} \left(\Delta(\xi_{2^i m}, \xi_{2^{i+1} m}) > \frac{s}{\sigma} \left(2 \frac{\ln(24\alpha_i)}{2^i m} + \frac{x}{2^{i/2}} \right) \right) \leq 2e^{-\alpha_i/2} + e^{-2^{i/2-4} x m}, \quad (3.41)$$

$$\mathbb{P} \left(\Delta(\xi_{2^K m}, 2^{\geq}) > \frac{s}{\sigma} \left(2 \frac{\ln(s_{\geq 2})}{2^K m} + \frac{x}{2^{K/2}} \right) \right) \leq 2e^{-\alpha_K/2} + e^{-2^{K/2-4} x m}, \quad (3.42)$$

$$\mathbb{P} \left(\Delta(2^{\geq}, \infty) > 3 + \frac{s}{\sigma} x + \frac{\ln(N)}{\ln((s - 1)/s_1)} \right) \leq e^{-xN/\sigma}. \quad (3.43)$$

Now let

$$A := 3 \frac{s}{\sigma} x + 1 + \sum_{i=0}^{K-1} \frac{s}{\sigma} \left(2 \frac{\ln(24\alpha_i)}{2^i m} + \frac{x}{2^{i/2}} \right) + \frac{s}{\sigma} \left(2 \frac{\ln(s_{\geq 2})}{2^K m} + \frac{x}{2^{K/2}} \right) + 3 + \frac{s}{\sigma} x + \frac{\ln(N)}{\ln((s - 1)/s_1)},$$

and let

$$B := 5e^{-\alpha_0/2} + \sum_{i=0}^K \left(2e^{-\alpha_i/2} + e^{-2^{i/2-4} x m} \right) + e^{-xN/\sigma}.$$

So that by (3.31), and (3.40)+(3.41)+(3.42)+(3.43),

$$\mathbb{P}(H(T) \geq A) \leq B, \quad (3.44)$$

and it only remains to upper bound A and B .

Toward upper bounding B , note that $x \mapsto \mathbb{E}[\tilde{\mu}[0, x]]$ is increasing, so $x \mapsto \mathcal{X}_x$ is increasing, and thus for every $0 \leq i \leq K-1$,

$$\alpha_i = 2^i m \mathcal{X}_{2^{i+1}m} \geq 2^i (m \mathcal{X}_{2m}) = 2^i \alpha_0.$$

Then by definition of m , $xm = \mathcal{X}_{2m}m = \alpha_0$, and by assumption, $x\mathbb{E}[\tilde{\mu}[0, x]] \geq 200$, so by Lemma 3.7.4, $\alpha_0 = x\mathbb{E}[\tilde{\mu}[0, x]]/2 \geq 50$. Hence, by standard comparisons with geometric sums,

$$\sum_{i=0}^K \left(2e^{-\alpha_i/2} + e^{-2^{i/2-4}xm} \right) \leq \sum_{i=0}^K \frac{1}{2^i} \left(2e^{-\alpha_0/2} + e^{-\alpha_0/16} \right) \leq 90e^{-\alpha_0/16}.$$

Furthermore note that, $\alpha_0 = x\mathbb{E}[\tilde{\mu}[0, x]]/2 \leq x\mathbb{E}[\tilde{\mu}[0, \infty]]/2 = xN/(2\sigma)$. So $e^{-xN/\sigma} \leq e^{-\alpha_0/16}$ and

$$B = 5e^{-\alpha_0/2} + \sum_{i=0}^K \left(2e^{-\alpha_i/2} + e^{-2^{i/2-4}xm} \right) + e^{-xN/\sigma} \leq 100e^{-\alpha_0/16}.$$

Therefore, by Lemma 3.7.4,

$$B \leq 100e^{-x\mathbb{E}[\tilde{\mu}[0, x]]/32}. \quad (3.45)$$

We now upper bound A . First, some straightforward inequalities using $x \geq 10$, $s \leq \sigma$, $s \geq s_1 + 2$ (by $N \geq 2$), $2^{K+1}m > N/(8\sigma)$ and $s \geq 10$ gives:

$$A \leq 100 \frac{s}{\sigma} \left(x + \frac{\sigma}{s} \frac{\ln(N)}{\ln(s/s_1)} + \frac{\sigma}{N} \ln(s \geq 2) \right) + 2 \frac{s}{\sigma} \sum_{i=0}^{K-1} \frac{\ln(24\alpha_i)}{2^i m}. \quad (3.46)$$

Then it remains to upper bound $\sum_{i=0}^{K-1} \frac{\ln(48\alpha_i)}{2^i m}$. To this end, we may assume that $K \geq 1$ since otherwise the sum is null. It directly follows that $2^K m \leq N/(8\sigma)$. Moreover, we have by Lemma 3.7.3 (b) and since for every $1 \leq i \leq s$, $\alpha_i \geq 50$,

$$\sum_{i=0}^{K-1} \frac{\ln(48\alpha_i)}{2^i m} \leq 2 \sum_{i=0}^{K-1} \frac{\ln(\alpha_i)}{2^i m} = 2 \sum_{i=0}^{K-1} \frac{\ln(\mathcal{X}_{2^{i+1}m} 2^i m)}{2^i m} \leq 4 \sum_{i=0}^{K-1} \frac{\ln(\mathcal{X}_{2^{i+1}m})}{2^i m}. \quad (3.47)$$

Next, we compare the last sum with the integral that appears in Theorem 3.4.4 (a):

$$\begin{aligned} \int_{\mathcal{X}_{2m}}^{\mathcal{X}_{m2K}} \frac{dl}{l\mathbb{E}[\tilde{\mu}[0, l]]} &= \sum_{i=1}^{K-1} \int_{\mathcal{X}_{m2^i}}^{\mathcal{X}_{m2^{i+1}}} \frac{dl}{l\mathbb{E}[\tilde{\mu}[0, l]]} \\ &\geq \sum_{i=1}^{K-1} \int_{\mathcal{X}_{m2^i}}^{\mathcal{X}_{m2^{i+1}}} \frac{dl}{lm2^{i+1}} \\ &= \sum_{i=1}^{K-1} \frac{\ln(\mathcal{X}_{m2^{i+1}}) - \ln(\mathcal{X}_{m2^i})}{2^{i+1}m} \\ &\geq \sum_{i=1}^{K-1} \frac{\ln(\mathcal{X}_{2^{i+1}m})}{2^{i+2}m} - \frac{\ln(\mathcal{X}_{2m})}{m}. \end{aligned}$$

So by (3.47), and since $x = \mathcal{X}_{2m} \geq 10$ and $m \geq 1/6$ (by Lemma 3.7.3 (c)),

$$\sum_{i=0}^{K-1} \frac{\ln(48\alpha_i)}{2^i m} \leq 16 \int_x^{\mathcal{X}_{m2K}} \frac{dl}{l\mathbb{E}[\tilde{\mu}[0, l]]} + 4 \frac{\ln(x)}{m} \leq 16 \int_x^{\mathcal{X}_{m2K}} \frac{dl}{l\mathbb{E}[\tilde{\mu}[0, l]]} + 24x.$$

Hence, by Lemma 3.7.4,

$$\sum_{i=0}^{K-1} \frac{\ln(48\alpha_i)}{2^i m} \leq 32 \int_x^{\mathcal{X}_{m2^K}} \frac{dl}{l\mathbb{E}[\bar{\mu}[0, l]]} + 24x. \quad (3.48)$$

Now let $t := \inf \{l \in \mathbb{N}, \mathbb{E}[\mu[0, l]] \geq N/2\}$. Note that $t \geq (s/\sigma)\mathcal{X}_{m2^K} \geq 1$, since $2^K m \leq \frac{N}{8\sigma}$ and since by Lemma 3.7.4,

$$\mathbb{E} \left[\mu \left[0, \frac{s}{\sigma} \mathcal{X}_{m2^K} \right] \right] = \sigma \mathbb{E}[\bar{\mu}[0, \mathcal{X}_{m2^K}]] \leq 2\sigma \mathbb{E}[\tilde{\mu}[0, \mathcal{X}_{m2^K}]] = m2^{K+1} \leq N/4.$$

Therefore by (3.48) and $x \geq 1$,

$$\sum_{i=0}^{K-1} \frac{\ln(48\alpha_i)}{2^i m} \leq 32 \int_1^{t\sigma/s} \frac{dl}{l\mathbb{E}[\bar{\mu}[0, l]]} + 24x. \quad (3.49)$$

Finally (3.46) and (3.49) gives,

$$A \leq 200 \frac{s}{\sigma} \left(x + \frac{\sigma}{N} \ln(s_{\geq 2}) + \int_1^{t\sigma/s} \frac{dl}{l\mathbb{E}[\bar{\mu}[0, l]]} + \frac{\sigma}{N} \frac{\ln(N)}{\ln(s/s_1)} \right), \quad (3.50)$$

To sum up, (3.50), (3.45) and (3.44) yield together the desired inequality. \square

3.7.3 GHP convergence of \mathcal{D} -trees: Proof of Theorem 3.4.3 (a)

Recall the assumptions of Theorem 3.4.3 (a): $\mathcal{D}_n \Rightarrow \Theta$, $\mathcal{M}^{\mathcal{D}_n} \rightarrow 0$, and Assumption 3.7. Note from Lemma 3.7.5 (b) that $\mu^\Theta = \infty$. Recall that in this case by Theorem 3.4.2 (b), the following convergence holds weakly in a GP sens:

$$(\mathcal{V}^{\mathcal{D}_n}, (\sigma^{\mathcal{D}_n}/s^{\mathcal{D}_n})d^{\mathcal{D}_n}, \mathcal{M}^{\mathcal{D}_n}) \longrightarrow (\mathcal{T}^\Theta, d^\Theta, \mathfrak{p}^\Theta). \quad (3.51)$$

We need to prove that (3.51) also holds in a GHP sens. To do so, by Lemma 3.3.3, and since \mathfrak{p}^Θ have a.s. support T^Θ (see Theorem 2.3.1), it is enough to prove that (3.51) holds in a GH sens. To this end, the main idea is to show that (i) the first branches of \mathcal{D}_n -trees converge in a GH sens, and that (ii) \mathcal{D}_n -trees are close from their first branches.

Toward (i), it is straightforward to check from Theorem 3.4.1 (b) that we have for every $k \in \mathbb{N}$,

$$\left(T_{Y_k^{\mathcal{D}_n}}^{\mathcal{D}_n}, (\sigma^{\mathcal{D}_n}/s^{\mathcal{D}_n})d^{\mathcal{D}_n}, \mathcal{M}^{\mathcal{D}_n} \right) \xrightarrow{\text{WGH}} \left(T_{Y_k^\Theta}^\Theta, d^\Theta, \mathfrak{p}^\Theta \right). \quad (3.52)$$

Indeed, one can construct both \mathcal{D} -trees and ICRT in $(\mathbb{R}^+)^{\mathbb{N}}$ using a dimension for each branches (see Aldous [10]), and since by Theorem 3.4.1 (b) the cuts and glue points converge, it directly follows that the subtrees obtained from the first branches converge. We omit the straightforward details.

Toward (ii), we adapt the proof of Theorem 3.4.4 (a) to prove the following lemma:

Lemma 3.7.10. *For every $\varepsilon > 0$,*

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow +\infty} \mathbb{P} \left(\frac{\sigma^{\mathcal{D}_n}}{s^{\mathcal{D}_n}} d_H \left(T_{Y_k^{\mathcal{D}_n}}^{\mathcal{D}_n}, T^{\mathcal{D}_n} \right) > \varepsilon \right) = 0.$$

Before proving Lemma 3.7.10, let us explain why it implies together with (3.52), Theorem 3.4.3 (a).

Proof of Theorem 3.4.3 (a). Fix $\varepsilon > 0$. Let $k \in \mathbb{N}$. By the Skorohod representation theorem, we may assume that the convergence (3.52) holds a.s. for the GH topology.

Then, by Lemma 3.7.10, if $k \in \mathbb{N}$ is large enough, we have for every $n \in \mathbb{N}$ large enough,

$$\mathbb{P} \left(\frac{\sigma^{\mathcal{D}_n}}{s^{\mathcal{D}_n}} d_H \left(T_{Y_k^{\mathcal{D}_n}}, T^{\mathcal{D}_n} \right) > \varepsilon \right) \leq \varepsilon.$$

Also, since by Lemma 3.7.5 (c) the Θ -ICRT is a.s. compact, if $k \in \mathbb{N}$ is large enough,

$$\mathbb{P} \left(d_H \left(\mathcal{T}_{Y_k^\Theta}^\Theta, \mathcal{T}^\Theta \right) > \varepsilon \right) \leq \varepsilon.$$

Thus, given the a.s. GH convergence in (3.52), if $k \in \mathbb{N}$ is large enough, we have for every $n \in \mathbb{N}$ large enough,

$$\mathbb{P} \left(d_{\text{GH}} \left((T^{\mathcal{D}_n}, (\sigma^{\mathcal{D}_n}/s^{\mathcal{D}_n})d^{\mathcal{D}_n}, \mathcal{M}^{\mathcal{D}_n}), (\mathcal{T}^\Theta, d^\Theta, \mathfrak{p}^\Theta) \right) > 2\varepsilon \right) \leq 2\varepsilon.$$

Finally, since $\varepsilon > 0$ is arbitrary, Theorem 3.4.3 (a) follows. \square

Our goal for the rest of the section is to prove Lemma 3.7.10. First, let us introduce some notations. For every $\mathcal{D} \in \Omega_{\mathcal{D}}$ let

$$*^{\mathcal{D}} := \frac{\sigma^{\mathcal{D}}}{N^{\mathcal{D}}} \ln(s_{\geq 2}^{\mathcal{D}}) + \int_1^{t^{\mathcal{D}} \sigma^{\mathcal{D}}/s^{\mathcal{D}}} \frac{dl}{l \mathbb{E}[\tilde{\mu}^{\mathcal{D}}[0, l]]} + \frac{\sigma^{\mathcal{D}}}{N^{\mathcal{D}}} \frac{\ln(N^{\mathcal{D}})}{\ln(s^{\mathcal{D}}/s_1^{\mathcal{D}})}.$$

And for every $\mathcal{D} \in \Omega_{\mathcal{D}}$, $x > 0$ let

$$\mathcal{L}_x^{\mathcal{D}} := Y_i^{\mathcal{D}} \quad \text{where} \quad i := \min\{j \in \mathbb{N}, \mu[0, Y_j] > \sigma \mathbb{E}[\tilde{\mu}[0, x]]/2\}. \quad (3.53)$$

By adapting the proof of Theorem 3.4.4 (a) we get the following result.

Lemma 3.7.11. *Fix $\mathcal{D} \in \Omega_{\mathcal{D}}$. Let $x \geq t > 0$. If $x \geq 10$, $s \geq 10$, $N \geq 2$, $x \leq \sigma/8$, and $t \mathbb{E}[\tilde{\mu}[0, x]] > 200$ then*

$$\mathbb{P} \left(d_H(T_{\mathcal{L}_x}, T) > 3 + 200 \frac{s}{\sigma} (t + *) \right) \leq 200 e^{-t \mathbb{E}[\tilde{\mu}[0, x]]/32} + e^{-tN/\sigma}.$$

Proof. We keep the notations of the proof of Theorem 3.4.4 (a). Notably, $d_H(T_{\mathcal{L}_x}, T) = \Delta(\xi_m, \infty)$. First by Lemmas, 3.7.7 (b), 3.7.8 (b), 3.7.9 respectively, we have for every $0 \leq i \leq K-1$,

$$\begin{aligned} \mathbb{P} \left(\Delta(\xi_{2^i m}, \xi_{2^{i+1} m}) > \frac{s}{\sigma} \left(2 \frac{\ln(24\alpha_i)}{2^i m} + \frac{t}{2^{i/2}} \right) \right) &\leq 2e^{-\alpha_i/2} + e^{-2^{i/2-4} t m}. \\ \mathbb{P} \left(\Delta(\xi_{2^K m}, 2^{\geq}) > \frac{s}{\sigma} \left(2 \frac{\ln(s_{\geq 2})}{2^K m} + \frac{t}{2^{K/2}} \right) \right) &\leq 2e^{-\alpha_K/2} + e^{-2^{K/2-4} t m}. \\ \mathbb{P} \left(\Delta(2^{\geq}, \infty) > 3 + \frac{s}{\sigma} t + \frac{\ln(N)}{\ln((s-1)/s_1)} \right) &\leq e^{-tN/\sigma}. \end{aligned}$$

Then note that

$$d_H(T_{\mathcal{L}_x}, T) = \Delta(\xi_m, \infty) \leq \sum_{i=0}^{K-1} \Delta(\xi_{2^i m}, \xi_{2^{i+1} m}) + \Delta(\xi_{2^K m}, 2^{\geq}) + \Delta(2^{\geq}, \infty),$$

so we can upper bound and conclude as in the proof of Theorem 3.4.4 (a). We omit the details. \square

We then deduce Lemma 3.7.10 from Lemma 3.7.11.

Proof of Lemma 3.7.10: Fix $t > 0$. Fix $x \in \mathbb{R}^+$ such that $x \geq 10$, $x \geq t$ and $t\mathbb{E}[\mu^\Theta[0, x]] > 200$. Note that for every n large enough, $s^{\mathcal{D}_n} \geq 10$, $N^{\mathcal{D}_n} \geq 2$, and $x \leq \sigma^{\mathcal{D}_n}/8$. By Lemma 3.7.5 (a), we have for every n large enough $t\tilde{\mu}^{\mathcal{D}_n}[0, x] > 200$. Therefore, by Lemma 3.7.11, for every n large enough,

$$\mathbb{P}\left(d_H\left(T_{\mathcal{L}_x^{\mathcal{D}_n}}^{\mathcal{D}_n}, T^{\mathcal{D}_n}\right) > 3 + 200 \frac{s^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} (t + *^{\mathcal{D}_n})\right) \leq 200e^{-t\mathbb{E}[\tilde{\mu}^{\mathcal{D}_n}[0, x]]/32} + e^{-tN^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n}}. \quad (3.54)$$

Then by Assumption 3.7, we have $*^{\mathcal{D}_n} \rightarrow 0$. Also since $\mathcal{D}_n \Rightarrow \Theta$, we have $s^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n} \rightarrow \infty$ and by Assumption 3.7 (b), we have $N^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n} \rightarrow \infty$. So by (3.54) and Lemma 3.7.5 (a),

$$\limsup_{n \rightarrow \infty} \mathbb{P}\left(d_H\left(T_{\mathcal{L}_x^{\mathcal{D}_n}}^{\mathcal{D}_n}, T^{\mathcal{D}_n}\right) > 300 \frac{s^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} t\right) \leq 200e^{-t\mathbb{E}[\mu^\Theta[0, x]]/32}. \quad (3.55)$$

Furthermore it directly follows from Theorem 3.4.1 (b) that $(\frac{\sigma^{\mathcal{D}_n}}{s^{\mathcal{D}_n}} \mathcal{L}_x^{\mathcal{D}_n})_{n \in \mathbb{N}}$ converges weakly, and that for all $k \in \mathbb{N}$ $\frac{\sigma^{\mathcal{D}_n}}{s^{\mathcal{D}_n}} Y_k^{\mathcal{D}_n} \rightarrow Y_k^{\mathcal{D}_n}$ weakly as $n \rightarrow \infty$. Thus, since a.s. $Y_k^\Theta \rightarrow \infty$ (see Chapter 2),

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\mathcal{L}^{\mathcal{D}_n} < Y_k^{\mathcal{D}_n}) = 0. \quad (3.56)$$

Thus, by (3.55) and (3.56),

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}\left(d_H\left(T_{Y_k^{\mathcal{D}_n}}^{\mathcal{D}_n}, T^{\mathcal{D}_n}\right) > 300 \frac{s^{\mathcal{D}_n}}{\sigma^{\mathcal{D}_n}} t\right) \leq 200e^{-t\mathbb{E}[\mu^\Theta[0, x]]/32}.$$

Finally, since $\mathbb{E}[\mu^\Theta[0, \infty]] = \infty$, taking $x \rightarrow \infty$ in (3.7.3) yields the desired results. \square

3.8 Additional remarks

3.8.1 The space $\Omega := \Omega_{\mathcal{D}} \cup \Omega_{\mathcal{P}} \cup \Omega_{\Theta}$, its topology and applications

The aim of this section is twofold: explain why our main results on \mathcal{D} -trees imply our main results on \mathcal{P} -trees and ICRT, and prove that Lévy trees are ICRT with random parameters.

To this end, let us define on Ω a topology coherent with Assumptions 3.1, 3.2, 3.3, 3.4. For every $\mathcal{D} \in \Omega_{\mathcal{D}}$ and $i \in \mathbb{N}$, let $\theta_i^{\mathcal{D}} := d_i^{\mathcal{D}}/\sigma^{\mathcal{D}}$ and let $m^{\mathcal{D}} := \sigma^{\mathcal{D}}/s^{\mathcal{D}}$. For every $\mathcal{P} \in \Omega_{\mathcal{P}}$ and $i \in \mathbb{N}$, let $\theta_i^{\mathcal{P}} := p_i^{\mathcal{P}}/\sigma^{\mathcal{P}}$, $N^{\mathcal{P}} := \infty$ and let $m^{\mathcal{P}} := \sigma^{\mathcal{P}}$. For every $\Theta \in \Omega_{\Theta}$, let $N^\Theta := \infty$ and let $m^\Theta := 0$. For every sequence $\{\Lambda_n\}_{n \in \mathbb{N}}$ in Ω and $\Lambda \in \Omega$, we say that $\Lambda_n \rightarrow^\Omega \Lambda$ if and only if $N^{\Lambda_n} \rightarrow \infty$, $m^{\Lambda_n} \rightarrow m^\Lambda$, and for every $i \in \mathbb{N}$, $\theta_i^{\Lambda_n} \rightarrow \theta_i^\Lambda$.

Lemma 3.8.1. *For every $(\mathcal{D}_n) \in \Omega_{\mathcal{D}}^{\mathbb{N}}$, $(\mathcal{P}_n) \in \Omega_{\mathcal{P}}^{\mathbb{N}}$, $(\Theta_n) \in \Omega_{\Theta}^{\mathbb{N}}$, $\mathcal{P} \in \Omega_{\mathcal{P}}$, $\Theta \in \Omega_{\Theta}$, we have:*

- (a) $\mathcal{D}_n \Rightarrow \mathcal{P} \iff \mathcal{D}_n \rightarrow^\Omega \mathcal{P}$.
- (b) $\mathcal{D}_n \Rightarrow \Theta \iff \mathcal{D}_n \rightarrow^\Omega \Theta$.
- (c) $\mathcal{P}_n \Rightarrow \Theta \iff \mathcal{P}_n \rightarrow^\Omega \Theta$.
- (d) $\Theta_n \Rightarrow \Theta \iff \Theta_n \rightarrow^\Omega \Theta$.
- (e) $\Omega_{\mathcal{D}}$ is dense on Ω .

Proof. Toward (a), if $\mathcal{D}_n \rightarrow^\Omega \mathcal{P}$ then for every $i \in \mathbb{N}$, $d_i^{\mathcal{D}_n}/s^{\mathcal{D}_n} = \theta_i^{\mathcal{D}_n} m^{\mathcal{D}_n} \rightarrow \theta_i^{\mathcal{P}} m^{\mathcal{P}} = p_i$. Also, since $N^{\mathcal{D}_n} \rightarrow \infty$, $s^{\mathcal{D}_n} \rightarrow \infty$. Hence, $\mathcal{D}_n \Rightarrow \mathcal{P}$.

On the other hand, if $\mathcal{D}_n \Rightarrow \mathcal{P}$, then it directly follows from Fubini's Theorem that

$$(m^{\mathcal{D}})^2 = (\sigma^{\mathcal{D}}/s^{\mathcal{D}})^2 = \sum_{i=1}^{\infty} \frac{(d_i^{\mathcal{D}})(d_i^{\mathcal{D}} - 1)}{(s^{\mathcal{D}})^2} \longrightarrow \sum_{i=1}^{\infty} p_i^2 = (\sigma^{\mathcal{P}})^2 = (m^{\mathcal{P}})^2.$$

It then follows that for every $i \in \mathbb{N}$,

$$\theta_i^{\mathcal{D}_n} = (d_i^{\mathcal{D}_n} / s^{\mathcal{D}_n}) / m^{\mathcal{D}_n} \longrightarrow (p_i^{\mathcal{P}}) / m^{\mathcal{P}} = \theta_i^{\mathcal{P}}.$$

Also since $d_1^{\mathcal{D}_n} \rightarrow \infty$, we have $N^{\mathcal{D}_n} \rightarrow \infty$. So $\mathcal{D}_n \rightarrow^\Omega \mathcal{P}$.

The proof of (b), (c), (d) is similar. We omit the details.

Toward (e). Let us show that $\Omega_{\mathcal{P}}$ is included in $\bar{\Omega}_D$, the adherence of Ω_D . Fix $\mathcal{P} \in \Omega_{\mathcal{P}}$. Let $(\mathcal{D}_n)_{n \in \mathbb{N}} \in \Omega_D^{\mathbb{N}}$ such that $s^{\mathcal{D}_n} \sim n$ and such that for every $i \in \mathbb{N}$ and $n \in \mathbb{N}$ large enough, $d_i = \lfloor p_i n \rfloor$. Note that $\mathcal{D}_n \Rightarrow \mathcal{P}$. Hence, since \mathcal{P} is arbitrary $\Omega_{\mathcal{P}} \subset \bar{\Omega}_D$. Similarly, $\Omega_{\Theta} \subset \bar{\Omega}_D$. \square

Let us rewrite Theorem 3.4.1 using \rightarrow^Ω , to deduce Theorem 3.4.1 (c) (d) from (a) (b). First, for every $\mathcal{D} \in \Omega_D$, let $\mathbb{P}_{\text{SB}}^{\mathcal{D}}$ denote the distribution of $(\sigma^{\mathcal{D}} / s^{\mathcal{D}}, 1 / \sigma^{\mathcal{D}}) \Upsilon^{\mathcal{D}}$ in \mathbb{K}_{SB} (see Section 3.3.1). For every $\mathcal{P} \in \Omega_{\mathcal{P}}$, let $\mathbb{P}_{\text{SB}}^{\mathcal{P}}$ denote the distribution of $(\sigma^{\mathcal{P}}, 1 / \sigma^{\mathcal{P}}) \Upsilon^{\mathcal{P}}$. For every $\Theta \in \Omega_{\Theta}$, let $\mathbb{P}_{\text{SB}}^{\Theta}$ denote the distribution of Υ^{Θ} . Theorem 3.4.1 is equivalent to the following lemma.

Lemma 3.8.2 (Theorem 3.4.1). *The map $\Lambda \rightarrow \mathbb{P}_{\text{SB}}^{\Lambda}$ is continuous for the weak SB topology.*

Proof. The result directly follows from the density of Ω_D and Theorem 3.4.1 (a), (b). \square

Similarly let us rewrite Theorem 3.4.2. Let $\Omega_{\mathcal{D}, \mathcal{M}}$ be the set of couple $(\mathcal{D}, \mathcal{M})$ such that $\mathcal{D} \in \Omega_D$ and \mathcal{M} is a probability measure with support on $\mathcal{V}^{\mathcal{D}}$. Let $\Omega_{\mathcal{P}, \mathcal{M}}$ be the set of couple $(\mathcal{P}, \mathcal{M})$ such that $\mathcal{P} \in \Omega_{\mathcal{P}}$ and \mathcal{M} is a probability measure with support on $\mathcal{V}^{\mathcal{P}}$. Let $\Omega_{\Theta, \mathcal{M}}$ be the set of couples (Θ, \mathcal{M}) such that $\Theta \in \Omega_{\Theta}$, $\mu^{\Theta} = \infty$ and \mathcal{M} is the null measure on $\{V_i\}_{i \in \mathbb{N}}$. Let $\Omega_{\Lambda, \mathcal{M}} := \Omega_{\mathcal{D}, \mathcal{M}} \cup \Omega_{\mathcal{P}, \mathcal{M}} \cup \Omega_{\Theta, \mathcal{M}}$. Recall that $\Omega_{\mathcal{P}}$ denotes the set of probability measure on $\{V_i\}_{i \in \mathbb{N}}$, and note that $\Omega_{\Lambda, \mathcal{M}} \subset \Omega \times \Omega_{\mathcal{M}}$. So we may equip $\Omega_{\Lambda, \mathcal{M}}$ with the product topology.

Then, for every $(\mathcal{D}, \mathcal{M}) \in \Omega_{\mathcal{D}, \mathcal{M}}$, let $\mathbb{P}_{\text{GP}}^{\mathcal{D}, \mathcal{M}}$ denote the distribution of $(\mathcal{V}^{\mathcal{D}}, (\sigma^{\mathcal{D}} / s^{\mathcal{D}}) d^{\mathcal{D}}, \mathcal{M})$ in $\Omega_{\Lambda, \mathcal{M}}$. For every $(\mathcal{P}, \mathcal{M}) \in \Omega_{\mathcal{P}, \mathcal{M}}$, let $\mathbb{P}_{\text{GP}}^{\mathcal{P}, \mathcal{M}}$ denotes the distribution of $(\{B_i^{\mathcal{P}}\}_{i \in \mathbb{N}}, \sigma^{\mathcal{P}} d^{\mathcal{P}}, \mathcal{M})$. For every $(\Theta, \mathcal{M}) \in \Omega_{\Theta, \mathcal{M}}$, let $\mathbb{P}_{\text{GP}}^{\Theta, \mathcal{M}}$ denotes the distribution of $(\mathcal{T}^{\Theta}, d^{\Theta}, \mathfrak{p}^{\Theta})$.

Theorem 3.4.2 is equivalent to the following lemma.

Lemma 3.8.3 (Theorem 3.4.2). *The map $(\Lambda, \mathcal{M}) \rightarrow \mathbb{P}_{\text{GP}}^{\Lambda, \mathcal{M}}$ is continuous for the weak GP topology.*

Proof. First, since Ω_D is dense on Ω , it is straightforward to check that $\Omega_{\mathcal{D}, \mathcal{M}}$ is dense on $\Omega_{\Lambda, \mathcal{M}}$. The result then follows from Theorem 3.4.2 (a), (b). \square

We now provides some direct applications of our main results to trees with random degree sequence, and notably Galton-Watson trees. First, note that by Lemma 3.8.1, Ω is a Polish space. Similarly note that, $\Omega_{\Lambda, \mathcal{P}}$ is also a Polish space. Then, by classical results on Polish spaces, and by Lemmas 3.8.1, 3.8.2, 3.8.3 we have the following lemma.

Lemma 3.8.4. *Let $\Omega^{\mathbb{P}}$ be the set of random variables on $(\Omega, \rightarrow^\Omega)$, and let $\Omega_{\Lambda, \mathcal{M}}^{\mathbb{P}}$ be the set of random variables on $\Omega_{\Lambda, \mathcal{M}}$. We have the following assertions:*

- (a) *If $(X^n)_{n \in \mathbb{N}}$ is a sequence in $(\Omega^{\mathbb{P}})$ which converges weakly toward $X \in \Omega^{\mathbb{P}}$ then $\mathbb{P}_{\text{SB}}^{X^n}$ converges weakly toward \mathbb{P}_{SB}^X for the SB topology.*
- (b) *If $(X^n, \mathcal{M}^n)_{n \in \mathbb{N}}$ is a sequence in $(\Omega_{\Lambda, \mathcal{M}}^{\mathbb{P}})$ which converges weakly toward $(X, \mathcal{M}) \in \Omega_{\Lambda, \mathcal{M}}^{\mathbb{P}}$ then $\mathbb{P}_{\text{GP}}^{X^n, \mathcal{M}^n}$ converges weakly toward $\mathbb{P}_{\text{GP}}^{X, \mathcal{M}}$ for the GP topology.*

Proof. Since Ω and $\Omega_{\Lambda, \mathcal{P}}$ are Polish spaces, (a) directly from Lemma 3.8.2, and (b) directly follows from Lemma 3.8.3. We omit the straightforward details. \square

Remarks. • Le Gall and Duquesne [50, 49] prove that in some specific cases the degree distribution of Galton-Watson converge (in order to study their Luckasiewicz walk). In those cases, by Lemma 3.8.4 (a), (b), Galton-Watson trees converge toward a "mix" of ICRT.

• Furthermore, Le Gall and Duquesne [50, 49] also proves that Lévy trees appears as GP limits of Galton-Watson trees. Thus, by unicity of the limits, and by Lemma 3.8.4 (b), this proves that Lévy trees "are" ICRT with random parameters.

• Finally, although the GP topology is far more popular than the new SB topology, (a) can be used to study Galton-Watson trees in the condensation cases, in which there is no GP convergence.

To conclude the section, we briefly explain how to adapt the proof of Theorem 3.4.3 (a) and Theorem 3.4.4 (a) to prove Theorem 3.4.3 (b) (c) and Theorem 3.4.4 (b) (c). Those results cannot be proved directly by the topological argument introduced in this section. However, they can be proved by adapting the proof in Section 3.7 by using Algorithm 3.4 instead of using Algorithm 3.5. The proof goes the same way with several simplifications, up to two arguments. To avoid doing the same work twice, let us briefly focus on those argument and omit the details for the rest of the proof.

The first one lies in the proof of Lemma 3.7.7 (a) where we explicitly used the symmetry of the leaves of \mathcal{D} -trees to prove (3.37). The second lies in the proof of Lemma 3.7.8 (a) where we used the Algorithm 3.6 to prove (3.38). Instead of using such arguments for \mathcal{P} -trees and ICRT, one can use the density of $\Omega_{\mathcal{D}}$ on Ω and take the limits of (3.37) and (3.38) by using Theorem 3.4.1 to prove similar equations.

3.8.2 Re-rooting principle for ICRT

We present here the foundations of a powerful computation tool for ICRT: the re-rooting principle. Although, it is not used in this paper, it will be extensively used in a forthcoming paper [25]. It is suitably explained here since it requires a slight modification of Theorem 3.4.1 (b) and of its proof.

The main idea is that the ICRT is "invariant under re-rooting". To be more precise, for every $k \in \mathbb{N}$, let the ICRT rooted at X_k be the ICRT "conditioned" on having $X_k = 0$. More rigorously, let the ICRT rooted at X_k be the tree defined by letting $X_k = 0$ in the first line of Algorithm 3.4 instead of its normal value, and then by defining every other variables the same way. By "invariance under re-rooting of the ICRT", we mean every equality in distribution between the ICRT and the re-rooted ICRT. To prove such an equality, a natural approach is to simply use invariance under re-rooting in the discrete setting and then to take the limit.

To this end, we can not use our previous results on \mathcal{D} -trees as \mathcal{D} -trees are not invariant under re-rooting. Indeed, the "graph structure" of a \mathcal{D} -tree conditioned on its roots depends on its root. For this reason, let us briefly adapt some of our main results to trees with pure degree sequence.

Recall that a tree T have pure degree sequence $(d_i)_{1 \leq i \leq s}$ if and only if T have vertices $(V_i)_{1 \leq i \leq s}$, and for every $1 \leq i \leq s$, V_i have degree d_i (not d_i children). For every $\mathcal{D} \in \Omega_{\mathcal{D}}$, a pure \mathcal{D} -tree is a uniform tree with pure degree sequence $(d_1^{\mathcal{D}} + 1, \dots, d_s^{\mathcal{D}} + 1, 1)$. We have the following construction for pure \mathcal{D} -tree:

Algorithm 3.7. Stick-breaking construction of a pure \mathcal{D} -tree rooted at V_k .

- Let $A^{\mathcal{D},k} = (A_i^{\mathcal{D},k})_{0 \leq i \leq s-1} := (V_k, A_1^{\mathcal{D}}, A_2^{\mathcal{D}}, \dots, A_{s-1}^{\mathcal{D}})$.
- Let $T_0^{\mathcal{D}} := (\{A_0\}, \emptyset)$ then for every $1 \leq i \leq s$ let

$$T_i^{\mathcal{D}} := \begin{cases} T_{i-1} \cup \{A_{i-1}, A_i\} & \text{if } A_i \notin T_{i-1}. \\ T_{i-1} \cup \{A_{i-1}, L_{\inf\{k, L_k \notin T_{i-1}\}}\} & \text{if } A_i \in T_{i-1} \text{ or } i = s. \end{cases}$$

- Let $T^{\mathcal{D},k}$ denote the rooted tree (T_s, V_k) .

Lemma 3.8.5. *For every $\mathcal{D} \in \Omega_{\mathcal{D}}$ and $1 \leq k \leq s$, $T^{\mathcal{D},k}$ is a pure \mathcal{D} -tree rooted at V_k .*

Proof. Note that a pure \mathcal{D} -tree rooted at V_k is a $(d_1, \dots, d_{k-1}, d_k + 1, d_{k+1}, \dots, d_s, 1) = \mathcal{D}'$ -tree conditioned on having root V_k . The result then follows from Algorithm 3.1 and Proposition 3.2.1. \square

Let us define a list of notations for pure \mathcal{D} -tree rooted at V_k similar to those for \mathcal{D} -trees: For every $1 \leq i \leq s$ let $(X_i^{\mathcal{D},k}) = \inf\{0 \leq j \leq s-1, A_j^{\mathcal{D},k} = V_i\}$. Then let $\mu^{\mathcal{D},k} = \sum_{i=1}^s \delta_{X_i}(d_i - 1)$. Let $(Y_i^{\mathcal{D},k})_{1 \leq i \leq N^{\mathcal{D}}+1}$ be the indexes of repetitions in $(A^{\mathcal{D},k})_{0 \leq k \leq s-1}$. Then for $1 \leq i \leq N^{\mathcal{D}} + 1$, let $Z_i^{\mathcal{D},k}$ denotes the smallest integer z such that $A_z^{\mathcal{D},k} = A_{Y_i^{\mathcal{D},k}}$. By convention, for $i > N^{\mathcal{D}} + 1$ let $Y_i^{\mathcal{D},k} = Z_i^{\mathcal{D},k} = +\infty$. And finally let $\Upsilon_i^{\mathcal{D},k} = ((X_i^{\mathcal{D},k}, Y_i^{\mathcal{D},k}, Z_i^{\mathcal{D},k})_{i \in \mathbb{N}}, \mu^{\mathcal{D},k})$.

We use the superscript \mathcal{D}, k to indicate that we work with pure \mathcal{D} -tree rooted at V_k . Also, to distinguish re-rooted ICRT from ICRT, we use the superscript Θ, k for the variables defined for the Θ -ICRT "conditioned" on having $X_k = 0$.

We have the following generalisation of Theorem 3.4.1 (b) for pure \mathcal{D} -tree rooted at V_k .

Proposition 3.8.6. *Let $(\mathcal{D}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{D}}^{\mathbb{N}}$ and let $(k_n)_{n \in \mathbb{N}} \in \mathbb{N}^{\mathbb{N}}$ such that for every $n \in \mathbb{N}$, $k_n \leq s^{\mathcal{D}_n}$. The following convergences hold weakly for the SB topology (see Section 3.3.1)*

- (a) *If $\mathcal{D}_n \Rightarrow \Theta$ and $k_n \rightarrow \infty$, then $(\sigma^{\mathcal{D}_n}/s^{\mathcal{D}_n}, 1/\sigma^{\mathcal{D}_n})\Upsilon^{\mathcal{D}_n, k_n} \xrightarrow{WSB} \Upsilon^{\Theta}$.*
- (b) *If $\mathcal{D}_n \Rightarrow \Theta$ and $k_n \rightarrow k$, then $(\sigma^{\mathcal{D}_n}/s^{\mathcal{D}_n}, 1/\sigma^{\mathcal{D}_n})\Upsilon^{\mathcal{D}_n, k_n} \xrightarrow{WSB} \Upsilon^{\Theta, k}$.*

Proof. One can directly adapt the proof of Theorem 3.4.1 (b) by letting in Algorithm 3.5 $\tilde{X}_{k_n} := 0$ instead of its normal value. We omit the details as nothing else fundamentally changes. \square

Let us now give a quick example of the re-rooting principle using Proposition 3.8.6.

Lemma 3.8.7. *Let $\Theta \in \Omega_{\Theta}$ such that $\mu^{\Theta} = \infty$, and let $k \in \mathbb{N}$. Assume that a.s. there exists a probability measure $p^{\mathcal{D},k}$ on $\mathcal{T}^{\Theta, k}$ such that $\frac{1}{n} \sum_{i=1}^n \delta_{Y_i^{\Theta, k}} \rightarrow p^{\Theta, k}$, then we have the following equality in distribution for the Gromov–Prokhorov topology:*

$$(\mathcal{T}^{\Theta}, d^{\Theta}, \mathbf{p}^{\Theta}) \stackrel{(d)}{=} (\mathcal{T}^{\Theta, k}, d^{\Theta, k}, \mathbf{p}^{\Theta, k}).$$

Proof. First by Lemma 3.8.1 there exists $(\mathcal{D}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{D}}^{\mathbb{N}}$ such that $\mathcal{D}_n \Rightarrow \Theta$. We may furthermore assume that for every $n \in \mathbb{N}$, $s^{\mathcal{D}_n} \geq k$. Then recall that for every $\mathcal{D} \in \Omega_{\mathcal{D}}$, $(L_i^{\mathcal{D}})_{i \leq N^{\mathcal{D}}}$ denotes the leaves of $\mathcal{V}^{\mathcal{D}}$. Also for every $\mathcal{D} \in \Omega_{\mathcal{D}}$ and $i \leq s^{\mathcal{D}}$ let $d^{\mathcal{D}, i}$ denotes the graph distance on $\mathcal{T}^{\mathcal{D}, i}$.

Now, by Proposition 3.8.6 (a) we have the following convergence in distribution (see e.g. Section 3.6.2 where we detail this implication for \mathcal{D} -trees),

$$(d^{\mathcal{D}_n, s^{\mathcal{D}_n}}(L_i^{\mathcal{D}_n}, L_j^{\mathcal{D}_n}))_{i, j \in \mathbb{N}} \xrightarrow{(d)} (d^{\Theta}(Y_i^{\Theta}, Y_j^{\Theta}))_{i, j \in \mathbb{N}}.$$

Similarly by Proposition 3.8.6 (b),

$$(d^{\mathcal{D}_n, k}(L_i^{\mathcal{D}_n}, L_j^{\mathcal{D}_n}))_{i, j \in \mathbb{N}} \xrightarrow{(d)} (d^{\Theta, k}(Y_i^{\Theta, k}, Y_j^{\Theta, k}))_{i, j \in \mathbb{N}}.$$

Then, since the matrix distance is invariant under re-rooting, we have for every $n \in \mathbb{N}$,

$$(d^{\mathcal{D}_n, s^{\mathcal{D}_n}}(L_i^{\mathcal{D}_n}, L_j^{\mathcal{D}_n}))_{i, j \leq N^{\mathcal{D}_n}} \stackrel{(d)}{=} (d^{\mathcal{D}_n, k}(L_i^{\mathcal{D}_n}, L_j^{\mathcal{D}_n}))_{i, j \leq N^{\mathcal{D}_n}}.$$

Therefore,

$$(d^{\Theta}(Y_i^{\Theta}, Y_j^{\Theta}))_{i, j \in \mathbb{N}} \stackrel{(d)}{=} (d^{\Theta, k}(Y_i^{\Theta, k}, Y_j^{\Theta, k}))_{i, j \in \mathbb{N}}. \quad (3.57)$$

Finally the assumption of the lemma together with Lemma 3.3.2 yield the desired result. \square

Remark. The assumption of Lemma 3.8.7 can be proved by adapting the proof of Theorem 2.3.1 and Proposition 2.3.2. However, it would require far too much work to be done here properly.

Acknowledgment Thanks are due to Nicolas Broutin for many interesting conversations and advices on this paper, and notably for having asking for an upper bound for the height of \mathcal{D} -trees.

3.9 Appendix

3.9.1 Proof that Algorithm 3.1 and Algorithm 3.6 sample \mathcal{D} -trees

Proof of Proposition 3.2.1. Fix $\mathcal{D} \in \Omega_{\mathcal{D}}$. For every \mathcal{D} -tuple A let $(T_i^A)_{1 \leq i \leq s}$ and T^A be the graphs constructed by Algorithm 3.1 from entry (\mathcal{D}, A) . We prove that $A \mapsto T^A$ is a bijection between \mathcal{D} -tuple and tree with degree sequence \mathcal{D} .

First, for every \mathcal{D} -tuple A , note by induction that for every $1 \leq i \leq s$, T_i^A is connected. Furthermore a quick enumeration shows that for every $1 \leq i \leq s$, V_i have d_i children in T^A . Hence, T^A is a tree with degree sequence \mathcal{D} .

Next it is well known that there are $\frac{s!}{d_1! \dots d_s!}$ trees with degree sequence \mathcal{D} and that there are also $\frac{s!}{d_1! \dots d_s!}$ \mathcal{D} -tuples. Hence, it is enough to prove that the map $A \mapsto T^A$ is injective.

Let A be a \mathcal{D} -tuple. Then for every $1 \leq i < s^{\mathcal{D}} - 1$ let $k_i := \inf\{k \in \mathbb{N}, L_k \notin T_i^A\}$. Note that A_1 is the root of T^A , and that for every $1 \leq i < s^{\mathcal{D}} - 1$ the following assertions hold:

- If L_{k_i} is not a child of A_i then A_{i+1} is a child of A_i and the edge (A_i, A_{i+1}) "disconnects" the root from L_{k_i} .
- If L_{k_i} is a child of A_i then A_{i+1} is the vertex in $\{A_1, \dots, A_i\}$ which is the closest to L_{k_i+1} .

As a result given T^A and (A_1, \dots, A_i) , the vertex A_{i+1} is entirely determined. Hence, the map $A \mapsto T^A$ is injective. \square

Proof of Proposition 3.6.1. The proof is similar to the proof of Proposition 3.2.1. So we keep similar notations and skip some details to focus on the points which really differ.

Fix $\mathcal{D} \in \Omega_{\mathcal{D}}$, and fix $W = \{W_i\}_{1 \leq i \leq s^{\mathcal{D}}}$ an arbitrary permutation of $\mathcal{V}^{\mathcal{D}}$. For every \mathcal{D} -tuple A , let $(T_i^A)_{1 \leq i \leq s}$ and T^A be the graphs constructed by Algorithm 3.6 from entry (\mathcal{D}, W, A) . We prove that $A \mapsto T^A$ is a bijection between \mathcal{D} -tuple and tree with degree sequence \mathcal{D} .

First for similar reasons, for every \mathcal{D} -tuple A , T^A is a tree with degree sequence \mathcal{D} . So it is enough to prove that the map $A \mapsto T^A$ is injective.

Let A be a \mathcal{D} -tuple. Then for every $1 \leq i < s^{\mathcal{D}} - 1$ let $k_i := \inf\{k \in \mathbb{N}, W_k \notin T_i^A\}$. Note that A_1 is the root of T^A , and that for every $1 \leq i < s^{\mathcal{D}} - 1$ the following assertions holds.

- If W_{k_i} is not a child of A_i then A_{i+1} is a child of A_i and the edge (A_i, A_{i+1}) "disconnects" the root from W_{k_i} .
- If W_{k_i} is a child of A_i then either $A_{i+1} = W_{k_i}$ or $A_{i+1} \in T_i^A$. In both cases, A_{i+1} is the vertex in $T_i^A \cup \{W_{k_i}\}$ which is the closest to W_{k_i+1} .

In the second assertion, note that $k_{i+1} = \inf\{k \in \mathbb{N}, W_k \notin T_i^A \cup \{W_i\}\}$ does not depend on the choice of A_{i+1} . As a result given T^A and (A_1, \dots, A_i) , the vertex A_{i+1} is entirely determined. Hence, the map $A \mapsto T^A$ is injective. \square

3.9.2 GP and GH convergence imply GHP convergence: Proof of Lemma 3.3.3

Beforehand, let us introduce the covering numbers. For every metric space (X, d) , and $\varepsilon > 0$ let $\mathcal{N}_{\varepsilon}(X, d)$ be the minimal number of closed balls of radius ε to cover X . Note that if (X, d) and (X', d') are isometric spaces then for every $\varepsilon > 0$, $\mathcal{N}_{\varepsilon}(X, d) = \mathcal{N}_{\varepsilon}(X', d')$, hence for every $\varepsilon > 0$, $\mathcal{N}_{\varepsilon}$ is well defined on \mathbb{K}_{GH} . Furthermore, for every $\varepsilon > 0$ $\mathcal{N}_{\varepsilon}$ is a measurable function on \mathbb{K}_{GH} .

Now, let $((X^n, d^n, p^n))_{n \in \mathbb{N}}$ and (X, d, p) be random measured metric spaces defined as in Lemma 3.3.3. That is, $([X^n, d^n, p^n])_{n \in \mathbb{N}}$ and $[X, d, p]$ are GHP measurable, $[X^n, d^n, p^n] \xrightarrow{\text{WGP}} [X, d, p]$, $[X^n, d^n, p^n] \xrightarrow{\text{WGH}} [X, d]$, and p have a.s. support X . It directly follows from the GH convergence that:

- (i) The diameter of $[X^n, d^n]$ converges weakly as $n \rightarrow \infty$ toward the diameter of $[X, d]$.
- (ii) For every $\varepsilon > 0$, $(\mathcal{N}_\varepsilon[X^n, d^n])_{n \in \mathbb{N}}$ is tight (see Burago Burago Ivanov [?] Section 7.4).

Hence, by [1] Theorem 2.4, $([X^n, d^n, p^n])_{n \in \mathbb{N}}$ is tight in a GHP sens.

Now, let $[X', d', p']$ be a GHP subsequential limit of $([X^n, d^n, p^n])_{n \in \mathbb{N}}$. It is enough to show that necessarily $[X, d, p] =^{(d, \text{GHP})} [X', d', p']$.

On the one hand, since $[X^n, d^n] \rightarrow^{\text{WGH}} [X, d, p]$ and $[X^n, d^n, p^n] \rightarrow^{\text{WGHP}} [X, d, p]$ along a suitable sequence, we have $[X, d] =^{(d, \text{GH})} [X', d']$. Hence, we have

$$\forall \varepsilon > 0, \quad \mathcal{N}_\varepsilon([X, d]) =^{(d)} \mathcal{N}_\varepsilon([X', d']). \quad (3.58)$$

On the other hand, since $[X^n, d^n, p^n] \rightarrow^{\text{WGP}} [X, d, p]$ and $[X^n, d^n, p^n] \rightarrow^{\text{WGHP}} [X, d, p]$ along a suitable sequence, we have $[X, d, p] =^{(d, \text{GP})} [X', d', p']$. So, we may assume that a.s. $d_{\text{GP}}([X, d, p], [X', d', p']) = 0$. So, by definition of the GP topology, a.s. there exists a metric space S and isometric embeddings $\phi : X \rightarrow S$, $\phi' : X' \rightarrow S'$ such that $d_P(\phi_* p, \phi'_* p') = 0$, and thus $\phi_* p = \phi'_* p'$. Hence, if supp denotes the support of a measure, a.s. $\phi(\text{supp}(p)) = \phi'(\text{supp}(p'))$. Therefore, since a.s. $\text{supp}(p) = X$, we have a.s.

$$\forall \varepsilon > 0, \quad \mathcal{N}_\varepsilon(X) = \mathcal{N}_\varepsilon(\text{supp}(p)) = \mathcal{N}_\varepsilon(\text{supp}(p')) \leq \mathcal{N}_\varepsilon(X'). \quad (3.59)$$

Finally, note that (3.58) and (3.59) implies together that a.s. $\text{supp}(p') = X'$. Thus, since a.s. $X = \text{supp}(p)$ and $\phi(p) = \phi(p')$, we have a.s. $\phi(X) = \phi'(X')$. Thus, since a.s. $\phi(p) = \phi(p')$ and $\phi(X) = \phi'(X')$, we have a.s. $d_{\text{GHP}}(X, X') = 0$.

Chapter 4

Limit of connected multigraph with fixed degree sequence

*"You don't write because you want to say something;
you write because you've got something to say."*

F. Scott Fitzgerald

This chapter is adapted from [27].

Motivated by the scaling limits of the connected components of the configuration model, we study uniform connected multigraphs with fixed degree sequence \mathcal{D} and with surplus k . We call those random graphs (\mathcal{D}, k) -graphs. We prove that, for every $k \in \mathbb{N}$, under natural conditions of convergence of the degree sequence, (\mathcal{D}, k) -graphs converge toward either (\mathcal{P}, k) -graphs or (Θ, k) -ICRG (inhomogeneous continuum random graphs). We prove similar results for (\mathcal{P}, k) -graphs and (Θ, k) -ICRG, which have applications to multiplicative graphs. Our approach relies on two algorithms, the cycle-breaking algorithm, and the stick-breaking construction of \mathcal{D} -tree that we introduced in a recent paper. From those algorithms we deduce a biased construction of (\mathcal{D}, k) -graph, and we prove our results by studying this bias.

Contents

4.1	Introduction	110
4.1.1	Motivations	110
4.1.2	Overview of the proof	111
4.2	Notions of convergence	112
4.2.1	Gromov–Prokhorov (GP) topology	112
4.2.2	Gromov–Hausdorff (GH) topology	113
4.2.3	Pointed Gromov–Hausdorff (GH^n) topology	113
4.2.4	Extension to pseudo metric spaces	113
4.3	Constructions of \mathcal{D}-trees, \mathcal{P}-trees and Θ-ICRT	113
4.3.1	\mathcal{D} -trees	113
4.3.2	\mathcal{P} -trees	114
4.3.3	ICRT	114
4.4	Constructions of (\mathcal{D}, k)-graphs, (\mathcal{P}, k)-graphs and (Θ, k)-ICRG	115
4.4.1	Generic gluing and cycle-breaking of discrete multigraphs (see Figure 4.3)	115
4.4.2	(\mathcal{D}, k) -graph	116
4.4.3	(\mathcal{P}, k) -graph	117
4.4.4	(Θ, k) -ICRG	117

4.5 Main results	118
4.5.1 The bias does not diverge	118
4.5.2 Gromov–Prokhorov convergence	118
4.5.3 Gromov–Hausdorff convergence	119
4.6 Study of the bias	120
4.6.1 Proof of Proposition 4.5.1 in the typical case	120
4.6.2 Proof of Proposition 4.5.1 when there are many vertices of degree 2	124
4.6.3 Bias of \mathcal{P} -trees and ICRT	127
4.7 Proof of the main theorems	128
4.7.1 Proof of Theorem 4.5.2	128
4.7.2 Proof of Theorem 4.5.3	129
4.8 Configuration model and multiplicative graphs	129
4.8.1 Definitions	129
4.8.2 Multiplicative multigraphs as local limit of the configuration model	130
4.8.3 Connections with (\mathcal{D}, k) -graphs and (\mathcal{P}, k) -graphs	132
4.9 Appendix A: \mathbb{R}-tree reconstruction problem	133
4.10 Appendix B: \square_c is continuous	135

4.1 Introduction

The present work is a continuation of our previous paper (Chapter 3), where we introduced a stick-breaking construction for \mathcal{D} -trees (uniform tree with fixed degree sequence \mathcal{D}) to prove that, under natural conditions, \mathcal{D} -trees converge, in a GP and a GHP sens, toward either \mathcal{P} -trees or ICRT. Here, we derive from Chapter 3 similar limits for graph versions of those trees, which have applications to multiplicative graphs and to the configuration model.

4.1.1 Motivations

Computer scientists have introduced multiplicative graphs [41, 73, 34] and the configuration model [20, 31] as natural generalizations of the Erdős–Rényi model. They are studied for 2 main reasons: first many tools introduced for the Erdős–Rényi model can also be used to study those graphs, then those models seems closer to real life network thanks to the "inhomogeneity in their degree distribution" (see e.g. Newman [72]). For those reasons, they are currently great models to study the evolution of random networks.

A natural question for any model of evolution is to study their potential phase transitions. It appears that those graphs have an intriguing phase transition where a giant component gets born. We refer the reader to [46] Chapter 1 and references therein for an elaborate discussion of the nature of this transition, and an overview of the literature it generated.

From the point of view of precise asymptotics, a main goal is to study the geometry of the connected components of those graphs in the critical regime. To this end, Addario-Berry, Broutin, and Goldschmidt [4] have developed a general approach in the case of the Erdős–Rényi model. This approach is divided in two main steps:

- First one encodes the random graphs into stochastic processes, and study those processes to deduce several limits for relevant quantities of the largest connected components such as the size, surplus, degrees. This has been noticed in the ground-breaking work of Aldous [12].
- Then, one use those convergences to reduce the problem to a study of a single connected component conditioned on those quantities.

This approach has been further developed for multiplicative graphs and the configuration model in many different regimes. We refer the reader to [4, 24, 23] for the homogeneous case, [42, 47, 54] for the power law case, and [36, 37] for a unified approach for multiplicative graphs. In this paper, we focus on solving (b), under what we believe to be the weakest assumptions. So we reduce the study of the largest connected components to solving (a), which tends to be simpler.

Moreover, we give a universal point of view on those models which can be summarized into the three following points: we describe multiplicative graphs as degenerate configuration model, we extend the unified point of view of Broutin, Duquesne, and Minmin [36, 37] to the configuration model, and we remove the omnipresent randomness assumption in the degree sequence.

4.1.2 Overview of the proof

Fix $k \in \mathbb{N}$. Fix $\{V_i\}_{i \in \mathbb{N}}$ a set of vertices. We say that a multigraph G have degree sequence $\mathcal{D} = (d_1, \dots, d_s)$ if G has vertices (V_1, \dots, V_s) and for every $1 \leq i \leq s$, V_i has degree $d_i + 1$. (This shift of $+1$ will be convenient to simplify many expressions, and be coherent with Chapter 3.) The surplus of a connected multigraph (V, E) is $|E| - |V| + 1$, and is, informally, the number of edges that one needs to delete to transform a multigraph into a tree. A (\mathcal{D}, k) -graph is a uniform connected multigraph with degree sequence \mathcal{D} and surplus k .

Our goal is to study the connected components of the configuration model conditioned on having degree sequence \mathcal{D} and surplus k , which are close from (\mathcal{D}, k) -graphs (see Lemma 4.8.3). To this end, we rely on two algorithms: the stick-breaking construction of \mathcal{D} -trees of Chapter 3, along with the cycle-breaking algorithm introduced by Addario-Berry, Broutin, Goldschmidt, and Miermont [5] which we invert to construct (\mathcal{D}, k) -graph by adding k edges to a biased \mathcal{D} -tree.

We use the cycle breaking algorithm in the following form. Take a connected multigraph with surplus k , repeat k times: choose an edge uniformly among all the edges that can be removed without disconnecting the graph, then cut this edge in the middle. By doing so, we add $2k$ named leaves $(\star_i)_{1 \leq i \leq 2k}$, and keep the degrees of $(V_i)_{i \in \mathbb{N}}$. Note that to invert this algorithm we can intuitively repair the broken edges by gluing the different pairs in $(\star_i)_{1 \leq i \leq 2k}$. Note that however this algorithm is not a bijection, since for each multigraph there are many corresponding trees. To bypass this, we bias each tree by the probability that they were obtained by their corresponding multigraph. This way, we construct a (\mathcal{D}, k) -graph from a biased \mathcal{D} -tree with k additional edges.

Thus, to study the geometry of a (\mathcal{D}, k) -graph, it is enough to study jointly the geometry of a \mathcal{D} -tree, the positions of $(\star_i)_{1 \leq i \leq 2k}$, and the previous bias which is a function of $(d(\star_i, \star_j))_{1 \leq i, j \leq 2k}$. Therefore, it is enough to study precisely the distance matrix between specific vertices of a \mathcal{D} -tree. If the bias was a continuous function of this matrix, then our main results would directly follow from Chapter 3 since the GP convergence of \mathcal{D} -trees implies the convergence of this matrix. However, some extra care is needed since the bias diverges when $(\star_i)_{1 \leq i \leq 2k}$ are close.

Therefore, we need to prove that $(\star_i)_{1 \leq i \leq 2k}$ cannot be too close. More precisely we show, using the structure of \mathcal{D} -trees and of the bias, that it is enough to lower bound $(d(\star_0, \star_i))_{1 \leq i \leq k}$ where \star_0 is a root leaf. We then use our construction of \mathcal{D} -trees, also introduced independently by Addario-Berry, Donderwinkel, Maazoun, and Martin in [6], to lower bound those distances using the k first repetitions in a random tuple.

Finally, since the bias is a function of the subtree spanned by $(\star_i)_{1 \leq i \leq 2k}$, it is also a function of the first branches of the stick-breaking construction. This allows us to consider the limit of the bias, to directly construct the limits of (\mathcal{D}, k) -graphs by biasing the \mathcal{P} -trees and ICRT, introduced by Aldous, Camarri and Pitman [16, 40], and then by gluing the k first pair of leaves.

Plan of the paper: In Section 4.2 we introduce the topologies that we are using. In Section 4.3, we construct \mathcal{D} -trees, \mathcal{P} -trees, Θ -ICRT. In Section 4.4, we construct (\mathcal{D}, k) -graphs, (\mathcal{P}, k) -graphs, and (Θ, k) -ICRG. We state our main results in Section 4.5. We study the bias in section 4.6. We deduce our main results in Section 4.7. Finally we discuss in Section 4.8 some connections between (\mathcal{D}, k) -graph, (\mathcal{P}, k) -graphs, the configuration model, and multiplicative graphs.

Notations: Throughout the paper, similar variables for, \mathcal{D} -trees, (\mathcal{D}, k) -graphs, \mathcal{P} -trees, (\mathcal{P}, k) -graphs, Θ -ICRT, (Θ, k) -ICRG share similar notations. To avoid any ambiguity, the models that we are using and their parameters are indicated by superscripts \mathcal{D} , (\mathcal{D}, k) , \mathcal{P} , (\mathcal{P}, k) , Θ , (Θ, k) . We often drop those superscripts when the context is clear.

Acknowledgment Thanks are due to Nicolas Broutin for many advices on the configuration model and on multiplicative graphs.

4.2 Notions of convergence

4.2.1 Gromov–Prokhorov (GP) topology

A measured metric space is a triple (X, d, μ) such that (X, d) is a Polish space and μ is a Borel probability measure on X . Two such spaces (X, d, μ) , (X', d', μ') are called isometry-equivalent if there exists an isometry $f : X \rightarrow X'$ such that if $f_*\mu$ is the image of μ by f then $f_*\mu = \mu'$. Let \mathbb{K}_{GP} be the set of isometry-equivalent classes of measured metric space. Given a measured metric space (X, d, μ) , we write $[X, d, \mu]$ for the isometry-equivalence class of (X, d, μ) and frequently use the notation X for either (X, d, μ) or $[X, d, \mu]$.

We now recall the definition of the Prokhorov’s distance. Consider a metric space (X, d) . For every $A \subset X$ and $\varepsilon > 0$ let $A^\varepsilon := \{x \in X, d(x, A) < \varepsilon\}$. Then given two (Borel) probability measures μ, ν on X , the Prokhorov distance between μ and ν is defined by

$$d_P(\mu, \nu) := \inf\{\varepsilon > 0: \mu\{A\} \leq \nu\{A^\varepsilon\} \text{ and } \nu\{A\} \leq \mu\{A^\varepsilon\}, \text{ for all Borel set } A \subset X\}.$$

The Gromov–Prokhorov (GP) distance is an extension of the Prokhorov’s distance: For every $(X, d, \mu), (X', d', \mu') \in \mathbb{K}_{\text{GP}}$ the Gromov–Prokhorov distance between X and X' is defined by

$$d_{\text{GP}}((X, d, \mu), (X', d', \mu')) := \inf_{S, \phi, \phi'} d_P(\phi_*\mu, \phi'_*\mu'),$$

where the infimum is taken over all metric spaces S and isometric embeddings $\phi : X \rightarrow S$, $\phi' : X' \rightarrow S$. d_{GP} is indeed a distance on \mathbb{K}_{GP} and $(\mathbb{K}_{\text{GP}}, d_{\text{GP}})$ is a Polish space (see e.g. [1]).

We use another convenient characterization of the GP topology: For every measured metric space (X, d^X, μ^X) let $(x_i^X)_{i \in \mathbb{N}}$ be a sequence of i.i.d. random variables of common distribution μ^X and let $M^X := (d^X(x_i^X, x_j^X))_{(i,j) \in \mathbb{N}^2}$. We recall Lemma 3.3.2 (see also [66]),

Lemma 4.2.1. *Let $(X^n)_{n \in \mathbb{N}} \in \mathbb{K}_{\text{GP}}^{\mathbb{N}}$ and let $X \in \mathbb{K}_{\text{GP}}$. Let $(y_i^X)_{i \in \mathbb{N}}$ be a sequence of random variables on X and let $N^X := (d^X(y_i^X, y_j^X))_{(i,j) \in \mathbb{N}^2}$. If*

$$M^{X^n} \xrightarrow{(d)} N^X \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^n \delta_{y_i^X} \xrightarrow{(d)} \mu^X,$$

then $X^n \xrightarrow{GP} X$.

4.2.2 Gromov–Hausdorff (GH) topology

Let \mathbb{K}_{GH} be the set of isometry-equivalent classes of compact metric space. For every metric space (X, d) , we write $[X, d]$ for the isometry-equivalent class of (X, d) , and frequently use the notation X for either (X, d) or $[X, d]$.

For every metric space (X, d) , the Hausdorff distance between $A, B \subset X$ is given by

$$d_H(A, B) := \inf\{\varepsilon > 0, A \subset B^\varepsilon, B \subset A^\varepsilon\}.$$

The Gromov–Hausdorff distance between $(X, d), (X', d') \in \mathbb{K}_{\text{GH}}$ is given by

$$d_{\text{GH}}((X, d), (X', d')) := \inf_{S, \phi, \phi'} (d_H(\phi(X), \phi'(X'))),$$

where the infimum is taken over all metric spaces S and isometric embeddings $\phi : X \rightarrow S$, $\phi' : X' \rightarrow S$. d_{GH} is indeed a distance on \mathbb{K}_{GH} and $(\mathbb{K}_{\text{GH}}, d_{\text{GH}})$ is a Polish space (see e.g. [1]).

4.2.3 Pointed Gromov–Hausdorff (GH^n) topology

Let $n \in \mathbb{N}$. Let $(X, d, (x_1, \dots, x_n))$ and $(X', d', (x'_1, \dots, x'_n))$ be metric spaces, each equipped with an ordered sequence of n distinguished points (we call such spaces n -pointed metric spaces). We say that these two n -pointed metric spaces are isometric if there exists an isometry ϕ from (X, d) to (X', d') such that for every $1 \leq i \leq n$, $\phi(x_i) = x'_i$.

Let \mathbb{K}_{GH}^n be the set of isometry-equivalent classes of compact metric space. As before, we write $[X, d, (x_1, x_2, \dots, x_n)]$ for the isometry-equivalent class of $(X, d, (x_1, \dots, x_n))$, and denote either by X when there is little chance of ambiguity.

The n -pointed Gromov–Hausdorff distance between $X, X' \in \mathbb{K}_{\text{GH}}^n$ is given by

$$d_{\text{GH}}^n((X, d, (x_1, \dots, x_n)), (X', d', (x'_1, \dots, x'_n))) := \inf_{S, \phi, \phi'} (d_H(\phi(X), \phi'(X'))),$$

where the infimum is taken over all metric spaces S and isometric embeddings $\phi : X \rightarrow S$, $\phi' : X' \rightarrow S$ such that for every $1 \leq i \leq n$, $\phi(x_i) = \phi'(x'_i)$. d_{GH}^n is indeed a distance on \mathbb{K}_{GH}^n and $(\mathbb{K}_{\text{GH}}^n, d_{\text{GH}}^n)$ is a Polish space (see [5] Section 2.1).

4.2.4 Extension to pseudo metric spaces

Note that the previous topologies naturally extends to pseudo metric spaces. Indeed, one may say that a pseudo metric space (X, d) is isometry-equivalent to the metric space given by quotienting X by the equivalent relation $d(a, b) = 0$ (see Burago, Burago, Ivanov [39] for details.) It is then enough to extend the equivalent classes to pseudo metric spaces.

4.3 Constructions of \mathcal{D} -trees, \mathcal{P} -trees and Θ -ICRT

4.3.1 \mathcal{D} -trees

Recall that a sequence (d_1, \dots, d_s) is a degree sequence of a tree if and only if $\sum_{i=1}^s d_i = s - 2$, and by convention $d_1 \geq d_2 \geq \dots \geq d_s$. Let $\Omega_{\mathcal{D}}$ be the set of such sequences.

For convenience issue, we want to label our leaves on a set $\{\star_i\}_{i \in \mathbb{N}}$ disjoint from $\{V_i\}_{i \in \mathbb{N}}$. So let us slightly change our definition of \mathcal{D} -trees. Note that a tree with degree sequence \mathcal{D} must have $N^{\mathcal{D}} + 2 := \sum_{i=1}^s \mathbf{1}_{d_i=0}$ leaves. We say that a tree T is a \mathcal{D} -tree if it is uniform among all tree with vertices $\{V_i\}_{i: d_i > 0} \cup \{\star_i\}_{0 \leq i \leq N+1}$ and such that for every i with $d_i > 0$, $\deg(V_i) = d_i + 1$.

We now recall the construction of \mathcal{D} -trees from Chapter 3. For simplicity, for every graph $G = (V, E)$ and edge $e = \{v_1, v_2\}$, $G \cup e$ denotes the graph $(V \cup \{v_1, v_2\}, E \cup \{e\})$.

Algorithm 4.1 (Algorithm 3.7). *Stick-breaking construction of a \mathcal{D} -tree $T^{\mathcal{D}}$ (see Figure 4.1).*

- Let $A^{\mathcal{D}} = (A_i^{\mathcal{D}})_{1 \leq i \leq s-1}$ be a uniform \mathcal{D} -tuple (tuple such that $\forall i \in \mathbb{N}$, V_i appears d_i times).
- Let $T_1^{\mathcal{D}} := (\{\star_0, A_1\}, \{\{\star_0, A_1\}\})$ then for every $2 \leq i \leq s$ let

$$T_i^{\mathcal{D}} := \begin{cases} T_{i-1} \cup \{A_{i-1}, A_i\} & \text{if } A_i \notin T_{i-1}. \\ T_{i-1} \cup \{A_{i-1}, \star_{\inf\{k, \star_k \notin T_{i-1}\}}\} & \text{if } A_i \in T_{i-1} \text{ or } i = s. \end{cases}$$

- Let $T^{\mathcal{D}} = T_s$.

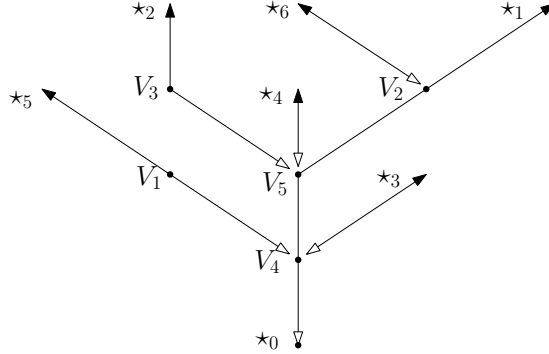


Figure 4.1: Stick breaking construction of a \mathcal{D} -tree with $\mathcal{D} = (1, 2, 1, 3, 3, 0, 0, \dots)$ and $(A_i^{\mathcal{D}})_{1 \leq i \leq s-1} = (V_4, V_5, V_2, V_5, V_3, V_4, V_5, V_4, V_1, V_2)$. The exploration starts at \star_0 then follows the white-black arrow toward \star_1 , then jumps at \star_5 to follow the path toward \star_2 and so on...

4.3.2 \mathcal{P} -trees

Let $\{V_{\infty, i}\}_{i \in \mathbb{N}}$ be a set of vertices disjoint with $\{V_i\}_{i \in \mathbb{N}}$ and $\{\star_i\}_{i \geq 0}$. Let $\Omega_{\mathcal{P}}$ be the set of sequence $(p_i)_{i \in \mathbb{N} \cup \{\infty\}}$ in \mathbb{R}^+ such that $\sum_{i=1}^{\infty} p_i + p_{\infty} = 1$, $p_1 > 0$ and $p_1 \geq p_2 \geq \dots$. For every $\mathcal{P} \in \Omega_{\mathcal{P}}$, the \mathcal{P} -tree is the random tree constructed as follows:

Algorithm 4.2. Definition of the \mathcal{P} -tree for $\mathcal{P} \in \Omega_{\mathcal{P}}$.

- Let $(A_i^{\mathcal{P}})_{i \in \mathbb{N}}$ be a family of i.i.d. random variables such that for all $i \in \mathbb{N}$, $\mathbb{P}(A_1^{\mathcal{P}} = V_i) = p_i$.
- For every $i \in \mathbb{N}$, let $B_i^{\mathcal{P}} = A_i$ if $A_i \in \mathbb{N}$, and let $B_i^{\mathcal{P}} = V_{\infty, i}$ otherwise.
- Let $T_1^{\mathcal{P}} := (\{\star_0, B_1\}, \{\{\star_0, B_1\}\})$ then for every $i \geq 2$ let

$$T_i^{\mathcal{P}} := \begin{cases} T_{i-1} \cup \{B_{i-1}, B_i\} & \text{if } B_i \notin T_{i-1}. \\ T_{i-1} \cup \{B_{i-1}, \star_{\inf\{k, \star_k \notin T_{i-1}\}}\} & \text{if } B_i \in T_{i-1}. \end{cases}$$

- Let $T^{\mathcal{P}} := \bigcup_{n \in \mathbb{N}} T_n$.

Remark. Usually, the leaves $\{\star_i\}_{i \in \mathbb{N}}$ are omitted in the formal definition of \mathcal{P} -trees. We consider them to clarify the intuition that they are degenerate \mathcal{D} -trees with an infinite number of leaves.

4.3.3 ICRT

First let us introduce a generic stick breaking construction. It takes for input two sequences in \mathbb{R}^+ called cuts $\mathbf{y} = (y_i)_{i \in \mathbb{N}}$ and glue points $\mathbf{z} = (z_i)_{i \in \mathbb{N}}$, which satisfy

$$\forall i < j, \quad y_i < y_j \quad ; \quad y_i \longrightarrow \infty \quad ; \quad \forall i \in \mathbb{N}, \quad z_i \leq y_i, \quad (4.1)$$

and creates a \mathbb{R} -tree (loopless geodesic metric space) by recursively "gluing" segment $(y_i, y_{i+1}]$ on position z_i , (see Figure 4.2) or rigorously, by constructing a consistent sequence of distances $(d_n)_{n \in \mathbb{N}}$ on $([0, y_n])_{n \in \mathbb{N}}$.

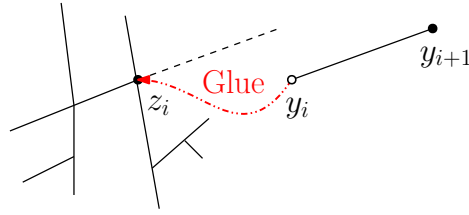


Figure 4.2: A typical step of the stick-breaking construction: the "gluing" of $(y_i, y_{i+1}]$ at z_i .

Algorithm 4.3. Generic stick-breaking construction of \mathbb{R} -tree.

- Let d_0 be the trivial metric on $[0, 0]$.
- For each $i \geq 0$ define the metric d_{i+1} on $[0, y_{i+1}]$ such that for each $x \leq y$:

$$d_{i+1}(x, y) := \begin{cases} d_i(x, y) & \text{if } x, y \in [0, y_i] \\ d_i(x, z_i) + |y - y_i| & \text{if } x \in [0, y_i], y \in (y_i, y_{i+1}] \\ |x - y| & \text{if } x, y \in (y_i, y_{i+1}], \end{cases}$$

where by convention $y_0 := 0$ and $z_0 := 0$.

- Let d be the unique metric on \mathbb{R}^+ which agrees with d_i on $[0, y_i]$ for each $i \in \mathbb{N}$.
- Let $\text{SB}(\mathbf{y}, \mathbf{z})$ be the completion of (\mathbb{R}^+, d) .

Now, let Ω_Θ be the space of sequences $(\theta_i)_{i \in \{0\} \cup \mathbb{N}}$ in \mathbb{R}^+ such that $\sum_{i=0}^{\infty} \theta_i^2 = 1$ and such that $\theta_1 \geq \theta_2 \geq \dots$. For every $\Theta \in \Omega_\Theta$, the Θ -ICRT is the random \mathbb{R} -tree constructed as follows:

Algorithm 4.4. Construction of Θ -ICRT (Algorithm 2.1)

- Let $(X_i)_{i \in \mathbb{N}}$ be a family of independent exponential random variables of parameter $(\theta_i)_{i \in \mathbb{N}}$.
- Let μ be the measure on \mathbb{R}^+ defined by $\mu = \theta_0^2 dx + \sum_{i=1}^{\infty} \delta_{X_i} \theta_i$.
- Let $(Y_i, Z_i)_{i \in \mathbb{N}}$ be a Poisson point process on $\{(y, z) \in \mathbb{R}^{+2} : y \geq z\}$ of intensity $dy \times d\mu$.
- Let $\mathbf{Y} := (Y_i)_{i \in \mathbb{N}}$ and let $\mathbf{Z} := (Z_i)_{i \in \mathbb{N}}$. Let $(Y_0, Z_0) := (0, 0)$.
- The Θ -ICRT is defined as $(T, d) = \text{SB}(\mathbf{Y}, \mathbf{Z})$. (see Algorithm 4.3)

4.4 Constructions of (\mathcal{D}, k) -graphs, (\mathcal{P}, k) -graphs and (Θ, k) -ICRG

4.4.1 Generic gluing and cycle-breaking of discrete multigraphs (see Figure 4.3)

In the entire section, $G = (V, E)$ denotes a multigraph. Let $\text{cyc}(G)$ the set of all edges $e \in E$ such that $G \setminus \{e\} := (V, E \setminus \{e\})$ is connected. (For multiples edges the operation \setminus only remove one edge at a time.) Let $\square(G) := \#\text{cyc}(G)$.

For every leaves $L_1 \neq L_2 \in G$, we define the operation of gluing L_1 and L_2 in G as follows: For every leaf $L \in G$, let the father of L be the only vertex $F \in G$ such that $(F, L) \in G$. Let F_1, F_2 be the father of L_1, L_2 . The multigraph obtained by gluing L_1 and L_2 in G is

$$\mathcal{G}_{L_1, L_2}(G) := (V / \{L_1, L_2\}, E \setminus \{\{F_1, L_1\}, \{F_2, L_2\}\} \cup \{\{F_1, F_2\}\}),$$

and intuitively corresponds to the graph obtained by fusing $\{F_1, L_1\}$ and $\{F_2, L_2\}$.

Similarly, for every leaves $L_1 \neq L_2 \neq \dots, L_{2k-1} \neq L_{2k}$, the multigraph obtained by gluing L_1 and L_2, \dots, L_{2k-1} and L_{2k} in G is

$$\mathcal{G}_{(L_i)_{1 \leq i \leq 2k}}(G) = \mathcal{G}_{(L_1, L_2), (L_3, L_4), \dots, (L_{2k-1}, L_{2k})}(G) := \mathcal{G}_{L_1, L_2} \circ \mathcal{G}_{L_3, L_4} \circ \dots \circ \mathcal{G}_{L_{2k-1}, L_{2k}}(G).$$

Note that this multigraph does not depend on the order in which we glue the different leaves.

Now recall Section 4.1.2. Let us give a formal definition of the cycle-breaking algorithm:

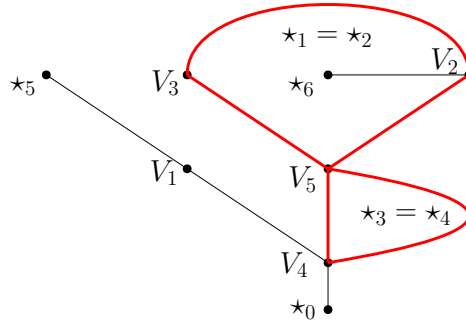


Figure 4.3: Gluing leaves of the tree T from Figure 4.1 to form a graph G with surplus 2. $\text{cyc}(G)$ is red. $\square(G) = 5$. $G = \mathcal{G}_{(\star_1, \star_2, \star_3, \star_4)}(T)$. Also, $\mathbb{P}(\text{CB}(G) = T) = \frac{2!}{2^2 \square(G \setminus \{V_4, V_5\}) \square(G)} = \frac{2}{2^2 \star_3 \star_5}$.

Algorithm 4.5. Cycle-breaking of a multigraph $G = (V, E)$ with $V \subset (V_i)_{i \in \mathbb{N}}$ and surplus k .

- For $1 \leq i \leq k$, let $e_i = (W_{2i+1}, W_{2i+2})$ be a uniform oriented edge in $\text{cyc}(G \setminus \{e_j\}_{1 \leq j < i})$.
- Let $\text{CB}(G) := (V \cup \{\star_i\}_{1 \leq i \leq k}, (E \setminus \{e_i\}_{1 \leq i \leq k}) \cup \{\{W_i, \star_{2k+1-i}\}\}_{1 \leq i \leq 2k})$.

To simplify our notations for every multigraph $G = (V, E)$ and $v, w \in V$, we write $\#_{v,w}(G)$ for the number of edges $\{v, w\}$ in G . Also, let $\circ(G) := \prod_{v \in V} 2^{\#_{v,v}(G)} \prod_{v,w \in V} \#_{v,w}(G)!$.

Lemma 4.4.1. For every connected multigraph G with $V \subset \{V_i\}_{i \in \mathbb{N}}$ and surplus k , we have:

- (a) $\text{CB}(G)$ is almost surely a tree with vertices $V \cup \{\star_i\}_{1 \leq i \leq 2k}$.
- (b) For every $v \in V$, $\deg_{\text{CB}(G)}(v) = \deg_G(v)$. For every $1 \leq i \leq 2k$, \star_i is a leaf in $\text{CB}(G)$.
- (c) Almost surely, $\mathcal{G}_{(\star_1, \star_2), \dots, (\star_{2k-1}, \star_{2k})}(\text{CB}(G)) = G$.
- (d) For every tree T satisfying (a) (b),

$$\mathbb{P}(\text{CB}(G) = T) = \frac{\circ(G)}{2^k \prod_{i=1}^k \square(G \setminus \{e_j\}_{1 \leq j < i})}. \quad (4.2)$$

Proof. (a) and (b) follows from a quick enumeration. (c) is easy to prove from the definition of \mathcal{G} . (d) follows from an induction. Indeed, the right hand side of (4.2) is just the product over each steps of the probability that (W_{2i+1}, W_{2i+2}) satisfies $\{W_{2i+1}, \star_{2k-2i}\}, \{W_{2i+2}, \star_{2k-2i-1}\} \in T$. \square

4.4.2 (\mathcal{D}, k) -graph

Note that (d_1, \dots, d_s) is a degree sequence of a connected multigraph with surplus k if and only if $\sum_{i=1}^s d_i = s + k - 2$, and by convention $d_1 \geq d_2 \geq \dots \geq d_s$. Note that by adding $2k$ numbers 0, this holds if and only if $(d_1, \dots, d_s, 0, \dots, 0) \in \Omega_{\mathcal{D}}$.

For convenience issue, let us slightly extend our definition of (\mathcal{D}, k) -graph. For $\mathcal{D} \in \Omega_{\mathcal{D}}$ with $N^{\mathcal{D}} \geq 2k$ we say that G is a (\mathcal{D}, k) -graph if it is uniform among all multigraph with vertices $\{V_i\}_{i: d_i > 0} \cup \{\star_i\}_{i \in \{0\} \cup \{2k+1, \dots, N^{\mathcal{D}}+1\}}$ and such that for every i with $d_i > 0$, $\deg(V_i) = d_i + 1$. The following result follows from Lemma 4.4.1 and constructs a (\mathcal{D}, k) -graph from a biased \mathcal{D} -tree.

Lemma 4.4.2. Let $T^{\mathcal{D}, k}$ be a random tree. Assume that for every tree T such that: T have vertices $\{V_i\}_{1 \leq i \leq s} \cup \{\star_i\}_{1 \leq i \leq 2k}$, for every $1 \leq i \leq s$ $\deg_T(V_i) = d_i + 1$, and $\{\star_i\}_{1 \leq i \leq 2k}$ are leaves of T ,

$$\mathbb{P}(T^{\mathcal{D}, k} = T) \propto \frac{\circ(\mathcal{G}_{(\star_i)_{1 \leq i \leq 2k}}(T))}{\prod_{i=1}^k \square(\mathcal{G}_{(\star_1, \star_2), \dots, (\star_{2i-1}, \star_{2i})}(T))}, \quad (4.3)$$

where \propto stands for proportional. Then $\mathcal{G}_{(\star_i)_{1 \leq i \leq 2k}}(T^{\mathcal{D}, k})$ is a (\mathcal{D}, k) -graph.

To simplify our notations, we write for every $i \in \mathbb{N}$, $\square_i(\cdot) := \square(\mathcal{G}_{(\star_1, \star_2), \dots, (\star_{2i-1}, \star_{2i})}(\cdot))$ and $\square_{\square, k}(\cdot) := \circ(\mathcal{G}_{(\star_i)_{1 \leq i \leq 2k}}(T)) / \prod_{i=1}^k \square_i(\cdot)$. So that the right hand side of (4.3) is $\square_{\square, k}(T)$.

4.4.3 (\mathcal{P}, k) -graph

Since \mathcal{P} -trees appear at the limit of \mathcal{D} -trees, it is natural to adapt Lemma 4.4.2 to construct limits for (\mathcal{D}, k) -graphs from \mathcal{P} -trees. Thus we informally define the (\mathcal{P}, k) -graph as a \mathcal{P} -tree biased by (4.2) where we glued $\{\star_{2i-1}, \star_{2i}\}_{1 \leq i \leq k}$. Below we formally define (\mathcal{P}, k) -graph.

Fix $\mathcal{P} \in \Omega_{\mathcal{P}}$. First note that Algorithm 4.2 can be seen as a function AB (Aldous–Bröder) which takes a tuple $A^{\mathcal{P}}$ in $\Omega_{\text{AB}} := (\{V_i\}_{i \in \mathbb{N} \cup \{\infty\}})^{\mathbb{N}}$ and send a tree $T^{\mathcal{P}}$. We equip Ω_{AB} with the weak topology and let \mathcal{B}_{AB} be the Borel algebra of this space. Also, we equip Ω_{AB} with the distribution $\mathbb{P}^{\mathcal{P}}$ of $(A_i^{\mathcal{P}})_{i \in \mathbb{N}}$, and complete the space so that event of measure null for $\mathbb{P}^{\mathcal{P}}$ are measurable.

Then note that $\square_{\square, k}(\text{AB})$ is a measurable function from Ω_{AB} to \mathbb{R}^+ since it is locally constant on the subspace of tuple that have at least $2k$ repetitions. Also, note that $\square_{\square, k}(\text{AB}) \leq (k+1)!2^k$. Thus we may define $\mathbb{P}^{\mathcal{P}, k}$ on $(\Omega^{\text{AB}}, \mathcal{B}^{\text{AB}})$ such that for every Borel space $B \in \mathcal{B}_{\text{AB}}$,

$$\mathbb{P}^{\mathcal{P}, k}(B) = \mathbb{E}[\mathbf{1}_{A^{\mathcal{P}} \in B} \square_{\square, k}(\text{AB}(A^{\mathcal{P}}))] / \mathbb{E}[\square_{\square, k}(\text{AB}(A^{\mathcal{P}}))].$$

Now let $A^{\mathcal{P}, k}$ be a random variable with distribution $\mathbb{P}^{\mathcal{P}, k}$. Then let $T^{\mathcal{P}, k} := \text{AB}(A^{\mathcal{P}, k})$. The (\mathcal{P}, k) -graph is the random graph $G^{\mathcal{P}, k} := \mathcal{G}_{(\star_i)_{1 \leq i \leq 2k}}(T^{\mathcal{D}, k}) \setminus \{\star_i\}_{i \in \mathbb{N}}$.

4.4.4 (Θ, k) -ICRG

Since Θ -ICRT appear as the limit of \mathcal{D} -trees it is natural to adapt Lemma 4.4.2 to construct limits for (\mathcal{D}, k) -graphs from Θ -ICRT. Thus we informally define (Θ, k) -ICRG as Θ -ICRT biased by (4.2) where we glued $\{\star_{2i-1}, \star_{2i}\}_{1 \leq i \leq k}$. Below we formally define (Θ, k) -ICRG. We stay rudimentary and refer to Chapter 3 of [39] or to the \mathbb{R} -graph theory of [5] for more details.

First we formally define the gluing of two points: For every pseudo metric space (X, d) and $x_1, x_2 \in X$ let $\mathfrak{G}_{x, y}((X, d))$ be the pseudo metric space (X, d') where for every $a_1, a_2 \in X$,

$$d'(a_1, a_2) := \inf\{d(a_1, a_2); d(a_1, x_1) + d(a_2, x_2); d(a_1, x_2) + d(a_2, x_1)\}.$$

Also for every $k \in \mathbb{N}$ and $x_1, x_2, \dots, x_{2k} \in X$ let

$$\mathfrak{G}_{(x_i)_{1 \leq i \leq 2k}}((X, d)) = \mathfrak{G}_{(x_1, x_2), \dots, (x_{2k-1}, x_{2k})}((X, d)) := \mathfrak{G}_{x_1, x_2} \circ \mathfrak{G}_{x_3, x_4} \circ \dots \circ \mathfrak{G}_{x_{2k-1}, x_{2k}}((X, d)).$$

One can check that $\mathfrak{G}_{(x_i)_{1 \leq i \leq 2k}}((X, d))$ does not depends on the order in $(x_{2i-1}, x_{2i})_{1 \leq i \leq k}$.

Recall Section 4.3.3. Let $\mathbb{K}_{\mathbf{y}\mathbf{z}}$ be the set of couples of sequences \mathbf{y} and \mathbf{z} satisfying (4.1). In Section 4.3.3 we defined the stick breaking construction as a function $\text{SB} : (\mathbf{y}, \mathbf{z}) \in \mathbb{K}_{\mathbf{y}\mathbf{z}} \rightarrow \text{SB}(\mathbf{y}, \mathbf{z})$.

For every $n \in \mathbb{N}$ and $(\mathbf{y}, \mathbf{z}) = ((y_i)_{i \in \mathbb{N}}, (z_i)_{i \in \mathbb{N}}) \in \mathbb{K}_{\mathbf{y}\mathbf{z}}$ let $\text{cyc}_n(\mathbf{y}, \mathbf{z})$ be the set of $x \in \mathbb{R}$ such that $\mathfrak{G}_{(y_i)_{1 \leq i \leq 2n}}(\text{SB}(\mathbf{y}, \mathbf{z})) \setminus \{x\}$ is connected. Note that $\text{cyc}_n(\mathbf{y}, \mathbf{z})$ is a finite union of interval so is measurable. Let $\square_n(\mathbf{y}, \mathbf{z})$ be its Lebesgue measure. Note that $\square_n(\mathbf{y}, \mathbf{z})$ only depends on $\{y_i\}_{1 \leq i \leq n}, \{z_i\}_{1 \leq i \leq n}$, and is a measurable function of $(\{y_i\}_{1 \leq i \leq n}, \{z_i\}_{1 \leq i \leq n})$ (see Lemma 4.10.2). Let $\square_{\square, k}(\mathbf{y}, \mathbf{z}) := 1 / \prod_{n=1}^k \square_n(\mathbf{y}, \mathbf{z})$.

Let \mathbb{M} be the set of all positive locally finite measure on \mathbb{R}^+ . Let $\mathbb{K}_{\text{SB}} := \mathbb{M} \times \mathbb{K}_{\mathbf{y}\mathbf{z}}$. We equip \mathbb{K}_{SB} with the weak topology and let \mathcal{B}_{SB} be the Borel algebra of this space. Let $\Theta \in \Omega_{\Theta}$. We will prove in Lemma 4.6.14 that $\mathbb{E}[\square_{\square, k}(\mathbf{Y}^{\Theta}, \mathbf{Z}^{\Theta})] < \infty$. Thus we may define $\mathbb{P}^{\Theta, k}$ on $(\mathbb{K}_{\text{SB}}, \mathcal{B}_{\text{SB}})$ such that for every Borel space $B \in \mathcal{B}_{\text{SB}}$,

$$\mathbb{P}^{\Theta, k}(B) = \mathbb{E}[\mathbf{1}_{A^{\mathcal{P}} \in B} \square_{\square, k}(\text{SB}(\mathbf{Y}^{\Theta}, \mathbf{Z}^{\Theta}))] / \mathbb{E}[\square_{\square, k}(\mathbf{Y}^{\Theta}, \mathbf{Z}^{\Theta})].$$

Now let $(\mu^{\Theta, k}, \mathbf{Y}^{\Theta, k}, \mathbf{Z}^{\Theta, k})$ be a random variable with distribution $\mathbb{P}^{\Theta, k}$. Let $\mathbf{Y}^{\Theta, k} = (Y_i^{\Theta, k})_{i \in \mathbb{N}}$. Then let $(T^{\Theta, k}, \bar{d}^{\Theta, k}) := \text{SB}(\mathbf{Y}^{\Theta, k}, \mathbf{Z}^{\Theta, k})$. The (Θ, k) -ICRG is the random pseudo metric space $(G^{\Theta, k}, d^{\Theta, k}) := \mathfrak{G}_{(Y_i^{\Theta, k})_{1 \leq i \leq 2k}}(T^{\Theta, k}, \bar{d}^{\Theta, k})$.

4.5 Main results

In this section $(\mathcal{D}_n)_{n \in \mathbb{N}}$, $(\mathcal{P}_n)_{n \in \mathbb{N}}$, $(\Theta_n)_{n \in \mathbb{N}}$ denote fixed sequences in $\Omega_{\mathcal{D}}$, $\Omega_{\mathcal{P}}$, Ω_{Θ} respectively. For every $\mathcal{D} = (d_1, \dots, d_{s^{\mathcal{D}}}) \in \Omega_{\mathcal{D}}$, let $(\sigma^{\mathcal{D}})^2 := \sum_{i=1}^s d_i(d_i - 1)$ then let $\lambda^{\mathcal{D}} := \sigma^{\mathcal{D}}/s^{\mathcal{D}}$. Also, for every $\mathcal{P} = (p_i)_{i \in \mathbb{N} \cup \{\infty\}}$ let $s^{\mathcal{P}} := \max\{i \in \mathbb{N} \cup \{\infty\} : p_i > 0\}$ and let $(\sigma^{\mathcal{P}})^2 = \sum_{i=1}^{\infty} (p_i)^2$. We always work under one of the following regimes:

Assumption 4.1 ($\mathcal{D}_n \Rightarrow \mathcal{P}$). For all $i \geq 1$, $d_i^{\mathcal{D}_n}/s^{\mathcal{D}_n} \rightarrow p_i^{\mathcal{P}}$ and $s^{\mathcal{D}_n} \rightarrow \infty$.

Assumption 4.2 ($\mathcal{D}_n \Rightarrow \Theta$). For all $i \geq 1$, $d_i^{\mathcal{D}_n}/\sigma^{\mathcal{D}_n} \rightarrow \theta_i^{\Theta}$ and $d_1^{\mathcal{D}_n}/s^{\mathcal{D}_n} \rightarrow 0$.

Assumption 4.3 ($\mathcal{P}_n \Rightarrow \Theta$). For all $i \geq 1$, $p_i^{\mathcal{P}_n}/\sigma^{\mathcal{P}_n} \rightarrow \theta_i^{\Theta}$ and $p_1^{\mathcal{P}_n} \rightarrow 0$.

Assumption 4.4 ($\Theta_n \Rightarrow \Theta$). For all $i \geq 1$, $\theta_i^{\Theta_n} \rightarrow \theta_i^{\Theta}$.

A few words on \Rightarrow . One can put a topology on $\Omega := \Omega_{\mathcal{D}} \cup \Omega_{\mathcal{P}} \cup \Omega_{\Theta}$ such that \Rightarrow corresponds with the notion of convergence on Ω . This has several advantages (see Section 3.8.1 for details). First (Ω, \Rightarrow) is a Polish space. Moreover, our results can be seen as continuity results for the function which associate to a set of parameters a metric space. Hence, our results can be used to study graph with random degree distributions. Furthermore $\Omega_{\mathcal{D}}$ is dense on Ω . So our results on (\mathcal{D}, k) -graphs imply the others.

4.5.1 The bias does not diverge

As explained previously in the introduction, our approach relies entirely on the stick breaking construction of Chapter 3 and on the study of the bias corresponding to the cycle-breaking construction. More precisely given the following result, our main results follows from the results of Section 3.4.

Proposition 4.5.1. For every $x, m \in \mathbb{R}^+$ let $h_m := x \mathbf{1}_{x \geq m}$. We have,

$$\lim_{m \rightarrow \infty} \max_{\mathcal{D} \in \Omega_{\mathcal{D}} : N^{\mathcal{D}} \geq 2k} \mathbb{E} \left[h_m \left(\frac{\square_{\square, k}(T^{\mathcal{D}})}{(\lambda^{\mathcal{D}})^k} \right) \right] = 0.$$

4.5.2 Gromov–Prokhorov convergence

First let us specify the measures that we consider. Let $\Omega_{\mathcal{M}}$ the set of measures on $\{V_i\}_{i \geq 1} \cup \{\star_i\}_{i \geq 0}$. We say that a sequence $(\mathbf{p}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{M}}^{\mathbb{N}}$ converges toward $\mathbf{p} \in \Omega_{\mathcal{M}}$ if $\max_{i \in \mathbb{N}} |\mathbf{p}_n(V_i) - \mathbf{p}(V_i)| \rightarrow 0$ and $\max_{i \in \mathbb{N}} |\mathbf{p}_n(\star_i) - \mathbf{p}(\star_i)| \rightarrow 0$. In the whole paper, for every $\mathcal{D} \in \Omega_{\mathcal{D}}$, $\mathbf{p}^{\mathcal{D}, k}$ denote a probability measure with support on $\mathcal{V}^{\mathcal{D}, k} := \{V_i, i : d_i \geq 1\} \cup \{\star_i, i \in \{0\} \cup \{2k+1, \dots, N^{\mathcal{D}}+1\}\}$. Similarly, for every $\mathcal{P} \in \Omega_{\mathcal{P}}$, $\mathbf{p}^{\mathcal{P}}$ denote a probability measure with support on $\mathcal{V}^{\mathcal{P}} := \{V_i^{\mathcal{P}}\}_{i: p_i > 0}$. Also, we sometimes let 0 denote the null measure.

Then we recall the probability measure on ICRT of Chapter 2. To simplify our expressions, we write $\mu^{\Theta} = \infty$ when either $\theta_0^{\Theta} > 0$ or $\sum_{i=1}^{\infty} \theta_i^{\Theta} = \infty$, (since $\mu^{\Theta} = \infty$ iff a.s. $\mu^{\Theta}[0, \infty] = \infty$).

Definition (Proposition 2.3.2). Let $\Theta \in \Omega_{\Theta}$ be such that $\mu^{\Theta} = \infty$. Almost surely, as $n \rightarrow \infty$, $\frac{1}{n} \sum_{i=1}^n \delta_{Y_i^{\Theta}}$ converges weakly toward a probability measure \mathbf{p}^{Θ} on \mathcal{T}^{Θ} .

Remark. When $\mu^{\Theta} < \infty$, $\frac{1}{n} \sum_{i=1}^n \delta_{Y_i^{\Theta}}$ does not converge. For this reason, although we prove the convergence of the distance matrices, one cannot define a proper measure for the GP convergence.

Then let us define a probability measure on $G^{\Theta, k}$. It directly follows from Proposition 2.3.2, that a.s. $\sum_{i=1}^n \delta_{Y_i^{\Theta, k}}$ converges weakly toward a probability measure $\mathbf{p}^{\Theta, k}$ on $T^{\Theta, k}$. Since convergence in $T^{\Theta, k}$ imply convergence in $G^{\Theta, k}$, it still makes sense to define $\mathbf{p}^{\Theta, k}$ on $G^{\Theta, k}$.

We now state the main result of this section. In what follows, $d^{\mathcal{D}, k}$ is the graph distance on $G^{\mathcal{D}, k}$ and similarly $d^{\mathcal{P}, k}$ is the graph distance on $G^{\mathcal{P}, k}$.

Theorem 4.5.2. *The following convergences hold weakly for the GP topology*

(a) *If $\mathcal{D}_n \Rightarrow \mathcal{P}$ and $\mathfrak{p}^{\mathcal{D}_n, k} \rightarrow \mathfrak{p}^{\mathcal{P}}$ then*

$$\left(G^{\mathcal{D}_n, k}, d^{\mathcal{D}_n, k}, \mathfrak{p}^{\mathcal{D}_n, k}\right) \xrightarrow{WGP} (G^{\mathcal{P}, k}, d^{\mathcal{P}, k}, \mathfrak{p}^{\mathcal{P}}).$$

(b) *If $\mathcal{D}_n \Rightarrow \Theta$, $\mathfrak{p}^{\mathcal{D}_n, k} \rightarrow 0$, and $\mu^\Theta = \infty$ then*

$$\left(G^{\mathcal{D}_n, k}, \lambda^{\mathcal{D}_n} d^{\mathcal{D}_n, k}, \mathfrak{p}^{\mathcal{D}_n, k}\right) \xrightarrow{WGP} (G^{\Theta, k}, d^{\Theta, k}, \mathfrak{p}^{\Theta, k}).$$

(c) *If $\mathcal{P}_n \Rightarrow \Theta$, $\mathfrak{p}^{\mathcal{P}_n} \rightarrow 0$, and $\mu^\Theta = \infty$ then*

$$\left(G^{\mathcal{P}_n, k}, \sigma^{\mathcal{P}_n} d^{\mathcal{P}_n, k}, \mathfrak{p}^{\mathcal{P}_n}\right) \xrightarrow{WGP} (G^{\Theta, k}, d^{\Theta, k}, \mathfrak{p}^{\Theta, k}).$$

(d) *If $\Theta_n \Rightarrow \Theta$, $\mu^{\Theta_n} = \infty$ for every $n \in \mathbb{N}$, and $\mu^\Theta = \infty$ then*

$$(G^{\Theta_n, k}, d^{\Theta_n, k}, \mathfrak{p}^{\Theta_n, k}) \xrightarrow{WGP} (G^{\Theta, k}, d^{\Theta, k}, \mathfrak{p}^{\Theta, k}).$$

4.5.3 Gromov–Hausdorff convergence

GH convergence requires additional assumptions. In Chapter 3 we gave quantitative assumptions. Here, we simply state rudimentary assumptions. We proved in Section 3.7.3 that the assumptions of Section 3.4.3 imply the followings. To simplify the notations, for every tree (and every \mathbb{R} -tree) T and $v_1, \dots, v_a \in T$, we write $T(\{v_i\}_{1 \leq i \leq a})$ for the subtree spanned by v_1, \dots, v_a .

Assumption 4.5. For every $\varepsilon > 0$,

$$\lim_{a \rightarrow \infty} \limsup_{n \rightarrow +\infty} \mathbb{P}(\lambda^{\mathcal{D}_n} d_H(T^{\mathcal{D}_n}(\{\star_i\}_{0 \leq i \leq a}), T^{\mathcal{D}_n}) > \varepsilon) = 0,$$

Assumption 4.6. For every $\varepsilon > 0$,

$$\lim_{a \rightarrow \infty} \limsup_{n \rightarrow +\infty} \mathbb{P}(\sigma^{\mathcal{P}_n} d_H(T^{\mathcal{P}_n}(\{\star_i\}_{0 \leq i \leq a}), T^{\mathcal{P}_n}) > \varepsilon) = 0,$$

Assumption 4.7. For every $\varepsilon > 0$,

$$\lim_{a \rightarrow \infty} \limsup_{n \rightarrow +\infty} \mathbb{P}(d_H(T^{\Theta_n}(\{Y_i^{\Theta_n}\}_{0 \leq i \leq a}), T^{\Theta_n}) > \varepsilon) = 0.$$

Theorem 4.5.3. *The following convergences hold weakly for the GH-topology.*

(a) *If $\mathcal{D}_n \Rightarrow \Theta$, $\mathfrak{p}^{\mathcal{D}_n} \rightarrow 0$, and Assumption 4.5 is satisfied then*

$$\left(G^{\mathcal{D}_n, k}, \lambda^{\mathcal{D}_n} d^{\mathcal{D}_n, k}\right) \xrightarrow{WGH} (G^{\Theta, k}, d^{\Theta, k}).$$

(b) *If $\mathcal{P}_n \Rightarrow \Theta$, $\mathfrak{p}^{\mathcal{P}_n} \rightarrow 0$, and Assumption 4.6 is satisfied then*

$$\left(G^{\mathcal{P}_n, k}, \sigma^{\mathcal{P}_n} d^{\mathcal{P}_n, k}\right) \xrightarrow{WGH} (G^{\Theta, k}, d^{\Theta, k}).$$

(c) *If $\Theta_n \Rightarrow \Theta$, and Assumption 4.7 is satisfied then*

$$(G^{\Theta_n, k}, d^{\Theta_n, k}) \xrightarrow{WGH} (G^{\Theta, k}, d^{\Theta, k}).$$

Remark. • Unlike the assumptions of Chapter 3, Assumptions 4.5, 4.6, 4.7 are necessary.

• By Lemma 3.3.3, the GHP convergence follows from the GP and GH convergence and the fact that, since p^Θ have a.s. full support on \mathcal{T}^Θ (Theorem 2.3.1), $p^{\Theta, k}$ have a.s. full support on $G^{\Theta, k}$.

4.6 Study of the bias

4.6.1 Proof of Proposition 4.5.1 in the typical case

Recall that for every $x, m \in \mathbb{R}^+$, $h_m = x \mathbf{1}_{x \geq m}$. Recall the definitions of $(\square_i)_{1 \leq i \leq k}$ and $\square_{\square, k}$ from section 4.4.2. For every $\mathcal{D} \in \Omega_{\mathcal{D}}$ with $N^{\mathcal{D}} \geq 2k$ and $m \in \mathbb{R}^+$ let

$$f^{\mathcal{D}}(m) := \mathbb{E} \left[h_m \left(\frac{\square_{\square, k}(T^{\mathcal{D}})}{(\lambda^{\mathcal{D}})^k} \right) \right].$$

In this section we estimate $f^{\mathcal{D}}$ under the additional assumption $2N^{\mathcal{D}} \geq s^{\mathcal{D}}/\sigma^{\mathcal{D}}$, which is satisfied when there are not too many vertices with degree 2.

Proposition 4.6.1. *There exists $c, C > 0$ such that for every $\mathcal{D} \in \Omega_{\mathcal{D}}$ with $N^{\mathcal{D}} \geq \max(2k, s^{\mathcal{D}}/(2\sigma^{\mathcal{D}}))$, and $m > 0$, we have $f^{\mathcal{D}}(m) \leq Cm^{-c}$.*

Our proof is organized as follows: We first upper bound $\square_{\square, k}$. Then we use Hölder's inequality to upper bound $f^{\mathcal{D}}(\varepsilon)$ with the numbers of leaves in some open balls around \star_0 . Then we use Algorithm 4.1 to upper bound those numbers with $(Y_i)_{1 \leq i \leq k}$. Finally we use the continuum \mathcal{D} -tree construction of Chapter 3 to study $(Y_i)_{1 \leq i \leq k}$ through random Poisson point process.

Let $d^{\mathcal{D}}$ be the graph distance in $T^{\mathcal{D}}$. Let $d'^{\mathcal{D}}(\cdot, \cdot) := \lambda^{\mathcal{D}} d^{\mathcal{D}}(\cdot, \cdot)$. We have:

Lemma 4.6.2. *Let $C = 2^{2k}(k+1)!$. For every $\varepsilon > 0$, for every $\mathcal{D} \in \Omega_{\mathcal{D}}$ with $N^{\mathcal{D}} \geq 2k$,*

$$f^{\mathcal{D}}(C\varepsilon^{-k})/(kC) \leq g^{\mathcal{D}}(\varepsilon) := \mathbb{E} \left[\frac{\mathbf{1}_{d'(\star_1, \star_2) \leq \varepsilon}}{\prod_{i=1}^k d'(\star_{2i-1}, \star_{2i})} \right].$$

Proof. First by definition of $\square_{\square, k}$, $\square_{\square, k} \leq (k+1)!2^k / \prod_{i=1}^k \square_i$. Then note for every $1 \leq i \leq k$ that $\square_i(T^{\mathcal{D}}) \geq d(\star_{2i-1}, \star_{2i}) - 1 \geq d(\star_{2i-1}, \star_{2i})/2$. Indeed, the path between the father of \star_{2i-1} and the father of \star_{2i} , together with the edge connecting those two fathers, forms a cycle. Thus,

$$f^{\mathcal{D}}(C\varepsilon^{-k})/C \leq \mathbb{E} \left[\frac{\mathbf{1}_{\prod_{i=1}^k d'(\star_{2i-1}, \star_{2i}) \leq \varepsilon^k}}{\prod_{i=1}^k d'(\star_{2i-1}, \star_{2i})} \right].$$

The desired result then follows from the symmetry of the leaves $(\star_i)_{1 \leq i \leq 2k}$. (That is the fact that permuting the label of the leaves of $T^{\mathcal{D}}$ independently of $T^{\mathcal{D}}$ does not change the law of $T^{\mathcal{D}}$.) \square

For the rest of the section $\varepsilon > 0$, and \mathcal{D} are fixed. We have to estimate $\prod_{i=1}^k d'(\star_{2i-1}, \star_{2i})$. However, it is hard to estimate since it depends on k separate parts of the tree. For this reason, we instead upper bound g with the numbers of leaves in some open balls around \star_0 . For every $n \geq 1$, let M_n be the proportion of leaves $L \in T \setminus \{\star_0\}$ such that $2^{-n-1} < d'(\star_0, L) \leq 2^{-n}$ and let M_0 be the proportion such that $d'(\star_0, L) > 1/2$. Let $K_\varepsilon := \inf\{n \in \mathbb{N}, 2^{-n} \geq \varepsilon\}$. We have:

Lemma 4.6.3. *There exists $C > 0$ which depends only on k such that,*

$$g(\varepsilon) \leq C \mathbb{E} \left[\sum_{n=K_\varepsilon}^{\infty} 2^{nk} n^{3k} M_n^k \right]^{1/k} \mathbb{E} \left[\sum_{n=0}^{\infty} 2^{nk} n^{3k} M_n^k \right]^{(k-1)/k}.$$

Proof. In this proof C denotes a real depending only on k which may vary from line to line. First, let $(L_i)_{1 \leq i \leq 2k}$ be uniform random variables in $\{\star_i\}_{0 \leq i \leq N+1}$. Note that by symmetry of the leaves,

$$g(\varepsilon) = \mathbb{E} \left[\frac{\mathbf{1}_{d'(L_1, L_2) \leq \varepsilon}}{\prod_{i=1}^k d'(L_{2i-1}, L_{2i})} \middle| \forall i \neq j, L_i \neq L_j \right].$$

Then by roughly speaking slightly changing $(L_i)_{1 \leq i \leq 2k}$ such that some equalities may hold,

$$\begin{aligned} g(\varepsilon) &\leq C \mathbb{E} \left[\frac{\mathbf{1}_{d'(L_1, L_2) \leq \varepsilon}}{\prod_{i=1}^k d'(L_{2i-1}, L_{2i})} \middle| \forall 1 \leq i \leq k, L_{2i-1} \neq L_{2i} \right] \\ &= C \mathbb{E} \left[\mathbb{E} \left[\frac{\mathbf{1}_{d'(L_1, L_2) \leq \varepsilon}}{d'(L_1, L_2)} \middle| L_1, L_1 \neq L_2, T \right] \prod_{i=2}^k \mathbb{E} \left[\frac{1}{d'(L_{2i-1}, L_{2i})} \middle| L_{2i-1}, L_{2i-1} \neq L_{2i}, T \right] \right]. \end{aligned}$$

Furthermore, by Hölder's inequality, and by symmetry of the leaves,

$$\begin{aligned} g(\varepsilon)^k &\leq C \mathbb{E} \left[\mathbb{E} \left[\frac{\mathbf{1}_{d'(L_1, L_2) \leq \varepsilon}}{d'(L_1, L_2)} \middle| L_1, L_1 \neq L_2, T \right]^k \right] \prod_{i=2}^k \mathbb{E} \left[\mathbb{E} \left[\frac{1}{d'(L_{2i-1}, L_{2i})} \middle| L_{2i-1}, L_{2i-1} \neq L_{2i}, T \right]^k \right] \\ &= C \mathbb{E} \left[\mathbb{E} \left[\frac{\mathbf{1}_{d'(\star_0, L_2) \leq \varepsilon}}{d'(\star_0, L_2)} \middle| \star_0 \neq L_2, T \right]^k \right] \mathbb{E} \left[\mathbb{E} \left[\frac{1}{d'(\star_0, L_2)} \middle| \star_0 \neq L_2, T \right]^k \right]^{k-1}. \end{aligned}$$

Therefore, we have by definition of $(M_n)_{n \in \{0\} \cup \mathbb{N}}$,

$$g^{\mathcal{D}}(\varepsilon)^k \leq C \mathbb{E} \left[\left(\sum_{i=K_\varepsilon}^{\infty} 2^n M_n \right)^k \right] \mathbb{E} \left[\left(\sum_{i=0}^{\infty} 2^n M_n \right)^k \right]^{k-1}. \quad (4.4)$$

If $k = 1$ the desired results follow from (4.4). If $k \geq 2$ then we have a.s., by Hölder's inequality,

$$\sum_{i=K_\varepsilon}^{\infty} 2^n M_n \leq \left(\sum_{i=K_\varepsilon}^{\infty} (2^n n^3 M_n)^k \right)^{1/k} \left(\sum_{i=N}^{\infty} \left(\frac{1}{n^3} \right)^{k/(k-1)} \right)^{(k-1)/k},$$

and similarly for $\sum_{i=0}^{\infty} 2^n M_n$. And the desired result follows from (4.4). \square

Recall Section 4.3.1. We now upper bound for $n \in \mathbb{N}$, $\mathbb{E}[M_n^k]$ using Algorithm 4.1. Recall the definition of $A^{\mathcal{D}}$. Let Y_1, Y_2, \dots be the indexes such that $A_i^{\mathcal{D}} \in \{A_1^{\mathcal{D}}, \dots, A_{i-1}^{\mathcal{D}}\}$.

Lemma 4.6.4. *For every $n \in \mathbb{N}$,*

$$\mathbb{E}[M_n^k] \leq k^k \sum_{a=1}^k \frac{1}{N^{k-a}} \mathbb{P} \left(Y_a \leq \frac{a}{2^n} \frac{s}{\sigma} \right).$$

Proof. First, let $(L_i)_{1 \leq i \leq 2k}$ be uniform random variables in $\{\star_i\}_{1 \leq i \leq N+1}$. By definition of M_n^k ,

$$\mathbb{E}[M_n^k] = \mathbb{P} \left[\frac{1}{2^{n+1}} < d'(\star_0, L_1), \dots, d'(\star_0, L_k) \leq \frac{1}{2^n} \right].$$

Then we want distinct leaves to use Algorithm 4.1. To this end, we develop the right hand side above by distinguishing the cases of equality. Let $\mathfrak{P}(k)$ be the set of partition of $\{1, \dots, k\}$. For every $I = \{I_1, \dots, I_a\} \in \mathfrak{P}(k)$, let \mathfrak{E}_I be the event that for every $x, y \in \{1, \dots, k\}$, $L_x = L_y$ iff they are in the same I_i . For every $I \subset \{1, \dots, k\}$ let $m_I := \min(I)$. We have,

$$\begin{aligned} \mathbb{E}[M_n^k] &= \sum_{I=\{I_1, \dots, I_a\} \in \mathfrak{P}(k)} \mathbb{P} \left[\mathfrak{E}_I, \frac{1}{2^{n+1}} < d'(\star_0, L_1), \dots, d'(\star_0, L_k) \leq \frac{1}{2^n} \right] \\ &= \sum_{\{I_1, \dots, I_a\} \in \mathfrak{P}(k)} \frac{1}{(N+1)^{k-a}} \mathbb{P} \left[L_{m_{I_1}} \neq \dots \neq L_{m_{I_a}}, \frac{1}{2^{n+1}} < d'(\star_0, L_{m_{I_1}}), \dots, d'(\star_0, L_{m_{I_a}}) \leq \frac{1}{2^n} \right]. \end{aligned}$$

Then by symmetry of the leaves,

$$\mathbb{E} \left[M_n^k \right] = \sum_{\{I_1, \dots, I_a\} \in \mathfrak{P}(k)} \frac{1}{(N+1)^{k-a}} \mathbb{P} \left[\frac{1}{2^{n+1}} < d'(\star_0, \star_1), \dots, d'(\star_0, \star_a) \leq \frac{1}{2^n} \right].$$

So since there is at most k^k partitions of $\{1, \dots, k\}$,

$$\mathbb{E} \left[M_n^k \right] \leq k^k \sum_{a=1}^k \frac{1}{N^{k-a}} \mathbb{P} \left[\frac{1}{2^{n+1}} < d'(\star_0, \star_1), \dots, d'(\star_0, \star_a) \leq \frac{1}{2^n} \right]. \quad (4.5)$$

Finally we use Algorithm 4.1. It is direct from the construction that, writing $Y_0 = 0$,

$$Y_a = \sum_{i=1}^a (Y_i - Y_{i-1}) \leq \sum_{i=1}^a (d(\star_0, \star_i) - 1) \leq (s/\sigma) \sum_{i=1}^a d'(\star_0, \star_i).$$

So the desired results follows from (4.5). \square

We now upper bound Y_a using a part of the continuum \mathcal{D} -tree construction of Algorithm 3.5:

- Let $(X_i)_{1 \leq i \leq s}$ be a family of independent exponential random variables of parameter $(d_i/\sigma)_{1 \leq i \leq s}$.
- Let μ be the measure on \mathbb{R}^+ defined by $\mu = \sum_{i=1}^s \delta_{X_i} (d_i - 1) / \sigma$.
- Let $(\hat{Y}_i)_{i \in \mathbb{N}}$ be a Poisson point process on \mathbb{R}^+ of rate $\mu[0, y] dy$.
- Let $(E_i)_{1 \leq i \leq s-1}$ be a family of exponential random variables of mean $(\sigma/(s-i))_{1 \leq i \leq s-1}$.

By Lemma 3.5.2 there exists a coupling such that Y_a is independent of $(E_i)_{1 \leq i \leq s-1}$ and such that a.s. $\sum_{i=1}^{Y_a} E_i \leq \hat{Y}_a$. Moreover, we have:

Lemma 4.6.5. *For every $a, n \in \mathbb{N}$ with $n \leq s/2$,*

$$\mathbb{P}(Y_a \leq n) \leq \mathbb{P}(\hat{Y}_a \leq 4n\sigma/s)/2.$$

Proof. Fix $n \leq s/2$. It is easy to check from basic estimates on the Gamma distribution that,

$$\mathbb{P} \left(\sum_{i=1}^n E_i \leq 4n(\sigma/s) \right) \geq 1/2.$$

So since Y_a and $(E_i)_{1 \leq i \leq s-1}$ are independent,

$$\mathbb{P} \left(\hat{Y}_a \leq 4n \frac{\sigma}{s} \right) \geq \mathbb{P} \left(\sum_{i=1}^{Y_a} E_i \leq 4n \frac{\sigma}{s} \right) \geq \mathbb{P} \left(Y_a \leq n, \sum_{i=1}^n E_i \leq 4n \frac{\sigma}{s} \right) \geq \frac{1}{2} \mathbb{P}(Y_a \geq n). \quad \square$$

Hence, to upper bound Y_a it is enough to upper bound \hat{Y}_a . To this end, we first upper bound μ .

Lemma 4.6.6. *For every $a \in \mathbb{N}$,*

- (a) *For every $x, t > 0$, $\mathbb{P}(\mu[0, x] > t) \leq x/t$.*
- (b) *For every $0 \leq x \leq 1 \leq t$, $\mathbb{P}(\mu[0, x] > t) \leq e^{-t/4}$.*

Proof. Note that by definition of μ , $(X_i)_{1 \leq i \leq s}$ and σ ,

$$\mathbb{E}[\mu[0, x]] = \sum_{i=1}^s \frac{d_i - 1}{\sigma} \mathbb{P}(X_i > x) \leq \sum_{i=1}^s \frac{d_i - 1}{\sigma} \frac{x d_i}{\sigma} \leq x.$$

So (a) follows from Markov's inequality. Also $\mu[0, x]$ is a sum of independent random variables bounded by 1 so (b) follows from Bernstein's inequality (see [32] Section 2.8). \square

Lemma 4.6.7. *For every $a \in \mathbb{N}$ and $0 \leq x \leq e^{-9}$, $\mathbb{P}(\hat{Y}_a \leq x) \leq 3x^{a+1}(-4a \log x)^a$.*

Proof. By definition of $(\hat{Y}_i)_{i \in \mathbb{N}}$, conditionally on μ , $\max\{i \in \mathbb{N}, Y_i \leq x\}$ is a Poisson random variable of mean $\int_0^x \mu[0, t] dt \leq x\mu[0, x]$. So, by basic inequalities on the Poisson distributions,

$$\mathbb{P}(\hat{Y}_a \leq x) = \mathbb{E}[\mathbb{P}(\hat{Y}_a \leq x | \mu)] \leq \mathbb{E}[(x\mu[0, x])^a]. \quad (4.6)$$

Then we have by integration by part and Lemma 4.6.6,

$$\begin{aligned} \mathbb{E}[\mu[0, x]^a] &= \int_0^\infty \mathbb{P}(\mu[0, x] \geq t)(at^{a-1} dt) \\ &\leq \int_0^x at^{a-1} dt + \int_x^{-4 \log x} (x/t)(at^{a-1} dt) + \int_{-4 \log x}^\infty e^{-t/4}(at^{a-1} dt) \\ &\leq 3x(-4a \log x)^a, \end{aligned}$$

using basic calculus for the last inequality. This concludes the proof. \square

Proof of Proposition 4.6.1. We now complete our upper bound for $f(\mathcal{D})$. In this proof, c, C denote reals which depend only on k and which may vary from line to line. First by Lemmas 4.6.7 and 4.6.5 we have for every $1 \leq a \leq k$ and $0 \leq x < 1/16$,

$$\mathbb{P}(Y_a \leq xs/\sigma) \leq Cx^{a+1}(-\log(x))^c.$$

Then by Lemma 4.6.4, and $2N \geq s/\sigma$, for every $n \in \mathbb{N}$ with $(s/\sigma)/2^n \geq 1$,

$$\begin{aligned} \mathbb{E}[M_n^k] &\leq k^k \sum_{a=1}^k \frac{1}{N^{k-a}} \mathbb{P}\left(Y_a \leq \frac{a}{2^n} \frac{s}{\sigma}\right) \\ &\leq k^k \sum_{a=1}^k \left(\frac{2\sigma}{s}\right)^{k-a} C \left(\frac{a}{2^n}\right)^{a+1} n^c \\ &\leq \frac{Cn^c}{2^{(k+1)n}}. \end{aligned} \quad (4.7)$$

Note that (4.7) naturally extends to the $n \in \mathbb{N}$ with $(s/\sigma)/2^n < 1$ since for those n almost surely for $1 \leq a \leq k$, we have $Y_a \geq a > \frac{a}{2^n} \frac{s}{\sigma}$.

Next, since $K_\varepsilon = \inf\{n \in \mathbb{N}, 2^{-n} \geq \varepsilon\}$,

$$\mathbb{E} \left[\sum_{i=K_\varepsilon}^\infty 2^{nk} n^{3k} M_n^k \right] \leq C 2^{-K_\varepsilon c} \leq C(2\varepsilon)^c,$$

and

$$\mathbb{E} \left[\sum_{i=0}^\infty 2^{nk} n^{3k} M_n^k \right] \leq C.$$

Thus by Lemma 4.6.3,

$$g^{\mathcal{D}}(\varepsilon) \leq C \mathbb{E} \left[\sum_{i=0}^\infty 2^{nk} n^{3k} M_n^k \right]^{1/k} \mathbb{E} \left[\sum_{i=K_\varepsilon}^\infty 2^{nk} n^{3k} M_n^k \right]^{(k-1)/k} \leq C(2\varepsilon)^c. \quad (4.8)$$

Finally, Proposition 4.6.1 follows from Lemma 4.6.2. \square

Along the way by (4.8) we have the following result, which we extend in the next section.

Lemma 4.6.8. *There exists $c, C > 0$ which depends only on k such that for every $\varepsilon > 0$, $\mathcal{D} \in \Omega_{\mathcal{D}}$ with $N^{\mathcal{D}} \geq \max(2k, s^{\mathcal{D}}/(2\sigma^{\mathcal{D}}))$, $g^{\mathcal{D}}(\varepsilon) \leq C\varepsilon^c$.*

4.6.2 Proof of Proposition 4.5.1 when there are many vertices of degree 2

This section is organized as follows. We first detail how to remove or add vertices of degree 2. We then prove from those constructions a connection between the \mathcal{D} -trees that do not have any vertex of degree 2 and the others. Finally we use this connection to prove Proposition 4.5.1.

First for every graph G and $x \in G$, we call x an edgepoint if x have degree 2. A simple way to remove the edgepoints is to shortcut them: Formally if $T = (V, E)$ is a tree, then ∇T be the tree (V', E') such that $V' = \{v \in V, \deg_T(v) \neq 2\}$ and for every $v, w \in V'$, $\{v, w\} \in E'$ iff there exists a path between v and w that only pass by v, w and vertices of degree 2. Note that ∇ keep the degrees: for every $v \in T$ with $\deg_v(T) \neq 2$, we have $v \in \nabla T$ and $\deg_T(v) = \deg_{\nabla T}(v)$.

Remark. One may extends ∇ to general graph. However, the natural way to preserves the degrees is to work with multigraph. We avoid this issue by working with trees.

Reciprocally one may construct any tree by adding some edgepoints along the oriented edges of a tree without edgepoint: For every $T = (V, E)$ let $(\vec{e}_i(T))_{1 \leq i \leq \#E}$ be some fixed oriented edges of T such that each edge of E appears in one and only one direction. Let $((W_{i,j})_{1 \leq j \leq r_i})_{1 \leq i \leq \#E}$ be some vertices that are not in V . For every $1 \leq i \leq \#E$ let $(W_{i,0}, W_{i,r_i+1}) := \vec{e}_i(G)$. Let

$$\Delta(T, ((W_{i,j})_{1 \leq j \leq r_i})_{1 \leq i \leq \#E}) := (V \cup \{W_{i,j}\}_{1 \leq i \leq \#E, 1 \leq j \leq r_i}, \{\{W_{i,j}, W_{i,j+1}\}\}_{1 \leq i \leq \#E, 0 \leq j \leq r_i}).$$

We now use Δ, ∇ to study \mathcal{D} -trees. Beforehand, let us introduce some notations. For every $\mathcal{D} = (d_1, \dots, d_s) \in \Omega_{\mathcal{D}}$, let $s_{\geq 2}^{\mathcal{D}} := \#\{a \in \mathbb{N}, d_a \geq 2\}$, let $s_{\geq 1}^{\mathcal{D}} = \#\{a \in \mathbb{N}, d_a \geq 1\}$, and let $s_1^{\mathcal{D}} := \#\{a \in \mathbb{N}, d_a = 1\}$. Also let $\nabla \mathcal{D}$ be the sequence $(d_1, d_2, \dots, d_{s_{\geq 2}^{\mathcal{D}}}, d_{s_{\geq 2}^{\mathcal{D}}+1}, \dots, d_s)$.

Also we say that $((W_{i,j})_{1 \leq j \leq r_i})_{1 \leq i \leq n}$ is an ordered partition of size $n \in \mathbb{N}$ of a finite set E iff for $1 \leq i \leq n$, $r_i \in \{0\} \cup \mathbb{N}$, and $(i, j) \mapsto W_{i,j}$ is a bijection from $\{1 \leq i \leq n, 1 \leq j \leq r_i\}$ to E . We have the following connections between \mathcal{D} -trees and $\nabla \mathcal{D}$ -trees:

Lemma 4.6.9. *Let $\mathcal{D} \in \Omega_{\mathcal{D}}$. Let W be a uniform ordered partition of size $s^{\nabla \mathcal{D}} - 1$ of $\{V_i\}_{i: d_i^{\mathcal{D}}=1}$. Then, a) $\nabla(T^{\mathcal{D}})$ is a $\nabla \mathcal{D}$ -tree, and b) $\Delta(T^{\nabla \mathcal{D}}, W)$ is a \mathcal{D} -tree*

Proof. First note that $\nabla(\Delta(T^{\nabla \mathcal{D}}, W)) = T^{\nabla \mathcal{D}}$, since this tree is obtained by adding some edgepoint on $T^{\nabla \mathcal{D}}$, which do not have edgepoint, then by removing all edgepoint. So b) imply a).

Toward b), simply note that Δ may be seen as a bijection from trees with degree sequence $\nabla \mathcal{D}$ and ordered partition of size $s^{\nabla \mathcal{D}} - 1$ of $\{V_i^{\mathcal{D}}\}_{d_i=1}$ toward trees with degree sequence \mathcal{D} . (Indeed, one may recover the initial tree by applying Δ and then read the ordered partition by, roughly speaking, following each oriented edges of the initial tree on the image tree.) \square

We now prove Proposition 4.5.1. To this end, it is enough to remove the assumption $2N^{\mathcal{D}} \geq s^{\mathcal{D}}/\sigma^{\mathcal{D}}$ of Proposition 4.6.1. Note that it is satisfied when $s_1^{\mathcal{D}} = 0$ since in this case, $\sigma^{\mathcal{D}} \geq 1$ and $s^{\mathcal{D}} = N^{\mathcal{D}} + s_{\geq 2}^{\mathcal{D}} \leq 2N^{\mathcal{D}}$. For this reason, our goal for the rest of the section will be to prove the following result, which together with Lemmas 4.6.8 and 4.6.2 yields Proposition 4.5.1.

Proposition 4.6.10. *Recall the definition of g from Lemma 4.6.2. There exists $C > 0$, which depends only on k , such that for every $\mathcal{D} \in \Omega_{\mathcal{D}}$ with $N^{\mathcal{D}} \geq 2k$ and $\varepsilon > 0$,*

$$g^{\mathcal{D}}(\varepsilon) \leq C\varepsilon \left(\int_{\varepsilon}^1 g^{\nabla \mathcal{D}}(\delta)/\delta^2 d\delta + k g^{\nabla \mathcal{D}}(1) + 1 \right).$$

To this end, it is enough to lower bound $(d^{\mathcal{D}}(\star_{2i-1}, \star_{2i}))_{1 \leq i \leq k}$ using $(d^{\nabla \mathcal{D}}(\star_{2i-1}, \star_{2i}))_{1 \leq i \leq k}$. To do so, by Lemma 4.6.9 (b), it suffices to study uniform ordered partitions. More precisely, we have to lower bound the cardinal of the sets of those partitions, which corresponds to the numbers of edgepoint added on each edge. This is done in the following lemma.

Lemma 4.6.11. *Let $((W_{i,j})_{1 \leq j \leq R_i})_{1 \leq i \leq n}$ be a uniform ordered partition of size n of a finite set E .*

- (a) *$(R_i)_{1 \leq i \leq n}$ is uniform among all set of integers such that $\sum_{i=1}^n R_i = \#E$.*
- (b) *Let $(S_i)_{1 \leq i \leq n}$ be independent geometric random variables of mean $\#E/n$ conditioned on $\sum_{i=1}^n S_i \leq \#E$. Then there exists a coupling between $(R_i)_{1 \leq i \leq n}$ and $(S_i)_{1 \leq i \leq n}$ such that almost surely for every $1 \leq i \leq n$, $R_i \geq S_i$.*

Proof. Toward (a), simply note that given $(R_i)_{1 \leq i \leq n}$, there are exactly $\#E!$ possible ways to label $((W_{i,j})_{1 \leq j \leq R_i})_{1 \leq i \leq n}$ to form an ordered partition of size n of E . Then (b) is an easy exercise. \square

Next, in order to use the independency of Lemma 4.6.11 (b), we will use the following lemma:

Lemma 4.6.12. *Let T be a tree. Assume that $(\star_i)_{1 \leq i \leq 2k}$ are leaves of T . For every $1 \leq i \leq k$ let \mathcal{E}_i be the set of edges that are on the minimal path between \star_{2i-1} and \star_{2i} . Then there exists $(\mathcal{E}'_i)_{1 \leq i \leq k}$ disjoint subsets of $(\mathcal{E}_i)_{1 \leq i \leq k}$ such that for every $1 \leq i \leq k$, $\#\mathcal{E}'_i \geq \max(\#\mathcal{E}_i/k, 2)$.*

Proof. Consider the following informal construction of $(\mathcal{E}'_i)_{1 \leq i \leq k}$:

- First let for $1 \leq i \leq k$, $\mathcal{E}'_i := \{\{\star_{2i-1}, F_{2i-1}\}, \{\star_{2i}, F_{2i}\}\}$, where for $1 \leq i \leq 2k$, F_i is the father of \star_i in T .
- Then while $\bigcup_{i=1}^k \mathcal{E}'_i \neq \bigcup_{i=1}^k \mathcal{E}_i$:
 - For $1 \leq i \leq k$: If possible add to \mathcal{E}'_i an arbitrary edge in \mathcal{E}_i that is not yet in $\bigcup_{j=1}^k \mathcal{E}'_j$.

It is easy to check that $(\mathcal{E}'_i)_{1 \leq i \leq k}$ are disjoint subsets of $(\mathcal{E}_i)_{1 \leq i \leq k}$. Also for $1 \leq i \leq k$, $\#\mathcal{E}'_i \geq 2$. Finally a quick enumeration gives that at the end of the algorithm $\#\mathcal{E}'_i \geq \#\mathcal{E}_i/k$. \square

Proof of Proposition 4.6.10. Let $\varepsilon > 0$. Let $\mathcal{D} \in \Omega_{\mathcal{D}}$. Let W be a uniform ordered partition of size $s_{\geq 2}^{\mathcal{D}}$ of $\{V_i^{\mathcal{D}}\}_{i:d_i=1}$ and independent of $T^{\nabla \mathcal{D}}$. Let $d^{\nabla \mathcal{D}, W}$ be the graph distance on $\Delta(T^{\nabla \mathcal{D}}, W)$. Then by Lemma 4.6.9 (b), $\Delta(T^{\nabla \mathcal{D}}, W)$ is a \mathcal{D} -tree. So, by definition of g , it is enough upper bound

$$G^{\mathcal{D}}(\varepsilon, T^{\nabla \mathcal{D}}) := \mathbb{E} \left[\frac{\mathbf{1}_{\lambda^{\mathcal{D}} d^{\nabla \mathcal{D}, W}(\star_1, \star_2) \leq \varepsilon}}{\prod_{i=1}^k (\lambda^{\mathcal{D}} d^{\nabla \mathcal{D}, W}(\star_{2i-1}, \star_{2i}))} \middle| T^{\nabla \mathcal{D}} \right]. \quad (4.9)$$

To this end, let us use Lemmas 4.6.11 and 4.6.12. Let \mathcal{E} be the set of edges of $T^{\nabla \mathcal{D}}$. Let $(S_e)_{e \in \mathcal{E}}$ be independent geometric random variables of mean $s_1^{\mathcal{D}}/\#\mathcal{E}$ conditioned on $\sum_{e \in \mathcal{E}} S_e \leq s_1^{\mathcal{D}}$. For $1 \leq i \leq k$ let \mathcal{E}_i be the set of edges that are on the minimal path between \star_{2i-1} and \star_{2i} in $T^{\nabla \mathcal{D}}$. By definition of Δ , and by Lemma 4.6.11, note that, there exists a coupling between W and $(S_e)_{e \in \mathcal{E}}$ such that a.s. for $1 \leq i \leq 2k$,

$$d^{\nabla \mathcal{D}, W}(\star_{2i-1}, \star_{2i}) \geq \sum_{e \in \mathcal{E}_i} (1 + S_e). \quad (4.10)$$

Then, by Lemma 4.6.12, let $(\mathcal{E}'_i)_{1 \leq i \leq k}$ be disjoint subsets of $(\mathcal{E}_i)_{1 \leq i \leq k}$ such that for every $1 \leq i \leq k$, $\#\mathcal{E}'_i \geq \max(\#\mathcal{E}_i/k, 2)$. It directly follows from (4.10) that a.s. for $1 \leq i \leq 2k$,

$$d^{\nabla \mathcal{D}, W}(\star_{2i-1}, \star_{2i}) \geq \sum_{e \in \mathcal{E}'_i} (1 + S_e).$$

Therefore,

$$G^{\mathcal{D}}(\varepsilon, T^{\nabla \mathcal{D}}) \leq \mathbb{E} \left[\frac{\mathbf{1}_{\lambda^{\mathcal{D}} \sum_{e \in \mathcal{E}'_1} (1 + S_e) \leq \varepsilon}}{\prod_{i=1}^k (\lambda^{\mathcal{D}} \sum_{e \in \mathcal{E}'_i} (1 + S_e))} \middle| T^{\nabla \mathcal{D}} \right].$$

Hence, if $(S'_e)_{e \in \mathcal{E}}$ are independent geometric random variables of mean $s_1^{\mathcal{D}}/\#\mathcal{E}$,

$$G^{\mathcal{D}}(\varepsilon, T^{\nabla \mathcal{D}}) \leq \frac{1}{\mathbb{P}(\sum_{e \in \mathcal{E}} S'_e \leq s_1^{\mathcal{D}} | T^{\nabla \mathcal{D}})} \mathbb{E} \left[\frac{\mathbf{1}_{\lambda^{\mathcal{D}} \sum_{e \in \mathcal{E}'_1} (1 + S'_e) \leq \varepsilon}}{\prod_{i=1}^k (\lambda^{\mathcal{D}} \sum_{e \in \mathcal{E}'_i} (1 + S'_e))} \middle| T^{\nabla \mathcal{D}} \right].$$

Then note that there exists a constant $C < \infty$ that does not depends on k, \mathcal{D} such that a.s. $\mathbb{P}(\sum_{e \in \mathcal{E}} S'_e \leq s_1^{\mathcal{D}} | T^{\Delta \mathcal{D}}) \leq 1/C$. So, since $(\mathcal{E}'_i)_{1 \leq i \leq k}$ are disjoint and $(S'_e)_{e \in \mathcal{E}}$ are independent,

$$G^{\mathcal{D}}(\varepsilon, T^{\nabla \mathcal{D}}) \leq C (\lambda^{\mathcal{D}})^{-k} \mathbb{E} \left[\frac{\mathbf{1}_{\sum_{e \in \mathcal{E}'_1} (1+S'_e) \leq \varepsilon/\lambda^{\mathcal{D}}}}{\sum_{e \in \mathcal{E}'_1} (1+S'_e)} \middle| T^{\nabla \mathcal{D}} \right] \prod_{i=2}^k \mathbb{E} \left[\frac{1}{\sum_{e \in \mathcal{E}'_i} (1+S'_e)} \middle| T^{\nabla \mathcal{D}} \right].$$

Therefore we have using Lemma 4.6.13 below, and the fact that for every $1 \leq i \leq k$, $\#\mathcal{E}'_i \geq 2$,

$$G^{\mathcal{D}}(\varepsilon, T^{\nabla \mathcal{D}}) \leq C(2e)^{-k} (\lambda^{\mathcal{D}})^{-k} \min \left(1, \frac{e\varepsilon/\lambda^{\mathcal{D}}}{\#\mathcal{E}'_1(1+s_1^{\mathcal{D}}/\#\mathcal{E})} \right) \prod_{i=1}^k \frac{1}{\#\mathcal{E}'_i(1+s_1^{\mathcal{D}}/\#\mathcal{E})}. \quad (4.11)$$

Next, let us rewrite (4.11). First, note that for every $1 \leq i \leq k$,

$$\#\mathcal{E}'_i \geq \#\mathcal{E}_i/k = d^{\nabla \mathcal{D}}(\star_{2i-1}, \star_{2i})/k.$$

Also,

$$1 + \frac{s_1^{\mathcal{D}}}{\#\mathcal{E}} = 1 + \frac{s_1^{\mathcal{D}}}{s^{\nabla \mathcal{D}} - 1} = \frac{s^{\nabla \mathcal{D}} + s_1^{\mathcal{D}} - 1}{s^{\nabla \mathcal{D}} - 1} = \frac{s^{\mathcal{D}} - 1}{s^{\nabla \mathcal{D}} - 1} \geq \frac{s^{\mathcal{D}}}{s^{\nabla \mathcal{D}}} = \frac{\lambda^{\mathcal{D}}}{\lambda^{\nabla \mathcal{D}}},$$

noting for the last equality that $\sigma^{\mathcal{D}} = \sigma^{\nabla \mathcal{D}}$. Then by elementary calculus it is easy to prove that,

$$\begin{aligned} \min \left(1, \frac{e\varepsilon/\lambda^{\mathcal{D}}}{\#\mathcal{E}'_1(1+s_1^{\mathcal{D}}/\#\mathcal{E})} \right) &\leq ke \min \left(1, \frac{\varepsilon}{\lambda^{\nabla \mathcal{D}} d^{\nabla \mathcal{D}}(\star_1, \star_2)} \right) \\ &= ke\varepsilon \int_{\varepsilon}^{\infty} \mathbf{1}_{\lambda^{\nabla \mathcal{D}} d^{\nabla \mathcal{D}}(\star_1, \star_2) \leq \delta} \frac{d\delta}{\delta^2}. \end{aligned}$$

Therefore by (4.11),

$$G^{\mathcal{D}}(\varepsilon, T^{\nabla \mathcal{D}}) \leq C(2ek)^{k+1} (\lambda^{\nabla \mathcal{D}})^{-k} \varepsilon \int_{\varepsilon}^{\infty} \mathbf{1}_{\lambda^{\nabla \mathcal{D}} d^{\nabla \mathcal{D}}(\star_1, \star_2) \leq \delta} \frac{d\delta}{\delta^2} \prod_{i=1}^k \frac{1}{d^{\nabla \mathcal{D}}(\star_{2i-1}, \star_{2i})}.$$

Finally by taking the expectation and by Fubini's theorem, we have,

$$\mathbb{E}[G^{\mathcal{D}}(\varepsilon, T^{\nabla \mathcal{D}})] \leq C(2ek)^{k+1} (\lambda^{\nabla \mathcal{D}})^{-k} \varepsilon \int_{\varepsilon}^{\infty} \mathbb{E} \left[\mathbf{1}_{\lambda^{\nabla \mathcal{D}} d^{\nabla \mathcal{D}}(\star_1, \star_2) \leq \delta} \prod_{i=1}^k \frac{1}{d^{\nabla \mathcal{D}}(\star_{2i-1}, \star_{2i})} \right] \frac{d\delta}{\delta^2},$$

which yields by definition of G and g ,

$$g^{\mathcal{D}}(\varepsilon) \leq C(2ek)^{k+1} \varepsilon \int_{\varepsilon}^{\infty} g^{\nabla \mathcal{D}}(\delta) \frac{d\delta}{\delta^2}. \quad (4.12)$$

To conclude the proof, note that for $\delta \geq 1$,

$$\begin{aligned} g^{\nabla \mathcal{D}}(\delta) &\leq \mathbb{E} \left[\prod_{i=1}^k \frac{1}{\lambda^{\nabla \mathcal{D}} d^{\nabla \mathcal{D}}(\star_{2i-1}, \star_{2i})} \right] \\ &\leq \mathbb{E} \left[1 + \sum_{j=1}^k \mathbf{1}_{\lambda^{\nabla \mathcal{D}} d^{\nabla \mathcal{D}}(\star_{2j-1}, \star_{2j}) \leq 1} \prod_{i=1}^k \frac{1}{\lambda^{\nabla \mathcal{D}} d^{\nabla \mathcal{D}}(\star_{2i-1}, \star_{2i})} \right] = 1 + kg^{\nabla \mathcal{D}}(1). \end{aligned}$$

So the desired result follows from (4.12). \square

Lemma 4.6.13. *Let $n \geq 2$, $m \geq 0$. Let $(S_i)_{1 \leq i \leq n}$ be independent geometric random variables of mean m . Then,*

$$\mathbb{E} \left[\frac{1}{\sum_{i=1}^n (1 + S_i)} \right] \leq \frac{2e}{n(1+m)}.$$

Also, for every $\varepsilon > 0$,

$$\mathbb{E} \left[\frac{\mathbf{1}_{\sum_{i=1}^n (1+S_i) \leq \varepsilon}}{\sum_{i=1}^n (1+S_i)} \right] \leq \frac{2e}{n(1+m)} \min \left(1, \frac{e\varepsilon}{(1+m)n} \right).$$

Proof. Note that $\sum_{i=1}^n (1 + S_i)$ is the time needed to get n success for Bernoulli trials that hold with probability $1/(1+m)$. Thus for every $x > 0$,

$$\mathbb{P} \left(\sum_{i=1}^n (1 + S_i) \leq x \right) \leq \binom{\lfloor x \rfloor}{n} \frac{1}{(1+m)^n} \leq \left(\frac{x}{1+m} \right)^n / n! \leq \left(\frac{ex}{(1+m)n} \right)^n.$$

It directly follows by integration by part that,

$$\begin{aligned} \mathbb{E} \left[\frac{1}{\sum_{i=1}^n (1 + S_i)} \right] &= \int_0^\infty \mathbb{P} \left(\sum_{i=1}^n (1 + S_i) \leq x \right) x^{-2} dx \\ &\leq \int_0^{(1+m)n/e} \left(\frac{ex}{(1+m)n} \right)^n x^{-2} dx + \int_{(1+m)n/e}^\infty x^{-2} dx \\ &= \frac{e}{(1+m)n(n-1)} + \frac{e}{(1+m)n} \\ &\leq \frac{2e}{(1+m)n}. \end{aligned}$$

The second inequality is proved in a similar way. \square

4.6.3 Bias of \mathcal{P} -trees and ICRT

Recall the definitions of section 4.4.2 and 4.4.4 of $(\square_i)_{1 \leq i \leq k}$ and $\square_{\square,k}$. Recall that for every $x, m \in \mathbb{R}^+$, $h_m : x \mapsto \mathbf{1}_{x \geq mx}$.

Lemma 4.6.14. *We have the following assertions:*

(a)

$$\lim_{m \rightarrow \infty} \max_{\mathcal{P} \in \Omega_{\mathcal{P}}} \mathbb{E} \left[h_m \left(\square_{\square,k}(T^{\mathcal{P}}) / (\sigma^{\mathcal{P}})^k \right) \right] = 0.$$

b)

$$\lim_{m \rightarrow \infty} \max_{\Theta \in \Omega_{\Theta}} \mathbb{E} \left[h_m \left(\square_{\square,k}(\mathbf{Y}^{\Theta}, \mathbf{Z}^{\Theta}) \right) \right] = 0.$$

Proof. We focus only on (a) since (b) can be proved in the exact same way. Fix $\mathcal{P} \in \Omega_{\mathcal{P}}$. Let $(\mathcal{D}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{D}}^{\mathbb{N}}$ such that $\mathcal{D}_n \Rightarrow \mathcal{P}$ (see the start of Section 4.5 or see Section 3.8.1 for existence). By Theorem 3.4.1, we have the following weak convergence,

$$(d^{\mathcal{D}_n}(\star_i, \star_j))_{1 \leq i, j \leq 2k} \xrightarrow{(d)} (d^{\mathcal{P}}(\star_i, \star_j))_{1 \leq i, j \leq 2k}.$$

Then by Lemma 4.10.2 (see also [5] Corollary 6.6), $\square_{\square,k}(T^{\mathcal{D}_n})$ converges weakly toward $\square_{\square,k}(T^{\mathcal{P}})$ as $n \rightarrow \infty$. Furthermore, by Fubini's Theorem,

$$(\lambda^{\mathcal{D}_n})^2 = (\sigma^{\mathcal{D}_n} / s^{\mathcal{D}_n})^2 = \sum_{i=1}^{\infty} \frac{(d_i^{\mathcal{D}_n})(d_i^{\mathcal{D}_n} - 1)}{(s^{\mathcal{D}_n})^2} \rightarrow \sum_{i=1}^{\infty} p_i^2 = (\sigma^{\mathcal{P}})^2.$$

Therefore, for every $m \geq 0$,

$$\limsup \mathbb{E}[h_m(\square_{\square,k}(T^{\mathcal{D}_n}) / (\lambda^{\mathcal{D}_n})^k)] \geq \mathbb{E}[h_{m+1}(\square_{\square,k}(T^{\mathcal{P}}) / (\sigma^{\mathcal{P}})^k)]. \quad (4.13)$$

Finally, Proposition 4.5.1 concludes the proof. \square

4.7 Proof of the main theorems

Theorems 4.5.2 and 4.5.3 directly follows from three thing: the trees converges, the operation of gluing leaves is a continuous map, and the bias converge. In this section, we precise the proofs.

4.7.1 Proof of Theorem 4.5.2

Proof of Theorem 4.5.2 (a). Let $(\mathcal{D}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{D}}^{\mathbb{N}}$ and $\mathcal{P} = (p_i)_{i \in \mathbb{N} \cup \{\infty\}} \in \Omega_{\mathcal{P}}$ such that $\mathcal{D}_n \Rightarrow \mathcal{P}$. Let $a \in \mathbb{N}$ such that $p_a > 0$. For all $1 \leq i \leq a$ let $W_i = V_i$. For all $1 \leq i \leq 2k$, let $W_{a+i} := \star_i$. By Theorem 3.4.1 (a), it is easy to check that we have the following joint convergence,

$$(d^{\mathcal{D}_n}(W_i, W_j))_{1 \leq i, j \leq a+2k} \xrightarrow{(d)} (d^{\mathcal{P}}(W_i, W_j))_{1 \leq i, j \leq a+2k}, \quad (4.14)$$

writing $d^{\mathcal{D}_n}$ for the graph distance on $T^{\mathcal{D}_n}$, and $d^{\mathcal{P}}$ for the graph distance on $T^{\mathcal{P}}$.

Then by Kolmogorov representation theorem, we may assume that (4.14) holds a.s. Furthermore, since we work with discrete trees, note that a.s. for every n large enough equality holds in (4.14). Hence, by Lemma 4.10.2 a.s. for every n large enough $\square_{\square, k}(T^{\mathcal{D}_n}) = \square_{\square, k}(T^{\mathcal{P}})$. Thus, by dominated convergence, for any continuous bounded function $f : \mathbb{R}^{(a+2k)^2} \rightarrow \mathbb{R}^+$,

$$\frac{\mathbb{E}[f(d^{\mathcal{D}_n}(W_i, W_j))_{1 \leq i, j \leq a+2k} \square_{\square, k}(T^{\mathcal{D}_n})]}{\mathbb{E}[\square_{\square, k}(T^{\mathcal{D}_n})]} \longrightarrow \frac{\mathbb{E}[f(d^{\mathcal{P}}(W_i, W_j))_{1 \leq i, j \leq a+2k} \square_{\square, k}(T^{\mathcal{P}})]}{\mathbb{E}[\square_{\square, k}(T^{\mathcal{P}})]}.$$

Therefore, writing $\bar{d}^{\mathcal{D}_n, k}$ for the graph distance on $T^{\mathcal{D}_n, k}$ and $\bar{d}^{\mathcal{P}, k}$ for the graph distance on $T^{\mathcal{P}, k}$,

$$(\bar{d}^{\mathcal{D}_n, k}(W_i, W_j))_{1 \leq i, j \leq a+2k} \xrightarrow{(d)} (\bar{d}^{\mathcal{P}, k}(W_i, W_j))_{1 \leq i, j \leq a+2k}. \quad (4.15)$$

Finally by gluing $(\star_1, \star_2), \dots, (\star_{2k-1}, \star_{2k})$, which is a continuous map for the matrix distance,

$$(d^{\mathcal{D}_n, k}(V_i, V_j))_{1 \leq i, j \leq a} \xrightarrow{(d)} (d^{\mathcal{P}, k}(V_i, V_j))_{1 \leq i, j \leq a}.$$

And Theorem 4.5.2 (a) follows from Lemma 4.2.1. \square

Proof of Theorem 4.5.2 (b). Let $(\mathcal{D}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{D}}^{\mathbb{N}}$ such that $\mathcal{D}_n \Rightarrow \Theta \in \Omega_{\Theta}$. For every $n \in \mathbb{N}$ let $\mathbf{p}^{\mathcal{D}_n, k}$ be a probability measure on $\mathcal{V}^{\mathcal{D}_n, k}$ such that $\mathbf{p}^{\mathcal{D}_n, k} \rightarrow 0$. For every $n \in \mathbb{N}$ and $1 \leq i \leq 2k$, let $W_i^{\mathcal{D}_n} := \star_i$. Also, let $(W_i^{\mathcal{D}_n})_{i > 2k}$ be a family of independent random variables with law $\mathbf{p}^{\mathcal{D}_n, k}$. Fix $a > 2k$. By Theorem 3.4.2 (b) and Lemma 14, we have

$$\left(\lambda^{\mathcal{D}_n} d^{\mathcal{D}_n}(W_i^{\mathcal{D}_n}, W_j^{\mathcal{D}_n}) \right)_{1 \leq i, j \leq a} \xrightarrow{(d)} (d^{\Theta}(Y_i^{\Theta}, Y_j^{\Theta}))_{1 \leq i, j \leq a}. \quad (4.16)$$

Then by Kolmogorov representation theorem we may assume that (4.16) holds almost surely. Hence, by Lemma 4.10.2 a.s. $\square_{\square, k}(T^{\mathcal{D}_n}) / (\lambda^{\mathcal{D}_n})^k \rightarrow \square_{\square, k}(Y^{\Theta}, Z^{\Theta})$ as $n \rightarrow \infty$. Thus, by Proposition 4.5.1 and dominated convergence, we have for all continuous bounded function $f : \mathbb{R}^{a^2} \rightarrow \mathbb{R}$,

$$\frac{\mathbb{E}[f((\lambda^{\mathcal{D}_n} d^{\mathcal{D}_n}(W_i^{\mathcal{D}_n}, W_j^{\mathcal{D}_n}))_{1 \leq i, j \leq a}) \square_{\square, k}(T^{\mathcal{D}_n})]}{\mathbb{E}[\square_{\square, k}(T^{\mathcal{D}_n})]} \longrightarrow \frac{\mathbb{E}[f((d^{\Theta}(Y_i^{\Theta}, Y_j^{\Theta}))_{1 \leq i, j \leq a}) \square_{\square, k}(Y^{\Theta}, Z^{\Theta})]}{\mathbb{E}[\square_{\square, k}(Y^{\Theta}, Z^{\Theta})]}.$$

Therefore,

$$(\lambda^{\mathcal{D}_n} \bar{d}^{\mathcal{D}_n, k}(W_i^{\mathcal{D}_n}, W_j^{\mathcal{D}_n}))_{1 \leq i, j \leq a} \xrightarrow{(d)} (\bar{d}^{\Theta, k}(Y_i^{\Theta}, Y_j^{\Theta}))_{1 \leq i, j \leq a}.$$

Finally by gluing the k first pair of vertices, which is a continuous map for the matrix distance,

$$(\lambda^{\mathcal{D}_n} d^{\mathcal{D}_n, k}(W_i^{\mathcal{D}_n}, W_j^{\mathcal{D}_n}))_{2k+1 \leq i, j \leq a} \xrightarrow{(d)} (d^{\Theta, k}(Y_i^{\Theta}, Y_j^{\Theta}))_{2k+1 \leq i, j \leq a}.$$

And Theorem 4.5.2 (a) follows from Lemma 4.2.1. \square

Proof of Theorem 4.5.2 (c,d). Since \mathbb{K}_{GP} is a Polish space, and $\Omega_{\mathcal{D}}$ is dense on (Ω, \Rightarrow) , the results follows from Theorem 4.5.2 (a,b) (see Section 3.8.1). Also, they can be proved similarly. \square

4.7.2 Proof of Theorem 4.5.3

Proof of Theorem 4.5.3 (a). Let $(\mathcal{D}_n)_{n \in \mathbb{N}} \in \Omega_{\mathcal{D}}^{\mathbb{N}}$ such that $\mathcal{D}_n \Rightarrow \Theta \in \Omega_{\Theta}$. By Theorem 3.4.2 (b),

$$(\lambda^{\mathcal{D}_n} d^{\mathcal{D}_n}(\star_i, \star_j))_{i,j \in \mathbb{N}} \xrightarrow{(d)} (d^{\Theta}(Y_i^{\Theta}, Y_j^{\Theta}))_{1 \leq i,j \in \mathbb{N}}.$$

Thus, by Lemma 4.9.3 for every $a \in \mathbb{N}$, we have for the a -pointed GH topology (see Section 4.2.3),

$$(T^{\mathcal{D}_n}(\{\star_i\}_{1 \leq i \leq a}), \lambda^{\mathcal{D}_n} d^{\mathcal{D}_n}, \{\star_i\}_{1 \leq i \leq a}) \xrightarrow{\text{WGH}^a} (T^{\Theta}(\{Y_i^{\Theta}\}_{1 \leq i \leq a}), d^{\Theta}, \{Y_i^{\Theta}\}_{1 \leq i \leq a}).$$

Therefore, by Assumption 4.5, we have for the $2k$ -pointed GH topology,

$$(T^{\mathcal{D}_n}, \lambda^{\mathcal{D}_n} d^{\mathcal{D}_n}, \{\star_i\}_{1 \leq i \leq 2k}) \xrightarrow{\text{WGH}^{2k}} (T^{\Theta}, d^{\Theta}, \{Y_i^{\Theta}\}_{1 \leq i \leq 2k}). \quad (4.17)$$

Then, by Skorohod representation theorem we may assume that the above convergence holds almost surely. Thus by Lemma 4.10.2 a.s. $\square_{\square,k}(T^{\mathcal{D}_n}) \rightarrow \square_{\square,k}(\mathbf{Y}^{\Theta}, \mathbf{Z}^{\Theta})$. Then for every continuous bounded function f on $\mathbb{K}_{\text{GH}}^{2k}$ we have by Proposition 4.5.1 and dominated convergence,

$$\frac{\mathbb{E}[f(T^{\mathcal{D}_n}, \lambda^{\mathcal{D}_n} d^{\mathcal{D}_n}, \{\star_i\}_{1 \leq i \leq 2k}) \square_{\square,k}(T^{\mathcal{D}_n})]}{\mathbb{E}[\square_{\square,k}(T^{\mathcal{D}_n})]} \rightarrow \frac{\mathbb{E}[f(T^{\Theta}, d^{\Theta}, \{Y_i^{\Theta}\}_{1 \leq i \leq 2k}) \square_{\square,k}(\mathbf{Y}^{\Theta}, \mathbf{Z}^{\Theta})]}{\mathbb{E}[\square_{\square,k}(\mathbf{Y}^{\Theta}, \mathbf{Z}^{\Theta})]}.$$

Therefore,

$$(T^{\mathcal{D}_n,k}, \lambda^{\mathcal{D}_n} \bar{d}^{\mathcal{D}_n,k}, \{\star_i\}_{1 \leq i \leq 2k}) \xrightarrow{\text{WGH}^{2k}} (T^{\Theta,k}, d^{\Theta,k}, \{Y_i^{\Theta,k}\}_{1 \leq i \leq 2k}).$$

Finally since the gluing of k pair of point is a continuous operation for the $2k$ -pointed GH topology the desired result follows. \square

Proof of Theorem 4.5.3 (b,c). The results can be proved in the exact same way. \square

4.8 Configuration model and multiplicative graphs

The main objective of this section is to explain the connections between the configuration model and multiplicative graphs, and between those models and (\mathcal{D}, k) -graphs and (\mathcal{P}, k) -graphs.

4.8.1 Definitions

For every multigraph G on $\{V_i\}_{i \in \mathbb{N}}$ and $i, j \in \mathbb{N}$ let $\#_{i,j}(G)$ be the number of edges $\{V_i, V_j\}$ in G . So that a multigraph on $\{V_i\}_{i \in \mathbb{N}}$ may be seen as a matrix.

We call a function $f : I \mapsto I$ a matching if $f \circ f = \text{Id}$ and for every $x \in I$, $f(x) \neq x$. Let Ω_{CM} be the set of decreasing sequence (d_1, \dots, d_s) in $\{0\} \cup \mathbb{N}$ such that $\sum_{i=1}^s d_i$ is even.

Algorithm 4.6. Construction of the configuration model from $\mathcal{D} = (d_1, \dots, d_s) \in \Omega_{\text{CM}}$:

- Let $f = (f_1, f_2)$ be a uniform matching of $\{(i, j)\}_{1 \leq i \leq s, 1 \leq j \leq d_i}$.
- The configuration model is the random multigraph $\text{CM}^{\mathcal{D}}$ with vertices $(V_i)_{1 \leq i \leq s}$ and such that for every $1 \leq i \leq s$, $\#_{i,i}(\text{CM}^{\mathcal{D}}) := \frac{1}{2} \sum_{a=1}^{d_i} \mathbf{1}_{f_1(i,a)=i}$ and for $1 \leq i \neq j \leq s$,

$$\#_{i,j}(\text{CM}^{\mathcal{D}}) := \sum_{a=1}^{d_i} \mathbf{1}_{f_1(i,a)=j} = \sum_{a=1}^{d_j} \mathbf{1}_{f_1(j,a)=i}.$$

Let Ω_{MG} be the set of sequence $(\lambda, p_1, \dots, p_s)$ in \mathbb{R}^{+*} with $p_1 \geq \dots \geq p_s$.

Algorithm 4.7. Construction of the multiplicative graph from $\mathcal{P} = (\lambda, p_1, \dots, p_s) \in \Omega_{\text{MG}}$:

- Let $(X_{i,j}^{\mathcal{P}})_{1 \leq i \neq j \leq s}$ be independent Bernoulli random variables with mean $1 - e^{-\lambda p_i p_j}$.
- The multiplicative graph is the random graph $\text{MG}^{\mathcal{P}}$ with vertices (V_1, \dots, V_s) and with edges $\{1 \leq i, j \leq s : X_{i,j} = 1\}$.

Next, we introduce multiplicative multigraphs, which are augmented multiplicative graphs.

Algorithm 4.8. Construction of the multiplicative multigraph from $\mathcal{P} = (\lambda, p_1, \dots, p_s) \in \Omega_{\text{MG}}$:

- Let $(N_{i,j}^{\mathcal{P}})_{1 \leq i, j \leq s}$ be independent Poisson random variables, such that for every $1 \leq i \leq s$, $N_{i,i}^{\mathcal{P}}$ have mean $\lambda p_i^2/2$ and for every $1 \leq i \neq j \leq s$, $N_{i,j}^{\mathcal{P}}$ have mean $\lambda p_i p_j$.
- The multiplicative multigraph is the random multigraph $\text{MG}^{\mathcal{P}+}$ with vertices $(V_i)_{1 \leq i \leq s}$ and such that for every $1 \leq i, j \leq s$, $\#_{i,j}(\text{MG}^{\mathcal{P}+}) := N_{i,j}^{\mathcal{P}}$.

Lemma 4.8.1. *There exists a coupling such that $\text{MG}^{\mathcal{P}}$ is the graph obtained from $\text{MG}^{\mathcal{P}+}$ by removing all its multi-edge. That is, for every $i \neq j$, $\{i, j\}$ is an edge of $\text{MG}^{\mathcal{P}}$ iff $\#_{i,j}(\text{MG}^{\mathcal{P}+}) \geq 1$.*

Proof. It is easy to check that there exists a coupling such that a.s. for every $1 \leq i \neq j \leq s$ $X_{i,j}^{\mathcal{P}} = 0$ iff $N_{i,j}^{\mathcal{P}} = 0$. The result follows. \square

4.8.2 Multiplicative multigraphs as local limit of the configuration model

Lemma 4.8.2. *Let $\mathcal{P} = (\lambda, p_1, \dots, p_s) \in \Omega_{\text{MG}}$. For $n \in \mathbb{N}$, let $\mathcal{D}^n = (d_i^n)_{1 \leq i \leq s^n} \in \Omega_{\text{CM}}$. If $s^n \rightarrow \infty$, and for every $1 \leq i \leq s$, $d_i^n \sim \sqrt{s^n} \lambda p_i$, and for every $n \in \mathbb{N}$, $i > s$, $d_i^n = 1$. Then,*

$$(\#_{i,j}(\text{CM}^{\mathcal{D}^n}))_{1 \leq i, j \leq s} \xrightarrow{(d)} (\#_{i,j}(\text{MG}^{\mathcal{P}+}))_{1 \leq i, j \leq s}.$$

Remark. From this result, one may see the LIFO-queue algorithm of Broutin, Duquesne, Wang [37, 36] as a limit of a recursive construction, based on a DFS exploration, of a uniform matching.

Proof. Let $(\mathcal{D}^n)_{n \in \mathbb{N}}$ and \mathcal{P} be as in the statement. For $n \in \mathbb{N}$, let $f^n = (f_1^n, f_2^n)$ a uniform matching of $\{(i, j)\}_{1 \leq i \leq s^n, 1 \leq j \leq d_i^n}$. We may assume that $\text{CM}^{\mathcal{D}^n}$ is constructed from f^n by Algorithm 4.6. The main idea is that for n large enough $\{f_1^n(i, j)\}_{1 \leq i \leq s^n, 1 \leq j \leq d_i^n}$ are mostly independent. Since Poisson random variables appears as the limits of Bernoulli trials this explain the convergence. From there, there are many standard ways to justify the convergence.

Below we briefly present a method based on random point process. We let the reader refer to Kallenberg [57] Section 4 for more details on convergence of point process. Let ν^n be the random measure on $\mathbb{K} := \{(i, j)\}_{1 \leq i \leq j \leq s} \times \mathbb{R}^2$ defined by

$$\nu^n := \sum_{1 \leq i < j \leq s} \sum_{\substack{1 \leq a \leq d_i \\ 1 \leq b \leq d_j}} \mathbf{1}_{f(i,a)=(j,b)} \delta_{(i,j,a/\sqrt{s^n}, b/\sqrt{s^n})} + \sum_{1 \leq i \leq n} \sum_{1 \leq a < b \leq d_i} \mathbf{1}_{f(i,a)=(i,b)} \delta_{(i,i,a/\sqrt{s^n}, b/\sqrt{s^n})}.$$

It is enough to prove that $\{\nu^n\}_{n \in \mathbb{N}}$ converges vaguely toward a Poisson point process of rate

$$d\nu := \sum_{1 \leq i \leq j \leq s} \mathbf{1}_{0 \leq x \leq \lambda p_i} \mathbf{1}_{0 \leq y \leq \lambda p_j} \delta_{i,j} dx dy + \sum_{1 \leq i \leq n} \mathbf{1}_{0 \leq x \leq y \leq \lambda p_i} \delta_{i,i} dx dy. \quad (4.18)$$

Indeed, provided this convergence, the desired result directly follows by integration over $dx dy$.

To this end, first note that for every $n \in \mathbb{N}$, writing $m_n := \sum_{i=1}^{s^n} d_i^n$,

$$\begin{aligned} \mathbb{E}[\nu^n(\mathbb{K})] &= \sum_{1 \leq i < j \leq s} \sum_{\substack{1 \leq a \leq d_i \\ 1 \leq b \leq d_j}} \mathbb{P}(f(i, a) = (j, b)) + \sum_{1 \leq i \leq n} \sum_{1 \leq a < b \leq d_i} \mathbb{P}(f(i, a) = (i, b)). \\ &= \sum_{1 \leq i < j \leq s} \frac{d_i d_j}{m^n} + \sum_{1 \leq i \leq n} \sum_{1 \leq a < b \leq d_i} \frac{d_i^2/2}{m^n} \\ &\rightarrow \sum_{1 \leq i < j \leq s} \lambda p_i p_j + \sum_{1 \leq i \leq n} \sum_{1 \leq a < b \leq d_i} \lambda p_i^2/2 \end{aligned}$$

where the last inequality comes from the assumptions of the lemma on $(\mathcal{D}_n)_{n \in \mathbb{N}}$. Thus, $\{\nu^n\}_{n \in \mathbb{N}}$ is tight for the vague topology. Let ν be a sub-sequential limit of $\{\nu^n\}_{n \in \mathbb{N}}$.

By a similar computation, for every $1 \leq i < j \leq s$, and $0 \leq a \leq a' \leq \lambda p_i$, $0 \leq b \leq b' \leq \lambda p_j$,

$$\mathbb{E}[\nu(\{i, j\} \times [a, a'] \times [b, b'])] = \lim_{n \rightarrow \infty} \mathbb{E}[\nu^n(\{i, j\} \times [a, a'] \times [b, b'])] = \lambda(a' - a)(b' - b).$$

And for every $1 \leq i \leq s$, $0 \leq a \leq a' \leq b \leq b' \leq p_i$.

$$\mathbb{E}[\nu(\{i, i\} \times [a, a'] \times [b, b'])] = \lim_{n \rightarrow \infty} \mathbb{E}[\nu^n(\{i, i\} \times [a, a'] \times [b, b'])] = \lambda(a' - a)(b' - b).$$

Next, we prove that ν satisfies the independency criterium. Beforehand let us introduce some notations. Let $v(\cdot, \cdot)$ be the covariance of two random variables. Let

$$S := \left\{ (i, j, a, b) \in \mathbb{N}^4 : 1 \leq i < j \leq s, \frac{1 \leq a \leq d_i}{1 \leq b \leq d_j} \right\} \cup \left\{ (i, i, a, b) \in \mathbb{N}^4 : 1 \leq i \leq s, 1 \leq a < b \leq d_i \right\}.$$

For every $K_1, K_2 \subset \mathbb{K}$ disjoint compact set, for every $n \in \mathbb{N}$, $\text{Cov}(\nu^n(K_1), \nu^n(K_2))$ equal

$$\sum_{\substack{(i, j, a, b) \in S \\ (i', j', a', b') \in S}} \mathbf{1}_{(i, j, \frac{a}{\sqrt{s^n}}, \frac{b}{\sqrt{s^n}}) \in K_1} \mathbf{1}_{(i', j', \frac{a'}{\sqrt{s^n}}, \frac{b'}{\sqrt{s^n}}) \in K_2} \text{Cov}(\mathbf{1}_{f(i, a) = (j, b)}, \mathbf{1}_{f(i', a') = (j', b')}).$$

Then, by distinguishing whether it is possible to have both $f(i, a) = (j, b)$ and $f(i', a') = (j', b')$, note that in the last sum there are $O(\#S)$ terms that are equal to $0 - 1/(m^n)^2$, $O((\#S)^2)$ terms that are equal to $1/(m^n)(m^n - 2) - (1/m^n)^2 = O(1/(m^n)^3)$, and the others that are null. Therefore,

$$|\text{Cov}(\nu^n(K_1), \nu^n(K_2))| = O(\#S)O(1/(m^n)^2) + O((\#S)^2)O(1/(m^n)^3) = O(1/m^n) \rightarrow 0.$$

Since the last convergence hold for every disjoint compact $K_1, K_2 \subset \mathbb{K}$, we have that for every disjoint compact $K'_1, K'_2 \subset \mathbb{K}$, $\text{Cov}(\nu(K'_1), \nu(K'_2)) = 0$.

Finally, to prove that ν is a Poisson point process of rate (4.18) it is enough to check that a.s. for every $x \in \mathbb{K}$, $\nu(x) \in \{0, 1\}$. To this end, one may adapt the previous argument to show that there exists $C > 0$, such that for every $x \in \mathbb{K}$, $\varepsilon > 0$, writing $B(x, \varepsilon)$ for the closed ball centered at x of radius ε for $\|\cdot\|_\infty$, if $B(x, \varepsilon)$ does not intersect $\{(i, i, 1/2, 1/2)\}_{1 \leq i \leq s}$ then

$$\mathbb{E}[\nu(B(x, \varepsilon))(\nu(B(x, \varepsilon)) - 1)] \leq C\varepsilon^2.$$

This implies the desired property, and so concludes the proof. \square

4.8.3 Connections with (\mathcal{D}, k) -graphs and (\mathcal{P}, k) -graphs

Recall that for every multigraph G on $\{V_i\}_{i \in \mathbb{N}}$, $\circ(G) := \prod_{i \in \mathbb{N}} 2^{\#i, i(G)} \prod_{i, j \in \mathbb{N}} \#i, j(G)!$.

Lemma 4.8.3. *Let $k \in \mathbb{N}$ we have the following assertions:*

- (a) *Let $\mathcal{D} = (d_1, \dots, d_s) \in \Omega_{\text{CM}}$ such that $\sum_{i=1}^s d_i = 2s + k - 2$. Then $\text{CM}^{\mathcal{D}}$ biased by $\circ(\text{CM}^{\mathcal{D}})$ and conditioned at being connected is a $((d_1 - 1, \dots, d_s - 1), k)$ -graph.*
- (b) *Let $\mathcal{W} = (\lambda, w_1, \dots, w_s) \in \Omega_{\text{MG}}$. For every $1 \leq i \leq s$, let $p_i := w_i / \sum_{j=1}^s w_j$. Let $\mathcal{P} = (p_1, \dots, p_s, 0, 0, \dots)$. Then $\text{MG}^{\mathcal{W}+}$ biased by $\circ(\text{MG}^{\mathcal{W}+})$ and conditioned at being connected and having surplus k is a (\mathcal{P}, k) -graph.*

Remark. The bias is not really important as typically those graphs are studied in a regime where with high probability the multigraph is a graph. Also removing this bias only remove the term $\circ(\mathcal{G}_{(\star_i)_{1 \leq i \leq 2k}}(T))$ in Section 4.4.2 which does not change our proofs.

Proof. (a) is a classic and is easy to obtain from a quick enumeration. So we focus on (b). The main idea is that, on the one hand multiplicative multigraph are limits of the configuration model, and on the other hand (\mathcal{P}, k) -graph are limits of (\mathcal{D}, k) -graph. Thus by identification, (b) follows. Let us detail:

Fix $k, \mathcal{W}, \mathcal{P} \in \Omega_{\text{MG}}$ as in (b). Let $(\mathcal{D}^n)_{n \in \mathbb{N}}$ be a sequence of Ω_{CM} as in Lemma 4.8.2. Then write $\text{CM}^{\mathcal{W}, k}$ for the random multigraph $\text{MG}^{\mathcal{W}+}$ biased by $\circ(\text{MG}^{\mathcal{W}})$ and conditioned at being connected and having surplus k . Also, write for $n \in \mathbb{N}$, $\text{CM}^{\mathcal{D}^n, k}$ for the random multigraph $\text{CM}^{\mathcal{D}^n}$ biased by $\circ(\text{CM}^{\mathcal{D}^n})$ conditioned on the fact that the subgraph of $\text{CM}^{\mathcal{D}^n}$ on $(V_i)_{1 \leq i \leq s}$ is connected and have surplus k . By Lemma 4.8.2, we have,

$$(\#_{i,j}(\text{CM}^{\mathcal{D}^n, k}))_{1 \leq i, j \leq s} \xrightarrow{(d)} (\#_{i,j}(\text{MG}^{\mathcal{W}, k}))_{1 \leq i, j \leq s}. \quad (4.19)$$

Then, for every $n \in \mathbb{N}$ let $S^n + 2s$ be the number of vertices that are in the connected component of $(V_i)_{1 \leq i \leq s}$ in $\text{CM}^{\mathcal{D}^n}$. Then let $\mathcal{D}^{n-} := (d_1^n, \dots, d_s^n, 1, \dots, 1)$ with S^n number 1 at the end. It is well known that for every $n \in \mathbb{N}$, conditioned on S^n , $\text{CM}^{\mathcal{D}^n, k}$ have the same law as $\text{CM}^{\mathcal{D}^{n-}}$ (where the vertices outside $(V_i)_{1 \leq i \leq s}$ in $\text{CM}^{\mathcal{D}^n}$ have been relabeled). More precisely,

$$(\#_{i,j}(\text{CM}^{\mathcal{D}^n, k}))_{1 \leq i, j \leq s} \stackrel{(d)}{=} (\#_{i,j}(\text{CM}^{\mathcal{D}^{n-}}))_{1 \leq i, j \leq s}.$$

Therefore, it directly follows from (4.19), that if for $n \in \mathbb{N}$, $\text{CM}^{\mathcal{D}^{n-}, k}$ be the random multigraph $\text{CM}^{\mathcal{D}^{n-}}$ biased by $\circ(\text{CM}^{\mathcal{D}^{n-}})$ and conditioned at being connected, then

$$(\#_{i,j}(\text{CM}^{\mathcal{D}^{n-}, k}))_{1 \leq i, j \leq s} \xrightarrow{(d)} (\#_{i,j}(\text{MG}^{\mathcal{W}, k}))_{1 \leq i, j \leq s}. \quad (4.20)$$

Next let for $n \in \mathbb{N}$, $\mathcal{D}_n \in \Omega_{\mathcal{D}}$ be the sequence $(d_1^n - 1, \dots, d_s^n - 1, 0, \dots, 0, 0, \dots, 0)$ where we added $S^n + 2k$ numbers 0 at the end. We have by (a) for every $n \in \mathbb{N}$,

$$(\#_{i,j}(G^{\mathcal{D}_n, k}))_{1 \leq i, j \leq s} \stackrel{(d)}{=} (\#_{i,j}(\text{CM}^{\mathcal{D}^{n-}, k}))_{1 \leq i, j \leq s}$$

Therefore by (4.20),

$$(\#_{i,j}(G^{\mathcal{D}_n, k}))_{1 \leq i, j \leq s} \xrightarrow{(d)} (\#_{i,j}(\text{MG}^{\mathcal{W}, k}))_{1 \leq i, j \leq s}. \quad (4.21)$$

Finally note that $\mathcal{D}_n \Rightarrow \mathcal{P}$. So, by (4.15) and Lemma 4.9.3, as $n \rightarrow \infty$ the subtree of $T^{\mathcal{D}_n, k}$ spanned by $\{V_i\}_{1 \leq i \leq s} \cup \{\star_i\}_{1 \leq i \leq 2k}$ converges weakly toward the subtree of $T^{\mathcal{P}, k}$ spanned by the same vertices. Therefore, we have by gluing $(\star_1, \star_2), \dots, (\star_{2k-1}, \star_{2k})$, then counting the edges,

$$(\#_{i,j}(G^{\mathcal{D}_n, k}))_{1 \leq i, j \leq s} \xrightarrow{(d)} (\#_{i,j}(G^{\mathcal{P}, k}))_{1 \leq i, j \leq s}.$$

And (4.21) concludes the proof. \square

To conclude the section let us compute the law of (\mathcal{P}, k) -graph.

Lemma 4.8.4. *Let $k \in \mathbb{N}$. Let $(p_1, \dots, p_s, 0, 0, \dots) \in \Omega_{\mathcal{P}}$. We have for every connected multigraph G on $\{V_i\}_{1 \leq i \leq s}$ with surplus k , writing α for proportional,*

$$\mathbb{P}(G^{\mathcal{P},k} = G) \propto \prod_{1 \leq i,j \leq s} (p_i p_j)^{\#_{i,j}(G)}.$$

Proof. Keep the notations of Lemma 4.8.3 (b). By definition of $\Omega_{\text{MG}}^{\mathcal{W}+}$, we have

$$\mathbb{P}(\Omega_{\text{MG}}^{\mathcal{W}+} = G) = \prod_{1 \leq i < j \leq s} \frac{(\lambda p_i p_j)^{\#_{i,j}(G)} e^{-\lambda p_i p_j}}{\#_{i,j}(G)!} \prod_{1 \leq i \leq s} \frac{(\lambda p_i^2 / 2)^{\#_{i,i}(G)} e^{-\lambda p_i^2 / 2}}{\#_{i,i}(G)!}.$$

So the result follows from Lemma 4.8.3 (b). \square

Remark. • When $k = 0$ the result is well known and is a classical definition for \mathcal{P} -trees.
• When the weight of the edges is not multiplicative, one can still construct similar multigraphs. Moreover, Lemma 4.8.4 is still true in this case. For $k = 0$, this relates those models with the general spanning trees constructed by Aldous–Bröder algorithm [9, 35].

4.9 Appendix A: \mathbb{R} -tree reconstruction problem

Recall that a \mathbb{R} -tree is a loopless geodesic metric space. If \mathcal{T} is a \mathbb{R} -tree, we say that $x \in \mathcal{T}$ is a leaf of \mathcal{T} if $\mathcal{T} \setminus \{x\}$ is connected. Let (\mathcal{T}, d) be a \mathbb{R} -tree with leaves $\{\star_i\}_{1 \leq i \leq N}$. In this section we reconstruct a \mathbb{R} -tree isometric to \mathcal{T} from $(d_{i,j})_{1 \leq i,j \leq N} := (d(\star_i, \star_j))_{1 \leq i,j \leq N}$.

For every $a, b \in \mathcal{T}$ let $\llbracket a, b \rrbracket$ be the geodesic path between a and b . Since \mathcal{T} is a \mathbb{R} -tree note that for every $1 \leq a \neq b \neq c \leq N$ there exists a unique vertex $\star_{a,b,c}$ in $\llbracket \star_a, \star_b \rrbracket \cap \llbracket \star_a, \star_c \rrbracket \cap \llbracket \star_b, \star_c \rrbracket$.

Lemma 4.9.1. *For every $1 \leq a \neq b \neq c \leq N$, $2d(\star_a, \star_{a,b,c}) = d_{a,b} + d_{a,c} - d_{b,c}$.*

Proof. Note that $d_{a,b} = d(\star_a, \star_{a,b,c}) + d(\star_{a,b,c}, \star_b)$, and similarly $d_{b,c} = d(\star_b, \star_{a,b,c}) + d(\star_{a,b,c}, \star_c)$ and $d_{a,c} = d(\star_a, \star_{a,b,c}) + d(\star_{a,b,c}, \star_c)$. The desired equality follows by sum. \square

To reconstruct \mathcal{T} we reconstruct recursively for $1 \leq n \leq N$ the subtree spanned by $\{\star_i\}_{1 \leq i \leq n}$, which is $\mathcal{T}_n := \bigcup_{1 \leq i,j \leq n} \llbracket \star_i, \star_j \rrbracket$. It is easy to check that for $1 \leq n \leq N$, (\mathcal{T}_n, d) is a \mathbb{R} -tree. Moreover, note that $\mathcal{T}_{n+1} = \mathcal{T}_n \cup \llbracket W_n, \star_{n+1} \rrbracket$, where W_n is the closest point from \star_{n+1} on \mathcal{T}_n . Therefore, it is enough to reconstruct $(W_i)_{1 \leq i < N}$ and $(d(W_n, \star_{n+1}))_{1 \leq n < N}$. This suggest the following construction. Below, $(\beta_i)_{i \in \mathbb{N}}$ is the canonical base of $\mathbb{R}^{+\mathbb{N}}$.

Algorithm 4.9. Reconstruction of a \mathbb{R} -tree on $(\mathbb{R}^{+\mathbb{N}}, \|\cdot\|_\infty)$ from $M = (d_{i,j})_{1 \leq i,j \leq N}$.

- Let $\star_1^M := 0$. Let $\mathcal{T}_1^M = (\{\star_1^M\}, \emptyset)$.
- Let $\star_2^M := (0, d(\star_1^M, \star_2^M))$. Let $\mathcal{T}_2^M := \{x\beta_1, 0 \leq x \leq d(\star_1^M, \star_2^M)\}$.
- For every $2 \leq n < N$:
 - Let $1 \leq b_n^M \neq c_n^M \leq n$ be the smallest integers (for some predetermined order) that minimize $d_{n+1,b_n^M} + d_{n+1,c_n^M} - d_{b_n^M,c_n^M}$.
 - Let W_n^M be the vertex of \mathcal{T}_n^M at distance $d_{b_n^M,n+1} + d_{b_n^M,c_n^M} - d_{n+1,c_n^M}$ of $\star_{b_n^M}^M$ and at distance $d_{c_n^M,n+1} + d_{c_n^M,b_n^M} - d_{n+1,b_n^M}$ of $\star_{c_n^M}^M$. (See below for existence and unicity.)
 - Let $\mathcal{T}_{n+1}^M := \mathcal{T}_n^M \cup \{W_n^M + x\beta_n, 0 \leq x \leq d_{n+1,b_n^M} + d_{n+1,c_n^M} - d_{b_n^M,c_n^M}\}$.
- Let $\mathcal{T}^M := \mathcal{T}_N^M$.

Remark. The idea of constructing subtrees on $(\mathbb{R}^{+\mathbb{N}}, d_\infty)$ comes from Aldous [10].

Lemma 4.9.2. *Let (\mathcal{T}, d) be a \mathbb{R} -tree with leaves $\{\star_i\}_{1 \leq i \leq N}$. Let $M = (d(\star_i, \star_j))_{1 \leq i, j \leq N}$. Then:*

- a) *For every $1 \leq n \leq N$, \mathcal{T}_n^M is well defined.*
- b) *$(\mathcal{T}, d, \star_1, \dots, \star_N)$ and $(\mathcal{T}^M, d_\infty, \star_1^M, \dots, \star_N^M)$ are isometric (see Section 4.2.3).*

Proof. We prove by induction that for $1 \leq n \leq N$, \mathcal{T}_n^M is well defined and that $(\mathcal{T}_n, d, \star_1, \dots, \star_n)$ and $(\mathcal{T}_n^M, d_\infty, \star_1^M, \dots, \star_n^M)$ are isometric. First if $n = 1$ or $n = 2$ then the result is obvious. Then let $2 \leq n < N$ such that \mathcal{T}_n^M is well defined and such that there exists an isometry ϕ_n from $(\mathcal{T}_n, d, \star_1, \dots, \star_n)$ to $(\mathcal{T}_n^M, d_\infty, \star_1^M, \dots, \star_n^M)$.

Recall that $\mathcal{T}_n = \bigcup_{1 \leq i, j \leq n} \llbracket \star_i, \star_j \rrbracket$, and that W_n is the closest point from \star_{n+1} on \mathcal{T}_n . So there exist $1 \leq b_n \neq c_n \leq n$ such that $W_n \in \llbracket \star_{b_n}, \star_{c_n} \rrbracket$. Hence $W_n = \star_{n+1, b_n, c_n}$. Then by Lemma 4.9.1,

$$d(\star_{n+1}, W_n) = d(\star_{n+1}, \star_{n+1, b_n, c_n}) = d_{n+1, b_n} + d_{n+1, c_n} - d_{b_n, c_n}. \quad (4.22)$$

Also, by Lemma 4.9.1, since W_n is the closest point from \star_{n+1} on \mathcal{T}_n ,

$$d(\star_{n+1}, W_n) \leq \min_{1 \leq b \leq c \leq n} d(\star_{n+1}, \star_{n+1, b, c}) = \min_{1 \leq b \leq c \leq n} d_{n+1, b} + d_{n+1, c} - d_{b, c}.$$

Therefore, we may assume that $b_n = b_n^M$ and that $c_n = c_n^M$.

Furthermore, since $W_n = \star_{n+1, b_n, c_n}$,

$$d(W_n, \star_{b_n}) = d_{b_n, n+1} + d_{b_n, c_n} - d_{n+1, c_n} \quad ; \quad d(W_n, \star_{c_n}) = d_{c_n, n+1} + d_{c_n, b_n} - d_{n+1, b_n}. \quad (4.23)$$

W_n is the only vertex of \mathcal{T}_n satisfying (4.23). Indeed, any vertex V satisfying (4.23) must also satisfy

$$d(V, \star_{b_n}) + d(V, \star_{c_n}) = d(\star_{b_n}, \star_{c_n}),$$

and so must be W_n , the only vertex of $\llbracket \star_{b_n}, \star_{c_n} \rrbracket$ at distance $d_{b_n, n+1} + d_{b_n, c_n} - d_{n+1, c_n}$ of \star_{b_n} .

Then, by definition of ϕ_n and (4.23), $\phi_n(W_n)$ is the only vertex of \mathcal{T}_n^M satisfying

$$d_\infty(\phi_n(W_n), \star_{b_n}^M) = d_{b_n, n+1} + d_{b_n, c_n} - d_{n+1, c_n} \quad ; \quad d_\infty(\phi_n(W_n), \star_{c_n}^M) = d_{c_n, n+1} + d_{c_n, b_n} - d_{n+1, b_n}.$$

Therefore, W_n^M and thus \mathcal{T}_n^{M+1} are well defined.

Finally recall that $\mathcal{T}_{n+1} = \mathcal{T}_n \cup \llbracket W_n, \star_{n+1} \rrbracket$. Then by definition of \mathcal{T}_{n+1}^M , (4.22) and $\phi_n(W_n) = W_n^M$, we have $\mathcal{T}_{n+1}^M = \mathcal{T}_n^M \cup \{\phi_n(W_n) + x\beta_n, 0 \leq x \leq d(W_n, \star_{n+1})\}$. Also both union are disjoint, so one can extend ϕ_n to an isometry ϕ_{n+1} from \mathcal{T}_{n+1} to \mathcal{T}_{n+1}^M such that for every $x \in \llbracket W_n, \star_{n+1} \rrbracket$, $\phi_{n+1}(x) := \phi_n(W_n) + d(W_n, x)\beta_n$. This concludes the proof. \square

We now prove a corollary, which we use to prove Theorem 4.5.3.

Lemma 4.9.3. *Let $((\mathcal{T}_n, d_n))_{n \in \mathbb{N}}$ be a sequence of \mathbb{R} -trees with leaves $\{\star_i^n\}_{1 \leq i \leq N}$. Assume that*

$$\forall 1 \leq i \neq j \leq N, \quad d_n(\star_i^n, \star_j^n) \longrightarrow d_{i,j} \in \mathbb{R}^{+*}.$$

Then there exist a unique N -pointed \mathbb{R} -tree $(\mathcal{T}, d, (\star_1, \dots, \star_N))$ up to isometry such that for every $1 \leq i, j \leq N$, $d(\star_i, \star_j) = d_{i,j}$. Moreover, $((\mathcal{T}_n, d_n, (\star_i^n)_{1 \leq i \leq N}))_{n \in \mathbb{N}}$ converges for the N -pointed Gromov–Hausdorff topology (see Section 4.2.3) toward $(\mathcal{T}, d, (\star_i)_{1 \leq i \leq N})$.

Proof. First uniqueness follows from Lemma 4.9.2. Let us prove existence. For every $n \in \mathbb{N}$ let $M^n := (d^n(\star_i^n, \star_j^n))_{1 \leq i, j \leq N}$. Similarly let $M = (d_{i,j})_{1 \leq i, j \leq N}$. Note that for every $n \in \mathbb{N}$,

$$\mathcal{T}^{M^n} \subset \left\{ x \in \mathbb{R}^{+N} \times \{0\}^N, d_\infty(0, x) \leq \max_{1 \leq j \leq N} d^n(\star_1^n, \star_j^n) \right\}. \quad (4.24)$$

Thus for every $1 \leq m < N$, $\{W_m^{M^n}\}_{n \in \mathbb{N}}$ is tight.

Let $(n_i)_{i \in \mathbb{N}}$ be an increasing sequence of integer such that for every $1 \leq m < N$, $(W_m^{M_{n_i}})_{i \in \mathbb{N}}$ converges toward W_m^∞ . Then, intuitively, the whole Algorithm 4.9 converges. More precisely, $\mathcal{T}^{M_{n_i}}$ converges for the Hausdorff distance toward a \mathbb{R} -tree \mathcal{T} that is constructed from Algorithm 4.9 with entry M and where for $1 \leq m < N$, W_m^M is replaced by W_m^∞ . Furthermore, for every $1 \leq m \leq N$, $(\star_m^{M_{n_i}})_{i \in \mathbb{N}}$ converges toward \star_m^∞ which is also obtained from the same algorithm.

Then it is easy to check that the leaves of \mathcal{T} are $(\star_m^\infty)_{1 \leq m \leq N}$, and that for every $1 \leq i, j \leq N$,

$$d(\star_i^\infty, \star_j^\infty) = \lim_{m \rightarrow \infty} d(\star_i^{n_m}, \star_j^{n_m}) = \lim_{m \rightarrow \infty} d_{n_m}(\star_i^n, \star_j^n) = d_{i,j}.$$

Therefore (\mathcal{T}, d) satisfies the properties described in the lemma.

Finally, let us prove the convergence. First, the right-hand side of (4.24) is compact so $(\mathcal{T}^{M_n})_{n \in \mathbb{N}}$ is a tight sequence for the Hausdorff topology. Then from any converging subsequence of $(\mathcal{T}^{M_n})_{n \in \mathbb{N}}$ we may further extract such that $(W_m^{M_{n_i}})_{i \in \mathbb{N}}$ converges. It then follows from the first part of the proof that $(\mathcal{T}^{M_n})_{n \in \mathbb{N}}$ converges for the Hausdorff distance toward \mathcal{T} . Finally by Lemma 4.9.2 for every $n \in \mathbb{N}$, $(\mathcal{T}^{M_n}, d_\infty, (\star_i^M)_{1 \leq i \leq N})$ and $(\mathcal{T}_n, d_n, (\star_i^n)_{1 \leq i \leq N})$ are isometric. The desired convergence follows. \square

4.10 Appendix B: \square_c is a continuous function of the matrix distance

Recall Section 4.4.4. Let us extend \square_i to general \mathbb{R} -trees. Note that for every \mathbb{R} -tree (\mathcal{T}, d) , one may define a Borel measure λ on \mathcal{T} such that for every $a, b \in \mathcal{T}$, $\lambda[a, b] = d(a, b)$. By analogy with \mathbb{R} we call λ the Lebesgue measure. For $c \in \mathbb{N}$, if $\{\star_i\}_{1 \leq i \leq 2c}$ are leaves of \mathcal{T} , we let $\text{cyc}_c(\mathcal{T})$ be the set of all $x \in \mathbb{R}$ such that $\mathfrak{G}_{(\star_i)_{1 \leq i \leq 2c}}(\mathcal{T}) \setminus \{x\}$ is connected. By Lemma 4.10.1 below $\text{cyc}_c(\mathcal{T})$ is measurable. Let $\square_c(\mathcal{T})$ be its Lebesgue measure.

It is easy to check that this definition of \square_c extends the definition of \square_c described in Section 4.4.4 and informally equals $\square_c + c$ where \square_c is defined in the discrete setting in Section 4.4.2. The goal of this section is to prove a continuity result for \square_c .

Lemma 4.10.1. *For every $c \in \mathbb{N}$, for every \mathbb{R} -tree (\mathcal{T}, d) , if $\{\star_i\}_{1 \leq i \leq 2c}$ are leaves of \mathcal{T} then*

$$\text{cyc}_c(\mathcal{T}) = \bigcup_{1 \leq b \leq c} [\star_{2b-1}, \star_{2b}].$$

Proof. On the one hand, for every $b \leq c$, $[\star_{2b-1}, \star_{2b}] \subset \text{cyc}(\mathfrak{G}_{(\star_i)_{1 \leq i \leq 2c}}(\mathcal{T}))$ since $[\star_{2b-1}, \star_{2b}]$ is a cycle in $\mathfrak{G}_{(y_i)_{1 \leq i \leq 2c}}(\mathcal{T})$ (a geodesic path that have the same starting and ending point).

On the other hand, let $x \in \mathcal{T} \setminus \bigcup_{1 \leq b \leq c} [\star_{2b-1}, \star_{2b}]$. If $\mathcal{T} \setminus \{x\}$ is connected then $x \notin \text{cyc}_c(\mathcal{T})$ since $\mathfrak{G}_{(y_i)_{1 \leq i \leq 2c}}(\mathcal{T}) \setminus \{x\}$ is also connected. Otherwise $\mathcal{T} \setminus \{x\}$ is disconnected. Let $\mathcal{T}_1, \mathcal{T}_2$ be the two connected components of $\mathcal{T} \setminus \{x\}$. For every $1 \leq b \leq 2c$ note that since $x \notin [\star_{2b-1}, \star_{2b}]$, either $\star_{2b-1}, \star_{2b} \in \mathcal{T}_1$ or $\star_{2b-1}, \star_{2b} \in \mathcal{T}_2$. Therefore, by induction, for every $1 \leq b \leq c$, $\mathfrak{G}_{(y_i)_{1 \leq i \leq 2b}}(\mathcal{T}) \setminus \{x\}$ is still disconnected. In other words, $x \notin \text{cyc}_c(\mathcal{T})$. \square

Lemma 4.10.2. *Let $c \in \mathbb{N}$. There exists a continuous function $f_c : \mathbb{R}^{2c \times 2c}$, such that for every \mathbb{R} -tree (\mathcal{T}, d) such that $\{\star_i\}_{1 \leq i \leq 2c}$ are leaves of \mathcal{T} ,*

$$\square_c((\mathcal{T}, d)) = f_c((d(\star_i, \star_j))_{1 \leq i, j \leq 2c}).$$

Furthermore for every $\lambda \in \mathbb{R}^+$, $\square_c((\mathcal{T}, \lambda d)) = \lambda \square_c((\mathcal{T}, d))$.

Proof. Fix $c \in \mathbb{N}$. note that \square_c is invariant under isometry so Lemma 4.9.2 imply that f_c exists. Also, the scaling property is straightforward from the initial definition since rescaling d rescale the Lebesgue measure. Thus, it remains to prove the continuity property.

To this end, we prove an explicit formula for $\square_c((\mathcal{T}, d)) - \square_{c-1}((\mathcal{T}, d))$ using Lemma 4.10.1. Let $M := (d_{i,j})_{1 \leq i,j \leq 2c} := (d(\star_i, \star_j))_{1 \leq i,j \leq 2c}$. Since (\mathcal{T}, d) is a \mathbb{R} -tree we may define φ_c as the unique isometry from $[0, d_{2i-1, 2i}]$ to $[\star_{2i-1}, \star_{2i}]$ such that $\varphi_c(0) = \star_{2i-1}$ and $\varphi_c(d_{2i-1, 2i}) = \star_{2i}$. We have by the transport formula,

$$\begin{aligned} \square_c((\mathcal{T}, d)) - \square_{c-1}((\mathcal{T}, d)) &= \lambda \left([\star_{2i-1}, \star_{2i}] \setminus \bigcup_{1 \leq b < c} \mathcal{E}_{[\star_{2b-1}, \star_{2b}]} \right) \\ &= \int_0^{d_{2i-1, 2i}} \prod_{b=1}^{c-1} \mathbf{1}_{\varphi_c(x) \notin [\star_{2b-1}, \star_{2b}]} dx. \end{aligned} \quad (4.25)$$

Then, since \mathcal{T} is a \mathbb{R} -tree, for every $1 \leq b < c$, $[\star_{2b-1}, \star_{2b}] \cap [\star_{2c-1}, \star_{2c}]$ is a segment. For every $1 \leq b < c$, let I_b be the real interval such that $x \in I_b$ iff $\varphi_c(x) \in [\star_{2b-1}, \star_{2b}]$. Intuitively, by (4.25) it is enough to show that for $1 \leq b < c$, I_b may be seen as a continuous function of M . Indeed, this would directly imply that $f_c(M) - f_{c-1}((d_{i,j})_{1 \leq i,j \leq 2c-2})$ is continuous. And the desired result would then follow by induction.

Thus let us fix $1 \leq b < c$, and let us compute I_b . For every $a, b, c \in \mathcal{T}$, $\star_{a,b,c}$ let be the unique vertex in $[\star_a, \star_b] \cap [\star_a, \star_c] \cap [\star_b, \star_c]$. Since \mathcal{T} is a \mathbb{R} -tree, note that

$$d(\star_{2c-1}, \star_{2c-1, 2c, 2b-1}) \neq d(\star_{2c-1}, \star_{2c-1, 2c, 2b}) \implies I_b = [d(\star_{2c-1}, \star_{2c-1, 2c, 2b-1}), d(\star_{2c-1}, \star_{2c-1, 2c, 2b})]^+,$$

where for $x, y \in \mathbb{R}$, $[x, y]^+ := [\min(x, y), \max(x, y)]$. Also note that

$$d(\star_{2c-1}, \star_{2c-1, 2c, 2b-1}) = d(\star_{2c-1}, \star_{2c-1, 2c, 2b}) \implies I_b \in \{\emptyset, \{d(\star_{2c-1}, \star_{2c-1, 2c, 2b})\}\}.$$

Moreover, by Lemma 4.9.1,

$$d(\star_{2c-1}, \star_{2c-1, 2c, 2b-1}) = d_{2c-1, 2c} + d_{2c-1, 2b-1} - d_{2c, 2b-1},$$

and

$$d(\star_{2c-1}, \star_{2c-1, 2c, 2b}) = d_{2c-1, 2c} + d_{2c-1, 2b} - d_{2c, 2b}.$$

Therefore I_b may be seen as a continuous function of M . Finally (4.25) concludes the proof. \square

Chapter 5

Looptree, Fennec, and Snake of ICRT.

"Having imagination, it takes you an hour to write a paragraph that, if you were unimaginative, would take you only a minute. Or you might not write the paragraph at all."

Franklin P. Adams

This chapter is adapted from [30].

We introduce a new theory of plane \mathbb{R} -tree, to define plane ICRT (inhomogeneous continuum random tree), and its looptree, fennec (a Gaussian free field on the looptree), and snake. We prove that a.s. the looptree is compact, and that a.s. the fennec and snake are continuous. We compute the looptree's fractal dimensions, and the fennec and snake's Hölder exponent. Alongside, we define a Gaussian free field on the ICRT, and prove a condition for its continuity. In a companion paper [26], we prove that the looptrees, fennecs, and snakes of trees with fixed degree sequence converge toward the looptrees, fennecs and snakes of ICRT.

Contents

5.1	Introduction	138
5.1.1	Motivations and overview of the results	138
5.1.2	Theory of plane \mathbb{R} -tree: an overview	139
5.1.3	Stick breaking and the chaining method	139
5.2	Model and main results.	140
5.2.1	Basic notions on \mathbb{R} -trees and plane \mathbb{R} -trees.	140
5.2.2	ICRT, and plane ICRT	140
5.2.3	The ICRT looptree	141
5.2.4	The ICRT fennec	142
5.2.5	Left, Front, Right (see Figure 5.2)	143
5.2.6	The contour path, and the ICRT snake	144
5.3	Recalls from Section 2.4	145
5.4	Basic properties on \mathcal{L}.	146
5.4.1	Proof of the compactness of \mathcal{L} .	146
5.4.2	Proof of Proposition 5.2.2	148
5.5	Preliminary results on left, front, right	149
5.5.1	Properties of $\curvearrowright, \triangleleft$.	149
5.5.2	Generic properties on $p_{\curvearrowright}, p_{\nabla}, p_{\curvearrowleft}$	149
5.5.3	Extension of \hat{p}_l to \mathcal{L}	150
5.5.4	Continuous extension of p_{\curvearrowright} to \mathcal{L}	151

5.5.5	Some preliminary results on $p_{l,\curvearrowright}$	152
5.6	Construction and Hölder continuity of the contour path.	154
5.6.1	Construction of the contour path \mathcal{C}_l on \mathcal{L}_l	154
5.6.2	Construction of the contour path \mathcal{C}	155
5.6.3	Proof of Theorem 5.2.8	157
5.6.4	Hölder continuity of \mathcal{C} : Proof of Theorem 5.2.10	157
5.7	Fractal dimensions of the looptree	158
5.7.1	Definitions of the fractal dimensions	158
5.7.2	Upper bound on the Minkowski dimensions	159
5.7.3	The rebranching principle.	159
5.7.4	Lower bound on the Hausdorff and Packing dimensions	161
5.8	Proof of Theorem 5.2.4	163
5.9	Proof and extension of Theorem 5.2.5	165
5.10	Appendix A: Proof of Lemma 5.2.7, and measurability	167
5.11	Appendix B: Hölder continuity of a Gaussian free field.	170
5.12	Appendix C: Supremum of empirical process.	171

5.1 Introduction

5.1.1 Motivations and overview of the results

We construct the looptree, "fennec" (a gaussian free field on the looptree), and snake of the ICRT. We also compute the fractal dimensions (Minkowski, Packing, Hausdorff) of the looptree, and the Hölder exponents of the fennec and snake. In a companion paper [26] we show that these objects are the scaling limits of the looptrees, fennecs, and snakes of " \mathcal{D} -trees" (uniform rooted trees with fixed degree sequence \mathcal{D}).

Informally, a looptree, introduced by Curien, Kortchemski [44], is constructed by replacing each vertex of a tree by a cycle of size proportional to its degree, while keeping the tree structure. Then the fennec (for field+snake) is a Gaussian free field on the looptree. Finally, the snake is the real process obtained by turning around the looptree clockwise, and reading the value of the fennec. Those definitions are made formal in Section 5.2, using a new theory of plane \mathbb{R} -tree.

Those objects are mainly motivated by scaling limits of maps with fixed face degree sequence. Indeed, the bijections of Bouttier, Di Francesco, Guitter [33] and Janson-Stefánsson [55] yield together a bijection between those maps and \mathcal{D} -trees with a discrete fennec. It is now well known that the convergence of the snakes implies the tightness of the maps. However, although Le Gall [62] developed a general approach to prove the universality, this question remains open in general. In the stable case, this key problem is under active investigation by Curien, Miermont, Riera [45]. We refer to Marzouk [68, 67, 69] for elaborate discussions on the subject.

Let us already mention, that independently, Marzouk [69] also proves scaling limits of looptrees and snakes of \mathcal{D} -trees toward objects constructed from processes with exchangeable increments. We strongly believe that both approaches are useful to study the limits, which thus coincide, from the point of views of processes with exchangeable increments, and stick-breaking constructions.

In this direction, our theory of plane \mathbb{R} -tree builds a bridge between those two points of view. Indeed, we extend many discrete notions to \mathbb{R} -trees, to construct several real processes directly from the trees. As a result, those processes can now be studied from stick-breaking constructions. This completes the pioneer ideas of Le Gall (see e.g. [61]), which allow to construct and study trees from real process. By analogy with this theory, we construct the height process and Lukasiewicz walk of ICRT, and we will study them in a forthcoming paper [25].

5.1.2 Theory of plane \mathbb{R} -tree: an overview

A discrete plane tree is a rooted tree with an ordering of the children of each vertex. Aldous [11] extended this notion to binary continuum trees, by using signs $-/+$, to construct the height process. The height process is then constructed by Duquesne [48] for general order.

However an order is not enough to construct several objects, and notably the Lukasiewicz walk. This walk was essential in the work of Le Gall, Le Jan [64, 63], and Le Gall, Duquesne [50, 49] to study Lévy trees. For \mathcal{P} -trees and ICRT, Aldous, Miermont, Pitman [14, 15], developed a similar theory based on processes with exchangeable increments, but those processes are far less understood.

Still their depth first walk of \mathcal{P} -trees led us to two intuitions: First, \mathcal{P} -trees are infinite. This is confirmed in Chapter 3 since \mathcal{P} -trees are the limits of \mathcal{D} -trees in the "condensation case" (when the largest degrees have the same order as the total degree, see Theorem 3.4.1 (a), 3.4.2 (a)). Then, this process can be seen as a Lukasiewicz walk. This interpretation of this "half-discrete-half continuum" tree led us to the theory of plane \mathbb{R} -tree below:

First we define a notion of angle, which can be seen as numbering the children of each vertex. With those angles, we can rewrite the discrete definitions of the looptree and fennec for the ICRT. Then, we define a notion of left and right, which we use to define a contour path in the looptree: morally start at the root, then turn around each cycle clockwise, and stop after a complete circuit. Finally we construct many processes, by composing this path with some functions on the looptree.

5.1.3 Stick breaking and the chaining method

Our approach is based on the stick-breaking construction of the ICRT from Chapter 2, which is adapted from Aldous, Camarri, Pitman [16, 40]. Stick-breaking constructions, first introduced by Aldous [10], generate a \mathbb{R} -tree and are separated in two steps:

- the line \mathbb{R}^+ is first cut into the segments ("sticks" or "branches") $[0, Y_1], (Y_1, Y_2], (Y_2, Y_3] \dots$
- then for every $i \in \mathbb{N}$ the segment $(Y_i, Y_{i+1}]$ is glued at position $Z_i \leq Y_i$.

In Chapter 2 we study the compactness and dimensions of ICRT. We now use similar methods, which can be split into simple topological/logical arguments, and many uses of the chaining method.

This method has found many applications in concentration theory (see e.g. Talagrand [79], or [32] Chapter 13), and to study random metric spaces, and notably stick-breaking constructions (see e.g. Aldous [10] ; Amini, Devroye, Griffiths, Olver [17] ; Curien, Haas [43] ; Sénizergues [77]).

Let us explain its main principle: The goal is to estimate the max of a function f on a space S . To this end, consider a sequence of increasing subspaces¹ $(S_i)_{i \in \mathbb{N}}$ of S "approximating" S , and for every $i \in \mathbb{N}$, a projection $p_i : S_{i+1} \rightarrow S_i$. The main idea is that if $(S_i)_{i \in \mathbb{N}}$ are properly chosen:

$$\max_{x \in S} f(x) \leq \sum_{i \in \mathbb{N}} \max_{x \in S_{i+1}} (f(x) - f(p_i(x))).$$

As a result, one can decompose a complex estimate into many simpler ones. Moreover, when $(S_i)_{i \in \mathbb{N}}$ are well chosen, it tends to give optimal bounds².

In most of our proofs S is the looptree, and $(S_i)_{i \in \mathbb{N}}$ are the sub-looptrees obtained after gluing a certain number of branches. Then f can be the distance between a vertex and a fixed set to prove compactness, or compute fractal dimensions. f can also be some partial sums to prove uniform convergence, or continuity. In Chapter 3, we also used it to prove the tightness of \mathcal{D} -trees.

In this paper we often re-decompose for every $i \in \mathbb{N}$, the estimate of $\max_{x \in S_{i+1}} f(x) - f(p_i(x))$: On the one hand, we estimate the maximum number of branches separating $x \in S_{i+1}$ from S_i . On the other hand, we estimate how f vary on each branches. Finally we multiply the worst cases.

¹For more complex algorithms one may want to consider general $(X_i)_{i \in \mathbb{N}}$ and a family of functions f_i on X_i .

²Reverse bounds tends to be much harder to prove. See e.g. Section 2.6.3 for ICRT in the non compact case.

Plan of the paper: In Section 5.2, we define our objects and state our main results. We also deduce the Hölder continuity of the fennec and snake from our other results. In Section 5.3, we recall several technical results from Chapter 2. In Section 5.4, we prove the compactness of the looptree. In Section 5.5, we study the notions of left and right. In Section 5.6, we construct the contour path and prove its Hölder continuity. In Section 5.7, we compute the fractal dimensions of the looptree. In Section 5.8, we study the Gaussian free field on the ICRT. In Section 5.9, we prove that the fennec is well defined, and extend its definition to other fields. Both Sections 5.8, 5.9 can be read right after Section 5.3.

Acknowledgment I am grateful to Cyril Marzouk for the interesting discussions we got at CIRM.

5.2 Model and main results.

5.2.1 Basic notions on \mathbb{R} -trees and plane \mathbb{R} -trees.

A *Polish space* is a separable, complete, metric space. A \mathbb{R} -tree is a geodesic, loopless, Polish space (see Le Gall [61]). A *rooted \mathbb{R} -tree* is a \mathbb{R} -tree with a distinguished vertex.

For every \mathbb{R} -tree \mathcal{T} , and $x, y \in \mathcal{T}$, let $\llbracket x, y \rrbracket$ denote the geodesic path between x and y . The *closest common ancestor* of $x, y \in \mathcal{T}$ is the vertex $x \wedge y \in \llbracket \rho, x \rrbracket \cap \llbracket \rho, y \rrbracket$ which maximizes $d(\rho, z)$. For every $x \in \mathcal{T}$, the *degree* $\deg(x)$ of x in \mathcal{T} is the number of connected components of $\mathcal{T} \setminus \{x\}$.

An *angle function* on a rooted \mathbb{R} -tree (\mathcal{T}, d, ρ) is a function $u : \mathcal{T}^2 \rightarrow [0, 1]$ such that:

- For all $x \in \mathcal{T}$, $u_{x,\rho} = u_{x,x} = 0$.
- For all $x \in \mathcal{T}$, $y, z \in \mathcal{T} \setminus \{x\}$, $u_{x,y} = u_{x,z}$ iff y and z are connected in $\mathcal{T} \setminus \{x\}$.

A *plane \mathbb{R} -tree* is a rooted \mathbb{R} -tree equipped with an angle function.

To avoid measurability issues, we further assume in the definition that a plane \mathbb{R} -tree is *balanced*: for every $x, y \in \mathcal{T}$ if $\deg(x) = 2$ then $u_{x,y} \in \{0, 1/2\}$.

5.2.2 ICRT, and plane ICRT

We first introduce a generic stick breaking construction. It takes for input two sequences in \mathbb{R}^+ called cuts $\mathbf{y} = (y_i)_{i \in \mathbb{N}}$ and glue points $\mathbf{z} = (z_i)_{i \in \mathbb{N}}$, which satisfy

$$\forall i < j, \quad y_i < y_j \quad ; \quad y_i \longrightarrow \infty \quad ; \quad \forall i \in \mathbb{N}, \quad z_i \leq y_i,$$

and creates an \mathbb{R} -tree by recursively "gluing" segment $(y_i, y_{i+1}]$ on position z_i , or rigorously, by constructing a consistent sequence of distances $(d_n)_{n \in \mathbb{N}}$ on $([0, y_n])_{n \in \mathbb{N}}$.

Algorithm 5.1. Generic stick-breaking construction of \mathbb{R} -tree.

- Let d_0 be the trivial metric on $[0, 0]$.
- For each $i \geq 0$ define the metric d_{i+1} on $[0, y_{i+1}]$ such that for each $x \leq y$:

$$d_{i+1}(x, y) := \begin{cases} d_i(x, y) & \text{if } x, y \in [0, y_i] \\ d_i(x, z_i) + |y - y_i| & \text{if } x \in [0, y_i], y \in (y_i, y_{i+1}] \\ |x - y| & \text{if } x, y \in (y_i, y_{i+1}] \end{cases}$$

where by convention $y_0 := 0$ and $z_0 := 0$.

- Let d be the unique metric on \mathbb{R}^+ which agrees with d_i on $[0, y_i]$ for each $i \in \mathbb{N}$.
- Let $\text{SB}(\mathbf{y}, \mathbf{z})$ be the completion of (\mathbb{R}^+, d) .

Let Ω be the space of sequences $\{\theta_i\}_{i \in \mathbb{N}}$ in \mathbb{R}^+ with $\sum_{i=0}^{\infty} \theta_i^2 = 1$ and $\theta_1 \geq \theta_2 \geq \dots$. The ICRT of parameter $\Theta \in \Omega$ is the random \mathbb{R} -tree constructed via the following algorithm.

Algorithm 5.2. Construction of Θ -ICRT (Algorithm 2.1)

- Let $\mathbf{X} = (X_i)_{i \in \mathbb{N}}$ a family of independent exponential random variables of parameter $(\theta_i)_{i \in \mathbb{N}}$.
- Let μ be the measure on \mathbb{R}^+ defined by $\mu = \theta_0^2 dx + \sum_{i=1}^{\infty} \delta_{X_i} \theta_i$.
- For each $l \in \mathbb{R}^+$ let μ_l be the restriction of μ to $[0, l]$, and let $\mathbf{p}_l := \mu_l / \mu[0, l]$.
- Let $\mathbf{Y} = (Y_i)_{i \in \mathbb{N}}$ be a Poisson point process on \mathbb{R}^+ of rate $\mu[0, l] dl$.
- Let $\mathbf{Z} = (Z_i)_{i \in \mathbb{N}}$ be a family of independent random variables with laws $(\mathbf{p}_{Y_i})_{i \in \mathbb{N}}$.
- The Θ -ICRT is defined as $(\mathcal{T}, d_{\mathcal{T}}) = \text{SB}(\mathbf{Y}, \mathbf{Z})$ (see Algorithm 5.1).

Remarks. • When $\theta_0 = 1$, the ICRT is the Brownian CRT.

- When $\theta_0 = 0$ and $\sum_{i=1}^{\infty} \theta_i < \infty$, \mathcal{T} "is" a \mathcal{P} -tree with a modified distance (see Section 3.5.2).
- When Θ is random and corresponds to the jumps and brownian part of a Lévy bridge, the ICRT "is" a Lévy tree. (Equal in GP distribution, by unicity of the limit of \mathcal{D} -trees, see Section 3.8.1.)
- Morally, $(X_i)_{i \in \mathbb{N}}$ and $(\theta_i)_{i \in \mathbb{N}}$ corresponds to the vertices of highest degrees with their degrees. On the other hand, θ_0 corresponds to vertices with small degrees.

We root ICRT at 0. Recall that to define a plane \mathbb{R} -tree, we need an angle function U . This is equivalent to define for each $x \in \mathcal{T}$, and for each connected component C of $\mathcal{T} \setminus \{x\}$ with $0 \notin C$, the value of $U_{x,y} \in [0, 1]$ for a unique $y \in C$. So, the following algorithm a.s. does well define an angle function on the ICRT. Also, the last line below insure that a.s. $(\mathcal{T}, d_{\mathcal{T}}, 0, U)$ is balanced.

Algorithm 5.3. Construction of the uniform angle function U on the ICRT.

- Let $(U_{X,i})_{i \in \mathbb{N}}, (U_{Z,i})_{i \in \mathbb{N}}$ be independent uniform random variables in $[0, 1]$.
- Let U be the unique angle function on \mathcal{T} such that:
 - For every $i \in \mathbb{N}$, $U(X_i, Y_{\inf\{a \in \mathbb{N}: Y_a > X_i\}}) = U_{X,i}$
 - For every $i \in \mathbb{N}$, $U(Z_i, Y_{i+1}) = U_{Z,i}$.
 - For every $x \in \mathbb{R}^+ \setminus \bigcup_{i \in \mathbb{N}} \{X_i, Y_i\}$, we have $U(x, Y_{\inf\{a \in \mathbb{N}: Y_a > x\}}) = 1/2$.

5.2.3 The ICRT looptree

To extend the discrete setting, we want to replace each vertex by a loop. So we define \mathcal{L} as $\mathcal{T} \times [0, 1]$ with a proper pseudo-distance $d_{\mathcal{L}}$ corresponding to the cycles.

To define $d_{\mathcal{L}}$ we need the sizes of the cycles, which in the discrete correspond to the degrees. For ICRT, although the degrees are infinite, the only vertices with high degrees are $(X_i)_{i \in \mathbb{N}}$ and their degrees are morally proportional to $(\theta_i)_{i \in \mathbb{N}}$.

Thus we may define \mathcal{L} by concatenating some cycles of perimeter $(\theta_i)_{i \in \mathbb{N}}$. Actually, this is not enough, since we forget the vertices of small degrees. Morally their degrees corresponds to $\theta_0 dl$. Then by concatenating the corresponding cycles we get segments of length $\theta_0^2 / 4dl$. (The factor $1/4$ is the mean distance between two points in a cycle of perimeter 1.)

So, we formally define the ICRT looptree as follows: Let \mathbf{c} be the distance in the torus $[0, 1]$. For every $x, y \in \mathcal{T}$, $u \in [0, 1]$, let $U_{x,y,u} = U_{x,y}$ if $x \neq y$ and let $U_{x,y,u} = u$ otherwise. We define a pseudo-distance $d_{\mathcal{L}}$ on $\mathcal{T} \times [0, 1]$ such that for all $(x, u), (y, v) \in \mathcal{T} \times [0, 1]$ (see Figure 5.1),

$$d_{\mathcal{L}}((x, u), (y, v)) := \frac{\theta_0^2}{4} d_{\mathcal{T}}(x, y) + \sum_{i \in \mathbb{N}} \theta_i \mathbf{c}(U_{X_i, x, u}, U_{X_i, y, v}). \quad (5.1)$$

Finally let $(\mathcal{L}, d_{\mathcal{L}})$ be the completion of the pseudo-metric space $(\mathcal{T} \times [0, 1], d_{\mathcal{L}})$.

Theorem 5.2.1. *Almost surely $d_{\mathcal{L}}$ is finite on $\mathcal{T} \times [0, 1]$, and $(\mathcal{L}, d_{\mathcal{L}})$ is compact.*

Remark. When $\theta_1 = 1$, \mathcal{L} is a cycle of size 1. When $\theta_0 = 1$, \mathcal{L} is the Brownian CRT.

Sadly, several of our notions are initially defined in $\mathbb{R}^+ \times [0, 1]$ or in $\mathcal{T} \times [0, 1]$, but not in \mathcal{L} . So, to avoid any issues, we extend the above notions to \mathcal{L} :

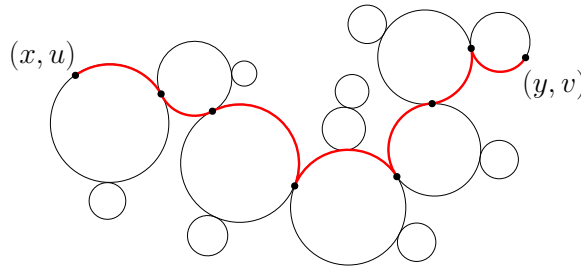


Figure 5.1: A continuum looptree \mathcal{L} . The geodesic between $(x, u), (y, v) \in \mathcal{L}$ is red. The distance between (x, u) and (y, v) is the sum of the length of all potential cords that lies in this geodesic. This is a simplified picture, since \mathcal{L} usually has infinitely many cycles which can have size null.

Proposition 5.2.2. *Almost surely the following assertions hold:*

- (a) \mathcal{L} is the completion of $(\mathbb{R}^+ \times [0, 1], d_{\mathcal{L}})$.
- (b) If \mathcal{T} is compact, then $\mathcal{L} = \mathcal{T} \times [0, 1]$, and $p_{\mathcal{T}, \mathcal{L}} : (x, u) \in \mathcal{L} \rightarrow x \in \mathcal{T}$ is continuous.
- (c) For every $i \in \mathbb{N}$ with $\theta_i > 0$, $\alpha \in \mathcal{T} \times [0, 1] \mapsto U_{X_i, \alpha}$ extends to a continuous function from $(\mathcal{L}, d_{\mathcal{L}})$ to $([0, 1], \mathfrak{c})$.
- (d) For every $\alpha, \beta \in \mathcal{L}$, writing when $\theta_0 > 0$ $d_{\mathcal{T}} : (x, u), (y, v) \in \mathcal{L} \mapsto d_{\mathcal{T}}(x, y)$,

$$d_{\mathcal{L}}(\alpha, \beta) = \frac{\theta_0^2}{4} d_{\mathcal{T}}(\alpha, \beta) + \sum_{i \in \mathbb{N}} \theta_i \mathfrak{c}(U_{X_i, \alpha}, U_{X_i, \beta}).$$

Remark. If \mathcal{T} is not compact, then \mathcal{L} morally contains the "ends of the infinite branches" of \mathcal{T} (see Section 5.5.4). Those points are dense in \mathcal{L} (cf Lemma 2.6.6), so $p_{\mathcal{T}, \mathcal{L}}$ is nowhere continuous.

We now give the fractal dimensions of \mathcal{L} , whose definitions are recalled in Section 5.7.1. Let,

$$\underline{\mathfrak{d}} := 1 + \liminf_{l \rightarrow \infty} \frac{\log \mathbb{E}[\mu[0, l]]}{\log l} \quad ; \quad \bar{\mathfrak{d}} := 1 + \limsup_{l \rightarrow \infty} \frac{\log \mathbb{E}[\mu[0, l]]}{\log l}. \quad (5.2)$$

Remark. $l \mapsto l\mathbb{E}[\mu[0, l]]$ is an analog of the Laplace exponent ψ for Lévy processes (see [15] and Lemma 2.6.1). Also, by Lemma 5.3.1 (b), $1 \leq \underline{\mathfrak{d}} \leq \bar{\mathfrak{d}} \leq 2$.

Theorem 5.2.3. *Almost surely the upper Minkowski dimension, and Packing dimension of \mathcal{L} are $\bar{\mathfrak{d}}$. Almost surely the lower Minkowski dimension, and Hausdorff dimension of \mathcal{L} are $\underline{\mathfrak{d}}$*

5.2.4 The ICRT fennec

To mimic the discrete setting, we want to define the ICRT fennec \mathfrak{F} as a Gaussian free field (a random function) on \mathcal{L} . We construct it explicitly by mimicking our construction of the looptree.

First, to deal with vertices of small degree we construct a Gaussian free field on \mathcal{T} . To this end, we adapt Algorithm 5.1: Let $\mathfrak{B} : \mathbb{R}^+ \rightarrow \mathbb{R}$ be a Brownian motion. Then define inductively \mathfrak{G} on \mathbb{R}^+ such that for every $i \in \mathbb{N}$ and $Y_i < x \leq Y_{i+1}$,

$$\mathfrak{G}(x) := \mathfrak{G}(Z_i) + \mathfrak{B}(x) - \mathfrak{B}(Y_i). \quad (5.3)$$

To construct the fennec \mathfrak{F} , we need to show that \mathfrak{G} extends to a continuous function on \mathcal{T} if $\theta_0 > 0$. We actually prove the much stronger result:

Theorem 5.2.4. *Almost surely \mathfrak{G} extends to a continuous function on \mathcal{T} if*

$$\int_0^\infty \frac{dl}{l \sqrt{\mathbb{E}[\mu[0, l]]}} < \infty. \quad (5.4)$$

Remark. By Theorem 2.3.1, \mathcal{T} is a.s. compact iff $\int^\infty dl / (\mathbb{E}[\mu[0, l]]) < \infty$, so \mathfrak{G} is well defined on most compact ICRT. Also, we believe one can adapt Section 2.6.3 to show that (5.4) is necessary. Moreover, in [49] Chapter 4.5, Duquesne and Le Gall prove the equivalence for Lévy trees.

We now adapt (5.1). Beforehand, recall that if $\theta_0 > 0$, by Proposition 5.2.2 (b), that $\mathcal{L} = \mathcal{T} \times [0, 1]$ and that $p_{\mathcal{L}, \mathcal{T}} : (x, u) \in \mathcal{L} \rightarrow x$ is continuous. Then the fennec is

$$\mathfrak{F} : \alpha \in \mathcal{L} \mapsto \frac{\theta_0}{\sqrt{6}} \mathfrak{G} \circ p_{\mathcal{L}, \mathcal{T}}(\alpha) + \sum_{i=1}^{\infty} \sqrt{\theta_i} \mathfrak{B}_i(U_{X_i, \alpha}). \quad (5.5)$$

Theorem 5.2.5. *Almost surely the sum in (5.5) converges uniformly on \mathcal{L} , so \mathfrak{F} is continuous.*

We will actually prove a stronger statement where the functions $(\mathfrak{B}_i)_{i \in \mathbb{N}}$ are replaced by more general random functions under a moment condition for their maximums (see Section 5.9). We believe that this extension may have applications to study more complex fields on \mathcal{D} -trees.

A direct corollary of Theorem 5.2.5 is that \mathfrak{F} is indeed a Gaussian free field on \mathcal{L} :

Proposition 5.2.6. *Almost surely, conditionally on $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, U$, for every $\alpha, \beta \in \mathcal{L}$, $\mathfrak{F}(\alpha) - \mathfrak{F}(\beta)$ is Gaussian with variance*

$$d'_{\mathcal{L}}(\alpha, \beta) := \frac{\theta_0^2}{6} d_{\mathcal{T}}(\alpha, \beta) + \sum_{i \in \mathbb{N}} \theta_i |U_{X_i, \alpha} - U_{X_i, \beta}| (1 - |U_{X_i, \alpha} - U_{X_i, \beta}|).$$

Remark. $\frac{1}{2} d_{\mathcal{L}} \leq d'_{\mathcal{L}} \leq d_{\mathcal{L}}$ so $d'_{\mathcal{L}}$ is equivalent to $d_{\mathcal{L}}$.

5.2.5 Left, Front, Right (see Figure 5.2)

In this section $(\mathcal{T}, d, \rho, u)$ denote an arbitrary plane \mathbb{R} -tree. For every $x \in \mathcal{T}$, $(y, w) \in \mathcal{T} \times [0, 1]$ let $u_{x, y, w} := u_{x, y}$ if $x \neq y$ and let $u_{x, y, w} := w$ otherwise. For all $\alpha = (x, v), \beta = (y, w) \in \mathcal{T} \times [0, 1]$, we say that α is at the left of β (or β is at the right of α) and write $\alpha \curvearrowright \beta$ if $u_{x \wedge y, x, v} < u_{x \wedge y, y, w}$. We say that β is in front of α and write $\alpha \triangleleft \beta$ if $x \in \llbracket 0, y \rrbracket$ and $u_{x, y, w} = v$.

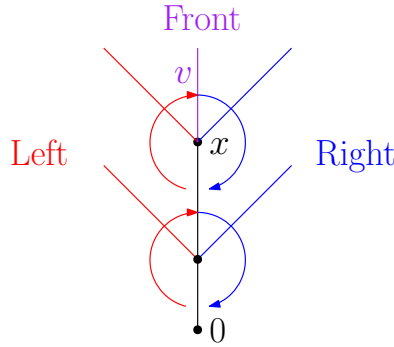


Figure 5.2: A vertex $(x, v) \in \mathcal{T} \times [0, 1]$ with its left (red), front (purple), right (blue) are represented.

Lemma 5.2.7. *Let \prec be the binary relation defined on $\mathcal{T} \times [0, 1]$ such that for all $\alpha, \beta \in \mathcal{T} \times [0, 1]$:*

$$\alpha \prec \beta \iff (\alpha \curvearrowright \beta) \text{ or } (\alpha \triangleleft \beta).$$

Then \prec is a total order relation on $\mathcal{T} \times [0, 1]$ and is called the contour order.

Proof. See Appendix 5.10. □

Finally let ν be a σ -finite borel measure on \mathcal{T} . Let $\nu_{\mathcal{L}} := \mu \times \mathbf{1}_{l \in [0, 1]} dl$. The mass on the left, front, right of $\alpha \in \mathcal{T} \times [0, 1]$ are denoted by, (see Appendix 5.10 for definiteness)

$$\nu_{\curvearrowright}(\alpha) := \nu_{\mathcal{L}}\{\beta : \beta \curvearrowright \alpha\} \quad ; \quad \nu_{\triangleleft}(\alpha) := \nu_{\mathcal{L}}\{\beta : \alpha \triangleleft \beta\} \quad ; \quad \nu_{\curvearrowleft}(\alpha) := \nu_{\mathcal{L}}\{\beta : \beta \curvearrowleft \alpha\}.$$

5.2.6 The contour path, and the ICRT snake

Recall that for $l \in \mathbb{R}^+$, μ_l is the restriction of μ to $[0, l]$, and $\mathbf{p}_l = \mu_l / \mu[0, l]$. Also by Proposition 2.3.2 a.s. $(\mathbf{p}_l)_{l \in \mathbb{R}^+}$ converges weakly toward a probability measure \mathbf{p} . We prove that a.s. \mathbf{p}_\wedge extends to a function continuous at $\mathcal{L} \setminus (\mathbb{R}^+ \times [0, 1])$ (see Section 5.5.3), and that \mathbf{p}_\wedge has an "inverse": Below $\sim_{\mathcal{L}}$ denote the equivalent relation on \mathcal{L} such that for every $\alpha, \beta \in \mathcal{L}$, $\alpha \sim_{\mathcal{L}} \beta$ iff $d_{\mathcal{L}}(\alpha, \beta) = 0$.

Theorem 5.2.8. *Almost surely there exists a continuous function $\mathfrak{C} : [0, 1] \mapsto \mathcal{L}$ such that for every $\alpha \in \mathcal{L}$, $\mathfrak{C}(\mathbf{p}_\wedge(\alpha)) \sim_{\mathcal{L}} \alpha$. We call \mathfrak{C} the contour path on \mathcal{L} (see Figure 5.6).*

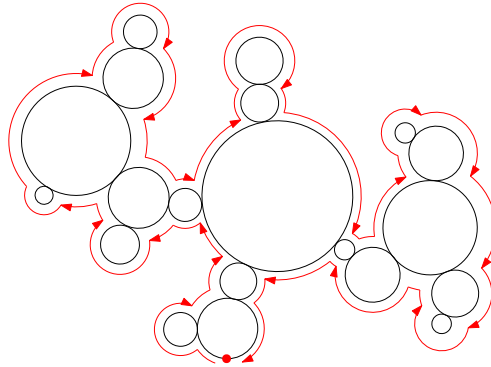


Figure 5.3: A continuum looptree \mathcal{L} , with its contour path $\mathfrak{C} : [0, 1] \mapsto \mathcal{L}$ in red. The path starts at the root $(0, 0)$ then "turn around" each cycle clockwise. It is continuous, "surjective", not "injective".

The density of $\{\mathbf{p}_\wedge(\alpha), \alpha \in \mathcal{L}\}$ (see Lemma 5.5.3) implies that \mathfrak{C} is unique up to $\sim_{\mathcal{L}}$. Moreover it also implies with the existence of \mathfrak{C} that:

Proposition 5.2.9. *Almost surely for every continuous function $F : \mathcal{L} \mapsto \mathbb{R}$, $F \circ \mathfrak{C}$ is the unique continuous function f such that for every $\alpha \in \mathcal{L}$, $f(\mathbf{p}_\wedge(\alpha)) = F(\alpha)$.*

In particular, we define the ICRT snake as $\mathcal{Z} := \mathfrak{F} \circ \mathfrak{C}$.

Remark. Similarly, if \mathcal{T} is a.s. compact, we can define the height process of the ICRT as follows. First by Proposition 5.2.2 (b), $\mathcal{L} = \mathcal{T} \times [0, 1]$, and $p_{\mathcal{T}, \mathcal{L}} : (x, u) \in \mathcal{L} \rightarrow x$ is continuous. Then we define the contour path on \mathcal{T} as $\mathfrak{C}_{\mathcal{T}} := p_{\mathcal{T}, \mathcal{L}} \circ \mathfrak{C}$. This path is a.s. continuous by continuity of \mathfrak{C} . Finally, the height process is

$$\mathfrak{H} : x \in [0, 1] \mapsto d_{\mathcal{T}}(0, \mathfrak{C}_{\mathcal{T}}(x)).$$

Also by first defining

$$\mathfrak{L} : (x, u) \in \mathcal{L} \mapsto \frac{\theta_0^2}{2} d_{\mathcal{T}}(0, x) + \sum_{i \in \mathbb{N}: X_i \in [0, x]} \theta_i (1 - U_{X_i, x, u}),$$

we may define the Lukasiewicz walk as $\mathfrak{X} := \mathfrak{L} \circ \mathfrak{C}$. We will study \mathfrak{H} and \mathfrak{X} in [25].

We now consider the Hölder continuity. Our first point is that for any continuous function F , the Hölder continuity of $F \circ \mathfrak{C}$ can be deduced from the Hölder continuity of F and \mathfrak{C} . Moreover:

Theorem 5.2.10. *Recall (5.2). A.s. \mathfrak{C} is Hölder continuous with any exponent smaller than $1/\bar{\delta}$.*

Remark. The bound is optimal. Indeed, if \mathfrak{C} is α -Hölder continuous then \mathcal{L} has Minkowski upper dimension at most $1/\alpha$ (see Lemma 5.7.3).

Thus, since by Proposition 5.2.6, a.s. \mathfrak{F} is a Gaussian free field on $(\mathcal{L}, d'_{\mathcal{L}})$, which have by Theorem 5.2.3 finite upper Minkowski dimension, we deduce (see Lemma 5.11.1):

Theorem 5.2.11. *Almost surely \mathfrak{F} is Hölder continuous with any exponent smaller than $1/2$, and \mathcal{Z} is Hölder continuous with any exponent smaller than $1/2\bar{\delta}$.*

5.3 Recalls from Section 2.4

We prove the following results in Section 2.4: (Actually in Chapter 2 we focus on the case where $\theta_0 > 0$ or $\sum_{i=1}^{\infty} \theta_i = \infty$, but the proof is exactly the same in the complementary case. Also (d) is slightly modified from Lemma 2.4.6 using Lemma 2.4.5.)

Lemma 5.3.1. *The following assertions hold a.s.:*

- a) *The map $l \mapsto \mathbb{E}[\mu[0, l]]$ is concave.*
- b) *As $l \rightarrow \infty$, $\mu[0, l] \sim \mathbb{E}[\mu[0, l]] = \theta_0^2 l + o(l)$.*
- c) *For every l large enough, $\#([0, l] \cap \{Y_i\}_{i \in \mathbb{N}}) \leq 2\mu[0, l]l \leq 2l^2$.*
- d) *For every i large enough, $\mu[Y_{i-1}, Y_i] \leq \log(Y_i)^2/Y_i$.*

We now adapt Lemma 2.4.7. First let us define a metric, which morally counts the number of branches between two points. To do so, we adapt Algorithm 5.1: First let $d_{\mathfrak{N},0}$ be the trivial metric on $[0, 0]$. Then for every $i \geq 0$, define $d_{\mathfrak{N},i+1}$ as the metric on $[0, Y_{i+1}]$ such that for every $x < y$,

$$d_{\mathfrak{N},i+1}(x, y) := \begin{cases} d_{\mathfrak{N},i}(x, y) & \text{if } x, y \in [0, Y_i] \\ d_{\mathfrak{N},i}(x, Z_i) + 1 & \text{if } x \in [0, y_i], y \in (Y_i, Y_{i+1}] \\ 1 & \text{if } x, y \in (Y_i, Y_{i+1}] \end{cases}$$

Finally let $d_{\mathfrak{N}}$ be the unique metric on \mathbb{R}^+ which agrees with $d_{\mathfrak{N},i}$ on $[0, Y_i]$ for every $i \in \mathbb{N}$.

Lemma 5.3.2. *Almost surely, for every n large enough, for every $x \leq 2^{n+1}$, $d_{\mathfrak{N}}(x, [0, 2^n]) \leq 4n$.*

Proof. We adapt the proof of Lemma 2.4.7. Let $\mathcal{F} := \sigma(\mu, \{Y_i\}_{i \in \mathbb{N}})$. Let z be a random variable \mathcal{F} -measurable in $[0, 2^{n+1}]$. Let us follow the geodesic path between z and $[0, 2^n]$ (see Figure 5.4): Let $z_0 := z$, then for every $i \geq 0$, let $k_i := \max\{k \in \mathbb{N} : Y_k < z_i\}$ and let $y_i := Y_{k_i}$, and let $z_{i+1} := Z_{k_i}$. Finally note that $d_{\mathcal{N}}(x, [0, 2^n]) = T := \inf\{t \in \mathbb{N}, z_t \leq 2^n\}$.

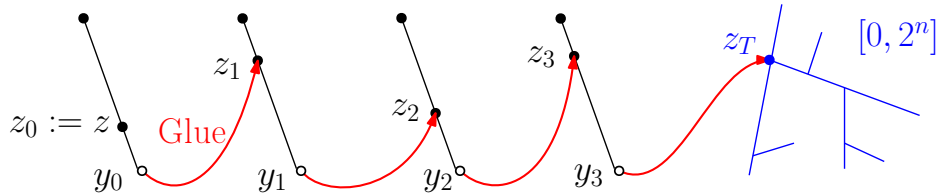


Figure 5.4: A typical construction of $(z_i, y_i)_{i \in \mathbb{N}}$. Each segment represents a branch.

Then for every $i \geq 0$ let $\mathcal{F}_i := \sigma(\mathcal{F}, z_0, \dots, z_i)$. Recall that conditionally on \mathcal{F} , $(Z_i)_{i \in \mathbb{N}}$ are independent so $((y_i, z_i), \mathcal{F}_i)_{i \geq 0}$ is a Markov chain. Also T is a stopping time for this Markov chain. Moreover by definition of $(Z_i)_{i \in \mathbb{N}}$, if $z_i \geq 2^n$,

$$\mathbb{P}(z_{i+1} \leq 2^n | \mathcal{F}_i) = \frac{\mu[0, 2^n]}{\mu[0, y_i]} \geq \frac{\mu[0, 2^n]}{\mu[0, z]} \geq \frac{\mu[0, 2^n]}{\mu[0, 2^{n+1}]} \geq 1/3,$$

where the last inequality holds a.s. for every n large enough by Lemma 5.3.1 (a) (b). Thus,

$$\mathbb{P}(T \geq 4n) = \mathbb{P}(z_{4n-1} > 2^n) \leq (2/3)^{4n-1}.$$

Therefore by an union bound,

$$\mathbb{P}(\exists 1 \leq i \leq 2^{2n+3} : Y_i \leq 2^n, d_{\mathfrak{N}}(Y_i, [0, 2^{n+1}]) \geq 4n | \mathcal{F}) = O(1/n^2).$$

Hence, by the Borel–Cantelli Lemma, and by Lemma 5.3.1 (c), a.s. for every n large enough,

$$\max\{d_{\mathfrak{N}}(Y_i, [0, 2^n]) : i \in \mathbb{N}, Y_i \leq 2^{n+1}\} \leq 4n - 1.$$

Finally simply note that $d_{\mathfrak{N}}(\cdot, [0, 2^n])$ is constant among branches. □

5.4 Basic properties on \mathcal{L} .

5.4.1 Proof of the compactness of \mathcal{L} .

We first show an upper bound on $d_{\mathcal{L}}$. Then we introduce the projections on $\mathcal{L}_l := [0, l] \times [0, 1]$. Then we show that for every $l \in \mathbb{R}^+$, $(\mathcal{L}_l, d_{\mathcal{L}})$ is a.s. compact. Then we upper bound for $n \in \mathbb{N}$, $d_H(\mathcal{L}_{2^n}, \mathcal{L}_{2^{n+1}})$. Finally, we show that \mathcal{L} is a.s. compact with the chaining method.

Lemma 5.4.1. *A.s. for every $(x, u), (y, v) \in \mathcal{T} \times [0, 1]$, we have $d_{\mathcal{L}}((x, u), (y, v)) \leq \mu[x, y]$. Also for every $x \in \mathcal{T}$, $(y, v) \in \mathcal{T} \times [0, 1]$, we have $d_{\mathcal{L}}((x, U_{x,y,v}), (y, v)) \leq \mu[x, y]$.*

Proof. We focus on the second assertion as the first can be proved similarly. Let $u = U_{x,y,v}$. Recall that by definition, writing \mathfrak{c} for the distance on the torus $[0, 1]$,

$$d_{\mathcal{L}}((x, u), (y, v)) = \frac{\theta_0^2}{4} d_{\mathcal{T}}(x, y) + \sum_{i \in \mathbb{N}} \theta_i \mathfrak{c}(U_{X_i, x, u}, U_{X_i, y, v}).$$

Then note that for every $i \in \mathbb{N}$ such that $X_i \notin [x, y]$, x, y are connected in $\mathcal{T} \setminus \{X_i\}$, hence $U_{X_i, x, u} = U_{X_i, x} = U_{X_i, y} = U_{X_i, y, v}$. Also $U_{x, x, u} = u = U_{x, y, v}$. Therefore,

$$d_{\mathcal{L}}((x, u), (y, v)) \leq \theta_0^2 d_{\mathcal{T}}(x, y) + \sum_{i \in \mathbb{N}: X_i \in [x, y]} \theta_i = \mu[x, y]. \quad \square$$

Remark. Note that Lemma 5.3.1 (c) (d) implies $\mu[Y_i, Y_{i+1}] = O(i^{-1/2+o(1)})$. Also by Lemma 5.4.1 the diameter of the branches of the looptree is bounded by $(\mu[Y_i, Y_{i+1}])_{i \in \mathbb{N}}$. Thus we are close from the compact setting of Curien and Haas [43] and of Sénizergues [77].

We now introduce the projections on \mathcal{T} and on \mathcal{L} (see Figure 5.5). For all $x \in \mathcal{T}$, and $l \in \mathbb{R}^+$, let $\bar{\rho}_l$ be the *projection of x on $\mathcal{T}_l := [0, l]$* , that is the unique $z \in \mathcal{T}_l$ which minimizes $d_{\mathcal{T}}(x, z)$.

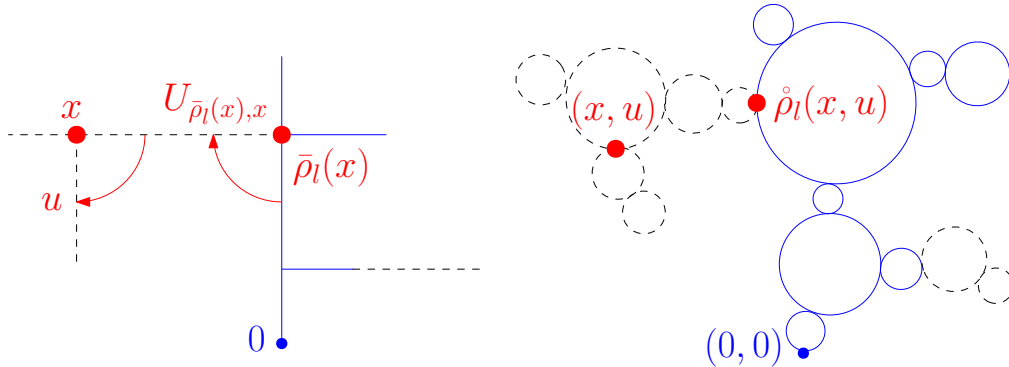


Figure 5.5: Two representations of $\hat{\rho}_l$, from the tree on the left, and from the looptree on the right. \mathcal{T}_l and \mathcal{L}_l are in blue. $\mathcal{T} \setminus \mathcal{T}_l$ and $\mathcal{L} \setminus \mathcal{L}_l$ are in black dashed. Here $U_{\bar{\rho}_l(x), x, u} = U_{\bar{\rho}_l(x), x}$ since $\bar{\rho}_l(x) \neq x$.

Lemma 5.4.2. *For every $(x, u) \in \mathcal{T} \times [0, 1]$ and $l \in \mathbb{R}^+$, $\beta \in \mathcal{L}_l \mapsto d_{\mathcal{L}}((x, u), \beta)$ reach its minimum at $\hat{\rho}_l(x, u) := (\bar{\rho}_l(x), U_{\bar{\rho}_l(x), x, u})$. We call $\hat{\rho}_l(x, u)$ the *projection of x on \mathcal{L}_l* .*

Proof. Note that, since \mathcal{T} is a \mathbb{R} -tree, for every $y \in [0, l]$, for every $z \in [\bar{\rho}_l(x), x]$, $U_{z, y} = 0$. Thus, for every $(y, u) \in \mathcal{L}_l$, since $d_{\mathcal{T}}(x, y) \geq d_{\mathcal{T}}(x, \bar{\rho}_l(x))$,

$$d_{\mathcal{L}}((x, u), (y, v)) \geq \frac{\theta_0^2}{4} d_{\mathcal{T}}(x, \bar{\rho}_l(x)) + \sum_{i: X_i \in [\bar{\rho}_l(x), x]} \theta_i \mathfrak{c}(U_{X_i, x, u} = d_{\mathcal{L}}(x, \hat{\rho}_l(x)), \quad (5.6)$$

where the last inequality is obtained by a small modification of the proof of Lemma 5.4.1. \square

Lemma 5.4.3. Fix $l > 0$. Almost surely, $d_{\mathcal{L}}$ is finite on $\mathcal{L}_l = [0, l] \times [0, 1]$, and (\mathcal{L}_l, d_l) is compact.

Proof. The finiteness of $d_{\mathcal{L}}$ directly follows from Lemma 5.4.1 since for every $(x, u), (y, v) \in \mathcal{L}_l$, $\llbracket x, y \rrbracket \subset [0, l]$, and since almost surely μ is locally finite. So let us focus on the compactness. Let $((x_n, u_n))_{n \in \mathbb{N}} \in (\mathcal{L}_l, d_{\mathcal{L}})^{\mathbb{N}}$. Up to extraction, we may assume that $(x_n)_{n \in \mathbb{N}}$ converges for $d_{\mathcal{T}}$ toward $x \in [0, l]$. Then note that, for every $X \neq x$ and for every n large enough, x_n and x are connected in $\mathcal{T} \setminus \{X\}$, so $U_{X, x, u} = U_{X, x_n, v}$. Furthermore, up to extraction, we may also assume that $U_{x, x_n, u_n} \rightarrow u \in [0, 1]$ as $n \rightarrow \infty$. Therefore, by dominated convergence, a.s.

$$d_{\mathcal{L}}((x_n, u_n), (x, u)) = \frac{\theta_0^2}{4} d_{\mathcal{T}}(x_n, x) + \sum_{i \in \mathbb{N}: X_i \leq l} \theta_i \mathfrak{c}(U_{X_i, x, u}, U_{X_i, y, v}) \rightarrow 0. \quad \square$$

We now show that $(\mathcal{L}_{2^n})_{n \in \mathbb{N}}$ is a Cauchy sequence for d_H .

Lemma 5.4.4. The following assertions hold a.s. for every n large enough:

- a) $\max_{x \in [0, 2^{n+1}]} \min_{y \in [0, 2^n]} \mu \llbracket y, x \rrbracket \leq 4n^3/2^n$.
- b) $d_H(\mathcal{L}_{2^n}, \mathcal{L}_{2^{n+1}}) \leq 4n^3/2^n$.

Proof. Let $(x, u) \in \mathcal{L}_{2^{n+1}}$. By lemma 5.4.1,

$$d_{\mathcal{L}}((x, u), \bar{\rho}_{2^n}(x, u)) \leq \mu \llbracket \bar{\rho}_{2^n}(x), x \rrbracket.$$

Also, a.s. for every n large enough, by Lemma 5.3.2 $\llbracket \bar{\rho}_{2^n}(x), x \rrbracket$ morally consists in the union of a part of at most $4n$ branches of the form $]Y_i, Y_{i+1}]$, which have by Lemma 5.3.1 mass at most $n^2/2^n$. Hence $\mu \llbracket \bar{\rho}_{2^n}(x), x \rrbracket \leq 4n^3/2^n$. Since $(x, u) \in \mathcal{L}_{2^{n+1}}$ is arbitrary, this concludes the proof. \square

We will reuse the following intermediate result in [25]:

Lemma 5.4.5. Almost surely $x \mapsto \mu \llbracket 0, x \rrbracket$ is bounded on \mathcal{T} .

Proof. First by Lemma 5.4.4 (a) almost surely, for every n large enough,

$$\max_{x \in [0, 2^{n+1}]} \mu \llbracket 0, x \rrbracket \leq \max_{x \in [0, 2^n]} \mu \llbracket 0, x \rrbracket + 4n^3/2^n.$$

Thus since $\sum 4n^3/2^n < \infty$, $x \mapsto \mu \llbracket 0, x \rrbracket$ is bounded on \mathbb{R}^+ . Moreover for every $x \in \mathcal{T} \setminus \mathbb{R}^+$, by monotone convergence, and since $\mu(\mathcal{T} \setminus \mathbb{R}^+) = 0$, as $y \rightarrow x$, $y \in \llbracket 0, x \rrbracket$,

$$\mu \llbracket 0, y \rrbracket \longrightarrow \mu \llbracket 0, x \rrbracket = \mu \llbracket 0, x \rrbracket. \quad (5.7) \quad \square$$

Proof of Theorem 5.2.1. By Lemma 5.4.5 and (5.7), for every $x \in \mathcal{T}$, $\min_{y \in \mathbb{R}^+} \mu \llbracket y, x \rrbracket = 0$. Hence, by Lemma 5.4.1, $d_H(\mathcal{T} \times [0, 1], \mathbb{R}^+ \times [0, 1]) = 0$. So, since \mathcal{L} is the completion of $(\mathcal{T} \times [0, 1], d_{\mathcal{L}})$, we have $d_H(\mathcal{L}, \mathbb{R}^+ \times [0, 1]) = 0$.

Then by Lemma 5.4.4, $(\mathcal{L}_{2^n})_{n \in \mathbb{N}}$ is a Cauchy sequence for the Hausdorff pseudo-distance. Also, by Lemma 5.4.3, for every $n \in \mathbb{N}$, \mathcal{L}_{2^n} is compact. Hence, since \mathcal{L} is complete, $(\mathcal{L}_{2^n})_{n \in \mathbb{N}}$ converges toward a compact subset of \mathcal{L} . This subset contains $\mathbb{R}^+ \times [0, 1]$. And since $d_H(\mathcal{L}, \mathbb{R}^+ \times [0, 1]) = 0$, it is at distance 0 of \mathcal{L} . Therefore, \mathcal{L} is compact. \square

In passing note that we also prove the following result, which we will reuse in [26].

Proposition 5.4.6. Almost surely as $l \rightarrow \infty$, $d_H(\mathcal{L}_l, \mathcal{L}) \rightarrow 0$.

5.4.2 Proof of Proposition 5.2.2

Toward (a), simply recall that $d_H(\mathcal{L}, \mathbb{R}^+ \times [0, 1]) = 0$.

Toward (b), to show that $\mathcal{T} \times [0, 1] = \mathcal{L}$ it is enough to prove that $(\mathcal{T} \times [0, 1], d_{\mathcal{L}})$ is compact. Let $((x_n, u_n))_{n \in \mathbb{N}}$ be a sequence in $\mathcal{T} \times [0, 1]$. Since \mathcal{T} is compact, we may assume up to extraction that $x_n \rightarrow x \in \mathcal{T}$, and $U_{x, x_n, u_n} \rightarrow u \in [0, 1]$. Let us prove $(x_n, u_n) \rightarrow (x, u)$. There is two cases:

Either $x \notin \mathbb{R}^+$. In this case, for every $l \in \mathbb{R}^+$, for every n large enough, x and x_n are connected in $\mathcal{T} \setminus [0, l]$. And it follows, by Lemma 5.4.2, that for every n large enough,

$$d_{\mathcal{L}}((x_n, u_n), (x, u)) \leq 2d_H([0, l] \times [0, 1], \mathcal{L}),$$

which converges a.s. to 0 as $l \rightarrow \infty$, by Proposition 5.4.6.

Or $x \in \mathbb{R}^+$. In this case, for every $l \geq x$, as $n \rightarrow \infty$, by Lemma 5.4.2, $\hat{\rho}_l(x_n, u_n) \rightarrow (x, u)$. Thus a.s.

$$\limsup_{n \rightarrow \infty} d_{\mathcal{L}}((x_n, u_n), (x, u)) \leq d_H(\mathcal{L}_l, \mathcal{L}) \xrightarrow{l \rightarrow \infty} 0.$$

Then we show that $p_{\mathcal{T}, \mathcal{L}} : (x, u) \in \mathcal{T} \times [0, 1] \mapsto x$ is continuous on $\mathcal{T} \times [0, 1]$. We argue by contradiction. Assume that there exists $(x_n, u_n) \in \mathcal{L}^{\mathbb{N}}$ such that as $n \rightarrow \infty$, $(x_n, u_n) \rightarrow (x, u) \in \mathcal{L}$ but $x_n \not\rightarrow x$. Since \mathcal{T} is compact we may assume up to extraction that $x_n \rightarrow y \in \mathcal{T}$.

Since $x \neq y$, there exists $a < b \in \mathbb{R}^+ \setminus \{x, y\}$ such that $[a, b] \subset]x, y[$. Then by density of μ on \mathbb{R}^+ (see proof of Theorem 2.3.1), either $\theta_0 > 0$, so

$$\liminf_{n \in \mathbb{N}} d_{\mathcal{L}}((x_n, u_n), (x, u)) \geq \frac{\theta_0^2}{4} d_{\mathcal{T}}(y, x) > 0,$$

or there exists $i \in \mathbb{N}$ with $\theta_i > 0$ and $X_i \in [a, b] \setminus \{x \wedge y\}$, so for every $n \in \mathbb{N}$ large enough,

$$d_{\mathcal{L}}((x_n, u_n), (x, u)) \geq \theta_i U_{X_i} > 0,$$

where the last inequality holds almost surely. Both cases contradict $d_{\mathcal{L}}((x_n, u_n), (x, u)) \rightarrow 0$.

Toward (c), note that by definition of $d_{\mathcal{L}}$, for every $i \in \mathbb{N}$ with $\theta_i > 0$, $\alpha \mapsto U_{X_i, \alpha}$ is $1/\theta_i$ Lipschitz from $\mathcal{T} \times [0, 1]$ to $([0, 1], \mathfrak{c})$, and so extends by continuity on \mathcal{L} .

Toward (d), by (c) and definition of $d_{\mathcal{L}}$ on $\mathcal{T} \times [0, 1]$, it is enough to show that as $n \rightarrow \infty$, $\alpha, \beta \in \mathcal{L}^2 \mapsto \sum_{i=1}^n \theta_i \mathfrak{c}(U_{X_i, \alpha}, U_{X_i, \beta})$ converges uniformly as $n \rightarrow \infty$. To this end, first note that by dominated convergence, for every $l \in \mathbb{R}^+$, as $n, m \rightarrow \infty$, a.s. (see proof of Lemma 5.4.1)

$$\Delta_{n, m, l} := \max_{\alpha, \beta \in \mathcal{L}_l} \sum_{i=n}^m \theta_i \mathfrak{c}(U_{X_i, \alpha}, U_{X_i, \beta}) \leq \sum_{i=n}^{\infty} \theta_i \mathbf{1}_{X_i \in [0, l]} \rightarrow 0.$$

Then, by definition of $d_{\mathcal{L}}$, and the triangular inequality, for every $l \in \mathbb{R}^+$, as $n, m \rightarrow \infty$, a.s.

$$\begin{aligned} \Delta_{n, m} &:= \max_{\alpha, \beta \in \mathcal{T} \times [0, 1]} \sum_{i=n}^m \theta_i \mathfrak{c}(U_{X_i, \alpha}, U_{X_i, \beta}) \\ &\leq 2d_H(\mathcal{L}_l, \mathcal{L}) + \max_{\alpha, \beta \in \mathcal{T} \times [0, 1]} \sum_{i=n}^m \theta_i \mathfrak{c}(U_{X_i, \hat{\rho}_l(\alpha)}, U_{X_i, \hat{\rho}_l(\beta)}) \\ &\leq 2d_H(\mathcal{L}_l, \mathcal{L}) + \Delta_{n, m, l} \\ &\rightarrow 2d_H(\mathcal{L}_l, \mathcal{L}). \end{aligned} \tag{5.8}$$

By Proposition 5.4.6 a.s. $d_H(\mathcal{L}_l, \mathcal{L}) \rightarrow 0$. Thus by (5.8) as $n, m \rightarrow \infty$ a.s. $\Delta_{n, m} \rightarrow 0$. The maximum in (5.8) is then directly extended to \mathcal{L} by (c). This yields the desired uniform convergence.

5.5 Preliminary results on left, front, right

In this section is we prove the technical results necessary to construct and study \mathfrak{C} in Section 5.6. In Section 5.5.1, 5.5.2 we prove generic results on \curvearrowright , \triangleleft , and on $\mathfrak{p}_{\curvearrowright}$, \mathfrak{p}_{∇} , $\mathfrak{p}_{\curvearrowleft}$. In Section 5.5.3, 5.5.4 we extend respectively $(\hat{\rho}_l)_{l>0}$ and $\mathfrak{p}_{\curvearrowright}$ to \mathcal{L} . In Section 5.5.5 we estimate $\mathfrak{p}_{l,\curvearrowright}$.

5.5.1 Properties of \curvearrowright , \triangleleft .

Recall Section 5.2.5. The following lemma is proved in Appendix 5.10:

Lemma 5.5.1. \triangleleft is a partial order. \curvearrowright is a strict partial order. \triangleleft raises \prec (see Figure 5.6): for every $\beta \curvearrowright \gamma \in \mathcal{T} \times [0, 1]$, for every $\alpha \in \mathcal{T} \times [0, 1]$, $\gamma \triangleleft \alpha \Rightarrow \beta \curvearrowright \alpha$ and $\beta \triangleleft \alpha \Rightarrow \alpha \curvearrowright \gamma$.

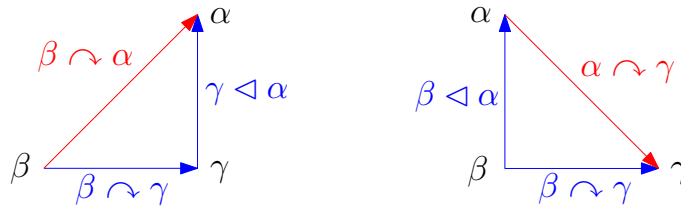


Figure 5.6: The cases of Lemma 5.5.1. The assumptions are blue. The conclusions are red.

Lemma 5.5.2. Recall $\bar{\rho}, \hat{\rho}$ from section 5.4.1. For every $l > 0$, $\alpha = (x, a), \beta = (y, b) \in \mathcal{T} \times [0, 1]$:

- (a) If $\hat{\rho}_l(\alpha) \curvearrowright \hat{\rho}_l(\beta)$ then $\alpha \curvearrowright \beta$.
- (b) If $\hat{\rho}_l(\alpha) \triangleleft \hat{\rho}_l(\beta)$ and $\hat{\rho}_l(\alpha) \neq \hat{\rho}_l(\beta)$ then $\alpha = \hat{\rho}_l(\alpha) \triangleleft \beta$.

Proof. First by definition of $\hat{\rho}$, $\hat{\rho}_l(\alpha) = (\bar{\rho}_l(x), U_{\bar{\rho}_l(x), \alpha}) \triangleleft \alpha$ and similarly $\hat{\rho}_l(\beta) \triangleleft \beta$. (a) follows since \triangleleft raises \curvearrowright . Toward (b), $\hat{\rho}_l(\alpha) \triangleleft \alpha$, and $\hat{\rho}_l(\alpha) \triangleleft \hat{\rho}_l(\beta)$, so either:

- $x = \bar{\rho}_l(x)$.
- $\bar{\rho}_l(y) = \bar{\rho}_l(x)$ and then by $\hat{\rho}_l(\alpha) \triangleleft \hat{\rho}_l(\beta)$, we have $\hat{\rho}_l(\alpha) = \hat{\rho}_l(\beta)$.
- $x, \bar{\rho}_l(y)$ are connected in $\mathcal{T} \setminus \{\bar{\rho}_l(x)\}$ which is absurd by definition of $\bar{\rho}$.

Hence, $x = \bar{\rho}_l(x)$. So since $\hat{\rho}_l(\alpha) \triangleleft \alpha$, by definition of \triangleleft , $\alpha = \hat{\rho}_l(\alpha)$. Finally $\alpha \triangleleft \hat{\rho}_l(\beta) \triangleleft \beta$. \square

5.5.2 Generic properties on $\mathfrak{p}_{\curvearrowright}$, \mathfrak{p}_{∇} , $\mathfrak{p}_{\curvearrowleft}$

In this section $(\mathcal{T}, d, \rho, u)$ denotes a plane \mathbb{R} -tree, and \mathfrak{p} denotes a probability Borel measure on \mathcal{T} .

Lemma 5.5.3. The following assertions hold:

- (a) For every $\alpha \in \mathcal{T} \times [0, 1]$, $\mathfrak{p}_{\curvearrowright}(\alpha) + \mathfrak{p}_{\nabla}(\alpha) + \mathfrak{p}_{\curvearrowleft}(\alpha) = 1$.
- (b) $\mathfrak{p}_{\curvearrowright}$ is increasing for \prec .
- (c) If A is a random variable with law $\mathfrak{p} \times \mathbf{1}_{l \in [0, 1]} dl$, then almost surely $\mathfrak{p}_{\nabla}(A) = 0$.
- (d) With the same notations, $\mathfrak{p}_{\curvearrowright}(A)$ is uniform in $[0, 1]$.
- (e) $\{\mathfrak{p}_{\curvearrowright}(\alpha)\}_{\alpha \in \mathcal{T} \times [0, 1]}$ is dense in $[0, 1]$.

Proof. Toward (a), note that for every $\alpha \in \mathcal{T} \times [0, 1]$, we have the following partition,

$$\mathcal{T} \times [0, 1] = \{\beta : \beta \neq \alpha, \beta \triangleleft \alpha\} \cup \{\beta : \alpha \curvearrowright \beta\} \cup \{\beta : \alpha \triangleleft \beta\} \cup \{\beta : \beta \curvearrowright \alpha\}. \quad (5.9)$$

Then by Fubini's theorem since for every $y \in \mathcal{T}$, $\#\{v \in [0, 1], (y, v) \triangleleft \alpha\} \leq 1$, we have $\mu_{\mathcal{L}}\{\beta : \beta \triangleleft \alpha\} = 0$. (a) follows from (5.9).

Toward (b), by Lemma 5.5.1, for every $\alpha \prec \beta \in \mathcal{L}$, $\{\gamma : \gamma \curvearrowright \alpha\} \subset \{\gamma : \gamma \curvearrowright \beta\}$.

Toward (c), let A, A' be independent random variables with law $\mu_{\mathcal{L}}$. Since \mathcal{T} is separable, and balanced, the events $A \curvearrowright A', A \triangleleft A'$ are measurable (see Lemma 5.10.4). The measurability of the other random variables in the proof is then due to Fubini's Theorem.

Then note that a.s. $\mathbf{p}_{\nabla}(A) = \mathbb{P}[A \triangleleft A' | A]$, and by (a) a.s. $\mathbf{p}\{\alpha : \alpha \triangleleft A'\} = \mathbb{P}[A \triangleleft A' | A'] = 0$. Hence, by Fubini's theorem, a.s. $\mathbf{p}_{\nabla}(A) = 0$.

Toward (d), let $(A_i)_{i \in \mathbb{N}}$ be a family of independent random variables with law $\mathbf{p} \times \mathbf{1}_{l \in [0,1]} dl$. By the weak law of large number a.s.

$$\frac{1}{n} \sum_{i=1}^n \mathbb{P}(A_i \curvearrowright A_1 | A_1) \xrightarrow{n \rightarrow \infty} \mathbf{p}_{\curvearrowright}(A_1). \quad (5.10)$$

Furthermore, by (b) for every $i \neq j \in \mathbb{N}$,

$$\mathbb{P}(A_i \triangleleft A_j) = \mathbb{E}[\mathbb{P}(A_i \triangleleft A_j | A_i)] = \mathbb{E}[\mathbf{p}_{\nabla}(A_i)] = 0.$$

Hence, since \prec is a total order, a.s. for every $i, j \in \mathbb{N}$ either $A_i \curvearrowright A_j$ or $A_j \curvearrowright A_i$. Moreover, the law of $(A_i)_{i \in \mathbb{N}}$ is invariant by permutations. Hence for every $n \in \mathbb{N}$, $\#\{i \leq n : A_i \curvearrowright A_1\}$ is uniform in $\{0, 1, 2, \dots, n-1\}$. Finally (d) follows from (5.10). (e) directly follows from (d). \square

5.5.3 Extension of $\hat{\rho}_l$ to \mathcal{L}

Recall notations $\bar{\rho}, \hat{\rho}$ from section 5.4.1. Although for all $l > 0$, $\hat{\rho}_l$ is continuous on $\mathcal{T} \times [0, 1]$, we do not want to extend it by continuity. Indeed, a continuous extension is only defined up to $\sim_{\mathcal{L}}$, while $\curvearrowright, \triangleleft$ are not. So we prove instead that $\hat{\rho}_l$ is locally constant around each vertex of $\mathcal{L} \setminus (\mathcal{T} \times [0, 1])$, and extend it naturally. First, we prove that $\hat{\rho}$ is piecewise constant:

Lemma 5.5.4. *For every $l > 0$, for every C connected component of $\mathcal{T} \setminus [0, l]$, $\bar{\rho}$ is constant on C , and $\hat{\rho}$ is constant on $C \times [0, 1]$.*

Proof. The first assertion is classic, so we leave it to the reader as an exercise. Let $x, y \in C$ and let $a, b \in [0, 1]$. $\bar{\rho}(x) = \bar{\rho}(y)$. Since $x, y \neq \bar{\rho}(x)$, $U_{\bar{\rho}_l(x), x, a} = U_{\bar{\rho}_l(x), x}$ and $U_{\bar{\rho}_l(x), y, b} = U_{\bar{\rho}_l(x), y}$. Also since x, y are connected in $\mathcal{T} \setminus \{\bar{\rho}(x)\}$, we have $U_{\bar{\rho}_l(x), x} = U_{\bar{\rho}_l(x), y}$. So $U_{\bar{\rho}_l(x), x, a} = U_{\bar{\rho}_l(y), y, b}$. Therefore, $\hat{\rho}(x, a) = \hat{\rho}(y, b)$. \square

For every $\alpha \in \mathcal{L}$, and $\varepsilon > 0$ let $B(\alpha, \varepsilon)$ denote the open ball for $d_{\mathcal{L}}$ of center α and radius ε .

Lemma 5.5.5. *For every $l > 0$, for every $\alpha \in \mathcal{L}$, $\hat{\rho}_l$ is constant on $(\mathcal{T} \times [0, 1]) \cap B(\alpha, d_{\mathcal{L}}(\alpha, [0, l])/2)$.*

Proof. We argue by contradiction. Let $l > 0$, $\alpha \in \mathcal{L}$, $\beta = (y, b), \gamma = (z, c) \in \mathcal{T} \times [0, 1]$. Assume that $\hat{\rho}_l(\beta) \neq \hat{\rho}_l(\gamma)$, and $d_{\mathcal{L}}(\beta, \alpha) < d_{\mathcal{L}}(\alpha, [0, l])/2$, and $d_{\mathcal{L}}(\gamma, \alpha) < d_{\mathcal{L}}(\alpha, [0, l])/2$. Let $D = d_{\mathcal{L}}(\alpha, [0, l])$. By the triangular inequality,

$$d(\beta, [0, l]) > D/2 \quad ; \quad d(\gamma, [0, l]) > D/2 \quad ; \quad d_{\mathcal{L}}(\beta, \gamma) < D. \quad (5.11)$$

Moreover, since $\hat{\rho}_l(\beta) \neq \hat{\rho}_l(\gamma)$, by Lemma 5.5.4, y, z are disconnected in $\mathcal{T} \setminus [0, l]$. Hence, $\llbracket y, \bar{\rho}_l(y) \rrbracket \cap \llbracket \bar{\rho}_l(z), z \rrbracket = \emptyset$. As a result, writing

$$S_{\beta} := \frac{\theta_0^2}{4} d(y, \bar{\rho}_l(y)) + \sum_{i: X_i \in \llbracket y, \bar{\rho}_l(y) \rrbracket} \theta_i \mathbf{c}(U_{X_i, \beta}, U_{X_i, \gamma}), \quad (5.12)$$

and similarly for S_{γ} , we have $d_{\mathcal{L}}(\beta, \gamma) \geq S_{\beta} + S_{\gamma}$.

Also, for every $X_i \in \llbracket y, \bar{\rho}_l(y) \rrbracket$, $\bar{\rho}_l(y)$ and z are connected in $\mathcal{T} \setminus \{X_i\}$ so $U_{X_i, \gamma} = U_{X_i, \hat{\rho}_l(\beta)}$. Hence, by (5.12), (5.6), $S_{\beta} = d_{\mathcal{L}}(\beta, \hat{\rho}_l(\beta))$. and similarly for S_{γ} . Thus, since $d_{\mathcal{L}}(\beta, \gamma) \geq S_{\beta} + S_{\gamma}$,

$$d_{\mathcal{L}}(\beta, \gamma) \geq d_{\mathcal{L}}(\beta, \hat{\rho}_l(\beta)) + d_{\mathcal{L}}(\gamma, \hat{\rho}_l(\gamma)).$$

This contradicts (5.11). \square

Lemma 5.5.6. *For every $0 < r \leq s$, $\dot{\rho}_r \circ \dot{\rho}_s = \dot{\rho}_s \circ \dot{\rho}_r = \dot{\rho}_r$.*

Proof. First since $\dot{\rho}_r$ have value in $\mathcal{L}_r \subset \mathcal{L}_s$ and $\dot{\rho}_s$ is the identity on \mathcal{L}_s , we have $\dot{\rho}_s \circ \dot{\rho}_r = \dot{\rho}_r$. Then, since $\dot{\rho}_r$, and $\dot{\rho}_s$ are locally constant around each vertex of $\mathcal{L} \setminus \mathcal{T} \times [0, 1]$, and since $\mathcal{T} \times [0, 1]$ is dense, it is enough to show that $\dot{\rho}_r \circ \dot{\rho}_s = \dot{\rho}_r$ on $\mathcal{T} \times [0, 1]$.

Let $(x, a) \in \mathcal{T} \times [0, 1]$. By definition of $\bar{\rho}$, $\bar{\rho}_r \circ \bar{\rho}_s(x)$ is the vertex z in $\llbracket 0, \bar{\rho}_s \rrbracket \cap [0, r]$ which maximizes $d_{\mathcal{T}}(0, z)$. Moreover, since \mathcal{T} is a \mathbb{R} -tree, $\llbracket \bar{\rho}_s, x \rrbracket \subset \mathcal{T} \setminus [0, s]$. Hence, since $r \leq s$, $\bar{\rho}_r \circ \bar{\rho}_s(x)$ is the vertex z in $\llbracket 0, x \rrbracket \cap [0, r]$ which maximizes $d_{\mathcal{T}}(0, z)$, which is $\bar{\rho}_r(x)$. Finally by definition of $\dot{\rho}$, we have $\dot{\rho}_r \circ \dot{\rho}_s(x, a) \triangleleft \dot{\rho}_s(x, a) \triangleleft (x, a)$. So,

$$\dot{\rho}_r \circ \dot{\rho}_s(x, a) = (\bar{\rho}_r(x), U_{\bar{\rho}_r(x), x, a}) = \dot{\rho}_r(x, a). \quad \square$$

5.5.4 Continuous extension of $\mathfrak{p}_{\curvearrowright}$ to \mathcal{L}

Lemma 5.5.7. *Recall the notation $\mathfrak{p}_{\mathcal{L}} := \mathfrak{p} \times \mathbf{1}_{x \in \mu[0, 1]} dx$. The map*

$$\alpha \in \mathcal{L} \mapsto \mathfrak{p}_{\mathcal{L}} \left(\bigcup_{l > 0} \{ \beta \in \mathcal{T} \times [0, 1], \dot{\rho}_l(\beta) \curvearrowright \dot{\rho}_l(\alpha) \} \right) \quad (5.13)$$

is well defined and coincide with $\mathfrak{p}_{\curvearrowright}$ on $\mathcal{T} \times [0, 1]$.

Remark. One may see each $\alpha \in \mathcal{L} \setminus (\mathcal{T} \times [0, 1])$ as the "end" of the infinite branch (see Figure 5.7) $\{p_{\mathcal{T}, \mathcal{L}} \circ \dot{\rho}_l(\alpha), l \in \mathbb{R}^+\}$, and the relations of $\mathcal{T} \times [0, 1]$ may be extended to those missing points.

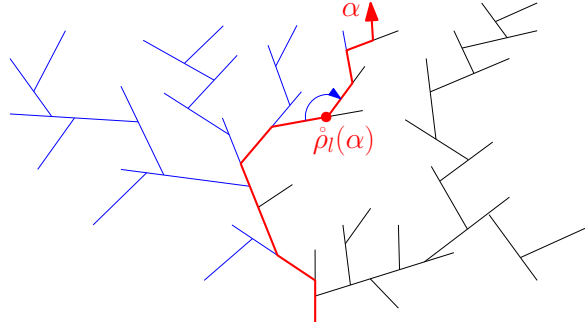


Figure 5.7: A simplified non compact ICRT, with a spinal representation of $\alpha \in \mathcal{L} \setminus (\mathcal{T} \times [0, 1])$. The infinite branch $\{p_{\mathcal{T}, \mathcal{L}} \circ \dot{\rho}_l(\alpha), l \in \mathbb{R}^+\}$ is in red. (5.13) estimates \mathfrak{p} of the blue part, on the left.

Proof. First by Lemmas 5.5.2 and 5.5.6, for every $l \leq l'$, $\beta \in \mathcal{T} \times [0, 1]$ such that $\dot{\rho}_l(\beta) \curvearrowright \dot{\rho}_l(\alpha)$ we have $\dot{\rho}_l \circ \dot{\rho}_{l'}(\beta) \curvearrowright \dot{\rho}_l \circ \dot{\rho}_{l'}(\alpha)$ so $\dot{\rho}_{l'}(\beta) \curvearrowright \dot{\rho}_{l'}(\alpha)$. Thus $(\{\beta \in \mathcal{T} \times [0, 1], \dot{\rho}_l(\beta) \curvearrowright \dot{\rho}_l(\alpha)\})_{l > 0}$ is increasing. Also, those sets are measurable (see Appendix 5.10). Hence (5.13) is well defined.

Next let $\alpha = (x, a), \beta = (y, b) \in \mathcal{T} \times [0, 1]$. If there exists $l > 0$ such that $\dot{\rho}_l(\beta) \curvearrowright \dot{\rho}_l(\alpha)$ then by Lemma 5.5.2, $\beta \curvearrowright \alpha$. Reciprocally, assume that $\beta \curvearrowright \alpha$ and either $x \in \mathbb{R}^+$ or $x \neq y$. Then $y \wedge x \in \mathbb{R}^+$, $\bar{\rho}_{x \wedge y}(x) = x$ and $\bar{\rho}_{x \wedge y}(y) = y$. Also, since $\beta \curvearrowright \alpha$, we have $U_{x \wedge y, y, b} < U_{x \wedge y, x, a}$. So,

$$\dot{\rho}_{x \wedge y}(y, b) = (x \wedge y, U_{x \wedge y, y, b}) \curvearrowright (x \wedge y, U_{x \wedge y, x, a}) = \dot{\rho}_{x \wedge y}(x, a). \quad (5.14)$$

As a result, for every $(x, a) \in \mathcal{T} \times [0, 1]$,

$$\{\beta, \beta \curvearrowright (x, a)\} \subset \bigcup_{l > 0} \{\beta, \dot{\rho}_l(\beta) \curvearrowright \dot{\rho}_l(x, a)\} \subset \{\beta, \beta \curvearrowright (x, a)\} \cup \{x\} \times [0, 1],$$

and the first inclusion is an equality when $x \in \mathbb{R}^+$. Finally for every $x \in \mathcal{T} \setminus \mathbb{R}^+$, $\mathfrak{p}(x) = 0$. (Indeed, if $\theta_0 = 0$, $\sum \theta_i < \infty$ then \mathfrak{p} have support $\{X_i\}_{i \in \mathbb{N}} \subset \mathbb{R}^+$. In the other case see Theorem 2.3.1.) So $\mathfrak{p}_{\curvearrowright}$ coincide with (5.13). \square

Lemma 5.5.8. Let $\mathbf{p}_\curvearrowright$ denote (5.13). $\mathbf{p}_\curvearrowright$ is continuous around each vertex of $\mathcal{L} \setminus \{\mathbb{R}^+ \times [0, 1]\}$.

Proof. Fix $\alpha \in \mathcal{L} \setminus \{\mathbb{R}^+ \times [0, 1]\}$. First, recall that the sets of (5.13) are increasing so, as $l \rightarrow \infty$,

$$\mathbf{p}_\mathcal{L}(\{\beta \in \mathcal{T} \times [0, 1], \dot{\rho}_l(\beta) \curvearrowright \dot{\rho}_l(\alpha)\}) \rightarrow \mathbf{p}_\curvearrowright(\alpha).$$

Furthermore, recall that by Lemma 5.5.5 $\dot{\rho}_l$ is locally constant, hence for every $l > 0$,

$$\liminf_{\beta \rightarrow \alpha} \mathbf{p}_\curvearrowright(\beta) \geq \mathbf{p}_\mathcal{L}(\{\beta \in \mathcal{T} \times [0, 1], \dot{\rho}_l(\beta) \curvearrowright \dot{\rho}_l(\alpha)\}) \xrightarrow{l \rightarrow \infty} \mathbf{p}_\curvearrowright(\alpha).$$

Let us show the other inequality. For every $\beta \in \mathcal{L}$ let

$$S_\beta := \bigcup_{l>0} \{\gamma \in \mathcal{T} \times [0, 1], \dot{\rho}_l(\gamma) \curvearrowright \dot{\rho}_l(\beta)\}.$$

Let $l > 0$. Let $\beta \in \mathcal{L}$ such that $\dot{\rho}_l(\beta) = \dot{\rho}_l(\alpha)$. Let $(z, c) \in S_\beta$. By Lemma 5.5.2 there exists $r > l$ such that $\dot{\rho}_r(z, c) \curvearrowright \dot{\rho}_r(\beta)$. Let $(y, b) = \dot{\rho}_r(\beta)$. By Lemma 5.5.6, $\dot{\rho}_l(y, b) = \dot{\rho}_l(\beta) = \dot{\rho}_l(\alpha)$.

If $y \wedge z > \bar{\rho}_l(y)$, then y and z are connected in $\mathcal{T} \setminus \{\bar{\rho}_l(y)\}$. So $\bar{\rho}_l(y) \in \llbracket 0, z \rrbracket$ and $U_{\bar{\rho}_l(y), y, b} = U_{\bar{\rho}_l(y), z, c}$. Also, $\dot{\rho}_l(y, b) = (\bar{\rho}_l(y), U_{\bar{\rho}_l(y), y, b})$. Hence, $\dot{\rho}_l(\alpha) = \dot{\rho}_l(y, b) \triangleleft (z, c)$.

If $y \wedge z \leq \bar{\rho}_l(y)$, then by Lemma 5.5.6,

$$\dot{\rho}_{y \wedge z}(y, b) = \dot{\rho}_{y \wedge z} \circ \dot{\rho}_l(y, b) = \dot{\rho}_{y \wedge z} \circ \dot{\rho}_l(\alpha) = \dot{\rho}_{y \wedge z}(\alpha).$$

Moreover, we have $\dot{\rho}_r(z, c) \curvearrowright (y, b)$ so by Lemma 5.5.3 $(z, c) \curvearrowright (y, b)$. Then by (5.14), we have $\dot{\rho}_{y \wedge z}(z, c) \curvearrowright \dot{\rho}_{y \wedge z}(y, b)$. Hence, $\dot{\rho}_{y \wedge z}(z, c) \curvearrowright \dot{\rho}_{y \wedge z}(\alpha)$. And $(z, c) \in S_\alpha$.

To sum up, for all $l > 0$, and $\beta \in \mathcal{L}$ such that $\dot{\rho}_l(\beta) = \dot{\rho}_l(\alpha)$,

$$S_\beta \subset S_\alpha \cup \{(z, c) \in \mathcal{T} \times [0, 1], \dot{\rho}_l(\alpha) \triangleleft (z, c)\}.$$

Let I_l denote the right most set above. By Lemma 5.5.5, to show that $\limsup_{\beta \rightarrow \alpha} \mathbf{p}_\curvearrowright(\beta) \leq \mathbf{p}_\curvearrowright(\alpha)$, it suffices to show that as $l \rightarrow \infty$, $\mathbf{p}_\mathcal{L}(I_l) \rightarrow 0$. First since \triangleleft is an order, and since for every $l \leq l'$, $\dot{\rho}_l(\alpha) \triangleleft \dot{\rho}_{l'}(\alpha)$, $(I_l)_{l>0}$ is decreasing. Then for every $z \in \mathcal{T}$, $\bar{\rho}_l \rightarrow z$ as $l \rightarrow \infty$. So for every $(y, b) \neq (z, c) \in \mathcal{T} \times [0, 1]$ such that $y \neq z$ or $z \in \mathbb{R}^+$, for every l large enough $\dot{\rho}_l(y, b) \neq \dot{\rho}_l(z, c)$. Thus, $\bigcap_{l>0} I_l$ is included in a set of the form $\{z\} \times [0, 1]$ with $z \in \mathcal{T} \setminus \mathbb{R}^+$, or of the form $\{(z, c)\}$. Finally recall that for every $z \in \mathcal{T} \setminus \mathbb{R}^+$, $\mathbf{p}(z) = 0$. Hence, $\mathbf{p}_\mathcal{L}(\bigcap_{l>0} I_l) = 0$. This shows the other desired inequality and thus concludes the proof. \square

5.5.5 Some preliminary results on $\mathbf{p}_{l, \curvearrowright}$

For every ν, σ -finite Borel measure on \mathcal{T} and $\alpha, \beta \in \mathcal{T} \times [0, 1]$, let $\nu_{\curvearrowright}(\alpha, \beta)$ denote $\nu_{\curvearrowright}(\beta) - \nu_{\curvearrowright}(\alpha)$.

Lemma 5.5.9. Almost surely for every $l > 0$, $\alpha, \beta \in \mathcal{L}_l = [0, l] \times [0, 1]$, $|\mu_{l, \curvearrowright}(\alpha, \beta)| \geq d_\mathcal{L}(\alpha, \beta)$. (See Figure 5.8.)

Proof. Let $\alpha = (x, a), \beta = (y, b) \in [0, l] \times [0, 1]$. By symmetry we may assume $(x, a) \prec (y, b)$. By Lemma 5.10.1 (a), (b), $\{\gamma \in \mathcal{L}_l, \gamma \curvearrowright \alpha\} \subset \{\gamma \in \mathcal{L}_l, \gamma \curvearrowright \beta\}$. Then writing

$$S := \{\gamma \in \mathcal{L}_l, \gamma \curvearrowright \beta\} \setminus \{\gamma \in \mathcal{L}_l, \gamma \curvearrowright \alpha\},$$

we have $\mu_{l, \curvearrowright}(\alpha, \beta) = \mu(S)$.

For every $z \in \llbracket x \wedge y, x \rrbracket$, $c \in (U_{z, x, a}, 1)$, we have $x \wedge z = z$, and $U_{z, x, a} < c = U_{z, z, c}$. So $(x, a) \curvearrowright (z, c)$. Also, z and x are connected in $\mathcal{T} \setminus \{x \wedge y\}$ so $z \wedge y = x \wedge y$, and $U_{z \wedge y, x, a} = U_{x \wedge y, x, a} < U_{x \wedge y, y, b}$. Hence, $(z, c) \curvearrowright (y, b)$. Therefore,

$$S_1 := \{(z, c) : z \in \llbracket x \wedge y, x \rrbracket, c \in (U_{z, x, a}, 1)\} \subset S.$$

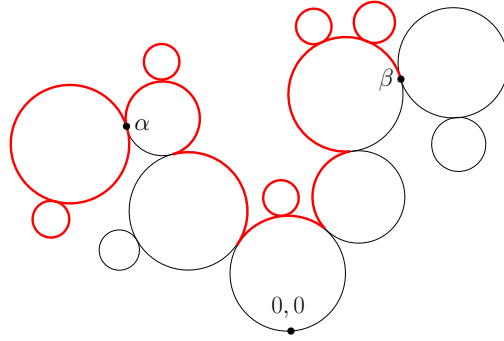


Figure 5.8: An informal proof of Lemma 5.5.9: $\alpha \prec \beta$. The set $S = \{\gamma, \gamma \curvearrowright \beta\} \setminus \{\gamma, \gamma \curvearrowright \alpha\}$ is red. (It can be obtained by turning clockwise from α to β .) Its total length is $\mu_{l,\curvearrowright}(\alpha, \beta)$. Note that this set contains a path between α and β . This path have length at least $d_{\mathcal{L}}(\alpha, \beta)$.

Similarly,

$$S_2 := \{(x \wedge y, c) : c \in (U_{x \wedge y, x, a}, U_{x \wedge y, y, b})\} \subset S,$$

and,

$$S_3 := \{(z, c) : z \in \llbracket x \wedge y, y \rrbracket, c \in (0, U_{z, y, b})\} \subset S.$$

Moreover S_1, S_2, S_3 are disjoint. And writing \mathfrak{c} for the distance on the torus $[0, 1]$, since for every $z \in \llbracket x \wedge y, x \rrbracket$, $U_{z, y, b} = 0$,

$$\mu(S_1) = \frac{\theta_0^2}{2} d_{\mathcal{T}}(x \wedge y, x) + \sum_{i: X_i \in \llbracket x \wedge y, x \rrbracket} (1 - U_{z, x, a}) \geq \frac{\theta_0^2}{4} d_{\mathcal{T}}(x \wedge y, x) + \sum_{i: X_i \in \llbracket x \wedge y, x \rrbracket} \theta_i \mathfrak{c}(U_{X_i, x, a}, U_{X_i, y, b}).$$

Similarly, $\mu(S_2) \geq \mu\{x \wedge y\} \mathfrak{c}(U_{x \wedge y, x, a}, U_{x \wedge y, y, b})$, and

$$\mu(S_3) \geq \frac{\theta_0^2}{4} d_{\mathcal{T}}(x \wedge y, y) + \sum_{i: X_i \in \llbracket x \wedge y, y \rrbracket} \theta_i \mathfrak{c}(U_{X_i, x, a}, U_{X_i, y, b}).$$

Finally, by sum, and since for every $i \in \mathbb{N}$, such that $X_i \notin \llbracket x, y \rrbracket$, $U_{X_i, x, a} = U_{X_i, y, b}$,

$$\mu(S) \geq \mu(S_1) + \mu(S_2) + \mu(S_3) \geq \frac{\theta_0^2}{4} d_{\mathcal{T}}(x, y) + \sum_{i=1}^{\infty} \theta_i \mathfrak{c}(U_{X_i, x, a}, U_{X_i, y, b}) = d_{\mathcal{L}}(\alpha, \beta). \quad \square$$

Lemma 5.5.10. *Almost surely for every $\alpha \in \mathcal{T} \times [0, 1]$, as $l \rightarrow \infty$, $\mathfrak{p}_{l,\curvearrowright}(\alpha) \rightarrow \mathfrak{p}_{\curvearrowright}(\alpha)$.*

Proof. When $\theta_0 = 0$ and $\sum_{i=1}^{\infty} \theta_i < \infty$, we have $\mu(\mathbb{R}^+) < \infty$. As a result, $\mu_l \rightarrow \mu$ in total variation. So $\mathfrak{p}_l \mathbf{1}_{x \in [0, 1]} dx \rightarrow \mathfrak{p} \mathbf{1}_{x \in [0, 1]} dx$ in total variation. The desired result follows.

When $\theta_0 \neq 0$ or $\sum_{i=1}^{\infty} \theta_i = \infty$, some extra care is needed since $(\mathfrak{p}_l)_{l \in \mathbb{R}^+}$ only converges weakly. For every $(x, a) \in \mathcal{T} \times [0, 1]$, let $S(x, a) := \{\beta \in \mathcal{T} \times [0, 1], \beta \curvearrowright (x, a)\}$. Let

$$S_1(x, a) := \{z \in \mathcal{T} \setminus \llbracket 0, x \rrbracket, U_{x \wedge z, z} < U_{x \wedge z, x, a}\}.$$

Note that $S_1(x, a)$ is a Borel set in $(\mathcal{T}, d_{\mathcal{T}})$ as an union of connected component (see Appendix 5.10). Moreover, by definition of \curvearrowright ,

$$S_1(x, a) \times [0, 1] \subset S(x, a) \subset (S_1(x, a) \cup \llbracket 0, x \rrbracket) \times [0, 1].$$

Hence, for every $x, a \in \mathcal{T} \times [0, 1]$

$$\mathfrak{p}_l(S_1(x, a)) \leq \mathfrak{p}_{l,\curvearrowright}(x) \leq \mathfrak{p}_l(S_1(x, a) \cup \llbracket 0, x \rrbracket), \quad (5.15)$$

and similarly,

$$\mathbf{p}(S_1(x, a)) \leq \mathbf{p}_{\curvearrowright}(x) \leq \mathbf{p}(S_1(x, a) \cup \llbracket 0, x \rrbracket), \quad (5.16)$$

Moreover, note that $S_1(x, a)$ is an open set of \mathcal{T} as a union of open connected components. Similarly, note that $S_1(x, a) \cup \llbracket 0, x \rrbracket$ is a closed set since its complementary is a union of open connected components. Furthermore, by Theorem 2.3.1, $\mathbf{p}(\mathbb{R}^+) = 0$, and \mathbf{p} has no atoms, so $\mathbf{p}(\llbracket 0, x \rrbracket) = 0$. As a result, by Portmanteau Theorem,

$$\mathbf{p}_l(S_1(x, a)) \rightarrow \mathbf{p}(S_1(x, a)) \quad \text{and} \quad \mathbf{p}_l(S_1(x, a) \cup \llbracket 0, x \rrbracket) \rightarrow \mathbf{p}(S_1(x, a) \cup \llbracket 0, x \rrbracket) = \mathbf{p}(S_1(x, a)). \quad (5.17)$$

The desired result follows by (5.15), (5.16), (5.17). \square

We now adapt Lemma 2.5.1 to estimate precisely the evolution of $\mathbf{p}_{l, \curvearrowright}$.

Lemma 5.5.11. *Let $\mathbf{U}_X := (U_{X,i})_{i \in \mathbb{N}}$. Almost surely $(\mu, \mathbf{Y}, \mathbf{U}_X)$ satisfies the following property. For all a large enough, conditionally on $\mathcal{F}_a := \sigma(\mu, \mathbf{Y}, (Z_i, U_{Z,i})_{i < a}, \mathbf{U}_X)$, for every $\alpha, \beta \in \mathcal{L}_{Y_a}$: If $\mu_{l, \curvearrowright}(\alpha, \beta) \geq (\log^6 Y_a)/Y_a$ then with probability at least $1 - 1/Y_a^5$, for every $b \geq a$,*

$$\left(1 - \frac{1}{\log Y_a}\right) \mathbf{p}_{Y_a, \curvearrowright}(\alpha, \beta) \leq p_{Y_b, \curvearrowright}(\alpha, \beta) \leq \left(1 + \frac{1}{\log Y_a}\right) \mathbf{p}_{Y_a, \curvearrowright}(\alpha, \beta).$$

Remark. With $\alpha = (0, 0)$, for all $a \in \mathbb{N}$, since $\mathbf{p}_{Y_a, \curvearrowright}(0, 0) = 0$, we have $\mathbf{p}_{Y_a, \curvearrowright}(\alpha, \beta) = \mathbf{p}_{Y_a, \curvearrowright}(\beta)$.

Proof. By Lemma 5.5.3 (b), $\alpha \prec \beta$. So, by Lemma 5.10.1 (a), (b), $\{\gamma : \gamma \curvearrowright \alpha\} \subset \{\gamma : \gamma \curvearrowright \beta\}$. Then by construction for every $i \in \mathbb{N}$, conditionally on \mathcal{F}_i , $\gamma_i := (Z_i, U_{Z,i})$ have law $\mathbf{p}_{Y_i} \times [0, 1]$. Moreover, for every $i \geq a$, writing $I_i :=]Y_i, Y_{i+1}] \times [0, 1]$, conditionally on \mathcal{F}_i a.s.

- With probability $\mathbf{p}_{Y_i, \curvearrowright}(x)$, $\gamma_i \curvearrowright \alpha$. Then by Lemma 5.10.2, for every $\delta \in I_i$, $\delta \curvearrowright \alpha$.
- With probability $\mathbf{p}_{Y_i, \nabla}(x)$, $\alpha \triangleleft \gamma_i$. Then by Lemma 5.10.1 (d), for every $\delta \in I_i$, $\alpha \triangleleft \delta$.
- With probability $\mathbf{p}_{Y_i, \curvearrowright}(x)$, $\alpha \curvearrowright \gamma_i$. Then by Lemma 5.10.1 (b), for every $\delta \in I_i$, $\alpha \curvearrowright \delta$.

And similarly for β .

Therefore, by Lemma 5.5.3 (a), and $\{\gamma : \gamma \curvearrowright \alpha\} \subset \{\gamma : \gamma \curvearrowright \beta\}$, for every $i \geq a$, a.s.

$$\mathbb{P}(\mu_{Y_{i+1}, \curvearrowright}(\alpha, \beta) = \mu_{Y_i, \curvearrowright}(\alpha, \beta) + \mu(Y_i, Y_{i+1}] | \mathcal{F}_i) = p_{Y_i, \curvearrowright}(\alpha, \beta),$$

and,

$$\mathbb{P}(\mu_{Y_{i+1}, \curvearrowright}(\alpha, \beta) = \mu_{Y_i, \curvearrowright}(\alpha, \beta) | \mathcal{F}_i) = 1 - p_{Y_i, \curvearrowright}(\alpha, \beta).$$

As a result, $(\mathbf{p}_{Y_i, \curvearrowright}(\alpha, \beta), \mathcal{F}_i)_{i \geq a}$ is a Pólya urn in the sense of Lemma 2.8.1. And we can conclude exactly as in the proof of Lemma 2.5.1. \square

5.6 Construction and Hölder continuity of the contour path.

5.6.1 Construction of the contour path \mathfrak{C}_l on \mathcal{L}_l .

Lemma 5.6.1. *Recall that $\sim_{\mathcal{L}}$ denotes the metric equivalence on $(\mathcal{L}, d_{\mathcal{L}})$. A.s. for every $l > 0$ there exists a $\mu[0, l]$ -Lipschitz function $\mathfrak{C}_l : [0, 1] \mapsto \mathcal{L}_l$ such that for every $\alpha \in \mathcal{L}_l$, $\mathfrak{C}_l(\mathbf{p}_{l, \curvearrowright}(\alpha)) \sim_{\mathcal{L}} \alpha$.*

Remark. Since by Lemma 5.5.3 (e), $\{\mathbf{p}_{\curvearrowright}(\alpha), \alpha \in \mathcal{L}_l\}$ is dense, \mathfrak{C}_l is unique up to $\sim_{\mathcal{L}}$.

Proof. A.s. for every $l > 0$ the following holds: Let $S_l := \{\mathbf{p}_{\curvearrowright}(\alpha), \alpha \in \mathcal{L}_l\}$. For every $u \in S_l$, we may chose $\mathfrak{C}_l(u) \in \mathcal{L}_l$ such that $\mathbf{p}_{\curvearrowright}(\mathfrak{C}_l(u)) = u$. Then by Lemma 5.5.9, for every $\alpha, \beta \in \mathcal{L}_l$,

$$d_{\mathcal{L}}(\alpha, \beta) \leq |\mu_{\curvearrowright}(\alpha) - \mu_{\curvearrowright}(\beta)| = \mu[0, l] |\mathbf{p}_{\curvearrowright}(\alpha) - \mathbf{p}_{\curvearrowright}(\beta)|. \quad (5.18)$$

Thus for every $\alpha \in \mathcal{L}_l$, $d_{\mathcal{L}}(\alpha, \mathfrak{C}_l(\mathbf{p}_{\curvearrowright}(\alpha))) = 0$. Also, by (5.18), \mathfrak{C}_l is $\mu[0, l]$ -Lipschitz on S_l . Furthermore by Lemma 5.5.3 (e), S_l is dense. Thus, by compactness of $(\mathcal{L}_l, d_{\mathcal{L}})$ (see Lemma 5.4.3), \mathfrak{C}_l extends to a $\mu[0, l]$ -Lipschitz function on $[0, 1]$. \square

5.6.2 Construction of the contour path \mathfrak{C} .

For every $f, g : [0, 1] \mapsto \mathcal{L}$, let $d_\infty(f, g) := \max_{x \in [0, 1]} d_{\mathcal{L}}(f(x), g(x))$. In this section we prove:

Proposition 5.6.2. $(\mathfrak{C}_{Y_a})_{a \in \mathbb{N}}$ is almost surely a Cauchy sequence for d_∞ .

Since \mathcal{L} is a.s. compact, this directly implies that $(\mathfrak{C}_{Y_a})_{a \in \mathbb{N}}$ converges uniformly, and we define \mathfrak{C} as its limit. Our proof is mainly constructive, with several topological arguments along the way.

First, since almost surely as $a \rightarrow \infty$, $\mu[0, Y_a]Y_a/(\log^6 Y_a) \rightarrow \infty$, a.s. there exists for every $a \in \mathbb{N}$ large enough $(l_a, n_a) \in \mathbb{R}^+ \times \mathbb{N}$ such that

$$1 \leq \frac{Y_a}{\log^6 Y_a} \mu[0, Y_a] l_a \leq 2 \quad \text{and} \quad 2n_a l_a = 1. \quad (5.19)$$

we take (l_a, n_a) such that n_a is the largest possible.

Then for every $a \in \mathbb{N}$, let $\alpha_{a,0} := (0, 0)$, $\alpha_{a,n_a} = (0, 1)$ and for every $0 < i < n_a$, let $\alpha_{a,i} \in \mathcal{L}_{Y_a}$ such that

$$\mathfrak{p}_{Y_a, \curvearrowright}(\alpha_{a,i}) \in [2il_a, (2i+1)l_a]. \quad (5.20)$$

$\alpha_{a,i}$ exists by Lemma 5.5.3 (e), and by Lemma 5.5.3 (d), we can sample $\alpha_{a,i}$ in a measurable way. Also note that for every $a \in \mathbb{N}$, since $(0, 0)$ is the minimum for \prec and since $(0, 1)$ is the maximum for \prec ,

$$\mathfrak{p}_{Y_a, \curvearrowright}(0, 0) = 0 \quad \text{and} \quad \mathfrak{p}_{Y_a, \curvearrowright}(0, 1) = 1.$$

Thus, by Lemma 5.5.3 (c) and (5.20), a.s. for every $a \in \mathbb{N}$ large enough,

$$(0, 0) = \alpha_{a,0} \prec \alpha_{a,1} \cdots \prec \alpha_{a,n_a} = (0, 1). \quad (5.21)$$

Next, by Lemma 5.3.1 (a), (b), a.s. as $a \rightarrow \infty$, $n_a = O(Y_a^2)$. Also by Lemma 5.3.1 (c), a.s. $i^2 = O(Y_i)$, so a.s. $\sum_{a=1}^\infty Y_a^{-3} < \infty$. Therefore by the Borel–Cantelli Lemma and Lemma 5.5.11, for every $a \leq b$ large enough, and $0 \leq i < n_a$, by (5.19), (5.20),

$$l_a/2 \leq \mathfrak{p}_{Y_b, \curvearrowright}(\alpha_{a,i}, \alpha_{a,i+1}) \leq 4l_a. \quad (5.22)$$

In particular for every $a \leq b$ large enough, $(\mathfrak{p}_{Y_b, \curvearrowright}(\alpha_{a,i}))_{0 \leq i \leq n_a}$ is strictly increasing. Hence we may define $\psi_{a,b} : [0, 1] \mapsto [0, 1]$ such that for every $0 \leq i \leq n_a$,

$$\psi_{a,b}(\mathfrak{p}_{Y_a, \curvearrowright}(\alpha_{a,i})) = \mathfrak{p}_{Y_b, \curvearrowright}(\alpha_{a,i}),$$

and such that for every $0 \leq i < n_a$, $\psi_{a,b}$ is linear in $[\alpha_{a,i}, \alpha_{a,i+1}]$. And by (5.22), $\psi_{a,b}$ is strictly increasing and continuous.

Lemma 5.6.3. Almost surely for every $a \in \mathbb{N}$ large enough, $(\psi_{a,b})_{b \geq a}$ converges uniformly toward a strictly increasing continuous function ψ_a . And a.s. $(\psi_{a,b}^{-1})_{b \geq a}$ converges uniformly toward ψ_a^{-1} .

Proof. By Lemma 5.5.10 a.s. for every $a \in \mathbb{N}$ large enough, for every $0 \leq i \leq n_a$, as $b \rightarrow \infty$,

$$\mathfrak{p}_{Y_b, \curvearrowright}(\alpha_{a,i}) \rightarrow \mathfrak{p}_{\curvearrowright}(\alpha_{a,i}). \quad (5.23)$$

So if ψ_a is the function such that for every $0 \leq i \leq n_a$,

$$\psi_{a,b}(\mathfrak{p}_{Y_a, \curvearrowright}(\alpha_{a,i})) = \mathfrak{p}_{\curvearrowright}(\alpha_{a,i}),$$

and such that for every $0 \leq i < n_a$, ψ_a is linear in $[\alpha_{a,i}, \alpha_{a,i+1}]$. Then almost surely for every $a \in \mathbb{N}$ large enough, $(\psi_{a,b})_{b \geq a}$ converges uniformly toward ψ_a .

Moreover, by (5.22) and (5.23), a.s. for every $a \in \mathbb{N}$ large enough, for every $0 \leq i < n_a$,

$$l_a/2 \leq \psi_a(\mathfrak{p}_{Y_a, \curvearrowright}(\alpha_{a,i+1})) - \psi_a(\mathfrak{p}_{Y_a, \curvearrowright}(\alpha_{a,i})) \leq 4l_a. \quad (5.24)$$

Hence, ψ_a is also a strictly increasing continuous function. Finally by (5.23), (5.24), and the linearity, a.s. for every $a \in \mathbb{N}$ large enough, $(\psi_{a,b}^{-1})_{b \geq a}$ converges uniformly toward ψ_a^{-1} . \square

Lemma 5.6.4. *Almost surely for every $a \leq b$ large enough,*

$$d_\infty(\mathfrak{C}_{Y_a} \circ \psi_{a,b}^{-1}, \mathfrak{C}_{Y_b}) \leq d_H(\mathcal{L}_{Y_a}, \mathcal{L}) + 6 \log^6(Y_a)/Y_a.$$

Remark. This result does not depend on the choice of $(\mathfrak{C}_{Y_a})_{a \in \mathbb{N}}$, since they are unique up to $\sim_{\mathcal{L}}$.

Before proving Lemma 5.6.4 let us explain why it implies Proposition 5.6.2.

Proof of Proposition 5.6.2. First by Proposition 5.4.6 a.s. $d_H(\mathcal{L}_{Y_a}, \mathcal{L}) \rightarrow 0$. Hence, by Lemma 5.6.4 a.s.

$$\lim_{a \rightarrow \infty} \max_{b \geq a} d_\infty(\mathfrak{C}_{Y_a} \circ \psi_{a,b}^{-1}, \mathfrak{C}_{Y_b}) = 0.$$

Then by the triangular inequality, a.s.

$$\lim_{a \rightarrow \infty} \limsup_{b, c \rightarrow \infty} d_\infty(\mathfrak{C}_{Y_c}, \mathfrak{C}_{Y_b}) \leq \lim_{a \rightarrow \infty} \limsup_{b, c \rightarrow \infty} d_\infty(\mathfrak{C}_{Y_a} \circ \psi_{a,c}^{-1}, \mathfrak{C}_{Y_a} \circ \psi_{a,b}^{-1}).$$

Finally by Lemma 5.6.3 and continuity of \mathfrak{C}_{Y_a} the right hand side above is almost surely null. \square

Proof of Lemma 5.6.4. Almost surely for every $a \leq b$ large enough, $\beta \in \mathcal{L}_b$ the following holds: Since \prec is an order, by (5.21) there exists $0 \leq i < n_a$ such that $\alpha_{a,i} \prec \beta \prec \alpha_{a,i+1}$. Then recall from Lemma 5.4.2 that $\dot{\rho}_{Y_a}(\beta)$ denote the projection of β on \mathcal{L}_{Y_a} . Then note that by Lemma 5.5.2, $\alpha_{a,i} \prec \dot{\rho}_{Y_a}(\beta) \prec \alpha_{a,i+1}$. Thus, by Lemma 5.5.3 (c),

$$\mathfrak{p}_{Y_a, \curvearrowright}(\alpha_{a,i}) \leq \mathfrak{p}_{Y_a, \curvearrowright} \circ \dot{\rho}_{Y_a}(\beta) \leq \mathfrak{p}_{Y_a, \curvearrowright}(\alpha_{a,i+1}). \quad (5.25)$$

Also, since $\alpha_{a,i} \prec \beta \prec \alpha_{a,i+1}$,

$$\mathfrak{p}_{Y_b, \curvearrowright}(\alpha_{a,i}) \leq \mathfrak{p}_{Y_b, \curvearrowright}(\beta) \leq \mathfrak{p}_{Y_b, \curvearrowright}(\alpha_{a,i+1}).$$

Then by applying $\psi_{a,b}^{-1}$ which is strictly increasing we get,

$$\mathfrak{p}_{Y_a, \curvearrowright}(\alpha_{a,i}) \leq \psi_{a,b}^{-1} \circ \mathfrak{p}_{Y_b, \curvearrowright}(\beta) \leq \mathfrak{p}_{Y_a, \curvearrowright}(\alpha_{a,i+1}). \quad (5.26)$$

Therefore by (5.25), (5.26), (5.20), and (5.19),

$$|\psi_{a,b}^{-1} \circ \mathfrak{p}_{Y_b, \curvearrowright}(\beta) - \mathfrak{p}_{Y_a, \curvearrowright} \circ \dot{\rho}_{Y_a}(\beta)| \leq 3l_a \leq 6 \log^6(Y_a)/(Y_a \mu[0, Y_a]).$$

Then since by Lemma 5.6.1, \mathfrak{C}_{Y_a} is a.s. $\mu[0, Y_a]$ -Lipschitz,

$$d_{\mathcal{L}}(\mathfrak{C}_{Y_a} \circ \psi_{a,b}^{-1} \circ \mathfrak{p}_{Y_b, \curvearrowright}(\beta), \mathfrak{C}_{Y_a} \circ \mathfrak{p}_{Y_a, \curvearrowright} \circ \dot{\rho}_{Y_a}(\beta)) \leq 6 \log^6(Y_a)/Y_a.$$

Also, by using the definition of \mathfrak{C}_a , we have, $\mathfrak{C}_a \circ \mathfrak{p}_{Y_a, \curvearrowright} \circ \dot{\rho}_{Y_a}(\beta) \sim_{\mathcal{L}} \dot{\rho}_{Y_a}(\beta)$. Hence,

$$d_{\mathcal{L}}(\mathfrak{C}_{Y_a} \circ \psi_{a,b}^{-1} \circ \mathfrak{p}_{Y_b, \curvearrowright}(\beta), \dot{\rho}_{Y_a}(\beta)) \leq 6 \log^6(Y_a)/Y_a.$$

Then using $d_{\mathcal{L}}(\dot{\rho}_{Y_a}(\beta), \beta) \leq d_H(\mathcal{L}_{Y_a}, \mathcal{L})$ (see Lemma 5.4.2),

$$d_{\mathcal{L}}(\mathfrak{C}_{Y_a} \circ \psi_{a,b}^{-1} \circ \mathfrak{p}_{Y_b, \curvearrowright}(\beta), \beta) \leq d_H(\mathcal{L}_{Y_a}, \mathcal{L}) + 6 \log^6(Y_a)/Y_a.$$

Finally by using the definition of \mathfrak{C}_a , we have $\beta \sim_{\mathcal{L}} \mathfrak{C}_{Y_b} \circ \mathfrak{p}_{Y_b, \curvearrowright}(\beta)$, so,

$$d_{\mathcal{L}}(\mathfrak{C}_{Y_a} \circ \psi_{a,b}^{-1} \circ \mathfrak{p}_{Y_b, \curvearrowright}(\beta), \mathfrak{C}_{Y_b} \circ \mathfrak{p}_{Y_b, \curvearrowright}(\beta)) \leq d_H(\mathcal{L}_{Y_a}, \mathcal{L}) + 6 \log^6(Y_a)/Y_a. \quad (5.27)$$

To conclude the proof by (5.27), almost surely for every $a \leq b \in \mathbb{N}$ large enough, writing $S_b := \{\mathfrak{p}_{Y_b, \curvearrowright}(\gamma), \gamma \in \mathcal{L}_{Y_b}\}$,

$$\max_{x \in S_b} d_{\mathcal{L}}(\mathfrak{C}_{Y_a} \circ \psi_{a,b}^{-1}(x), \mathfrak{C}_{Y_b}(x)) \leq d_H(\mathcal{L}_{Y_a}, \mathcal{L}) + 6 \log^6(Y_a)/Y_a.$$

The maximum is then extended to $[0, 1]$, by density of S_b (see Lemma 5.5.3 (d)), since $\mathfrak{C}_{Y_a}, \psi_{a,b}^{-1}, \mathfrak{C}_{Y_b}$ are almost surely continuous. \square

5.6.3 Proof of Theorem 5.2.8

Recall that by Lemmas 5.5.7, 5.5.8, \mathbf{p}_{\cap} extends to a function continuous at each point of $\mathcal{L} \setminus (\mathbb{R}^+ \times [0, 1])$. We need to show that almost surely for every $\alpha \in \mathcal{L}$, $\mathfrak{C} \circ \mathbf{p}_{\cap}(\alpha) \sim_{\mathcal{L}} \alpha$. To this end, by continuity of \mathbf{p}_{\cap} and \mathfrak{C} at $\mathcal{L} \setminus (\mathbb{R}^+ \times [0, 1])$, and by density of $\mathbb{R}^+ \times [0, 1]$ (see Proposition 5.2.2 (a)), it is enough to show the desired result on $\mathbb{R}^+ \times [0, 1]$.

Almost surely for every $\alpha \in \mathbb{R}^+ \times [0, 1]$ the following holds: First by definition of $(\mathfrak{C}_l)_{l>0}$, for every l large enough, $\mathfrak{C}_l \circ \mathbf{p}_{l,\cap}(\alpha) \sim_{\mathcal{L}} \alpha$. Then by Lemma 5.5.10, a.s. $\mathbf{p}_{l,\cap}(\alpha) \rightarrow \mathbf{p}_{\cap}(\alpha)$. Hence, since a.s. $(\mathfrak{C}_l)_{l>0}$ converges uniformly toward \mathfrak{C} , as $l \rightarrow \infty$,

$$\mathfrak{C}_l \circ \mathbf{p}_{l,\cap}(\alpha) \rightarrow \mathfrak{C} \circ \mathbf{p}_{\cap}(\alpha)$$

Therefore, $\mathfrak{C} \circ \mathbf{p}_{\cap}(\alpha) \sim_{\mathcal{L}} \alpha$.

5.6.4 Holder continuity of \mathfrak{C} : Proof of Theorem 5.2.10

The next result is more precise than Theorem 5.2.10, and we will use it in the next section to estimate the Minkowski lower box dimension of \mathcal{L} .

Lemma 5.6.5. *Almost surely, for every $n \in \mathbb{N}$ large enough, for every $s, t \in [0, 1]$,*

$$d_{\mathcal{L}}(\mathfrak{C}(s), \mathfrak{C}(t)) \leq 13|s - t| \mathbb{E}[\mu[0, 2^n]] + 13n^6 2^{-n}.$$

Proof. We keep the notations of Section 5.6.2. By taking $b \rightarrow \infty$ in Lemma 5.6.4, by Lemma 5.6.3, a.s. for every a large enough,

$$d_{\infty}(\mathfrak{C}_{Y_a} \circ \psi_a^{-1}, \mathfrak{C}) \leq d_H(\mathcal{L}_{Y_a}, \mathcal{L}) + 6 \log^6(Y_a)/Y_a.$$

Hence, by the triangular inequality, a.s. for every a large enough, for every $s, t \in [0, 1]$,

$$d_{\mathcal{L}}(\mathfrak{C}(s), \mathfrak{C}(t)) \leq d_{\mathcal{L}}(\mathfrak{C}_{Y_a} \circ \psi_a^{-1}(s), \mathfrak{C}_{Y_a} \circ \psi_a^{-1}(t)) + 2d_H(\mathcal{L}_{Y_a}, \mathcal{L}) + 12 \log^6(Y_a)/Y_a.$$

Moreover, by Lemma 5.6.1, \mathfrak{C}_{Y_a} is $\mu[0, Y_a]$ -Lipschitz, and by (5.20), (5.24), ψ_a^{-1} is 6-Lipschitz. Thus,

$$d_{\mathcal{L}}(\mathfrak{C}(s), \mathfrak{C}(t)) \leq 6|s - t| \mu[0, Y_a] + 2d_H(\mathcal{L}_{Y_a}, \mathcal{L}) + 12l_a \mu[0, Y_a]. \quad (5.28)$$

Next, as a corollary of Lemma 2.4.5, a.s. for every n large enough, there exists $a \in \mathbb{N}$, such that $Y_a \in [2^n, 2^{n+1}]$. So, a.s. for every $n \in \mathbb{N}$ large enough, by (5.28), for every $s, t \in [0, 1]$,

$$d_{\mathcal{L}}(\mathfrak{C}(s), \mathfrak{C}(t)) \leq 6|s - t| \mu[0, 2^{n+1}] + 2d_H(\mathcal{L}_{2^n}, \mathcal{L}) + 12n^6 2^{-n}.$$

The desired result follows from Lemma 5.4.4 (b), and by Lemma 5.3.1 (a) (b). \square

Proof of Theorem 5.2.10. Let $\delta > \bar{\delta}$. By (5.2), $\mathbb{E}[\mu[0, 2^n]] = o(2^{n(\beta-1)})$. So, for every $n \in \mathbb{N}$ large enough, for every $s, t \in [0, 1]$ with $2^{-n\beta} \leq |s - t| \leq 13n^6 2^{-n\beta}$,

$$d_{\mathcal{L}}(\mathfrak{C}(s), \mathfrak{C}(t)) \leq n^7 2^{-n} \leq -\log(|s - t|)^8 |s - t|^{1/\beta}.$$

Since $\beta > \bar{\delta}$ is arbitrary, the desired result follows. \square

5.7 Fractal dimensions of the looptree

5.7.1 Definitions of the fractal dimensions

In this section X denotes a pseudo-metric space.

Definition. (*Minkowski dimensions*) For every $\varepsilon > 0$, an ε -set of X is a finite subset S of X such that $d_H(S, X) \leq \varepsilon$. For every $\varepsilon > 0$ let N_ε be the smallest size of a ε -set of X . Define the Minkowski lower box and upper box dimensions respectively by

$$\underline{\dim}(X) := \liminf_{\varepsilon \rightarrow 0} \frac{\log N_\varepsilon}{-\log \varepsilon} \quad \text{and} \quad \overline{\dim}(X) := \limsup_{\varepsilon \rightarrow 0} \frac{\log N_\varepsilon}{-\log \varepsilon}.$$

Definition. (*Packing dimension*) For every $s \geq 0$ and $A \subset X$ let

$$P_0^s(A) := \limsup_{\delta \rightarrow 0} \left\{ \sum_{i \in I} \text{diam}(B_i)^s \mid \{B_i\}_{i \in I} \text{ are disjoint balls } B(x, r) \text{ with } x \in A \text{ and } 0 < r \leq \delta \right\}.$$

and

$$P^s(X) := \inf \left\{ \sum_{i=1}^{\infty} P_0^s(A_i) \mid X \subset \bigcup_{i=1}^{\infty} A_i \right\}.$$

Then P^s is a decreasing function of s , and we define the packing dimension of X as

$$\dim_P(X) := \sup\{s, P^s(X) < \infty\}.$$

Definition. (*Hausdorff dimension*) For every $s, r \geq 0$ write

$$H_r^s(X) := \inf_{\text{diam}(A_i) \leq r} \left\{ \sum_{i=1}^{\infty} \text{diam}(A_i)^s \mid X \subseteq \bigcup_{i=1}^{\infty} A_i \right\}.$$

The Hausdorff dimension of X is defined by

$$\dim_H(X) := \sup \left\{ s, \sup_{r \in \mathbb{R}^+} H_r^s(X) < \infty \right\}.$$

Remark. Although, the above dimensions are usually considered for metric spaces, it is easy to check that they are exactly the same for a pseudo-metric space and for its quotient. For this reason, the below results still apply here.

To compute the packing dimension and Hausdorff dimension we will use the following extension of Theorem 6.9, and Theorem 6.11 from [70]. ([70] deals with subsets of Euclidian space, but the same arguments hold for every pseudo-metric space.)

Lemma 5.7.1. *Let p be a Borel probability measure on X and $s \in \mathbb{R}^+$.*

- a) *If p -almost everywhere $\limsup(\log p(B(x, \varepsilon)))/(\log \varepsilon) \geq s$ as $\varepsilon \rightarrow 0$, then $\dim_P(X) \geq s$.*
- b) *If p -almost everywhere $\liminf(\log p(B(x, \varepsilon)))/(\log \varepsilon) \geq s$ as $\varepsilon \rightarrow 0$, then $\dim_H(X) \geq s$.*

We have well-known inequalities (see e.g. Chapter 3 of Falconer [53]):

Lemma 5.7.2. *For every pseudo-metric space X we have*

$$\dim_H(X) \leq \underline{\dim}(X) \leq \overline{\dim}(X) \quad \text{and} \quad \dim_H(X) \leq \dim_P(X) \leq \overline{\dim}(X).$$

So we only need to upper bound $\underline{\dim}(\mathcal{L})$, $\overline{\dim}(\mathcal{L})$, and to lower bound $\dim_H(\mathcal{L})$ and $\dim_P(\mathcal{L})$.

5.7.2 Upper bound on the Minkowski dimensions

To upper bound the Minkowski dimensions we use the contour path \mathfrak{C} . By Theorem 5.2.8 for every $\alpha \in \mathcal{L}$, $\mathfrak{C}(\mathfrak{p}_{\wedge}(\alpha)) \sim_{\mathcal{L}} \alpha$. So $d_H(\mathfrak{C}([0, 1], \mathcal{L})) = 0$. As a result, for every $x > 0$, $\{ix, 0 \leq i \leq 1/x\}$ is a ε_x -set of \mathcal{L} , where $\varepsilon_x := \max_{s,t} |\mathfrak{C}(s) - \mathfrak{C}(t)|$. So for every $\alpha > 0$, if \mathfrak{C} is α Hölder continuous,

$$\overline{\dim}(\mathcal{L}) \leq \limsup_{x \rightarrow 0} \frac{\log(1/x + O(1))}{-\log(\varepsilon_x)} \leq \limsup_{x \rightarrow 0} \frac{\log(1/x + O(1))}{(1 + o(1)) \log(1/x^\alpha)} = 1/\alpha.$$

Therefore by Theorem 5.2.10 a.s. $\overline{\dim}(\mathcal{L}) \leq \bar{\mathfrak{d}}$. Also,

Lemma 5.7.3. *Almost surely for every $\alpha > 1/\overline{\dim}(\mathcal{L})$, \mathfrak{C} is not α Hölder continuous.*

We now consider $\underline{\dim}(\mathcal{L})$. By (5.2), there exists $(n_i)_{i \in \mathbb{N}}$ such that $\mathbb{E}[\mu[0, 2^{n_i}]] = o(2^{n_i}(\mathfrak{d}-1+o(1)))$. So by Lemma 5.2.10, a.s. for every i large enough, $s, t \in [0, 1]$ with $|s - t| \leq 2^{-n_i}/\mathbb{E}[\mu[0, 2^{n_i}]]$,

$$d_{\mathcal{L}}(\mathfrak{C}(s), \mathfrak{C}(t)) \leq 13|s - t|\mathbb{E}[\mu[0, 2^{n_i+1}]] + 13n_i^6 2^{-n_i} \leq n_i^7 2^{-n_i}.$$

Therefore, using the same ε -sets as before, a.s. as $i \rightarrow \infty$,

$$\underline{\dim}(\mathcal{L}) \leq \liminf_{i \rightarrow \infty} \frac{\log(2^{n_i} \mathbb{E}[\mu[0, 2^{n_i}]])}{-\log(n_i^7 2^{-n_i})} \leq \frac{n_i(1 + \mathfrak{d} - 1 + o(1))}{n_i(1 + o(1))} \leq \mathfrak{d}.$$

5.7.3 The rebranching principle.

We want to lower bound the dimensions of \mathcal{L} with Lemma 5.7.1. To this end, we morally needs to lower bound the distance between two random vertices in \mathcal{L} . In this section we show that it morally suffices to lower bound $d_{\mathcal{L}}((0, 0), (Y_1, 0))$. Our starting point is the rebranching principle, which we use here as follows: Let by convention $Y_0 = 0$.

Proposition 5.7.4. *For every permutation σ of $\{0\} \cup \mathbb{N}$, we have the following joint equality in distribution,*

$$\begin{aligned} (d_{\mathcal{T}}(Y_i, Y_j))_{i,j \geq 0} &\stackrel{(d)}{=} (d_{\mathcal{T}}(Y_{\sigma(i)}, Y_{\sigma(j)}))_{i,j \geq 0}, \\ \left(\mathbf{1}_{X_k \in \llbracket Y_i, Y_j \rrbracket} \right)_{i,j \geq 0, k \in \mathbb{N}} &\stackrel{(d)}{=} \left(\mathbf{1}_{X_k \in \llbracket Y_{\sigma(i)}, Y_{\sigma(j)} \rrbracket} \right)_{i,j \geq 0, k \in \mathbb{N}}. \end{aligned}$$

Proof. We briefly recall some discrete notions of Chapter 3. Let $(V_i)_{i \in \mathbb{N}}$ be a set of vertices. We say that $\mathcal{D} = (d_1, \dots, d_n)$ is a (pure) degree sequence if $\sum_{i=1}^n d_i = 2n - 2$ and if $d_1 \geq d_2 \geq \dots \geq d_n$.

For every degree sequence $\mathcal{D} = (d_1, \dots, d_n)$, we say that a tree T have degree sequence \mathcal{D} if T has vertices $(V_i)_{1 \leq i \leq n}$ and for every $1 \leq i \leq n$, V_i has degree d_i . For every degree sequence \mathcal{D} let $T^{\mathcal{D}}$ denote a uniform tree with degree sequence \mathcal{D} . Also let $L_0^{\mathcal{D}}, L_1^{\mathcal{D}}, \dots$ be the leaves of $T^{\mathcal{D}}$. Those leaves are determinist since they are the vertices of degree 1. We root $T^{\mathcal{D}}$ at $L_0^{\mathcal{D}}$.

For every tree T , let d_T denote the graph distance in T . Also, for every $A, B, C \in T$, we write $(A, T) \in \langle B, C \rangle$ if A lies in the path between B and C in T . By Proposition 3.8.6 (b), and by Lemma 3.8.1 (e), there exists $(\mathcal{D}_n)_{n \in \mathbb{N}}$ some degree sequences, such that the number of leaves and vertices of degree at least 2 diverges, and such that for every $a, b \in \mathbb{N}$ the following weak convergences hold jointly

$$(d_{T^{\mathcal{D}_n}}(L_i^{\mathcal{D}_n}, L_j^{\mathcal{D}_n}))_{0 \leq i, j \leq a} \longrightarrow (d_{\mathcal{T}}(Y_i, Y_j))_{0 \leq i, j \leq a}, \quad (5.29)$$

$$\left(\mathbf{1}_{(V_k, T^{\mathcal{D}_n}) \in \langle L_i^{\mathcal{D}_n}, L_j^{\mathcal{D}_n} \rangle} \right)_{0 \leq i, j \leq a, k \leq b} \longrightarrow \left(\mathbf{1}_{X_k \in \llbracket Y_i, Y_j \rrbracket} \right)_{0 \leq i, j \leq a, k \leq b}. \quad (5.30)$$

The main principle behind the rebranching principle is that for every n large enough the laws of the left hand sides of (5.29) and (5.30) are invariant under the permutation of the leaves. So by limit the right hand sides must also be invariant under those permutations. \square

Our goal is now to deduce from Proposition 5.7.4 an identity for $d_{\mathcal{L}}$. First let us introduce some topological notions. Recall that a.s. \mathbf{p} is a probability Borel measure on \mathcal{T} . Let $\mathbf{p}_{\mathcal{L}} := \mathbf{p} \times \mathbf{1}_{x \in [0,1]} dx$. Let $\mathcal{B}_{\mathcal{L}}$ be the product topology on $\mathcal{T} \times [0, 1]$. Let $\mathcal{F} := \sigma(\mathbf{X}, \mathbf{Y}, \mathbf{Z}, (U_{X,i})_{i \in \mathbb{N}}, (U_{Z,i})_{i \in \mathbb{N}})$. Let $\mathcal{B}_{\mathbb{R}}$ be the Borel topology on \mathbb{R} . Finally, for every $\alpha \in \mathcal{T} \times [0, 1]$, $\varepsilon > 0$, let $B(\alpha, \varepsilon)$ denote the open ball of center α of radius ε for $d_{\mathcal{L}}$ in $\mathcal{T} \times [0, 1]$.

Keep in mind that we are using two levels of randomness. On the one hand, we work on a completed probability space $(\Omega, \mathcal{F}, \mathbb{P})$. On the other hand, we work at $\omega \in \Omega$ fixed, on the random probability space $(\mathcal{T} \times [0, 1], \mathcal{B}_{\mathcal{L}}, \mathbf{p}_{\mathcal{L}})$.

Lemma 5.7.5. *Almost surely the following assertions hold:*

- (a) *The map $\alpha, \beta \mapsto d_{\mathcal{L}}(\alpha, \beta)$ is $(\mathcal{B}_{\mathcal{L}} \times \mathcal{B}_{\mathcal{L}}, \mathcal{B}_{\mathbb{R}})$ -measurable.*
- (b) *For every $(x, u) \in \mathcal{T} \times [0, 1]$, $\varepsilon > 0$, the open ball $B((x, u), \varepsilon)$ is $\mathcal{B}_{\mathcal{L}}$ -measurable.*
- (c) *For every $\varepsilon > 0$, the map $\alpha \in \mathcal{T} \times [0, 1] \mapsto \mathbf{p}_{\mathcal{L}}(B(\alpha, \varepsilon))$ is $(\mathcal{B}_{\mathcal{L}}, \mathcal{B}_{\mathbb{R}})$ -measurable.*

Proof. Toward (a), for every $i \in \mathbb{N}$, the map $y \in \mathcal{T} \mapsto U_{X_i, y}$ is continuous on $\mathcal{T} \setminus \{X_i\}$, since it is locally constant. Then $v \mapsto U_{X_i, X_i, v}$ is continuous. Hence $(y, v) \in \mathcal{T} \times [0, 1] \mapsto U_{X_i, y, v}$ is $(\mathcal{B}_{\mathcal{L}}, \mathcal{B}_{\mathbb{R}})$ -measurable. By definition of $d_{\mathcal{L}}$, (a) follows by sum and composition.

Then by (a) the map $(x, u), (y, v) \mapsto \mathbf{1}_{d_{\mathcal{L}}((x, u), (y, v)) \leq \varepsilon}$ is also $(\mathcal{B}_{\mathcal{L}} \times \mathcal{B}_{\mathcal{L}}, \mathcal{B}_{\mathbb{R}})$ -measurable. (b), (c) follows by Fubini's Theorem. \square

Next let \mathcal{M} be the set of probability distribution on \mathbb{R}^+ . We equip \mathcal{M} with the weak topology. Recall that \mathcal{M} is a Polish space. Let $\mathcal{B}_{\mathcal{M}}$ be the Borel topology on \mathcal{M} . For every probability distribution ν on $\mathcal{T} \times [0, 1]$ $\mathcal{B}_{\mathcal{L}}$ -measurable we construct $d_{\mathcal{L}} \star \nu \in \mathcal{M}$ as follows: Let α, β be two random variables with law ν . By Lemma 5.7.5 (a), $d_{\mathcal{L}}(\alpha, \beta)$ is a random variable $(\mathcal{B}_{\mathcal{L}} \times \mathcal{B}_{\mathcal{L}}, \mathcal{B}_{\mathbb{R}})$ -measurable. Let $d_{\mathcal{L}} \star \nu \in \mathcal{M}$ denote its probability distribution.

Lemma 5.7.6. *Let α be a random variable with law $\mathbf{p}_{\mathcal{L}}$. We have:*

- a) *For every $i, j \geq 0$, $d_{\mathcal{L}}((Y_i, 0), (Y_j, 0))$ is a random variable $(\mathcal{F}, \mathcal{B}_{\mathbb{R}})$ -measurable.*
- b) *Almost surely,*

$$d_{\mathcal{L}} \star \left(\frac{1}{n} \sum_{i=1}^n \delta_{Y_i, 0} \right) \xrightarrow{\text{weakly}} d_{\mathcal{L}} \star \mathbf{p}_{\mathcal{L}}.$$

- c) *$d_{\mathcal{L}} \star \mathbf{p}_{\mathcal{L}}$ is a random variable $(\mathcal{F}, \mathcal{B}_{\mathcal{M}})$ -measurable.*

Proof. Toward (a), let $i, j \in \mathbb{N}$. Note that $d_{\mathcal{T}}(Y_i, Y_j)$ is $(\mathcal{F}, \mathcal{B}_{\mathbb{R}})$ measurable since by Algorithm 5.1 it is a sum of measurable random variables. Also by Algorithm 5.1, for every $k \in \mathbb{N}$, $\mathbf{1}_{X_k \in [0, Y_i]}$ is $(\mathcal{F}, \mathcal{B}_{\mathbb{R}})$ -measurable. Then by the construction of the uniform angle function U in Algorithm 5.3, either $X_j \notin [0, Y_i]$ so $U_{X_k, Y_i, 0} = 0$, or $X_j \in [0, Y_i]$ and then $U_{X_k, Y_i, 0} = U_{X_k}$. Hence $U_{X_k, Y_i, 0}$ is $(\mathcal{F}, \mathcal{B}_{\mathbb{R}})$ -measurable. Similarly, $U_{X_k, Y_j, 0}$ is $(\mathcal{F}, \mathcal{B}_{\mathbb{R}})$ -measurable. (a) follows by definition of $d_{\mathcal{L}}$.

Toward (b), recall that by Theorem 2.3.1, a.s. $\frac{1}{n} \sum_{i=1}^n \delta_{Y_i}$ converges weakly toward \mathbf{p} for the weak topology on \mathcal{T} . Thus, a.s. if $(V_i)_{i \in \mathbb{N}}$ is a family of independent uniform random variables in $[0, 1]$, $\frac{1}{n} \sum_{i=1}^n \delta_{Y_i, V_i}$ converges weakly toward $\mathbf{p}_{\mathcal{L}}$. Hence, a.s.

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \delta_{Y_i, V_i} \times \delta_{Y_j, V_j} \xrightarrow{\text{weakly}} \mathbf{p}_{\mathcal{L}} \times \mathbf{p}_{\mathcal{L}}.$$

It directly follows by Lemma 5.7.5 that a.s.

$$d_{\mathcal{L}} \star \left(\frac{1}{n} \sum_{i=1}^n \delta_{Y_i, V_i} \right) \xrightarrow{\text{weakly}} d_{\mathcal{L}} \star \mathbf{p}_{\mathcal{L}}. \quad (5.31)$$

Finally note that almost surely $\{Y_i\}_{i \in \mathbb{N}} \cap \{X_i\}_{i \in \mathbb{N}} = \emptyset$. So almost surely for every $i, j \in \mathbb{N}$,

$$d_{\mathcal{L}}((Y_i, 0), (Y_j, 0)) = d_{\mathcal{L}}((Y_i, V_i), (Y_j, V_j)).$$

(b) follows from (5.31).

Toward (c), by (a) the left hand side of (b) is $(\mathcal{F}, \mathcal{B}_{\mathcal{M}})$ measurable. So (c) follows from (b). \square

Next, for every $f : \mathbb{R}^+ \mapsto \mathbb{R}$ bounded continuous and $M \in \mathcal{M}$ let $f(M) := \int_{\mathbb{R}^+} f(x) dM(x)$. Recall that $M \mapsto f(M)$ is continuous. So, by Lemma 5.7.6 (c), $f(d_{\mathcal{L}} \star p_{\mathcal{L}})$ is $(\mathfrak{F}, \mathcal{B}_{\mathbb{R}})$ measurable.

Proposition 5.7.7. *For every bounded continuous function $f : \mathbb{R}^+ \mapsto \mathbb{R}$,*

$$\mathbb{E}[f(d_{\mathcal{L}} \star p_{\mathcal{L}})] = \mathbb{E}[f(d_{\mathcal{L}}((0, 0), (Y_1, 0)))].$$

Proof. By Lemma 5.7.6 (b) and by Portmanteau's Theorem, it suffices to prove that for every $i \neq j$, $d_{\mathcal{L}}((Y_i, 0), (Y_j, 0))$ has the same law as $d_{\mathcal{L}}((0, 0), (Y_1, 0))$. Fix $i \neq j$. Since almost surely $\{Y_i\}_{i \in \mathbb{N}} \cap \{X_i\}_{i \in \mathbb{N}} = \emptyset$, by definition of $d_{\mathcal{L}}$, writing c for the distance on the torus $[0, 1]$,

$$d_{\mathcal{L}}((Y_i, 0), (Y_j, 0)) = d_{\mathcal{T}}(Y_i, Y_j) + \sum_{k \in \mathbb{N}} \theta_k c(U_{X_k, Y_i}, U_{X_k, Y_j}). \quad (5.32)$$

Then for every $k \in \mathbb{N}$ with $X_k \notin [Y_i, Y_j]$, Y_i and Y_j are connected in $\mathcal{T} \setminus \{X_k\}$ so $U_{X_k, Y_i} = U_{X_k, Y_j}$. Moreover, by the construction of the uniform angle function U in Algorithm 5.3, note that conditionally on $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$, $(c(U_{X_k, Y_i}, U_{X_k, Y_j}))_{k: X_k \in [Y_i, Y_j]}$ is a family of uniform random variables in $[0, 1/2]$.

As a result, by (5.32), conditionally on $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$, $d_{\mathcal{L}}((Y_i, 0), (Y_j, 0)) - d_{\mathcal{T}}(Y_i, Y_j)$ is a sum of uniform random variables in $([0, \theta_k/2])_{k: X_k \in [Y_i, Y_j]}$. And similarly for $d_{\mathcal{L}}((0, 0), (Y_1, 0)) - d_{\mathcal{T}}(0, Y_1)$. The desired result follows from Proposition 5.7.4. \square

5.7.4 Lower bound on the Hausdorff and Packing dimensions

To simplify the notations, let us write $d_{\mathcal{L}}(0, Y_1)$ for $d_{\mathcal{L}}((0, 0), (Y_1, 0))$. Let

$$\underline{d}' := \liminf_{\varepsilon \rightarrow 0} \frac{\log \mathbb{P}(d_{\mathcal{L}}(0, Y_1) < \varepsilon)}{\log \varepsilon} \quad ; \quad \bar{d}' := \limsup_{\varepsilon \rightarrow 0} \frac{\log \mathbb{P}(d_{\mathcal{L}}(0, Y_1) < \varepsilon)}{\log \varepsilon}. \quad (5.33)$$

Remark. Keep in mind that many inequalities get reversed since \log is negative on $(0, 1)$.

Lemma 5.7.8. *Almost surely $\dim_H(\mathcal{L}) \geq \underline{d}'$ and $\dim_P(\mathcal{L}) \geq \bar{d}'$.*

Proof. First note that $\mathcal{T} \times [0, 1] \subset \mathcal{L}$, so

$$\dim_H(\mathcal{T} \times [0, 1], d_{\mathcal{L}}) \leq \dim_H(\mathcal{L}) \quad \text{and} \quad \dim_P(\mathcal{T} \times [0, 1], d_{\mathcal{L}}) \leq \dim_P(\mathcal{L}).$$

Then recall Section 5.7.3. For every $n \in \mathbb{N}$, let $f_n : x \mapsto \max(\min(1, 2 - 2^n x), 0)$. Let α be a uniform random variable with law $p_{\mathcal{L}}$. We have by Proposition 5.7.7, for every $n \in \mathbb{N}$,

$$\mathbb{E}[f_n(d_{\mathcal{L}} \star p_{\mathcal{L}})] = \mathbb{E}[f(d_{\mathcal{L}}(0, Y_1))] \leq \mathbb{P}(d_{\mathcal{L}}(0, Y_1) < 2^{-n+1}).$$

Then by the Borel–Cantelli Lemma and Markov's inequality, a.s. for every n large enough,

$$f_n(d_{\mathcal{L}} \star p_{\mathcal{L}}) \leq n^2 \mathbb{P}(d_{\mathcal{L}}(0, Y_1) < 2^{-n+1}).$$

Also, note that a.s. if α is a random variable with law $p_{\mathcal{L}}$, so writing $\mathbb{E}_{\mathcal{L}}$ for the expectation with respect to $p_{\mathcal{L}}$, $\mathbb{E}_{\mathcal{L}}(B(\alpha, 2^{-n})) \leq f_n(d_{\mathcal{L}} \star p_{\mathcal{L}})$. Hence, by the Borel–Cantelli Lemma and Markov's inequality, a.s. $p_{\mathcal{L}}$ a.s. for every n large enough,

$$B(\alpha, 2^{-n}) \leq n^4 \mathbb{P}(d_{\mathcal{L}}(0, Y_1) < 2^{-n+1}).$$

The desired result follows from Lemma 5.7.1. \square

Lemma 5.7.9. Recall (5.2), (5.33). We have $\underline{\mathfrak{d}}' \geq \underline{\mathfrak{d}}$ and $\overline{\mathfrak{d}}' \geq \overline{\mathfrak{d}}$.

Remark. Since a.s. $\dim(\mathcal{L}) \leq \underline{\mathfrak{d}}$, $\overline{\dim}(\mathcal{L}) \leq \overline{\mathfrak{d}}$, by Lemmas 5.7.2, 5.7.8, we have $\underline{\mathfrak{d}}' \leq \underline{\mathfrak{d}}$ and $\overline{\mathfrak{d}}' \leq \overline{\mathfrak{d}}$.

Proof. In order to simplify the main proof, let us first deal with the case $\theta_0 > 0$. By Lemma 5.3.1 (b), $\mathbb{E}[\mu[0, l]] \sim \theta_0^2 l$ as $l \rightarrow \infty$. So $\underline{\mathfrak{d}} = \overline{\mathfrak{d}} = 2$. Also, $d_{\mathcal{L}}(0, Y_1) \geq \theta_0^2/4Y_1$ so for every $\varepsilon > 0$,

$$\mathbb{P}(d_{\mathcal{L}}(0, Y_1) \leq \varepsilon) \leq \mathbb{P}(Y_1 \leq 4\varepsilon/\theta_0^2).$$

Then, since $\{Y_i\}_{i \in \mathbb{N}}$ is a Poisson point process of rate $\mu[0, x]dx$, as $\varepsilon \rightarrow 0$, by Lemma 5.3.1 (a),

$$\mathbb{P}(Y_1 \leq \varepsilon) = \mathbb{E}[\mathbb{P}(Y_1 \leq \varepsilon | \mu)] \leq \varepsilon \mathbb{E}[\mu[0, \varepsilon]] = O(\varepsilon^2).$$

As a result, $\underline{\mathfrak{d}}' \geq 2 = \underline{\mathfrak{d}}$ and $\overline{\mathfrak{d}}' \geq 2 = \overline{\mathfrak{d}}$. So we may assume henceforth that $\theta_0 = 0$.

Next, note that a.s. $d_{\mathcal{L}}(0, Y_1) = \sum_{0 \leq X_i \leq Y_1} \theta_i \mathfrak{c}(U_{X_i})$. Thus, writing for every $n \in \mathbb{N}$, $\Delta_n : x \mapsto \sum_{0 \leq X_i \leq Y_1} \min(\theta_i \mathfrak{c}(U_{X_i}), 2^{-n})$, we have $d_{\mathcal{L}}(0, Y_1) \leq \Delta_n(Y_1)$. Then writing for every $n \in \mathbb{N}$, $\varepsilon_n := 16n^2/\mathbb{E}[\mu[0, 2^n]]$, we have, since $x \mapsto \Delta_n(x)$ is increasing,

$$\mathbb{P}(d_{\mathcal{L}}(0, Y_1) \leq 2^{-n}) \leq \mathbb{P}(\Delta_n(\varepsilon_n) \leq 2^{-n}) + \mathbb{P}(Y_1 \leq \varepsilon_n, \Delta_n(Y_1) \leq 2^{-n}). \quad (5.34)$$

Our proof consists in estimating both terms of the right hand side above.

Remark. Let us morally explain why we work with Δ_n and not $d_{\mathcal{L}}$. We want to estimate, for $\varepsilon > 0$ small, the typical $d_{\mathcal{L}}$ -distance between two random vertices ε - $d_{\mathcal{T}}$ -close from each other. It appears, that the few vertices ε - $d_{\mathcal{T}}$ -close from the vertices of high degrees X_1, X_2, \dots , tends to be $d_{\mathcal{L}}$ -far from each other. As a result, the moments of $d_{\mathcal{L}}(0, Y_1) \mathbf{1}_{Y_1 \leq \varepsilon}$ are highly biased toward the moments of the typical $d_{\mathcal{L}}$ -distance between two vertices ε - $d_{\mathcal{T}}$ -close from X_1, X_2, \dots . To avoid this bias, we use $(\Delta_n)_{n \in \mathbb{N}}$ to truncate the $d_{\mathcal{L}}$ -distance between the vertices $d_{\mathcal{T}}$ -close from X_1, X_2, \dots .

Now, recall that $(X_i)_{i \in \mathbb{N}}$ are independent exponential random variables of parameter $(\theta_i)_{i \in \mathbb{N}}$. We get with elementary computations, for every $0 \leq x \leq 1$, $n \in \mathbb{N}$,

$$\mathbb{E}[\Delta_n(x)] = \sum_{i=1}^{\infty} \mathbb{P}(X_i \leq x) \mathbb{E}[\min(\theta_i \mathfrak{c}(U_{X_i}), 2^{-n})] \geq \frac{1}{8} \sum_{i=1}^{\infty} \theta_i x \min(\theta_i, 2^{-n}). \quad (5.35)$$

And, for every $n \in \mathbb{N}$,

$$\mathbb{E}[\mu[0, 2^n]] = \sum_{i=1}^{\infty} \mathbb{P}(X_i \leq 2^n) \theta_i = \sum_{i=1}^{\infty} (1 - e^{-\theta_i 2^n}) \theta_i \leq \sum_{i=1}^{\infty} \min(1, \theta_i 2^n) \theta_i. \quad (5.36)$$

Hence, by (5.35) and (5.36), for every $n \in \mathbb{N}$, $0 \leq x \leq 1$, $\mathbb{E}[\Delta_n(x)] \geq x \mathbb{E}[\mu[0, 2^n]] 2^{-n-3}$. In particular, for every $n \in \mathbb{N}$ such that $\varepsilon_n \leq 1$, $\mathbb{E}[\Delta_n(\varepsilon_n)] \geq 2n^2 2^{-n}$. Also, note that Δ_n is a sum of independent random variables bounded by 2^{-n} . So, by Bernstein's inequality (see e.g. [32] Section 2.7 (2.10)) for every $n \in \mathbb{N}$ with $\varepsilon_n \leq 1$, $\mathbb{P}(\Delta_n(\varepsilon_n) \leq 2^{-n}) \leq e^{-n^2/6}$. Furthermore, by Lemma 5.3.1 (c), $\mathbb{E}[\mu[0, 2^n]] = O(2^n)$. So, as $n \rightarrow \infty$ with $\varepsilon_n \leq 1$,

$$\mathbb{P}(\Delta_n(\varepsilon_n) \leq 2^{-n}) = O(n^2 2^{-n} / \mathbb{E}[\mu[0, 2^n]]). \quad (5.37)$$

Toward upper bounding the right most term of (5.34), let us recall the construction of Aldous, Camarri, Pitman [?] of the ICRT in the simple case where $\theta_0 = 0$. Let $((A_{i,j})_{j \geq 0})_{i \in \mathbb{N}}$ be a family of independent Poisson point processes of intensity $(\theta_i)_{i \in \mathbb{N}}$ on \mathbb{R}^+ . By Lemma 2.2.1 that there exists a coupling such that a.s.

$$((X_i)_{i \in \mathbb{N}}, \{Y_i, Z_i\}_{i \in \mathbb{N}}) = ((A_{i,0})_{i \in \mathbb{N}}, \{A_{i,j}, A_{i,0}\}_{i,j \in \mathbb{N}}).$$

In particular $Y_1 \in \{A_{i,1}\}_{i \in \mathbb{N}}$. So, by an union bound, for every $x \geq 0$,

$$\begin{aligned} \mathbb{P}(Y_1 \leq x, \Delta_n(Y_1) \leq 1/2^n) &\leq \sum_{i \in \mathbb{N}} \mathbb{P}(A_{i,1} \leq x, \Delta_n(Y_1) \leq 2^{-n}) \\ &\leq \sum_{i \in \mathbb{N}} \mathbb{P}(A_{i,1} \leq x, \theta_i U_{X_i} \leq 2^{-n}) \\ &\leq \sum_{i \in \mathbb{N}} x^2 \theta_i^2 \min(1, 2^{-n}/\theta_i). \end{aligned}$$

Also, by directly adapting (5.36), we get for every $n \in \mathbb{N}$, $\mathbb{E}[\mu[0, 2^n]] \geq (1/4) \sum_{i=1}^{\infty} \min(1, \theta_i 2^n) \theta_i$. So, for every $x \geq 0$,

$$\mathbb{P}(Y_1 \leq x, \Delta_n(Y_1) \leq 1/2^n) \leq 4x^2 \mathbb{E}[\mu[0, 2^n]] 2^{-n}. \quad (5.38)$$

In particular, since $\varepsilon_n = 16n^2/\mathbb{E}[\mu[0, 2^n]]$,

$$\mathbb{P}(Y_1 \leq \varepsilon_n, \Delta_n(Y_1) \leq 1/2^n) \leq 4\varepsilon_n^2 \mathbb{E}[\mu[0, 2^n]] 2^{-n} = O(n^4 2^{-n}/\mathbb{E}[\mu[0, 2^n]]). \quad (5.39)$$

To sum up, by (5.34), (5.37), (5.39), we have as $n \rightarrow \infty$ with $\varepsilon_n \leq 1$,

$$\mathbb{P}(d_{\mathcal{L}}((0, Y_1) \leq 2^{-n}) = O(n^4 2^{-n}/\mathbb{E}[\mu[0, 2^n]]). \quad (5.40)$$

Finally, we have as $n \rightarrow \infty$ with $\varepsilon_n \geq 1$, using, $Y_1 \leq A_{1,1}$, (5.38), and $\varepsilon_n = 16n^2/\mathbb{E}[\mu[0, 2^n]]$,

$$\begin{aligned} \mathbb{P}(d_{\mathcal{L}}(0, Y_1) \leq 2^{-n}) &\leq \mathbb{P}(Y_1 > n^2) + \mathbb{P}(Y_1 \leq n^2, \Delta_n(Y_1) \leq 2^{-n}) \\ &\leq \mathbb{P}(A_{1,1} > n^2) + \mathbb{P}(Y_1 \leq n^2, \Delta_n(Y_1) \leq 2^{-n}) \\ &= O(n^2 e^{-\theta_1 n^2} + n^4 \mathbb{E}[\mu[0, 2^n]] 2^{-n}) \\ &= O(n^8 2^{-n}/\mathbb{E}[\mu[0, 2^n]]). \end{aligned} \quad (5.41)$$

The desired result follows from (5.40) (5.41) and the definitions of $\bar{\mathfrak{d}}, \mathfrak{d}$ in (5.2), and of $\bar{\mathfrak{d}}', \mathfrak{d}'$ in (5.33). \square

5.8 Proof of Theorem 5.2.4

Recall the definition of \mathfrak{G} from (5.3). Note that \mathfrak{G} is continuous on any bounded interval of \mathbb{R}^+ . We need to prove that it extends to a continuous map on \mathcal{T} . To this end, we use the chaining method. Let $\mathcal{F} = \sigma(\mu, \mathbf{Y}, \mathbf{Z})$. For all $x \in \mathcal{T}$, $y \in \mathbb{R}^+$, recall that $\bar{\rho}_y(x)$ denote the projection of x on $[0, y]$. Also to simplify the notations, for every $x, y \in \mathbb{R}^+$ let $d_{\mathfrak{G}}(x, y) := \max_{z \in \llbracket x, y \rrbracket} |\mathfrak{G}(x) - \mathfrak{G}(z)|$.

Remark. Although $d_{\mathfrak{G}}$ is a (pseudo) distance on \mathcal{T} , we will not study the topology of $(\mathcal{T}, d_{\mathfrak{G}})$.

Lemma 5.8.1. *a) For every x, y random variables in \mathbb{R}^+ \mathfrak{F} -measurable and $t \geq 0$, almost surely*

$$\mathbb{P}\left(d_{\mathfrak{G}}(x, y) > t\sqrt{d_{\mathcal{T}}(x, y)} \middle| \mathcal{F}\right) \leq 4e^{-t^2/2}.$$

b) For every $x, y, t \geq 0$,

$$\mathbb{P}\left(d_{\mathfrak{G}}(\bar{\rho}_y(x), x) > t^2/\sqrt{\mu[0, y]}\right) \leq 5e^{-t^2/4}.$$

Proof. Toward (a). Let $\gamma : [0, d_{\mathcal{T}}(x, y)] \mapsto \llbracket x, y \rrbracket$ be the geodesic from x to y . Note that by definition of \mathfrak{G} , conditionally on \mathcal{F} , $\mathfrak{G} \circ \gamma - \mathfrak{G}(x)$ is a brownian motion. So (a) directly follows from standard results on Brownian motions.

Toward (b), by Lemma 2.4.7 for every $x, y, t \geq 0$,

$$\mathbb{P}(d_{\mathcal{T}}(\bar{\rho}_y(x), x) > t^2/\mu[0, y]) \leq e^{-t^2/4}.$$

The desired result follows using (a) and an union bound. \square

We may assume below that $\theta_0 > 0$ or $\sum_{i=1}^{\infty} \theta_i = \infty$, since otherwise $\mathbb{E}[\mu(\mathbb{R}^+)] < \infty$ and the assumption of Theorem 5.2.4 is not satisfied. Then by Lemma 5.3.1 (a), we may define for $n \in \mathbb{N}$, \mathcal{X}_n as the unique real such that $\mathbb{E}[\mu[0, \mathcal{X}_n]] = e^n$. The next result is adapted from Lemma 2.6.2, which is used to prove the compactness of ICRT and upper bound its fractal dimensions.

Lemma 5.8.2. *Almost surely for every n large enough:*

$$\max_{x \in [0, \mathcal{X}_n]} |\mathfrak{G}(x) - \mathfrak{G}(\bar{\rho}_{\mathcal{X}_{n-1}}(x))| \leq 21 \log(\mathcal{X}_n) e^{-(n-1)/2}.$$

Proof. We first adapt the proof of Lemma 2.6.2, to morally deal with vertices far from $\{Y_i\}_{i \in \mathbb{N}}$. For every $n \in \mathbb{N}$, and $x \in \mathbb{R}^+$ let $E_n(x)$ be the event $d_{\mathfrak{G}}(\bar{\rho}_{\mathcal{X}_{n-1}}(x), x) > 20 \log(\mathcal{X}_n) e^{-(n-1)/2}$. By Fubini's Theorem, Lemma 5.8.1 (b) and $\mu[0, \mathcal{X}_{n-1}] = e^{n-1}$, we have

$$\mathbb{E} \left[\int_0^{\mathcal{X}_n} \mathbf{1}_{E_n(x)} dx \right] = \int_0^{\mathcal{X}_n} \mathbb{P}(E_n(x)) dx \leq 5 \mathcal{X}_n \exp(-5 \log \mathcal{X}_n) = 5 \mathcal{X}_n^{-4}.$$

Furthermore by Lemma 5.3.1 (b), $e^n = O(\mathcal{X}_n)$ so $\sum \mathcal{X}_n^{-1} < \infty$. Hence by Markov's inequality and the Borel–Cantelli Lemma, for every n large enough,

$$\int_0^{\mathcal{X}_n} \mathbf{1}_{E_n(x)} dx < \mathcal{X}_n^{-3}.$$

Note that it implies that, for every n large enough, for every $x \in [0, \mathcal{X}_{2n}]$, $y \in [0, \mathcal{X}_{2n}]$, with $x \in \llbracket p_{\mathcal{X}_{n-1}}(y), y \rrbracket$ and $d_{\mathcal{T}}(x, y) \geq \mathcal{X}_n^{-3}$,

$$|\mathfrak{G}(x) - \mathfrak{G}(\bar{\rho}_{\mathcal{X}_{n-1}}(x))| \leq 20 \log(\mathcal{X}_n) e^{-(n-1)/2}, \quad (5.42)$$

since otherwise for every $z \in \llbracket x, y \rrbracket$ we would have $E_n(z)$.

Next, let $\Lambda_n := (\{Y_i\}_{i \in \mathbb{N}} \cap [\mathcal{X}_{n-1}, \mathcal{X}_n]) \cup \{\mathcal{X}_n\}$. Note that for $x \in [0, \mathcal{X}_n]$, if there exists $y \in \Lambda_n$ such that $x \in \llbracket \bar{\rho}_{\mathcal{X}_{n-1}}(y), y \rrbracket$ and $d_{\mathcal{T}}(x, y) > \mathcal{X}_n^{-3}$, then (5.42) holds. So, writing for $y \in \Lambda_n$, $B(y, \mathcal{X}_n^{-3})$ for the closed ball of center y and radius ε , and $\gamma_{n,y} := \llbracket \bar{\rho}_{\mathcal{X}_{n-1}}(y), y \rrbracket \cap B(y, \mathcal{X}_n^{-3})$, it remains to estimate \mathfrak{G} on the set $\bigcup_{y \in \Lambda_n} \gamma_{n,y}$.

More precisely, by (5.42), it is enough to show that a.s. for every n large enough and $y \in \Gamma_n$,

$$M_{n,y} := \max_{x \in \gamma_{n,y}} |\mathfrak{G}(x) - \mathfrak{G}(y)| \leq \log(\mathcal{X}_n) e^{-(n-1)/2}. \quad (5.43)$$

To this end, note that for every $y \in \Gamma_n$, $\gamma_{n,y}$ is a geodesic with extremities \mathfrak{F} -measurable, and of length at most \mathcal{X}_n^{-3} . Thus, by Lemma 5.8.1 (a) and an union bound a.s.,

$$\mathbb{P} \left(\max_{y \in \Gamma_n} M_{n,y} > \log(\mathcal{X}_n) e^{-(n-1)/2} \middle| \mathcal{F} \right) \leq \#\Gamma_n 4e^{-\mathcal{X}_n^3 \log(\mathcal{X}_n)^2 e^{-(n-1)}}. \quad (5.44)$$

Then by Lemma 5.3.1 (b), $e^n = O(\mathcal{X}_n)$, and by Lemma 5.3.1 (c), $\#\Gamma_n = O(\mathcal{X}_n^2)$. Thus a.s. the right hand side of (5.44) is summable. Finally (5.43) follows from the Borel–Cantelli Lemma. \square

Proof of Theorem 5.2.4. Beforehand, let us rewrite the assumption. We have,

$$\int_{\mathcal{X}_1}^{\infty} \frac{dl}{l\sqrt{\mathbb{E}[\mu[0, l]]}} = \sum_{n=1}^{\infty} \int_{\mathcal{X}_n}^{\mathcal{X}_{n+1}} \frac{dl}{l\sqrt{\mathbb{E}[\mu[0, l]]}} \geq \sum_{n=1}^{\infty} \int_{\mathcal{X}_n}^{\mathcal{X}_{n+1}} \frac{dl}{le^{(n+1)/2}} \geq \sum_{n=2}^{\infty} \frac{\log \mathcal{X}_n}{e^{(n+9)/2}} - \log \mathcal{X}_1.$$

So the assumption implies that $\sum \log \mathcal{X}_n e^{-n/2} < \infty$. (It is actually equivalent.)

Then for every $n \in \mathbb{N}$ let $\mathfrak{G}_n : x \in \mathcal{T} \mapsto \mathfrak{G}(\bar{\rho}_{\mathcal{X}_{n-1}}(x))$. Note that for every $n \in \mathbb{N}$, since \mathfrak{G} is continuous on $([0, 2^n], d_{\mathcal{T}})$, \mathfrak{G}_n is continuous. Moreover, by Lemma 5.8.2, a.s. for every n large enough $\|\mathfrak{G}_n - \mathfrak{G}_{n-1}\|_{\infty} \leq 21 \log \mathcal{X}_n e^{-(n-1)/2}$. Therefore, by the assumption of Theorem 5.2.4, $\{\mathfrak{G}_n\}_{n \in \mathbb{N}}$ is a.s. a Cauchy sequence of continuous function, and so converges uniformly toward a continuous function $\tilde{\mathfrak{G}}$ on \mathcal{T} . Finally note that $\tilde{\mathfrak{G}}$ extends \mathfrak{G} . \square

5.9 Proof and extension of Theorem 5.2.5

We consider $(\mathfrak{D}_i)_{i \in \mathbb{N}}$ independent random functions from $[0, 1]$ to \mathbb{R} independent with $\sigma(\mathbf{X}, \mu, \mathbf{Y}, \mathbf{Z}, U)$. We assume that:

Assumption 5.1. For every $i \in \mathbb{N}$, $\mathfrak{D}_i(0) = 0$ and $\mathbb{E}[\mathfrak{D}_i(U_{X_i})] = 0$.

Assumption 5.2. There exists $\kappa \geq 2$ such that for every $i \in \mathbb{N}$, $\mathbb{E}[\|\mathfrak{D}_i\|_{\infty}^{\kappa}] \leq 1$, and

$$\mathbb{E}[\mu[0, l]] = O(l^{\kappa/2-1} \log(l)^{-5\kappa}). \quad (5.45)$$

Remarks. The assumptions are not necessary, and our method may be extended to other functions. However, they naturally appear in the study of field on \mathcal{D} -trees: (we will give more details in [26])

- The first assumption is called locally centered. One can always split each \mathfrak{D}_i , into a "constant" $u \mapsto \mathbf{1}_{u \neq 0} c_i$, and a centered function. However, dealing with constants require different techniques.
- The second assumption is an improvement of the so called $4 + \varepsilon$ moment assumption. Indeed by Lemma 5.3.1 (a), $\mathbb{E}[\mu[0, 2^n]] = O(2^n)$, so when $\kappa > 4$, (5.45) holds.
- In particular if μ corresponds to the α -stable trees, $\mu[0, l] \sim c.l^{\alpha-1}$, and we can consider $\kappa > 2\alpha$.

Under Assumptions 5.1, 5.2, which are satisfied by $(\mathfrak{B}_i)_{i \in \mathbb{N}}$ (up to a constant), we have:

Proposition 5.9.1. *Almost surely $\sum_{i=1}^{\infty} \sqrt{\theta_i} \mathfrak{D}_i(U_{X_{i,\alpha}})$ converges uniformly on \mathcal{L} .*

For every $n, m \in \mathbb{N}$ let

$$\mathfrak{S}_{n,m} := \sum_{i=n}^m \sqrt{\theta_i} \mathfrak{D}_i(U_{X_{i,\alpha}}).$$

To prove the uniform convergence, we need to show that almost surely as $N \rightarrow \infty$,

$$\Delta_N := \max_{\alpha \in \mathcal{L}} \max_{n, m \geq N} |\mathfrak{S}_{n,m}(\alpha)| \rightarrow 0.$$

Note that for every $n, m \in \mathbb{N}$, $|\mathfrak{S}_{n,m}(\alpha)|$ reach its maximum on $\bigcup_{i=n}^m \{X_i\} \times [0, 1] \subset \mathbb{R}^+ \times [0, 1]$. So,

$$\Delta_N = \max_{(x,u) \in \mathbb{R}^+ \times [0,1]} \max_{n, m \geq N} |\mathfrak{S}_{n,m}(x, u)|. \quad (5.46)$$

To show that a.s. $\Delta_N \rightarrow 0$ we use the chaining method (see Section 5.1.3): For every $i \in \mathbb{N}$, let

$$\Delta_{N,i} := \max_{Y_i < x \leq Y_{i+1}, u \in [0,1]} \max_{n, m \geq N} |\mathfrak{S}_{n,m}(x, u) - \mathfrak{S}_{n,m}(Z_i, U_{Z_i, Y_{i+1}})|$$

Then define \mathfrak{M}_N by induction on $([0, Y_i])_{i \in \mathbb{N}}$ such that $\mathfrak{M}_N(0) = 0$, and such that for every $i \geq 0, x \in]Y_i, Y_{i+1}]$,

$$\mathfrak{M}_N(x) := \mathfrak{M}_N(Z_i) + \Delta_{N,i}. \quad (5.47)$$

The triangular inequality implies by induction that for every $i \in \mathbb{N}$, for every $x \in [0, Y_i]$, $u \in [0, 1]$, $\max_{n,m \geq N} |\mathfrak{S}_{n,m}(x, u)| \leq \mathfrak{M}_N(x)$, so

$$\Delta_N \leq \max_{x \in \mathbb{R}^+} \mathfrak{M}_N(x). \quad (5.48)$$

Our method consists in estimating $(\Delta_{N,i})_{N,i \in \mathbb{N}}$, and then deduce from (5.47) and Lemma 5.3.2 an estimate on $\max \mathfrak{M}_N$.

To simplify the notations we write $\mathcal{F} := \sigma(\mathbf{X}, \mu, \mathbf{Y}, \mathbf{Z})$ and $\mu^N := \sum_{i \geq N} \theta_i \delta_{X_i}$.

Lemma 5.9.2. *There exists $C_\kappa > 0$ which depends only on κ such that for every $N, i \in \mathbb{N}$, $t > 0$,*

$$\mathbb{P}(\Delta_{N,i} \geq t | \mathcal{F}) \leq C_\kappa \left(\frac{\mu^N(Y_i, Y_{i+1}]}{t^2} \right)^{\kappa/2}.$$

Proof. We work conditionally on \mathcal{F} . First for every $x \in (Y_i, Y_{i+1}]$ and $j \in \mathbb{N}$ with $X_j \notin \{Z_i\} \cup (Y_i, x]$, Z_i and x are connected in $\mathcal{T} \setminus \{X_j\}$ so $U_{X_j, Z_i} = U_{X_j, x}$. Similarly, $U_{Z_i, x} = U_{Z_i, Y_{i+1}}$. Also, by definition of U , for every $j \in \mathbb{N}$ with $X_j \in (Y_i, x)$, $U_{X_j, x} = U_{X_j, Y_{i+1}} = U_{X_j}$. Also, $U_{x, x, 0} = 0$. Therefore, for every $n, m \in \mathbb{N}$, and $x \in (Y_i, Y_{i+1}]$,

$$\mathfrak{S}_{n,m}(x, 0) - \mathfrak{S}_{n,m}(Z_i, U_{Z_i, Y_{i+1}}) = \sum_{j \in \mathbb{N}: Y_i < X_j < x} \sqrt{\theta_j} \mathfrak{D}_j(U_{X_j}).$$

Moreover by definition of \mathfrak{D} , U and Assumption 5.1, $(\mathfrak{D}_j(U_{X_j}))_{j \in \mathbb{N}}$ are independent and centered. Also by Assumption 5.2, for every $j \in \mathbb{N}$, $\mathbb{E}[\|\mathfrak{D}_j\|_\infty^\kappa] \leq 1$. So by Lemma 5.12.1, there exists $C > 0$, which depends only on ε , such that for every $N \in \mathbb{N}$, and $t > 0$, a.s.

$$\mathbb{P} \left(\max_{n,m \geq N} \max_{Y_i < x \leq Y_{i+1}} |\mathfrak{S}_{n,m}(x, 0) - \mathfrak{S}_{n,m}(Z_i, U_{Z_i, Y_{i+1}})| \geq t \middle| \mathcal{F} \right) \leq C \left(\frac{\mu^N(Y_i, Y_{i+1}]}{t^2} \right)^{\kappa/2}. \quad (5.49)$$

Furthermore, for every $x \in (Y_i, Y_{i+1}]$, $u \in [0, 1]$, with $x \notin \{X_j\}_{j \in \mathbb{N}}$, for every $n, m \in \mathbb{N}$, $\mathfrak{S}_{n,m}(x, u) = \mathfrak{S}_{n,m}(x, 0)$. And for every $j \in \mathbb{N}$,

$$\max_{n,m \geq N} \max_{u \in [0,1]} |\mathfrak{S}_{n,m}(X_j, u) - \mathfrak{S}_{n,m}(X_j, 0)| = \mathbf{1}_{j \geq N} \sqrt{\theta_j} \max_{u \in [0,1]} \mathfrak{D}_j(u).$$

Then by Assumption 5.2, and by Markov's inequality, for every $j \in \mathbb{N}$, $t > 0$,

$$\mathbb{P} \left(\sqrt{\theta_j} \max_{u \in [0,1]} \mathfrak{D}_j(u) > t \right) \leq (\theta_j/t^2)^{\kappa/2}.$$

Therefore by an union bound, and by convexity of $x \mapsto x^{\kappa/2}$, for every $t > 0$,

$$\begin{aligned} \mathbb{P} \left(\max_{n,m \geq N} \max_{Y_i < x \leq Y_{i+1}} \max_{u \in [0,1]} |\mathfrak{S}_{n,m}(x, u) - \mathfrak{S}_{n,m}(x, 0)| \geq t \middle| \mathcal{F} \right) &\leq \sum_{j \geq N} \left(\frac{\theta_j}{t^2} \right)^{\kappa/2} \mathbf{1}_{Y_i < X_j \leq Y_{i+1}} \\ &\leq \left(\frac{\mu^N(Y_i, Y_{i+1}]}{t^2} \right)^{\kappa/2}. \end{aligned} \quad (5.50)$$

Finally the desired inequality follows from (5.49), (5.50), and an union bound. \square

Lemma 5.9.3. *A.s. for every n large enough, for every $i \in \mathbb{N}$ with $(Y_i, Y_{i+1}] \cap [2^n, 2^{n+1}] \neq \emptyset$,*

$$\Delta_{1,i} \leq \lambda_n := n^2 \mu[0, 2^n]^{1/\kappa} 2^{n/\kappa - n/2}.$$

Proof. A.s. for every n large enough and $i \in \mathbb{N}$ with $(Y_i, Y_{i+1}] \cap [2^n, 2^{n+1}] \neq \emptyset$, we have by Lemma 5.3.1 (d)

$$\mu(Y_i, Y_{i+1}] \leq \log(Y_{i+1})^2 / Y_{i+1} \leq n^2 / 2^n.$$

Then by Lemma 5.3.1 (c), then (a) and (b), a.s. for every n large enough,

$$\#\{i \in \mathbb{N}, (Y_i, Y_{i+1}] \cap [2^n, 2^{n+1}] \neq \emptyset\} \leq 2\mu[0, 2^{n+1}]2^{n+1} \leq 9\mu[0, 2^n]2^n.$$

Therefore, by Lemma 5.9.2 a.s. for every n large enough, writing E_n for the event that there exists $i \in \mathbb{N}$ such that $(Y_i, Y_{i+1}] \cap [2^n, 2^{n+1}] \neq \emptyset$, and such that $\Delta_{1,i} > \lambda_n$,

$$\mathbb{P}(E_n | \mathcal{F}) \leq 9C_\kappa \mu[0, 2^n]2^n \left(\frac{n^2/2^n}{n^4 \mu[0, 2^n]^{2/\kappa} 2^{2n/\kappa - n}} \right)^{\kappa/2} = 9C_\kappa n^{-\kappa}.$$

The desired result follows by the Borel–Cantelli Lemma. \square

Proof of Proposition 5.9.1. Recall that we need to prove that a.s. as $N \rightarrow \infty$, $\|\mathfrak{M}_N\|_\infty \rightarrow 0$. First, by Lemma 5.3.2 a.s. for every $n \in \mathbb{N}$ large enough, every $x \in [0, 2^{n+1}]$ is separated from $[0, 2^n]$ by at most $4n$ branches. So, by Lemma 5.9.3 a.s. for every n large enough, for every $N \in \mathbb{N}$,

$$\max_{x \leq 2^{n+1}} \mathfrak{M}_N(x) \leq \max_{x \leq 2^n} \mathfrak{M}_N(x) + 4n\lambda_n$$

Hence, a.s. for every n large enough, for every $N \in \mathbb{N}$,

$$\max_{x \in \mathbb{R}^+} \mathfrak{M}_N(x) \leq \max_{x \leq 2^n} \mathfrak{M}_N(x) + \sum_{i=n}^{\infty} 4i\lambda_i. \quad (5.51)$$

Then note that for every $i \in \mathbb{N}$, $(\Delta_{N,i})_{N \in \mathbb{N}}$ is decreasing. So by Lemma 5.9.2, since μ is a.s. locally finite, a.s. as $N \rightarrow \infty$, $\Delta_{N,i} \rightarrow 0$. Therefore a.s. for every $n \in \mathbb{N}$, as $N \rightarrow \infty$,

$$\max_{x \leq 2^n} \mathfrak{M}_N(x) \leq \sum_{i: Y_i \leq 2^n} \Delta_{N,i} \rightarrow 0. \quad (5.52)$$

Moreover by Lemma 5.3.1 (a), and by Assumption 5.2, a.s. as $n \rightarrow \infty$,

$$\sum_{i \geq n} 4i\lambda_i = \sum_{i \geq n} 4i^3 \mu[0, 2^i]^{1/\kappa} 2^{i/\kappa - i/2} \sim \sum_{i \geq n} 4i^3 \mathbb{E}[\mu[0, 2^i]]^{1/\kappa} 2^{i/\kappa - i/2} \rightarrow 0. \quad (5.53)$$

To conclude the proof, by (5.51), (5.52), (5.53) almost surely as $N \rightarrow \infty$, $\|\mathfrak{M}_N\|_\infty \rightarrow 0$. \square

5.10 Appendix A: Proof of Lemma 5.2.7, and measurability

Since the proofs of this section consist in checking some list of cases, we strongly advise the reader to draw those different cases to make the reading less tedious. Also, we use the symbol \ddagger to mean that we proved that a case is absurd. Recall Sections 5.2.1, 5.2.5. Let $(\mathcal{T}, d, \rho, u)$ denote a plane \mathbb{R} -tree. Let $\mathcal{L} := \mathcal{T} \times [0, 1]$. Let ν denote a σ -finite measure on \mathcal{T} .

First \prec is reflexive since for every $\alpha \in \mathcal{T} \times [0, 1]$, $\alpha \prec \alpha$. Also all elements are comparable: For every $\alpha = (x, a), \beta = (y, b) \in \mathcal{T} \times [0, 1]$ such that neither $\alpha \prec \beta$ nor $\beta \prec \alpha$, $U_{x \wedge y, x, y} = U_{x \wedge y, x, y}$. So $x = x \wedge y$ or $y = x \wedge y$, since otherwise x, y would be connected in $\mathcal{T} \setminus \{x \wedge y\}$. In the first case $\alpha \prec \beta$. In the second case, $\beta \prec \alpha$.

Proof that \prec is antisymmetric. Let $\alpha = (x, a), \beta = (y, b) \in \mathcal{T} \times [0, 1]$. Assume that $\alpha \prec \beta$ and $\beta \prec \alpha$, let us prove that $\alpha = \beta$. We have either:

- $\alpha \curvearrowright \beta, \beta \curvearrowright \alpha$: So $u_{x \wedge y, x, a} < u_{x \wedge y, y, b}$ and $u_{x \wedge y, y, b} < u_{x \wedge y, x, a}$. \ddagger
- $\alpha \curvearrowright \beta, \beta \triangleleft \alpha$: By $\beta \nabla \alpha, y \triangleleft x, u_{y, x, a} = b$. Then by $x \wedge y = y$ and $\alpha \curvearrowright \beta, u_{y, x, a} < u_{y, y, b}$. \ddagger
- $\alpha \triangleleft \beta, \beta \curvearrowright \alpha$: This case is similar by symmetry. \ddagger
- $\alpha \triangleleft \beta, \beta \triangleleft \alpha$: $x \in \llbracket 0, y \rrbracket$ and $y \in \llbracket 0, x \rrbracket$ so $x = y$. Moreover, $a = u_{x, y, b} = b$. \square

Toward proving that \prec is transitive, it is enough to show that:

Lemma 5.10.1. *For every $\alpha = (x, a), \beta = (y, b), \gamma = (z, c) \in \mathcal{T} \times [0, 1]$:*

- (a) *If $\alpha \curvearrowright \beta, \beta \curvearrowright \gamma$, then $\alpha \curvearrowright \gamma$.*
- (b) *If $\alpha \curvearrowright \beta, \beta \triangleleft \gamma$, then $\alpha \curvearrowright \gamma$.*
- (c) *If $\alpha \triangleleft \beta, \beta \curvearrowright \gamma$, then $\alpha \curvearrowright \gamma$ or $\alpha \triangleleft \gamma$.*
- (d) *If $\alpha \triangleleft \beta, \beta \triangleleft \gamma$, then $\alpha \triangleleft \gamma$.*

Proof. Toward (a) ($\alpha \curvearrowright \beta, \beta \curvearrowright \gamma$) note that $x \wedge y, y \wedge z \in \llbracket 0, y \rrbracket$, so we have either:

- $x \wedge y \in \llbracket 0, y \wedge z \rrbracket$: So $x \wedge z = x \wedge y$. Also, y and z are connected in $\mathcal{T} \setminus \{x \wedge y\}$, so $u_{x \wedge y, y, b} = u_{x \wedge y, z, c}$. Then by $\alpha \curvearrowright \beta, u_{x \wedge y, x, a} < u_{x \wedge y, y, b}$. So $u_{x \wedge z, x, a} < u_{x \wedge z, z, c}$ and $x \curvearrowright z$.
- $y \wedge z \in \llbracket 0, x \wedge y \rrbracket$: Similarly, $x \wedge z = y \wedge z$, and $u_{y \wedge z, x, a} = u_{y \wedge z, y, b}$. Then by $\beta \curvearrowright \gamma$, we have $u_{y \wedge z, y, b} < u_{y \wedge z, z, c}$. So $u_{x \wedge z, x, a} < u_{x \wedge z, z, c}$ and $x \curvearrowright z$.
- $x \wedge y = y \wedge z$:
 - Either $x \wedge z = x \wedge y$, and then $u_{x \wedge y, x, a} < u_{x \wedge y, y, b}$ and $u_{y \wedge z, y, b} < u_{y \wedge z, z, c}$. So $u_{x \wedge z, x, a} < u_{x \wedge z, z, c}$. Therefore $x \curvearrowright z$.
 - Or $x \wedge z \neq x \wedge y$, so x and z are connected in $\mathcal{T} \setminus \{x \wedge y\}$. So $u_{x \wedge y, x, a} = u_{x \wedge y, z, c}$. However, by $\alpha \curvearrowright \beta, u_{x \wedge y, x, a} < u_{x \wedge y, y, b}$ and by $\beta \curvearrowright \gamma, u_{y \wedge z, y, b} < u_{y \wedge z, z, c}$. \ddagger

Toward (b), ($\alpha \curvearrowright \beta$, and $\beta \triangleleft \gamma$). We have either:

- $x = y$: By $\alpha \curvearrowright \beta, a < b$. Also, by $\beta \triangleleft \gamma, y \in \llbracket 0, z \rrbracket$, and $u_{y, z, c} = b > a$. Thus, $\alpha \curvearrowright \gamma$.
- $x \neq y, x \wedge y = y$: By $\alpha \curvearrowright \beta, u_{y, x, a} < u_{y, y, b} = b$. Also, by $\beta \triangleleft \gamma, u_{y, z, c} = b$. Thus, x and z are in different components of $\mathcal{T} \setminus \{y\}$, so $x \wedge z = y$. Hence, $\alpha \curvearrowright \gamma$.
- $x \neq y, x \wedge y \neq y$: $y \triangleleft z$ so y and z are connected in $\mathcal{T} \setminus \{x \wedge y\}$, thus $x \wedge y = x \wedge z$ and $u_{x \wedge y, y, b} = u_{x \wedge y, z, c}$. Then by $\alpha \curvearrowright \beta, u_{x \wedge y, x, a} < u_{x \wedge y, y, b}$. Hence, $\alpha \curvearrowright \gamma$.

Toward (c), we have $\alpha \triangleleft \beta$, and $\beta \curvearrowright \gamma$ so $x, y \wedge z \in \llbracket 0, y \rrbracket$. Then we have either:

- $y \wedge z \in \llbracket x, y \rrbracket$: So y, z are connected in $\mathcal{T} \setminus \{x\}$. Thus $u_{x, z, c} = u_{x, y, b}$. Then by $\alpha \triangleleft \beta, u_{x, y, b} = a$. Hence $\alpha \triangleleft \gamma$.
- $y \wedge z = x$: $x \wedge z = x$. By $\alpha \triangleleft \beta, u_{x, y, b} = a$, and by $\beta \curvearrowright \gamma, u_{x, y, b} < u_{x, z, c}$. So $\alpha \curvearrowright \gamma$.
- $y \wedge z \in \llbracket 0, x \rrbracket$: Then by $x \in \llbracket 0, y \rrbracket, x \wedge z = y \wedge z$, and x and y are connected in $\mathcal{T} \setminus \{y \wedge z\}$. Thus $u_{y \wedge z, x, a} = u_{y \wedge z, y, b}$. Then by $\beta \curvearrowright \gamma, u_{y \wedge z, y, b} < u_{y \wedge z, z, c}$. Hence, $x \curvearrowright z$.

Toward (d), ($\alpha \triangleleft \beta, \beta \triangleleft \gamma$), we have either:

- $x = y$: By $\alpha \triangleleft \beta, b = u_{x, y, b} = a$ so $\alpha = \beta$ and $\alpha \triangleleft \gamma$.
- $x \neq y$: By $x \triangleleft y$ and $y \triangleleft z, y$ and z are connected in $\mathcal{T} \setminus \{x\}$. So $u_{x, z, c} = u_{x, y, b} = a, \alpha \triangleleft \gamma$. \square

Also by a simple symmetry argument, (change the angles by $u \mapsto 1 - u$) by Lemma 5.10.1 (b),

Lemma 5.10.2. *With the same notations, if $\beta \triangleleft \gamma, \beta \curvearrowright \alpha$ then $\gamma \curvearrowright \alpha$.*

Lemma 5.10.3. *$\nu_{\curvearrowright}, \nu_{\nabla}, \nu_{\triangleleft}$ are well defined.*

Proof. First let $\mathcal{M} := \{x \in \mathcal{T}, \deg(x) > 2\} \cup \{\rho\}$. \mathcal{M} is countable since \mathcal{T} is separable. Then for every $x \in \mathcal{T}$, and $a \in [0, 1]$, let

$$\mathcal{C}_{x, a} := \{y \in \mathcal{T} \setminus \{x\}, u_{x, y} = a\}.$$

This set is open so measurable. Also, for every $x \in \mathcal{T}$, and $a \in]0, 1[$ let $\mathcal{C}_{x, \leq a} := \bigcup_{b \leq a, b \neq 0} \mathcal{C}_{x, b}$. Define similarly, $\mathcal{C}_{x, < a}, \mathcal{C}_{x, \geq a}, \mathcal{C}_{x, > a}$. Since \mathcal{T} is separable all those sets are a countable union of open set and so are measurable.

Then by simply rewriting the definition of \curvearrowright , for every $\alpha = (x, u) \in \mathcal{L}$,

$$\begin{aligned} \{\beta \in \mathcal{L}, \beta \curvearrowright \alpha\} = & \left(\bigcup_{z \in \mathcal{M} \cap]0, x[} \mathcal{C}_{z, < u_{z, x}} \times [0, 1] \right) \cup \left(\bigcup_{z \in \mathcal{M} \cap]0, x[} \{z\} \times [0, u_{z, x}) \right) \\ & \cup (([0, x[\setminus \mathcal{M}) \times [0, 1/2)) \cup (\{x\} \times [0, u)) \cup \mathcal{C}_{x, < u}. \end{aligned}$$

So $\{\beta \in \mathcal{L}, \beta \curvearrowright \alpha\}$ is measurable as a countable union of measurable set. Similarly,

$$\begin{aligned} \{\beta \in \mathcal{L}, \alpha \curvearrowright \beta\} = & \left(\bigcup_{z \in \mathcal{M} \cap]0, x[} \mathcal{C}_{z, > u_{z, x}} \times [0, 1] \right) \cup \left(\bigcup_{z \in \mathcal{M} \cap]0, x[} \{z\} \times (u_{z, x}, 1] \right) \\ & \cup (([0, x[\setminus \mathcal{M}) \times (1/2, 1]) \cup (\{x\} \times (u, 1]) \cup \mathcal{C}_{x, > u}, \end{aligned}$$

is also measurable. Finally, $\{\beta \in \mathcal{L}, \alpha \triangleleft \beta\} = \mathcal{C}_{x, u} \cup \{(x, u)\}$ is measurable. \square

Also for every $\alpha \in \mathcal{L}$,

$$\{\gamma \in \mathcal{L} : \gamma \triangleleft \alpha\} = \left(\bigcup_{z \in \mathcal{M} \cap]0, x[} \{z, u_{z, x}\} \right) \cup (([0, x[\setminus \mathcal{M}) \times \{1/2\})$$

is measurable.

Lemma 5.10.4. $\{\alpha, \beta \in \mathcal{L}, \alpha \triangleleft \beta\}, \{\alpha, \beta \in \mathcal{L}, \alpha \curvearrowright \beta\}$ are measurable sets for the product topology.

Proof. First note that $\bigcup_{x, y \in \mathcal{T}}]x, y[$ is dense on \mathcal{T} , so since \mathcal{M} is countable, $\{x \in \mathcal{T}, \deg(x) = 2\}$ is dense. Then since \mathcal{T} is separable, there is a countable dense set \mathcal{N} of \mathcal{T} , and we may assume that $\mathcal{N} \subset \{x \in \mathcal{T}, \deg(x) = 2\}$. Moreover, note that for every $x \neq y \in \mathcal{T}$, $]x, y[\cap (\mathcal{M} \cup \mathcal{N}) \neq \emptyset$.

Then, let for $z \in \mathcal{T}$, $\mathcal{C}'_{z, 0} := \{(x, a) \in \mathcal{T} \times [0, 1] : (x, a) \triangleleft (z, 0), x \neq z\}$. The set

$$\{\alpha, \beta \in \mathcal{L}, \alpha \triangleleft \beta\} = \{\{x, a\} \times \{x, a\}, x \in \mathcal{T}, a \in [0, 1]\} \cup \left(\bigcup_{z \in \mathcal{M} \cup \mathcal{N}} \mathcal{C}'_{z, 0} \times \mathcal{C}_{z, \leq 1} \right),$$

is measurable as a countable union of measurable set.

Also, by considering several cases depending on the position of $x \wedge y$, we get: (The two last cases are here to deal with the cases where $x \in]0, y[$ or $y \in]0, x[$.)

$$\begin{aligned} \{\alpha, \beta \in \mathcal{L}, \alpha \curvearrowright \beta\} = & \{\{x, u\} \times \{x, v\}, x \in \mathcal{T}, u < v\} \cup \left(\bigcup_{z \in \mathcal{M}, u \in \mathbb{Q}} \mathcal{C}_{z, < u} \times \mathcal{C}_{z, > u} \right) \\ & \cup \left(\bigcup_{z \in \mathcal{N}} \{\alpha, \alpha \curvearrowright (z, 1/2)\} \times \{\beta, (z, 1/2) \triangleleft \beta\} \right) \cup \left(\bigcup_{z \in \mathcal{N}} \{\alpha, (z, 1/2) \triangleleft \alpha\} \times \{\beta, (z, 1/2) \curvearrowright \beta\} \right). \end{aligned}$$

Thus $\{\alpha, \beta \in \mathcal{L}, \alpha \curvearrowright \beta\}$ is measurable as a countable union of measurable set. \square

5.11 Appendix B: Hölder continuity of a Gaussian free field.

If (X, d) is a metric space, we call a Gaussian free field on X , a random function $\mathfrak{F} : X \mapsto \mathbb{R}^+$ such that for every $x \in X$, $\mathfrak{F}(x)$ is measurable, and such that for every $x, y \in X$, $\mathfrak{F}(x) - \mathfrak{F}(y)$ is a Gaussian random variable with variance $d(x, y)$.

Lemma 5.11.1. *Let (X, d) be a metric space. Let \mathfrak{F} be a Gaussian free field on X . If \mathfrak{F} is almost surely continuous and (X, d) have finite upper Minkowski dimension then almost surely \mathfrak{F} is Hölder continuous with any exponent smaller than $1/2$.*

Proof. Since (X, d) have finite upper Minkowski dimension, there exists $k \in \mathbb{N}$ such that for every $n \in \mathbb{N}$, there exists $a_{n,1}, \dots, a_{n,k^n} \in X$ such that $\max_{x \in X} d(x, \{a_{n,1}, \dots, a_{n,k^n}\}) \leq 2^{-n}$.

For every $n \in \mathbb{N}$, let E_n denote the following event:

$$E_n := \left\{ \forall i \leq k^n, \forall j \leq k^{n+1}, d(a_{n,i}, a_{n+1,j}) \leq 2^{9-n} \Rightarrow |\mathfrak{F}(a_{n,i}) - \mathfrak{F}(a_{n+1,j})| \leq kn2^{-n/2} \right\}.$$

By an upper bound, and by definition of a Gaussian free field, we have for every $n \in \mathbb{N}$,

$$\mathbb{P}(E_n) \leq k^n k^{n+1} e^{-(8kn2^{-n/2})^2 / (2 \cdot 2^{9-n})} = k^{2n+1} e^{-k^2 n^2 / 2^{18}}.$$

Thus, since the right hand side is summable, by the Borel–Cantelli Lemma, almost surely for every $n \in \mathbb{N}$ large enough we have E_n .

The rest of the proof is deterministic. Assume that there exists $N \in \mathbb{N}$ such that for every $n \geq N$ we have E_n , and that \mathfrak{F} is continuous. Let $x, y \in X$ such that $d(x, y) \leq 2^{-N-9}$. Let $n \in \mathbb{N}$ such that $2^{-n+1} < d(x, y) \leq 2^{-n}$. Since \mathfrak{F} is continuous we can consider $m \in \mathbb{N}$, $m > n$ such that for every $z \in X$ with $d(x, z) \leq 2^{-m}$, $|\mathfrak{F}(x) - \mathfrak{F}(z)| \leq 2^{-n/2}$, and similarly for y . Let by induction $x_{m+1}, x_m, x_{m-1}, \dots, x_n$ such that $x_{m+1} = x$, and such that for every $m \geq i \geq n$, $x_i \in \{a_{i,1}, \dots, a_{i,k^i}\}$, and $d(x_i, x_{i+1}) \leq 2^{-i}$. Define similarly y_{m+1}, \dots, y_n . We have

$$d(x_n, y_{n+1}) \leq \sum_{i=n}^m d(x_{i+1}, x_i) + d(x, y) + \sum_{i=n+1}^m d(y_i, y_{i+1}) \leq \sum_{i=n}^m 2^{-i} + 2^{-n} + \sum_{i=n+1}^m 2^{-i} \leq 2^{2-n}.$$

So by the triangular inequality, then by definition of m and by E_n, E_{n+1}, \dots, E_m ,

$$\begin{aligned} |\mathfrak{F}(x) - \mathfrak{F}(y)| &\leq |\mathfrak{F}(x_{m+1}) - \mathfrak{F}(x_m)| + \sum_{i=n}^{m-1} |\mathfrak{F}(x_{i+1}) - \mathfrak{F}(x_i)| \\ &\quad + |\mathfrak{F}(x_n) - \mathfrak{F}(y_{n+1})| + \sum_{i=n+1}^{m-1} |\mathfrak{F}(y_i) - \mathfrak{F}(y_{i+1})| + |\mathfrak{F}(y_m) - \mathfrak{F}(y_{m+1})| \\ &\leq 2^{-n/2} + \sum_{i=n}^{m-1} ki2^{-i/2} + kn2^{-n/2} + \sum_{i=n+1}^{m-1} ki2^{-i/2} + 2^{-n/2} \\ &\leq 7^{71} kn^7 2^{-n/2} \\ &\leq 9^{91} k \frac{d(x, y)^{1/2}}{\log(d(x, y))^7}, \end{aligned}$$

using $d(x, y) \geq 2^{-n+1}$ for the last inequality. So, since (x, y) are arbitrary with $d(x, y) \leq 2^{-N-9}$, \mathfrak{F} is almost surely Hölder continuous with any exponent smaller than $1/2$. \square

5.12 Appendix C: Supremum of empirical process.

Lemma 5.12.1. *Let $n \in \mathbb{N}$, $\kappa \geq 2$, $(x_i)_{i \leq n} \in \mathbb{R}^n$. Let $(X_i)_{i \leq n}$ be a family of independent centered, random variables. For every $k \leq n$ let $S_k := x \in \mathbb{R} \mapsto \sum_{i \leq k} \mathbf{1}_{x \leq x_i} X_i$. Assume that for every $i \leq n$, $v_i := \mathbb{E}[X_i^\kappa]^{2/\kappa} < \infty$, and let $V := \sum_{i \leq n} v_i$. Then for every $t > 0$,*

$$P_n := \mathbb{P} \left(\sup_{i \leq n} \|S_i\|_\infty > t \right) \leq C_\kappa \left(\sqrt{V}/t \right)^\kappa,$$

where C_κ is a constant which depends only on κ .

Remark. By taking $n \rightarrow \infty$, the previous lemma also holds when $n = \infty$.

Proof. Fix $\kappa \geq 2$. We work by induction with trivial initialization $n = 0$. Fix $C \in \mathbb{R}^+$, which we chose later. Assume that Lemma 5.12.1 holds for every $m < n$ with the constant C . Let $\varepsilon > 0$, which we chose later. Let $k \in \mathbb{N}$ be the smallest integer such that $\sum_{i \leq k} v_i \geq \frac{1}{2} \sum_{i \leq n} v_i$. Note that,

$$P_n \leq \mathbb{P} \left(\sup_{i < k} \|S_i\|_\infty > t \right) + \mathbb{P} (\|S_k\|_\infty > \varepsilon t) + \mathbb{P} \left(\sup_{k < i \leq n} \|S_i - S_k\|_\infty > (1 - \varepsilon)t \right).$$

Then by induction, and by definition of k ,

$$P_n \leq C \left(\sqrt{V/2}/t \right)^\kappa + \mathbb{P} (\|S_k\|_\infty > \varepsilon t) + C \left(\sqrt{V/2}/((1 - \varepsilon)t) \right)^\kappa. \quad (5.54)$$

It remains to estimate $\|S_k\|_\infty$. To this end, first note that $(|S_k(x)|^\kappa)_{x \in \mathbb{R}}$ is a sub-martingale, so by Doob's inequality,

$$\mathbb{P} (\|S_k\|_\infty > \varepsilon t) \leq \max_{x \in \mathbb{R}} \frac{\mathbb{E}[|S_k(x)|^\kappa]}{(\varepsilon t)^\kappa} = \frac{\mathbb{E}[|\sum_{i=1}^k X_i|^\kappa]}{(\varepsilon t)^\kappa}. \quad (5.55)$$

Furthermore by Marcinkiewicz's inequality (see [32] Section 15.4), writing $c_\kappa := 2^{\kappa+1} (2\kappa)^{\kappa/2}$,

$$\mathbb{E} \left[\left| \sum_{i=1}^k X_i \right|^\kappa \right] \leq c_\kappa \mathbb{E} \left[\left(\sum_{i=1}^k X_i^2 \right)^{\kappa/2} \right],$$

and by Minkowski's inequality

$$\mathbb{E} \left[\left(\sum_{i=1}^k X_i^2 \right)^{\kappa/2} \right] \leq c_\kappa \left(\sum_{i=1}^k \mathbb{E} \left[(X_i^2)^{\kappa/2} \right]^{2/\kappa} \right)^{\kappa/2} = c_\kappa V^{\kappa/2}.$$

Therefore, by (5.54) and (5.55),

$$P_n \leq \left(C \frac{1 + 1/(1 - \varepsilon)^\kappa}{2^\kappa} + c_\kappa / \varepsilon^\kappa \right) \left(\sqrt{V}/t \right)^\kappa.$$

Finally some quick computations show that there exists C, ε such that $(C \frac{1 + 1/(1 - \varepsilon)^\kappa}{2^\kappa} + c_\kappa / \varepsilon^\kappa) \leq C$, and the desired result follows by induction. \square

Remark. If $\kappa \geq 4$, one may chose $\varepsilon = 1/3$, $C_\kappa = 2.3^\kappa c_\kappa \leq 4.9^\kappa \kappa^{\kappa/2}$.

We now prove an extension of Lemma 5.12.1, that will actually be used in [26] to study non locally centered fennec on \mathcal{D} -trees. We could also have used it to further extend Proposition 5.9.1.

Lemma 5.12.2. *Let $(\lambda_i)_{i \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. Let $(X_i)_{i \in \mathbb{N}}$ be independent real random variables with density $(f_i)_{i \in \mathbb{N}}$. Let $a \leq b$. Let $S : x \mapsto \sum_{i=1}^{\infty} \lambda_i (\mathbf{1}_{X_i \leq x} - \mathbb{P}(X_i \leq x))$. Assume that the previous series converges uniformly on $[a, b]$.*

(a) *Assume that $(\lambda_i)_{i \in \mathbb{N}}$ is decreasing. Let,*

$$C := \sum_{q=1}^{\infty} |\lambda_{2^q}| \quad ; \quad D := 2 \log(2) \sum_{q=1}^{\infty} q |\lambda_{2^q}|$$

$$W := \sum_{i=1}^{\infty} \lambda_i^2 (b-a) \max_{a \leq x \leq b} f_i(x) \quad ; \quad E := \sum_{i=1}^{\infty} \frac{|\lambda_i|}{i} (b-a) \max_{a \leq x \leq b} f_i(x).$$

For every $t > 0$,

$$\mathbb{P} \left(\sup_{a \leq x \leq b} |S(x) - S(a)| > 6\sqrt{W}t + (t+1)C + D + 2E \right) \leq e^{3-t}.$$

(b) *Let,*

$$V := \sum_{i=1}^{\infty} \lambda_i^2 \quad ; \quad M := (b-a) \max_{i \in \mathbb{N}} \max_{a \leq x \leq b} f_i(x).$$

For every $t > 0$,

$$\mathbb{P} \left(\sup_{a \leq x \leq b} |S(x) - S(a)| > 7\sqrt{V}(t+M) \right) \leq e^{3-t}.$$

Remarks. • Although the bound is written in a general setting, the bound is actually bad to deal with the i such that $|\lambda_i|$ is large and $\mathbb{P}(X_i \in [a, b])$ is small. To avoid this issue, we advise to replace, up to reordering, $(\lambda_i)_{i \in \mathbb{N}}$ by $(\lambda_i \mathbf{1}_{X_i \in [a, b]})_{i \in \mathbb{N}}$, and have some a priori bounds on C, D, W, E or V, M .

Proof. We first explain why (b) follows from (a). In (b) we may assume up to reordering that $(\lambda_i)_{i \in \mathbb{N}}$ is decreasing. First $\sqrt{W}t \leq \sqrt{MV}t \leq \sqrt{V}(t+M)/2$. Then by Cauchy–Schwartz’s inequality, $E \leq M\sqrt{V} \left(\sum_{i=1}^{\infty} i^{-2} \right)^{1/2}$. Also by Cauchy–Schwartz inequality,

$$C \leq |\lambda_1| + |\lambda_2| + \frac{1}{2}(|\lambda_3| + |\lambda_4|) + \dots \leq \sqrt{V} \left(1 + \sum_{q=0}^{\infty} 2^q \cdot (1/2^q)^2 \right)^{1/2} = \sqrt{3V}$$

Similarly, $D \leq 3\sqrt{V}$. Also we may assume $t > 3$ since otherwise the result is trivial so $D \leq t\sqrt{V}$. (b) follows by sum.

Toward (a), for every $i \in \mathbb{N}$, let $U_i : x \mapsto \lambda_i (\mathbf{1}_{x \geq X_i} - \mathbb{P}(X_i \leq x))$. For every $I \subset \mathbb{N}$ let $S_I = \sum_{i \in I} U_i$. Also for every $x \in \mathbb{N}$ let $S_{<x} := \sum_{i < x} U_i$ and let $S_{\geq x} := \sum_{i \geq x} U_i$. Note that for every $I \subset \mathbb{N}$, $x, y \in \mathbb{R}^+$, $S_I(x) - S_I(y)$ is a sum of independent centered random variables. Also for every $i \geq 0$, $x \leq y \in [a, b]$, $q \geq 2$,

$$\mathbb{E}[|\lambda_i(U_i(x) - U_i(y))|^q] \leq |\lambda_i|^q \mathbb{P}(x \leq X_i \leq y) \leq |\lambda_i|^q (y-x) \max_{z \in [a, b]} f_i(z).$$

So by Bernstein’s inequality (see e.g. [32] Chapter 2.8) for every $I \subset \mathbb{N}$, $x \leq y \in [a, b]$, $t > 0$,

$$\mathbb{P} \left(|S_I(y) - S_I(x)| > \sqrt{2(y-x)/(b-a)W}t + t \max_{i \in I} |\lambda_i| \right) \leq 2e^{-t}. \quad (5.56)$$

Finally, since for every $x \in \mathbb{R}$, $S_I(x)$ is $(X_i)_{i \in I}$ measurable, by Fubini's theorem, (5.56) also holds if x, y are random variables in $[a, b]$ independent with $(X_i)_{i \in I}$.

Next, for every $q \geq 0$ let,

$$\Gamma_q := \{a + (b - a)k2^{-q}, 0 \leq k < 2^q\} \cup (\{X_i, 1 \leq i < 2^q\} \cap [a, b]),$$

and let $p_q : x \in [a, b] \mapsto \max\{z \in \Gamma_q, x > z\}$. Fix $0 \leq q \leq m$, $x \in [a, b]$. We have,

$$S_{<2^q}(x) - S_{<2^q}(p_q(x)) = \sum_{i < 2^q} \lambda_i (\mathbf{1}_{X_i \leq x} - \mathbb{P}(X_i \leq x)) - \sum_{i < 2^q} \lambda_i (\mathbf{1}_{X_i \leq p_q(x)} - \mathbb{P}(X_i \leq p_q(x)|p_q(x))).$$

Then, since for every $i < 2^q$, $X_i \notin (p_q(x), x]$,

$$S_{<2^q}(x) - S_{<2^q}(p_q(x)) = \sum_{i < 2^q} \lambda_i (\mathbb{P}(X_i \leq p_q(x)|p_q(x)) - \mathbb{P}(X_i \leq x)).$$

Thus, by definition of $(f_i)_{i \in \mathbb{N}}$, and since $|x - p_q(x)| \leq 2^{-q}(b - a)$,

$$|S_{<2^q}(x) - S_{<2^q}(p_q(x))| \leq \sum_{i < 2^q} 2^{-q}(b - a)|\lambda_i| \max_{a \leq y \leq b} f_i(y). \quad (5.57)$$

Hence, writing δ_q for the right hand side above,

$$|S(x) - S(p_q(x))| \leq \delta_q + |S_{\geq 2^q}(x) - S_{\geq 2^q}(p_q(x))|. \quad (5.58)$$

Now fix $t > 0$. For all $q \geq 0$, $i \geq 2^q$, note that $X_i, p_q(X_i)$ are independent with $(X_i)_{i \in \{2^q, \dots, n\} \setminus \{i\}}$, and that $X_i \in [a, b]$ implies $|p_q(X_i) - X_i| \leq 2^{-q}(b - a)$. So by (5.56),

$$\mathbb{P}\left(X_i \in [a, b], |S_{\{2^q, 2^q+1, \dots\} \setminus \{i\}}(p_q(X_i)) - S_{\{2^q, 2^q+1, \dots\} \setminus \{i\}}(X_i)| > \sqrt{2^{1-q}Wt} + t|\lambda_{2^q}|\right) \leq 2e^{-t}.$$

Then, note that

$$||S_{\{2^q, 2^q+1, \dots\} \setminus \{i\}}(p_q(X_i)) - S_{\{2^q, 2^q+1, \dots\} \setminus \{i\}}(X_i)| - |S_{\geq 2^q}(p_q(X_i)) - S_{\geq 2^q}(X_i)|| \leq |\lambda_i| \leq |\lambda_{2^q}|.$$

So, using (5.58),

$$\mathbb{P}\left(X_i \in [a, b], |S(p_q(X_i)) - S(X_i)| > \sqrt{2^{1-q}Wt} + (t+1)|\lambda_{2^q}| + \delta_q\right) \leq 2e^{-t}.$$

We get a similar bound by replacing X_i above by $x \in [a, b]$ determinist. Hence, by an union bound,

$$\mathbb{P}\left(\exists x \in \Gamma_q, |S(p_q(x)) - S(x)| > \sqrt{2^{1-q}Wt} + (t+1)|\lambda_{2^q}| + \delta_q\right) \leq 2^{q+2}e^{-t}.$$

So by writing $t_q := \sqrt{2^{1-q}W(t + 2q \log(2))} + (t + 2q \log(2) + 1)|\lambda_{2^q}| + \delta_q$,

$$\mathbb{P}(\exists x \in \Gamma_q, |S(p_q(x)) - S(x)| > t_q) \leq 2^{2-q}e^{-t}.$$

Finally by an union bound, we have the following event with probability at least $1 - 2^4e^{-t}$:

$$\forall q \geq 0, \forall x \in \Gamma_q, |S(p_q(x)) - S(x)| \leq t_q. \quad (5.59)$$

It remains to estimate $\sup_{a \leq x \leq y \leq b} |S(x) - S(y)|$ under (5.59), which we now assume. First since $\Gamma_0 \subset \Gamma_1 \subset \dots \subset \Gamma_m$, for every $q \leq r$, $p_q \circ p_r = p_q$. So for every $x \in \Gamma_m$, by (5.59),

$$|S(x) - S(a)| \leq \sum_{q=0}^{m-1} |S(p_{q+1}(x)) - S(p_q(x))| \leq \sum_{q=0}^{m-1} t_q.$$

Also, since $S(x)$ is a.s. càdlàg (right continuous with left limit) and continuous at b , and since $\Gamma := \bigcup_{q \geq 0} \Gamma_q$ is dense on $[a, b]$,

$$\max_{x \in [a, b]} |S(x) - S(a)| \leq \max_{x \in \Gamma} |S(x) - S(a)|.$$

Therefore it remains to estimate $\sum_{q \geq 0} t_q$. We have with elementary computations:

$$\sum_{q \geq 0} \delta_q = \sum_{q \geq 0} \sum_{i < 2^q} 2^{-q} (b-a) |\lambda_i| \max_{a \leq y \leq b} f_i(y) \leq (b-a) \sum_{i \geq 1} \frac{2}{i} |\lambda_i| \max_{a \leq y \leq b} f_i(y) = 2E.$$

Then,

$$\sum_{q \geq 0} (1 + (t + 2q \log(2))) |\lambda_{2^q}| = (t+1)C + D$$

Also, note that the result is trivial if $t < 3$ so we may assume $t \geq 3$. Then, with basic analysis,

$$\sum_{q \geq 0} \sqrt{2^{1-q}(t + 2q \log(2))} \leq \sqrt{t} \sum_{q \geq 0} \sqrt{2^{1-q}(3 + 2q \log(2))/3} \leq 6\sqrt{t}$$

The desired result follows by sum. □

Lemma 5.12.3. *We keep the notations of Lemma 5.12.2. Let $n \in \mathbb{N}$. For every $k \leq n$, let*

$$S_k : x \mapsto \sum_{i=1}^k \lambda_i (\mathbf{1}_{X_i \leq x} - \mathbb{P}(X_i \leq x)).$$

For every $t > 0$,

$$P_n := \mathbb{P} \left(\sup_{k \leq n} \sup_{a \leq x \leq b} |S_k(x) - S_k(a)| > 42\sqrt{V}(t + M) \right) \leq e^{6-t}.$$

Remark. By taking $n \rightarrow \infty$, the previous result also holds if $n = \infty$.

Proof. The proof is similar to the proof of Lemma 5.12.1. We work by induction with trivial initialization $n = 0$. Assume that Lemma 5.12.3 holds for every $m < n$. For every real function f let $\|f\|_{[a,b]}$ denote $\max_{a \leq x \leq b} |f(x) - f(a)|$. Fix $t > 0$. We may assume $t > 6$ since otherwise the result is trivial. Let $U = 7\sqrt{V}(t + M)$. Let $k \in \mathbb{N}$ be the smallest integer such that $\sum_{i \leq k} v_i \geq \frac{1}{2} \sum_{i \leq n} v_i$. Note that,

$$P_n \leq \mathbb{P} \left(\sup_{i < k} \|S_i\|_{[a,b]} > 5U \right) + \mathbb{P} (\|S_k\|_{[a,b]} > U) + \mathbb{P} \left(\sup_{k < i \leq n} \|S_i - S_k\|_{[a,b]} > 5U \right).$$

Then, using $t > 6$, by elementary computations $5U \geq 42\sqrt{V/2}(t + 1 + M)$. Hence, by induction, by Lemma 5.12.2, and by definition of k ,

$$P_n \leq 2.e^{6-(t+1)} + e^{3-t} \leq e^{6-t}.$$

The desired result follows by induction. □

Bibliography

- [1] R. Abraham, J.-F. Delmas, and P. Hoscheit. A note on gromov-hausdorff-prokhorov distance between (locally) compact measure spaces. *Electron. J. Probab.*, 18(14, 21.), 2013.
- [2] L. Addario-Berry. Tails bounds for the height and width of a random tree with a given degree sequence. *Random structures and Algorithms*, 41:253–261, 2012.
- [3] L. Addario-Berry and J. Barrett. Random tree-weighted graphs. arXiv:2008.12167.
- [4] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probab. Theory Relat. Fields*, 152(3-4):367–406, 2012.
- [5] 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, 2017.
- [6] L. Addario-Berry, S. Donderwinkel, M. Maazoun, and J. Martin. A new proof of Cayley’s formula. arxiv:2107.09726, 2021.
- [7] L. Addario-Berry, L. Devroye, and S. Janson. Sub-gaussian tail bounds for the width and height of conditioned galton-watson trees. *Ann. Probab.*, 41:1072–187, 2013.
- [8] D. Aldous. The continuum random tree II. an overview? *Stochastic analysis*, 167:23–70, 1990.
- [9] D. Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *Siam J. Discrete Math.*, 3(4):450–465, 1990.
- [10] D. Aldous. The continuum random tree I. *Ann. Probab.*, 19:1–28, 1991.
- [11] D. Aldous. The continuum random tree III. *Ann. Probab.*, 21:248–289, 1993.
- [12] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25(2):812–854, 1997.
- [13] D. Aldous and V. Limic. The entranceboundary of the multiplicative coalescent. *Electron. J. Probab.*, 3(3), 1998.
- [14] D. Aldous, G. Miermont, and J. Pitman. The exploration process of inhomogeneous continuum random trees, and an extension of Jeulin’s local time identity. *Probab. Theory Related Fields*, 129(2):182–218, 2004.
- [15] D. Aldous, G. Miermont, and J. Pitman. Weak convergence of random p-mappings and the exploration process of inhomogeneous continuum random trees. *Probab. Theory Related Fields*, 133(1):1–17, 2005.
- [16] 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.

- [17] O. Amini, L. Devroye, S. Griffiths, and N. Olver. Explosion and linear transit times in infinite trees. *Probability Theory and Related Fields*, 167:325–347, 2017.
- [18] O. Angel, D. A. Croydon, S. Hernandez-Torres, and D. Shiraishi. Scaling limits of the three-dimensional uniform spanning tree and associated random walk. arXiv:2003.09055.
- [19] E. Archer, A. Nachmias, and M. Shalev. The GHP scaling limit of uniform spanning trees in high dimensions. arXiv:2112.01203, 2021.
- [20] E. Bender and E. Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A*, 24(3):296–307, 1978.
- [21] N. Berestycki. *Recent progress in coalescent theory*. Ensaios Matematicos, Volume 16, 1-193, 2009.
- [22] J. Bertoin. *Random fragmentation and coagulation processes*. Cambridge University Press, Cambridge, 2006.
- [23] S. Bhamidi, N. Broutin, S. Sen, and X. Wang. Scaling limits of random graph models at criticality: Universality and the basin of attraction of the Erdős-Rényi random graph. arXiv:1411.3417, 2014.
- [24] S. Bhamidi, R. Van Der Hofstad, and S. Sen. The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs. *Probab. Theory Relat. Fields*, 170:387–474, 2018.
- [25] A. Blanc-Renaudie. A few notes on ICRT excursion. (In preparation).
- [26] A. Blanc-Renaudie. Limit of looptree, fennec, and snake of trees with fixed degree sequence. (In preparation).
- [27] A. Blanc-Renaudie. Limit of connected multigraphs with fixed degree sequence. arXiv:2112.07725, dec 2021.
- [28] A. Blanc-Renaudie. Limit of trees with fixed degree sequence. arxiv:2110.03378, oct 2021.
- [29] A. Blanc-Renaudie. Compactness and fractal dimension of inhomogeneous continuum random trees. *Probab. Theory Relat. Fields*, 2022. <https://doi.org/10.1007/s00440-022-01138-9>.
- [30] A. Blanc-Renaudie. Looptree, Fennec, and Snake of ICRT. arXiv:2203.10891, March 2022.
- [31] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.
- [32] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities. A Nonasymptotic Theory of Independence*. Oxford university press, 2013.
- [33] J. Bouttier Bouttier, P. Di Francesco, and E. Guitter. Planar maps as labeled mobiles. *Electron. J. Combin.*, 11(1):Research Paper 69, 27, 2004.
- [34] T. Britton, M. Deijfen, and A. Martin-Löf. Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.*, 124(6):1377–1397, 2006.
- [35] A. Broder. Generating random spanning trees. In *Proc. 30'th IEEE Symp. Found. Comp. Sci*, pages 442–447, 1989.

- [36] N. Broutin, T. Duquesne, and M. Wang. Limits of multiplicative inhomogeneous random graphs and Lévy trees: The continuum graphs. *Ann. Appl. Probab.* (to appear) [arxiv.org:1804.05871](https://arxiv.org/abs/1804.05871).
- [37] N. Broutin, T. Duquesne, and M. Wang. Limits of multiplicative inhomogeneous random graphs and Lévy trees: Limit theorems. *PTRF*, 2021.
- [38] N. Broutin and J-F. Marckert. Asymptotics of trees with a prescribed degree sequence. *Random structures and Algorithms*, 44:290–316, 2014.
- [39] 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.
- [40] M. Camarri and J. Pitman. Limit distributions and random trees derived from the birthday problem with unequal probabilities. *Electron. J. Probab.*, 5(2), 2000.
- [41] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Ann. Comb.*, 6(2):125–145, 2002.
- [42] G. Conchon-Kerjan and C. Goldschmidt. The stable graph: the metric space of a critical random graph with i.i.d power-law degrees. [arxiv:2002.04954](https://arxiv.org/abs/2002.04954).
- [43] N. Curien and B. Haas. Random trees constructed by aggregation. *Annales de l’Institut Fourier*, 67(5):1963–2001, 2017.
- [44] N. Curien and I. Kortchemski. Random stable looptrees. *Electron. J. Probab.*, 19:no. 108, 35, 2014.
- [45] N. Curien, G. Miermont, and A. Riera. Private communication. In preparation.
- [46] S. Dhara. *Critical Percolation on Random Networks with Prescribed Degrees*. PhD thesis, Technische Universiteit Eindhoven, [arXiv:1809.03634](https://arxiv.org/abs/1809.03634), 2018.
- [47] S. Dhara, R. van der Hofstad, J. S.H. van Leeuwen, and S. Sen. Heavy-tailed configuration models at criticality. *Ann. Inst. H. Poincaré Probab. Statist.*, 56(3):1515 – 1558, August 2020.
- [48] T. Duquesne. The coding of compact real trees by real valued functions. [arXiv:0604106](https://arxiv.org/abs/0604106), 2006.
- [49] T. Duquesne and J-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Asterisque*, 281, 2002.
- [50] T. Duquesne and J-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Relat. Fields*, 131:553–603, 2005.
- [51] P. Erdős and P. Rényi. On random graphs i. *Publ. Math.*, pages 290–297, 1959.
- [52] S.N. Evans and J. Pitman. Construction of markovian coalescent. *Annales de l’I. H. P.*, 34:339–383, 1998.
- [53] K. Falconer. *Fractal Geometry. Mathematical Foundations and Applications*. Wiley, New York, 2nd edition, 2003.
- [54] C. Goldschmidt, B. Haas, and D. Sénizergues. Stable graphs: distributions and line-breaking construction. [arxiv.org:1811.06940](https://arxiv.org/abs/1811.06940).

- [55] S. Janson and S. O. Stefánsson. Scaling limits of random planar maps with a unique large face. *Ann. Probab.*, 43(3):1045–1081, 2015.
- [56] S. Janson. The loop-erased random walk and the uniform spanning tree on the four-dimensional discrete torus. *Probab. Theory Related Fields*, 144(3-4):319–370, 2009.
- [57] O. Kallenberg. *Random Measures, Theory and Applications*. Springer, 2010.
- [58] I. Kortchemski. Sub-exponential tail bounds for conditioned stable bienaymé–galton–watson trees. *Probab. Theory Relat. Fields*, 168:1–40, 2017.
- [59] G. Lawler, O. Schramm, and W. Werner. Conformal invariance of planar loop-erased random walks and uniform spanning trees [mr2044671]. In *Selected works of Oded Schramm. Volume I*, 2, Sel. Works Probab. Stat., pages 931–987. Springer, New York, 2011.
- [60] J-F. Le Gall. Spacial branching processes, random snakes and partial differential equations. In *Lectures in Mathematics ETH Zürich.*, Birkh"auser, Boston., 1999.
- [61] J-F. Le Gall. Random trees and applications. *Probability Surveys*, 2:245–311, 2005.
- [62] J-F. Le Gall. Uniqueness and universality of the Brownian map. *Ann. Probab.*, 41(4):2880–2960, 2013.
- [63] J-F. Le Gall and J-F. Le Jan. Branching processes in levy processes: Laplace functionals of snakes and superprocesses. *Ann. Probab.*, 26:1407–1432, 1998.
- [64] J-F. Le Gall and J-F. Le Jan. Branching processes in levy processes: The exploration process. *Ann. Probab.*, 26:213–252, 1998.
- [65] J-F. Le Gall and G. Miermont. Scaling limits of random trees and planar maps. *Probab and statistical physics in two or more dimensions*, 15:155–211, 2012.
- [66] W. Löhr. Equivalence of gromov-prokhorov and gromov’s \square_λ -metric on the space of metric measure spaces. *Electron. C. Probab.*, 26(1):213–252, 2013.
- [67] C. Marzouk. On scaling limits of random trees and maps with a prescribed degree sequence. arXiv:1903.06138.
- [68] C. Marzouk. Scaling limits of random bipartite planar maps with a prescribed degree sequence. *Random structures and Algorithms*, 3:448–503, 2018.
- [69] C. Marzouk. Scaling limits of random looptrees and bipartite plane maps with prescribed large faces. arXiv:2202.08666, 2022.
- [70] P. Mattila. *Geometry of Sets and Measures in Euclidian Spaces. Fractals and Rectifiability*. Cambridge studies in advanced mathematics, 1995.
- [71] G. Miermont. Aspects of random maps. Saint-Flour lecture notes.
- [72] M. Newman. The structure and function of complex networks. *SIAM review*, 45:167–256, 2003.
- [73] I. Norros and H. Reittu. On a conditionally Poissonian graph process. *Adv. in Appl. Probab.*, 38(1):59–75, 2006.

- [74] Y. Peres and D. Revelle. Scaling limits of the uniform spanning tree and loop-erased random walk on finite graphs. *arXiv:math/0410430*, 2005.
- [75] O. Safsafi. *Arbres couvrants minimums aléatoires inhomogènes, propriétés et limite*. PhD thesis, LPSM, March 2021.
- [76] O. Schramm. Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.*, 118:221–288, 2000.
- [77] D. Sénizergues. Random gluing of metric spaces. *Ann. Probab.*, 47(6):3812–3865, 2019.
- [78] D. Sénizergues. *Structures arborescentes aléatoires : recollements d’espaces métriques et graphes stables*. PhD thesis, Université Sorbonne Paris cité., 2019.
- [79] Michel Talagrand. *The generic chaining*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005. Upper and lower bounds of stochastic processes.
- [80] C. Villani. *Optimal transport, old and new*. Springer, 2006.
- [81] D. B. Wilson. Generating random spanning trees more quickly than the cover time. *In Proceedings of the Twenty-eighth Annual ACM Symposium on the Theory of Computing (Philadelphia, PA, 1996)*, pages 296–303, 1996.