École doctorale **Sciences, Technologies, Santé – Galilée** (ED 146)
Laboratoire **LAGA – Laboratoire Analyse Géométrie Applications** (UMR 7539)

THÈSE DE DOCTORAT

de
L'UNIVERSITÉ SORBONNE PARIS CITÉ
PRÉPARÉE À L'UNIVERSITÉ PARIS 13

Spécialité : **Mathématiques**

Structures arborescentes aléatoires : recollements d'espaces métriques et graphes stables

Présentée par

Delphin SÉNIZERGUES

Soutenue publiquement le 28 juin 2019 devant le jury composé de :

M. Romain Abraham	(Université d'Orléans)	Examinateur
M. Nicolas Broutin	(Sorbonne Université)	Rapporteur
M. Nicolas Curien	(Université Paris-Saclay)	Directeur de thèse
M. Thomas Duquesne	(Sorbonne Université)	Examinateur
Mme Bénédicte Haas	(Université Paris 13)	Directrice de thèse
M. Remco van der Hofstad	(Eindhoven University of Technology)	Rapporteur
M. Yueyun Hu	(Université Paris 13)	Examinateur
M. Philippe Marchal	(Université Paris 13)	Examinateur

Remerciements

Voici venu le moment fatidique de remercier tous les gens qui ont compté pour moi pendant le déroulement de cette thèse. Ces gens étant nombreux et le-dit moment coïncidant curieusement avec la nuit précédant l'impression du manuscrit, je pense qu'il est probable que j'oublie des noms, je vous prie donc par avance de m'excuser.

Mes premiers remerciements vont bien sûr à mes « papa » et « maman » de thèse, Nicolas et Bénédicte, pour avoir encadré mes premiers pas dans le monde de la recherche en mathématiques, pour avoir su superviser mon travail tout en me laissant la liberté d'aller explorer mes propres pistes et avoir inlassablement relu et corrigé les versions successives des différents chapitres qui composent cette thèse. Merci Nicolas pour ton optimisme, ton flot intarissable d'idées et ton énergie sans limite. Merci Bénédicte pour ta patience et ta disponibilité, pour m'avoir rassuré les fois où les maths n'avançaient pas et m'avoir encouragé les fois où mes idées te semblaient prometteuses. Vos styles respectifs se complètent parfaitement et je n'aurais pas pu espérer mieux !

Merci à Christina, avec qui je partage la paternité d'un des chapitres de cette thèse (ainsi qu'avec Bénédicte), sans conteste celui avec le plus bel anglais ! C'était un plaisir de travailler avec toi, autant sous ta supervision pour mon stage à Oxford que pour cette collaboration. Merci aussi à mes autres collaborateurs Siggi et Benedikt, grâce à qui j'ai eu la chance d'aller à Reykjavik et de voir un petit bout d'Islande.

Il est temps maintenant de remercier le jury : je remercie donc chaleureusement Nicolas Broutin et Remco van der Hofstad d'avoir accepté la lourde tâche de relire mon manuscrit, ainsi que les autres membres du jury Romain Abraham, Thomas Duquesne, Yueyun Hu et Philippe Marchal de braver la canicule pour venir m'écouter. (Malheureusement Philippe, je n'ai pas pu réserver de cave à champagne pour la soutenance comme tu me le préconisais !)

Au cours de ces années de thèse, j'ai eu beaucoup de plaisir de croiser régulièrement en conférence au CIRM, à Saint-Flour, aux journées cartes et ailleurs, tout ce petit monde des gens qui étudient arbres, cartes, graphes aléatoires et avec qui j'ai toujours plaisir à discuter, parmi lesquels Mickaël, Nathan, Cyril, Jérémie, Igor, Olivier, Paul, Tom et bien d'autres. Bien sûr, un thésard ne passe pas toute sa vie en conférence et j'ai bien sûr beaucoup apprécié mon temps passé au contact des gens du LAGA. Je tiens à exprimer d'abord ma reconnaissance à l'équipe administrative, Yolande, Isabelle et maintenant Frédéric qui la remplace, pour leur efficacité et leur bonne humeur. Il me faut aussi remercier Clément et Laurent pour lesquels j'ai eu le plaisir de donner des TDs, ainsi que le reste de l'équipe proba-stats, parmi lesquels Bastien ainsi que Jean-Stéphane, mon tuteur. Que seraient ces années de thèse sans la présence des autres doctorants du LAGA avec qui on se retrouve sans faute à 12h au 12, et avec qui on continue à l'heure du café de débattre du dernier épisode de Game of Thrones, du niveau des étudiants ou des méfaits du capitalisme ? Parmi eux, Anna-Laura, Nicolas, Daniel, Marion qui ont déjà quitté le nid, Mattia, Tom, Carlos et Pierre qui sont encore là mais plus pour longtemps. Enfin il y a ceux qui n'ont pas encore fini et qui seront encore là l'année prochaine pour assurer la relève :

Hugo, Marta, Ana, Thuy, Ayoub, Safaa, Mouna, Nelly. Mention spéciale à Jean-Michel avec qui j'ai co-organisé le séminaire des doctorants cette année et qui se donne du mal régulièrement pour m'expliquer des notions de catégories dans le transilien H, et ce malgré mes faibles aptitudes dans ce domaine.

Je suis reconnaissant envers tous les enseignants de mathématiques qui m'ont permis d'en arriver jusque là, en particulier mon professeur de première année de prépa Guillaume Roussel dont les cours limpides m'ont confirmé mon attrait pour cette matière et m'ont donné une base solide pour réussir les concours. Merci aussi à Raphaël Cerf, mon tuteur à l'ENS, grâce à qui je sais faire correctement un exposé au tableau noir. Ces années de scolarité n'ont pas été riches qu'en termes de maths et j'y ai rencontré un tas de gens intéressants qu'il sera difficile de citer de façon exhaustive. Commençons par la prépa, avec Pierre, Jérémy, Édouard, pour tous ces moments à traîner dans une chambre d'internat tout en engloutissant des goûters à n'importe quelle heure de la journée et Lucie, pour tout ce qu'on a vécu ensemble. Ensuite, parmi tous les gens rencontrés à Paris, toute la bande qui se retrouve à Arêches, avec skis, raquettes ou de quoi terraformer Mars : Olivier, Raphaël, Antoine, Claire, Léa, Sophie, Alice, Michel, Hélène, Yaël, Vincent, Aurore, Émilie, Corentin, Julie. Bien sûr, quelques noms manquent à cette liste car j'ai décidé de ne citer chaque personne qu'une fois : je n'oublie pas mes colocs passés et présents, les cinq fantastiques François, Aymeric, Ambre, Budz et Jacko, qui ont subi toutes ces années avec patience mes blagues de matheux et ma tendance générale à augmenter l'entropie. Enfin, il me faut citer ces autres gens rencontrés à l'ENS dans une salle de cours ou un bar en sous-sol, et qui ont contribué à rendre ce lieu formidable à mes yeux, parmi lesquels Anne-Fleur, Tristan, Salomé, Rémi, Élodie, Babeth, Mathilde, Hugo, Jannaud et j'en passe. Merci à vous tous, ces années n'auraient pas été les mêmes sans vous.

Ils m'ont permis de sortir la tête de l'ENS et de continuer ma passion en arrivant à Paris : merci à l'En Avant de Paris, où j'ai plaisir à retrouver à l'entraînement tous les gymnastes du groupe élite. Mention spéciale au coach Jean-Paul, à notre champion Louis et à mes coéquipiers de cette saison Adrien, William, Sven, Mathias et Ludo pour l'ambiance et la réussite de l'équipe. Merci aussi aux gyms du CAB qui m'accueillent de temps en temps pour un entraînement dans leur salle quand je passe par là et que ça me fait plaisir de voir continuer à progresser. Je n'oublie pas bien sûr dans tout ça mes amis de Bègles et de Bordeaux, qui me font me sentir encore chez moi là-bas alors que je vis à Paris depuis bientôt 7 ans : Clément, Simon, Théo, Matthias, Maxime, Gabriel, vous êtes top les gars !

Enfin, merci à ma famille, Grandpa, Grandma, Mané, Jonas, Loup, et particulièrement Papa et Maman, pour m'avoir toujours encouragé dans tout ce que j'entreprends.

Table des matières

Chapitre 1

Introduction

Dans cette introduction, on commence par évoquer le contexte et des résultats classiques dans l'étude des espaces métriques aléatoires, dont beaucoup concernent des modèles d'arbres. On présentera ensuite les différents travaux de cette thèse, dont le thème central est l'étude d'espaces métriques aléatoires qui ne sont pas des arbres mais dont la structure y est fortement apparentée. Dans la Section 1.2, on présente une façon aléatoire de recoller une suite d'espaces métriques itérativement, en attachant à chaque étape de la procédure un nouveau bloc sur la structure construite jusque là. Sous certaines conditions sur les blocs que l'on agglomère, on calcule la dimension de Hausdorff de la structure obtenue et son expression est surprenante !

Ensuite en Section 1.3 on étudie certaines propriétés asymptotiques (degrés, hauteur, profil) de deux modèles d'arbres discrets construits récursivement, les arbres récursifs pondérés et les arbres à attachement préférentiel affine à poids initiaux. Les premiers encodent la structure discrète sous-jacente aux recollements des blocs dans la construction précédente, et les seconds ont un rôle similaire pour des modèles de graphes discrets construits de façon analogue. Cette connexion est exploitée dans les preuves des résultats exposés en Section 1.4, qui concernent des limites d'échelle de certains modèles de graphes discrets vers des espaces métriques continus construits par recollement itératif.

On finira en Section 1.5 par décrire des résultats obtenus avec Christina Goldschmidt et Bénédicte Haas à propos de la composante α-stable à surplus fixé. Cet espace métrique aléatoire apparaît comme la limite d'échelle des grandes composantes connexes de modèles de configuration critique à queue lourde. Cet objet est presque un arbre à l'exception d'un nombre fini de cycles dont nous étudions la structure.

Les contributions originales de cette thèse sont les résultats encadrés.

Sommaire

1.1 Graphes et géométrie aléatoire

1.1.1 Des graphes partout

Des graphes, pour quoi faire ?

De façon classique, un graphe G est une structure définie à partir d'un ensemble de sommets, qui peuvent deux à deux être reliés ou non par une arête. La notion de graphe est extrêmement utile dans des domaines variés car elle peut émerger dès qu'on étudie un ensemble fini d'objets (les sommets) et une relation symétrique entre ces objets (deux sommets sont reliés si les objets sont en relation). Dans les mathématiques de la modélisation, les sommets et arêtes peuvent représenter une réalité concrète : un réseau physique (serveurs reliés par des câbles), un réseau social (individus reliés par des relations d'amitiés), une généalogie (individus reliés par filiation), un modèle génétique (allèles d'un même gène reliés s'ils ne diffèrent que d'une mutation)... Ce type de structures réelles a motivé l'introduction de nombreux modèles de graphes aléatoires, dont on compare les propriétés avec celles des réseaux réels observés.

Ils peuvent aussi servir à discrétiser un problème en servant d'espace ambiant : pour des raisons techniques, on peut avoir envie de travailler sur le graphe $\mathbb{Z}^3$ (où deux sommets sont reliés s'ils sont à distance 1) comme approximation de l'espace $\mathbb{R}^3$. Ils peuvent ainsi servir comme réseau sous-jacent pour étudier des phénomènes physiques : on peut citer la marche aléatoire pour modéliser le mouvement d'une particule, la percolation pour modéliser des trous dans un milieu poreux ou le modèle d'Ising pour l'aimantation d'un matériau ferromagnétique.

Tous les graphes sont géométriques

En fait, tous les graphes (connexes) peuvent être vus comme des espaces géométriques sans avoir besoin d'être plongés dans un espace euclidien : pour tout couple de points x et y dans un graphe G, on peut définir une distance (dite *distance de graphe*) entre ces deux points par la longueur du plus court chemin dans le graphe qui permet de joindre x à y. Ce point de vue fait de l'ensemble des sommets d'un graphe un *espace métrique*. Grâce à cela, il est possible de considérer des processus initialement étudiés dans l'espace euclidien $\mathbb{Z}^d$ comme la marche aléatoire, la percolation ou le modèle d'Ising, sur des graphes avec une géométrie bien différente (par exemple un arbre binaire infini) et de se poser des questions similaires sur, respectivement, l'éloignement du marcheur aléatoire à son point de départ, la taille des composantes connexes ou la distance typique des corrélations.

Il est naturel de généraliser ce point de vue en étudiant ce type de processus dans une *géométrie aléatoire*, c'est-à-dire sur des graphes sous-jacents définis de façon aléatoire. Cela peut provenir de considérations concrètes comme la résilience d'un réseau sous des pannes aléatoires de transmission (percolation par arêtes sur un graphe aléatoire représentant le réseau), ou d'un intérêt plus « mathématique », celui de comparer et de classifier les comportements que peut avoir un même processus défini dans différents cadres.

Cela motive l'étude approfondie des propriétés géométriques de tels graphes aléatoires. La plupart de ces modèles dépendent d'un paramètre n qui mesure la taille du graphe obtenu (n sera typiquement le nombre de sommets) et on s'intéresse à la géométrie des objets obtenus quand n est très grand. Cette étude est en partie comparative : sur les nombreux modèles décrits dans toutes la littérature il arrive souvent que certains modèles aient des propriétés locales différentes mais que leur géométrie à grande échelle soit similaire. Les physiciens parlent souvent d'*universalité* pour décrire ce type de phénomène et une grande partie du travail dans ce domaine est de classifier les modèles en différentes classes d'universalité, dont on étudie et compare les propriétés.

1.1.2 Propriétés à grande échelle et topologie de Gromov–Hausdorff

De même qu'il paraît raisonnable de dire que l'espace $\delta \cdot \mathbb{Z}^d$ approxime bien l'espace euclidien $\mathbb{R}^d$ lorsque δ est petit, on aimerait bien exprimer le fait que nos graphes aléatoires convenablement renormalisés approximent un certain espace limite, lui-même potentiellement aléatoire. Pour cela nous allons introduire la topologie de Gromov–Hausdorff, qui donne une signification rigoureuse à la notion de convergence d'espaces métriques compacts.

Ce point de vue a été introduit en 1975 par David Edwards [57], puis développé par Mikhail Gromov [65] dans les années 1980 pour démontrer un théorème sur les groupes à croissance polynomiale. À partir des années 2000, ce formalisme a été repris par les probabilistes afin d'exprimer des convergences de graphes. Depuis, définir et étudier des variables aléatoires à valeurs « espaces métriques » représente un domaine très actif des probabilités, et c'est dans ce contexte que s'inscrit cette thèse.

Tout graphe fini muni de sa distance de graphe est un espace métrique compact, décrivons donc une façon de comparer de tels objets.

Distance de Hausdorff entre compacts d'un même espace. Lorsque deux compacts sont deux sous-ensembles d'un même espace, on peut les comparer en utilisant la distance suivante, introduite par Félix Haudorff [73] : la *distance de Hausdorff* entre deux parties compactes non-vides A et B d'un espace métrique (E,d) est définie par

$$\mathrm{d}_{\mathrm{H}}^{E}(A,B) = \inf\left\{\epsilon > 0 \,\middle|\, A \subset B^{(\epsilon)} \quad \text{et} \quad B \subset A^{(\epsilon)}\right\},$$

où $A^{(\epsilon)} := \{x \in E \mid d(x,A) < \epsilon\}$ et $B^{(\epsilon)}$ défini de la même manière sont les ϵ-grossissements des ensembles correspondants. Pour cette distance, deux compacts d'un même espace sont proches s'ils sont presque « superposés ». L'idée ensuite pour comparer deux compacts « abstraits » sera de plonger tous les deux isométriquement dans le même espace et de comparer leurs images. De cette façon, deux compacts seront proches s'ils sont presque « superposables ».

Distance de Gromov–Hausdorff entre compacts abstraits. Afin de formaliser cette idée, on utilisera l'espace d'Urysohn $(\mathcal{U},\delta)$ qui l'unique espace métrique complet séparable à vérifier une certaine propriétés d'extension des isométries (voir [78] pour une construction) et qui est une espace assez « gros » pour que tout espace métrique compact soit isométrique à un de ses sous-ensembles. On peut maintenant introduire la distance de Gromov–Hausdorff : si (X,d) et (X',d') sont deux espaces métriques compacts non-vides, leur distance de Gromov–Hausdorff est donnée par

$$\mathrm{d}_{\mathrm{GH}}((X,d),(X',d')) = \inf_{\phi:X\to\mathcal{U},\phi':X'\to\mathcal{U}} \left\{\mathrm{d}_{\mathrm{H}}^{\mathcal{U}}(\phi(X),\phi'(X'))\right\},$$

où la borne inférieure est prise sur l'ensemble des plongements isométriques ϕ et ϕ' de X et X' dans l'espace d'Urysohn $\mathcal{U}$.

On peut définir d'autres distances plus fines à l'aide de la même idée afin de prendre également en compte de la structure supplémentaire et ainsi comparer des compacts munis d'une mesure finie et/ou d'un (ou plusieurs) point distingué. Cela donne lieu à la *distance de Gromov–Hausdorff–Prokhorov* dans le cas mesuré, et on peut rajouter la mention *pointée* (resp. k-pointée) pour signifier qu'on tient compte d'un (resp. k) point distingué.

Espace de Gromov–Hausdorff. La « distance » de Gromov–Hausdorff présentée au-dessus n'en est pas tout à fait une, puisque deux espaces différents peuvent avoir une distance nulle entre eux s'ils sont isométriques. D'autre part, on pourrait aussi s'inquiéter du fait que les espaces métriques compacts ne forment pas un ensemble à proprement parler. Pour contourner ces problèmes, nous ne considérons que l'ensemble des parties compactes (munies de la distance induite) de l'espace d'Urysohn et quotientons cet espace par la relation d'équivalence « être isométrique ». On appelle l'espace obtenu l'*espace de Gromov–Hausdorff* et dans ce manuscrit, on le notera $\mathbb{M}$. Pour cette définition, les éléments de $\mathbb{M}$ sont techniquement des classes d'isométrie de compacts de $\mathcal{U}$, mais grâce aux propriétés de l'espace d'Urysohn, n'importe que espace métrique compact est isométrique à sous-ensemble de $\mathcal{U}$; tout espace métrique compact non-vide est donc bien représenté par un élément de $\mathbb{M}$. On fera souvent l'abus de ne pas faire la différence entre un compact et sa classe d'isométrie. L'espace $(\mathbb{M}, \mathrm{d}_{\mathrm{GH}})$ est lui-même un espace métrique dont on peut montrer qu'il est séparable et complet (voir par exemple [1]). On a donc un cadre agréable pour étudier des variables aléatoires à valeurs dans cet espace.

Limite d'échelle de modèles discrets. De façon générale, on dit ici qu'un modèle discret aléatoire admet une limite d'échelle si ce modèle convenablement renormalisé tend vers un modèle continu, lorsque la taille du modèle discret tend vers l'infini. Dans notre cadre ce sont les distances qu'on renormalise : un modèle discret de graphes vus comme des espaces métriques (G_n, d_n) admet une limite d'échelle quand $n \to \infty$ s'il existe une suite (λ_n) telle que la limite en loi suivante ait lieu dans l'espace $\mathbb{M}$,

$$(G_n, \lambda_n \cdot \mathrm{d}_n) \xrightarrow[n\to\infty]{(\mathrm{d})} (\mathcal{G}, \mathrm{d}), \tag{1.1}$$

pour une variable aléatoire $(\mathcal{G}, \mathrm{d})$ à valeurs dans $\mathbb{M}$.

1.1.3 Que sait-on faire ?

Maintenant que ce cadre est défini, c'est un domaine à part entière que d'étudier des convergences du type (1.1) pour une grande variété de modèles. Pour prouver une telle convergence, il faut en général montrer la tension des lois de $(G_n, \lambda_n \cdot \mathrm{d}_n)$ ainsi qu'identifier la limite $(\mathcal{G}, \mathrm{d})$. Cela suppose en particulier de pouvoir donner une description explicite de la limite, ce qui est parfois complexe. Une grande partie des objets étudiés jusqu'à maintenant sont codés par des processus (par exemple continus à valeurs réelles). En particulier, si

$$(\mathcal{G}, \mathrm{d}) = f(X) \tag{1.2}$$

avec $f : (C([0\,,\infty], \mathbb{R}), \|\cdot\|_\infty) \to \mathbb{M}$ une fonction continue et X un processus réel, une des méthodes pour prouver la limite d'échelle (1.1) est de montrer que

$$(G_n, \lambda_n \cdot \mathrm{d}_n) \overset{(\mathrm{d})}{=} f(X_n)$$

pour une suite de processus X_n qui tend en loi vers X.

On est également intéressé par l'étude des propriétés de l'espace aléatoire obtenu à la limite. Ces propriétés peuvent être obtenues en utilisant l'approximation discrète donnée justement par la convergence (1.1), mais à l'inverse il arrive également que des propriétés puissent être directement étudiées grâce à la description qu'on a de l'objet dans le continu et que cela puisse renseigner sur le comportement asymptotique du modèle discret.

Les espaces que l'on comprend sont presque tous des arbres

Les graphes aléatoires les mieux compris sont les arbres, dont on pourrait dire que leur structure est essentiellement unidimensionnelle. Citons des exemples liés aux travaux de cette thèse.

Arbres de Bienaymé-Galton-Watson. Les arbres de Bienaymé-Galton-Watson sont un des modèles les plus étudiés d'arbres aléatoires discrets. Ils correspondent à la généalogie d'un processus de branchement très simple qui a été introduit par Irénée-Jules Bienaymé en 1845, puis étudié à nouveau en 1875 par Francis Galton et Henri William Watson, dans le but de modéliser l'extinction de noms de familles nobles. Si on se donne μ une mesure de probabilité sur $\mathbb{N}$, un arbre de Bienaymé-Galton-Watson τ de loi de reproduction μ code la généalogie du processus suivant : à l'instant 0, un individu est présent et à chaque étape $n \geq 1$, tous les individus présents au temps précédent meurent et chacun laisse place à un nombre de descendants tiré selon la loi μ, indépendamment les uns des autres.

Convergence. Lorsque la loi de reproduction μ est de moyenne 1 et de variance finie $0 < \sigma^2 < \infty$, les arbres de Bienaymé-Galton-Watson de loi de reproduction μ conditionnés à avoir un nombre n de sommets, renormalisés par $n^{-1/2}$, convergent en loi lorsque $n \to \infty$ vers un multiple du même objet universel limite : l'arbre brownien d'Aldous.

Théorème 1 (Aldous, [10])**.** *Soit μ une loi sur $\mathbb{N}$ d'espérance 1 et de variance $0 < \sigma^2 < \infty$. On note T un arbre de Bienaymé-Galton-Watson de loi de reproduction μ. Pour tout $n \geq 1$ pour lequel $\mathbb{P}(|T| = n) > 0$, on définit T_n comme une version de T conditionnée à avoir exactement n sommets, que l'on munit de sa distance de graphe d_{gr}. Alors*

$$\left(T_n, n^{-1/2} \cdot \mathrm{d}_{gr}\right) \underset{n\to\infty}{\longrightarrow} (\mathcal{T}, \sigma^{-1} \cdot d),$$

où la convergence a lieu pour la topologie de Gromov–Hausdorff, le long des n pour lesquels T_n est bien défini.

Arbre brownien d'Aldous. L'arbre brownien d'Aldous $(\mathcal{T}, d)$ qui apparaît à la limite a une description de la forme (1.2), c'est-à-dire qu'il est codé par un processus aléatoire réel. En effet, on peut le construire à partir de $\mathbf{e} = (\mathbf{e}(t))_{t\in[0,1]}$ l'excursion brownienne normalisée de la façon suivante : on définit d'abord une pseudo-distance d sur $[0\,,1]$ par

$$d(s,t) := 2 \cdot \left(\mathbf{e}(s) + \mathbf{e}(t) - 2 \cdot \min_{u\in[s\wedge t, s\vee t]} \mathbf{e}(u)\right).$$

On définit ensuite la relation d'équivalence $\sim$ sur $[0\,,1]$ telle que $s \sim t$ si et seulement si $d(s,t) = 0$. En identifiant ensemble les points qui sont en relation, c'est-à-dire en considérant l'ensemble $\mathcal{T} = [0\,,1]/\sim$, muni de la distance induite sur le quotient que l'on note encore d, on obtient l'arbre brownien d'Aldous $(\mathcal{T}, d)$.

Construction line-breaking. En fait, cette construction n'est pas la première donnée par David Aldous, qui a au départ [9] défini cet objet comme un sous-ensemble de $\ell^1 := \left\{(x_n)_{n\geq 1} \in \mathbb{R}^{\mathbb{N}^*} \;\middle|\; \sum_{n=1}^{\infty} |x_n| < \infty\right\}$, à partir d'une construction qu'il nomme « line-breaking » et qu'il exprime comme suit : soit $C_1, C_2, \dots$ les points d'un processus de Poisson sur $\mathbb{R}^+$ d'intensité $t \mapsto t dt$, rangés dans l'ordre croissant. Ces points « découpent » la demi-droite $\mathbb{R}_+$ en intervalles de la forme $[C_i\,, C_{i+1})$. En commençant avec le segment $[0\,, C_1)$, on construit un arbre récursivement en attachant à la i-ième étape le segment $[C_i\,,C_{i+1})$ à un point B_i choisi uniformément sur l'arbre construit jusque là. À chaque fois qu'une branche est ainsi accrochée, on la construit le long d'une coordonnée de ℓ^1 non encore utilisée, de telle façon à ce que la distance ℓ^1 entre deux points de l'arbre soit toujours égale à la longueur qu'il faut parcourir dans l'arbre pour aller de l'un à l'autre. L'arbre brownien est obtenu comme l'adhérence de la partie ainsi construite après une infinité d'étapes. Le fait que ces deux constructions coïncident est aussi un résultat d'Aldous, voir [10]. Plus tard dans la Section 1.2, nous allons nous intéresser plus particulièrement à cette construction et ses généralisations.

Arbres stables. Lorsque la loi de reproduction μ n'a admet pas de variance mais admet une queue de distribution de la forme $\mu([k\,,\infty)) \sim k^{-\alpha}$ pour $\alpha \in (1\,,2)$, le Théorème 1 n'est plus valide mais un résultat similaire [52] de Thomas Duquesne assure que les arbres T_n renormalisés cette fois par $n^{-\frac{\alpha-1}{\alpha}}$ convergent vers un multiple d'un autre arbre continu : l'arbre α-stable. Celui-ci a été introduit par Thomas Duquesne et Jean-François Le Gall [53] en s'appuyant sur des résultats antérieurs de Jean-François Le Gall et Yves Le Jan [89]. Sa définition est similaire à celle de l'arbre brownien, sauf que l'excursion brownienne normalisée est remplacée par un certain processus de hauteur associé à une excursion d'un processus de Lévy α-stable spectralement positif par une construction que nous ne détaillerons pas ici.

Arbres de fragmentation. D'autres arbres définis dans le continus ne le sont pas forcément à partir d'une excursion d'un processus naturel : c'est le cas des *arbres de fragmentations*, introduits par Bénédicte Haas et Grégory Miermont [68], qui codent la généalogie de fragments qui se disloquent aléatoirement selon un *processus de fragmentation auto-similaire*, à valeurs dans $\mathcal{S}^{\downarrow} = \{(s_1, s_2, \dots) \mid s_1 \geq s_2 \geq \dots \geq 0, \sum_{i=1}^{\infty} s_i \leq 1\}$. Les deux modèles décrits ci-dessus peuvent également être décrits comme des arbres de fragmentation.

Arbre couvrant minimal. Citons également l'arbre $\mathscr{M}$ obtenu par Louigi Addario-Berry, Nicolas Broutin, Christina Goldschmidt et Grégory Miermont [6] comme limite d'échelle de l'arbre couvrant minimal du graphe complet à n sommets. Celui-ci n'a pas de codage par une fonction explicite mais peut être compris à travers un continuum d'approximations par des arbres continus dont la loi est relativement explicite. Cette absence de représentation par un processus en fait tout de même un objet assez mal compris.

Autres exemples bien compris qui ne sont pas des arbres

Citons quelques exemples d'autres graphes et espaces métriques aléatoires qui sont aussi bien compris et qui ne sont pas pour autant des arbres.

Graphe d'Erdős-Rényi. Le graphe d'Erdős-Rényi $G(n,p)$ est un des modèles les plus simples de graphes aléatoire. Il s'agit d'un graphe à n sommets entre lesquels chacune des $\binom{n}{2}$ arêtes possibles est présente avec une probabilité p, indépendamment des autres. Un des premiers résultats sur la géométrie de ce modèle fait intervenir la taille de ses composantes connexes et est dû à Paul Erdős et Alfréd Rényi [58] : si on choisit $p = \frac{c}{n}$ pour une certaine constante $c > 0$, alors on observe un transition de phase (un changement qualitatif de comportement) pour la valeur $c = 1$. Si $c < 1$ alors les plus grandes composantes sont de tailles logarithmiques en n, alors que pour $c > 1$, la plus grande composante est de taille proportionnelle à n.

Limite d'échelle d'un graphe d'Erdős-Rényi critique. Lorsque le modèle est étudié au point critique $c = 1$, on peut observer des comportements intéressants. David Aldous [8] a montré que si p est dans « la fenêtre critique » i.e. $p = \frac{1}{n} + \frac{\lambda}{n^{4/3}}$ pour une constante $\lambda \in \mathbb{R}$, alors les tailles des composantes prises dans l'ordre décroissant convergent en loi après normalisation par $n^{-2/3}$ vers une suite aléatoire de réels positifs qui correspondent à des longueurs d'excursions d'un certain processus brownien. Louigi Addario-Berry, Nicolas Broutin et Christina Goldschmidt ont montré [5] qu'en considérant les composantes connexes comme des espaces métriques, on obtient une convergence dans la limite d'échelle en normalisant les distances par $n^{-1/3}$ vers une suite d'espaces métriques aléatoires. Ceux-ci sont presque des arbres, à un nombre fini de cycles près, et peuvent être décrits comme des arbres browniens biaisés munis d'un nombre fini d'identification de points, mais aussi à partir d'une construction du type « line-breaking » en partant d'une structure cyclique [4].

Un modèle bidimensionnel : les cartes planaires. Les cartes planaires sont des graphes planaires munis d'un plongement sur la sphère, à déformation près. Leurs propriétés ont notamment été étudiées par Tutte dans les années 1960 pour prouver le fameux théorème des quatre couleurs. Depuis les années 1980, ces objets ont éveillé un intérêt chez les physiciens qui, dans le cadre d'une théorie de la gravitation en deux dimensions, s'intéressent à des résultats asymptotiques sur de grandes cartes aléatoires.

Les modèles les plus simples de cartes aléatoires consistent à prendre une carte uniformément parmi un ensemble de cartes dont le nombre d'arêtes est fixé (cartes générales, triangulations, quadrangulations etc.). Pour ces modèles bidimensionnels, des résultats de limites d'échelle sont aussi prouvés et un objet universel apparaît à la limite : *la carte brownienne*, qui est homéomorphe presque sûrement à une sphère. Le domaine de l'étude des cartes aléatoires est actuellement très actif et c'est à peu près le seul exemple d'espace métrique aléatoire qui ne soit pas essentiellement un arbre et pour lequel on a accès à des propriétés fines de la limite : volume et « périmètre » des boules, structure des géodésiques etc...

1.1.4 Des arbres augmentés

En dehors des cartes et des arbres, peu d'autres modèles continus sont compris et un des objectifs des travaux de cette thèse est d'étudier des constructions qui sortent légèrement de ce cadre, tout en étant reliées à une structure d'arbre.

Une première façon de réaliser un tel modèle est d'imiter la construction line-breaking d'Aldous en collant des espaces métriques plus généraux que de simples segments ; c'est sur les propriétés de cette construction que portent les résultats énoncés en Section 1.2. Plus précisément, si (X, d) et (X', d') sont deux espaces métriques compacts munis chacun d'un point distingué x (resp. x'), leur recollement *ponctuel* est obtenu en identifiant leurs points distingués respectifs : on considère l'ensemble

$$X \sqcup X'/(x \sim x'), \tag{1.3}$$

qu'on munit d'une distance δ qui respecte les distances dans X et X', i.e. $\delta_{|X\times X} = d$ et $\delta_{|X'\times X'} = d'$ et telle que pour deux points $a \in X$ et $b \in X'$ on ait $\delta(a, b) = d(a, x) + d'(x', b)$. On effectue ce type de recollement de façon itérative en collant ponctuellement à chaque étape un nouveau « bloc » sur un point choisi aléatoirement sur la structure déjà construite. Le théorème principal concerne le calcul de la *dimension de Hausdorff* des structures ainsi obtenues après une infinité d'étapes.

Dans la Section 1.3, on énonce des résultats sur la structure d'arbre discret qui apparaît dans le processus précédent comme le plan selon lequel nos blocs ont été recollés ensemble et qu'on appelle un *arbre récursif pondéré*. On s'intéresse en particulier à la hauteur, au profil et aux degrés des sommets dans cet arbre. On verra qu'un autre modèle, celui d'*arbre à attachement préférentiel* peut être identifié comme un exemple d'arbre récursif pondéré et cela est crucial pour les travaux présentés dans le Section 1.4.

En effet, en Section 1.4, on s'intéresse à des processus de croissance de graphes discrets qu'on peut interpréter comme des graphes (qui évoluent au cours du temps) collés le long d'un arbre à attachement préférentiel. Grâce aux résultats de la Section 1.3, on prouve qu'ils admettent une limite d'échelle qui se construit par des recollements itératifs comme étudiés dans la Section 1.2.

Enfin en Section 1.5, on énonce des résultats sur la structure de la *composante α-stable*. C'est un espace métrique aléatoire qui peut être défini à partir d'un arbre α-stable biaisé dans lequel on identifie aléatoirement un nombre fini de paires de points, créant ainsi des cycles. C'est la contrepartie α-stable des limites d'échelle browniennes des composantes d'un graphe d'Erdős–Rényi critique ; elle apparaît comme la limite d'échelle des composantes connexes d'un modèle de graphe discret à degrés i.i.d. dont la loi est critique et à queue lourde. Un des résultats de la Section 1.5 donne une construction de cet objet à partir, encore une fois, d'une construction par recollements itératifs comme décrite en Section 1.2.

1.1.5 Recoller des espaces le long de l'arbre d'Ulam

En fait, tous nos recollements d'espaces métriques le long d'un arbre, que ce soit un arbre récursif pondéré ou un arbre à attachement préférentiel, peuvent être interprétés dans un formalisme commun, celui de décoration aléatoire sur l'arbre d'Ulam, développé dans le Chapitre 4. Présentons-en ici un bref aperçu.

Décorations sur l'arbre d'Ulam. On considère l'arbre d'Ulam, défini de façon usuelle comme l'ensemble des mots finis sur l'alphabet $\mathbb{N}^*$,

$$\mathbb{U} = \bigcup_{n\geq 0} (\mathbb{N}^*)^n. \tag{1.4}$$

On dit que $\mathcal{D} = (\mathcal{D}(u))_{u\in\mathbb{U}}$ est une *décoration sur l'arbre d'Ulam* si pour tout $u \in \mathbb{U}$,

$$\mathcal{D}(u) = (D_u, \mathrm{d}_u, \rho_u, (x_{ui})_{i\geq 1}),$$

est un espace métrique compact enraciné, d'ensemble sous-jacent D_u, distance d_u, racine ρ_u et muni d'une suite de points distingués $(x_{ui})_{i\geq 1} \in D_u$. On utilise ici la convention que pour tout $i \in \mathbb{N}^*$ l'écriture ui désigne l'élément de $\mathbb{U}$ construit en ajoutant la lettre i à la fin du mot u.

Recollement des décorations. Dans ce cadre, pour toute décoration $\mathcal{D}$, on définit l'espace métrique $\mathscr{G}(\mathcal{D})$, qu'on obtient informellement en considérant

$$\left(\bigsqcup_{u\in\mathbb{U}} D_u\right) / \sim$$

où la relation $\sim$ est définie par $(\rho_{ui} \sim x_{ui})$ pour tout $u \in \mathbb{U}$ et tout $i \in \mathbb{N}$, puis en prenant la complétion de l'espace obtenu. Les distances se calculent de manière similaire au cas du recollement ponctuel de deux espaces, décrit en (1.3), et correspondent à celles du « quotient métrique » par la relation $\sim$ au sens de [32].

Cette construction permet donc de donner un sens clair à ce qu'on appelle un recollement d'espaces métriques le long d'un arbre discret. Tous nos modèles aléatoires de recollements d'espaces métriques, itératifs ou non, pourront en fait s'obtenir comme le recollement $\mathscr{G}(\mathcal{D})$ d'une décoration *aléatoire* $\mathcal{D}$. Le diagramme suivant résume les liens entre les objets étudiés dans les différents chapitre de la thèse.

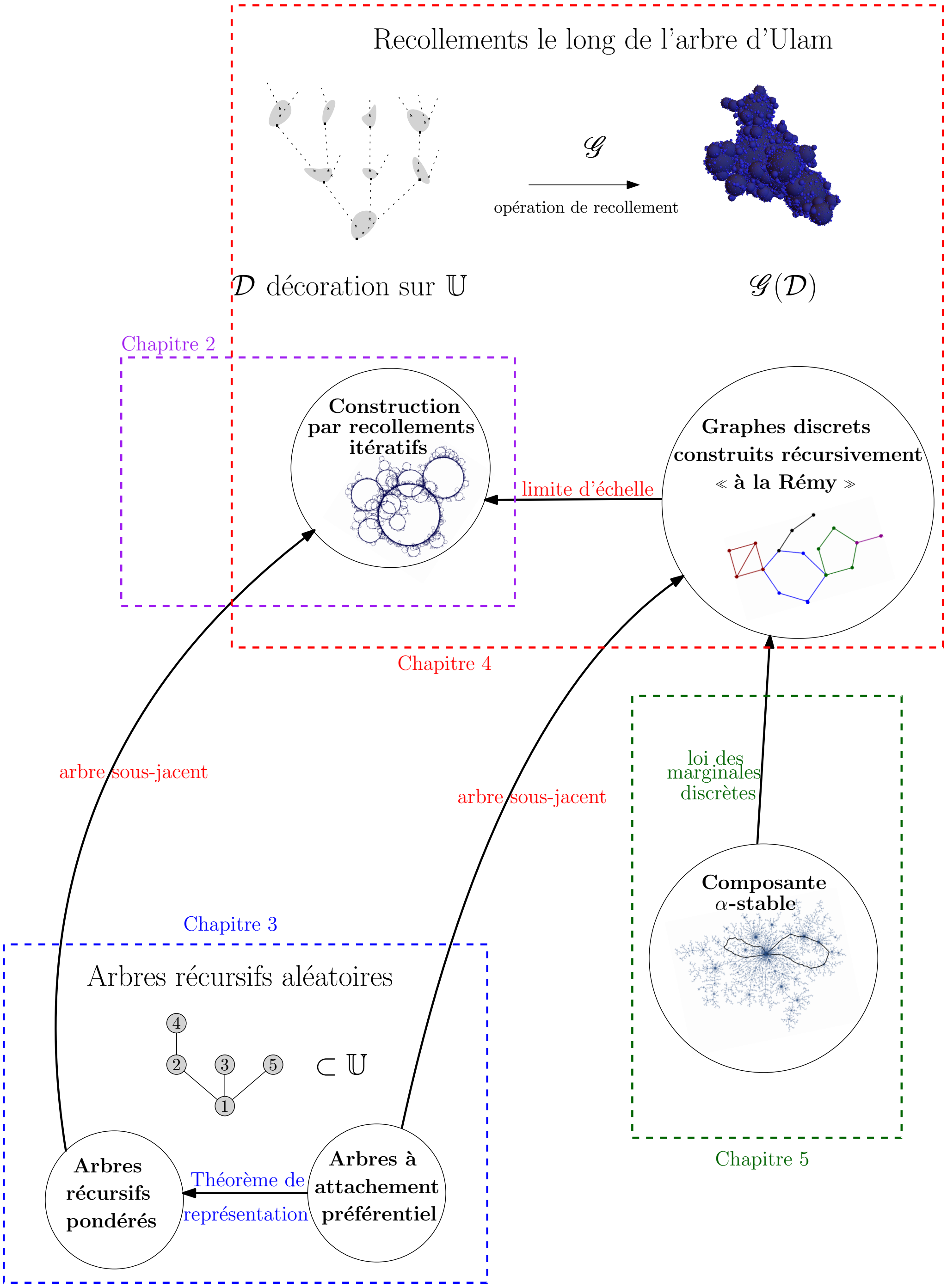


FIGURE 1.1 – Diagramme récapitulatif des objets étudiés dans la thèse

1.2 Recollements itératifs d'espaces métriques

Cette section expose les résultats prouvés dans le Chapitre 2, dans lequel on étudie une construction d'espaces métriques par recollements successifs de blocs.

1.2.1 Une construction par recollement successifs

Donnons nous une suite $((\mathbf{b}_n, \mathbf{d}_n, \boldsymbol{\rho}_n, \boldsymbol{\nu}_n))_{n\geq 1}$ d'espaces métriques pointés et mesurés, que nous appellerons les *blocs* de notre construction. Ici pour tout $n \geq 1$, $\mathbf{b}_n$ est un ensemble non-vide, que l'on munit de la distance $\mathbf{d}_n$ pour laquelle il est compact, d'un point distingué $\boldsymbol{\rho}_n$ que l'on appelle sa *racine*, et de $\boldsymbol{\nu}_n$ une mesure Borélienne finie (non-nulle si $n = 1$).

Décrivons une façon de recoller aléatoirement ces espaces les uns sur les autres de manière itérative. On construit ainsi une suite $(\mathcal{T}_n)_{n\geq 1}$ d'espaces métriques compacts qui représentent l'objet obtenu à chaque étape du processus. On commence avec $\mathcal{T}_1 := \mathbf{b}_1$. Ensuite à chaque étape $n \geq 1$, on obtient $\mathcal{T}_{n+1}$ en collant ponctuellement la racine du bloc $\mathbf{b}_{n+1}$ sur un point X_n de $\mathcal{T}_n$ choisi aléatoirement. Ce point est choisi sous (une version normalisée de) la mesure finie $\mu_n = \boldsymbol{\nu}_1 + \cdots + \boldsymbol{\nu}_n$ obtenue comme la somme des mesures portées par les blocs $\mathbf{b}_1, \dots, \mathbf{b}_n$ vus comme des sous-parties $\mathcal{T}_n$. Comme la suite $(\mathcal{T}_n)_{n\geq 1}$ ainsi définie est croissante pour l'inclusion, on peut introduire

$$\mathcal{T}^* := \bigcup_{n\geq 1} \mathcal{T}_n,$$

l'union croissante de ces objets, qui est d'une certaine façon l'objet obtenu une fois qu'on a fini de recoller tous les blocs ensemble. Afin d'étudier un espace dont les propriétés ne sont pas trop pathologiques, on considère en fait la complétion $\mathcal{T} := \overline{\mathcal{T}^*}$. Dans le cas où $\mathcal{T}$ est compact, cela est en fait assez naturel, vu notre cadre de travail, puisque dans ce cas $\mathcal{T}$ est simplement la limite de la suite $(\mathcal{T}_n)_{n\geq 1}$ au sens de Gromov–Hausdorff.

Comportement de $\mathcal{T}$: le cas des segments

Tout l'enjeu du travail concernant cette construction est de décrire les propriétés géométriques de $\mathcal{T}$ en fonction des propriétés des blocs. On s'intéressera principalement à la compacité et à la dimension de Hausdorff. En fait, cette construction a déjà été étudiée par Nicolas Curien et Bénédicte Haas [41] dans le cas d'un recollement de segments, créant ainsi un arbre. Dans notre formalisme, leur construction correspond à prendre des blocs $((\mathbf{b}_n, \mathbf{d}_n, \boldsymbol{\rho}_n, \boldsymbol{\nu}_n))_{n\geq 1}$ où pour chaque $n \geq 1$, le n-ième bloc est un segment d'une certaine longueur a_n muni de sa distance euclidienne, enraciné en une de ses extrémités et munis de la mesure de Lebesgue. Dans leur cas, ils ont supposé que la suite $(a_n)_{n\geq 1}$ des longueurs des différents segments avait un comportement en puissance décroissante $a_n \approx n^{-\alpha}$, pour un réel $\alpha > 0$, et sous ces hypothèses ils ont prouvé que l'arbre $\mathcal{T}$ obtenu est presque sûrement compact et on calculé sa dimension de Hausdorff. Lorsque $\alpha \leq 1$ la dimension de $\mathcal{T}$ est presque sûrement α^{-1} alors que pour $\alpha > 1$ cette dimension est toujours 1. L'ensemble intéressant à étudier en terme de dimension est en fait l'ensemble $\mathcal{L} = (\mathcal{T} \backslash \mathcal{T}^*)$, puisque la dimension de $\mathcal{T}^*$ est toujours celle d'une union dénombrable de segments, c'est-à-dire 1. Pour cet ensemble-là, ils prouvent qu'on a bien $\dim_{\mathrm{H}}(\mathcal{L}) := \frac{1}{\alpha}$ presque sûrement pour tout $\alpha > 0$.

Hypothèses sur les blocs que l'on considère

On veut donc adapter nos hypothèses pour qu'elles généralisent celles utilisées dans le cas des segments. Afin de rester dans un cadre similaire, on peut supposer que tous les blocs sont de forme semblable, à un facteur de dilatation près, de façon à ce que la taille des blocs décroisse encore en une puissance décroissante de leur indice. On pourrait par exemple considérer des sphères d'une certaine dimension d et imposer que la suite de leurs rayons se comporte en $n^{-\alpha}$ pour un $\alpha > 0$. Dans ce cas précis, il serait naturel de considérer que la mesure totale portée par chacune des sphères soit encore la mesure de Lebesgue et décroisse donc aussi comme une puissance $n^{-\alpha d}$ de l'indice n.

Afin de se placer dans un contexte plus général, on va « découpler » la dilatation des distances et de la mesure, de telle façon à ce que la masse totale du n-ième bloc soit d'un ordre $n^{-\beta}$ pour un paramètre β choisi de façon indépendante du paramètre α. S'il est indispensable d'utiliser des tailles de blocs qui tendent vers 0 si on veut espérer une limite compacte, rien n'oblige à ce qu'il en soit de même pour leur masse et on autorise donc le paramètre β à être un réel quelconque (et donc possiblement négatif). On va aussi autoriser les blocs à avoir des formes aléatoires, en supposant tout de même que ces formes sont les mêmes « en loi ».

Plus précisément, nos blocs seront réalisés à partir de copies indépendantes et identiquement distribuées d'un espace métrique compact pointé muni d'une mesure *de probabilité* $(\mathsf{B}, \mathsf{D}, \rho, \nu)$, que l'on appellera le *bloc sous-jacent*. On considère une suite $((\mathsf{B}_n, \mathsf{D}_n, \rho_n, \nu_n))_{n\geq 1}$ de variables aléatoires i.i.d. avec la distribution de $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ et on considère les blocs

$$\forall n \geq 1, \quad (\mathbf{b}_n, \mathbf{d}_n, \boldsymbol{\rho}_n, \boldsymbol{\nu}_n) := (\mathsf{B}_n, \lambda_n \cdot \mathsf{D}_n, \rho_n, w_n \cdot \nu_n), \tag{1.5}$$

pour une suite $(\lambda_n)_{n\geq 1}$ de *facteurs de dilatation* et une suite $(w_n)_{n\geq 1}$ de *poids.*

On fait l'hypothèse que les suites $(\lambda_n)_{n\geq 1}$ et $(w_n)_{n\geq 1}$ sont choisies de telle façon à ce que pour des réels $\alpha > 0$ et $\beta \in \mathbb{R}$, on ait

$$\lambda_n \approx n^{-\alpha} \quad \text{and} \quad w_n \approx n^{-\beta} \qquad \text{quand } n \to \infty,$$

dans un sens faible. Essentiellement, on demande d'avoir des majorations du type $\lambda_n \leq n^{-\alpha+o(1)}$ et $w_n \leq n^{-\beta+o(1)}$ pour n grand et veut quand même s'assurer que la proportion asymptotique des n pour lesquels on a simultanément $\lambda_n \geq n^{-\alpha-\epsilon}$ et $w_n \geq n^{-\beta-\epsilon}$ soit strictement positive, pour tout $\epsilon > 0$ fixé. Pour des raisons techniques, le cas $\beta < 1$, le cas $\beta > 1$ et le cas $\beta = 1$ sont étudiés sous des hypothèses différentes, que nous ne détaillons pas ici.

On supposera que le bloc sous-jacent $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ avec lequel on travaille se comporte de façon d-dimensionnelle (au sens de la dimension fractale), pour un certain $d \in [0, \infty)$. Cela est formalisé par l'Hypothèse (H_d), pour l'énoncé de laquelle nous renvoyons au chapitre correspondant. Cette hypothèse a été choisie de telle façon à être satisfaite pour beaucoup d'objets aléatoires connus. Par exemple l'arbre brownien d'Aldous satisfait cette hypothèse pour $d = 2$ et la carte brownienne pour $d = 4$.

Comme dans le cas des segments, l'ensemble intéressant à étudier en terme de dimension est $\mathcal{L} := (\mathcal{T} \setminus \mathcal{T}^*)$ qu'on appelle l'ensemble des *feuilles* de la structure. Notre théorème principal donne la valeur presque sûre de la dimension de $\mathcal{L}$ en fonction des paramètres α, β et d.

Théorème 1. *Sous les hypothèses précédentes, la structure $\mathcal{T}$ qui résulte de la construction décrite est p.s. compacte et la dimension de Hausdorff de $\mathcal{L}$ est donnée p.s. par*

$$\begin{aligned}\dim_{\mathrm{H}}(\mathcal{L}) &= \frac{2\beta - 1 - 2\sqrt{(\beta-1)(\beta-\alpha d)}}{\alpha}, \quad \text{si } \beta > 1 \text{ et } \alpha < \frac{1}{d}, \\ &= \frac{1}{\alpha} \quad \text{sinon.}\end{aligned}$$

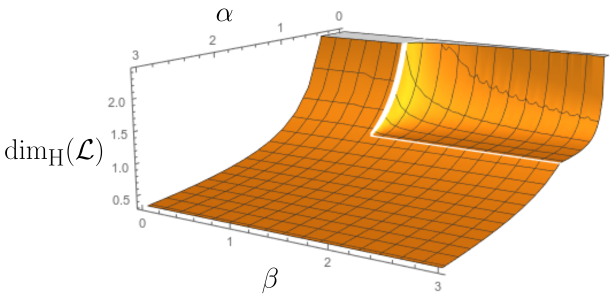


FIGURE 1.2 – Le graphe représente la dimension de Hausdorff de $\mathcal{L}$ donnée par le Théorème 1 en fonction de α et β, pour $d = 1$.

Avant de passer aux idées des preuves, faisons quelques remarques sur le théorème et sur l'étonnante formule donnée pour la dimension de Hausdorff. D'abord, pour d fixé, l'expression donnée pour $\beta > 1$ et $\alpha < \frac{1}{d}$ peut être réécrite en

$$d + \frac{(\sqrt{\beta - \alpha d} - \sqrt{\beta - 1})^2}{\alpha}.$$

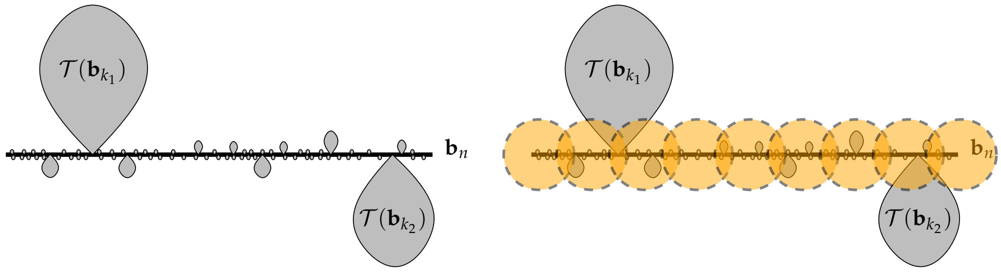
Sous cette forme, on voit immédiatement que cette expression est toujours plus grande que d et un calcul rapide montre aussi qu'elle est aussi plus petite $\frac{1}{\alpha}$. Elle est également décroissante en α et en β sur le domaine considéré et tend vers $\frac{1}{\alpha}$ quand $\beta \to 1$ pour $\alpha > 0$ fixé. Finalement la fonction $(\alpha, \beta) \mapsto \dim_H(\mathcal{L})$ est continue sur tout le domaine $\mathbb{R}_+^* \times \mathbb{R}$. Elle est représentée en Figure 1.2.

Ensuite, on peut remarquer le changement qualitatif de comportement pour autour de la valeur $\beta = 1$: pour $\beta > 1$ la dimension de $\mathcal{L}$ dépend de la géométrie du bloc sous-jacent à travers la valeur de d, sa dimension, alors qu'elle n'en dépend pas du tout quand $\beta < 1$. Une des raisons de cette transition est la sommabilité (ou non) de la suite des poids (w_n). Lorsque $\beta < 1$ celle-ci définit une série divergente et il s'en suit que chaque bloc $\mathbf{b}_n$ n'attire qu'un nombre asymptotiquement négligeable d'autres blocs et sa géométrie n'a donc pas d'effet sur celle de $\mathcal{L}$. Au contraire lorsque $\beta > 1$ et que cette suite est donc sommable, une proportion asymptotiquement positive des blocs s'attache directement sur n'importe quel bloc $\mathbf{b}_n$ donné. Dans ce cas la forme de $\mathcal{L}$ « épouse » beaucoup plus celle des blocs et la géométrie de ceux-ci a donc une réelle influence sur celle de $\mathcal{L}$.

1.2.2 Idées de la preuve

La preuve de ce théorème se décline en plusieurs parties. Tout d'abord, comme il est standard pour le calcul de dimensions de Hausdorff, on prouve séparément des majorants et des minorants pour la dimension, en exhibant des recouvrement explicites (pour la majoration) ou bien en construisant une mesure de probabilité à laquelle on peut appliquer le principe de distribution de masse, pour la minoration. Ensuite, la plupart des preuves que nous proposons ne sont pas valides pour tous les paramètres α et β à la fois, ce qui multiplie les techniques utilisées. Décrivons rapidement les points-clés de chacune des preuves.

Majoration en $\frac{1}{\alpha}$ valide pour tous paramètres. Tout d'abord, un premier argument permet de montrer que sous les hypothèses du théorème (et même sous des hypothèses bien plus



(a) La sous-structure $\mathcal{T}(\mathbf{b}_n)$

(b) Un recouvrement de $\mathbf{b}_n$ avec des petites boules

FIGURE 1.3 – Explication de la procédure de raffinement d'un recouvrement, les blocs sont représentés ici par des segments

générales), la dimension de Hausdorff de $\mathcal{L}$ est presque sûrement plus petite que $\frac{1}{\alpha}$, quels que soient les paramètres α et β. Cela est prouvé en fournissant des recouvrements explicites de $\mathcal{L}$.

Pour tout $n \geq 1$, on note $\mathcal{T}(\mathbf{b}_n)$ la sous-structure issue du bloc $\mathbf{b}_n$ dans $\mathcal{T}$, c'est-à-dire l'adhérence de l'ensemble des blocs qui ont été greffés au-dessus de $\mathbf{b}_n$. Ces sous-ensembles vont nous permettre de fournir des recouvrements de $\mathcal{L}$ puisqu'il est assez immédiat de remarquer qu'ils recouvrent $\mathcal{L}$. La majoration de la dimension provient de l'asymptotique suivante $\mathrm{diam}(\mathcal{T}(\mathbf{b}_n)) \leq n^{-\alpha+o(1)}$ qui assure que sous nos conditions la taille de la sous-structure issue d'un bloc est encore d'une taille comparable à celle du bloc, c'est-à-dire qu'elle décroit en $n^{-\alpha}$.

Majoration plus subtile quand $\beta > 1$ et $\alpha > 1/d$. Lorsque $\beta > 1$ la majoration précédente n'est pas optimale, alors qu'elle l'est lorsque $\beta < 1$. Une des raisons pour ce changement qualitatif de comportement en $\beta = 1$ provient de la remarque suivante : lorsque $\beta < 1$ on a $w_n \approx n^{-\beta}$ et donc le poids cumulé $W_n := \sum_{i=1}^n w_i$ est tel que $W_n \approx n^{1-\beta}$, donc le poids relatif du bloc $\mathbf{b}_n$ à l'instant où il vient d'être greffé est de l'ordre de $\frac{w_n}{W_n} \approx n^{-1}$. Cela indique que l'indice du premier bloc collé sur $\mathbf{b}_n$ sera de l'ordre de n et donc le diamètre de celui-ci sera du même ordre de grandeur $n^{-\alpha}$ que celui de $\mathbf{b}_n$. Au contraire, dans le cas où $\beta > 1$, la suite des poids est sommable, et donc ce poids relatif est de l'ordre de $n^{-\beta}$. Le premier bloc à s'attacher à $\mathbf{b}_n$ a donc un indice de l'ordre de n^β et donc un diamètre de l'ordre de $n^{-\alpha\beta}$ ce qui est négligeable devant celui de $\mathbf{b}_n$. Cela laisse donc penser qu'il pourrait être plus efficace de recouvrir $\mathcal{T}(\mathbf{b}_n)$ avec beaucoup de sous-ensembles d'un ordre de grandeur plus petit, plutôt que de prendre $\mathcal{T}(\mathbf{b}_n)$ tout entier.

On applique cette stratégie pour obtenir des recouvrements de $\mathcal{T}(\mathbf{b}_n)$ de plus en plus fins. Supposons que l'on ait à disposition un recouvrement de chacun des $\mathcal{T}(\mathbf{b}_n)$ pour $n \geq 1$. Alors on peut en définir un nouveau de la façon suivante : pour chaque $n \geq 1$, on recouvre d'abord le bloc $\mathbf{b}_n$ avec des boules d'un rayon très petit par rapport à la taille de $\mathbf{b}_n$, puis on utilise le recouvrement précédent pour recouvrir les parties (en nombre fini) de $\mathcal{T}(\mathbf{b}_n)$ qui ne sont pas contenues dans ces boules. Cela conduit à un nouveau recouvrement de $\mathcal{T}(\mathbf{b}_n)$. La majoration finale est obtenue en itérant cette construction à l'infini à partir du recouvrement naïf décrit au paragraphe précédant. La difficulté consiste à choisir à chaque itération le bon ordre de grandeur pour le rayon des boules, de façon à obtenir la meilleure borne possible sur la dimension de Hausdorff.

Minorations. Afin de produire de minorations sur la dimension d'un espace, une des méthodes est d'exhiber une mesure de probabilité ν sur cet espace de telle façon à ce que pour ν-presque tout x la masse de la boule centrée en x satisfasse $\nu(\mathrm{B}(x,r)) \leq r^{s+o(1)}$. Dans ce cas, s est un minorant pour la dimension de Hausdorff de l'espace, on appelle cela le *principe de distribution de masse*. On utilise des mesures différentes selon la valeur du paramètre β.

Minoration quand $\beta < 1$. Dans le cas où $\beta < 1$, une mesure arrive de façon naturelle. On note $\bar{\mu}_n$ la mesure portée par les n premiers blocs, renormalisée de façon à en faire une mesure de probabilité. On peut montrer sous les hypothèses du théorème que cette mesure de probabilité converge étroitement vers une mesure $\bar{\mu}$ et que cette mesure est portée par $\mathcal{L}$ lorsque $\beta < 1$. Un argument de couplage permet de construire en même temps que la structure $\mathcal{T}$ un point Y de $\mathcal{L}$ qui a la distribution μ, conditionnellement à $\mathcal{T}$. La description que l'on a de la loi de Y nous permet de contrôler des quantités de la forme $\bar{\mu}(B(Y,r))$ pour $r \to 0$. Finalement, grâce au principe de distribution de masse, on obtient la majoration $\dim_{\mathrm{H}}(\mathcal{L}) \geq \frac{1}{\alpha}$.

Minoration quand $\beta > 1$. Lorsque $\beta > 1$, la mesure $\bar{\mu}$ est portée par $\mathcal{T}^*$ et donc ne peut pas servir à produire de minoration pour la dimension de $\mathcal{L}$. On construit donc une famille de mesures sur les feuilles de manière *ad hoc*, en espérant que l'une d'entre elles ait le comportement adapté. La construction de ces mesures dépend d'un paramètre $\gamma > 1$. On construit en fait d'abord une suite de mesures $(\pi_k)_{k\geq 0}$ de telle façon à ce que pour chaque $k \geq 0$, la mesure π_k soit portée par les blocs d'indices entre n^{γ^k} et $2n^{\gamma^k}$, où n est un grand entier fixé. Comme toutes ces mesures sont portées par le compact $\mathcal{T}$, cette suite admet au moins un sous-suite qui converge étroitement vers une mesure π, et on peut vérifier que celle-ci est portée par (un sous-ensemble de) $\mathcal{L}$. Cette mesure a un comportement « multi-échelle » et pour π-presque tout x, on arrive à contrôler la masse $\pi(\mathrm{B}(x,r))$ qu'elle donne aux boules de rayon r lorsque r tend vers 0 en l'approximant à chaque échelle $r \in [n^{-\alpha\gamma^{k+1}}, n^{-\alpha\gamma^k}]$ par la mesure $\pi_k(\mathrm{B}(x,r))$. En utilisant le principe de distribution de masse, cela donne une minoration sur la dimension de Hausdorff. La minoration du théorème est obtenue en optimisant sur le paramètre γ.

1.3 Arbres récursifs pondérés et à attachement préférentiel

Dans cette section, on présente les contributions du Chapitre 3 à l'étude de propriétés asymptotiques de deux familles d'arbres discrets construits par un processus itératif d'ajout de sommets. Ces modèles sont reliés à des limites d'échelle pour des graphes discrets construits par une procédure que l'on présentera en Section 1.4.

1.3.1 Deux modèles reliés

Arbres récursifs pondérés

Pour toute suite de réels positifs $(w_n)_{n\geq 1}$ avec $w_1 > 0$, on définit la loi $\mathrm{WRT}((w_n)_{n\geq 1})$ sur les suites croissantes d'arbres enracinés, dont l'objet aléatoire associé est appelé *l'arbre récursif pondéré de suite de poids* $(w_n)_{n\geq 1}$. Une suite d'arbres $(\mathtt{T}_n)_{n\geq 1}$ sous cette loi est définie récursivement partir de l'arbre $\mathtt{T}_1$ qui ne contient que le sommet racine u_1 : pour $n \geq 1$ l'arbre $\mathtt{T}_{n+1}$ est obtenu à partir de $\mathtt{T}_n$ en ajoutant le sommet u_{n+1} étiqueté $n+1$. Le parent de ce nouveau sommet est le sommet d'étiquette K_{n+1}, où K_{n+1} est pris aléatoirement parmi $\{1,\dots,n\}$ proportionnellement aux poids $w_1, \dots, w_n$

$$\forall k \in \{1, \dots, n\}, \qquad \mathbb{P}\left(K_{n+1} = k \mid \mathtt{T}_n\right) \propto w_k.$$

On permettra aussi d'utiliser des suites de poids aléatoires $(\mathsf{w}_n)_{n\geq 1}$ et dans ce cas la loi $\mathrm{WRT}((\mathsf{w}_n)_{n\geq 1})$ est la distribution de la suite aléatoire d'arbre obtenue par le mécanisme décrit ci-dessus, conditionnellement à la suite de poids $(\mathsf{w}_n)_{n\geq 1}$.

Les WRT sont une généralisation de l'*arbre récursif uniforme* (URT), que l'on obtient si la suite des poids est constante. Beaucoup de résultats ont été prouvés depuis les années 1970 pour l'URT ainsi que pour d'autres modèles similaires, voir [50] pour un survol. En ce qui concerne les WRT, peu de travaux [91, 74] concernent leur étude générale depuis son introduction par Konstantin Borovkhov et Vladimir Vatutin [30].

Les WRT apparaissent en fait de façon implicite dans les constructions par recollements itératifs : la structure $\mathcal{T}_n$ construite par recollement dans la section précédente avec une suite de poids (w_n) peut être interprétée comme un recollement des blocs $\mathbf{b}_1, \dots \mathbf{b}_n$ le long d'un arbre récursif pondéré utilisant la même suite de poids.

Arbres à attachement préférentiel à poids initiaux

Pour toute suite $(a_n)_{n\geq 1}$ de réels, avec $a_1 > -1$ et $a_n \geq 0$ pour $n \geq 2$, on définit un autre modèle de croissance d'arbre. La construction s'obtient comme au-dessus : $\mathtt{P}_1$ ne contient qu'un sommet u_1 et $\mathtt{P}_{n+1}$ s'obtient à partir de $\mathtt{P}_n$ en ajoutant un sommet u_{n+1} étiqueté $n+1$ et le parent du nouveau venu est le sommet d'étiquette J_{n+1}, où J_{n+1} est pris aléatoirement parmi $\{1, \dots, n\}$ de telle façon à ce que

$$\forall k \in \{1, \dots, n\}, \qquad \mathbb{P}\left(J_{n+1} = k \mid \mathtt{P}_n\right) \propto \deg^+_{\mathtt{P}_n}(u_k) + a_k,$$

où $\deg^+_{\mathtt{P}_n}(\cdot)$ désigne le nombre d'enfants dans l'arbre $\mathtt{P}_n$. Quand $n = 1$, par convention, le deuxième sommet u_2 est toujours un enfant de u_1, même dans le cas $-1 < a_1 \leq 0$ pour lequel notre définition de J_2 n'a pas de sens. La suite $(\mathtt{P}_n)_{n\geq 1}$ ainsi définie est appelée *arbre à attachement préférentiel affine à poids initiaux* $(a_n)_{n\geq 1}$ et on note sa loi $\mathrm{PA}((a_n)_{n\geq 1})$.

Cette loi sur les arbres est une variation des multiples modèles d'attachement préférentiel, qui ont été extensivement étudiés depuis que Albert-Lázló Barabási et Réka Albert [15] ont proposé que ce type de dynamique explique le comportement de réseaux complexes comme celui d'Internet, en particulier en terme de répartition des degrés dans le graphe. Citons le livre de Remco van der Hofstad [75] comme référence pour un modèle proche de celui de [15].

Dans notre cas, la motivation pour étudier ce modèle provient plutôt de l'analyse de certaines constructions de graphes aléatoires discrets, que l'on détaillera dans la Section 1.4. Les arbres $\mathtt{P}_n$ ainsi construits auront la même interprétation en tant que « plan de recollement » pour ces graphes discrets que $\mathtt{T}_n$ pour la structure continue $\mathcal{T}_n$.

Les PA sont des WRT !

Le résultat suivant est central dans notre étude des arbres à attachement préférentiel. C'est en fait l'observation qui a initié tout le travail exposé dans cette section et la suivante.

Théorème 2 (WRT-représentation des PA)**.** *À toute suite* $\mathbf{a} = (a_n)_{n\geq 1}$ *de fitnesses initiales,* *on associe la suite aléatoire* $(\mathsf{w}^{\mathbf{a}}_n)_{n\geq 1} = (\mathsf{W}^{\mathbf{a}}_n - \mathsf{W}^{\mathbf{a}}_{n-1})_{n\geq 1}$ *construite de façon à ce que*

$$\mathsf{W}^{\mathbf{a}}_0 = 0, \qquad \mathsf{W}^{\mathbf{a}}_1 = 1, \qquad \forall n \geq 2, \quad \mathsf{W}^{\mathbf{a}}_n = \prod_{k=1}^{n-1} \beta_k^{-1}, \tag{1.6}$$

où les $(\beta_k)_{k\geq 1}$ *sont indépendants et ont pour lois respectives* $\mathrm{Beta}(A_k + k, a_{k+1})$, $k \geq 1$. *Alors, les lois* $\mathrm{PA}((a_n)_{n\geq 1})$ *et* $\mathrm{WRT}((\mathsf{w}^{\mathbf{a}}_n)_{n\geq 1})$ *coïncident.*

Cette connexion entre les deux modèles provient de l'évolution des degrés dans la construction de $(\mathtt{P}_n)$, qui peut être couplée avec des processus d'urnes de Pólya. En fait, l'ensemble du processus $(\mathtt{P}_n)$ peut être codé dans l'évolution d'une infinité d'urnes de Pólya *indépendantes*. Le résultat du théorème est ensuite déduit du théorème de de Finetti appliqué à ces urnes.

Quel comportement asymptotique pour $(\mathsf{w}^{\mathbf{a}}_n)_{n\geq 1}$?

À partir de la description explicite de la loi de $(\mathsf{w}^{\mathbf{a}}_n)_{n\geq 1}$ donnée par le Théorème 2, il est facile d'obtenir des informations sur son comportement par des simples calculs de moments. On introduira la condition (H_c) qui dépend d'un paramètre $c > 0$ qui est satisfaite par une suite de poids (a_n) si

$$\sum_{i=1}^{n} a_i \underset{n\to\infty}{\bowtie} c \cdot n \tag{H_c}$$

où le signe $\bowtie$ est une comparaison asymptotique légèrement plus précise que $\sim$, que nous ne détaillons pas ici.

Lorsqu'une suite (a_n) satisfait (H_c) la suite aléatoire associée $(\mathsf{W}^{\mathbf{a}}_n)$ satisfait presque sûrement

$$\mathsf{W}^{\mathbf{a}}_n \underset{n\to\infty}{\bowtie} Z \cdot n^{\gamma}, \tag{1.7}$$

avec $\gamma = \frac{c}{c+1}$ et où Z est une variable aléatoire.

Pour cette raison, toutes les propriétés asymptotiques que nous allons montrer pour les WRT dont la suite de poids satisfait ce type de comportement asymptotique vont automatiquement s'appliquer aux arbres à attachement préférentiel dont la suite de poids initiaux satisfait (H_c).

1.3.2 Propriétés des WRT

Présentons les résultats obtenus sur les WRT, pour des suites déterministes $(w_n)_{n\geq 1}$ qui vérifient

$$W_n := \sum_{i=1}^{n} w_i \underset{n\to\infty}{\bowtie} C \cdot n^{\gamma}, \tag{1.8}$$

pour un réel strictement positif γ et une constante $C > 0$. On utilisera parfois les hypothèses plus fortes $(\square^p_\gamma)$ définies pour $p \in (1\,, 2]$, pour les définitions desquelles nous renvoyons au chapitre correspondant.

Degrés

Les degrés des sommets sont très simples à étudier dans un WRT puisque le degré d'un sommet fixé u_k évolue comme une somme de variables de Bernoulli indépendantes

$$\left(\deg^+_{\mathtt{T}_n}(u_k)\right)_{n\geq 1} \overset{(\mathrm{d})}{=} \left(\sum_{i=k}^{n-1} \mathbf{1}_{\left\{U_i \leq \frac{w_k}{W_i}\right\}}\right)_{n\geq 1},$$

où les $(U_i)_{i\geq 1}$ sont une suite i.i.d. de variables uniformes sur $(0\,,1)$. De cette écriture, si $\gamma < 1$, on obtient immédiatement la convergence

$$n^{-(1-\gamma)} \cdot (\deg^+_{\mathtt{T}_n}(u_1), \deg^+_{\mathtt{T}_n}(u_2), \dots) \underset{n\to\infty}{\longrightarrow} \frac{1}{C(1-\gamma)} \cdot (w_1, w_2, \dots),$$

au sens de la topologie produit. En fait si les w_n sont au plus de l'ordre de $n^{\gamma-1}$ quand $n \to \infty$ alors la convergence a aussi lieu dans un espace de suite ℓ^p approprié.

D'après le Théorème 2 et le comportement (1.7) on peut utiliser cette convergence dans le cas d'un arbre à attachement préférentiel $\mathrm{PA}(\mathbf{a})$ pour une suite $\mathbf{a} = (a_n)_{n\geq 1}$ qui satisfait (H_c). La convergence s'exprime alors comme

$$n^{-\frac{1}{c+1}} \cdot (\deg^+_{\mathtt{P}_n}(u_1), \deg^+_{\mathtt{P}_n}(u_2), \dots) \underset{n\to\infty}{\longrightarrow} (\mathsf{m}^{\mathbf{a}}_1, \mathsf{m}^{\mathbf{a}}_2, \dots), \tag{1.9}$$

où la suite $(\mathsf{m}^{\mathbf{a}}_1, \mathsf{m}^{\mathbf{a}}_2, \dots) := \frac{c+1}{Z} \cdot (\mathsf{w}^{\mathbf{a}}_1, \mathsf{w}^{\mathbf{a}}_2, \dots)$ est proportionnelle à la suite de poids mais où la constante de proportionnalité est aléatoire. La suite limite $(\mathsf{m}^{\mathbf{a}}_n)$ a une description naturelle comme les incréments d'une chaîne de Markov inhomogène, dont la loi est explicite pour certaines suites (a_n) particulières.

Cette convergence fait écho à des résultats similaires exprimés soit dans le cadre de l'attachement préférentiel soit dans le cadre d'urnes inhomogènes : on peut citer la série d'articles [99, 100, 101] par Erol Peköz, Adrian Röllin et Nathan Ross, ou celui de Philippe Marchal et Cyril Banderier [14] qui utilise des méthodes combinatoires.

Hauteur et profil

On s'intéresse également à la hauteur des sommets dans l'arbre $\mathtt{T}_n$. Pour tout $k \geq 0$, on pose

$$\mathbb{L}_n(k) := \#\{u \in \mathtt{T}_n \mid \mathrm{ht}(u) = k\} = \#\{i \leq n \mid \mathrm{ht}(u_i) = k\},$$

le nombre de sommets de $\mathtt{T}_n$ à hauteur k. La fonction $(k \mapsto \mathbb{L}_n(k))$ est appelée le profil de l'arbre $\mathtt{T}_n$ et a été étudiée par de nombreux auteurs pour l'arbre récursif uniforme [98, 93, 49, 51, 61, 77] et aussi pour les arbres à attachement préférentiel pour des suites (a_n) constantes par Zsolt Katona dans [85].

On peut aussi se demander quelle est la hauteur maximale atteinte de l'arbre. Dans le cas de l'URT et l'arbre à attachement préférentiel à poids initiaux constants, cela été étudié par Boris Pittel dans [105], qui montre que cette hauteur est presque sûrement équivalente à une constante fois $\log n$, lorsque $n \to \infty$, où la constante est déterminée par une certaine équation. Dans le cas de l'URT, elle s'avère être égale à e, la constante de Néper, et cela apparaissait déjà dans un travail de Luc Devroye [45].

Le théorème suivant exprime un comportement similaire dans le cas des WRT dont la suite des poids cumulée croit polynomialement. Pour ce résultat on aura besoin de supposer que les poids satisfont la condition $(\square^p_\gamma)$ pour un certain $p \in (1\,,2]$, que nous avons mentionnée au-dessus.

Théorème 3. *Supposons qu'il existe $\gamma > 0$ et $p \in (1\,,2]$ tels que la suite $(w_n)_{n\geq 1}$ satisfasse $(\square^p_\gamma)$. Alors, pour une suite $(\mathtt{T}_n)_{n\geq 1}$ de loi $\mathrm{WRT}((w_n)_{n\geq 1})$, on a*

$$\mathbb{L}_n(k) \underset{n\to\infty}{=} \frac{n}{\sqrt{2\pi\gamma\log n}} \exp\left\{-\frac{1}{2}\cdot\left(\frac{k-\gamma\log n}{\sqrt{\gamma\log n}}\right)^2\right\} + O\left(\frac{n}{\log n}\right), \tag{1.10}$$

où le terme d'erreur est uniforme en $k \geq 0$. De plus, il existe un intervalle $(z_-\,, z_+)$ contenant 0, où $z_+ := \sup\{z \in \mathbb{R} \mid 1 + \gamma(e^z - 1 - ze^z) = 0\}$, tel qu'on ait presque sûrement pour tout

$z \in (z_- \, , z_+)$

$$\mathbb{L}_n\left(\lfloor \gamma e^z \log n \rfloor\right) = n^{1+\gamma(e^z - 1 - ze^z) + o(1)}. \tag{1.11}$$

On a également la convergence presque sûre suivante

$$\frac{\mathrm{ht}(\mathtt{T}_n)}{\log n} \underset{n\to\infty}{\longrightarrow} \gamma \cdot e^{z_+}. \tag{1.12}$$

La première convergence (1.10) est une forme de théorème local limite pour le profil : elle indique que le profil de l'arbre est asymptotiquement gaussien d'espérance et de variance $\gamma \log n$. La normalisation indique en particulier que chacun des niveaux à hauteur k de l'ordre $\gamma \log n \pm \sqrt{\log n}$ comprend de l'ordre de $\frac{n}{\sqrt{\log n}}$ sommets. La deuxième expression (1.11) permet de connaître l'ordre de grandeur du nombre de sommets présents à d'autres niveaux : pour un niveau de l'ordre de $\gamma e^z \log n$ on trouve un nombre de sommets d'ordre $n^{1+\gamma(e^z-1-ze^z)+o(1)}$ tant que z reste dans le domaine $z \in (z_- \, , z_+)$. Lorsque $z \to z_+$, l'exposant $1 + \gamma(e^z - 1 - ze^z)$ du nombre de sommets à hauteur $\gamma e^z \log n$ tend vers 0. La dernière convergence (1.12) indique que cette hauteur $\gamma e^{z_+} \log n$ est bien la hauteur maximale atteinte par l'arbre.

Ce résultat est obtenu en suivant une méthode classique dans l'étude d'arbre à croissance logarithmique (voir [35, 36, 85]) : on étudie la transformée de Laplace du profil $z \mapsto \sum_{k=0}^{n} e^{zk} \mathbb{L}_n(k)$ sur un ouvert du plan complexe et on prouve que, correctement renormalisée, celle-ci converge presque sûrement vers une fonction analytique. On applique ensuite un théorème de Zakhar Kabluchko, Alexander Marynych et Henning Sulzbach [84, Théorème 2.1], qui à l'aide d'arguments précis d'inversion de Fourier permet d'obtenir une convergence très précise pour $\mathbb{L}_n$. À noter que le comportement obtenu en (1.11) assure en particulier l'existence de sommets de l'arbre à toute hauteur de l'ordre $\gamma e^{(z_+ - \epsilon)} \log n$. La convergence (1.12) est ensuite obtenue en montrant une majoration correspondante en utilisant des méthodes plus rudimentaires, à partir de l'*espérance* de la transformée de Laplace $z \mapsto \mathbb{E}\left[\sum_{k=0}^{n} e^{zk} \mathbb{L}_n(k)\right]$ et de l'inégalité de Chernoff.

Mesures

Enfin, on étudie également la convergence de certaines mesures portées par les arbres $(\mathtt{T}_n)_{n\geq 1}$. Une façon d'exprimer ces résultats est de considérer que les arbres que nous construisons sont en fait des arbres plans et sont donc des sous-ensembles croissants de $\mathbb{U}$, l'arbre d'Ulam (dont la définition est donnée plus haut en (1.4)). On peut définir plusieurs mesures de probabilités sur $\mathtt{T}_n$. Une première, μ_n est celle qu'on appelle la mesure de poids, qui charge chacun des sommets $\{u_1, \dots, u_n\}$ proportionnellement à son poids $\mu_n(u_k) \propto w_k$. Une seconde est la mesure uniforme sur les sommets $\{u_1, \dots, u_n\}$ de $\mathtt{T}_n$. Une troisième, dans le cadre des PA, est la mesure d'attachement préférentiel qui charge chaque sommet u_k de $\mathtt{T}_n$ proportionnellement à $a_k + \deg^+_{\mathtt{T}_n}(u_k)$.

Théorème 4. *Sous l'hypothèse* $\sum_{i=1}^{\infty} \left(\frac{w_n}{W_n}\right)^2$*, toutes ces mesures convergent vers une même mesure limite* μ *portée par* $\partial\mathbb{U}$ *l'ensemble des rayons infinis dans* $\mathbb{U}$.

La preuve de ce théorème utilise des résultats [103] de Robin Pemantle sur les urnes de Pólya généralisées inhomogènes en temps ainsi que de la théorie élémentaire des martingales discrètes. Ce résultat permettra dans le cadre développé dans la section suivante d'intégrer « gratuitement » des convergences de mesures à nos limites d'échelle au sens de Gromov–Hausdorff.

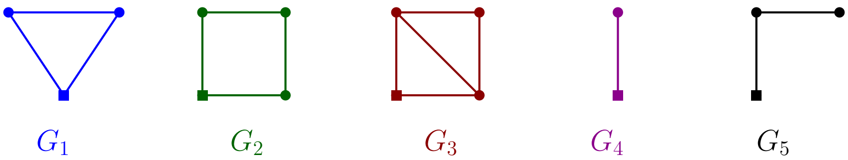
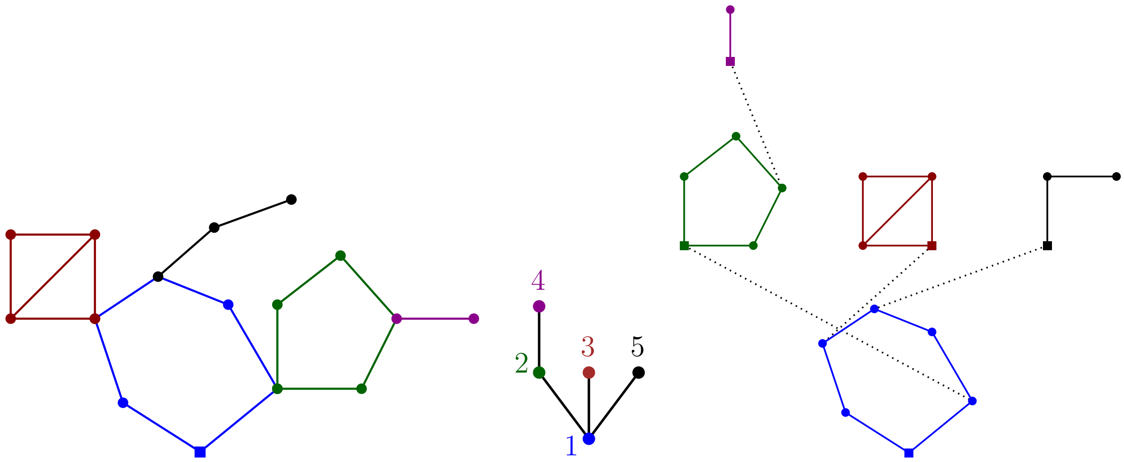


FIGURE 1.4 – Exemple de suite de graphes utilisée pour l'algorithme de Rémy



(a) Une réalisation de H_5, colorée selon la règle décrite

(b) L'arbre $\mathtt{P}_5$

(c) Le graphe H_5 vu comme un recollement de morceaux le long de $\mathtt{P}_5$

FIGURE 1.5 – Décomposition de H_5 le long de l'arbre à attachement préférentiel $\mathtt{P}_5$

1.4 Graphes discrets construits par recollements itératifs

Dans cette section, on présente les résultats du Chapitre 4. On s'intéresse à une classe de graphes aléatoires qui sont construits itérativement à la manière de l'algorithme de Rémy [107]. On montre que les modèles de cette classe admettent des limites d'échelle qui peuvent être décrites par un processus de recollements itératifs comme décrits dans la Section 1.2.

1.4.1 Des graphes construits par recollements discrets

Algorithme de Rémy généralisé

Considérons $(G_n)_{n\geq 1}$ une suite de graphes finis, connexes et enracinés. On construit une suite de graphes $(H_n)_{n\geq 1}$ récursivement comme suit. En partant de $H_1 = G_1$, on choisit à chaque étape une arête de H_n uniformément au hasard, on la divise en deux par l'addition d'un nouveau sommet en son milieu et on colle une copie de G_{n+1} au graphe obtenu en identifiant le sommet nouvellement créé au sommet racine de G_{n+1} pour ainsi obtenir H_{n+1}. C'est une généralisation de l'algorithme de Rémy [107] utilisé pour générer des arbres binaires uniformes en collant récursivement des graphes composés d'une seule arête. Dans la version originale, la suite $(n^{-1/2} \cdot H_n)$ converge vers un multiple de l'arbre brownien d'Aldous, en loi d'après les travaux d'Aldous, mais aussi presque sûrement [42]. Des versions de cette construction ont déjà été étudiées pour des suites périodiques particulières $(G_n)_{n\geq 1}$, par Bénédicte Haas et Robin Stephenson dans [71] et par Nathan Ross et Yutin Wen [109].

Décomposition le long d'un arbre à attachement préférentiel

Le point de départ de ce travail est la remarque suivante, illustrée par la Figure 1.5. En partant d'une suite arbitraire $(G_n)_{n\geq 1}$, par exemple celle représentée en Figure 1.4, on peut coupler la construction des graphes $(H_n)_{n\ \geq 1}$ avec celle d'une suite d'arbres $(\mathtt{P}_n)_{n\geq 1}$. Pour cela on peut considérer que chacun des G_n a des arêtes d'une certaine couleur, différente pour tous les n, et on impose la règle que lors de la duplication d'une arête, les arêtes produites sont de la même couleur que l'arête originale. De cette façon, l'arbre $\mathtt{P}_n$ représente les relations d'adjacence entre les couleurs (voir Figure 1.5) et il évolue comme un arbre à attachement préférentiel affine à poids initiaux donnés par les nombres d'arêtes des graphes (G_n). Ainsi les nombres d'arêtes dans chacune des parties colorées correspondent (à constante additive près) aux degrés des sommets dans un arbre à attachement préférentiel !

Limite d'échelle

Notons $\mathbf{a} = (a_n)_{n\geq 1}$ les nombres d'arêtes respectifs des $(G_n)_{n\geq 1}$. Grâce à la relation entre H_n et les degrés dans l'arbre à attachement préférentiel associé, on obtient le résultat suivant. On rappelle que l'hypothèse (H_c) qui dépend d'une constante $c > 0$ est définie dans la section précédente.

Théorème 5. *Supposons qu'il existe $c > 0$ et $c' < \frac{1}{c+1}$ tels que $(a_n)_{n\geq 1}$ satisfasse la condition (H_c) et que $a_n \leq (n+1)^{-c'+o(1)}$, alors la convergence suivante a lieu presque sûrement pour la topologie de Gromov–Hausdorff–Prokhorov*

$$(H_n, n^{\frac{-1}{c+1}}\,\mathrm{d_{gr}}, \mu_{\mathrm{unif}}) \underset{n\to\infty}{\longrightarrow} (\mathcal{H}, \mathrm{d}, \mu).$$

L'espace limite $(\mathcal{H}, \mathrm{d}, \mu)$, qui dépend de toute la suite $(G_n)_{n\geq 1}$, est naturellement décrit comme le résultat d'une construction par recollements itératifs comme décrit en Section 1.2. Les blocs $(\mathbf{b}_n)_{n\geq 1}$ utilisés pour cette construction sont des versions continues $(\mathcal{G}_n)_{n\geq 1}$ des graphes $(G_n)_{n\geq 1}$ normalisées de façon à ce que la longueur totale contenue dans chacun (somme des longueurs des arêtes) soit donnée par la suite $(\mathsf{m}_n^{\mathbf{a}})$ obtenue en (1.9) comme la limite des degrés des sommets des arbres $(\mathtt{P}_n)_{n\geq 1}$ dans leur ordre d'apparition.

Preuve et méthode générale

La preuve du Théorème 5 se base sur la décomposition décrite au-dessus en terme de sous-graphes recollés le long d'un arbre à attachement préférentiel. Cette description nous permet de considérer séparément les processus (indépendants) d'évolution de chacune des parties colorées, tout en contrôlant l'échelle de temps (donné par les degrés dans l'arbre $\mathtt{P}_n$) dans lesquelles elles évoluent. Le Théorème 6 énoncé dans la suite nous permet ensuite de conclure en vérifiant que chacune des parties colorées converge en un certain sens, et une propriété de relative compacité pour l'ensemble de la structure.

En fait, ce schéma de preuve peut s'adapter à d'autres suites de graphes aléatoires qui partagent cette structure de « graphes recollés le long d'un arbre à attachement préférentiel ». On ne détaillera pas ici leur construction mais les arbres construits par l'algorithme de Marchal [92], par l'α-modèle de Ford [60], par leur généralisation l'$\alpha - \gamma$-croissance [37] ainsi que les arbraboucles d'arbres à attachement préférentiel affine [40] peuvent être décomposés et étudiés de manière similaire.

Interprétation comme une recollement de décorations. Comme illustré en Figure 1.5, le graphe H_n peut être interprété comme le résultat $\mathscr{G}(\mathcal{D}^{(n)})$ du recollement d'une certaine décoration $\mathcal{D}^{(n)}$, dans le formalisme décrit en Section 1.1.5. L'arbre $\mathtt{P}_5$ peut être vu comme un sous-ensemble de l'arbre d'Ulam $\mathbb{U}$ donc la définition de $\mathcal{D}^{(n)}(u)$ pour $u \in \mathtt{P}_n$ est intuitivement claire d'après la figure (quitte à compléter la suite de points distingués de manière arbitraire). Afin de définir $\mathcal{D}^{(n)}$ sur $\mathbb{U}$ tout entier, on déclare simplement que $\mathcal{D}^{(n)}(u)$ est réduite à un point si $u \notin \mathtt{P}_n$. Grâce à cette représentation on peut utiliser les résultats suivants afin d'obtenir le résultat de limite d'échelle énoncé dans le Théorème 5.

Convergence. Pour prouver la convergence d'une suite d'espaces métriques construits comme $(\mathscr{G}(\mathcal{D}_n))_{n\geq 1}$ pour une suite $(\mathcal{D}_n)_{n\geq 1}$ de décorations, il suffira de vérifier une propriété de « convergence fini-dimensionnelle » associée à la propriété de relative compacité suivante

$$\inf_{\substack{\theta\subset\mathbb{U} \\ \theta \text{ arbre plan}}} \sup_{u\in\mathbb{U}} \left(\sum_{\substack{v\prec u \\ v\notin\theta}} \sup_{n\geq 1} \operatorname{diam}(\mathcal{D}_n(v)) \right) = 0. \tag{1.13}$$

Pour résumer cela on a le théorème suivant, issu du Chapitre 4

> **Théorème 6.** *Si une suite de famille de décorations $(\mathcal{D}_n)_{n\geq 1}$ est telle que pour tout $u \in \mathbb{U}$ on a la convergence suivante au sens de la topologie « Gromov–Hausdorff infiniment pointé »*
>
> $$\mathcal{D}_n(u) \underset{n\to\infty}{\longrightarrow} \mathcal{D}_\infty(u),$$
>
> *pour une décoration limite $\mathcal{D}_\infty$ et qu'elle vérifie la condition de relative compacité (1.13), alors on a la convergence*
>
> $$\mathscr{G}(\mathcal{D}_n) \underset{n\to\infty}{\longrightarrow} \mathscr{G}(\mathcal{D}_\infty) \quad \text{quand } n\to\infty,$$
>
> *au sens de Gromov–Hausdorff.*

Il est aussi possible de prendre en compte des mesures de probabilités sur les décorations, desquelles on peut prouver la convergence sous des hypothèses adaptées, améliorant ainsi cette convergence en une convergence au sens de Gromov–Hausdorff–Prokhorov.

1.5 La composante α-stable

Dans cette section, on présente des résultats issus du Chapitre 5 obtenus en collaboration avec Christina Goldschmidt et Bénédicte Haas, sur la géométrie d'un objet aléatoire appelé la composante α-stable.

1.5.1 Limite d'échelle de graphes à degrés i.i.d.

Décrivons d'abord un modèle de graphe aléatoire. Soient $D_1, D_2, \dots, D_n$ des variables aléatoires à valeurs dans $\mathbb{N}^*$ indépendantes et identiquement distribuées telles que $\mathbb{E}\left[D_1^2\right] < \infty$. On construit un graphe aléatoire sur l'ensemble de sommets $\{1, 2, \dots, n\}$. On travaille sur l'événement $\{\sum_{i=1}^n D_i \text{ est paire}\}$ et on définit G_n comme un graphe choisit uniformément au hasard parmi les graphes pour lesquels le sommet i a degré D_i pour tout $i \in [\![1\,,n]\!]$. S'il n'en existe pas, on peut décider arbitrairement que G_n n'a aucune arête, mais cela arrive avec une probabilité qui tend vers 0 lorsque $n \to \infty$, donc cela ne sera pas important pour les convergences qui vont suivre.

Michael Molloy et Bruce Reed [97] ont montré une transition de phase pour la taille des composantes connexes de ce graphe : si le paramètre $\nu := \mathbb{E}[D_1(D_1-1)]/\mathbb{E}[D_1]$ est strictement plus grand que 1, alors il existe une unique composante géante de taille proportionnelle à n, alors que si ν est plus petit que 1, il n'y a pas de composante géante. Dans le cas que l'on étudie, on choisit une distribution pour D_1 de telle façon à ce que l'on soit exactement au point critique de cette transition $\nu = 1$.

Loi des degrés. Pour les questions traitées dans ce travail, le cas $\mathbb{E}\left[D_1^3\right] < \infty$ a déjà été étudié [18] et tombe dans la même classe d'universalité que le graphe d'Erdős-Renyi critique $G(n,p)$ avec $p = 1/n$ traité par Louigi Addario-Berry, Nicolas Broutin et Christina Goldschmidt [5]. On s'intéresse donc au cas où ce moment troisième est infini et où la loi de D_1 a une queue polynomiale. Pour cela, on fixe $1 < \alpha < 2$ et on suppose que

$$\nu = 1 \quad \text{et} \quad \mathbb{P}(D_1 = k) \sim ck^{-2-\alpha} \quad \text{quand } k \to \infty, \tag{1.14}$$

où $c > 0$ est une constante. Dans ce cadre, Adrien Joseph [83] a montré que les tailles des plus grandes composantes connexes (nombre de sommets) sont d'ordre $n^{\frac{\alpha}{\alpha+1}}$.

Convergence des composantes connexes. On note $C_1^n, C_2^n, \ldots$ les composantes connexes de G_n, listées dans l'ordre décroissant de nombre de sommets, les égalités étant départagées de façon arbitraire. Les composantes sont vues comme des espaces métriques, chacune munie de la distance de graphe d_i^n. On munit aussi chacune des composantes de la mesure de comptage sur les sommets

$$\mu_i^n = \sum_{v \in C_i^n} \delta_v.$$

On note $s(C_i^n)$ le *surplus* de la composante C_i^n, c'est-à-dire le nombre d'arêtes qu'il faudrait enlever pour en faire un arbre. Le théorème suivant est (une version faible de celui) prouvé par Guillaume Conchon–Kerjan et Christina Goldschmidt dans [38].

Théorème 2. *Quand $n \to \infty$, on a la convergence en loi suivante pour la topologie Gromov–Hausdorff–Prokhorov produit*

$$\left(C_i^n, (a_D \cdot n^{-\frac{\alpha-1}{\alpha+1}}) \cdot \mathrm{d}_i^n, (b_D \cdot n^{-\frac{\alpha}{\alpha+1}}) \cdot \mu_i^n\right)_{i \geq 1} \xrightarrow[n\to\infty]{(\mathrm{d})} ((C_i, d_{C_i}, \mu_{C_i}))_{i \geq 1},$$

pour une suite aléatoire $((C_i, d_{C_i}, \mu_{C_i}))_{i \geq 1}$ que l'on appelle le graphe α-stable *et des constante a_D et b_D qui dépendent que de la loi des degrés.*

Ce résultat est un analogue α-stable du comportement brownien observé pour le graphe d'Erdős-Renyi critique dans [5].

Description de la limite. Le graphe α-stable est construit à partir d'une version biaisée d'un processus de Lévy α-stable spectralement positif. Il est constitué d'une suite d'espaces métriques mesurés qui sont des $\mathbb{R}$-*graphes* au sens de [6] i.e. ce sont localement des $\mathbb{R}$-arbres, mais ils peuvent aussi contenir des cycles. Il est possible de donner un sens au surplus d'une composante connexe limite, que l'on note $s(C_i)$, $i \geq 1$. Le Théorème 2 implique en particulier que

$$b_D \cdot n^{-\frac{\alpha}{\alpha+1}} \cdot (|C_1^n|, |C_2^n|, \ldots) \xrightarrow[n\to\infty]{(\mathrm{d})} (\mu_{C_1}(C_1), \mu_{C_2}(C_2), \ldots),$$

de façon jointe avec la convergence

$$(s(C_1^n), s(C_2^n), \dots) \xrightarrow[n\to\infty]{(\mathrm{d})} (s(C_1), s(C_2), \dots).$$

Les suites $(\mu_{C_1}(C_1), \mu_{C_2}(C_2), \dots)$ et $(s(C_1), s(C_2), \dots)$ s'expriment comme des fonctions explicites du sous-jacent. De plus, les composantes limites $(C_1, C_2, \dots)$ sont conditionnellement indépendantes sachant $(\mu_{C_1}(C_1), \mu_{C_2}(C_2), \dots)$ et $(s(C_1), s(C_2), \dots)$.

La composante α-stable à surplus s. Il existe une famille de lois, indexée par $s \geq 0$, d'espaces métriques mesurés $(\mathcal{G}^s, d^s, \mu^s)$, où μ^s est une mesure de probabilité, telle que pour tout $i \geq 1$, conditionnellement à ce que $\mu_{C_i}(C_i) = x$ et $s(C_i) = s$, on ait

$$\left(C_i, d_{C_i}, \mu_{C_i}\right) \overset{(\mathrm{d})}{=} \left(\mathcal{G}^s, x^{1-1/\alpha} \cdot d^s, x \cdot \mu^s\right). \tag{1.15}$$

Cette famille de loi ne dépend de la loi des degrés plus que par l'exposant $\alpha \in (1\,, 2]$ qui intervient dans sa queue de distribution. Pour $s = 0$, le graphe $(\mathcal{G}^s, d^s, \mu^s)$ est l'arbre α-stable. Informellement, pour $s \geq 1$, l'espace $(\mathcal{G}^s, d^s, \mu^s)$, est obtenu en choisissant s feuilles aléatoirement dans une version s-biaisée de l'arbre α-stable, construite par un changement de mesure adapté pour l'excursion du processus de Lévy α-stable codant l'arbre, puis en les recollant aléatoirement chacune sur un des points de branchement se trouvant le long de leur chemin les reliant à la racine, choisi de façon proportionnelle à leur « temps local à droite » qui se lit dans l'excursion.

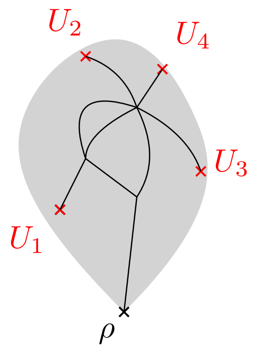
1.5.2 Propriétés de la composante α-stable de surplus s

Grâce à (1.15), l'étude des composantes du graphe α-stable se ramène simplement à celle de la famille des composantes à surplus fixé $(\mathcal{G}^s, d^s, \mu^s)$, $s \geq 0$. L'objectif de notre travail du Chapitre 5 a donc été de décrire et donner différentes constructions de ces composantes. Le cas brownien a déjà été étudié [4], et bien que nos résultats α-stables soient énoncés de manière similaire, nos preuves n'utilisent pas les mêmes outils. Dans le cas brownien, les auteurs s'appuient sur des résultats combinatoires pour comprendre la structure de la composante. Dans notre cas, on utilise la construction continue donnée par [38] à partir d'arbre stable biaisé. À partir de maintenant et pour le reste de la section, on considère une valeur de s fixée plus grande que 1.

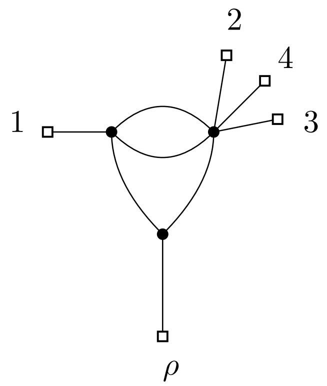
Structure cyclique et marginales discrètes

Noyau et marginales discrètes. Comme pour un graphe connexe discret, on peut décrire le graphe continu $\mathcal{G}^s$ comme une structure cyclique sur laquelle sont collés des sous-arbres. On appelle $\mathcal{K}^s$ le sous-ensemble des points qui constituent ces cycles non-triviaux auquel on ajoute le chemin qui relie cette structure à la racine[1]. On peut voir $\mathcal{K}^s$ comme un multigraphe (présence possible de boucles et arêtes multiples) enraciné muni de longueurs sur ses arêtes et on note K^s le multigraphe discret obtenu en oubliant les longueurs des arêtes, qu'on appelle le *noyau discret*. Une première information d'intérêt sur $\mathcal{G}^s$ serait la loi de son noyau discret. En fait, on calcule une information un peu plus précise ; pour $n \geq 0$, on peut tirer conditionnellement à $(\mathcal{G}^s, d^s, \mu^s)$ une suite de n feuilles $(U_i)_{1 \leq i \leq n}$ i.i.d. sous la mesure μ^s. On considère alors $\mathcal{G}_n^s$ l'ensemble constitué du noyau $\mathcal{K}^s$ ainsi que des chemins le reliant à chacune des n feuilles. De même, on note G_n^s le multigraphe obtenu en ne retenant que la forme combinatoire de $\mathcal{G}_n^s$ et les feuilles de G_n^s peuvent naturellement être numérotées de 1 à n dans leur ordre d'échantillonnage, voir Figure 1.6.

[1] Les arbres stables étant naturellement définis comme enracinés, la composante $\mathcal{G}^s$ hérite de ce point distingué et on l'appelle encore la racine



(a) L'ensemble $\mathcal{G}_4^2$ construit comme sous-ensemble de $\mathcal{G}^2$



(b) Le multigraphe discret associé G_4^2

FIGURE 1.6 – Définition de la marginale discrète G_4^2 de $\mathcal{G}^2$ à partir de feuilles uniformes U_1, U_2, U_3, U_4.

Distribution à n fixé. Afin de pouvoir exprimer la loi des marginales discrètes, on doit introduire quelques notations. Un multigraphe $G = (V, E)$ est ici vu comme un ensemble de sommets muni d'un *multi-ensemble* d'arêtes, correspondant au fait que chaque arête $e = \{u, v\}$ entre deux sommets u et v peut être présente avec une multiplicité, que l'on note $\text{mult}(e)$. On note $\text{supp}(E)$ l'ensemble des arêtes présentes avec une multiplicité supérieure à 1. Pour tout multigraphe $G = (V, E)$, on note $\text{sl}(G)$ son nombre de boucles (arêtes qui joignent un sommet à lui-même), et $I(G) \subset V$ son ensemble de sommets internes.

Pour $n \geq 0$, soit $\mathbb{M}_{s,n}$ l'ensemble des multigraphes connexes avec $n+1$ feuilles numérotées de 0 à n, surplus s et aucun sommet de degré 2 (les sommets internes ne sont pas étiquetés). On définit une suite de poids

$$w_0 := 1, \quad w_1 := 0, \quad w_2 := \alpha - 1, \quad w_k := (k-1-\alpha)\dots(2-\alpha)(\alpha-1), \quad \text{for } k \geq 3. \quad (1.16)$$

En voyant la racine comme une feuille d'étiquette 0, on peut voir G_n^s comme une variable aléatoire à valeurs dans $\mathbb{M}_{s,n}$. Le théorème suivant décrit sa loi.

Théorème 7. *Soit $n \geq 0$. Pour tout multigraphe connexe $G = (V, E) \in \mathbb{M}_{s,n}$,*

$$\mathbb{P}\left(\mathsf{G}_n^s = G\right) \propto \frac{\prod_{v \in I(G)} w_{\deg(v)-1}}{|\text{Sym}(G)| 2^{\text{sl}(G)} \prod_{e \in \text{supp}(E)} \text{mult}(e)!}.$$

En particulier cela donne la loi du noyau K^s quand $n = 0$.

Le facteur $|\text{Sym}(G)|$ au dénominateur compte le nombre de symétries de l'ensemble des sommets du graphe, et apparaît parce qu'on a considéré des sommets internes non-étiquetés. On renvoie le lecteur au chapitre correspondant pour plus de détails.

Distribution en tant que processus. En fait, on peut considérer la suite des marginales discrètes $(\mathsf{G}_n^s)_{n\geq 0}$ obtenue en échantillonnant un nombre croissant de feuilles dans la composante. Son évolution en tant que processus a une forme qui nous sera utile dans la suite : elle suit une version généralisée d'un algorithme proposé par Philippe Marchal [92] pour décrire l'évolution des marginales discrètes de l'arbre α-stable.

Une étape de cet algorithme aléatoire d'exprime de la façon suivant. Soit $G = (V, E) \in \mathbb{M}_{s,n}$ un multigraphe. On déclare que toutes les arêtes ont poids $\alpha - 1$, tous les sommets internes $u \in I(G)$ ont poids $\deg_G(u) - 1 - \alpha$ et les feuilles ont poids 0. On choisit une arête ou un sommet proportionnellement à son poids. Alors

- si c'est un sommet, on attache une arête qui lie ce sommet à une nouvelle feuille d'étiquette $n+1$,
- si c'est une arête, on attache une arête qui lie une nouvelle feuille d'étiquette $n+1$ à un sommet nouvellement créé qui divise cette arête en deux.

On dit qu'une suite de graphes évolue selon l'algorithme de Marchal si c'est un processus de Markov dont les transitions sont données par l'étape décrite au dessus. L'expression des lois des G_n^s permet de prouver le théorème suivant.

Théorème 8. *La suite $(\mathsf{G}_n^s)_{n\geq 0}$ évolue selon l'algorithme de Marchal.*

Description globale de la composante

La description de l'évolution de la suite $(\mathsf{G}_n^s)_{n\geq 0}$ dans le Théorème 8, en plus de la loi du noyau, nous permet en fait de récupérer toute la géométrie de la composante $\mathcal{G}^s$ grâce à la proposition suivante qui dérive par absolue continuité d'un résultat similaire pour l'arbre α-stable. On considère ici les graphes $(\mathsf{G}_n^s)_{n\geq 0}$ comme des espaces métriques, munis de la mesure uniforme sur leurs feuilles.

Proposition 9. *On a la convergence suivante*

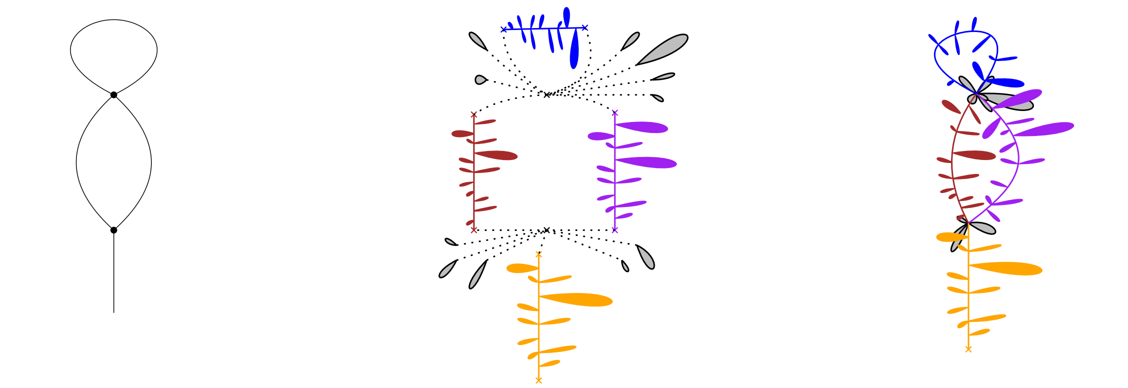
$$\frac{\mathsf{G}_n^s}{n^{1-1/\alpha}} \xrightarrow[n\to\infty]{\text{p.s.}} \alpha \cdot \mathcal{G}^s \tag{1.17}$$

pour la topologie de Gromov–Hausdorff–Prokhorov.

L'obtention de résultats sur $\mathcal{G}^s$ peut donc passer uniquement par l'étude des approximations discrètes $(\mathsf{G}_n^s)_{n\geq 0}$, pour laquelle on dispose d'outils qui incluent en particulier des couplages avec des modèles d'urnes. (C'est en fait un modèle de croissance qui tombe dans le cadre de la Section 1.2.) Cela permet de décrire deux constructions pour décrire la loi de $\mathcal{G}^s$, analogues à celles données dans le cas brownien [4].

Recollement le long du noyau. Une première façon de décrire $\mathcal{G}^s$ à partir du noyau discret K^s est la suivante : on considère une suite i.i.d. d'arbres α-stables aléatoirement normalisés selon une certaine loi, de telle façon à ce qu'on associe un arbre à chaque arête et une infinité à chaque sommet interne. La loi des normalisations est explicite en termes de lois de Dirichlet et de Poisson-Dirichlet, voir le Théorème 5.6 pour une expression explicite.

Tous ces arbres sont enracinés, mais dans les cas des arbres indexés par des arêtes, on tire en plus une seconde feuille distinguée selon la mesure naturelle portée par l'arbre. Ces arbres sont ensuite recollés selon le patron donné par le noyau : chaque arête est remplacée par l'arbre correspondant en identifiant sa racine et sa feuille distinguée à l'une et l'autre des extrémités de l'arête ; les arbres indexés par des sommets sont ensuite tous recollés sur cette structure en identifiant leur racine au sommet interne en question. Cette construction est illustrée en Figure 1.7.



(a) Une réalisation du noyau K^2

(b) Recollement d'arbres le long du noyau

(c) La composante obtenue $\mathcal{G}^2$

FIGURE 1.7 – Construction de $\mathcal{G}^2$ en recollant des arbres le long de son noyau

Construction itérative par recollements. Une seconde façon de décrire $\mathcal{G}^s$ est de partir du noyau continu $\mathcal{K}^s$ et d'y recoller itérativement des segments de longueurs aléatoires. Cette construction est une instance de processus de recollement itératif décrit dans la Section 1.2. Elle généralise celle donnée par Christina Goldschmidt et Bénédicte Haas [63] pour construire l'arbre α-stable.

De façon analogue à ce cas-là, une suite aléatoire $(\mathsf{m}_n)_{n\geq 1}$ associée à cette construction est définie comme les incréments d'une chaîne de Markov $(\mathsf{M}_n)_{n\geq 1}$ de type MLMC (Mittag-Leffler Markov chain) et joue le rôle des poids et des facteurs de normalisation dans le cadre énoncé dans la Section 1.2. Le premier bloc est une version de $\mathcal{K}^s$ munie d'une mesure qui charge toute sa longueur mais qui possède aussi des atomes de poids aléatoire sur ses sommets internes. Les blocs suivants sont constitué d'un segment muni d'une mesure qui charge sa longueur mais possède aussi un atome de masse aléatoire en l'extrémité en laquelle il est enraciné. Une version explicite cette construction est énoncée dans le Théorème 5.8 du Chapitre 5.

Chapter 2

Random gluing of metric spaces

This chapter is adapted from [112], accepted for publication in The Annals of Probability.

We construct random metric spaces by gluing together an infinite sequence of pointed metric spaces that we call *blocks*. At each step, we glue the next block to the structure constructed so far by randomly choosing a point on the structure and then identifying it with the distinguished point of the block. The random object that we study is the completion of the structure that we obtain after an infinite number of steps. In [41], Curien and Haas study the case of segments, where the sequence of lengths is deterministic and typically behaves like $n^{-\alpha}$. They proved that for $\alpha > 0$, the resulting tree is compact and that the Hausdorff dimension of its set of leaves is α^{-1}. The aim of this paper is to handle a much more general case in which the blocks are i.i.d. copies of the same random metric space, scaled by deterministic factors that we call $(\lambda_n)_{n\geq 1}$. We work under some conditions on the distribution of the blocks ensuring that their Hausdorff dimension is almost surely d, for some $d \geq 0$. We also introduce a sequence $(w_n)_{n\geq 1}$ that we call the *weights* of the blocks. At each step, the probability that the next block is glued onto any of the preceding blocks is proportional to its weight. The main contribution of this paper is the computation of the Hausdorff dimension of the set $\mathcal{L}$ of points which appear during the completion procedure when the sequences $(\lambda_n)_{n\geq 1}$ and $(w_n)_{n\geq 1}$ typically behave like a power of n, say $n^{-\alpha}$ for the scaling factors and $n^{-\beta}$ for the weights, with $\alpha > 0$ and $\beta \in \mathbb{R}$. For a large domain of α and β we have the same behaviour as the one observed in [41], which is that $\dim_{\mathrm{H}}(\mathcal{L}) = \alpha^{-1}$. However for $\beta > 1$ and $\alpha < 1/d$, our results reveal an interesting phenomenon: the dimension has a non-trivial dependence in α, β and d, namely

$$\dim_{\mathrm{H}}(\mathcal{L}) = \frac{2\beta - 1 - 2\sqrt{(\beta-1)(\beta-\alpha d)}}{\alpha}.$$

The computation of the dimension in the latter case involves new tools, which are specific to our model.

Contents

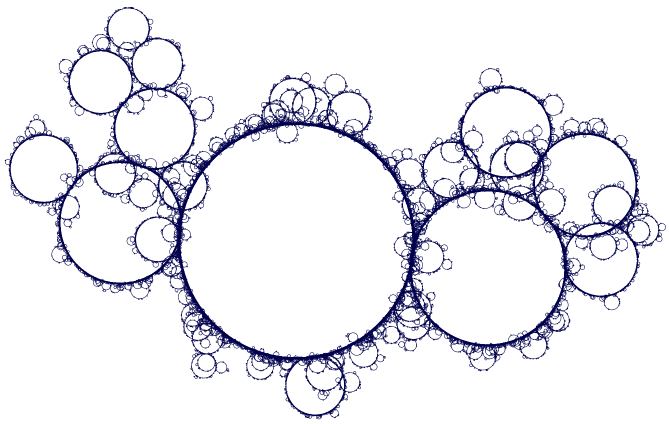


Figure 2.1 – Gluing of circles of radii $\lambda_n = n^{-3/5}$, with weights $w_n = n^{-3/2}$. The Hausdorff dimension of the resulting metric space is $(\frac{10}{3} - \sqrt{5})$.

2.1 Introduction

Let us recall Aldous' famous line-breaking construction of the Brownian CRT (Continuum Random Tree) in [9]. On the half-line $[0, \infty)$, consider $C_1, C_2, \ldots, C_n$ the points of a Poisson process with intensity tdt. Cut the half-line in closed intervals $[C_i, C_{i+1}]$, which we call *branches* (of length $C_{i+1} - C_i$). Starting from $[0, C_1]$, construct a tree by recursively gluing the branch $[C_i, C_{i+1}]$ to a random point chosen uniformly on the tree already constructed (i.e. under the normalised length measure). Aldous' Brownian CRT is the completion of the tree constructed after an infinite number of steps. This process can be generalised by using any arbitrary sequence (λ_n) for the length of the successive branches. This model was introduced and studied by Curien and Haas in [41], who proved that when $\lambda_n = n^{-\alpha+o(1)}$ for some $\alpha > 0$, the tree obtained is a.s. compact and has Hausdorff dimension $(1 \vee \alpha^{-1})$. In [12], Amini et. al. obtained a necessary and sufficient condition on the sequence (λ_n) for the almost sure compactness of the resulting tree, under the assumption that this sequence is non-increasing. In [66], Haas describes how the height of the tree explodes when $n \to \infty$ under the assumption that $\lambda_n \approx n^{\alpha}$, with $\alpha \geq 0$.

Our goal is to define a more general version of this model, in which the branches are replaced by arbitrary (and possibly random) measured metric spaces, and to investigate the compactness and the Hausdorff dimension of the resulting metric space. As we will see, in this broader context, a striking phenomenon (absent from [41]) pops up. In all this chapter we will work with

$$(\lambda_n)_{n\geq 1} \quad \text{and} \quad (w_n)_{n\geq 1},$$

two sequences of non-negative real numbers that will be the *scaling factors* and *weights* of the metric spaces that we glue. All the scaling factors $(\lambda_n)_{n\geq 1}$ are considered strictly positive, but the weights, except for the first one w_1, can possibly be null.

Definition of the model and main results Let us first present a simpler version of our construction, in which we construct a tree through an aggregation of segments. For now the branches, which we denote by $(\mathbf{b}_n)_{n\geq 1}$, are segments of length $(\lambda_n)_{n\geq 1}$, rooted at one end and endowed with the Lebesgue measure normalised so that their respective total measure is $(w_n)_{n\geq 1}$ (or endowed with the null measure for branches with vanishing weight). We then define a sequence $(\mathcal{T}_n)_{n\geq 1}$ of increasing trees by gluing those branches as follows. First, $\mathcal{T}_1 = \mathbf{b}_1$. Then, if $\mathcal{T}_n$ is constructed, we build $\mathcal{T}_{n+1}$ by first sampling a point X_n chosen proportionally to the measure μ_n obtained by aggregating the measures concentrated on the branches $\mathbf{b}_1, \dots, \mathbf{b}_n$ and then gluing $\mathbf{b}_{n+1}$ onto $\mathcal{T}_n$ by identifying its root with X_n. Let $\mathcal{T}^*$ be the increasing union of the trees $\mathcal{T}_n$ for $n \geq 1$ and $\mathcal{T}$ be the completion of $\mathcal{T}^*$. Note that if $(w_n) = (\lambda_n)$, this model coincides with the one studied in [41].

We can compute the Hausdorff dimension of the resulting tree in the case where (λ_n) and (w_n) behave like powers of n, say $\lambda_n = n^{-\alpha}$ and $w_n = n^{-\beta}$. We define $\mathcal{L} := (\mathcal{T} \setminus \mathcal{T}^*)$ to which we refer as the set of *leaves* of $\mathcal{T}$. In this particular case it coincides, up to a countable set, with the set of points x such that $\mathcal{T} \setminus \{x\}$ remains connected. In the above context a trivial consequence of our main theorem is that $\mathcal{T}$ is a.s. compact and

$$\begin{aligned} \dim_{\mathrm{H}}(\mathcal{L}) &= \frac{2\beta - 1 - 2\sqrt{(\beta-1)(\beta-\alpha)}}{\alpha} \quad \text{if } \beta > 1 \text{ and } \alpha < 1, \\ &= \frac{1}{\alpha} \quad \text{otherwise,} \end{aligned}$$

where $\dim_{\mathrm{H}}(X)$ stands for the Hausdorff dimension of the metric space X, see Section 2.A.2.

Note that, since we can check that the dimension of the *skeleton* $\mathcal{T}^*$ is always 1, we can recover the dimension of $\mathcal{T}$ as $\dim_{\mathrm{H}}(\mathcal{T}) = \max(1, \dim_{\mathrm{H}}(\mathcal{L}))$. We see that $\dim_{\mathrm{H}}(\mathcal{L}) = \frac{1}{\alpha}$ as in [41] for most values of β, however, a new phenomenon, absent from [41], happens in the case $\beta > 1$ (the sum of the weights is finite) and $\alpha < 1$ (the total length is infinite). In this case, the Hausdorff dimension of $\mathcal{T}$ depends in a non-trivial manner on α and β.

Now we want to generalise it to sequences $(\mathbf{b}_n)$ of more general metric spaces that we call *blocks*, which can be random and possibly more elaborate than just segments. Specifically, our blocks are based on the distribution of a random pointed measured compact metric space, $(\mathsf{B}, \mathsf{D}, \rho, \nu)$, with underlying set B, distance D, distinguished point ρ and endowed with a probability measure ν. We sometimes denote it B by abuse of notation when no confusion is possible and we refer to it as the *underlying random block*. We consider a sequence $((\mathsf{B}_n, \mathsf{D}_n, \rho_n, \nu_n))_{n\geq 1}$ of i.i.d. random variables with the distribution of $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ and define our blocks by setting

$$\forall n \geq 1, \quad (\mathbf{b}_n, \mathbf{d}_n, \boldsymbol{\rho}_n, \boldsymbol{\nu}_n) := (\mathsf{B}_n, \lambda_n \cdot \mathsf{D}_n, \rho_n, w_n \cdot \nu_n), \tag{2.1}$$

meaning that we dilate all the distances in the space B_n by the factor λ_n and scale the measure by w_n. We suppose that, $\lambda_n \approx n^{-\alpha}$ for some $\alpha > 0$, and $w_n \approx n^{-\beta}$ for some $\beta \in \mathbb{R}$, in some loose sense which we make precise in the sequel. For technical reasons, we have to separate the case $\beta < 1$, the case $\beta > 1$ and $\beta = 1$. This gives rise to the three hypotheses Hyp. $\bigcirc_{\alpha,\beta}$, Hyp. $\diamond_{\alpha,\beta}$ and Hyp. $\square_{\alpha,1}$. For any $d \in [0, \infty)$, we will introduce the Hypothesis H_d and suppose that the distribution of our underlying random block $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ satisfies this hypothesis for some $d \geq 0$. This hypothesis ensures that our random block exhibits a d-dimensional behaviour. We set out all these hypotheses just below the statement of our theorem.

Except in Section 2.2.1, we will always assume that the blocks are of the form (2.1). This is implicit in all our results.

In this extended setting, we can perform the same gluing algorithm and build a sequence $(\mathcal{T}_n)_{n\geq 1}$ of random compact metric spaces by iteratively gluing the root of $\mathbf{b}_{n+1}$ onto a point chosen in $\mathcal{T}_n$ according to the measure μ_n obtained as the sum of the measures of the blocks $\mathbf{b}_1, \dots, \mathbf{b}_n$. Again $\mathcal{T}^* = \bigcup_{n\geq 1} \mathcal{T}_n$ is called the *skeleton* of the construction and its completion is still denoted $\mathcal{T}$. See Figure 2.1 for a non-isometric, non-proper representation in the plane of a simulation of this model, with B chosen to be almost surely a circle of unit length. As for the case of segments, we refer to $\mathcal{L} = (\mathcal{T} \setminus \mathcal{T}^*)$ as the set of leaves of the construction. We can now state our main theorem.

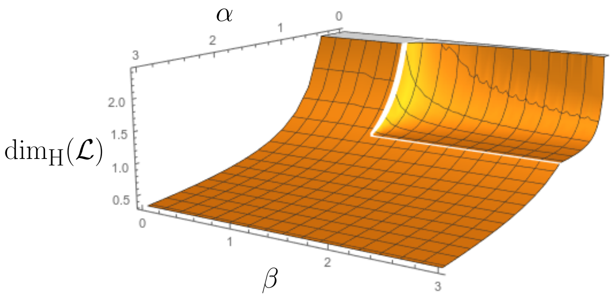


Figure 2.2 – The plot represents the Hausdorff dimension of the leaves as a function of α and β, the dimension d being fixed to 1. The expression obtained for $\beta > 1$ and $\alpha < \frac{1}{d}$ can be rewritten as $d + \dfrac{(\sqrt{\beta - \alpha d} - \sqrt{\beta - 1})^2}{\alpha}$ This expression is always larger than d and smaller than $\frac{1}{\alpha}$, and it is decreasing in α and β on the domain on which we consider it. When $\beta \to 1$, it converges to the value $\frac{1}{\alpha}$ so that the function $(\alpha, \beta) \mapsto \dim_H(\mathcal{L})$ is continuous on the domain $\mathbb{R}_+^* \times \mathbb{R}$.

Theorem 2.1. *Suppose that there exists $d \geq 0$, such that $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ satisfies Hypothesis H_d, and $\alpha > 0$ and $\beta \in \mathbb{R}$ such that the sequences (w_n) and (λ_n) satisfy either Hyp. $\diamond_{\alpha,\beta}$ or Hyp. $\bigcirc_{\alpha,\beta}$, or Hyp. $\square_{\alpha,1}$. Then, almost surely, the structure $\mathcal{T}$ resulting from the construction is compact, and*

$$\begin{aligned} \dim_{\mathrm{H}}(\mathcal{L}) &= \frac{2\beta - 1 - 2\sqrt{(\beta-1)(\beta - \alpha d)}}{\alpha}, \quad \text{if } \beta > 1 \text{ and } \alpha < \frac{1}{d}, \\ &= \frac{1}{\alpha} \quad \text{otherwise.} \end{aligned}$$

Remark that for $\beta > 1$ the dimension of the set $\mathcal{L}$ depends on the geometry of the underlying random block through d, its dimension. For $\beta \leq 1$, it is not the case, and actually the theorem remains true under much weaker hypotheses for the distribution of $(\mathsf{B}, \mathsf{D}, \rho, \nu)$, namely that ν is not almost surely concentrated on $\{\rho\}$, and that $\forall k \geq 0, \mathbb{E}\left[(\mathrm{diam}(\mathsf{B}))^k\right] < \infty$, where $\mathrm{diam}(\cdot)$ denotes the diameter of a metric space. We could even replace the assumption that the blocks $((\mathsf{B}_n, \mathsf{D}_n, \rho_n, \nu_n))_{n\geq 1}$ are i.i.d. by some weaker assumption but we do not do it for the sake of clarity. The proofs when $\beta \leq 1$ are quite short and the interested reader can easily generalise them to a more general setting.

Hypotheses of the theore. Let us define and discuss the precise hypotheses of our theorem. First, let us describe the assumptions that we make on the sequences (λ_n) and (w_n). We define

$$W_n = \sum_{k=1}^{n} w_k,$$

and for all $\epsilon > 0$, we set

$$G^\epsilon := \left\{ k \geq 1 \;\middle|\; w_k \geq k^{-\beta-\epsilon},\ \lambda_k \geq k^{-\alpha-\epsilon} \right\}, \tag{2.2}$$

and also

$$G_n^\epsilon := \left\{ k \in [\![n\, , 2n]\!] \;\middle|\; w_k \geq n^{-\beta-\epsilon},\ \lambda_k \geq n^{-\alpha-\epsilon} \right\}. \tag{2.3}$$

As said earlier, we separate the case $\beta < 1$, the case $\beta > 1$ and the case $\beta = 1$.

Hypothesis $\bigcirc_{\alpha,\beta}$. *We have $\alpha > 0$ and $\beta < 1$ and for all $n \geq 1$, $\lambda_n \leq n^{-\alpha+o(1)}$ and $w_n \leq n^{-\beta+o(1)}$. Furthermore $W_n = n^{1-\beta+o(1)}$ and for all $\epsilon > 0$,*

$$\liminf_{n\to\infty} \frac{\sum_{k=1}^n w_k \mathbf{1}_{\{k\in G^\epsilon\}}}{\sum_{k=1}^n w_k} > 0.$$

The last display ensures that for all $\epsilon > 0$, the set G^ϵ contains an asymptotically positive proportion of the total weight.

Hypothesis $\diamond_{\alpha,\beta}$. *We have $\alpha > 0$ and $\beta > 1$ and for all $n \geq 1$, $\lambda_n \leq n^{-\alpha+o(1)}$ and $w_n \leq n^{-\beta+o(1)}$. Furthermore, for all $\epsilon > 0$,*

$$\# G_n^\epsilon \underset{n\to\infty}{=} n^{1+o(1)}.$$

Under the stronger assumption $\lambda_n = n^{-\alpha+o(1)}$ and $w_n = n^{-\beta+o(1)}$, Hypothesis $\diamond_{\alpha,\beta}$ holds if $\beta > 1$ (resp. Hypothesis $\bigcirc_{\alpha,\beta}$, if $\beta < 1$). The case $\beta = 1$ is slightly different and in this case we set

Hypothesis $\square_{\alpha,1}$. *We have $\alpha > 0$ and $\beta = 1$ and for all $n \geq 1$, $\lambda_n \leq n^{-\alpha+o(1)}$ and $w_n \leq n^{-1+o(1)}$. Furthermore, for all $\epsilon > 0$,*

$$\frac{1}{\log\log\log N} \sum_{k=N}^{N^{1+\epsilon}} \frac{w_k}{W_k} \mathbf{1}_{\{k\in G^\epsilon\}} \underset{N\to\infty}{\longrightarrow} +\infty.$$

Note that this last hypothesis requires in particular that $W_n \to \infty$ as $n \to \infty$. Now let us define Hypothesis H_d, for any $d \geq 0$, which will ensure that our random underlying block has the appropriate d-dimensional behaviour.

Hypothesis H_d. *The law of the block $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ satisfies the following conditions:*

(i) $\bullet$*If $d = 0$, the block B is a finite metric space which is not a.s. reduced to a single point and such that the measure ν satisfies $\nu(\{x\}) > 0$, for all points $x \in \mathsf{B}$.*

$\bullet$*If $d > 0$, there exists an increasing function $\varphi : [0,1] \to [0, d/2]$, satisfying $\lim_{r\to 0} \varphi(r) = 0$, such that almost surely, there exists a (random) $r_0 \in (0,1)$ such that*

$$\forall r \in [0, r_0),\ \forall x \in \mathsf{B}, \qquad r^{d+\varphi(r)} \leq \nu(\mathrm{B}(x,r)) \leq r^{d-\varphi(r)}. \tag{$\star_{r_0}$}$$

(ii) *Let $N_r(\mathsf{B})$ be the minimal number of balls of radius r needed to cover B. Then*

$$\mathbb{E}[N_r(\mathsf{B})] \leq r^{-d+o(1)} \quad as \quad r \to 0.$$

(iii) *For all $k \geq 0$, we have $\mathbb{E}\left[\mathrm{diam}(\mathsf{B})^k\right] < \infty$.*

Here $\mathrm{B}(x,r)$ is the open ball centred at x with radius r and the notation $\mathrm{diam}(\mathsf{B})$ denotes the diameter of B, defined as the maximal distance between two points of B. The conditions (i) and (ii) ensure that the blocks that we glue together have dimension d. The condition (iii) ensures that the blocks cannot be too big. In this chapter, some results are stated under some weaker assumptions on the distribution of random block $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ and they are hence all still valid under Hypothesis H_d.

Motivations. The assumptions of Theorem 2.1 are rather general and various known models fall into our setting. First, let us cite two constructions that were already covered by the work presented in [41]. Of course we have Aldous' line-breaking construction of the CRT but let us also cite the work of Ross and Wen in [109], in which the authors study a discrete model of growing trees and prove that its scaling limit can be described as a line-breaking procedure à la Aldous using a Poisson process of intensity $t^l \mathrm{d}t$, with l an integer. The Hausdorff dimension of the resulting tree is then $(l+1)/l$. Our extended setting now also includes the *Brownian looptree*, defined in [40], which appears as the scaling limit of the so-called *discrete looptree* associated with the Barabási-Albert model. This random metric space also has a natural construction through an aggregation of circles, and our theorem proves that this object has almost surely Hausdorff dimension 2. These examples do not really use our theorem in its full generality since their underlying block is deterministic. In fact, Hypothesis H_d is very general and is satisfied (for the appropriate $d \geq 0$) by many distributions of blocks, including the Brownian CRT ($d = 2$), see [54], the Brownian map ($d = 4$), see [113, 88], the θ-stable trees ($d = \frac{\theta+1}{\theta}$), see [55]. Hence, our results can apply to a whole variety of such constructions, with a very general distribution of the blocks, and we are currently working on some examples in which this construction naturally arises as the limit of discrete models.

Indications on the proofs The computations of the dimension in Theorem 2.1 differ, depending on the assumptions we make on α and β, and always consist of an upper bound, that we derive by providing explicit coverings, and a lower bound that arises from the construction of a probability measure satisfying the assumptions of Frostman's lemma, see Lemma 2.20 in the Appendix for a statement.

If we just assume that the scaling factors are smaller than $n^{-\alpha+o(1)}$, we can prove that the dimension is bounded above by $\frac{1}{\alpha}$ for rather general behaviours of the weights. To do so, we adapt arguments from [41] to our new setting. The essential idea behind the proof is that the sub-structure descending from a block $\mathbf{b}_n$ has size $n^{-\alpha+o(1)}$, and so one only needs to cover every block $\mathbf{b}_n$ with a ball of radius $n^{-\alpha+o(1)}$ to cover the whole structure.

When $\alpha < \frac{1}{d}$ and $\beta > 1$, although the sub-structure descending from a block $\mathbf{b}_n$ may have diameter of order $n^{-\alpha+o(1)}$, we can also check that the index of the first block glued on block n has index roughly n^{β}, which is large compared to n. Hence the diameter of the substructure descending from $\mathbf{b}_n$ is essentially due to $\mathbf{b}_n$ itself. This gives us a hint that we can cover the whole substructure descending from the block $\mathbf{b}_n$, using a covering of $\mathbf{b}_n$ with balls that are really small compared to the size of $\mathbf{b}_n$, and that it would lead to a more optimal covering. In fact we use these two observations to recursively construct a sequence of finer and finer coverings, which lead to the optimal upper-bound. The idea of the proof is presented in more details in Section 2.4.2.

Concerning the lower bounds, for all values of α and β, we can define a natural probability measure $\bar{\mu}$ on $\mathcal{T}$ as the limit of (a normalised version of) the measure μ_n defined on $\mathcal{T}_n$ for every $n \geq 1$, see Section 2.3. In the case $\beta \leq 1$, this probability measure only charges the leaves of $\mathcal{T}$, and an application of Lemma 2.20 gives the lower bound $\frac{1}{\alpha}$.

For $\beta > 1$, the measure $\bar{\mu}$ does not charge the leaves and so the preceding argument does not work. We construct another measure as the sub-sequential limit of a sequence of measures (π_k) which are concentrated on sets of the form $(\mathcal{T}_{2n_k} \setminus \mathcal{T}_{n_k})$ with (n_k) chosen appropriately, see Section 2.5.2 for a presentation of the idea of the proof. The limiting measure is then concentrated on a strict subset of leaves and again, using Lemma 2.20 yields the appropriate lower bound.

Related constructions Let us also cite some other models that have been studied in the literature and which share some features with ours. First, the line-breaking construction of the scaling limit of critical random graphs in [4] by Addario-Berry, Broutin and Goldschmidt, that of the stable trees in [63] by Goldschmidt and Haas, and that of the stable graphs in [64] by Goldschmidt, Haas and Sénizergues, use a gluing procedure that is identical to ours. Their constructions are not directly handled by Theorem 2.1 but they fall in a slightly more general setting, for which our proofs still hold. In [30], Borovkov and Vatutin study a discrete tree constructed recursively, which corresponds to the "genealogical tree" of the blocks in our model. Last, in [106], Rembart and Winkel study the distribution of random trees that satisfy a self-similarity condition (in law). They provide an iterative construction of those trees in which infinitely many branches are glued at each step.

Plan of the chapter In Section 2.2, we give a rigorous definition of our model, set up some useful notation, and discuss some general properties. In the second section, we study the (normalised) natural measure $\bar{\mu}_n$ on $\mathcal{T}_n$ and prove that it converges to a measure $\bar{\mu}$ on $\mathcal{T}$ under suitable assumptions. In Section 2.4.1, we prove the almost sure compactness of $\mathcal{T}$ and some upper-bounds on its Hausdorff dimension under some relatively weak hypotheses. In Section 2.4.2, we develop a new (more involved) approach that allows us to obtain a better upper-bound for some parameters for which the former fails to be optimal. In Section 2.5, we prove the lower bounds that match the upper-bounds obtained in Section 2.4. It is again divided in two subsections, each providing a proof that is only valid for some choices of parameters α and β. The Appendix 2.A.2 contains a short reminder of basic properties concerning Hausdorff dimension. The Appendices 2.A.1, 2.A.3 and 2.A.4 contain some technical proofs that can be skipped at first reading.

Acknowledgements The author would like to thank the anonymous referees for their valuable comments which helped to improve the presentation of the manuscript.

2.2 General framework

In this section, we start by providing a precise definition of our model and then we investigate some of its general properties.

2.2.1 Construction

Consider $((\mathbf{b}_n, \mathbf{d}_n, \boldsymbol{\rho}_n, \boldsymbol{\nu}_n))_{n\geq 1}$ a sequence of compact pointed metric spaces endowed with a finite Borel measure. Recall from the introduction the heuristics of our recursive construction. We define $\mathcal{T}_1$ as the first block $\mathbf{b}_1$ endowed with its measure $\boldsymbol{\nu}_1$. Then at each step, we construct $\mathcal{T}_{n+1}$ from $\mathcal{T}_n$ by gluing the root of the block $\mathbf{b}_{n+1}$ to a random point $X_n \in \mathcal{T}_n$, which has distribution (a normalised version of) μ_n. The measure μ_{n+1} is defined as the sum of the measures μ_n and $\boldsymbol{\nu}_{n+1}$, the measure supported by $\mathbf{b}_{n+1}$. We define $\mathcal{T}^*$ as the increasing union of all the $\mathcal{T}_n$ for $n \geq 1$, and its completion is denoted $\mathcal{T}$. In the next paragraph, we describe formally how to construct such growing metric spaces as subsets of a larger ambient space. The definitions here are rather technical and the proofs in the paper do not use the details of the construction, so the reader can skip this part at first reading.

Embedded construction We consider (U, δ) the Urysohn space, and fix a point $u_0 \in U$. The space U is defined as the only Polish metric space (up to isometry) which has the following

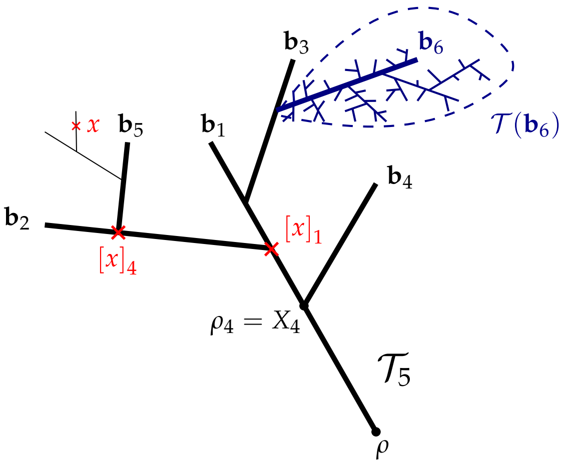


Figure 2.3 – Substructure descending from a set, and projection on a substructure, illustrating some notation introduced in the paper in the case of the gluing of segments.

extension property (see [78] for constructions and basic properties of U): given any finite metric space X, and any point $x \in X$, any isometry from $X \setminus \{x\}$ to U can be extended to an isometry from X to U. In the rest of the construction, we assume that the measured metric spaces $((\mathbf{b}_n, \mathbf{d}_n, \boldsymbol{\rho}_n, \boldsymbol{\nu}_n))_{n\geq 1}$ are all embedded in the space U and that their root is identified to u_0. From the properties of the Urysohn space, this is always possible (see Appendix 2.A.1 for a construction in the case of random blocks).

We introduce

$$\ell^1(U, u_0) := \left\{ (x_n)_{n\geq 1} \in U^{\mathbb{N}^*} \,\middle|\, \sum_{n=1}^{\infty} \delta(x_n, u_0) < +\infty \right\}.$$

If we endow $\ell^1(U, u_0)$ with the distance $\mathrm{d}((x_n)_{n\geq 1}, (y_n)_{n\geq 1}) = \sum_{n=1}^{\infty} \delta(x_n, y_n)$, it is an easy exercise to see that it makes this space Polish. We can now construct the $\mathcal{T}_n$ recursively, by $\mathcal{T}_1 = \{(x, u_0, u_0, \dots) \mid x \in \mathbf{b}_1\}$, and identifying $\mathcal{T}_1$ to the block $\mathbf{b}_1$, we set $\mu_1 = \boldsymbol{\nu}_1$. For $n \geq 1$, the point X_n is sampled according to $\bar{\mu}_n$ a normalised version of μ_n with total mass 1. The point X_n is of the form $\left(x_1^{(n)}, x_2^{(n)}, \dots, x_n^{(n)}, u_0, \dots\right)$ and we set

$$\mathcal{T}_{n+1} := \mathcal{T}_n \cup \left\{ \left(x_1^{(n)}, x_2^{(n)}, \dots x_n^{(n)}, x, u_0 \dots\right) \,\middle|\, x \in \mathbf{b}_{n+1} \right\}.$$

We set $\mu_{n+1} := \mu_n + \boldsymbol{\nu}_n$, where as in the preceding section, we see $\mathbf{b}_{n+1}$ as the corresponding subset of $\mathcal{T}_{n+1}$. Then $\mathcal{T}^* = \bigcup_{n\geq 1} \mathcal{T}_n$ and $\mathcal{T} = \overline{(\mathcal{T}^*)}$ is its closure in the space $(\ell^1(U, u_0), \mathrm{d})$. At the end $\mathcal{T}$ is a random closed subset of a Polish space.

In the rest of the chapter, we will not refer to this formal construction of $\mathcal{T}$ and we will identify $\mathbf{b}_n$ with the corresponding subset in $\mathcal{T}$. We recall the notation $W_n = \sum_{k=1}^{n} w_n$ for the total mass of the measure μ_n.

2.2.2 Some notation

Let us introduce some notation that will be useful in the sequel, some of which is illustrated in Figure 2.3. Recall that from now on, we always assume that the blocks are of the form (2.1).

• If (E, d, ρ) is a pointed metric space, and $x \in E$, we define $\mathrm{ht}(x)$, the height of x, as its distance to the root $\mathrm{d}(\rho, x)$. We also denote $\mathrm{ht}(E) = \sup_{x\in E} \mathrm{ht}(x)$, the height of E. Let us consider $(\mathsf{B}, \mathsf{D}, \rho, \nu)$, a random block of our model before scaling, and X a point of B which conditionally on $(\mathsf{B}, \mathsf{D}, \rho, \nu)$, has distribution ν. We denote

$$\mathcal{H} := \mathrm{ht}(X) = \mathsf{D}(\rho, X), \tag{2.4}$$

the height of a uniform random point in the block. Remark that Hypothesis H_d implies that $\mathbb{E}\left[\mathcal{H}^2\right] < \infty$, and that $\mathbb{P}(\mathcal{H} > 0) > 0$. Some of our results are stated under these weaker assumptions.

• Whenever we sample the point X_n under $\bar{\mu}_n$, we do it in the following way: first we sample K_n such that for all $1 \leq k \leq n$, $\mathbb{P}(K_n = k) = \frac{w_k}{W_n}$ and then, conditionally on $K_n = k$, the point X_n is chosen on the block $\mathbf{b}_k$ using the normalised version of the measure $\boldsymbol{\nu}_k$. Whenever $K_n = k$, we say that $\mathbf{b}_{n+1}$ is grafted onto $\mathbf{b}_k$ and write $\mathbf{b}_{n+1} \to \mathbf{b}_k$. Remark that this entails that $X_n \in \mathbf{b}_k$, but this condition is not sufficient in the case where X_n belongs to several blocks (which only happens if the measures carried by the blocks have atoms). We denote

$$\bar{\mu}_n^* := \text{law of } (K_n, X_n), \tag{2.5}$$

seen as a measure on $\bigsqcup_{k=1}^n \{k\} \times \mathbf{b}_k$. In this way, the random variables $((K_n, X_n))_{n\geq 1}$ are independent with respective distributions $(\bar{\mu}_n^*)_{n\geq 1}$. We remain loose on the fact that we sometimes consider the blocks as abstract metric spaces and at other times we see them as subsets of $\mathcal{T}$. It is implicit in the preceding discussion that everything is expressed conditionally on the sequence of blocks $(\mathbf{b}_n)_{n\geq 1}$.

• We simultaneously construct a sequence of increasing discrete trees $(\mathsf{T}_n)_{n\geq 1}$ by saying that for $n \geq 1$, the tree $\mathsf{T_n}$ has n nodes labelled 1 to n and i is a child of j if and only if $\mathbf{b}_i \to \mathbf{b}_j$. Also define T their increasing union. We denote $\prec$ the genealogical order on $\mathbb{N}^*$ induced by this tree. We denote $\mathrm{d}_{\mathsf{T}}(i,j)$ for the graph distance between the nodes with label i and j in this tree and $\mathrm{ht}_{\mathsf{T}}(\cdot)$ for their height.

• For $x \in \mathcal{T}$, we define $[x]_n$, the projection of x on $\mathcal{T}_n$, as the unique point y of $\mathcal{T}_n$ that minimizes the distance $\mathrm{d}(x,y)$.

• Similarly, for $k \geq 1$, we define $[k]_n$, the projection of k on T_n, as the unique node $i \leq n$ that minimizes the distance $\mathrm{d}_{\mathsf{T}}(i,k)$.

• If S is a subset of a block $\mathbf{b}_n$ for some $n \geq 1$ then we define $\mathcal{T}(S)$, the substructure descending from S as

$$\mathcal{T}(S) := \overline{S \cup \bigcup_{\substack{i \succ n \\ [X_{i-1}]_n \in S}} \mathbf{b}_i}.$$

If $S = \mathbf{b}_n$, this reduces to

$$\mathcal{T}(\mathbf{b}_n) = \overline{\bigcup_{i \succeq n} \mathbf{b}_i},$$

and we consider $(\mathcal{T}(\mathbf{b}_n), \mathrm{d}, \boldsymbol{\rho}_n)$ as a rooted metric space.

• Remark that if $x \in \mathcal{T}(\mathbf{b}_k)$ for some $k \geq 1$ then we have $[x]_k \in \mathbf{b}_k$ and more generally, for any $n \leq k$, we have $[x]_n \in \mathbf{b}_{[k]_n}$.

• We often use the little-o notation and denote $o(1)$ a deterministic function that tends to 0 when some parameter tends to 0 or ∞, depending on the context. For such functions that are random, we write instead $o^\omega(1)$.

2.2.3 Zero-One law for compactness, boundedness and Hausdorff dimension

The main properties of $\mathcal{T}$ and $\mathcal{L}$ that we study are compactness and Hausdorff dimension. One can check that some of these properties are constants almost surely by an argument using Kolmogorov's zero-one law.

Indeed, take the whole construction $\mathcal{T}$ and contract the compact subspace $\mathcal{T}_n$ into a single point. We can easily check that the resulting space is compact (resp. bounded) iff the former is compact (resp. bounded). Also, the subset $\mathcal{L}$ and its image after the contraction of $\mathcal{T}_n$ have the same Hausdorff dimension. Now remark that the space that we just described only depends on the randomness of the blocks and the gluings after n steps. Indeed, if we start at time n with a unique point with weight W_n and then follow the procedure by gluing recursively $\mathbf{b}_{n+1},\ \mathbf{b}_{n+2}, \dots,$ we get exactly the same space.

Hence, as this is true for all n, these properties only depend on the tail σ-algebra generated by the blocks and the gluings, and are therefore satisfied with probability 0 or 1.

Remark 2.2. *In the setting of [41], where the blocks are segments and the weights correspond to the lengths of those segments, the authors proved that the event of boundedness and compactness for $\mathcal{T}$ coincide almost surely. This is not the case in our more general setting: consider the case of branches with weights and lengths defined as*

$$\begin{aligned} w_n = 2^n,\ \lambda_n = 2^{-n} \quad &\textit{for } n \notin \left\{2^k \;\middle|\; k \in \mathbb{N}\right\}, \\ w_{2^k} = 1,\ \lambda_{2^k} = 1, \quad &\textit{for } k \in \mathbb{N}. \end{aligned}$$

In this case, an application of Borel-Cantelli lemma shows that a.s. for n large enough, no branch $\mathbf{b}_n$ is ever grafted onto a branch $\mathbf{b}_{2^k}$ for any k. It is then clear that the resulting tree is a.s. bounded since the sum of the lengths of the branches $\mathbf{b}_n$ for $n \notin \{2^k \mid k \in \mathbb{N}\}$ is finite, but it cannot be compact since there exists an infinite number of branches with length 1.

2.2.4 Monotonicity of Hausdorff dimension

Let us present an argument of monotonicity of the Hausdorff dimension of $\mathcal{L}$ with respect to the sequence (λ_n), on the event on which $\mathcal{T}$ is compact. Let (w_n) be a sequence of weights and (λ_n) and (λ'_n) be two sequences of scaling factors such that for all $n \geq 1$, we have $\lambda_n \geq \lambda'_n$. Suppose that $((\mathsf{B}_n, \mathsf{D}_n, \rho_n, \nu_n))_{n\geq 1}$ is a sequence of random compact metric spaces endowed with a probability measure. Then, let $\mathcal{T}$ (resp. $\mathcal{T}'$) be the structure constructed using the blocks $(\mathbf{b}_n, \mathbf{d}_n, \boldsymbol{\rho}_n, \boldsymbol{\nu}_n) = (\mathsf{B}_n, \lambda_n \cdot \mathsf{D}_n, \rho_n, w_n \cdot \nu_n)$, for $n \geq 1$, (resp. $(\mathbf{b}'_n, \mathbf{d}'_n, \boldsymbol{\rho}'_n, \boldsymbol{\nu}'_n) = (\mathsf{B}_n, \lambda'_n \cdot \mathsf{D}_n, \rho_n, w_n \cdot \nu_n)$). Note that since we use the same sequence of weights we can couple the two corresponding gluing procedures.

Let f be the application that maps each of the block $\mathbf{b}_n$ to the corresponding $\mathbf{b}'_n$. Recall here that we see the blocks as subsets of the structure. We can verify that $f : \mathcal{T}^* \longrightarrow (\mathcal{T}')^*$, is 1-Lipschitz. We can then extend uniquely f to a function $\hat{f} : \mathcal{T} \longrightarrow \mathcal{T}'$, which is also 1-Lipschitz. Suppose $\mathcal{T}$ is compact. Then its image $\hat{f}(\mathcal{T})$ is compact, hence closed in $\mathcal{T}'$. Since $(\mathcal{T}')^* \subset \hat{f}(\mathcal{T})$ and $(\mathcal{T}')^*$ is dense in $\mathcal{T}'$, we have $\hat{f}(\mathcal{T}) = \mathcal{T}'$ and so $\hat{f}$ is surjective. Now since $(\mathcal{T}')^* = \hat{f}(\mathcal{T}^*)$, we also have $\mathcal{L}' = \hat{f}(\mathcal{L})$, and since $\hat{f}$ is Lipschitz,

$$\dim_{\mathrm{H}}(\mathcal{L}') \leq \dim_{\mathrm{H}}(\mathcal{L}). \tag{2.6}$$

2.3 Study of a typical point

In this section we study the height of a typical point of $\mathcal{T}_n$, i.e. the distance from the root to a point sampled according to $\bar{\mu}_n$. The proofs in this section are really close to those of [41, Section 1], to which we refer for details.

2.3.1 Coupling with a marked point

We construct a sequence of points $(Y_n)_{n\geq 1}$ coupled with the sequence $(\mathcal{T}_n)_{n\geq 1}$ in such a way that for all $n \geq 1$, the point Y_n has distribution $\bar{\mu}_n$ conditionally on $\mathcal{T}_n$ and such that the distance from Y_n to the root is non-decreasing in n. For technical reasons, we in fact define a sequence $((J_n, Y_n))_{n\geq 1}$ such that for any $n \geq 1$, (J_n, Y_n) has distribution $\bar{\mu}_n^*$ conditionally on $(\mathsf{T}_n, \mathcal{T}_n)$, see (2.5). The properties of this construction are stated in the following lemma.

Lemma 2.3. *We can couple the construction of* $((\mathsf{T}_n, \mathcal{T}_n))_{n\geq 1}$ *with a sequence* $((J_n, Y_n))_{n\geq 1}$ *such that for all* $n \geq 1$,

(i) *we have* $J_n \in \{1, \dots, n\}$ *and* $Y_n \in \mathbf{b}_{J_n}$,

(ii) *conditionally on* $(\mathsf{T}_n, \mathcal{T}_n)$, *the couple* (J_n, Y_n) *has distribution* $\bar{\mu}_n^*$,

(iii) *for all* $1 \leq k \leq n$, $([J_n]_k, [Y_n]_k) = (J_k, Y_k)$.

Furthermore, under the assumption H_d*(iii), the sequence* $(Y_n)_{n\geq 1}$ *almost surely converges in* $\mathcal{T}$ *iff*

$$\sum_{n=1}^{\infty} \frac{\lambda_n w_n}{W_n} \mathbf{1}_{\{\lambda_n \leq 1\}} < \infty \quad \textit{and} \quad \sum_{n=1}^{\infty} \frac{w_n}{W_n} \mathbf{1}_{\{\lambda_n > 1\}} < \infty. \tag{2.7}$$

Note that if either

$$W_\infty := \sum_{n=1}^{\infty} w_n < \infty \quad \text{or} \quad \sum_{n=1}^{\infty} \frac{w_n \lambda_n}{W_n} < \infty,$$

then (2.7) is satisfied, and this is the case under the assumptions of Theorem 2.1. In this case we let

$$Y := \lim_{n\to\infty} Y_n.$$

Proof. Let $n \geq 2$. Conditionally on $\mathcal{T}_n$ and T_n, sample a couple (J_n, Y_n) under the measure $\bar{\mu}_n^*$. Then two cases may happen:

- with probability $1 - w_n/W_n$: we have $J_n < n$, so the point Y_n belongs to $\mathcal{T}_{n-1}$, that is $[Y_n]_{n-1} = Y_n$, and conditionally on this event $([J_n]_{n-1}, [Y_n]_{n-1})$ has the same distribution as (J_{n-1}, Y_{n-1}),
- with probability w_n/W_n : we have $J_n = n$. In this case the point Y_n is located on the last block $\mathbf{b}_n$ grafted on $\mathcal{T}_{n-1}$. Conditionally on this event (if $w_n > 0$), Y_n is distributed on this block under the measure $\boldsymbol{\nu}_n$ and the couple $([J_n]_{n-1}, [Y_n]_{n-1})$ is independent of the location of Y_n on the n-th block and has the same distribution as (J_{n-1}, Y_{n-1}).

From this observation we deduce that

$$(\mathcal{T}_{n-1}, \mathsf{T}_{n-1}, [J_n]_{n-1}, [Y_n]_{n-1}) = (\mathcal{T}_{n-1}, \mathsf{T}_{n-1}, J_{n-1}, Y_{n-1})$$

in distribution and more generally, $(\mathcal{T}_k, \mathsf{T}_k, [J_n]_k, [Y_n]_k) = (\mathcal{T}_k, \mathsf{T}_k, J_k, Y_k)$ in distribution for all $1 \leq k \leq n$.

Reversing this observation, we can construct a sequence $(J_n, Y_n)_{n\geq 1}$ (coupled to the K_n and X_n involved in the construction of $\mathcal{T}^*$) such that conditionally on $\mathcal{T}_n$ and T_n, the couple (J_n, Y_n) has distribution $\bar{\mu}_n^*$ and that for all $1 \leq k \leq n$, we have $([J_n]_k, [Y_n]_k) = (J_k, Y_k)$. To do so, we consider:

- a sequence $(U_n)_{n\geq 1}$ of uniform random variables on $(0\,,1)$,
- a sequence $(Z_n)_{n\geq 1}$ of points respectively sampled on $(\mathbf{b}_n)_{n\geq 1}$ with respective distribution (a normalised version of) the measure $(\boldsymbol{\nu}_n)_{n\geq 1}$ whenever it is non-zero, (set $Z_n = \boldsymbol{\rho}_n$ a.s. whenever $\boldsymbol{\nu}_n$ is trivial),
- a sequence $(I_n, P_n)_{n\geq 1}$, sampled with respective distributions $(\bar{\mu}_n^*)_{n\geq 1}$,

independently for all these random variables. Then we construct (K_n, X_n) and (J_n, Y_n) as follows. We set $(J_1, Y_1) = (1, Z_1)$. Then recursively for $n \geq 1$, we assume that X_{n-1} (if $n \neq 1$) and Y_n have been constructed:

- if $U_{n+1} \leq \frac{w_{n+1}}{W_{n+1}}$, then we set $(K_n, X_n) := (J_n, Y_n)$, $J_{n+1} := n+1$ and $Y_{n+1} := Z_{n+1}$,
- if $U_{n+1} > \frac{w_{n+1}}{W_{n+1}}$, then we set $(K_n, X_n) := (I_n, P_n)$, $J_{n+1} := J_n$ and $Y_{n+1} := Y_n$.

We can check that with this construction, for all $1 \leq k \leq n$, we have $([J_n]_k, [Y_n]_k) = (J_k, Y_k)$, the $(K_n, X_n)_{n\geq 1}$ are independent with the appropriate distribution and for all $n \geq 1$ conditionally on $\mathcal{T}_n$ and T_n the couple (J_n, Y_n) has distribution $\bar{\mu}_n^*$. Notice that the distance from Y_n to the root ρ is non-decreasing. Denoting $\mathcal{T}_0 = \{\rho\}$ and $Y_0 = \rho$, for all $0 \leq m \leq n$ we have

$$\mathrm{d}(Y_n, Y_m) = \mathrm{d}(Y_n, \mathcal{T}_m) = \sum_{k=m+1}^{n} \mathrm{d}(Z_k, \rho_k) \mathbf{1}_{\left\{U_k \leq \frac{w_k}{W_k}\right\}}, \tag{2.8}$$

which is equal in distribution to

$$\sum_{k=m+1}^{n} \lambda_k \mathcal{H}_k \mathbf{1}_{\left\{U_k \leq \frac{w_k}{W_k}\right\}},$$

where the $(\mathcal{H}_k)_{k\geq 1}$ are i.i.d., independent of the $(U_k)_{k\geq 1}$ and have the law of $\mathcal{H}$, see (2.4). Under H_d(iii) the random variable $\mathcal{H}$ has a finite second moment, and an application of Kolmogorov's three series theorem tells us that the almost sure convergence of $\sum_{k\geq 1} \lambda_k \mathcal{H}_k \mathbf{1}_{\left\{U_k \leq \frac{w_k}{W_k}\right\}}$ is equivalent to (2.7). In this case, $(Y_n)_{n\geq 1}$ is a Cauchy sequence in the complete space $\mathcal{T}$ and hence it converges. □

Also notice that with this construction, the discrete counterpart of (2.8) is

$$\mathrm{d}_{\mathsf{T}}(J_n, J_m) = \sum_{k=m+1}^{n} \mathbf{1}_{\left\{U_k \leq \frac{w_k}{W_k}\right\}} \quad \text{and so} \quad \mathrm{ht}_{\mathsf{T}}(J_n) = \sum_{k=2}^{n} \mathbf{1}_{\left\{U_k \leq \frac{w_k}{W_k}\right\}}. \tag{2.9}$$

Remark that for any $\theta \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\theta\, \mathrm{ht}_{\mathsf{T}}(J_n)\right)\right] &= \mathbb{E}\left[\exp\left(\theta \sum_{k=2}^{n} \mathbf{1}_{\left\{U_k \leq \frac{w_k}{W_k}\right\}}\right)\right] \\
&= \prod_{k=2}^{n} \left(\frac{W_k - w_k}{W_k} \cdot 1 + \frac{w_k}{W_k} e^{\theta}\right) \\
&= \exp\left(\sum_{k=2}^{n} \log\left(1 + (e^{\theta} - 1)\frac{w_k}{W_k}\right)\right) \\
&\leq \exp\left((e^{\theta} - 1) \sum_{k=2}^{n} \frac{w_k}{W_k}\right), \qquad (2.10)
\end{aligned}$$

where in the last line we use the inequality $\log(1+x) \leq x$, valid for all $x > -1$.

2.3.2 Convergence of the measure $\bar{\mu}_n$

Recall the definition of the random variable $\mathcal{H}$ in (2.4).

Proposition 2.4. *Assume that $\mathbb{E}\left[\mathcal{H}^2\right] < \infty$ and $\mathbb{P}\left(\mathcal{H} > 0\right) > 0$ and that (2.7) holds. Then almost surely there exists a probability measure $\bar{\mu}$ on $\mathcal{T}$ such that*

$$\bar{\mu}_n \underset{n\to\infty}{\longrightarrow} \bar{\mu} \qquad \textit{weakly.}$$

Furthermore, conditionally on $(\mathcal{T}, \bar{\mu})$, the point Y is distributed according to $\bar{\mu}$ almost surely. If $W_\infty < \infty$, then $\bar{\mu} = \frac{1}{W_\infty}\mu_\infty$, and $\bar{\mu}$ is concentrated on $\mathcal{T}^*$. If $W_\infty = \infty$, then $\bar{\mu}$ is concentrated on $\mathcal{L}$.

The proof of the last proposition is very similar to the proof of [41, Theorem 4], and is left to the reader. We can easily check that the assumptions of Proposition 2.4 are satisfied under the hypotheses of Theorem 2.1. We now state an additional lemma that will be useful later in this chapter.

Lemma 2.5. *Suppose that the assumptions of Proposition 2.4 hold, that $\bar{\mu}$ is concentrated on the set $\mathcal{L}$ and that the sequence of weights satisfies $\frac{w_n}{W_n} \leq n^{-1+o(1)}$. Then almost surely*

$$\bar{\mu}(\mathcal{T}(\mathbf{b}_n)) \leq n^{-1+o^{\omega}(1)},$$

where the random function $o^{\omega}(1)$ is considered as $n \to \infty$.

Proof. Let us introduce some notation. If $i \geq n$, we set $M_i^{(n)} := \bar{\mu}_i(\mathcal{T}(\mathbf{b}_n))$ the relative mass of the tree descending from $\mathbf{b}_n$ in $\mathcal{T}_i$. As i varies, this sequence of random variables evolves like one of Pemantle's time-dependent Pólya urns (see [103]) and is therefore a martingale. The topological boundary of $\mathcal{T}(\mathbf{b}_n)$ for the topology of $\mathcal{T}$ is either the empty set or the singleton $\{\boldsymbol{\rho}_n\}$, thus it has zero $\bar{\mu}$-measure[1]. It follows from Portmanteau theorem that the quantity of interest $\bar{\mu}(\mathcal{T}(\mathbf{b}_n))$ corresponds to $M_\infty^{(n)}$, the almost sure limit of this positive martingale. We can write

$$M_{i+1}^{(n)} = \left(\frac{W_i}{W_{i+1}}\right) M_i^{(n)} + \frac{w_{i+1}}{W_{i+1}} \mathbf{1}_{\left\{U_{i+1} \leq M_i^{(n)}\right\}},$$

[1] Indeed, under the assumptions of the lemma, $\bar{\mu}$ is carried on the leaves.

with $(U_i)_{i\geq 1}$ a sequence of i.i.d. random variables, uniform on $(0\,,1)$. We are going to show by induction on $k \geq 1$ that there exists a function $o(1)$ as $n \to \infty$ such that for all $i \geq n$, we have

$$\mathbb{E}\left[(M_i^{(n)})^k\right] \leq n^{-k+o(1)}.$$

Note that we use the notation $o(1)$ for all such functions, but that in this proof, the corresponding functions can depend on k but not on i.

• For $k = 1$, the result follows from the fact that $(M_i^{(n)})_{i\geq n}$ is a martingale and that almost surely $M_n^{(n)} \leq \frac{w_n}{W_n} \leq n^{-1+o(1)}$.

• Let $k \geq 2$. Suppose that the result is true for all $1 \leq l \leq k-1$. Then

$$\begin{aligned}
&\mathbb{E}\left[\left(M_{i+1}^{(n)}\right)^k \,\middle|\, M_i^{(n)}\right] \\
&= \mathbb{E}\left[\left(\left(\frac{W_i}{W_{i+1}}\right) M_i^{(n)} + \frac{w_{i+1}}{W_{i+1}} \mathbf{1}_{\left\{U_{i+1} \leq M_i^{(n)}\right\}}\right)^k \,\middle|\, M_i^{(n)}\right] \\
&= \left(\frac{W_i}{W_{i+1}}\right)^k \left(M_i^{(n)}\right)^k + \mathbb{E}\left[\sum_{l=0}^{k-1} \binom{k}{l} \left(\frac{W_i}{W_{i+1}}\right)^l \left(M_i^{(n)}\right)^l \left(\frac{w_{i+1}}{W_{i+1}} \mathbf{1}_{\left\{U_{i+1} \leq M_i^{(n)}\right\}}\right)^{k-l} \,\middle|\, M_i^{(n)}\right] \\
&= \left(\frac{W_i}{W_{i+1}}\right)^k \left(M_i^{(n)}\right)^k + \sum_{l=0}^{k-1} \binom{k}{l} \left(\frac{W_i}{W_{i+1}}\right)^l \left(M_i^{(n)}\right)^{l+1} \left(\frac{w_{i+1}}{W_{i+1}}\right)^{k-l} \\
&\leq \left(M_i^{(n)}\right)^k \left(\left(\frac{W_i}{W_{i+1}}\right)^k + k \cdot \frac{w_{i+1}}{W_{i+1}} \left(\frac{W_i}{W_{i+1}}\right)^{k-1}\right) + \sum_{l=0}^{k-2} \binom{k}{l} \left(M_i^{(n)}\right)^{l+1} \left(\frac{w_{i+1}}{W_{i+1}}\right)^{k-l}.
\end{aligned}$$

Now taking the expectation and using the fact that $\forall x \in [0\,,1]$, $(1-x)^k + k(1-x)^{k-1}x \leq 1$, we get, using the induction hypothesis,

$$\begin{aligned}
\mathbb{E}\left[(M_{i+1}^{(n)})^k\right] &\leq \mathbb{E}\left[(M_i^{(n)})^k\right] + \sum_{l=0}^{k-2} \binom{k}{l} \mathbb{E}\left[\left(M_i^{(n)}\right)^{l+1}\right] \left(\frac{w_{i+1}}{W_{i+1}}\right)^{k-l} \\
&\leq \mathbb{E}\left[(M_i^{(n)})^k\right] + \sum_{l=0}^{k-2} \binom{k}{l} n^{-(l+1)+o(1)} (i^{-1+o(1)})^{k-l}.
\end{aligned}$$

Using the last display in cascade we get that for all $i \geq n$:

$$\begin{aligned}
\mathbb{E}\left[(M_i^{(n)})^k\right] &\leq \mathbb{E}\left[(M_n^{(n)})^k\right] + \sum_{j=n}^{\infty} \sum_{l=0}^{k-2} \binom{k}{l} n^{-l-1+o(1)} j^{-k+l+o(1)} \\
&\leq \mathbb{E}\left[(M_n^{(n)})^k\right] + \sum_{l=0}^{k-2} \binom{k}{l} (n^{-l-1+o(1)}) \sum_{j=n}^{\infty} j^{-k+l+o(1)} \\
&\leq n^{-k+o(1)} + \sum_{l=0}^{k-2} \binom{k}{l} n^{-l-1+o(1)} n^{-k+l+1+o(1)} \\
&\leq n^{-k+o(1)}.
\end{aligned}$$

This finishes the proof by induction. This property passes to the limit in i by dominated convergence so, for all $n \geq 1$, we have $\mathbb{E}\left[(M_\infty^{(n)})^k\right] \leq n^{-k+o(1)}$. For N an integer and $\epsilon > 0$,

$$\begin{aligned}
\mathbb{P}\left(M_\infty^{(n)} \geq n^{-1+\epsilon}\right) &\leq n^{N-N\epsilon} \mathbb{E}\left[(M_\infty^{(n)})^N\right] \\
&\leq n^{-N\epsilon+o(1)}.
\end{aligned}$$

If we take N large enough, those quantities are summable and so, using the Borel-Cantelli lemma we get that with probability one, $M_\infty^{(n)} \leq n^{-1+\epsilon}$ for all n large enough. This completes the proof. $\square$

2.4 Upper-bounds and compactness for the (α, β)-model

In this section, we compute upper-bounds on the Hausdorff dimension of the set $\mathcal{L}$. We first prove Proposition 2.6, which tells us that, under the condition that $\lambda_n \leq n^{-\alpha+o(1)}$ for some $\alpha > 0$ and in a very general setting for the behaviour of the weights (w_n), the dimension is bounded above by $1/\alpha$. The techniques used in the proof are very robust, and do not depend on the geometry of the blocks nor on the sequence of weights. In a second step, in Proposition 2.7, we handle the more specific case where the underlying block satisfies Hypotheses H_d and that $\lambda_n \leq n^{-\alpha+o(1)}$ for some $0 < \alpha < 1/d$ and $w_n \leq n^{-\beta+o(1)}$ for some $\beta > 1$. In the proof of this proposition, a careful analysis allows us to refine some of the arguments of the previous proof and prove upper-bounds on the Hausdorff dimension of $\mathcal{L}$ that are below the "generic" value $1/\alpha$, given by Proposition 2.6. The techniques used for the proof are new and really take into account the behaviour of the weights and the geometry of the blocks.

2.4.1 Upper-bound independent of the weights and compactness

Notice that under H_d(iii), the underlying block $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ satisfies, for any $N > 0$,

$$\mathbb{P}\left(\operatorname{diam}(\mathsf{B}) \geq n^{\epsilon}\right) \leq \frac{\mathbb{E}\left[\operatorname{diam}(\mathsf{B})^N\right]}{n^{-N\epsilon}},$$

which is summable if N is large enough. Hence if (B_n) is an i.i.d. sequence with the same law as B, then using the Borel-Cantelli lemma we have almost surely,

$$\operatorname{diam}(\mathsf{B}_n) \leq n^{o^\omega(1)}. \tag{2.11}$$

Proposition 2.6. *Suppose that $\lambda_n \leq n^{-\alpha+o(1)}$, with $\alpha > 0$, and that $W_n \leq n^{\gamma}$ for some $\gamma > 0$ for all n. Suppose also that (2.11) holds. Then the tree-like structure $\mathcal{T}$ is almost surely compact and*

(i) $\mathrm{d_H}(\mathcal{T}_n, \mathcal{T}) \leq n^{-\alpha+o^\omega(1)}$,

(ii) $\dim_{\mathrm{H}}(\mathcal{L}) \leq \frac{1}{\alpha}$.

Since our model is invariant by multiplying all the weights by the same constant, we can always assume that $w_1 \leq 1$. Hence, the assumption on the weights in Proposition 2.6 is always satisfied if W_n grows at most polynomially in n, which is the case if Hyp. $\diamond_{\alpha,\beta}$, Hyp. $\square_{\alpha,1}$ or Hyp. $\bigcirc_{\alpha,\beta}$ is fulfilled, for any choice of $\alpha > 0$ and $\beta \in \mathbb{R}$.

Proof of Proposition 2.6. We start with point (i). First,

$$\mathrm{d_H}(\mathcal{T}_{2^i}, \mathcal{T}_{2^{i+1}}) \leq \sup_{2^i+1 \leq k \leq 2^{i+1}} \lambda_k \operatorname{diam}(\mathsf{B}_k) + \sup_{2^i+1 \leq k \leq 2^{i+1}} \mathrm{d}(\rho_k, \mathcal{T}_{2^i}).$$

For any $2^i \leq k \leq 2^{i+1} - 1$, the point ρ_{k+1} in the tree is identified with the point X_k, taken under the measure $\bar{\mu}_k$ on the tree $\mathcal{T}_k$. From our construction in Section 2.2.2, the point X_k belongs to some $\mathbf{b}_{K_k}$, and the couple (K_k, X_k) is sampled with measure $\bar{\mu}_k^*$. Bounding the contribution of every block along the ancestral line with their maximum, we get

$$\mathrm{d}(\rho_{k+1}, \mathcal{T}_{2^i}) = \mathrm{d}(X_k, \mathcal{T}_{2^i}) \leq \left(\sup_{2^i+1 \leq k \leq 2^i} \lambda_k \operatorname{diam}(\mathsf{B}_k) \right) \mathrm{ht}_{\mathsf{T}}(K_k). \tag{2.12}$$

Now using Lemma 2.24 in Appendix 2.A.4, we know that there exists a constant $C > 0$ such that $\sum_{i=1}^n \frac{w_i}{W_i} \leq C \log n$. Combining this with (2.10) (which holds for K_k because it has the same distribution as J_k) and the Markov inequality, we get for any $u > 0$,

$$\mathbb{P}\left(\mathrm{ht}_{\mathsf{T}}(K_n) \geq u \log n\right) \leq \exp\left((C(e-1) - u) \log n\right) = n^{C(e-1)-u}.$$

The last display is summable in n if we choose u large enough. Hence using the Borel-Cantelli lemma, we almost surely have $\mathrm{ht}_{\mathsf{T}}(K_n) \leq u \log n$ for n large enough. Hence, in (2.12) we have $\mathrm{ht}_{\mathsf{T}}(K_k) \leq (2^i)^{o^\omega(1)}$, where the symbol $o^\omega(1)$ denotes a random function of i that tends to 0 when $i \to \infty$. Combining this with (2.11) and the upper-bound on λ_n we get,

$$\mathrm{d}_{\mathrm{H}}(\mathcal{T}_{2^i}, \mathcal{T}_{2^{i+1}}) \leq (2^i)^{-\alpha + o^\omega(1)}.$$

Replacing i by k and summing the last display over $k \geq i$,

$$\sum_{k=i}^{\infty} \mathrm{d}_{\mathrm{H}}(\mathcal{T}_{2^k}, \mathcal{T}_{2^{k+1}}) \leq (2^i)^{-\alpha + o^\omega(1)},$$

hence the sequence of compact sets $(\mathcal{T}_{2^i})$ is a.s. a Cauchy sequence for Hausdorff distance between compacts of the complete space $\mathcal{T}$. So the sequence $(\mathcal{T}_n)_{n \geq 1}$ is also Cauchy because of the increasing property of the construction, and $\mathcal{T}$ is then almost surely compact. Moreover we have, a.s.

$$\mathrm{d}_{\mathrm{H}}(\mathcal{T}_{2^i}, \mathcal{T}) \leq (2^i)^{-\alpha + o^\omega(1)},$$

and this entails (i). Remark that since $\mathrm{ht}(\mathcal{T}(\mathbf{b}_n)) \leq \mathrm{d}_{\mathrm{H}}(\mathcal{T}_{n-1}, \mathcal{T})$, this implies that a.s. we have

$$\mathrm{ht}\left(\mathcal{T}(\mathbf{b}_n)\right) \leq n^{-\alpha + o^\omega(1)}. \tag{2.13}$$

We now prove point (ii). Let $\epsilon > 0$. From (2.13), the collection of balls $\mathrm{B}(\boldsymbol{\rho}_n, n^{-\alpha+\epsilon})$, for $n \geq N$, where N is an arbitrary number, is a covering of $\mathcal{L}$ whose maximal diameter tends to 0 as $N \to \infty$. Besides, if we fix δ, for N large enough, and $s > \frac{1}{\alpha - \epsilon}$, we have:

$$\mathcal{H}_s^\delta(\mathcal{L}) \leq \sum_{n=N}^{\infty} \operatorname{diam}(\mathrm{B}(\boldsymbol{\rho}_n, n^{-\alpha+\epsilon}))^s \leq \sum_{n=N}^{\infty} 2^s n^{(-\alpha+\epsilon)s} \underset{N \to \infty}{\longrightarrow} 0.$$

Hence for all such s, we have $\mathcal{H}_s(\mathcal{L}) = 0$ and so $\dim_{\mathrm{H}}(\mathcal{L}) \leq \frac{1}{\alpha - \epsilon}$. Letting $\epsilon \to 0$ finishes the proof. □

2.4.2 Upper-bound for $\alpha < 1/d$ and $\beta > 1$

Now let us study the specific case where the blocks satisfy Hypothesis H_d and that $\lambda_n \leq n^{-\alpha+o(1)}$ for some $0 < \alpha < 1/d$ and $w_n \leq n^{-\beta+o(1)}$ for some $\beta > 1$. The preceding Proposition 2.6 still holds but it is not optimal in this specific case. As in the previous proof we construct explicit coverings of the set $\mathcal{L}$ in order to bound its Hausdorff dimension. We construct them using an iterative procedure, which strongly depends on the dimension d and the exponent β. Starting from the covering given in the proof of Proposition 2.6, the procedure provides at each step a covering that is "better" in some sense than the preceding. In the limit, we prove the bound given in Proposition 2.7, which explicitly depends on β and d.

Proposition 2.7. *Suppose $0 < \alpha < \frac{1}{d}$ and $\beta > 1$ and that for all $n \geq 1$, $\lambda_n \leq n^{-\alpha+o(1)}$ and $w_n \leq n^{-\beta+o(1)}$. Suppose also that H_d(iii) and H_d(ii) hold for some $d \geq 0$. Then the Hausdorff dimension of $\mathcal{L}$ almost surely satisfies:*

$$\dim_{\mathrm{H}}(\mathcal{L}) \leq \frac{2\beta - 1 - 2\sqrt{(\beta-1)(\beta-\alpha d)}}{\alpha}.$$

For our purposes, we will work with countable sets of balls of $\mathcal{T}$, i.e. sets of the form

$$R = \{\mathrm{B}(x_i, r_i) \mid \forall i \geq 1,\ x_i \in \mathcal{T},\ r_i > 0\},$$

where $\mathrm{B}(x, r)$ denotes the open ball centred at x with radius r. Let us introduce some notation. If R is such a set of balls of $\mathcal{T}$, we say that R is a *covering* of the subset $X \subset \mathcal{T}$ if $X \subset \bigcup_{B \in R} B$. We can also define the s-volume of R as

$$\mathrm{Vol}_s(R) := \sum_{B \in R} \mathrm{diam}(B)^s.$$

In this way if the diameters of the balls that belong to R are bounded above by some $\delta > 0$, and R is a covering of X, then $\mathcal{H}_s^\delta(X) \leq \mathrm{Vol}_s(R)$, see Section 2.A.2 in the Appendix for the definition of $\mathcal{H}_s^\delta(X)$. Also, if R and R' are collections of balls and R covers X and R' covers X', then obviously $R \cup R'$ is a countable set of balls that covers $X \cup X'$ and for any s, we have

$$\mathrm{Vol}_s(R \cup R') \leq \mathrm{Vol}_s(R) + \mathrm{Vol}_s(R'). \tag{2.14}$$

In what follows, we construct random sets of balls and we prove that they are coverings of our set $\mathcal{L}$, which allows us to prove upper-bounds on the Hausdorff dimension of $\mathcal{L}$.

An idea of the proof

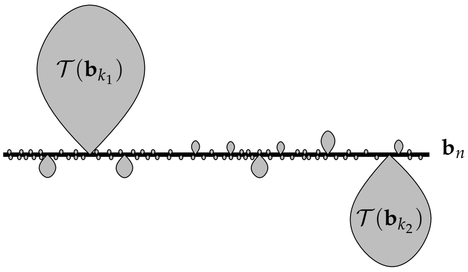
We briefly explain the idea of the proof before going into technicalities. The goal will be to provide a covering of each $\mathcal{T}(\mathbf{b}_n)$, for all n large enough. Since from the definition of $\mathcal{L}$, for any $N \geq 1$,

$$\mathcal{L} \subset \bigcup_{n \geq N} \mathcal{T}(\mathbf{b}_n), \tag{2.15}$$

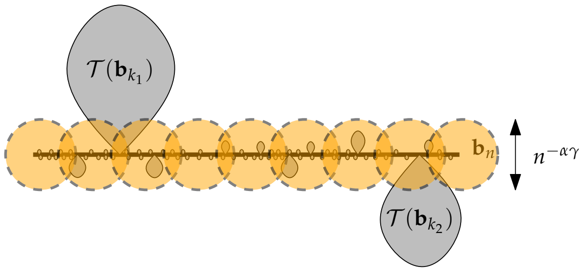
then the union over all n large enough of coverings of the $\mathcal{T}(\mathbf{b}_n)$ is indeed a covering of $\mathcal{L}$.

We recall how we derived the upper-bound $\frac{1}{\alpha}$ for the Hausdorff dimension of $\mathcal{L}$ in the proof of Proposition 2.6 (ii). The idea is to consider for every $n \geq 1$, a ball of radius $n^{-\alpha+\epsilon}$, say centred at $\boldsymbol{\rho}_n$. For n large enough, this ball covers $\mathcal{T}(\mathbf{b}_n)$ by (2.13). Thanks to (2.15), the set of balls $\{\mathrm{B}(\boldsymbol{\rho}_n, n^{-\alpha+\epsilon}) \mid n \geq N\}$, for any $N \geq 1$, is a covering of $\mathcal{L}$.

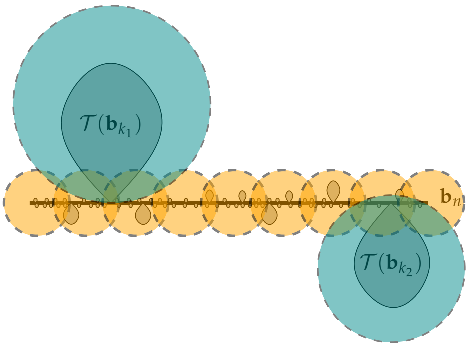
For $\beta \leq 1$, this covering is good because, as a block of index n has relative weight w_n/W_n which can be of order up to $n^{-1+o(1)}$ when it appears, the indices of the first blocks that are glued on $\mathbf{b}_n$ can have also an index of the order of n, and so a height of order up to $n^{-\alpha}$. On the contrary, if $\beta > 1$, we will see that the first block to be grafted on $\mathbf{b}_n$ has index roughly of order n^β, and so a height at most of order $n^{-\alpha\beta}$, which is very small compared to $n^{-\alpha}$. This gives us a hint that we can provide a "better" covering using a big number of smaller balls to cover $\mathbf{b}_n$ instead of just a "big" one, see Figure 2.4. We will use this rough idea to provide an algorithm that will construct finer and finer (random) coverings. Let us fix $\beta > 1$ from now on and take $s > d$, and explain informally how the algorithm works.



(a) The substructure $\mathcal{T}(\mathbf{b}_n)$



(b) A covering of $\mathbf{b}_n$ with small balls



(c) The remaining substructures are covered using the preceding step

Figure 2.4 – Explanation of Step 2 of the algorithm

Goal: At each step i of the algorithm, we want to construct for all $n \geq 1$ a set of balls $R^s_{n,i}$ such that, for n large enough, this set of balls is a covering of $\mathcal{T}(\mathbf{b}_n)$. Such a set of balls $R^s_{n,i}$ will have an s-volume of roughly $n^{f_i(s)}$, say. From step to step, we try to lower the s-volume of the set of balls constructed by the algorithm, which corresponds to lowering this exponent $f_i(s)$. Whenever we manage to get an exponent below -1, we stop the algorithm. We will see that it implies that the Hausdorff dimension of $\mathcal{L}$ is lower than or equal to s.

Step 1: The first step of the algorithm is deterministic and corresponds to what we did in the proof of Proposition 2.6. For each n we take a ball centred at $\boldsymbol{\rho}_n$ of radius roughly $n^{-\alpha}$ (in fact $n^{-\alpha+\epsilon}$ but let us not consider these technicalities for the moment). As seen before, for n large enough, it is a covering of $\mathcal{T}(\mathbf{b}_n)$. The s-volume of this covering is then of order $n^{-\alpha s}$. Denote $f_1(s) = -\alpha s$. If $f_1(s) < -1$, stop. Otherwise, proceed to step 2.

Step 2: As represented in Figure 2.4a, decompose $\mathcal{T}(\mathbf{b}_n)$ as

$$\mathcal{T}(\mathbf{b}_n) = \mathbf{b}_n \cup \bigcup_{\mathbf{b}_k \to \mathbf{b}_n} \mathcal{T}(\mathbf{b}_k).$$

Since the first block grafted on the block $\mathbf{b}_n$ typically has an index that is very large compared to n, we design a covering using smaller balls. We fix $\gamma > 1$ and decide to cover $\mathbf{b}_n$ with balls of size $n^{-\alpha\gamma}$, so that the blocks (and their descending substructure) of index $> n^\gamma$ are included in these balls, see Figure 2.4b. Since the blocks have dimension d, this covering uses roughly $\left(\frac{n^{-\alpha}}{n^{-\alpha\gamma}}\right)^d$ balls, each with s-volume $n^{-\alpha\gamma s}$. So the total volume used is around $n^{-\alpha d+\alpha\gamma d-\alpha\gamma s}$.

But doing so, we forgot to cover the blocks $\mathbf{b}_k$ such that $\mathbf{b}_k \to \mathbf{b}_n$ and $k \leq n^\gamma$. To take care of them, we use the preceding step of the algorithm and cover each of them with a ball of radius $k^{-\alpha}$, see Figure 2.4c. Recalling that $s \leq 1/\alpha$, we get that in expectation, these balls have a s-volume of order

$$\sum_{k=n+1}^{n^\gamma} \mathbb{P}(\mathbf{b}_k \to \mathbf{b}_n) k^{-\alpha s} \approx n^{-\beta} \sum_{k=n+1}^{n^\gamma} k^{-\alpha s} \approx n^{-\beta+\gamma(1-\alpha s)}.$$

Hence, the total s-volume used to cover $\mathcal{T}(\mathbf{b}_n)$ has order $n^{\max(-\beta+\gamma(1-\alpha s), -\alpha d+\alpha\gamma d-\alpha\gamma s)}$. Since we want to construct a covering having the smallest possible volume, we can optimize on γ the last exponent. Under our assumptions, one can check that it is minimal if we take $\gamma := \frac{\beta-\alpha d}{1-\alpha d} > 1$. We then get

$$\max(-\beta+\gamma(1-\alpha s), -\alpha d+\alpha\gamma d-\alpha\gamma s) = \frac{-\alpha d+\alpha\beta d-\alpha\beta s+\alpha^2 ds}{1-\alpha d} := f_2(s).$$

We can check that the new exponent $f_2(s)$ is smaller than $f_1(s) = -\alpha s$. Hence we can cover $\mathcal{T}(\mathbf{b}_n)$ with balls using a total s-volume of a lower order than the preceding step. If $f_2(s) < -1$, stop. Otherwise, proceed to step 3.

Step i: Now we recursively repeat the preceding step. Thanks to step $i-1$, we know that we can provide a covering of $\mathcal{T}(\mathbf{b}_n)$ for any n, using an s-volume of approximately $n^{f_{i-1}(s)}$. Now we fix a number $\gamma > 1$ and we cover the block $\mathbf{b}_n$ with balls of radius $n^{-\alpha\gamma}$. As in step 2, this covering has a s-volume of order $n^{-\alpha d+\alpha\gamma d-\alpha\gamma s}$. Then we take care of the $\mathbf{b}_k$ such that $\mathbf{b}_k \to \mathbf{b}_n$ and $k < n^\gamma$. To cover them we use step $i-1$, which ensures that we can do that for each k with an s-volume of roughly $k^{f_{i-1}(s)}$. Hence the expectation on the s-volume for all these balls is, if s is such that $f_{i-1}(s) \geq -1$,

$$\sum_{k=n+1}^{n^\gamma} \mathbb{P}(\mathbf{b}_k \to \mathbf{b}_n) k^{f_{i-1}(s)} \approx n^{-\beta} \sum_{k=n+1}^{n^\gamma} k^{f_{i-1}(s)} \approx n^{-\beta+\gamma(1+f_{i-1}(s))}.$$

We then choose the optimal $\gamma > 1$ that minimizes the maximum of the exponents

$$\max(-\alpha d+\alpha\gamma d-\alpha\gamma s, -\beta+\gamma(1+f_{i-1}(s))).$$

We denote the value for which the minimum is obtained by $\gamma_i(s)$, which depends on the parameter s. The first exponent is linearly decreasing with γ, the other one is linearly increasing, and their value for γ tending to 1, satisfy $-\alpha s > -\beta+1+f_{i-1}(s)$. Hence, the value of $\gamma_i(s)$ is the value for which the two of them are equal, and this value is strictly greater than 1. We call this minimal exponent $f_i(s)$. If $f_i(s) < -1$, stop. Otherwise, proceed to step $i+1$.

Upper-bound on Hausdorff dimension Now, suppose s is such that $f_i(s)$ is well-defined and $f_i(s) < -1$, for some $i \geq 1$. If we cover every $\mathcal{T}(\mathbf{b}_n)$ using the covering provided by step i of the algorithm, then the union of all those coverings covers $\mathcal{L}$. Furthermore, we only need to cover all the $\mathcal{T}(\mathbf{b}_n)$ for n sufficiently large to cover $\mathcal{L}$, so we can have a covering of $\mathcal{L}$ using arbitrarily small balls. Hence we get that for all $\delta > 0$, we have $\mathcal{H}_s^\delta(\mathcal{L}) < \sum_{n=1}^\infty n^{f_i(s)} < \infty$ and so $\mathcal{H}_s(\mathcal{L}) < \infty$, which proves that

$$\dim_{\mathrm{H}}(\mathcal{L}) \leq s.$$

This rough analysis is turned into a rigorous proof in what follows. We begin with elementary definitions and calculations that arise from what precedes.

Study of a sequence of functions

We begin by recursively defining the sequence of functions $(f_i)_{i\geq 1}$, together with a sequence $(s_i)_{i\geq 1}$ of real numbers.

Definition-Proposition 2.8. *We set* $s_0 := \infty$. *We define a sequence* $(f_i)_{i\geq 1}$ *of functions as follows. We set*

$$\forall s \in [d\,,\infty), \quad f_1(s) := -\alpha s,$$

and set $s_1 := \frac{1}{\alpha}$. *Then for all* $i \geq 1$, *we recursively define*

$$\forall s \in [d\,, s_i], \quad f_{i+1}(s) := \frac{\alpha(-d+\beta d-\beta s-f_i(s)d)}{1+f_i(s)+\alpha s-\alpha d}.$$

Define s_{i+1} *as the unique solution to the equation* $f_{i+1}(s) = -1$.

Before proving the validity of this definition, let us state some properties of this sequence of functions:

Proposition 2.9. *The following properties are satisfied:*

(i) *For all* $i \geq 1$, *the function* f_i *is continuous, strictly decreasing, and* $f_i(d) = -\alpha d$.

(ii) *For all* $i \geq 1$, *for all* $s \in (d\,, s_i]$, *we have* $f_{i+1}(s) < f_i(s)$.

(iii) *Let* $s_\infty := \frac{2\beta-1-2\sqrt{(\beta-1)(\beta-\alpha d)}}{\alpha}$. *Then we have for all* $s \in [d\,, s_\infty)$,

$$f_i(s) \underset{i\to\infty}{\longrightarrow} f_\infty(s),$$

where

$$f_\infty(s) = \frac{-(1+\alpha s)+\sqrt{1+2\alpha s+\alpha^2 s^2-4\alpha d+4\alpha\beta d-4\alpha\beta s}}{2}.$$

(iv) *For all* $s \in [d\,, s_\infty)$, *we have* $f_\infty(s) > -1$.

(v) *The sequence* $(s_i)_{i\geq 1}$ *is strictly decreasing and*

$$s_i \underset{i\to\infty}{\longrightarrow} s_\infty.$$

(vi) *For all* $i \geq 1$, *we have* $f_{i+1}(s_i) < -1$.

Proof. We define the function F on the set $\left\{(s,x) \in \mathbb{R}^2 \;\middle|\; d \leq s \leq \frac{1}{\alpha},\ x > \alpha d - \alpha s - 1\right\}$ by

$$F(s,x) = \frac{\alpha(-d+\beta d-\beta s-dx)}{1+x+\alpha s-\alpha d}.$$

We have for all $s > d$ and all $x > \alpha d - \alpha s - 1$,

$$\partial_x F(s,x) = \frac{\alpha(\beta-\alpha d)(s-d)}{(1+x+\alpha s-\alpha d)^2} > 0.$$

This shows that for all $s > d$, the function $F(s,\cdot)$ is strictly increasing, and also strictly concave since the derivative is strictly decreasing.
From these facts we can show by induction on i the points (i) and (ii) of Proposition 2.9, together with the validity of the definition of f_i and s_i, in Definition-Proposition 2.8.

• For $i = 1$, the function f_1 is well-defined, s_1 is indeed the unique solution to $f_1(s) = -1$ and the point (i) is satisfied. Moreover, $f_2(s)$ is well-defined for $s \in [d\,, s_1]$ by $f_2(s) = F(s, -\alpha s)$ and for all $s \in (d\,, s_1)$, we have

$$F(s, -\alpha s) + \alpha s = \frac{\alpha(\beta - 1)(d - s)}{1 - \alpha d} < 0,$$

which proves that (ii) holds for $i = 1$.

• By induction, if f_i and s_i are defined up to some $i \geq 1$ and satisfy (i), then one can verify that for all $s \in [d\,, s_i]$, the function f_{i+1} is well-defined by the formula:

$$f_{i+1}(s) = F(s, f_i(s)).$$

From the monotonicity of $F(s, \cdot)$ and f_i, this function is continuous and strictly decreasing. One can check that $F(d, x) = -\alpha d$ for any $x > -1$ so f_{i+1} satisfies (i). Then, if $i = 1$, the initialisation already gives us that (ii) holds. Otherwise, if $i \geq 2$, then using the induction hypothesis, for all $s \in (d\,, s_{i-1}]$ we have $f_i(s) < f_{i-1}(s)$. Using that $F(s, \cdot)$ is strictly increasing for $s > d$ we get that for all $s \in (d\,, s_i]$, $f_{i+1}(s) < f_i(s)$, and so (ii) holds. Since f_{i+1} is continuous and strictly decreasing and that $f_{i+1}(d) > -1$ and $f_{i+1}(s_i) < f_i(s_i) = -1$, then s_{i+1} is well-defined. This finishes our proof by induction.

Let us study at fixed $s > d$ the equation $F(s, x) = x$. We get the following second order equation:

$$x^2 + x(1 + \alpha s) + (\alpha d - \alpha\beta d + \alpha\beta s) = 0,$$

for which the discriminant is $\Delta_s = 1 + 2\alpha s + \alpha^2 s^2 - 4\alpha d + 4\alpha\beta d - 4\alpha\beta s$. We can evaluate this quantity at d and at $\frac{1}{\alpha}$. We get

$$\Delta_d = (\alpha d - 1)^2 > 0 \qquad \text{and} \qquad \Delta_{1/\alpha} = 4(\beta - 1)(\alpha d - 1) < 0.$$

We can check that it vanishes exactly at $s = s_\infty$ so that in the end, Δ_s is strictly positive on $[d\,, s_\infty)$, null at s_∞ and strictly negative on $(s_\infty\,, \frac{1}{\alpha}]$. Hence, the function $F(s, \cdot)$ has 2 (resp. 1, resp. 0) fixed points on the corresponding intervals.

The convergence (iii) is a consequence of the fact that for $s \in [d\,, s_\infty)$, the function $F(s, \cdot)$ is strictly increasing and concave, has exactly two fixed points and that the initial value $f_1(s)$ is greater than the smallest fixed point. This proves the convergence of the sequence $f_i(s)$ towards the greatest fixed point of $F(s, \cdot)$, the value of which can be computed using the equation above. The property of the limit (iv) can be checked by proving that for all $s \in [d\,, s_\infty]$, we have $f_\infty(s) \geq f_\infty(s_\infty) = \sqrt{(\beta - 1)(\beta - \alpha d)} - \beta > -1$.

Let us prove the point (v). According to property (ii), we have $f_i(s_{i+1}) > f_{i+1}(s_{i+1}) = -1$, and since f_i is decreasing, we get $s_{i+1} < s_i$. Hence the sequence $(s_i)_{i\geq 1}$ is strictly decreasing, bounded below by d, so it converges. Now let $s > s_\infty$. If the sequence $(f_i(s))_{i\geq 1}$ was well-defined for all $i \geq 1$, then for all $i \geq 1$ we would have $f_i(s) > -1$, so it would be decreasing, bounded below, hence it would have a limit, which would be a fixed point of $F(s, \cdot)$. It is impossible since $F(s, \cdot)$ has no fixed point, so the sequence is not well-defined for all $i \geq 1$ and so for i large enough, $s > s_i$. We conclude that $\lim_{i\to\infty} s_i \leq s_\infty$. If we had $\lim_{i\to\infty} s_i < s_\infty$, then it would contradict the property (iv). We conclude that indeed $\lim_{i\to\infty} s_i = s_\infty$.

The last property (vi) follows from property (ii). Indeed, we have $f_{i+1}(s_i) < f_i(s_i) = -1$. $\square$

Construction of the coverings

Let us provide a rigorous proof of our upper-bound, which follows the heuristics that we derived in the beginning of the section. Here, we distinguish two types of negligible functions, $o_n(1)$ and

$o_\epsilon(1)$. A function denoted $o_n(1)$ (resp. $o_\epsilon(1)$) is negligible as $n \to \infty$ (resp. as $\epsilon \to 0$) and does not depend on ϵ (resp. on n).

Proposition 2.10. *Fix $i \geq 1$ and $s \leq s_{i-1}$. For all $\epsilon > 0$, we can construct a set of balls $R^{s,\epsilon}_{n,i}$ simultaneously for all $n \geq 1$, such that the following holds.*

(i) *Almost surely, for n large enough, $R^{s,\epsilon}_{n,i}$ covers $\mathcal{T}(\mathbf{b}_n)$.*

(ii) *We have*

$$\mathbb{E}\left[\mathrm{Vol}_s(R^{s,\epsilon}_{n,i})\right] \leq n^{f_i(s)+o_n(1)+o_\epsilon(1)}.$$

(iii) *The diameter of the balls used are such that* $\max_{B \in R^{s,\epsilon}_{n,i}} \mathrm{diam}(B) \underset{n\to\infty}{\longrightarrow} 0.$

We will define the set of balls $R^{s,\epsilon}_{n,i}$ over the block $\mathbf{b}_n$ and its descendants in an algorithmic way, and each step of the algorithm only depends on the gluings that happen after time n. The proof of the upper-bound will directly follow from Proposition 2.10. Let us first state an elementary result, the proof of which is left to the reader. Note that we allow the function $o_\epsilon(1)$ to be infinite for large values of ϵ.

Lemma 2.11. *Let $\xi \geq -1$. Then for all $\gamma > 1$,*

$$\mathbb{E}\left[\sum_{k=n+1}^{n^\gamma} \mathbf{1}_{\{\mathbf{b}_k \to \mathbf{b}_n\}} k^{\xi+o_n(1)+o_\epsilon(1)}\right] \leq n^{-\beta+\gamma(\xi+1)+o_n(1)+o_\epsilon(1)}.$$

Proof of Proposition 2.10. Let $s > 0$ and $\epsilon > 0$. We prove the proposition by induction on i. The first set of balls that we build is the following: for each block $\mathbf{b}_n$, we cover the block with a ball of radius $n^{-\alpha+\epsilon}$, centred on the point $\boldsymbol{\rho}_n$. We write:

$$R^{s,\epsilon}_{n,1} = \{\mathrm{B}(\boldsymbol{\rho}_n, n^{-\alpha+\epsilon})\}.$$

According to (2.13), there exists a random N such that for all $n \geq N$, the set $R^{s,\epsilon}_{n,1}$ covers $\mathcal{T}(\mathbf{b}_n)$. The diameter of the ball of $R^{s,\epsilon}_{n,1}$ tends to 0 as $n \to \infty$. Besides we have,

$$\mathbb{E}\left[\mathrm{Vol}_s(R^{s,\epsilon}_{n,1})\right] \leq (2n)^{-\alpha s+\epsilon s} = n^{f_1(s)+o_n(1)+o_\epsilon(1)}.$$

The property is thus proved for $i = 1$.

Let $i \geq 1$ and $s < s_i$. Let us construct $\left(R^{s,\epsilon}_{n,i+1}\right)_{n\geq 1}$, using the previous step i. We set $\gamma_{i+1}(s) > 1$ a positive real number that we will choose later, and $\epsilon > 0$. We define $R^{s,\epsilon}_{n,i+1}$ as follows: it is the union over all the blocks $\mathbf{b}_k$ for $k < n^{\gamma_i(s)}$ that are grafted on the block $\mathbf{b}_n$, of their covering $R^{s,\epsilon}_{k,i}$ of the preceding step, together with the union of a deterministic set of balls that we define hereafter.

We want to cover $\mathbf{b}_n$ with balls of radius $n^{-\alpha\gamma_{i+1}(s)}$, which is equivalent to covering B_n with balls of radius $\lambda_n^{-1} n^{-\alpha\gamma_{i+1}(s)}$. Under Hypothesis H_d, for any $d \geq 0$, using Lemma 2.21 and Lemma 2.23 in Appendix 2.A.3, we can a.s. find a random collection $(x_m)_{1\leq m\leq M_r(\mathsf{B}_n)}$ of points of B_n such that the balls centred on those points with radius $r := \lambda_n^{-1} n^{-\alpha\gamma_{i+1}(s)}$ cover B_n, and such that $M_r(\mathsf{B}_n) \leq N_{r/4}(\mathsf{B}_n)$, where $N_r(\mathsf{B})$ is the minimal number of balls of radius r needed to cover B.

From the assumption on the sequence (λ_n), we have $r \geq n^{-\alpha\gamma_{i+1}(s)+\alpha+o_n(1)}$. Since $N_r(\mathsf{B}_n)$ is decreasing in r, using Hypothesis H_d(ii) we get that

$$\mathbb{E}\left[N_{r/4}(\mathsf{B}_n)\right] \leq n^{-\alpha d+\alpha\gamma_{i+1}(s)d+o_n(1)}.$$

In the end,

$$R^{s,\epsilon}_{n,i+1} := \left(\bigcup_{k \le n^{\gamma_{i+1}(s)} : \mathbf{b}_k \to \mathbf{b}_n} R^{s,\epsilon}_{k,i} \right) \cup \left\{ \mathrm{B}\left(x_m, n^{-\alpha\gamma_{i+1}(s)+\epsilon}\right) \;\middle|\; 1 \le m \le M_r(\mathsf{B}_n) \right\}.$$

Remark that for any $n \ge 1$, the set of balls $R^{s,\epsilon}_{n,i+1}$ is independent of the event $\{\mathbf{b}_n \to \mathbf{b}_k\}$ for any $k < n$. Now we compute the expectation of the s-volume of these sets of balls as follows, and use the preceding remark in the third line,

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{Vol}_s(R^{s,\epsilon}_{n,i+1})\right] \\
&= \mathbb{E}\left[\sum_{k=n+1}^{n^{\gamma_{i+1}(s)}} \mathbf{1}_{\{\mathbf{b}_k \to \mathbf{b}_n\}} \mathrm{Vol}_s(R^{s,\epsilon}_{k,i})\right] + \mathbb{E}\left[M_r(\mathsf{B}_n)\right]\left(2n^{(-\alpha\gamma_{i+1}(s)+\epsilon)}\right)^s \\
&\le \mathbb{E}\left[\sum_{k=n+1}^{n^{\gamma_{i+1}(s)}} \mathbf{1}_{\{\mathbf{b}_k \to \mathbf{b}_n\}} \mathbb{E}\left[\mathrm{Vol}_s(R^{s,\epsilon}_{k,i})\right]\right] + \mathbb{E}\left[N_{r/4}(\mathsf{B}_n)\right]\left(2n^{(-\alpha\gamma_{i+1}(s)+\epsilon)}\right)^s \\
&\le \mathbb{E}\left[\sum_{k=n+1}^{n^{\gamma_{i+1}(s)}} \mathbf{1}_{\{\mathbf{b}_k \to \mathbf{b}_n\}} k^{f_i(s)+o_n(1)+o_\epsilon(1)}\right] + n^{-\alpha d+\alpha d\gamma_{i+1}(s)+(-\alpha\gamma_{i+1}(s)+\epsilon)s+o_n(1)+o_\epsilon(1)} \\
&\le n^{-\beta+\gamma_{i+1}(s)(f_i(s)+1)+o_n(1)+o_\epsilon(1)} + n^{-\alpha d+\alpha d\gamma_{i+1}(s)-\alpha\gamma_{i+1}(s)s+o_n(1)+o_\epsilon(1)},
\end{aligned}$$

where in the last line we used Lemma 2.11 which applies because $s \le s_i$, hence $f_i(s) \ge -1$. We then take $\gamma_{i+1}(s) := \frac{\beta-\alpha d}{f_i(s)+1-\alpha d+\alpha s} > 1$, which yields

$$-\beta + \gamma_{i+1}(s)(f_i(s)+1) = -\alpha d + \alpha\gamma_{i+1}(s)d - \alpha\gamma_{i+1}(s)s = f_{i+1}(s).$$

We then have,

$$\mathbb{E}\left[\mathrm{Vol}_s(R^{s,\epsilon}_{n,i+1})\right] \le n^{f_{i+1}(s)+o_n(1)+o_\epsilon(1)}.$$

We can check that $\max_{B \in R^{s,\epsilon}_{n,i+1}} \mathrm{diam}(B) \underset{n\to\infty}{\longrightarrow} 0$, and that almost surely, for n large enough, the collections of balls $R^{s,\epsilon}_{n,i+1}$ are indeed coverings of $\mathcal{T}(\mathbf{b}_n)$ thanks again to (2.13). This finishes the proof. $\square$

We can now prove the main proposition of this section.

Proof of Proposition 2.7. Let $i \ge 1$. For $\epsilon > 0$ small enough, we use Proposition 2.10 to get a set of balls $(R^{s_i,\epsilon}_{n,i+1})_{n\ge1}$, which satisfies

$$\mathbb{E}\left[\mathrm{Vol}_{s_i}(R^{s_i,\epsilon}_{n,i+1})\right] \le n^{f_{i+1}(s_i)+o_n(1)+o_\epsilon(1)}.$$

From Proposition 2.9(vi), we have $f_{i+1}(s_i) < -1$, so we can choose ϵ small enough such that the exponent is eventually smaller than $\frac{f_{i+1}(s_i)-1}{2} < -1$ as $n \to \infty$. Then, for $N \ge 1$, we set $R_N = \bigcup_{n\ge N} R^{s_i,\epsilon}_{n,i+1}$. According to Proposition 2.10, the set of balls $R^{s_i,\epsilon}_{n,i+1}$ is a covering of $\mathcal{T}(\mathbf{b}_n)$ for n large enough and so R_N is a covering of $\mathcal{L}$, for all N. Since for any $\delta > 0$, we may choose N large enough so that $\max_{B\in R_N} \mathrm{diam}(B) < \delta$, we get

$$\mathcal{H}_{s_i}(\mathcal{L}) = \lim_{\delta\to0} \mathcal{H}^\delta_{s_i}(\mathcal{L}) \le \limsup_{N\to\infty} \mathrm{Vol}_{s_i}(R_N) \le \mathrm{Vol}_{s_i}(R_1) < +\infty, \quad \text{a.s.}$$

since

$$\mathbb{E}\left[\mathrm{Vol}_{s_i}(R_1)\right] \le \sum_{n=1}^{\infty} n^{\frac{f_{i+1}(s_i)-1}{2}+o_n(1)} < +\infty.$$

This shows that the Hausdorff dimension of $\mathcal{L}$ satisfies $\dim_{\mathrm{H}}(\mathcal{L}) \leq s_i$, almost surely. In the end, since the sequence $(s_i)_{i\geq 1}$ tends to s_∞, we conclude that almost surely,

$$\dim_{\mathrm{H}}(\mathcal{L}) \leq s_\infty = \frac{2\beta - 1 - 2\sqrt{(\beta-1)(\beta-\alpha d)}}{\alpha}. \qquad \square$$

2.5 Lower-bounds for the (α, β)-model

In this section we compute lower-bounds on the Hausdorff dimension of the set $\mathcal{L}$. We do that by constructing Borel measures on $\mathcal{L}$ that satisfy the assumptions of Frostman's lemma (Lemma 2.20 in Appendix 2.A.2). In the case where $\beta \leq 1$ we use the natural measure $\bar{\mu}$ on $\mathcal{T}$ which arises as the limit of the normalised weight measures on $\mathcal{T}_n$ (see Proposition 2.4). The case $\beta > 1$ is a bit more technical because the natural measure $\bar{\mu}$ is not concentrated on $\mathcal{L}$, so we have to construct another measure π, that we define as the subsequential limit of some well-chosen sequence of probability measures on $\mathcal{T}$.

2.5.1 Case $\beta \leq 1$ and use of the measure $\bar{\mu}$

In this subsection, we suppose that $\beta \leq 1$. Under the assumptions of Proposition 2.4, the sequences of measures $\bar{\mu}_n$ almost surely converges weakly to a measure $\bar{\mu}$, which is concentrated on the set of leaves $\mathcal{L}$. The existence of $\bar{\mu}$ will be useful for the proof of the next proposition. Recall the definition of the random variable $\mathcal{H}$ from (2.4) and the fact that the assumptions on $\mathcal{H}$ in the proposition are satisfied under Hypothesis H_d for any $d \geq 0$.

Proposition 2.12. *Suppose that Hypothesis $\bigcirc_{\alpha,\beta}$ or Hypothesis $\square_{\alpha,1}$ is satisfied. Suppose also that $\mathbb{E}\left[\mathcal{H}^2\right] < \infty$ and that $\mathbb{P}\left(\mathcal{H} > 0\right) > 0$. Then the Hausdorff dimension of $\mathcal{L}$ almost surely satisfies:*

$$\dim_{\mathrm{H}}(\mathcal{L}) \geq \frac{1}{\alpha}.$$

As we said earlier, the idea is to prove this lower bound on the dimension using Frostman's lemma: we will thus prove that almost surely, for $\bar{\mu}$-almost all leaves $x \in \mathcal{L}$, we have an upper bound of the type

$$\bar{\mu}(\mathrm{B}(x,r)) \leq r^{1/\alpha - \epsilon},$$

for r sufficiently small, and for all ϵ. An application of Lemma 2.20 will then finish our proof.

In order to prove this control on the masses of the balls, we will use two lemmas. The first one allows us to compare $\bar{\mu}(\mathrm{B}(x,r))$ with a quantity of the form $\bar{\mu}(\mathcal{T}(\mathbf{b}_n))$ for an appropriate n. The second one, Lemma 2.5, provides a good control of the quantities $\bar{\mu}(\mathcal{T}(\mathbf{b}_n))$ for large n, such that the combination of the two will provide the upper bound that we want. Let $\epsilon > 0$. Recall from (2.2) the definition of G^ϵ.

Lemma 2.13. *Set $n_0 = 2$ and $n_{k+1} = \lceil n_k^{1+\epsilon} \rceil$. Under the hypotheses of Proposition 2.12, almost surely for $\bar{\mu}$-almost every $x \in \mathcal{L}$, for all k large enough[2], there exists $n \in [\![n_k, n_{k+1}]\!] \cap G^\epsilon$ such that*

$$x \in \mathcal{T}(\mathbf{b}_n) \quad \text{and} \quad \mathrm{d}(x, \boldsymbol{\rho}_n) \geq n^{-\alpha - 2\epsilon}.$$

Proof. Note that in our setting, the hypothesis of Proposition 2.4 holds and so the random leaf Y constructed in Section 2.3 is defined a.s. Also, according to Proposition 2.4, conditionally

[2] The threshold depends on the realisation and on x.

on $(\mathcal{T}, \bar{\mu})$, the point Y has distribution $\bar{\mu}$. So it suffices to prove that the lemma holds for the random leaf Y. We recall

$$\mathrm{d}(Y_n, Y_m) \overset{\text{law}}{=} \sum_{k=m+1}^{n} \lambda_k \mathcal{H}_k \mathbf{1}_{\left\{U_k \leq \frac{w_k}{W_k}\right\}} \quad \text{and} \quad Y = \lim_{n\to\infty} Y_n.$$

Let us introduce a constant $c > 0$ and set

$$p := \mathbb{P}\left(\mathcal{H} > c\right).$$

For $\beta < 1$, thanks to our assumptions, we can fix c such that p is non-zero. We then have

$$\begin{aligned}
\mathbb{P}\left(\forall i \in [\![n_k\,, n_{k+1}]\!] \cap G^\epsilon,\ U_i > \frac{w_i}{W_i} \text{ or } \mathcal{H}_i < c\right) &= \prod_{\substack{i=n_k \\ i\in G^\epsilon}}^{n_k^{1+\epsilon}} \left(1 - p\frac{w_i}{W_i}\right) \\
&= \exp\left(\sum_{\substack{i=n_k \\ i\in G^\epsilon}}^{n_k^{1+\epsilon}} \log\left(1 - p\frac{w_i}{W_i}\right)\right) \\
&\leq \exp\left(-p\sum_{\substack{i=n_k \\ i\in G^\epsilon}}^{n_k^{1+\epsilon}} \frac{w_i}{W_i}\right) \\
&\underset{\text{Lem.2.25}}{\leq} \exp\left(-pC_\epsilon \log(n_k)\right).
\end{aligned}$$

To write the last line we use Lemma 2.25 in the Appendix and we can see that the last display is summable over k.

For the case $\beta = 1$, Hypothesis $\square_{\alpha,1}$ allows us to write

$$\mathbb{P}\left(\forall i \in [\![n_k\,, n_{k+1}]\!] \cap G^\epsilon,\ U_i > \frac{w_i}{W_i} \text{ or } \mathcal{H}_i < c\right) \leq \exp\left(-pf(k)\log\log\log(n_k)\right),$$

with a function $f(k)$ tending to infinity. Since $n_k \geq 2^{(1+\epsilon)^k}$, then $\log\log\log(n_k) \geq (1+o(1))\log k$ and the last display is also summable in k. In both cases, an application of the Borel-Cantelli lemma shows that we have almost surely, for k large enough,

$$\exists n \in [\![n_k\,, n_{k+1}]\!] \cap G^\epsilon, \qquad U_n \leq \frac{w_n}{W_n} \quad \text{and} \quad \mathcal{H}_n \geq c.$$

Since $n \in G^\epsilon$, we have $\lambda_n \geq n^{-\alpha-\epsilon}$. Combined with the fact that $\mathcal{H}_n \geq c$ we get

$$\mathrm{d}(\boldsymbol{\rho}_n, Y) \geq \lambda_n \mathcal{H}_n \geq cn^{-\alpha-\epsilon} \geq n^{-\alpha-2\epsilon}, \text{ for } n \text{ (or equivalenly } k) \text{ large enough.} \qquad \square$$

Proof of Proposition 2.12. Let $\epsilon > 0$. Let us fix a realisation of $\mathcal{T}$ and a leaf $x \in \mathcal{L}$ such that the conclusions of Lemma 2.13 and Lemma 2.5 hold. Note that thanks to Hypothesis $\bigcirc_{\alpha,\beta}$ or Hypothesis $\square_{\alpha,1}$, the condition of application of Lemma 2.5 are fulfilled. From the definition of n_{k+1} we have $n_k^{1+\epsilon} < n_{k+1} \leq n_k^{1+\epsilon} + 1$ and so $n_{k+1} \underset{k\to\infty}{=} n_k^{1+\epsilon+o(1)}$. We know from Lemma 2.13 that for all k large enough, there exists $n \in [\![n_k\,, n_{k+1}]\!]$ such that $x \in \mathcal{T}(\mathbf{b}_n)$ and

$$\mathrm{d}(\boldsymbol{\rho}_n, x) \geq n^{-\alpha-2\epsilon} \geq n_{k+1}^{-\alpha-2\epsilon} \geq n_k^{(1+\epsilon+o(1))(-\alpha-2\epsilon)}.$$

So if we take k large enough and $r \in [n_{k+2}^{-\alpha-2\epsilon}, n_{k+1}^{-\alpha-2\epsilon})$, then

$$\begin{aligned}\bar{\mu}(\mathrm{B}(x,r)) \leq \bar{\mu}(\mathcal{T}(\mathbf{b}_n)) \underset{\text{Lem. 2.5}}{\leq} n^{-1+\epsilon} \leq n_k^{-1+\epsilon} &= n_{k+2}^{\frac{-1+\epsilon}{(1+\epsilon)^2}+o(1)} \\ &\leq \left(r^{\frac{-1}{\alpha+2\epsilon}}\right)^{\frac{-1+\epsilon}{1+\epsilon}+o(1)} \\ &\leq r^{\frac{1}{\alpha}+g(\epsilon)+o(1)},\end{aligned}$$

with a function g tending to 0 as $\epsilon \to 0$.

Since the last display is true almost surely for all r sufficiently small, we use Lemma 2.20 (Frostman's lemma) to deduce that the Hausdorff dimension of $\mathcal{L}$ is a.s. larger than $\frac{1}{\alpha} + g(\epsilon)$. Taking $\epsilon \to 0$ we get that almost surely,

$$\dim_{\mathrm{H}}(\mathcal{L}) \geq \frac{1}{\alpha}. \qquad \square$$

2.5.2 Case $\beta > 1$ and construction of measures on the leaves

The following section is devoted to prove the following proposition.

Proposition 2.14. *Suppose that Hypothesis $\diamond_{\alpha,\beta}$ is satisfied and that the block* B *satisfies Hypothesis H_d for some $d \geq 0$. Then the Hausdorff dimension of $\mathcal{L}$ almost surely satisfies*

$$\begin{aligned}\dim_{\mathrm{H}}(\mathcal{L}) &\geq \frac{2\beta - 1 - 2\sqrt{(\beta-1)(\beta-\alpha d)}}{\alpha}, \quad \text{if } \alpha < \frac{1}{d}, \\ &\geq \frac{1}{\alpha} \quad \text{otherwise.}\end{aligned}$$

In the case $\beta > 1$, we cannot use the natural measure $\bar{\mu}$ to get a good lower bound on the Hausdorff dimension of $\mathcal{L}$ since, as stated in Proposition 2.4, the measure $\bar{\mu}$ does not charge the leaves. So the goal of this subsection is to artificially construct a probability measure concentrated on the leaves that will give us, using Frostman's lemma, the appropriate lower bound on the Hausdorff dimension, that is, the one matching with the upper bound derived in Section 2.4. The measure will be obtained as a sub-sequential limit of a sequence of measures concentrated on the blocks, and will only charge a strict subset of $\mathcal{L}$.

First, let us fix some notation. Recall the definition of G_n^ϵ in (2.3). It follows from Hypothesis $\diamond_{\alpha,\beta}$ that there exists a function $h(n)$ tending to 0 such that $\#G_n^{h(n)} = n^{1+o(1)}$. We choose such a function h and let

$$G_n := G_n^{h(n)}. \tag{2.16}$$

We will also use an increasing sequence of positive integers $(n_k)_{k\geq 0}$, such that for all $k \geq 0$, we have $n_{k+1} = \lceil n_k^\gamma \rceil$, with a fixed $\gamma > 1$, that we will optimise later. Also we suppose n_0 to be very large, with conditions that we will make explicit in what follows. For all $n \geq 1$, we set

$$\overline{\mathbf{b}}_n := \left\{ x \in \mathbf{b}_n \;\middle|\; \mathbf{d}_n(\boldsymbol{\rho}_n, x) > \frac{\mathrm{ht}(\mathbf{b}_n)}{2} \right\},$$

the "upper-half" of block $\mathbf{b}_n$. For technical reasons, we will only keep the blocks that behave reasonably well in our construction, see the forthcoming property (P_d), introduced in Section 2.5.2. We recursively define some random sets of integers. Let

$$\widetilde{G}_{n_0} = \{n \in G_{n_0} \mid \mathsf{B}_n \text{ satisfies property } (P_d)\}, \tag{2.17}$$

and for $k \geq 0$,

$$\widetilde{G}_{n_{k+1}} = \left\{ n \in G_{n_{k+1}} \,\middle|\, \mathsf{B}_n \text{ satisfies property } (P_d);\ \exists i \in \widetilde{G}_{n_k},\ \mathbf{b}_n \to \mathbf{b}_i,\ X_{n-1} \in \overline{\mathbf{b}}_i \right\}. \tag{2.18}$$

We then define for all $k \geq 0$,

$$\mathcal{B}_k = \bigcup_{n \in \widetilde{G}_{n_k}} \overline{\mathbf{b}}_n.$$

In other words, $\mathcal{B}_0$ is the union of all the upper-halves of the blocks $\mathbf{b}_n$, for n in G_{n_0} for which B_n behaves well, and $\mathcal{B}_{k+1}$ is defined to be the union of all the upper-halves of the blocks of index $n \in G_{n_{k+1}}$ that are grafted directly on $\mathcal{B}_k$, and such that B_n behaves well. Note that for the moment, $\mathcal{B}_k$ can be empty.

We define the measure $\sum_{n \in \widetilde{G}_{n_k}} \boldsymbol{\nu}_n$ and refer to it as the *mass measure*[3] on the k-th generation. To simplify notation we denote it by $|\cdot|$. We do not index it by k since the index for which we consider it is always clear from the context. We also define a sequence $(\pi_k)_{k\geq 0}$ of probability measures on $\mathcal{B}_k$ by

$$\pi_k := \frac{|\cdot \cap \mathcal{B}_k|}{|\mathcal{B}_k|},$$

the normalised mass measure on $\mathcal{B}_k$. Note that the sequence (π_k) is only well-defined on the event where $\mathcal{B}_k$ has non-zero mass for all k. In what follows we will ensure that it is the case for an event of strictly positive probability and only work conditionally on this event. Remark that, still conditionally on this event and on the event that $\mathcal{T}$ is compact, which has probability 1, the sequence $(\pi_k)_{k\geq 0}$ is a sequence of probability measures on a compact space, hence it admits at least one subsequential limit π for the Lévy-Prokhorov distance. We can check using [41, Lemma 17], which is essentially an application of the Portmanteau theorem, that π is concentrated on $\bigcap_{k\geq 0} \mathcal{T}(\mathcal{B}_k) \subset \mathcal{L}$.

Idea of the proof

Let us briefly explain how the measure π that we just constructed enables us to derive the appropriate lower bound for the Hausdorff dimension. We give the intuition for $\alpha < 1/d$; the idea for $\alpha > 1/d$ is very similar. We will be very rough for this sketch of proof and we keep the notation introduced above. Let us here forget that some blocks may not satisfy (P_d), and that we only deal with half-blocks.

Number of blocks in $\mathcal{B}_k$ Suppose that the number of blocks in $\mathcal{B}_k$ evolves like a power of n_k, say n_k^a. Then the total weight of $\mathcal{B}_k$ is $|\mathcal{B}_k| \approx n_k^a n_k^{-\beta}$, because all the blocks in $\mathcal{B}_k$ have weight $\approx n_k^{-\beta}$. Since the probability that any block with index in $[\![n_{k+1}, 2n_{k+1}]\!]$ is grafted on $\mathcal{B}_k$ is roughly $|\mathcal{B}_k|$, and since the number of blocks in $\mathcal{B}_{k+1}$ is roughly n_{k+1}^a, we have

$$n_k^{\gamma a} \approx n_{k+1}^a \approx |\mathcal{B}_k| \cdot n_{k+1} \approx n_k^a n_k^{-\beta} n_k^{\gamma}.$$

Hence we have $a = \frac{\gamma - \beta}{\gamma - 1}$, and so $|\mathcal{B}_k| \approx n_k^{\frac{\gamma(1-\beta)}{\gamma-1}}$.

[3]It may appear more natural to define the mass measure as simply $\sum_{n=1}^{\infty} \nu_n$ which gives $\mathcal{T}$ a finite mass. However, this definition would not have the property that for every set $S \subset \overline{\mathbf{b}}_n$, we have $|S| = \boldsymbol{\nu}_n(S)$. Indeed, when blocks can have an atom at their root, which is possible in the case $d = 0$, the contribution of the mass added by the roots of future blocks should be counted in $|S|$.

Estimation on π For each $k \geq 0$, the set $\mathcal{B}_k$ is made of blocks of size $\approx n_k^{-\alpha}$. Let us suppose that the quantities of the form $\pi(\mathrm{B}(x,r))$ are well-approximated by $\pi_k(\mathrm{B}(x,r))$, whenever $r \in [n_{k+1}^{-\alpha}, n_k^{-\alpha}]$. For x close to $\mathcal{B}_k$, such a ball typically intersects only one block of $\mathcal{B}_k$, with weight roughly $n_k^{-\beta}$, and since the block is d-dimensional, the ball covers a proportion $\left(r/n_k^{-\alpha}\right)^d$ of this block. So $|\mathrm{B}(x,r) \cap \mathcal{B}_k| \approx n_k^{-\beta}\left(r/n_k^{-\alpha}\right)^d$, and

$$\begin{aligned}\pi(\mathrm{B}(x,r)) \approx \pi_k(\mathrm{B}(x,r)) = \frac{|\mathrm{B}(x,r)\cap\mathcal{B}_k|}{|\mathcal{B}_k|} &\approx r^d n_k^{-\beta+\alpha d-\gamma\frac{\beta-1}{\gamma-1}} \\ &\approx r^d n_k^{\frac{1}{\gamma-1}(\gamma(\alpha d-1)+\beta-\alpha d)}.\end{aligned}$$

Then, if $\alpha < 1/d$, the last exponent is negative for γ large enough. For such γ, using $r > n_{k+1}^{-\alpha} \approx n_k^{-\alpha\gamma}$ yields

$$\pi(\mathrm{B}(x,r)) \leq r^{d-\frac{1}{\alpha\gamma(\gamma-1)}(\gamma(\alpha d-1)+\beta-\alpha d)}. \tag{2.19}$$

Optimisation We then choose γ such that the exponent $d - \frac{\gamma(\alpha d-1)-\alpha d+\beta}{\alpha\gamma(\gamma-1)}$ is maximal. We get the value $\frac{2\beta-1-2\sqrt{(\beta-1)(\beta-\alpha d)}}{\alpha}$ which matches our upper-bound.

Plan of the proof Our goal is now to make these heuristics rigorous. First we will make some precise estimation on how the mass of the blocks with indices in $\widetilde{G}_{n_k}$ is spread on subsets of $\mathcal{T}_{n_k-1}$. Then we will decompose each block of each $\mathcal{B}_k$ into subsets that we call fragments for which our preceding estimation holds. After that, we use this decomposition to control the behaviour of (π_k), and also how the measures π_k can approximate the limiting measure π. At the end we conclude by optimising over the parameters.

We will distinguish the two cases $d = 0$ and $d > 0$ and mostly work on the latter. We then explain quickly how the proof can be adapted to $d = 0$, in which fewer technicalities are involved.

Mass estimations

Before proving our main proposition, we have to state some technical lemmas that will allow us to control how regularly the mass of $\mathcal{B}_{k+1}$ is spread on $\mathcal{B}_k$. Let us now define the property (P_d), in a different way depending whether $d = 0$ or $d > 0$. Let $C > 0$ be a positive number and recall the definition of $(\star_{r_0})$ in Hypothesis H_d(i). For $d > 0$ we say that a pointed compact metric space endowed with a probability measure $(\mathbf{b}, \mathbf{d}, \boldsymbol{\rho}, \boldsymbol{\nu})$ satisfies (P_d) iff

$$\begin{cases} C^{-1} \leq \mathrm{ht}(\mathbf{b}) \leq C, \\ \boldsymbol{\nu}\left(\left\{x \in \mathbf{b} \;\middle|\; \mathbf{d}(\boldsymbol{\rho}, x) \geq \dfrac{\mathrm{ht}(\mathbf{b})}{2}\right\}\right) \geq C^{-1}, \\ (\mathbf{b}, \mathbf{d}, \boldsymbol{\rho}, \boldsymbol{\nu}) \text{ satisfies } (\star_{r_0}) \text{ with } r_0 = C^{-1}. \end{cases} \tag{P_d}$$

For $d = 0$ we say that $(\mathbf{b}, \mathbf{d}, \boldsymbol{\rho}, \boldsymbol{\nu})$ satisfies (P_0) iff $\mathbf{b}$ is finite and

$$\begin{cases} C^{-1} \leq \mathrm{ht}(\mathbf{b}) \leq C, \\ \#\mathbf{b} \leq C, \\ \forall x \in \mathbf{b},\; \nu(\{x\}) \geq C^{-1}. \end{cases} \tag{P_0}$$

In any case, under Hypothesis H_d(i), for any $d \geq 0$, we can choose C such that the underlying block $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ satisfies (P_d) with a positive probability $p > 0$, i.e.,

$$p = \mathbb{P}\left(\mathsf{B} \text{ satisfies } (P_d)\right) > 0.$$

From now on we fix such a constant C. We also set $\mathsf{M} := \nu\left(\left\{x \in \mathsf{B} \;\middle|\; \mathsf{D}(\rho, x) \geq \frac{\mathrm{ht}(\mathsf{B})}{2}\right\}\right)$, and $m := \mathbb{E}\left[\mathsf{M} \mid \mathsf{B} \text{ satisfies } (P_d)\right]$. We also denote by $\mathtt{M}$ a random variable with the law of M conditionally on the event $\{\mathsf{B} \text{ satisfies } (P_d)\}$.

Lemma 2.15. *Let S be a subset of some $\mathbf{b}_i$ with $i \leq n_k - 1$, measurable with respect to $\mathcal{F}_{n_k-1}$, the σ-field generated by the blocks and the gluings up to time $n_k - 1$. Let $\chi(S)$ be the total mass of the union of the sets $\left\{\overline{\mathbf{b}}_n \;\middle|\; n \in G_{n_k},\ \mathbf{b}_n \to \mathbf{b}_i,\ X_{n-1} \in S,\ \mathsf{B}_n \text{ satifies } (P_d)\right\}$, namely, the total mass of the half-blocks that are grafted on S with index in G_{n_k}, and such that the corresponding blocks satisfy property (P_d). Then for all $x \in [0, 1]$,*

$$\mathbb{P}\left(\left|\chi(S) - a_k|S|\right| > x a_k |S| \mid \mathcal{F}_{n_k-1}\right) \leq 2 \exp\left(-x^2 n_k^{1+o(1)} |S|\right),$$

where $a_k := pm \sum_{i \in G_{n_k}} \frac{w_i}{W_{i-1}}$ is such that $\mathbb{E}\left[\chi(S) \mid \mathcal{F}_{n_k-1}\right] = a_k|S|$ and the function $o(1)$ in the right-hand side does not depend on x.

This lemma roughly states that, for every subset $S \subset \mathbf{b}_i$ for $i \in G_{n_k}$, if the subset has enough mass to attract a substantial number of the blocks coming between time n_k and $2n_k$ then we have a good control on how the mass of $\mathcal{B}_k$ grafted on S can deviate from its expected value.

Proof. First we write $\chi(S)$ as

$$\chi(S) = \sum_{i \in G_{n_k}} \mathbf{1}_{\left\{U_i \leq \frac{|S|}{W_{i-1}}\right\}} \mathbf{1}_{\{\mathsf{B}_i \text{ satisfies } (P_d)\}} \mathsf{M}_i w_i,$$

where the (U_i) are independent uniform variables on $[0, 1]$, independent of everything else. Then we can compute

$$\begin{aligned}
\mathbb{E}\left[\chi(S) \mid \mathcal{F}_{n_k-1}\right] &= \sum_{i \in G_{n_k}} \frac{|S|}{W_{i-1}} p w_i \cdot \mathbb{E}\left[\mathsf{M}_i \mid \mathsf{B}_i \text{ satisfies } (P_d)\right] \\
&= |S| \cdot pm \left(\sum_{i \in G_{n_k}} \frac{w_i}{W_{i-1}}\right) = |S| \cdot a_k.
\end{aligned}$$

Let us bound the exponential moments of $\chi(S)$:

$$\begin{aligned}
\mathbb{E}\left[\exp(\theta \chi(S)) \mid \mathcal{F}_{n_k-1}\right] &= \prod_{i \in G_{n_k}} \left(\frac{p|S|}{W_{i-1}} \cdot \mathbb{E}\left[e^{\theta w_i \mathsf{M}_i} \;\middle|\; \mathsf{B}_i \text{ satisfies } (P_d)\right] + 1 - \frac{p|S|}{W_{i-1}}\right) \\
&= \prod_{i \in G_{n_k}} \left(1 + \frac{p|S|}{W_{i-1}} \left(\mathbb{E}\left[e^{\theta w_i \mathtt{M}}\right] - 1\right)\right) \\
&\leq \exp\left(p|S| \sum_{i \in G_{n_k}} \frac{1}{W_{i-1}} \left(\mathbb{E}\left[e^{\theta w_i \mathtt{M}}\right] - 1\right)\right),
\end{aligned}$$

where we have used the inequality $e^z \geq 1+z$, in the last line. Now,

$$\begin{aligned}
&\mathbb{E}\left[\exp(\theta(\chi(S) - a_k|S|)) \mid \mathcal{F}_{n_k-1}\right] \\
&\leq \exp\left(p|S| \sum_{i\in G_{n_k}} \frac{1}{W_{i-1}}(\mathbb{E}\left[e^{\theta w_i \mathtt{M}}\right] - 1) - \theta|S| \sum_{i\in G_{n_k}} \frac{pmw_i}{W_{i-1}} \right) \\
&\leq \exp\left(p|S| \left(\sum_{i\in G_{n_k}} \frac{1}{W_{i-1}}(\mathbb{E}\left[e^{\theta w_i \mathtt{M}}\right] - 1 - \theta m w_i) \right)\right) \\
&\leq \exp\left(p|S| \left(\sum_{i\in G_{n_k}} \frac{1}{W_{i-1}} c(\theta w_i)^2 \right)\right).
\end{aligned}$$

Here we used the fact that for $z \in [-1\,,1]$, we have $e^z \leq 1+z+3z^2$ and so

$$\mathbb{E}\left[e^{z\mathtt{M}}\right] - 1 - z\mathbb{E}\left[\mathtt{M}\right] \leq \mathbb{E}\left[1 + z\mathtt{M} + 3(z\mathtt{M})^2\right] - 1 - z\mathbb{E}\left[\mathtt{M}\right] \leq 3\mathbb{E}\left[\mathtt{M}^2\right] z^2 \leq cz^2,$$

for c a constant. Since we ask that $z \in [-1\,,1]$, the computation above is valid if we restrict ourselves to $|\theta| \leq (\sup_{i\in G_{n_k}} w_i)^{-1} = n_k^{\beta+o(1)}$. Note that we can use this inequality for negative values of θ. Hence for $x \in [0\,,1]$

$$\begin{aligned}
&\mathbb{P}\left(|\chi(S) - a_k|S|| > xa_k|S| \mid \mathcal{F}_{n_k-1}\right) \\
&\leq \mathbb{P}\left(\chi(S) - a_k|S| > xa_k|S| \mid \mathcal{F}_{n_k-1}\right) + \mathbb{P}\left(-(\chi(S) - a_k|S|) > xa_k|S| \mid \mathcal{F}_{n_k-1}\right) \\
&\leq \mathbb{P}\left(\exp\left(\theta(\chi(S) - a_k|S|)\right) > e^{\theta x a_k|S|} \,\Big|\, \mathcal{F}_{n_k-1}\right) + \mathbb{P}\left(\exp\left(-\theta\left(\chi(S) - a_k|S|\right)\right) > e^{\theta x a_k|S|} \,\Big|\, \mathcal{F}_{n_k-1}\right) \\
&\leq 2\exp\left(p|S| \left(\sum_{i\in G_{n_k}} \frac{1}{W_{i-1}} c(\theta w_i)^2 \right) - \theta x a_k |S| \right) \\
&= 2\exp\left(p|S| \left(\sum_{i\in G_{n_k}} \frac{1}{W_{i-1}} (c(\theta w_i)^2 - \theta x m w_i) \right)\right).
\end{aligned}$$

Taking $\theta = xn_k^{\beta-\epsilon}$ in the last inequality, which is possible for n_k large enough, this gives

$$\begin{aligned}
&\mathbb{P}\left(\left|\chi(S) - a_k|S|\right| > xa_k|S| \,\big|\, \mathcal{F}_{n_k-1}\right) \\
&\leq 2\exp\left(p|S|x^2 \left(\sum_{i\in G_{n_k}} \frac{1}{W_{i-1}} (n_k^{\beta-\epsilon} w_i)(cn_k^{\beta-\epsilon} w_i - m) \right)\right).
\end{aligned}$$

From our assumptions on the sequence (w_n), we have $n_k^{\beta-\epsilon}(\sup_{i\in G_{n_k}} w_i) \to 0$ and hence $(cn_k^{\beta-\epsilon} w_i - m)$ is eventually smaller than $-\frac{m}{2}$, uniformly for $i \in G_{n_k}$. Also $\frac{1}{W_n}$ is always greater than $\frac{1}{W_\infty}$. Combining this with the last display we get, for n_k large enough

$$\begin{aligned}
&\mathbb{P}\left(\left|\chi(S) - a_k|S|\right| > xa_k|S| \,\big|\, \mathcal{F}_{n_k-1}\right) \\
&\leq 2\exp\left(-x^2|S| \cdot \frac{pm}{2W_\infty} \# G_{n_k} (\inf_{i\in G_{n_k}} w_i) n_k^{\beta-\epsilon} \right) \\
&\leq 2\exp\left(-x^2|S| n_k^{1-\epsilon+o(1)} \right).
\end{aligned}$$

Now for every $\epsilon > 0$, this inequality is true for n_k large enough, so this proves the lemma. □

Let us also state another technical lemma, the proof of which is in the Appendix 2.A.4.

Lemma 2.16. *Suppose that Hypothesis $\diamond_{\alpha,\beta}$ is satisfied. For the sequence (a_i) defined in Lemma 2.15,*

$$\prod_{i=1}^{k} a_i = n_k^{\frac{\gamma(1-\beta)}{(\gamma-1)}+o(1)},$$

where the $o(1)$ is considered when $k \to \infty$.

Construction of fragments, case $d > 0$

Fragments of a random block Let us discuss how we can decompose a metric space into a partition of subsets that we call r-*fragments*, all of them having a diameter of order r. Suppose that the random block $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ comes with a sequence of random points $(X_n)_{n\geq 1}$, which are i.i.d. with law ν, conditionally on $(\mathsf{B}, \mathsf{D}, \rho, \nu)$, and that this block satisfies Hypothesis H_d(i), for some $d > 0$. The following lemma ensures that in this setting, we can construct a partition $F(\mathsf{B}, r) = (f_i^{(r)})_{1\leq i\leq N}$ of r-fragments of $(\mathsf{B}, \mathsf{D}, \rho, \nu)$, which have approximately equal diameter and measure. Recall the function φ defined in Hypothesis H_d(i), and the notation $N_r(\mathsf{B})$ for the minimal number of balls of radius r needed to cover B.

Lemma 2.17. *Suppose that $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ satisfies Hypothesis H_d(i) for some $d > 0$. For any $r \in [0, 1]$, we construct a finite partition of Borel subsets $(f_i^{(r)})_{1\leq i\leq N}$ of the block $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ in a deterministic way from $(\mathsf{B}, \mathsf{D}, \rho, \nu)$ and the sequence of random points $(X_n)_{n\geq 1}$. There exists two functions ψ and ϕ defined on the interval $[0, 1]$, which tend to 0 at 0 such that the following holds almost surely on the event $\{(\mathsf{B}, \mathsf{D}, \rho, \nu)$ satisfies $(\star_{r_0})\}$, for any $r_0 > 3r$:*

(i) *For all $1 \leq i \leq N$,*

$$\operatorname{diam}(f_i^{(r)}) \leq 2r \quad \textit{and} \quad \left(\frac{r}{4}\right)^{d+\varphi(r/4)} \leq \nu(f_i^{(r)}) \leq r^{d-\varphi(r)}.$$

(ii) *For all $r' < r_0/3$, we have*

$$\forall x \in \mathsf{B}, \quad \#\left\{1 \leq i \leq N \,\middle|\, \mathrm{B}(x, r') \cap f_i^{(r)} \neq \emptyset\right\} \leq (r \vee r')^{d+\psi(r\vee r')} \cdot r^{-d+\phi(r)}.$$

(iii) *The (random) number N of fragments satisfies*

$$N \leq N_{r/4}(\mathsf{B}) \quad \textit{and} \quad N \leq \left(\frac{r}{4}\right)^{-d-\varphi(r/4)}.$$

In this chapter, we use this construction on $(\mathsf{B}_n, \mathsf{D}_n, \rho_n, \nu_n)$, assuming that for all $n \geq 1$, a sequence $(X_{n,j})_{j\geq 1}$ is defined on the same probability space and that this sequence is i.i.d. with law ν_n, conditionally on $(\mathsf{B}_n, \mathsf{D}_n, \rho_n, \nu_n)$. For any $n \geq 1$ and $r > 0$, we denote $F(\mathsf{B}_n, r) = \left\{f_i^{(r)} \,\middle|\, 1 \leq i \leq N\right\}$ the partition of B_n into (random) r-fragments which is given by the lemma. The proof of Lemma 2.17 can be found in the Appendix.

Decomposition of $\mathbf{b}_n$ into fragments Fix a parameter $\eta \in (0, \frac{1}{d})$. We want to decompose every $\mathbf{b}_n$, for $n \in \widetilde{G}_{n_k}$, in fragments of size approximately $n_{k+1}^{-\eta}$. For that, it is sufficient to use $F(\mathsf{B}_n, \mathsf{r}_n)$ the decomposition of B_n in r_n-fragments with $\mathsf{r}_n = (\lambda_n^{-1} \cdot n_{k+1}^{-\eta})$, which is given by Lemma 2.17. Let us emphasise that these fragments are constructed as subsets of B_n, but we consider them as subsets of $\mathbf{b}_n$ in what follows, without changing notation. We define the set

F_k is as the collection of all these fragments coming from every $\mathbf{b}_n$ with $n \in \widetilde{G}_{n_k}$. We have of course

$$\bigcup_{f \in F_k} f = \bigcup_{n \in \widetilde{G}_{n_k}} \mathbf{b}_n. \tag{2.20}$$

In our construction, we decided to keep only the blocks that were sufficiently well-behaved with respect to some properties that will be useful now. Recall the definition of the random set $\widetilde{G}_{n_k}$ in equations (2.17) and (2.18). Remark that, from the definition of G_{n_k}, we have,

$$c_k := \min_{n \in G_{n_k}} \left(\lambda_n^{-1} \cdot n_{k+1}^{-\eta}\right) = n_k^{\alpha - \gamma\eta + o(1)} \quad \text{and} \quad C_k := \max_{n \in G_{n_k}} \left(\lambda_n^{-1} \cdot n_{k+1}^{-\eta}\right) = n_k^{\alpha - \gamma\eta + o(1)}.$$

If γ and η are such that $\gamma > \frac{\alpha}{\eta}$, then the last exponent is strictly negative, and so we can take n_0 sufficiently large so that $C_k < C^{-1}/3$, for all $k \geq 0$. For $n \in \widetilde{G}_{n_k}$, we know that B_n satisfies $(\star_{r_0})$ with $r_0 = C^{-1}$. Hence, for all $n \in \widetilde{G}_{n_k}$, we have $3\mathsf{r}_n \leq r_0$, and so the conclusions of Lemma 2.17 hold simultaneously for all the decompositions $F(\mathsf{B}_n, \mathsf{r}_n)$ for $n \in \widetilde{G}_{n_k}$.

Control on the mass and number of fragments Recall the function h that we defined in (2.16), which tends to 0 at infinity, and the function φ specified in Hypothesis H_d(i), which tends to 0 at 0. Thanks to Lemma 2.17, we get, for all $f \in F_k$ such that $f \subset \mathbf{b}_n$,

$$\begin{aligned} |f| = w_n \cdot \nu_n(f) &\underset{(2.16),\ \text{Lem.}2.17(i)}{\geq} n_k^{-\beta - h(n_k)} \cdot (\lambda_n^{-1} \cdot n_{k+1}^{-\eta})^{d + \varphi(n_k^{-\alpha d + \alpha + o(1)})} \\ &\geq n_k^{-\beta - h(n_k)} \cdot c_k^{d + \varphi(C_k)}. \end{aligned}$$

Note that the last quantity is deterministic and only depends on n_k, and so almost surely,

$$\begin{aligned} \min_{f \in F_k} |f| &\geq n_k^{-\beta - h(n_k)} \cdot c_k^{d + \varphi(C_k)} \\ &\geq n_k^{-\eta d \gamma + \alpha d - \beta + o(1)} = n_{k+1}^{-\eta d + \frac{1}{\gamma}(\alpha d - \beta) + o(1)}. \end{aligned} \tag{2.21}$$

Note that a similar computation using upper-bounds instead of lower-bounds also yields, almost surely,

$$\max_{f \in F_k} |f| \leq n_k^{-\beta + o(1)} \cdot C_k^{d - \varphi(c_k)} \leq n_k^{-\eta d \gamma + \alpha d - \beta + o(1)}, \tag{2.22}$$

where the right-hand side is deterministic. Also, from Lemma 2.17(iii), we get that the number of fragments obtained from the block $\mathbf{b}_n$ by that construction is bounded above by $(\mathsf{r}_n/4)^{-d - \varphi(\mathsf{r}_n/4)}$, with $\mathsf{r}_n = \lambda_n^{-1} \cdot n_{k+1}^{-\eta} = n_k^{\alpha - \gamma\eta + o(1)}$, and so at the end, the total number of fragments in F_k is bounded above by a deterministic quantity which grows at most polynomially in n_k.

Construction of fragments, case $d = 0$

In this case, we will consider the finite number of points of each block as a decomposition into fragments, hence we set $F_k = \left\{ \{x\} \;\middle|\; x \in \mathbf{b}_n,\ n \in \widetilde{G}_{n_k} \right\}$. Note that

$$\forall f \in F_k,\ f \subset \mathbf{b}_n, \quad |f| = w_n \cdot \nu_n(f) \underset{(2.16),\ (P_0)}{\geq} n_k^{-\beta - h(n_k)} \cdot C^{-1},$$

and so the equations (2.21) and (2.22) are still valid when $d = 0$, and also the number of fragments in F_k grows linearly, hence polynomially in n_k.

Using the mass estimates

Recall that we fixed a parameter $\eta \in (0, \frac{1}{d})$. We let $0 < \epsilon < (1 - \eta d)$. If γ and η are such that $\gamma > \frac{\beta - \alpha d}{1 - \eta d - \epsilon}$, then $-\eta d + \frac{1}{\gamma}(\alpha d - \beta) > -1 + \epsilon$. And so we get from (2.21) that $\min_{f \in F_k} |f| \geq n_{k+1}^{-1+\epsilon+o(1)}$. We can apply the result of Lemma 2.15 for every fragment $f \in F_k$, with $x = n_{k+1}^{-\epsilon/4}$,

$$\begin{aligned}
&\mathbb{P}\left(\left| \chi(f) - a_{k+1}|f| \right| > n_{k+1}^{-\epsilon/4} a_{k+1} |f| \;\middle|\; \mathcal{F}_{n_{k+1}-1} \right) \\
&\leq 2 \exp\left(- \left(n_{k+1}^{-\epsilon/4} \right)^2 n_{k+1}^{1+o(1)} \min_{f \in F_k} |f| \right) \\
&\leq 2 \exp\left(- n_{k+1}^{-\epsilon/2} n_{k+1}^{1+o(1)} n_{k+1}^{-1+\epsilon+o(1)} \right) \leq 2 \exp\left(- n_{k+1}^{\epsilon/4} \right) \quad \text{for } n_{k+1} \text{ large enough.}
\end{aligned}$$

For that, again, we impose that n_0 is large enough such that the last display is true for all k and for all f. Now we can sum this over all fragments,

$$\begin{aligned}
&\mathbb{E}\left[\sum_{k=0}^{\infty} \sum_{f \in F_k} \mathbb{P}\left(\left| \chi(f) - a_{k+1}|f| \right| > n_{k+1}^{-\epsilon/4} a_{k+1} |f| \;\middle|\; \mathcal{F}_{n_{k+1}-1} \right) \right] \\
&\leq \sum_{k=0}^{\infty} \mathbb{E}\left[\# F_k \right] \cdot 2 \exp\left(- n_{k+1}^{\epsilon/4} \right) \underset{n_0 \to \infty}{\longrightarrow} 0,
\end{aligned}$$

since $(\# F_k)$ is almost surely bounded by a deterministic quantity which grows at most polynomially in n_k. The same is true for the $\mathcal{B}_k$, i.e.,

$$\mathbb{E}\left[\sum_{k=0}^{\infty} \mathbb{P}\left(\left| \chi(\mathcal{B}_k) - a_{k+1}|\mathcal{B}_k| \right| > n_{k+1}^{-\epsilon/4} a_{k+1} |\mathcal{B}_k| \;\middle|\; \mathcal{F}_{n_{k+1}-1} \right) \right] \underset{n_0 \to \infty}{\longrightarrow} 0.$$

In the rest of Section 2.5.2, we will fix n_0 large enough and work on the event of large probability $\mathcal{E}$ on which we have, for all $k \geq 0$ and for all $f \in F_k$

$$\left| \chi(\mathcal{B}_k) - a_{k+1}|\mathcal{B}_k| \right| \leq n_{k+1}^{-\epsilon/4} a_{k+1} |\mathcal{B}_k| \quad \text{and} \quad \left| \chi(f) - a_{k+1}|f| \right| \leq n_{k+1}^{-\epsilon/4} a_{k+1} |f|. \tag{2.23}$$

Remark that thanks to Section 2.2.4, giving a lower bound of the Hausdorff dimension on a set of positive probability is enough to prove that the bound holds almost surely. Note that this construction depends on the parameters η and ϵ and γ. The parameters must satisfy

$$\eta \in (0, \frac{1}{d}), \quad \epsilon \in (0, 1 - \eta d), \quad \gamma > \max\left(\frac{\alpha}{\eta}, \frac{\beta - \alpha d}{1 - \eta d - \epsilon} \right), \tag{2.24}$$

and we can choose them in this particular order.

Control on the limiting measure

In this section, the values of η and ϵ and n_0 are fixed in such a way that the construction of the previous section holds. Note that everything in the section implicitly depends on those values. On the event $\mathcal{E}$, if we consider a fragment $f \in F_k$, we have a very good control on the values of $\pi_i(\mathcal{T}(f))$ for $i \geq k$. Indeed set

$$c_1 = \prod_{k=0}^{\infty} (1 - n_{k+1}^{-\frac{\epsilon}{4}}) \quad \text{and} \quad c_2 = \prod_{k=0}^{\infty} (1 + n_{k+1}^{-\frac{\epsilon}{4}}).$$

Remark that both c_1 and c_2 are strictly positive real numbers. Using in cascade the estimations (2.23) which hold on the event $\mathcal{E}$, we get that for $f \in F_k$ and $i \geq k$,

$$|\mathcal{T}(f) \cap \mathcal{B}_i| \leq c_2 |f| \left(\prod_{j=k+1}^{i} a_j \right). \tag{2.25}$$

In fact we can use the same argument for $\mathcal{B}_k$, which is not empty on the event $\mathcal{E}$. For k large enough we can write

$$|\mathcal{B}_i| \in |\mathcal{B}_k| \left(\prod_{j=k+1}^{i} a_j \right) \cdot [c_1\,, c_2]. \tag{2.26}$$

Remark that (2.26) combined with Lemma 2.16 yields that almost surely on $\mathcal{E}$,

$$n_k^{\frac{\gamma(1-\beta)}{(\gamma-1)}+o(1)} \leq |\mathcal{B}_k| \leq n_k^{\frac{\gamma(1-\beta)}{(\gamma-1)}+o(1)}, \tag{2.27}$$

where the upper and lower-bound are both deterministic. For the normalized mass measure π_i on $\mathcal{B}_i$, we have

$$\pi_i(\mathcal{T}(f)) = \frac{|\mathcal{T}(f) \cap \mathcal{B}_i|}{|\mathcal{B}_i|} \underset{(2.25),(2.26)}{\leq} \frac{c_2}{c_1} \cdot \frac{|f|}{|\mathcal{B}_k|}.$$

If π is a sub-sequential limit of the (π_k), using Portmanteau theorem (remark that π is concentrated on the leaves and the leaves of $\mathcal{T}(f)$ belong to the interior of $\mathcal{T}(f)$), we get

$$\pi(\mathcal{T}(f)) \leq \frac{c_2}{c_1} \cdot \frac{|f|}{|\mathcal{B}_k|}.$$

Then,

$$\max_{f \in F_k} \pi(\mathcal{T}(f)) \leq \frac{c_2}{c_1} \frac{1}{|\mathcal{B}_k|} \max_{f \in F_k} |f|.$$

We can now write, for all $r > 0$, $x \in \mathcal{T}$,

$$\pi(\mathrm{B}(x,r)) \leq \sum_{f \in F_k,\ f \cap \mathrm{B}(x,r) \neq \emptyset} \pi(\mathcal{T}(f)) \leq \#\{f \in F_k,\ f \cap \mathrm{B}(x,r) \neq \emptyset\} \cdot \max_{f \in F_k} \pi(\mathcal{T}(f)).$$

Putting everything together, we get

$$\pi(\mathrm{B}(x,r)) \leq \#\{f \in F_k,\ f \cap \mathrm{B}(x,r) \neq \emptyset\} \cdot \frac{c_2}{c_1} \frac{1}{|\mathcal{B}_k|} \max_{f \in F_k} |f|. \tag{2.28}$$

Control on the number of fragments intersecting a ball

From (2.28), we see that the last thing that we have to estimate is $\#\{f \in F_k,\ f \cap \mathrm{B}(x,r) \neq \emptyset\}$, the number of fragments of F_k that have a non-empty intersection with a ball of radius r. Since the measure π only charges $\bigcap_{k\geq 1} \mathcal{T}(\mathcal{B}_k)$, we are only interested in balls centred around points belonging to this set. Let us fix some notation again. For all $k \geq 0$, we set

$$\Delta_k := \inf \left\{ \mathrm{d}(x,y) \mid x \in \mathcal{B}_{k-1},\ y \in \mathcal{B}_k \right\},$$

the set distance between levels $\mathcal{B}_{k-1}$ and $\mathcal{B}_k$, for the integers k for which it is possible. On the event $\mathcal{E}$, this quantity is well-defined for all $k \geq 1$ and the following upper and lower-bounds are almost surely satisfied

$$n_k^{-\alpha+o(1)} = \frac{C^{-1}}{2} \min_{n \in G_{n_k}} \lambda_n \leq \Delta_k \leq C \max_{n \in G_{n_k}} \lambda_n = n_k^{-\alpha+o(1)}. \tag{2.29}$$

Now let us state a lemma.

Lemma 2.18. *Let $x \in \bigcap_{k\geq 1} \mathcal{T}(\mathcal{B}_k)$. For $k \geq 0$, we denote $x_k := [x]_{2n_k} \in \mathcal{B}_k$. If $\mathbf{b}_n$ is the block of $\mathcal{B}_k$ such that $x_k \in \mathbf{b}_n$, we have, $\forall r \in [0\,, \Delta_k]$,*

$$\{f \in F_k \mid f \cap \mathrm{B}(x, r) \neq \emptyset\} \subset \{f \in F_k \mid \ f \cap \mathrm{B}(x_k, r) \cap \mathbf{b}_n \neq \emptyset\}\,. \tag{2.30}$$

The proof of this lemma is simple and left to the reader. It tells us is that in fact, if r is small enough, then all the fragments $f \in F_k$ who intersect the ball of centre x and radius r belong to the same block. This will allow us in the sequel, combined with Lemma 2.17(ii), to bound the number of fragments involved, which is what we wanted.

Obtaining the lower-bound

In order to get the lower-bound on the Hausdorff dimension of $\mathcal{L}$ matching that of the theorem, we have to distinguish between the case $\alpha < \frac{1}{d}$ and the case $\alpha > \frac{1}{d}$. The case $\alpha = \frac{1}{d}$ can be recovered by a monotonicity argument, as seen in Section 2.2.4. Whenever $d = 0$, we have $\frac{1}{d} = +\infty$ and only the first case can happen.

Case $\beta > 1$ and $\alpha < 1/d$ We use the construction of Section 2.5.2 with $\eta = \alpha$. Recall (2.24) for the admissible parameters of the construction. In this case, if ϵ is fixed and small enough, the only condition on γ implied by (2.24) is $\gamma > \frac{\beta-\alpha d}{1-\alpha d-\epsilon}$ since $\gamma > \frac{\alpha}{\eta}$ reduces to $\gamma > 1$, which is already contained in the previous inequality because $\frac{\beta-\alpha d}{1-\alpha d-\epsilon} > 1$. We define

$$\Lambda_k := \Delta_k \wedge \left(\frac{\min_{n\in G_{n_k}} \lambda_n}{\log n_k}\right).$$

Using (2.29), we get $n_k^{-\alpha+o(1)} \leq \Lambda_k \leq n_k^{-\alpha+o(1)}$, almost surely on the event $\mathcal{E}$, where the functions denoted $o(1)$ in the lower and upper-bound are deterministic functions of k. Here the choice of $\frac{1}{\log n_k}$ is rather arbitrary and we could change it to any quantity that tends to 0 and is $n_k^{o(1)}$ as $n_k \to \infty$. Now we claim the following:

Lemma 2.19. *On the event $\mathcal{E}$, for all $k \geq 0$, for any $d \geq 0$, almost surely,*

$$\forall r \in [\Lambda_{k+1}\,, \Lambda_k], \quad \#\{f \in F_k \mid f \cap \mathrm{B}(x, r) \neq \emptyset\} \leq r^{d+o(1)} n_k^{\alpha\gamma d+o(1)}. \tag{2.31}$$

Note that the bounds Λ_k of the interval on which we consider r are random, but the upper bound given by (2.31) is deterministic.

Proof. For $r \in [\Lambda_{k+1}\,, \Lambda_k]$, we have $r \leq \Lambda_k \leq \Delta_k$ so using Lemma 2.18, with $x_k = [x]_{2n_k}$ and n such that $x_k \in \mathbf{b}_n$, we know that (2.30) holds. In the case $d > 0$, the fragments of F_k that come from $\mathbf{b}_n$ were constructed as fragments of B_n of size $\mathsf{r}_n := \lambda_n^{-1} n_{k+1}^{-\alpha}$. Recall that we denote the set of these fragments, seen as subsets of B_n, by $F(\mathsf{B}_n, \mathsf{r}_n)$ and denote the point of B_n corresponding to $x_k \in \mathbf{b}_n$ by $\mathsf{x_k}$. The analogue of the ball $\mathrm{B}(x_k, r)$ in B_n is then the ball of centre $\mathsf{x_k}$ and radius $\mathsf{r'}_n := \lambda_n^{-1} r$. From our definition of Λ_k we have

$$\mathsf{r'}_n = \lambda_n^{-1} r \leq \lambda_n^{-1} \frac{\min_{n\in G_{n_k}} \lambda_n}{\log n_k} \leq \frac{1}{\log n_k} \underset{k\to\infty}{\longrightarrow} 0,$$

as well as $\mathsf{r}_n := \lambda_n^{-1} n_{k+1}^{-\alpha} = n_k^{\alpha-\gamma\alpha+o(1)} \to 0$ when $k \to \infty$. Applying Lemma 2.17(ii) yields

$$\begin{aligned}
\#\{f \in F(\mathsf{B}_n, \mathsf{r}_n) \mid \ f \cap \mathrm{B}(\mathsf{x_k}, \mathsf{r}_n) \neq \emptyset\} &\underset{\text{Lem.2.17}(ii)}{\leq} (\mathsf{r}_n \vee \mathsf{r'}_n)^{d+\psi(\mathsf{r}_n \vee \mathsf{r'}_n)} \cdot \mathsf{r}_n^{-d+\phi(\mathsf{r}_n)} \\
&\leq \left((\lambda_n^{-1} n_{k+1}^{-\alpha}) \vee (\lambda_n^{-1} r)\right)^{d+o(1)} \cdot (\lambda_n^{-1} n_{k+1}^{-\alpha})^{-d+o(1)} \\
&\leq r^{d+o(1)} n_k^{\alpha\gamma d+o(1)},
\end{aligned}$$

and the last quantity is deterministic. Since any fragment in $\{f \in F_k \mid f \cap \mathrm{B}(x,r) \neq \emptyset\}$ corresponds to a fragment in $\{f \in F(\mathsf{B}_n, \mathsf{r}_n) \mid f \cap \mathrm{B}(\mathsf{x}_\mathsf{k}, \mathsf{r}_n) \neq \emptyset\}$, the cardinality of $\{f \in F_k \mid f \cap \mathrm{B}(x,r) \neq \emptyset\}$ is almost surely bounded above by the last display, which proves that (2.31) holds whenever $d > 0$. In the case $d = 0$, from our definition of fragments and the property (P_0), we easily have

$$\#\{f \in F_k \mid f \cap \mathrm{B}(x_k, r) \cap \mathbf{b}_n \neq \emptyset\} \leq C \leq r^{d+o(1)} n_k^{\alpha\gamma d + o(1)},$$

and so (2.31) also holds whenever $d = 0$. □

We can compute

$$\begin{aligned}
\pi(\mathrm{B}(x,r)) &\underset{(2.28)}{\leq} \#\{f \in F_k,\ f \cap \mathrm{B}(x,r) \neq \emptyset\} \cdot \frac{c_2}{c_1} \cdot \frac{1}{|\mathcal{B}_k|} \cdot \max_{f \in F_k} |f| \\
&\underset{(2.31),(2.27),(2.22)}{\leq} (r^{d+o(1)} n_k^{\alpha\gamma d + o(1)}) \cdot r^{o(1)} \cdot (n_k^{\frac{\gamma(\beta-1)}{(\gamma-1)} + o(1)}) \cdot (n_k^{-\alpha\gamma d + \alpha d - \beta + o(1)}) \\
&\leq r^{d+o(1)} \cdot n_k^{\frac{1}{(\gamma-1)}(\gamma\alpha d - \alpha d + \beta - \gamma + o(1))} \\
&\leq r^{d - \frac{1}{\alpha\gamma(\gamma-1)}(\gamma\alpha d - \alpha d + \beta - \gamma) + o(1)}.
\end{aligned}$$

In the last line we have used that $r > \Lambda_{k+1} \geq n_{k+1}^{-\alpha+o(1)} = n_k^{-\gamma\alpha+o(1)}$ and so $n_k > r^{-\frac{1}{\alpha\gamma}+o(1)}$ and the fact that $\gamma\alpha d - \alpha d + \beta - \gamma < 0$ because $\gamma > \frac{\beta - \alpha d}{1-\alpha d}$. Let us now maximise the quantity $d - \frac{\gamma\alpha d - \alpha d + \beta - \gamma}{\alpha\gamma(\gamma-1)}$ for $\gamma \in (\frac{\beta-\alpha d}{1-\alpha d}, +\infty)$. It is an easy exercise to see that the maximum is attained at $\bar{\gamma} = \frac{\beta - \alpha d + \sqrt{(\beta-1)(\beta-\alpha d)}}{1-\alpha d}$, with value

$$\frac{2\beta - 1 - 2\sqrt{(\beta-1)(\beta-\alpha d)}}{\alpha}.$$

If we fix ϵ small enough then, the value of γ that maximises the last display satisfies $\gamma > \frac{\beta-\alpha d}{1-\alpha d-\epsilon}$ and so, using this value to construct π, we get that on the event $\mathcal{E}$, for all $x \in \bigcap_{k\geq 1} \mathcal{T}(\mathcal{B}_k)$,

$$\pi(\mathrm{B}(x,r)) \leq r^{\frac{2\beta-1-2\sqrt{(\beta-1)(\beta-\alpha d)}}{\alpha} + o(1)},$$

which allows us to conclude using Lemma 2.20.

Case $\beta > 1$ and $\alpha > 1/d$ Here we suppose that $d > 0$. We fix $\eta < \frac{1}{d}$ which we suppose to be very close to $\frac{1}{d}$ and a small $\epsilon > 0$ and use the construction of Section 2.5.2 with these values, which satisfy (2.24) if we take $\gamma > \max\left(\frac{\beta-\alpha d}{1-\eta d-\epsilon}, \frac{\alpha}{\eta}\right)$.

- For $r \in [\Lambda_{k+1}, n_{k+1}^{-\eta}]$, we apply Lemma 2.19 to get

$$\#\{f \in F_k,\ f \cap \mathrm{B}(x,r) \neq \emptyset\} \leq n_k^{o(1)}.$$

Now we can use the upper-bound (2.28), replacing term by term

$$\begin{aligned}
\pi(\mathrm{B}(x,r)) &\underset{(2.28)}{\leq} \#\{f \in F_k,\ f \cap \mathrm{B}(x,r) \neq \emptyset\} \cdot \frac{c_2}{c_1} \frac{1}{|\mathcal{B}_k|} \max_{f \in F_k} |f| \\
&\underset{(2.27),(2.22)}{\leq} n_k^{o(1)} \cdot \frac{c_2}{c_1} n_k^{\frac{\gamma(\beta-1)}{(\gamma-1)} + o(1)} n_k^{-\eta d\gamma + \alpha d - \beta + o(1)} \\
&\leq n_k^{\frac{\gamma(\beta-1)}{(\gamma-1)} - \gamma\eta d + \alpha d - \beta + o(1)}.
\end{aligned}$$

Hence, using the fact that $r > \Lambda_{k+1} = n_k^{-\gamma\alpha+o(1)}$ and that the exponent in the last display is negative, we get

$$\begin{aligned}\pi(\mathrm{B}(x,r)) &\leq r^{-\frac{1}{\gamma\alpha}\cdot\left(\frac{\gamma(\beta-1)}{(\gamma-1)}-\gamma\eta d+\alpha d-\beta\right)+o(1)}\\ &\leq r^{\frac{\eta d}{\alpha}-\frac{1}{\alpha\gamma}\left(\frac{\gamma(\beta-1)}{(\gamma-1)}+\alpha d-\beta\right)+o(1)}.\end{aligned}$$

- For $r \in [n_{k+1}^{-\eta}, \Lambda_k]$, we have once again using Lemma 2.19,

$$\#\{f \in F_k,\ f \cap \mathrm{B}(x,r) \neq \emptyset\} \leq r^{d+o(1)} n_{k+1}^{\eta d+o(1)}.$$

Replacing in (2.28) yields

$$\begin{aligned}\pi(\mathrm{B}(x,r)) &\leq r^{d+o(1)} n_{k+1}^{\eta d+o(1)} \cdot \frac{c_2}{c_1} \frac{1}{n_k^{\frac{\gamma(1-\beta)}{(\gamma-1)}+o(1)}} n_k^{-\eta d\gamma+\alpha d-\beta+o(1)}\\ &\leq r^{d+o(1)} \cdot n_k^{\alpha d-1+\frac{\beta-1}{\gamma-1}+o(1)}.\end{aligned}$$

Since $r \leq \Lambda_k \leq n_k^{-\alpha+o(1)}$, we have $n_k \leq r^{-\frac{1}{\alpha}+o(1)}$. Since the quantity $\alpha d - 1 + \frac{\beta-1}{\gamma-1}$ is positive, we can write

$$\begin{aligned}\pi(\mathrm{B}(x,r)) &\leq r^{d+o(1)} \cdot r^{-\frac{1}{\alpha}\left(\alpha d-1+\frac{\beta-1}{\gamma-1}\right)+o(1)}\\ &\leq r^{d-\frac{\alpha d-1}{\alpha}-\frac{\beta-1}{\alpha(\gamma-1)}+o(1)}\\ &\leq r^{\frac{1}{\alpha}-\frac{\beta-1}{\alpha(\gamma-1)}+o(1)}.\end{aligned}$$

Now the result is obtained by taking $\epsilon \to 0$, $\eta \to \frac{1}{d}$ and $\gamma \to \infty$.

To conclude the proof of Proposition 2.14, we have to prove that in the case $\alpha = 1/d$, the dimension of $\mathcal{L}$ is bounded below by d. To that end, we use the monotonicity of the Hausdorff dimension of $\mathcal{L}$, with respect to the scaling factors (λ_n), proved in Section 2.2.4. Suppose the sequences (λ_n) and (w_n) satisfy Hypothesis $\diamond_{\alpha,\beta}$ for $\alpha = 1/d$. If for some $\epsilon > 0$, we set $\lambda'_n = n^{-\epsilon}\lambda_n$ for all $n \geq 1$, then the sequences (λ'_n) and (w_n) satisfy Hypothesis $\diamond_{\alpha+\epsilon,\beta}$. Now for $n \geq 1$, we have $\lambda_n \geq \lambda'_n$, and $\mathcal{T}$ is compact with probability 1 from Proposition 2.6. Hence (2.6) holds and so we have, a.s.

$$\dim(\mathcal{L}) \geq \frac{1}{\alpha+\epsilon}.$$

In the end, $\dim(\mathcal{L}) \geq d$.

Proof of Theorem 2.1. Use Proposition 2.6, Proposition 2.7 for the upper-bounds and Proposition 2.12 and Proposition 2.14 for the lower-bounds. □

2.A Appendix

2.A.1 Lifting to the Urysohn space

In this section, we prove that it is always possible to work with random measured metric spaces that are embedded in the Urysohn space. Let us first recall the definition of the Gromov-Hausdorff-Prokhorov distance. If (X,d) is a metric space, and $A \subset X$ then we denote $A^{(\epsilon)} := \{x \in X \mid d(x,A) < \epsilon\}$, the ϵ-fattening of A. Then the Hausdorff distance $\mathrm{d_H}$ on the set of non-empty compact subsets of X, is defined as

$$\mathrm{d_H}(K,K') := \inf\left\{\epsilon > 0 \,\middle|\, K \subset (K')^{\epsilon},\ K' \subset (K)^{\epsilon}\right\}.$$

Also we denote the so-called Lévy-Prokhorov distance on the Borel probability measures by

$$\mathrm{d_{LP}}(\nu,\nu') := \inf\left\{\epsilon>0 \;\middle|\; \forall A\in\mathcal{B}(X),\ \nu(A)\le \nu'\left((A)^\epsilon\right)+\epsilon \text{ and } \nu'(A)\le \nu\left((A)^\epsilon\right)+\epsilon\right\},$$

where $\mathcal{B}(X)$ is the set of Borel sets of X. Now let (X,d,ρ,ν) and $(X',\mathrm{d}',\rho',\nu')$ be two compact, rooted, metric spaces endowed with a probability measure. Their Gromov-Hausdorff-Prokhorov distance is defined as

$$\mathrm{d_{GHP}}\left((X,\mathrm{d},\rho,\nu),(X',\mathrm{d}',\rho',\nu')\right) := \inf_{E,\phi,\phi'} \max(\mathrm{d}(\rho,\rho'),\mathrm{d_H}(\phi(X),\phi(X')),\mathrm{d_{LP}}(\phi_*\nu,\phi'_*\nu')),$$

where the infimum is taken over all Polish spaces (E,δ) and all isometric embeddings $\phi:\ X\to E$ and $\phi':\ X'\to E$, of respectively X and X' into E. The notation $\phi_*\nu$ denotes the push-forward of the measure μ through the map ϕ. As it is, this is only a pseudo-distance and it becomes a distance on the set $\mathbb{K}$ of GHP-isometry (root and measure preserving isometry) classes of compact, rooted, metric spaces endowed with a probability measure, which from [1, Theorem 2.5], is a Polish space. We consider all our blocks as (possibly random) elements of the set $\mathbb{K}$.

We would like to see all the blocks as compact subsets of the same space. To that end, we consider the Urysohn space (U,δ), and fix a point $u_0\in U$. The space U is defined as the only Polish metric space (up to isometry) which has the following extension property (see [78] for constructions and basic properties of U): given any finite metric space X, and any point $x\in X$, any isometry from $X\setminus\{x\}$ to U can be extended to an isometry from X to U. This property ensures in particular that any separable metric space can be isometrically embedded into U. In what follows we will use the fact that if (K,d,ρ) is a rooted compact metric space, there exists an isometric embedding of K to U such that ρ is mapped to u_0. We set

$$\mathbb{K}(U) := \{(K,\nu)\mid K\subset U,\ K \text{ compact}, u_0\in K,\ \nu \text{ is a Borel measure and } \operatorname{supp}(\nu)\subset K\},$$

where $\operatorname{supp}(\nu)$ denotes the topological support of ν. We endow $\mathbb{K}(U)$ with the "Hausdorff-Prokhorov" distance

$$\mathrm{d_{HP}}((K,\nu),(K',\nu')) = \max\left(\mathrm{d_H}(K,K'),\mathrm{d_{LP}}(\nu,\nu')\right).$$

It is easy to see that $(\mathbb{K}(U),\mathrm{d_{HP}})$ is a Polish space. Now, we have a map $f:\mathbb{K}(U)\to\mathbb{K}$, which maps every (K,ν) to the isometry class of $(K,\delta_{|_K},u_0,\nu)$ in $\mathbb{K}$. This map is continuous and hence measurable. The properties of U ensure that f is surjective. Using a theorem of measure theory from [90], every probability distribution τ on $\mathbb{K}$ can be lifted to a probability measure σ on $\mathbb{K}(U)$, such that $f_*\sigma=\tau$. Hence, for all $n\geq 1$, we can have a version of $(\mathbf{b}_n,\mathbf{d}_n,\boldsymbol{\rho}_n,\boldsymbol{\nu}_n)=(\mathsf{B}_n,\lambda_n\cdot\mathsf{D}_n,\rho_n,w_n\cdot\nu_n)$ that is embedded in the space U.

2.A.2 Hausdorff dimension

We recall some notations and definitions that are in relation with Hausdorff dimension and that we use throughout the paper. Let (X,d) be a metric space and $\delta>0$. We say that the family $(O_i)_{i\in I}$ of subsets of X is a δ-cover of X if it is a covering of X, and the set I is at most countable and for all $i\in I$, the set O_i is such that its diameter satisfies $\operatorname{diam}(O_i)<\delta$. We set

$$\mathcal{H}_s^\delta(X) := \inf\left\{\sum_{i\in I}\operatorname{diam}(O_i)^s \;\middle|\; (O_i)_{i\in I} \text{ is a } \delta\text{-cover of } X\right\},$$

As this quantity increases when δ decreases to 0, we define its limit

$$\mathcal{H}_s(X) := \lim_{\delta\to 0}\mathcal{H}_s^\delta(X)\in[0,\infty],$$

the s-*dimensional Hausdorff measure* of X. Now the Hausdorff dimension of X is defined as

$$\dim_{\mathrm{H}}(X) := \inf\{s > 0 \mid \mathcal{H}_s(X) = 0\} = \sup\{s > 0 \mid \mathcal{H}_s(X) = \infty\}.$$

We refer to [59] for details. A useful tool for deriving lower-bounds on the Hausdorff dimension of a metric space is the so-called Frostman's lemma. In this chapter we use the following version.

Lemma 2.20 (Frostman's lemma)**.** *Let (X, d) be a metric space. If there exists a non-zero finite Borel measure μ on X and $s > 0$ such that for μ-almost every $x \in X$, we have*

$$\mu\left(\mathrm{B}(x, r)\right) \underset{r\to 0}{\leq} r^{s+o(1)},$$

then

$$\dim_{\mathrm{H}}(X) \geq s.$$

2.A.3 Decomposition into *fragments*

In this section, we prove Lemma 2.17. We first construct our fragments in a deterministic setting and then show how we can apply this to random blocks.

Decomposition of a deterministic block Let $(\mathbf{b}, \mathbf{d}, \boldsymbol{\rho}, \boldsymbol{\nu})$ be a (deterministic) pointed compact metric space endowed with a Borel probability measure. We are interested in how we can decompose $\mathbf{b}$ into a partition of subsets that all have approximately the same diameter r. For $r > 0$, we set

$$\mathcal{P}_r(\mathbf{b}) := \left\{\{x_1, x_2, \dots, x_n\} \subset \mathbf{b} \;\middle|\; n \geq 1 \text{ and } \forall i \neq j,\ \mathbf{d}(x_i, x_j) \geq \frac{r}{2}\right\}.$$

It is easy to verify that we can find $p = \{x_1, x_2, \dots, x_n\} \in \mathcal{P}_r(\mathbf{b})$ such that $\mathbf{b} \subset \bigcup_{i=1}^n \mathrm{B}(x_i, r)$ and the balls $\left(\mathrm{B}(x_i, \frac{r}{4})\right)_{1\leq i\leq n}$ are disjoint. Indeed, any $\frac{r}{2}$-net of $\mathbf{b}$ belongs to $\mathcal{P}_r(\mathbf{b})$ (they are the maximal elements of $\mathcal{P}_r(\mathbf{b})$ for the order relation of inclusion). We denote by $\mathcal{P}_r^*(\mathbf{b})$ the set

$$\mathcal{P}_r^*(\mathbf{b}) := \left\{\{x_1, x_2, \dots, x_n\} \in \mathcal{P}_r(\mathbf{b}) \;\middle|\; \mathbf{b} \subset \bigcup_{i=1}^n \mathrm{B}(x_i, r)\right\},$$

which is non-empty from what precedes. Considering the collection of balls of radius r with centres in $p \in \mathcal{P}_r^*(\mathbf{b})$ gives rise to a covering of $\mathbf{b}$ which is close to optimal in a sense specified by the following lemma. We recall the notation $N_r(\mathbf{b})$ which denotes the minimal number of balls of radius r needed to cover $\mathbf{b}$.

Lemma 2.21. *For any $p = \{x_1, x_2, \dots, x_n\} \in \mathcal{P}_r^*(\mathbf{b})$, we have $n \leq N_{r/4}(\mathbf{b})$.*

Proof. Let $p = \{x_1, x_2, \dots, x_n\} \in \mathcal{P}_r^*(\mathbf{b})$. Remark that, for any set S such that the union of the balls $\left(\mathrm{B}(s, \frac{r}{4})\right)_{s\in S}$ covers $\mathbf{b}$, each of the balls $\mathrm{B}(x_i, r/4)$, for $1 \leq i \leq n$, contains at least a point of S. Since those balls are disjoint, the cardinality of S is at least n. ☐

From any element $p = \{x_1, x_2, \dots, x_n\} \in \mathcal{P}_r^*(\mathbf{b})$, we can then construct a partition of $\mathbf{b}$, into subsets $(f_i)_{1\leq i\leq n}$ that we call fragments, and such that

$$\forall i \in [\![1\,, n]\!], \quad \mathrm{B}\left(x_i, \frac{r}{4}\right) \subset f_i \subset \mathrm{B}(x_i, r).$$

We define the (f_i) recursively as

$$\begin{cases} f_1 := \left\{ x \in \mathbf{b} \;\middle|\; \mathbf{d}(x, x_1) = \min_{1\le i\le n} \mathbf{d}(x, x_i) \right\} \\ f_{k+1} := \left\{ x \in \left(\mathbf{b} \setminus \bigcup_{i=1}^{k} f_i \right) \;\middle|\; \mathbf{d}(x, x_{k+1}) = \min_{1\le i\le n} \mathbf{d}(x, x_i) \right\}. \end{cases}$$

If we suppose that $\mathbf{b}$ satisfies the condition $(\star_{r_0})$, and that $r < r_0$ then, for $p = \{x_1, x_2, \dots, x_n\} \in \mathcal{P}_r^*(\mathbf{b})$, we get

$$n \left(\frac{r}{4}\right)^{d+\varphi(r)} \le \sum_{i=1}^{n} \boldsymbol{\nu}\left(\mathrm{B}\left(x_i, \frac{r}{4}\right)\right) \le \boldsymbol{\nu}(\mathbf{b}) = 1,$$

so that we have

$$n \le \left(\frac{r}{4}\right)^{-d-\varphi(r/4)}, \tag{2.32}$$

and also, for all $i \in [\![1\,, n]\!]$,

$$\operatorname{diam} f_i \le 2r \qquad \text{and} \qquad \left(\frac{r}{4}\right)^{d+\varphi(r/4)} \le \boldsymbol{\nu}(f_i) \le r^{d-\varphi(r)}. \tag{2.33}$$

Let us state another lemma.

Lemma 2.22. *Let $r_0 < 1$. Under the condition $(\star_{r_0})$, there exists two functions ψ and ϕ defined on $[0\,, r_0/3]$, which tend to 0 at 0 such that for all $r \in (0\,, r_0/3)$, for all $p = \{x_1, x_2, \dots, x_n\} \in \mathcal{P}_r^*(\mathbf{b})$ and fragments (f_i) constructed as above, we have*

$$\forall x \in \mathbf{b}, \forall r' \in (0\,, r_1), \quad \#\left\{1 \le i \le n \;\middle|\; \mathrm{B}(x, r') \cap f_i \ne \emptyset\right\} \le (r \vee r')^{d+\psi(r\vee r')} \cdot r^{-d+\phi(r)}.$$

Proof. Let $x \in \mathbf{b}$. If for an $i \in [\![1\,, n]\!]$, we have $y \in \mathrm{B}(x, r') \cap f_i \ne \emptyset$, then $\mathbf{d}(x, y) < r'$ and $\mathbf{d}(x_i, y) < r$ so $\mathbf{d}(x, x_i) < r + r'$, and so we get that $\mathrm{B}(x_i, r) \subset \mathrm{B}(x, r' + 2r)$. Then, using that $f_i \subset \mathrm{B}(x_i, r)$,

$$\left(\bigcup_{i:\ f_i \cap \mathrm{B}(x,r') \ne \emptyset} \mathrm{B}\left(x_i, \frac{r}{4}\right) \right) \subset \left(\bigcup_{i:\ f_i \cap \mathrm{B}(x,r') \ne \emptyset} f_i \right) \subset \mathrm{B}(x, r' + 2r).$$

We can use the condition $(\star_{r_0})$ to get, for all $r,\ r' \in [0\,, r_0/3]$,

$$\#\left\{1 \le i \le n \;\middle|\; \mathrm{B}(x, r') \cap f_i \ne \emptyset\right\} \cdot \left(\frac{r}{4}\right)^{d+\varphi(r/4)} \le (r' + 2r)^{d-\varphi(r'+2r)}.$$

And so,

$$\begin{aligned} \#\left\{1 \le i \le n \;\middle|\; \mathrm{B}(x, r') \cap f_i \ne \emptyset\right\} &\le 4^{d+\varphi(r/4)} \frac{(r' + 2r)^{d-\varphi(r'+2r)}}{r^{d+\varphi(r/4)}} \\ &\le 4^{d+\varphi(r/4)} \frac{(3(r \vee r'))^{d-\varphi(r'+2r)}}{r^{d+\varphi(r/4)}} \\ &\le 4^{d+\varphi(r/4)} 3^{d-\varphi(r'+2r)} (r \vee r')^{d-\varphi(r'+2r)} r^{-d-\varphi(r/4)} \\ &\le (r \vee r')^{d-\varphi(3(r\vee r'))} r^{-d-\varphi(r/4) + \frac{\log\left(12^{3d/2}\right)}{\log r}}, \end{aligned}$$

which proves the lemma. ∎

This lemma gives us an abstract result for the existence and the properties of these decompositions in fragments. The next paragraph explains a procedure to construct one using a sequence of i.i.d. random points, on a possibly random block.

Finding an element of $\mathcal{P}_r^*(\mathbf{b})$ Suppose that the measure $\boldsymbol{\nu}$ charges all open sets. Remark that this is almost surely true for our random block $(\mathsf{B},\mathsf{D},\rho,\nu)$ because it satisfies Hypothesis H_d(i). Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. random variables with law $\boldsymbol{\nu}$. Let us construct a random element of $\mathcal{P}_r^*(\mathbf{b})$ for some fixed $r>0$. Define the set E_n recursively as

$$E_1 := \{1\} \quad \text{and} \quad \begin{cases} E_{n+1} := E_n \quad \text{if} \quad X_{n+1} \in \bigcup_{i\in E_n} \mathrm{B}\left(X_i, \frac{r}{2}\right), \\ E_{n+1} := E_n \cup \{n+1\} \quad \text{otherwise.} \end{cases}$$

We set $E_\infty = \bigcup_{n\geq 1} E_n$. Note that from the construction, $\{X_i,\ i\in E_\infty\}\in\mathcal{P}_r(\mathbf{b})$.

Lemma 2.23. *Almost surely, we have $\{X_i,\ i\in E_\infty\}\in\mathcal{P}_r^*(\mathbf{b})$.*

Proof. The fact that the balls $\left(\mathrm{B}(X_i,\frac{r}{4})\right)_{i\in E_\infty}$ are disjoint follows directly from the construction. Now let $x\in\mathbf{b}$. Since $\boldsymbol{\nu}\left(\mathrm{B}\left(x,\frac{r}{4}\right)\right)>0$, by the Borel-Cantelli lemma there exists at least one n such that $X_n\in\mathrm{B}\left(x,\frac{r}{4}\right)$. If $n\in E_\infty$ then $x\in\mathrm{B}\left(X_n,\frac{r}{4}\right)$. Otherwise $n\notin E_\infty$, in which case there exists $k\leq n$ such that $X_n\in\mathrm{B}\left(X_k,\frac{r}{2}\right)$ and so $x\in\mathrm{B}\left(X_k,\frac{3}{4}r\right)$. In both cases

$$\mathrm{B}\left(x,\frac{r}{4}\right)\subset\bigcup_{i\in E_\infty}\mathrm{B}(X_i,r).$$

Since we can apply the same reasoning to every point of a dense sequence $(y_k)_{k\geq 1}$, the lemma is proved. ☐

Proof of Lemma 2.17. This is just a consequence of Lemma 2.21, Lemma 2.22 and Lemma 2.23 and equations (2.32) and (2.33), which almost surely apply to the random block $(\mathsf{B},\mathsf{D},\rho,\nu)$. ☐

2.A.4 Computations

Lemma 2.24. *Suppose that there exists $\gamma\geq 0$ such that for all $n\in\mathbb{N}$, $W_n\leq n^\gamma$. Then there exists a constant C such that*

$$\sum_{k=1}^{n}\frac{w_k}{W_k}\leq C\log n.$$

Proof. If the series $\sum w_k$ converges then the result is trivial so let us suppose that it diverges. For $k\geq 0$, we define $n_k := \inf\left\{i\geq 1 \,\middle|\, W_i\geq 2^k\right\}$ and write

$$\sum_{k=n_0}^{n}\frac{w_k}{W_k}\leq\sum_{i=0}^{\left\lceil\frac{\log W_n}{\log 2}\right\rceil}\sum_{k=n_i}^{n_{i+1}-1}\frac{w_k}{W_k}\leq\sum_{i=0}^{\left\lceil\frac{\log W_n}{\log 2}\right\rceil}\frac{1}{2^i}\sum_{k=n_i}^{n_{i+1}-1}w_k\leq\sum_{i=0}^{\left\lceil\frac{\log W_n}{\log 2}\right\rceil}\frac{2^{i+1}}{2^i}$$
$$\leq 2\left\lceil\frac{\log W_n}{\log 2}\right\rceil,$$

which grows at most logarithmically thanks to our assumption on the sequence (W_n). ☐

Lemma 2.25. *Let $\beta<1$ and assume that $w_n\leq n^{-\beta+o(1)}$ and $W_n=n^{1-\beta+o(1)}$ and that, for some $\epsilon>0$,*

$$\liminf_{n\to\infty}\frac{1}{W_n}\sum_{\substack{k=1\\k\in G^\epsilon}}^{n}w_k>0.$$

Then there exists a constant C_ϵ such that for N large enough we have

$$\sum_{\substack{k=N\\k\in G^\epsilon}}^{N^{1+\epsilon}}\frac{w_k}{W_k}\geq C_\epsilon\log N.$$

Proof. Let c be such that, for n large enough $\frac{1}{W_n}\sum_{k=1}^n w_k \mathbf{1}_{\{k\in G^\epsilon\}} > c$. Let $C := \frac{3}{c}$, note that $C > 1$ because $c \le 1$. For all $i \ge 1$, we set $k_i = \inf\left\{n \,\middle|\, W_n \ge C^i\right\}$. For all $i \ge 1$, we have $W_{k_i-1} \le C^i \le W_{k_i} \le C^i + w_{k_i}$. We get,

$$\begin{aligned}\sum_{k=k_i+1}^{k_{i+1}} w_k \mathbf{1}_{\{k\in G^\epsilon\}} \ge cW_{k_{i+1}} - W_{k_i} \ge cC^{i+1} - C^i - w_{k_i} &\ge C^i(cC - 1 - \frac{w_{k_i}}{C^i}) \\ &\ge C^i(1+o(1)),\end{aligned}$$

for i tending to infinity. Now for N a large integer, we set

$$I_N := \inf\{i \mid k_i \ge N\} = \left\lceil \frac{\log W_N}{\log C} \right\rceil \quad \text{and} \quad J_N := \sup\left\{i \,\middle|\, k_i \le N^{1+\epsilon}\right\} = \left\lfloor \frac{\log W_{\lfloor N^{1+\epsilon}\rfloor}}{\log C} \right\rfloor.$$

Then we compute

$$\begin{aligned}\sum_{k=N}^{N^{1+\epsilon}} \frac{w_k}{W_k}\mathbf{1}_{\{k\in G^\epsilon\}} &\ge \sum_{i=I_N}^{J_N-1}\sum_{k=k_i+1}^{k_{i+1}} \frac{w_k}{W_k}\mathbf{1}_{\{k\in G^\epsilon\}} \\ &\ge \sum_{i=I_N}^{J_N-1} \frac{1}{W_{k_{i+1}}} \sum_{k=k_i+1}^{k_{i+1}} w_k \mathbf{1}_{\{k\in G^\epsilon\}} \\ &\ge \sum_{i=I_N}^{J_N-1} \frac{1}{C^{i+1}(1+o(1))} C^i(1+o(1)) \\ &\ge \frac{J_N - I_N - 1}{C}(1+o(1)).\end{aligned}$$

We finish the proof by noting that, thanks to the hypothesis on the growth of W_n, the last display grows logarithmically in N. $\square$

Proof of Lemma 2.16. From our assumptions, it is easy to see that we have $a_k = n_k^{1-\beta+o(1)}$. For all k, we write

$$\log a_k = (1 - \beta + r_k)\log n_k,$$

with $r_k \to 0$ as $k \to \infty$. We write

$$\sum_{i=1}^k \log a_i = \sum_{i=1}^k (1-\beta+r_i)\log n_i = \sum_{i=0}^{k-1}(1-\beta+r_{k-i})\log n_{k-i}. \tag{2.34}$$

For any $k \ge 0$, from the recursive definition of the sequence (n_k), we have $n_{k+1} - 1 < n_k^\gamma \le n_{k+1}$, which entails $\log n_k = \frac{1}{\gamma}\log n_{k+1} + s_k$, with $|s_k| \le 1$. Using this recursively yields

$$\left|\log n_{k-i} - \frac{1}{\gamma^i}\log n_k\right| \le \frac{\gamma}{1-\gamma}.$$

Hence using (2.34) and the fact that $\log n_k$ grows exponentially in k,

$$\begin{aligned}\sum_{i=1}^k \log a_i &= \log n_k \left((1-\beta)\sum_{i=0}^{k-1}\frac{1}{\gamma^i} + \underbrace{\sum_{i=0}^{k}\frac{r_{k-i}}{\gamma^i}}_{\to 0} \right) + \underbrace{\sum_{i=0}^{k-1}(1-\beta+r_{k-i})\left(\log n_{k-i} - \frac{1}{\gamma^i}\log n_k\right)}_{=O(k)} \\ &= \log n_k \left(\frac{(1-\beta)\gamma}{\gamma-1} + o(1) \right),\end{aligned}$$

which proves the lemma. $\square$

Chapter 3

Geometry of weighted recursive and affine preferential attachment trees

This chapter is adapted from the work [110], in preparation.

We study two models of growing recursive trees. For both models, initially the tree only contains one vertex u_1 and at each time $n \geq 2$ a new vertex u_n is added to the tree and its parent is chosen randomly according to some rule. In the *weighted recursive tree*, we choose the parent u_k of u_n among $\{u_1, u_2, \dots, u_{n-1}\}$ with probability proportional to w_k, where $(w_n)_{n\geq 1}$ is some deterministic sequence that we fix beforehand. In the *affine preferential attachment tree with initial fitnesses*, the probability of choosing the same u_k is proportional to $a_k + \deg^+(u_k)$, where $\deg^+(u_k)$ denotes its current number of children, and the sequence of *initial fitnesses* $(a_n)_{n\geq 1}$ is deterministic and chosen as a parameter of the model.

We show that for any sequence $(a_n)_{n\geq 1}$, the corresponding preferential attachment tree has the same distribution as some weighted recursive tree with a *random* sequence of weights (with some explicit distribution). We prove almost sure convergences for some statistics associated to weighted recursive trees as time goes to infinity, such as degree sequence, height, profile and measures. Thanks to the connection between the two models, these results also apply to affine preferential attachment trees.

Contents

3.1 Introduction

The uniform recursive tree has been introduced in the 70's as an example of random graphs constructed by addition of vertices: starting from a tree with a single vertex, the vertices arrive one by one and the n-th vertex picks its parent uniformly at random from the $n-1$ already present vertices. Many properties of this tree were then investigated due to its particularly simple dynamics: number of leaves, profile, height, degrees, distribution of vertices into subtrees... We refer to [50] for an overview. A generalisation of the uniform recursive, the weighted recursive tree (WRT), was introduced in [30] in 2006. In this model, each vertex is assigned a non-negative weight, constant in time. When a newcomer randomly picks its parents, it does so with probability proportional to those weights. Although more general than the uniform recursive tree, WRT have attracted far fewer contributions, see e.g. [91, 74].

We also study another model of trees which we call the affine preferential attachment tree (PA) with initial fitnesses. In this tree every vertex has a fixed initial fitness, and the probability of picking any vertex to be the parent of a newcomer is proportional to its initial fitness plus its current number of children. This type of preferential attachment mechanism has been extensively studied in the last two decades because it shares some quantitative properties with real networks, see in particular the literature about Barabási-Albert model. Our motivation for studying such trees arises from the analysis of some growing random graphs, see the companion paper [111].

We shall see that using a de Finetti-type argument, preferential attachment trees can be seen as WRT with random weights. This will enable us to translate results obtained for WRT to corresponding results for PA.

3.1.1 Two related models of growing trees

Definitions. For any sequence of non-negative real numbers $(w_n)_{n\geq 1}$ with $w_1 > 0$, we define the distribution $\mathrm{WRT}((w_n)_{n\geq1})$ on sequences of growing rooted trees[1], which is called the *weighted recursive tree with weights* $(w_n)_{n\geq1}$. We construct a sequence of rooted trees $(\mathtt{T}_n)_{n\geq1}$ starting from $\mathtt{T}_1$ containing only one root-vertex u_1 and let it evolve in the following manner: the tree $\mathtt{T}_{n+1}$ is obtained from $\mathtt{T}_n$ by adding a vertex u_{n+1} with label $n+1$. The father of this new vertex is chosen to be the vertex with label K_{n+1}, where

$$\forall k \in \{1, \dots, n\}, \qquad \mathbb{P}\left(K_{n+1} = k \mid \mathtt{T}_n\right) \propto w_k.$$

In this definition, we also allow sequences of weights $(\mathsf{w}_n)_{n\geq1}$ that are random and in this case the distribution $\mathrm{WRT}((\mathsf{w}_n)_{n\geq1})$ denotes the law of the random sequence of trees obtained by the above process conditionally on $(\mathsf{w}_n)_{n\geq1}$, so that the obtained distribution is a mixture of WRT with deterministic sequence.

Similarly, for any sequence $(a_n)_{n\geq1}$ of real numbers, with $a_1 > -1$ and $a_n \geq 0$ for $n \geq 2$, we define another model of growing trees. The construction goes on as before: $\mathtt{P}_1$ containing only one root-vertex u_1 and $\mathtt{P}_{n+1}$ is obtained from $\mathtt{P}_n$ by adding a vertex u_{n+1} with label $n+1$ and the father of the newcomer is chosen to be the vertex with label J_{n+1}, where now

$$\forall k \in \{1, \dots, n\}, \qquad \mathbb{P}\left(J_{n+1} = k \mid \mathtt{P}_n\right) \propto \deg^+_{\mathtt{P}_n}(u_k) + a_k,$$

where $\deg^+_{\mathtt{P}_n}(\cdot)$ denotes the number of children in the tree $\mathtt{P}_n$. In the particular case where $n=1$, the second vertex u_2 is always defined as a child of u_1, even in the case $-1 < a_1 \leq 0$ for which the

[1]In fact, in the rest of the chapter we will see them as plane trees, see Section 3.1.2.

last display does not make sense. We call this sequence of trees an *affine preferential attachment tree with initial fitnesses* $(a_n)_{n\geq 1}$ and its law is denoted by $\mathrm{PA}((a_n)_{n\geq 1})$.

Here and in the rest of the paper, whenever we have any sequence of real numbers $(x_n)_{n\geq 1}$, we write $\mathbf{x} = (x_n)_{n\geq 1}$ in a bold font as a shorthand for the sequence itself, and $(X_n)_{n\geq 1}$ with a capital letter to denote the sequence of partial sums defined for all $n \geq 1$ as $X_n := \sum_{i=1}^{n} x_i$. In particular, we do so for sequences of initial fitnesses $(a_n)_{n\geq 1}$, for deterministic sequences of weights $(w_n)_{n\geq 1}$ and for random sequence of weights $(\mathsf{w}_n)_{n\geq 1}$.

Representation result. The following result gives a connection between these two models of growing trees. It is an analogue of the so-called "Pólya urn-representation" result described in [17, Theorem 2.1] or [28, Section 1.2] for related models.

Theorem 3.1 (WRT-representation of PA trees). *For any sequence* $\mathbf{a}$ *of initial fitnesses, we define the associated random sequence* $\mathsf{w}^{\mathbf{a}} = (\mathsf{w}_n^{\mathbf{a}})_{n\geq 1}$ *as*

$$\mathsf{w}_1^{\mathbf{a}} = \mathsf{W}_1^{\mathbf{a}} = 1 \qquad \text{and} \qquad \forall n \geq 2, \quad \mathsf{W}_n^{\mathbf{a}} = \prod_{k=1}^{n-1} \beta_k^{-1}, \tag{3.1}$$

where the $(\beta_k)_{k\geq 1}$ *are independent with respective distribution* $\mathrm{Beta}(A_k + k, a_{k+1})$*. Then, the distributions* $\mathrm{PA}(\mathbf{a})$ *and* $\mathrm{WRT}(\mathsf{w}^{\mathbf{a}})$ *coincide.*

Let us explain how this sequence $\mathsf{w}^{\mathbf{a}}$ can be read from the growth of the tree. For any sequence of weights $\mathbf{w}$ that satisfies

$$W_n \underset{n\to\infty}{\sim} C \cdot n^{\gamma}, \tag{3.2}$$

for some $\gamma \in (0\,,1)$ and a positive $C > 0$, it is easy to prove that the degree of vertices in a sequence of random trees $(\mathtt{T}_n)_{n\geq 1}$ with distribution $\mathrm{WRT}(\mathbf{w})$ are such that almost surely for all $k \geq 1$

$$\deg^+_{\mathtt{T}_n}(u_k) \underset{n\to\infty}{\sim} \frac{w_k}{C(1-\gamma)} \cdot n^{1-\gamma}. \tag{3.3}$$

From this observation, if the sequence $\mathsf{w}^{\mathbf{a}}$ has almost surely the behaviour (3.2), then we can retrieve it from the behaviour of degrees in the tree by taking the limit for all $k \geq 1$

$$\mathsf{w}_k^{\mathbf{a}} = \frac{\mathsf{w}_k^{\mathbf{a}}}{\mathsf{w}_1^{\mathbf{a}}} = \lim_{n\to\infty} \frac{\deg^+_{\mathtt{P}_n}(u_k)}{\deg^+_{\mathtt{P}_n}(u_1)} \quad \text{almost surely.}$$

As suggested by the last display, the result of the theorem is obtained by studying the evolution of the degrees in the preferential attachment model $(\mathtt{P}_n)_{n\geq 1}$. The key argument lies in the fact that we can describe the whole process using a sequence of *independent* Pólya urns, related to the degrees of the vertices. The theorem is then obtained by using de Finetti theorem for these urns.

In fact, and this is the content of Proposition 3.2 below, if A_n grows linearly as some $c \cdot n$ with some $c > 0$ then the sequence $(\mathsf{W}_n^{\mathbf{a}})$ indeed almost surely satisfies (3.2) for $\gamma = \frac{c}{c+1}$. This is done using moment computation under the explicit definition of $(\mathsf{W}_n^{\mathbf{a}})_{n\geq 1}$ given by the theorem. In the rest of this chapter, we also investigate other properties of the WRT under this type of assumptions for the sequence of weights, such as convergence of height, profile and measures carried on the tree. Thanks to this connection, our results will then also hold for PA trees under the assumption that A_n grows linearly.

Assumptions on the sequences. For two sequences (x_n) and (y_n) we say that

$$x_n \underset{n\to\infty}{\bowtie} y_n \qquad \text{if and only if} \qquad \exists \epsilon > 0,\ x_n \underset{n\to\infty}{=} y_n \cdot (1 + O(n^{-\epsilon})). \tag{3.4}$$

Our main assumption for sequences $\mathbf{a} = (a_n)_{n\geq 1}$ of initial fitnesses is the following (H_c), which is parametrised by some positive $c > 0$ and ensures that the initial fitness of vertices is c on average,i.e.,

$$A_n \underset{n\to\infty}{\bowtie} c \cdot n. \tag{H_c}$$

For sequences of weights $\mathbf{w} = (w_n)_{n\geq 1}$, we introduce the following hypothesis, which depends on a parameter $\gamma > 0$

$$W_n \underset{n\to\infty}{\bowtie} \mathrm{cst} \cdot n^{\gamma}. \tag{$\square_\gamma$}$$

The following proposition ensures in particular that our assumptions on sequences of initial fitnesses $\mathbf{a}$ translate to a power behaviour for the random sequence of cumulated weights $(\mathsf{W}_n^{\mathbf{a}})_{n\geq 1}$ defined in Theorem 3.1.

Proposition 3.2. *Suppose that there exists $c > 0$ such that $\mathbf{a}$ satisfies (H_c), then the random sequence $(\mathsf{w}_n^{\mathbf{a}})_{n\geq 1}$ defined in Theorem 3.1 almost surely satisfies ($\square_\gamma$) with*

$$\gamma = \frac{c}{c+1}.$$

If furthermore $\mathbf{a}$ is such that $a_n \leq (n+1)^{c'+o(1)}$ for some $c' \in [0,1)$, then almost surely $\mathsf{w}_n^{\mathbf{a}} \leq (n+1)^{c'-\frac{1}{c+1}+o_\omega(1)}$, where $o_\omega(1)$ is a random function of n which tends to 0 when $n \to \infty$.

Convergence of degrees using the WRT representation In the WRT with a deterministic sequence of weights that satisfy (3.2), the degree of one fixed vertex evolves as a sum of independent Bernoulli random variable and it is possible to handle it with elementary methods and obtain (3.3). Further calculations allow us to improve this statement to a convergence

$$n^{-(1-\gamma)} \cdot (\deg^+_{\mathtt{T}_n}(u_1), \deg^+_{\mathtt{T}_n}(u_2), \dots) \to \frac{1}{C(1-\gamma)}(w_1, w_2, \dots) \tag{3.5}$$

in an ℓ^p sense, for sequences $\mathbf{w}$ that satisfy some additional control. This is proved in Proposition 3.5.

Suppose that $\mathbf{a}$ satisfies (H_c). Applying this convergence to sequence of random trees $(\mathsf{P}_n)_{n\geq 1}$ which has distribution $\mathrm{PA}(\mathbf{a})$, using its WRT-representation provided by Theorem 3.1, together with Proposition 3.2, yields the following almost sure convergence to a random sequence, in the product topology,

$$n^{-\frac{1}{c+1}} \cdot (\deg^+_{\mathsf{P}_n}(u_1), \deg^+_{\mathsf{P}_n}(u_2), \dots) \underset{n\to\infty}{\longrightarrow} (\mathsf{m}_1^{\mathbf{a}}, \mathsf{m}_2^{\mathbf{a}}, \dots),$$

which also takes place in the space ℓ^p for all $p > \frac{c+1}{1-(c+1)c'}$ as soon as $a_n \leq n^{c'+o(1)}$, for some $0 \leq c' < \frac{1}{c+1}$. This improves some ℓ^p convergence proved in distribution in [100] for a related model, which we treat in Proposition 3.31.

Of course, thanks to our discussion above concerning the convergence of degrees, it is immediate that the sequence $(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1}$ is almost surely proportional to the sequence $(\mathsf{w}_n^{\mathbf{a}})_{n\geq 1}$ i.e.

$$(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1} = \frac{c+1}{Z} \cdot (\mathsf{w}_n^{\mathbf{a}})_{n\geq 1} \quad \text{a.s.},$$

where Z is the random variable such that $\mathsf{W}_n^{\mathbf{a}} \sim Z \cdot n^{\frac{c}{c+1}}$ almost surely as $n \to \infty$, which exists thanks to Proposition 3.2. Of course, even if $(\mathsf{W}_n^{\mathbf{a}})_{n\geq 1}$ was defined as a product of independent random variables, it is not the case for the sequence $(\mathsf{M}_n^{\mathbf{a}})_{n\geq 1}$ (defined as the sequence of partial sums of $(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1}$) since the random variable Z depends on the whole sequence $(\beta_n)_{n\geq 1}$ used in the definition of $(\mathsf{W}_n^{\mathbf{a}})_{n\geq 1}$. Nevertheless, the sequence still has the nice property of being an inhomogeneous Markov chain with a simple *backward* transition, characterised by the equality

$$\mathsf{M}_n^{\mathbf{a}} = \beta_n \cdot \mathsf{M}_{n+1}^{\mathbf{a}},$$

where β_n is independent of $\mathsf{M}_{n+1}^{\mathbf{a}}$ and has a $\mathrm{Beta}(A_n + n, a_{n+1})$ distribution. This is the content of Proposition 3.27.

Distribution of the limiting chain. For some specific choices of sequences $\mathbf{a}$, the distribution of the chain $(\mathsf{M}_n^{\mathbf{a}})_{n\geq 1}$ is explicit. Whenever $\mathbf{a}$ is of the form

$$\mathbf{a} = (a, b, b, b, \dots) \qquad \text{with } a > -1 \text{ and } b > 0,$$

we retrieve Goldschmidt and Haas' Mittag-Leffler Markov chain family, introduced in [63] and also studied by James [79]. The other case where the chain is explicit is when $\mathbf{a}$ is of the form

$$\mathbf{a} = (a, \underbrace{0, 0, \dots, 0}_{\ell-1}, m, \underbrace{0, 0, \dots, 0}_{\ell-1}, m, \dots) \qquad \text{with } a > -1 \text{ and } \ell, m \in \mathbb{N}.$$

In this case, the process $(\mathsf{M}_n^{\mathbf{a}})_{n\geq 1}$ is constant on the interval of the form $[\![1 + k\ell\,, (k+1)\ell]\!]$ and we define $\mathsf{N}_k^{\mathbf{a}} := \mathsf{M}_{(k-1)\ell+1}^{\mathbf{a}}$ for all $k \geq 1$. Then the sequence $\frac{\ell^{\frac{\ell}{m+\ell}}}{m+\ell} \cdot (\mathsf{N}_k^{\mathbf{a}})_{k\geq 1}$ has the *Product Generalised Gamma* distribution $\mathrm{PGG}\,(a, \ell, m)$, which we define in Section 3.5.1.

3.1.2 Other geometric properties of weighted random trees

Let us now state the convergence for other statistics of weighted random trees, namely profile, height and probability measures. Here we let $(\mathtt{T}_n)_{n\geq 1}$ be a sequence of trees evolving according to the distribution $\mathrm{WRT}(\mathbf{w})$ for some deterministic sequence $\mathbf{w}$ and state our results in this setting. Our results will also apply to random sequences of weights $\mathbf{w}$ that satisfy the assumptions of the theorems almost surely, they will hence apply to PA trees with appropriate sequences of initial fitnesses, thanks to Theorem 3.1 and Proposition 3.2.

Height and profile of WRT

Let

$$\mathbb{L}_n(k) := \#\left\{1 \leq i \leq n \mid \mathrm{ht}(u_i) = k\right\}$$

be the number of vertices of $\mathtt{T}_n$ at height k. The function $k \mapsto \mathbb{L}_n(k)$ is called the *profile* of the tree $\mathtt{T}_n$. The height of the tree is the maximal distance of a vertex to the root, which we can also express as $\mathrm{ht}(\mathtt{T}_n) := \max\left\{k \geq 0 \mid \mathbb{L}_n(k) > 0\right\}$. We are interested in the asymptotic behaviour of $\mathbb{L}_n$ and $\mathrm{ht}(\mathtt{T}_n)$ as $n \to \infty$.

In order to express our results, we need to introduce some quantities. For $\gamma > 0$, we define the function $f_\gamma : \mathbb{R} \to \mathbb{R}$ as

$$f_\gamma : z \mapsto f_\gamma(z) := 1 + \gamma\left(e^z - 1 - ze^z\right).$$

This function is increasing on $(-\infty\,,0]$ and decreasing on $[0\,,\infty)$ with $f_\gamma(-\infty) = 1-\gamma$ and $f_\gamma(0) = 1$ and $f_\gamma(\infty) = -\infty$. We define z_+ and z_- as

$$z_+ := \sup\{z\in\mathbb{R} \mid f_\gamma(z) > 0\} \quad \text{and} \quad z_- := \inf\{z\in\mathbb{R} \mid 1+\gamma(e^z-1) > 0\}. \tag{3.6}$$

We are going to assume that we work with a sequence $\mathbf{w}$ which satisfies the following assumption ($\square_\gamma^p$) for some $\gamma > 0$ and $p\in(1\,,2]$,

$$W_n \underset{n\to\infty}{\bowtie} \mathrm{cst}\cdot n^\gamma \qquad \text{and} \qquad \sum_{i=n}^{2n} w_i^p \le n^{1+(\gamma-1)p+o(1)}. \tag{$\square_\gamma^p$}$$

Thanks to Proposition 3.2, this property is almost surely satisfied for $\gamma = \frac{c}{c+1}$ by the random sequence $\mathbf{w}^{\mathbf{a}}$ for any sequence $\mathbf{a}$ of initial fitnesses satisfying $A_n \underset{n\to\infty}{\bowtie} c\cdot n$ and $a_n \le (n+1)^{o(1)}$.

Theorem 3.3. *Suppose that there exists $\gamma > 0$ and $p \in (1\,,2]$ such that the sequence $\mathbf{w}$ satisfies ($\square_\gamma^p$). Then, for a sequence of random trees $(\mathtt{T}_n)_{n\ge1} \sim \mathrm{WRT}(\mathbf{w})$, we have the almost sure asymptotics for the profile*

$$\mathbb{L}_n(k) \underset{n\to\infty}{=} \frac{n}{\sqrt{2\pi\gamma\log n}}\exp\left\{-\frac{1}{2}\cdot\left(\frac{k-\gamma\log n}{\sqrt{\gamma\log n}}\right)^2\right\} + O\left(\frac{n}{\log n}\right), \tag{3.7}$$

where the error term is uniform in $k\ge 0$. Also for any compact $K\subset(z_-\,,z_+)$ we have almost surely for all $z\in K$

$$\mathbb{L}_n\left(\lfloor\gamma e^z\log n\rfloor\right) = n^{f_\gamma(z)-\frac{1}{2}\frac{\log\log n}{\log n}+O\left(\frac{1}{\log n}\right)}, \tag{3.8}$$

where the error term is uniform in $z\in K$. Moreover, we have the almost sure convergence

$$\frac{\mathrm{ht}(\mathtt{T}_n)}{\log n} \underset{n\to\infty}{\longrightarrow} \gamma\cdot e^{z_+}. \tag{3.9}$$

The proof of this result follows the path used for many similar results for trees with logarithmic growth (see [35, 36, 85]): we study the Laplace transform of the profile $z\mapsto\sum_{k=0}^n e^{zk}\mathbb{L}_n(k)$ on an open domain of the complex plane and prove its convergence to some random analytic function when appropriately rescaled. Then, we apply [84, Theorem 2.1], which consists in a fine Fourier inversion argument and hence allows to obtain precise asymptotics for $\mathbb{L}_n$. The application of the theorem in its full generality proves a so-called Edgeworth expansion for $\mathbb{L}_n$, which we express here in a weaker form by equations (3.7) and (3.8). The convergence (3.7) expresses that the profile is asymptotically close to a Gaussian shape centred around $\gamma\log n$ and with variance $\gamma\log n$, so that a majority of vertices have a height of order $\gamma\log n$. The second equation (3.8) provides the behaviour of the number of vertices at a given height, for heights that are not necessarily close to $\gamma\log n$ (for which the preceding result ensures that there are of order $\frac{n}{\sqrt{\log n}}$ vertices per level). According to this result, at height $\lfloor\gamma e^z\log n\rfloor$ for any $z\in(z_-\,,z_+)$ there are of order $\frac{n^{f_\gamma(z)}}{\sqrt{\log n}}$ vertices. Remark that the exponent $f_\gamma(z)$ is continuous in z and tends to 0 when $z\to z_+$. Although this does not directly prove the convergence (3.9), it already provides a lower-bound for $\mathrm{ht}(\mathtt{T}_n)$ since it ensures that asymptotically there always exist vertices at height $\lfloor\gamma e^{(z_+-\epsilon)}\log n\rfloor$, for any small $\epsilon > 0$. The convergence of the height (3.9) can then be obtained by proving a corresponding upper-bound, which can be done using quite rough estimates.

This result includes the well-known asymptotics $\mathrm{ht}(\mathtt{T}_n)\sim e\log n$ as $n\to\infty$ for the uniform random tree, proved for example in [45, 105]. Using the connection of preferential attachment trees to weighted recursive trees given by Theorem 3.1, it also includes the case of preferential attachment trees with constant initial fitnesses, for which similar results were proved, in [105] for the height and in [85] for the asymptotic behaviour of the profile (3.7).

As a complement to this result, let us mention that there is another case when we can compute the asymptotic height of the tree, which corresponds to sequences $\mathbf{w}$ that grow fast to infinity. For any sequence of weights $\mathbf{w}$, a quantity of interest is $\sum_{i=2}^{n} \frac{w_i}{W_i}$, which is the expected height of a "typical" point. When this quantity grows faster than logarithmically, we have the almost sure convergence (see Proposition 3.25 in Section 3.3.3)

$$\lim_{n\to\infty} \frac{\mathrm{ht}(\mathtt{T}_n)}{\sum_{i=2}^{n} \frac{w_i}{W_i}} = 1,$$

which in some sense indicates that all the action takes place at the very tip of the tree.

Convergence of the weight measure

We also study the convergence of some natural probability measures defined on the trees $(\mathtt{T}_n)_{n\geq 1}$. This will prove useful for the applications developed in Chapter 4.

For this result it will be easier to work with plane trees. We introduce the Ulam-Harris tree $\mathbb{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$, where $\mathbb{N} := \{1, 2, \ldots\}$. Classically, a plane tree τ is defined as a non-empty subset of $\mathbb{U}$ such that

(i) if $v \in \tau$ and $v = ui$ for some $i \in \mathbb{N}$, then $u \in \tau$,

(ii) for all $u \in \tau$, there exists $\deg_\tau^+(u) \in \mathbb{N} \cup \{0\}$ such that for all $i \in \mathbb{N}$, $ui \in \tau$ iff $i \leq \deg_\tau^+(u)$.

We choose to construct our sequence $(\mathtt{T}_n)_{n\geq 1}$ of weighted recursive trees as plane trees by considering that each time a vertex is added, it becomes the right-most child of its parent. In this way the vertices $(u_1, u_2 \ldots)$ of the trees $(\mathtt{T}_n)_{n\geq 1}$, listed in order of arrival, form a sequence of elements of $\mathbb{U}$. In fact, from now on, we will always assume that we use this particular embedded construction, both for WRT and PA trees.

We also denote $\partial\mathbb{U} = \mathbb{N}^{\mathbb{N}}$, which can be interpreted as the set of infinite paths from the root to infinity, and write $\overline{\mathbb{U}} = \mathbb{U} \cup \partial\mathbb{U}$. We classically endow this set with the distance

$$d(u, v) = \exp(-\mathrm{ht}(u \wedge v))$$

where $u \wedge v$ denotes the most recent common ancestor of u and v in $\overline{\mathbb{U}}$.

For every $n \geq 1$, we define the measure μ_n on $\mathbb{U}$, which only charges the set $\{u_1, \ldots, u_n\}$ of vertices of $\mathtt{T}_n$, with for any $1 \leq k \leq n$,

$$\mu_n(u_k) = \frac{w_k}{W_n}. \tag{3.10}$$

We refer to μ_n as the *natural weight measure* on $\mathtt{T}_n$. The following theorem classifies the possible behaviours of (μ_n) for any weight sequence.

Theorem 3.4. *The sequence $(\mu_n)_{n\geq 1}$ converges almost surely weakly towards a limiting probability measure μ on $\overline{\mathbb{U}}$. There are three possible behaviours for μ:*

(i) *If $\sum_{i=1}^{\infty} w_i < \infty$, then μ is carried on $\mathbb{U}$.*

(ii) *If $\sum_{i=1}^{\infty} w_i = \infty$ and $\sum_{i=1}^{\infty} \left(\frac{w_i}{W_i}\right)^2 < \infty$, then μ is diffuse and supported on $\partial\mathbb{U}$.*

(iii) *If* $\sum_{i=1}^{\infty}\left(\frac{w_i}{W_i}\right)^2 = \infty$ *then* μ *is concentrated on one point of* $\partial\mathbb{U}$.

This convergence can be extended to other natural measures on the tree, such as the uniform measure on $\mathtt{T}_n$, or some "preferential attachment measure" which charges each vertex proportionally to some affine function of its degree. This is the content of Proposition 3.8.

3.1.3 Organisation of the chapter

The paper is organised as follows.

We first investigate some properties of weighted random trees $(\mathtt{T}_n)_{n\geq 1}$ with deterministic weight sequence $\mathbf{w}$. In Section 3.2.1 we first prove Proposition 3.5 which states the convergence of the degree sequence using elementary methods. Then in Section 3.2.2, we prove the weak convergence of the weight measure μ_n to some limit μ and describe three regimes for its behaviour. We also study other natural measures related to the sequence of trees $(\mathtt{T}_n)$ and prove that they also converge towards μ. For all these measures, our main tool is introducing martingales related to the mass of a subtree descending from a fixed vertex. This is the content of Theorem 3.4 and Proposition 3.8. In Section 3.3, we prove Theorem 3.3 about the convergence of the height and the profile of WRT. This is achieved by first proving the uniform convergence of a rescaled version of the Laplace transform of the profile on a complex domain, which is the content of Proposition 3.9. This ensures that we can use [84, Theorem 2.1] for the convergence of the profile. This convergence provides a lower-bound for the height of the tree; we then prove a matching upper-bound to obtain asymptotics for the height.

Then we switch to studying some sequence $(\mathtt{P}_n)_{n\geq 1}$ of preferential attachment trees with initial fitnesses $\mathbf{a}$. In Section 3.4, we present a proof of Theorem 3.1 using a coupling of the preferential attachment process with a sequence of Pólya urn processes and this establishes that $(\mathtt{P}_n)_{n\geq 1}$ can also be described as having distribution $\mathrm{WRT}(\mathsf{w}^{\mathbf{a}})$ for a random sequence $\mathsf{w}^{\mathbf{a}}$; we then prove Proposition 3.2 which relates the properties of $\mathsf{w}^{\mathbf{a}}$ to the ones of $\mathbf{a}$. We finish the section by stating and proving Proposition 3.27 in which we prove that the sequence $(\mathsf{M}_n^{\mathbf{a}})$ defined above as some random multiple of $(\mathsf{W}_n^{\mathbf{a}})$ is a Markov chain. In Section 3.5, we identify in Proposition 3.28 the distribution of the chain $(\mathsf{M}_n^{\mathbf{a}})$ for particular sequences $\mathbf{a}$ using moment identifications. We then present an application of this result to an other model of preferential attachment graph in Proposition 3.31.

Some technical results can be found in Appendix 3.A.

3.2 Measures and degrees in weighted random trees

In this section, we work with a sequence of trees $(\mathtt{T}_n)_{n\geq 1}$ that has distribution $\mathrm{WRT}(\mathbf{w})$ for a deterministic sequence $\mathbf{w}$. We start with two statistics of the tree that are quite easy to analyse, namely the sequence of degrees of the vertices of the tree and also some natural measures defined on the tree.

3.2.1 Convergence of the degree sequence

We start the section by proving convergence for the sequence of degrees of the vertices in their order of creation under the WRT model. We suppose here that the sequence of weights $\mathbf{w}$ is such that there exist constants $C > 0$ and $0 < \gamma < 1$ for which

$$W_k \underset{k\to\infty}{\sim} C \cdot k^{\gamma}. \tag{3.11}$$

We write $\deg^+_{\mathtt{T}_n}(u_k)$ for the out-degree of the vertex u_k in $\mathtt{T}_n$. For a fixed $k\geq 1$, we remark that, as a sequence of random variables indexed by $n\geq 1$, we have the equality in distribution

$$\left(\deg^+_{\mathtt{T}_n}(u_k)\right)_{n\geq 1} \overset{(\mathrm{d})}{=} \left(\sum_{i=k}^{n-1} \mathbf{1}_{\left\{U_i\leq \frac{w_k}{W_i}\right\}}\right)_{n\geq 1}, \tag{3.12}$$

with $(U_i)_{i\geq 1}$ a sequence of independent uniform variables in $(0\,,1)$. With this description of the distribution of the degrees of fixed vertices, only using some law of large numbers for the convergence and Chernoff bounds for the fluctuations we obtain the following result:

Proposition 3.5. *For a sequence of weights* $\mathbf{w}$ *satisfying* (3.11)*, the following holds.*

(i) *We have the almost sure pointwise convergence*

$$n^{-(1-\gamma)}\cdot\left(\deg^+_{\mathtt{T}_n}(u_1),\deg^+_{\mathtt{T}_n}(u_2),\dots\right) \underset{n\to\infty}{\longrightarrow} \frac{1}{(1-\gamma)C}\cdot(w_1,w_2,\dots). \tag{3.13}$$

(ii) *If the sequence furthermore satisfies* $w_k\leq (k+1)^{\gamma-1+c'+o(1)}$ *for some constant* $0\leq c'<1-\gamma$*, then there exists a function of* k *which goes to* 0 *as* $k\to\infty$*, also denoted* $o(1)$*, such that all* n *large enough, we have for all* $k\geq 1$

$$\deg^+_{\mathtt{T}_n}(u_k)\leq n^{1-\gamma}\cdot (k+1)^{\gamma-1+c'+o(1)}, \tag{3.14}$$

and the convergence (3.13) *holds almost surely in the space* ℓ^p *for all* $p>\frac{1}{1-\gamma-c'}$*.*

Proof. To prove (i), just remark that for any $k\geq 1$ such that $w_k\neq 0$, thanks to (3.11), we have

$$\sum_{i=k}^{n-1}\frac{w_k}{W_i}\underset{n\to\infty}{\sim} w_k\cdot \frac{n^{1-\gamma}}{C(1-\gamma)},$$

so thanks to the law of large numbers, we get that almost surely

$$\deg^+_{\mathtt{T}_n}(u_k)=\sum_{i=k}^{n-1} \mathbf{1}_{\left\{U_i\leq \frac{w_k}{W_i}\right\}} \underset{n\to\infty}{\sim} \sum_{i=k}^{n-1}\frac{w_k}{W_i} \underset{n\to\infty}{\sim} w_k\cdot\frac{n^{1-\gamma}}{(1-\gamma)C},$$

and hence $n^{-(1-\gamma)}\cdot \deg^+_{\mathtt{T}_n}(u_k)\to \frac{w_k}{(1-\gamma)C}$. For the indices k for which $w_k=0$, we of course have $\deg^+_{\mathtt{T}_n}(u_k)=0$ almost surely for all $n\geq 1$, and so the convergence also holds. This finishes the proof of (i).

For the second part of the statement, let us first compute

$$\begin{aligned}\mathbb{E}\left[\exp\left(\deg^+_{\mathtt{T}_n}(u_k)\right)\right]=\mathbb{E}\left[\exp\left(\sum_{i=k}^{n-1}\mathbf{1}_{\left\{U_i\leq\frac{w_k}{W_i}\right\}}\right)\right] &= \prod_{i=k}^{n-1}\left(1+(e-1)\frac{w_k}{W_i}\right)\\ &\leq \exp\left((e-1)w_k\sum_{i=k}^{n-1}\frac{1}{W_i}\right).\end{aligned}$$

Now let C' be a constant such that for all $n\geq 1$, we have $\sum_{i=1}^{n-1}\frac{1}{W_i}\leq C'\cdot n^{1-\gamma}$ (such a constant exists because of the assumption (3.11)). For all $k\geq 1$, we introduce

$$\xi_k := \max\left(2C'(e-1)w_k, k^{\gamma-1}\log^2(k+a)\right),$$

where the real number $a > 0$ is chosen in such a way that the function $x \mapsto x^{\gamma-1}\log(x+a)$ is decreasing on $\mathbb{R}_+^*$. Using Markov's inequality, we get for any integers k and n such that $n \geq k$

$$\begin{aligned}\mathbb{P}\left(\deg^+_{\mathtt{T}_n}(u_k) \geq \xi_k \cdot n^{1-\gamma}\right) &\leq \exp\left(-\xi_k \cdot n^{1-\gamma} + (e-1)w_k \sum_{i=k}^{n-1} \frac{1}{W_i}\right) \\ &\leq \exp\left(-\frac{1}{2} \cdot \xi_k \cdot n^{1-\gamma}\right).\end{aligned}$$

Using a union bound, the fact that $\deg^+_{\mathtt{T}_n}(u_k) = 0$ for any $k > n$, and the definition of ξ_k, we get that for all $n \geq 1$

$$\begin{aligned}\mathbb{P}\left(\exists k \geq 1, \quad \deg^+_{\mathtt{T}_n}(u_k) \geq \xi_k \cdot n^{1-\gamma}\right) &\leq \sum_{k=1}^{n} \exp\left(-\frac{1}{2} \cdot \xi_k \cdot n^{1-\gamma}\right) \\ &\leq n \cdot \exp\left(-\frac{1}{2} \cdot \log^2(n+a)\right).\end{aligned}$$

The last display is summable over all $n \geq 1$ and hence using the Borel-Cantelli lemma, we almost surely have for n large enough

$$\forall k \geq 1, \quad \deg^+_{\mathtt{T}_n}(u_k) \leq n^{1-\gamma} \cdot \xi_k.$$

We can conclude by noting that under our assumptions we have $\xi_k \leq (k+1)^{\gamma-1+c'+o(1)}$. The convergence in ℓ^p for $p > \frac{1}{1-\gamma-c'}$ is just obtained by dominated convergence using the pointwise convergence (3.13) and the ℓ^p domination (3.14). $\square$

3.2.2 Convergence of measures

The goal of this section is to prove Theorem 3.4, which concerns the convergence of the sequence of weight measures (μ_n) seen as measures on $\overline{\mathbb{U}}$. One of the key arguments is the fact that the weight of the subtree descending from a fixed vertex can be described using a generalised Pólya urn scheme, as studied by Pemantle [103]. We also prove Proposition 3.8, which states the weak convergence of other measures.

Convergence of the weight measure in $\overline{\mathbb{U}}$. Recall from the introduction the definition of the Ulam-Harris tree $\mathbb{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n$ and its completed version $\overline{\mathbb{U}} = \mathbb{U} \cup \partial\mathbb{U}$, which is endowed with the distance $d(u, v) = \exp\left(-\operatorname{ht}(u \wedge v)\right)$. For any $u \in \mathbb{U}$, we write $T(u) := \left\{uv \;\middle|\; v \in \overline{\mathbb{U}}\right\}$ the subtree descending from u. In $\overline{\mathbb{U}}$ there is an easy characterisation of the weak convergence of Borel measures, which a direct consequence of the Portmanteau theorem (see e.g. [27, Theorem 2.1]):

Lemma 3.6. *Let $(\pi_n)_{n\geq 1}$ be a sequence of Borel probability measures on $\overline{\mathbb{U}}$. Then $(\pi_n)_{n\geq 1}$ converges weakly to a probability measure π if and only if for any $u \in \mathbb{U}$,*

$$\pi_n(\{u\}) \to \pi(\{u\}) \quad \text{and} \quad \pi_n(T(u)) \to \pi(T(u)) \quad \text{as} \quad n \to \infty.$$

We are going to apply this criterion to our sequence $(\mu_n)_{n\geq 1}$, which, we recall, is defined in such a way that for all $n \geq 1$, the measure μ_n charges only the vertices $\{u_1, u_2, \ldots, u_n\}$ of the tree $\mathtt{T}_n$, and such that for any $1 \leq k \leq n$,

$$\mu_n(\{u_k\}) = \frac{w_k}{W_n}. \tag{3.15}$$

We can already see that if $(W_n)_{n\geq 1}$ converges to some W_∞ we have $\mu_n(\{u_k\}) \to \frac{w_k}{W_\infty}$ as $n \to \infty$, and in this case it is easy to verify that μ_n weakly converges to some limit μ which is such that $\mu(\{u_k\}) = \frac{w_k}{W_\infty}$. In this case $\mu(\mathbb{U}) = 1$ and so μ is carried on $\mathbb{U}$.

From now on, let us assume that $W_n \to \infty$ as $n \to \infty$. In this case we have $\mu_n(\{u_k\}) \to 0$ as $n \to \infty$. Now denote for every integers $n, k \geq 1$,

$$M_n^{(k)} := \mu_n(T(u_k)),$$

the proportion of the total mass above vertex u_k at time n. Remark that this quantity evolves as the proportion of red balls in a *time-dependent Pólya urn scheme* with weights $(w_i)_{i\geq k+1}$, see [103], starting at time k with W_{k-1} black balls and w_k red balls[2]. In particular, for all $n \geq k$,

$$\begin{aligned}\mathbb{E}\left[M_{n+1}^{(k)} \,\middle|\, \mathtt{T}_n\right] &= \frac{W_n}{W_{n+1}} \cdot M_n^{(k)} + \frac{w_{n+1}}{W_{n+1}} \cdot M_n^{(k)} \\ &= M_n^{(k)}.\end{aligned}$$

Hence for all $k \geq 1$, the sequence $(M_n^{(k)})_{n\geq k}$ is a martingale with value in $[0\,,1]$ so it converges almost surely to a limit $M_\infty^{(k)}$. Also, for any $u \in \mathbb{U}$ that does not receive a label in the process, the sequence $(\mu_n(T(u)))_{n\geq 1}$ (and also $(\mu_n(\{u\}))_{n\geq 1}$) is identically equal to zero. Hence we have convergence of $(\mu_n(\{u\}))_{n\geq 1}$ and $(\mu_n(T(u)))_{n\geq 1}$ for all $u \in \mathbb{U}$.

The last step in order to prove the weak convergence of $(\mu_n)_{n\geq 1}$ is to prove that the quantities that we obtain in the limit indeed define a probability measure on $\overline{\mathbb{U}}$. If for all $u \in \mathbb{U}$ we have

$$\lim_{n\to\infty} \mu_n(T(u)) = \sum_{i=1}^{\infty} \lim_{n\to\infty} \mu_n(T(ui)), \tag{3.16}$$

then it entails that $\mu_n \underset{n\to\infty}{\to} \mu$, where μ is the unique probability measure on $\overline{\mathbb{U}}$ such that for all $u \in \mathbb{U}$,

$$\mu(\{u\}) = 0 \quad \text{and} \quad \mu(T(u)) = \lim_{n\to\infty} \mu_n(T(u)).$$

For any $u \notin \{u_1, u_2, \ldots\}$, the equality (3.16) is immediate, so let us prove it for all u_k for $k \geq 1$. For any $n, k, i \geq 1$, let

$$M_n^{(k,i)} := \sum_{j=i+1}^{\infty} \mu_n(T(u_k j)) = \mu_n\left(T(u_k)\right) - \sum_{j=1}^{i} \mu_n\left(T(u_k j)\right).$$

Using what we just proved, we know that for any k, i, the quantity $M_n^{(k,i)}$ almost surely converges as $n \to \infty$ to some limit $M_\infty^{(k,i)}$. Proving (3.16) reduces to proving that for any $k \geq 1$, we almost surely have $M_\infty^{(k,i)} \underset{i\to\infty}{\to} 0$. By construction, the sequence $(M_\infty^{(k,i)})_{i\geq 1}$ is non-negative and non-increasing, hence it converges, so it suffices to prove that its limit is 0 almost surely.

We define $\tau^{(k,i)} := \inf\{n \geq 1 \mid u_n = u_k i\}$, the time when the vertex u_k receives its i-th child in the growth procedure. Remark that after this random time, the process $(M_n^{(k,i)})_{n\geq \tau^{(k,i)}}$ is a martingale because again, it evolves as the proportion of red balls in a time-dependent Pólya urn scheme, starting with w_k red balls and $W_{\tau^{(k,i)}}$ blacks balls. (If $\tau^{(k,i)}$ is infinite, then the sequence $(M_n^{(k,i)})_{n\geq 1}$ is identically 0.) Hence, using the crude bound $\tau^{(k,i)} \geq i$, which entails that $W_{\tau^{(k,i)}} \geq W_i$ almost surely, we get

$$\mathbb{E}\left[M_\infty^{(k,i)}\right] = \mathbb{E}\left[M_{\tau^{(k,i)}}^{(k,i)} \mathbf{1}_{\{\tau^{(k,i)}<\infty\}}\right] \leq \frac{w_k}{W_i} \underset{i\to\infty}{\to} 0,$$

[2]Those numbers of balls are not required to be integers.

hence $M_\infty^{(k,i)} \underset{i\to\infty}{\to} 0$ in L^1, so its almost sure limit is also 0. In the end, by Lemma 3.6, the sequence of measures (μ_n) almost surely converges weakly to a limit μ, and this measure only charges the set $\partial\mathbb{U}$.

Lemma 3.7. *Suppose that $\sum_{n=1}^\infty w_n = \infty$ so that μ is carried on $\partial\mathbb{U}$. Then either $\sum_{n=1}^\infty \left(\frac{w_n}{W_n}\right)^2 < \infty$ and then μ is almost surely diffuse or $\sum_{n=1}^\infty \left(\frac{w_n}{W_n}\right)^2 = \infty$ and then μ is carried on one point of $\partial\mathbb{U}$.*

Proof. For any $k \geq 1$ the process $(\mu_n(T(u_k))_{n\geq k}$ follows a so-called time-dependent Pólya urn scheme with weights $(w_n)_{n\geq k+1}$. By the work of Pemantle in [102], if we assume $\sum_{n=1}^\infty \left(\frac{w_n}{W_n}\right)^2 = \infty$ then the limiting proportion $\mu(T(u_k))$ almost surely belongs to the set $\{0,1\}$. This translates into the fact that $\mu(T(u)) \in \{0,1\}$ almost surely for any $u \in \mathbb{U}$, which entails that μ is almost surely carried on one leaf of $\partial\mathbb{U}$.

On the contrary, let us suppose that $\sum_{n=1}^\infty \left(\frac{w_n}{W_n}\right)^2 < \infty$ and prove that this entails that the limiting measure μ is diffuse almost surely. Consider the function $(\cdot \wedge \cdot) : \overline{\mathbb{U}} \times \overline{\mathbb{U}} \to \overline{\mathbb{U}}$ which associates to each couple (u,v) their most recent common ancestor $u \wedge v$ in the completed tree $\overline{\mathbb{U}}$. This function is continuous with respect to the distance d. Then, since $\mu_n \to \mu$ almost surely, we also have the almost sure weak convergence

$$(\cdot \wedge \cdot)_*(\mu_n \otimes \mu_n) \to (\cdot \wedge \cdot)_*(\mu \otimes \mu). \tag{3.17}$$

Let us fix $n \geq 1$ and let D_n and D_n' be two independent vertices taken under μ_n, conditionally on the tree $\mathtt{T}_n$. Then, the proof of [41, Lemma 3.8] ensures that

$$\mathbb{P}\left(D_n \wedge D_n' = u_k\right) = \left(\frac{w_k}{W_k}\right)^2 \cdot \prod_{i=k+1}^{n} \left(1 - \left(\frac{w_i}{W_i}\right)^2\right) \underset{k\to\infty}{\longrightarrow} p_k := \left(\frac{w_k}{W_k}\right)^2 \cdot \prod_{i=k+1}^{\infty} \left(1 - \left(\frac{w_i}{W_i}\right)^2\right).$$

Note that the obtained sequence $(p_k)_{k\geq 1}$ is a probability distribution, which thanks to the weak convergence (3.17) corresponds to the (annealed) distribution $p_k = \mathbb{P}\left(D_\infty \wedge D_\infty' = u_k\right)$, where D_∞ and D_∞' are two independent points taken under the measure μ, conditionally on μ. Now we can write

$$\mathbb{P}\left(d(D_\infty, D_\infty') \leq e^{-k}\right) = \mathbb{P}\left(\mathrm{ht}(D_\infty \wedge D_\infty') \geq k\right) \leq \sum_{i=k+1}^{\infty} p_i,$$

where the inequality is due to the fact that the vertices $u_1, u_2, \ldots, u_k$ have a height smaller than k. Hence $\mathbb{P}\left(d(D_\infty, D_\infty') = 0\right) \leq \lim_{k\to\infty} \mathbb{P}\left(d(D_\infty, D_\infty') \leq e^{-k}\right) = 0$. So, almost surely, two points taken independently under μ are different, and this ensures that μ is diffuse. □

In the end, we just finished the proof of Theorem 3.4.

Other sequences of measures We also study two other sequences of measures (η_n) and (ν_n) carried on the Ulam tree $\mathbb{U}$. For every $n \geq 2$, these measures only charge the vertices $\{u_1, u_2, \ldots, u_n\}$ in such a way that for any $1 \leq k \leq n$,

$$\eta_n(u_k) = \frac{b_k + \deg^+_{\mathtt{T}_n}(u_k)}{B_n + n - 1} \qquad \text{and} \qquad \nu_n(u_k) = \frac{1}{n},$$

where $(b_n)_{n\geq 1}$ is a sequence of real numbers such that $b_1 > -1$ and $b_n \geq 0$ for all $n \geq 2$. We write $B_n := \sum_{k=1}^n b_k$. We suppose that $B_n = O(n)$ and that there exists $\epsilon > 0$ such that $b_n = O(n^{1-\epsilon})$. The assumptions on the sequence $(b_n)_{n\geq 1}$ are chosen such that they are satisfied by a sequence $(a_n)_{n\geq 1}$ of initial fitnesses that satisfies (H_c) for some $c > 0$.

Proposition 3.8. *Under the assumptions $\sum_{n=1}^{\infty} w_n = \infty$ and $\sum_{n=1}^{\infty} \left(\frac{w_n}{W_n}\right)^2 < \infty$, the sequences $(\eta_n)_{n\geq 1}$ and $(\nu_n)_{n\geq 1}$ converge almost surely weakly towards the limiting measure μ on $\partial\mathbb{U}$ defined in Theorem 3.4.*

For the proof of this proposition, we are going to use Lemma 3.6 again, using appropriate martingales in order to handle the evolution of the measure of the subtree descending from every vertex $u \in \mathbb{U}$. We treat the two sequences of measures separately.

The degree measure. Consider the sequence $(\eta_n)_{n\geq 1}$ on $\overline{\mathbb{U}}$. Since the sequence $(W_n)_{n\geq 1}$ tends to infinity, we have $\eta_n(\{u\}) \to 0$ for every $u \in \mathbb{U}$. Indeed, using the equality in distribution (3.12) and Lemma 3.32 in the appendix, it is easy to see that either $\sum_{i=1}^{\infty} W_i^{-1} < \infty$ and in this case the degrees $\deg^+_{\mathtt{T}_n}(u_k)$ are eventually constant as $n \to \infty$; or $\sum_{i=1}^{\infty} W_i^{-1} = \infty$, in which case we have the almost sure asymptotic behaviour $\deg^+_{\mathtt{T}_n}(u_k) \sim w_k \cdot \sum_{i=k}^{n} W_i^{-1}$. In both cases, for all $k \geq 1$, we have $n^{-1} \deg^+_{\mathtt{T}_n}(u_k) \to 0$ almost surely as $n \to \infty$.

As in the preceding case, for all $k \geq n$ we let

$$N_n^{(k)} := \eta_n(T(u_k)).$$

Conditionally on $\mathtt{T}_n$, with probability $M_n^{(k)}$, the vertex u_{n+1} is grafted onto $T(u_k)$ and with complementary probability, it is not. So

$$\begin{aligned} N_{n+1}^{(k)} &= \frac{1}{B_{n+1}+n} \cdot \left((B_n + n - 1) \cdot N_n^{(k)} + b_{n+1} + 1\right) \quad \text{with probability } M_n^{(k)}, \\ &= \frac{B_n + n - 1}{B_{n+1}+n} \cdot N_n^{(k)} \quad \text{with probability } (1 - M_n^{(k)}). \end{aligned}$$

Now compute

$$\begin{aligned} \mathbb{E}\left[N_{n+1}^{(k)} - M_{n+1}^{(k)} \,\middle|\, \mathcal{F}_n\right] &= \frac{B_n + n - 1}{B_{n+1}+n} \cdot N_n^{(k)} + \frac{b_{n+1}+1}{B_{n+1}+n} \cdot M_n^{(k)} - M_n^{(k)} \\ &= \frac{B_n + n - 1}{B_{n+1}+n} \cdot \left(N_n^{(k)} - M_n^{(k)}\right). \end{aligned}$$

Hence, if we denote $X_n^{(k)} := (B_n + n - 1) \cdot \left(N_n^{(k)} - M_n^{(k)}\right)$, then the last computation shows that $\left(X_n^{(k)}\right)_{n\geq k}$ is a martingale for the filtration generated by $(\mathtt{T}_n)_{n\geq 1}$. More precisely we can write

$$X_{n+1}^{(k)} - X_n^{(k)} = \underbrace{\left(\frac{W_n}{W_{n+1}}(1 + b_{n+1}) - \frac{w_{n+1}}{W_{n+1}}(B_{n+1} + n)\right)}_{c_n} \cdot \left(\mathbf{1}_{\{u_{n+1} \in T(u_k)\}} - M_n^{(k)}\right),$$

hence we have

$$\mathbb{E}\left[X_{n+1}^{(k)} - X_n^{(k)} \,\middle|\, \mathtt{T}_n\right] = 0 \qquad \text{and} \qquad \mathbb{E}\left[\left(X_{n+1}^{(k)} - X_n^{(k)}\right)^2\right] \leq c_n^2.$$

Then, using [39, Theorem 1], we get that if

$$\sum_{n=k}^{\infty} n^{-2} c_n^2 < \infty \tag{3.18}$$

then $\frac{X_n^{(k)}}{n} \to 0$ a.s. as $n \to \infty$, which would prove that $N_n^{(k)} \longrightarrow M_\infty^{(k)}$ as $n \to \infty$. In our case, we can verify that (3.18) holds. Indeed, using the fact that we have assumed that $B_n = O(n)$ and $b_{n+1} = O(n^{1-\epsilon})$, we have

$$\begin{aligned}
n^{-2}c_n^2 &= n^{-2}\left(\frac{W_n}{W_{n+1}}(1+b_{n+1}) - \frac{w_{n+1}}{W_{n+1}}(B_{n+1}+n)\right)^2 \\
&\leq n^{-2} \cdot 3\left(1 + b_{n+1}^2 + \left(\frac{w_{n+1}}{W_{n+1}}(B_{n+1}+n)\right)^2\right) \\
&\leq 3n^{-2} + 3b_{n+1}^2 n^{-2} + \mathrm{cst} \cdot \left(\frac{w_{n+1}}{W_{n+1}}\right)^2,
\end{aligned}$$

which is summable under our assumptions. In the end, using Lemma 3.6, we have the almost sure convergence

$$\eta_n \longrightarrow \mu \qquad \text{weakly.}$$

The uniform measure on the vertices of $\mathtt{T}_n$. Consider the sequence (ν_n) on $\overline{\mathbb{U}}$. Fix $k \geq 1$. For any $n \geq k$ we can write $\nu_n(T(u_k)) = \frac{1}{n}\sum_{i=k}^n \mathbf{1}_{\{u_i \in T(u_k)\}}$. For any $i \geq k+1$, we have $\mathsf{p}_i := \mathbb{P}(u_i \in T(u_k) \mid \mathcal{F}_{i-1}) = \mu_{i-1}(T(u_k))$, which tends a.s. to some limit $\mu(T(u_k))$ as $i \to \infty$. Using Lemma 3.32 in the appendix, we have

$$\frac{\sum_{i=k+1}^n \mathbf{1}_{\{u_i \in T(u_k)\}}}{\sum_{i=k+1}^n \mathsf{p}_i} \underset{n\to\infty}{\longrightarrow} 1 \quad \text{a.s. on the event} \quad \left\{\sum_{i=k+1}^\infty \mathsf{p}_i = \infty\right\}$$

and also

$$\sum_{i=k+1}^n \mathbf{1}_{\{u_i \in T(u_k)\}} \quad \text{converges a.s. on the event} \quad \left\{\sum_{i=k+1}^\infty \mathsf{p}_i < \infty\right\}.$$

In both cases we get $\nu_n(T(u_k)) \underset{n\to\infty}{\to} \lim_{i\to\infty} \mathsf{p}_i = \mu(T(u_k))$ almost surely. We also have for any $k \geq 1$,

$$\nu_n(\{u_k\}) = \frac{1}{n} \underset{n\to\infty}{\to} 0 \qquad \text{and of course} \qquad \forall u \notin \{u_1, u_2, \dots\}, \forall n \geq 1,\ \nu_n(\{u\}) = \nu_n(T(u)) = 0,$$

so we can conclude using Lemma 3.6 that almost surely $\nu_n \underset{n\to\infty}{\to} \mu$ weakly.

3.3 Height and profile of WRT

The main goal of this section is to prove Theorem 3.3 which gives asymptotics for the profile and height of the tree. Recall that we denote

$$\mathbb{L}_n(k) := \#\{1 \leq i \leq n \mid \mathrm{ht}(u_i) = k\},$$

the number of vertices at height k in the tree $\mathtt{T}_n$. In order to get information on the sequence of functions $(k \mapsto \mathbb{L}_n(k))_{n\geq 1}$ we study their Laplace transform

$$z \mapsto \sum_{k=0}^\infty \mathbb{L}_n(k) e^{kz} = \sum_{i=1}^n e^{z\,\mathrm{ht}(u_i)} = n \cdot \int_{\mathbb{U}} e^{z\,\mathrm{ht}(u)} \mathrm{d}\nu_n(u), \tag{3.19}$$

where the last expression is given using an integral against the probability measure ν_n defined in Section 3.2.2 as the uniform measure on the vertices of $\mathtt{T}_n$. The key result in our approach is to prove the convergence of this sequence of analytic functions when appropriately rescaled, uniformly in z on an open neighbourhood of 0 in the complex plane. It then allows us to use [84, Theorem 2.1] and hence derive a convergence result for the profile. We actually start in Section 3.3.1 by studying the convergence of the similarly defined sequence of functions

$$z \mapsto \int_{\mathbb{U}} e^{z\,\mathrm{ht}(u)} \mathrm{d}\mu_n(u) = \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)}, \tag{3.20}$$

where we integrate with respect to the weight measure μ_n instead of the uniform measure ν_n as before. This one is easier to study because for every fixed $z \in \mathbb{C}$, it defines a martingale as n grows, up to some deterministic scaling. Then in Section 3.3.2, we make use of this first convergence and show that up to some deterministic multiplicative constant, the two sequences of integrals appearing in (3.19) and (3.20) are almost surely equivalent when n tends to infinity.

We work under some technical assumption for the sequence $\mathbf{w}$. Let us fix $\gamma > 0$ and suppose from now on that $\mathbf{w}$ satisfies the assumption ($\square_\gamma^p$) for some $p \in (1\,, 2]$, i.e.,

$$W_n \underset{n\to\infty}{\bowtie} \mathrm{cst} \cdot n^\gamma \qquad \text{and} \qquad \sum_{i=n}^{2n} w_n^p \le n^{1+(\gamma-1)p+o(1)}.$$

We let $\phi : z \mapsto \gamma(e^z - 1)$ be a function of a complex parameter z and let $z \mapsto N_n(z)$ be the following rescaled version of the Laplace transform of the profile

$$N_n(z) := n^{-(1+\phi(z))} \sum_{k=0}^{\infty} \mathbb{L}_n(k) e^{zk}.$$

The proposition below ensures that the sequence $(z \mapsto N_n(z))_{n\ge1}$ converges uniformly on all compact subsets of some open domain $\mathscr{D} \subset \mathbb{C}$ to some limiting function $z \mapsto N_\infty(z)$ which does not vanish anywhere on the set $\mathscr{D} \cap \mathbb{R}$, along with some more technical statements.

Proposition 3.9. *Suppose that the weight sequence* $\mathbf{w}$ *satisfies* ($\square_\gamma^p$) *for some* $\gamma > 0$ *and some* $p \in (1\,, 2]$. *Then there exists an open connected domain* $\mathscr{D} \subset \mathbb{C}$ *such that* $\mathscr{D} \cap \mathbb{R} = (z_-\,, z_+)$ *with* $z_- < 0$ *and* z_+ *is the largest real solution of the equation* $\gamma(ze^z - e^z + 1) - 1 = 0$ *and such that the following properties are satisfied.*

(i) *With probability* 1, *the sequence of random analytic functions* $(z \mapsto N_n(z))_{n\ge1}$ *converges uniformly on all compact subsets of* $\mathscr{D}$, *as* $n \to \infty$, *to some random analytic function* $z \mapsto N_\infty(z)$ *which satisfies* $\mathbb{P}(N_\infty(z) \neq 0 \text{ for all } z \in (z_-, z_+)) = 1$.

(ii) *For every compact set* $K \subset \mathscr{D}$ *and* $r \in \mathbb{N}$, *we can find an a.s. finite random variable* $C_{K,r}$ *such that for all* $n \in \mathbb{N}$,

$$\sup_{z\in K} |N_n(z) - N_\infty(z)| < C_{K,r}(\log n)^{-r}.$$

(iii) *For every compact set* $K \subset (z_-\,, z_+)$, *every* $0 < a < \pi$ *and* $r \in \mathbb{N}$,

$$\sup_{z\in K} \left[e^{-(1+\phi(z))\log n} \int_a^\pi \left| \sum_{k=0}^{\infty} \mathbb{L}_n(k) e^{z+iu} \right| \mathrm{d}u \right] = o\big((\log n)^{-r}\big) \qquad a.s. \text{ as } n \to \infty.$$

Under the results of Proposition 3.9 we can apply [84, Theorem 2.1] whose conclusions for the sequence $(k \mapsto \mathbb{L}(k))_{n\ge1}$ are the following. For any $k \ge 0$, $n \ge 1$ and $z \in (z_-\,, z_+)$, we denote

$$x_n(k; z) = \frac{k - \gamma e^z \log n}{\sqrt{\gamma e^z \log n}}.$$

Then, for every integer $r \geq 0$ and every compact set $K \subset (z_- , z_+)$, we have the convergence

$$(\log n)^{\frac{r+1}{2}} \cdot \sup_{k \in \mathbb{N}} \sup_{z \in K} \left| e^{zk-(1+\phi(z))\log n} \mathbb{L}_n(k) - \frac{N_\infty(z) e^{-\frac{1}{2} x_n(k;z)^2}}{\sqrt{2\pi\gamma e^z \log n}} \sum_{j=0}^{r} \frac{G_j(x_n(k);z)}{(\log n)^{j/2}} \right| \xrightarrow[n\to\infty]{\text{a.s.}} 0, \quad (3.21)$$

where for all $j \geq 0$, the (random) functions $G_j(x,z)$ are polynomials of degree at most 3 in x and are entirely determined from ϕ and N_∞, with $G_1 = 1$, see [84] for their complete definition. The asymptotics (3.7) and (3.8) stated in Theorem 3.3 follow from the last display. Indeed, (3.7) is obtained by letting $r = 0$ and $z = 0$ and using the fact that $N_\infty(0) = 1$ almost surely. For (3.8), we let $r = 0$, and use $k = \lfloor \gamma e^z \log n \rfloor$.

In Section 3.3.3, we complete the proof of Theorem 3.3 by computing the asymptotic behaviour of the height of the tree. Since the convergence of the profile already ensures that there are almost surely vertices at height $\gamma e^{(z_+ - \epsilon)} \log n$ for $\epsilon > 0$ small enough and all n large enough, it suffices to prove a corresponding upper-bound in order to finish proving the convergence (3.9) in Theorem 3.3.

3.3.1 Study of the Laplace transform of the weighted profile

We study the sequence $\left(z \mapsto \sum_{i=1}^n \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)}\right)_{n\geq 1}$. The following lemma is the starting point of our analysis:

Lemma 3.10. *For all $z \in \mathbb{C}$ and all $n \geq 1$, we have*

$$\mathbb{E}\left[\sum_{i=1}^{n+1} \frac{w_i}{W_{n+1}} e^{z\,\mathrm{ht}(u_i)} \,\middle|\, \mathtt{T}_n \right] = \left(1 + (e^z - 1)\frac{w_{n+1}}{W_{n+1}}\right) \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)}.$$

Proof. Recall that conditionally on $\mathtt{T}_n$, the $(n+1)$-st vertex u_{n+1} of $\mathtt{T}_{n+1}$ is a child of the vertex $u_{K_{n+1}}$, where $\mathbb{P}(K_{n+1} = k \mid \mathtt{T}_n) = \frac{w_k}{W_n}$. We compute

$$\sum_{i=1}^{n+1} \frac{w_i}{W_{n+1}} e^{z\,\mathrm{ht}(u_i)} = \frac{W_n}{W_{n+1}} \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)} + \frac{w_{n+1}}{W_{n+1}} \cdot e^z \cdot e^{z\,\mathrm{ht}(u_{K_{n+1}})}.$$

Taking conditional expectation with respect to $\mathtt{T}_n$ yields

$$\begin{aligned} \mathbb{E}\left[\sum_{i=1}^{n+1} \frac{w_i}{W_{n+1}} e^{z\,\mathrm{ht}(u_i)} \,\middle|\, \mathtt{T}_n \right] &= \frac{W_n}{W_{n+1}} \cdot \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)} + \frac{w_{n+1}}{W_{n+1}} \cdot e^z \cdot \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)} \\ &= \left(1 + (e^z - 1)\frac{w_{n+1}}{W_{n+1}}\right) \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)}. \end{aligned}$$

This concludes the proof. $\square$

Let J be an integer that we are going to fix later on. The last result ensures that if $z \in \mathbb{C}$ is such that $\forall i \geq J,\ 1 + (e^z - 1)\frac{w_i}{W_i} \neq 0$, then we can define for all $n \geq J$

$$C_n(z) := \prod_{i=J}^{n} \left(1 + (e^z - 1)\frac{w_i}{W_i}\right) \qquad \text{and} \qquad M_n(z) := \frac{1}{C_n(z)} \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)},$$

and the sequence $(M_n(z))_{n\geq J}$ is a martingale. We want to prove results about the asymptotic behaviour of $(z \mapsto M_n(z))_{n\geq J}$, uniformly in z on an appropriate domain. If J is fixed, then there exists parameters z with $\mathrm{Im}(z) = \pi \mod 2\pi$ for which the sequence $(C_n(z))_{n\geq J}$ takes the value

0. Due to our assumption ($\square_\gamma$) on the sequence $\mathbf{w}$, we know that $\frac{w_n}{W_n} \to 0$ as $n \to \infty$. If we restrict ourselves to a domain of the form $\{z \in \mathbb{C} \mid \mathrm{Re}(z) < x\}$ for some $x > 0$, then

$$\left|1 + (e^z - 1)\frac{w_n}{W_n}\right| \geq 1 - |e^z - 1| \cdot \frac{w_n}{W_n} \geq 1 - (e^x + 1) \cdot \frac{w_n}{W_n} \underset{n\to\infty}{\to} 1 > 0,$$

hence it suffices to take J large enough in order for the sequence $(C_n(z))_{n\geq J}$ to only take non-zero values for all $z \in \{\xi \in \mathbb{C} \mid \mathrm{Re}(\xi) < x\}$ and all $n \geq J$. In what follows we work on the domain

$$\mathscr{E} = \{z \in \mathbb{C} \mid \mathrm{Re}\, z < z_+\}\,,$$

where z_+ is as defined in Proposition 3.9. Using the preceding discussion, we fix $J \geq 1$ such that the sequence $z \mapsto (C_n(z))_{n\geq J}$ does not have any zero on $\mathscr{E}$, so that $z \mapsto (M_n(z))_{n\geq J}$ is well-defined for all $z \in \mathscr{E}$.

We introduce the following notation. Let $F(z,n)$ and $G(z,n)$ be two functions of a complex parameter z and an integer $n \in \mathbb{N}$. For $D \subset \mathbb{C}$ a domain of the complex plane we write

$$F(n,z) = O_D(G(n,z)) \qquad \text{(resp.} \qquad F(n,z) = o_D(G(n,z))) \tag{3.22}$$

to express the fact that $F(n,z)$ is a big (resp. small) o of $G(n,z)$ as $n \to \infty$, uniformly on all compact $K \subset D$.

Now, let us derive some information on the asymptotic behaviour of $C_n(z)$.

Lemma 3.11. *Suppose that* $\mathbf{w}$ *satisfies* ($\square_\gamma$). *Then there exists* $\epsilon > 0$ *and an analytic function* $z \mapsto c(z)$ *on* $\mathscr{E}$ *such that*

$$C_n(z) = \exp\bigl(\phi(z)\log n + c(z) + O_{\mathscr{E}}(n^{-\epsilon})\bigr).$$

Remark that the lemma implies that for any $z \in \mathscr{E}$, we have

$$|C_n(z)| \sim e^{\mathrm{Re}(c(z))} \cdot n^{\mathrm{Re}\,\phi(z)}$$

as $n \to \infty$. It is also immediate that $\mathbb{E}\left[\sum_{k=1}^n \frac{w_k}{W_n} e^{z\,\mathrm{ht}(u_k)}\right] = \mathbb{E}\left[M_J(z)\right] \cdot C_n(z)$ satisfies the same asymptotics up to a constant, as soon as z is such that $\mathbb{E}\left[M_J(z)\right] \neq 0$.

Before proving the lemma, we state the following result which follows from elementary calculus. Its proof can be found in the appendix.

Lemma 3.12. *Suppose that* (w_n) *satisfies* ($\square_\gamma$). *Then there exists* ϵ *such that*

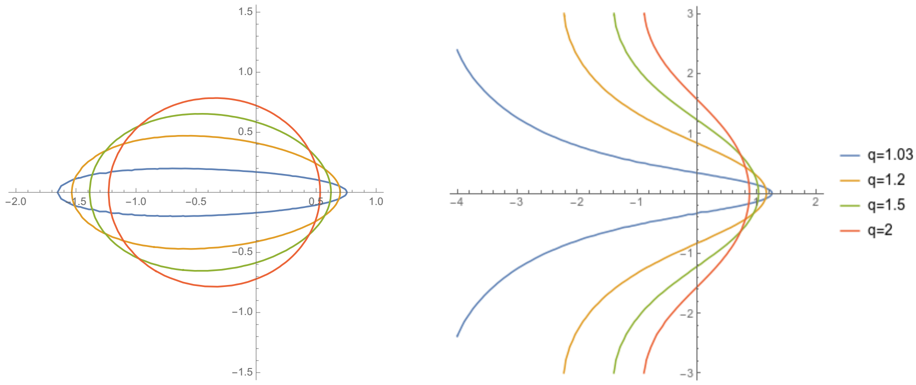
$$\sum_{i=n}^{+\infty}\left(\frac{w_i}{W_i}\right)^2 = O\bigl(n^{-\epsilon}\bigr) \quad \textit{and also} \quad \sum_{i=1}^{n}\frac{w_i}{W_i} = \gamma\log n + \mathrm{cst} + O\bigl(n^{-\epsilon}\bigr)\,.$$

Proof of Lemma 3.11. For any $z \in \mathbb{C}\setminus(-\infty\,, -1]$ we write $\mathrm{Log}(1+z)$ for a complex determination of the logarithm which coincides with $\sum_{i=1}^{\infty}\frac{(-1)^{n-1}}{n}z^n$ near 0. If we let

$$h(i,z) = \mathrm{Log}\left(1 + (e^z - 1)\frac{w_i}{W_i}\right) - (e^z - 1)\frac{w_i}{W_i},$$

then $|h(i,z)| = O_{\mathscr{E}}\left(\left(\frac{w_i}{W_i}\right)^2\right)$ is summable in i and the rest of the series is

$$\left|\sum_{i=n}^{\infty} h(i,z)\right| \leq \sum_{i=n}^{\infty}|h(i,z)| = O_{\mathscr{E}}\left(\sum_{i=n}^{\infty}\left(\frac{w_i}{W_i}\right)^2\right) = O_{\mathscr{E}}\bigl(n^{-\epsilon}\bigr)\,, \tag{3.23}$$



(a) $\gamma = 2$ (b) $\gamma = \frac{1}{2}$

Figure 3.1 – The boundary of $\mathscr{V}_q$ for some values of $q \in (1\,,2]$, plotted for $\gamma = 2$ and $\gamma = \frac{1}{2}$.

for some $\epsilon > 0$, thanks to Lemma 3.12. Then we write

$$C_n(z) = \prod_{i=J}^{n} \left(1 + (e^z - 1)\frac{w_i}{W_i}\right) = \exp\left((e^z - 1)\sum_{i=J}^{n} \frac{w_i}{W_i} + \sum_{i=J}^{n} h(i,z)\right)$$

which yields using (3.23) and Lemma 3.12

$$\begin{aligned} C_n(z) &= \exp\left((e^z - 1)(\gamma \log n + \mathrm{cst} + O_{\mathscr{E}}(n^{-\epsilon})) + \sum_{i=J}^{\infty} h(i,z) - \sum_{i=n+1}^{\infty} h(i,z)\right) \\ &= \exp\left(\phi(z)\log n + \underbrace{(e^z - 1)\cdot \mathrm{cst} + \sum_{i=J}^{\infty} h(i,z)}_{c(z)} + O_{\mathscr{E}}(n^{-\epsilon})\right), \end{aligned}$$

and $c(z)$ is an analytic function of z, which finishes the proof. $\square$

Convergence of the martingales $(M_n(z))_{n\geq 1}$. When the parameter z is a positive real number, the sequence $(M_n(z))_{n\geq 1}$ is a positive martingale and so it converges almost surely to some limit. We want to prove that these martingales converge almost surely and in L^1 for the largest possible range of parameters z. We recall that the weight sequence $\mathbf{w}$ satisfies $(\square_\gamma^p)$ for some fixed parameters $\gamma > 0$ and $p \in (1\,,2]$. We align our notation with that used in [36, Theorem 2.2] which states something similar to our forthcoming Proposition 3.14 for another model, the binary search tree.

For any $z \in \mathscr{E}$ and $q \in (1\,,p]$, we let

$$g(z,q) := \phi(q \operatorname{Re} z) - q \operatorname{Re}(\phi(z)) - q + 1 = \gamma(e^{q \operatorname{Re} z} - 1 - q \operatorname{Re}(e^z) + q) - q + 1. \tag{3.24}$$

For any $q \in (1\,,p]$, let $\mathscr{V}_q = \{z \in \mathscr{E} \mid g(z,q) < 0\}$, and denote

$$\mathscr{V} = \bigcup_{1<q\leq p} \mathscr{V}_q.$$

Lemma 3.13. *The domain $\mathscr{V}$ is an open domain of the complex plane and contains the open interval $I_\gamma := \{x \in \mathbb{R} \mid \gamma(xe^x - e^x + 1) - 1 < 0\}$ which contains* 0.

Proof. Of course $\mathscr{V}$ is open as a union of open sets. For any real x we have $g(x,1)=0$. So, if $\frac{\partial g}{\partial q}(x,1)<0$ then there exists $q>1$ for which $g(x,q)<0$. Since $\frac{\partial g}{\partial q}(x,1)=\gamma(xe^x-e^x+1)-1$, the set $\mathscr{V}$ contains the interval I_γ defined above. Since $\frac{\partial g}{\partial q}(0,1)=-1<0$, we have $0\in I_\gamma$. $\square$

Proposition 3.14. *The sequence of functions* $(z\mapsto M_n(z))_{n\geq J}$ *converges uniformly almost surely and in* L^1 *towards an analytic function* $z\mapsto M_\infty(z)$ *on every compact of* $\mathscr{V}$*. Furthermore, for any compact* $K\subset\mathscr{V}$*, there exists a real* $\epsilon(K)>0$ *such that almost surely*

$$|M_n(z)-M_\infty(z)|=O_K\left(n^{-\epsilon(K)}\right).$$

The proof of the proposition will follow from the next lemma, together with Lemma 3.34, stated in the appendix.

Lemma 3.15. *For any* $q\in(1\,,p]$,

$$\mathbb{E}\left[|M_n(z)|^q\right]=O_{\mathscr{E}}\left(n^{0\vee g(z,q)+o_{\mathscr{E}}(1)}\right), \tag{3.25}$$

and also

$$\mathbb{E}\left[|M_{2n}(z)-M_n(z)|^q\right]=O_{\mathscr{E}}\left(n^{(1-q)\vee g(z,q)+o_{\mathscr{E}}(1)}\right). \tag{3.26}$$

Proof. For any $q\in(1\,,p]$ and $n\geq J$, we write

$$M_{n+1}(z)-M_n(z)=M_n(z)\left(\frac{C_n(z)}{C_{n+1}(z)}-1\right)+\frac{1}{C_{n+1}(z)}\cdot\frac{w_{n+1}}{W_{n+1}}\cdot e^{z\,\mathrm{ht}(u_{n+1})}.$$

Taking the q-th power of the modulus on both sides and using the inequality $|a+b|^q\leq 2^q|a|+2^q|b|$, we get

$$\begin{aligned}&\mathbb{E}\left[|M_{n+1}(z)-M_n(z)|^q\right]\\&\leq\mathbb{E}\left[|M_n(z)|^q\right]\cdot 2^q\left|\frac{C_n(z)}{C_{n+1}(z)}-1\right|^q+2^q\frac{1}{|C_{n+1}(z)|^q}\left(\frac{w_{n+1}}{W_{n+1}}\right)^q\cdot\mathbb{E}\left[|e^z|^{q\,\mathrm{ht}(u_{n+1})}\right].\end{aligned} \tag{3.27}$$

Using Lemma 3.33 in the appendix, we have for any $n\geq J$,

$$\mathbb{E}\left[|M_{n+1}(z)|^q\right]\leq\mathbb{E}\left[|M_n(z)|^q\right]+2^q\cdot\mathbb{E}\left[|M_{n+1}(z)-M_n(z)|^q\right].$$

Using the last display and equation (3.27), we get a recurrence inequality of the form

$$\mathbb{E}\left[|M_{n+1}(z)|^q\right]\leq(1+a_n(z))\cdot\mathbb{E}\left[|M_n(z)|^q\right]+b_n(z), \tag{3.28}$$

where

$$a_n(z)=2^{2q}\left|\frac{C_n(z)}{C_{n+1}(z)}-1\right|^q\quad\text{and}\quad b_n(z)=2^{2q}\frac{1}{|C_{n+1}(z)|^q}\left(\frac{w_{n+1}}{W_{n+1}}\right)^q\cdot\mathbb{E}\left[|e^z|^{q\,\mathrm{ht}(u_{n+1})}\right].$$

Applying (3.28) in cascade we get

$$\mathbb{E}\left[|M_n(z)|^q\right]\leq\mathbb{E}\left[|M_J(z)|^q\right]\cdot\prod_{i=J}^{n-1}(1+a_i(z))\cdot\left(\sum_{i=J}^{n-1}b_i(z)\right). \tag{3.29}$$

Now notice that from our assumption on the sequence $(w_n)_{n\geq 1}$ we have

$$a_n(z) = 2^{2q}\left|\frac{C_n(z)}{C_{n+1}(z)} - 1\right|^q = 2^{2q}\left|\frac{1}{1+(e^z-1)\frac{w_{n+1}}{W_{n+1}}} - 1\right|^q = O_{\mathscr{E}}\left(\left(\frac{w_{n+1}}{W_{n+1}}\right)^q\right). \tag{3.30}$$

On the other hand, thanks to Lemma 3.11 we have

$$\begin{aligned} b_n(z) &= \mathrm{cst}\cdot\left(\frac{w_{n+1}}{W_{n+1}}\right)^q \cdot |C_{n+1}(z)|^{-q} \cdot e^{q\operatorname{Re} z} \cdot \mathbb{E}\left[\sum_{k=1}^{n} \frac{w_k}{W_n} e^{(q\operatorname{Re} z)\,\mathrm{ht}(u_k)}\right] \\ &= \left(\frac{w_{n+1}}{W_{n+1}}\right)^q \cdot O_{\mathscr{E}}\left(n^{-q\operatorname{Re}(\phi(z))}\right) \cdot O_{\mathscr{E}}\left(n^{\phi(q\operatorname{Re} z)}\right) \\ &= \left(\frac{w_{n+1}}{W_{n+1}}\right)^q \cdot O_{\mathscr{E}}\left(n^{g(z,q)-1+q}\right). \end{aligned} \tag{3.31}$$

We conclude using the following lemma which is an application of Hölder inequality using the assumption ($\square_\gamma^p$)

Lemma 3.16. *For any $q \in (1\,, p]$ we have* $\sum_{i=n}^{2n}\left(\frac{w_i}{W_i}\right)^q \leq n^{1-q+o(1)}$.

Together with (3.30), this proves that $(a_n(z))_{n\geq 1}$ is summable and so $\prod_{i=J}^{\infty}(1+a_i(z)) = O_{\mathscr{E}}(1)$. Also

$$\sum_{i=n}^{2n} b_i(z) = O_{\mathscr{E}}\left(n^{g(z,q)+o_{\mathscr{E}}(1)}\right),$$

and so $\sum_{i=J}^{n} b_i(z) = O_{\mathscr{E}}\left(n^{0\vee g(z,q)+o_{\mathscr{E}}(1)}\right)$. Replacing this in (3.29) finishes to prove (3.25). In order to prove (3.26), we use Lemma 3.33 again and write

$$\begin{aligned} \mathbb{E}\left[|M_{2n}(z)-M_n(z)|^q\right] &\leq 2^q \cdot \sum_{i=n}^{2n-1} \mathbb{E}\left[|M_{i+1}(z)-M_i(z)|^q\right] \\ &\underset{(3.25),(3.27)}{\leq} \sum_{i=n}^{2n-1}\left(a_i(z)\cdot O_{\mathscr{E}}\left(n^{0\vee g(z,q)+o_{\mathscr{E}}(1)}\right) + b_i(z)\right) \\ &\underset{(3.30),(3.31)}{\leq} \sum_{i=n}^{2n-1}\left(\frac{w_{n+1}}{W_{n+1}}\right)^q\left(O_{\mathscr{E}}\left(n^{0\vee g(z,q)+o_{\mathscr{E}}(1)}\right) + O_{\mathscr{E}}\left(n^{g(z,q)-1+q}\right)\right). \end{aligned}$$

Using Lemma 3.16 we get $\mathbb{E}\left[|M_{2n}(z)-M_n(z)|^q\right] = O_{\mathscr{E}}\left(n^{(1-q)\vee g(z,q)+o_{\mathscr{E}}(1)}\right)$ which finishes the proof of the lemma. ☐

Proof of Proposition 3.14. Any compact $K \subset \mathscr{V}_q$ can be covered by a finite number of $\mathscr{V}_q$. The convergence result is then an application of Lemma 3.34, stated in the appendix, on the domain $\mathscr{V}_q$ with $\alpha(z) = 0$ and, say $\delta(z) = -\frac{1}{2}g(z,q) > 0$. The limiting function is analytic as a uniform limit of analytic functions. ☐

Zeros of the limit. Now that we have proved that their exists a limiting function $z \mapsto M_\infty(z)$ defined on the domain $\mathscr{V}$, we are interested in the possible location of the zeros of this random function. In fact, the function $z \mapsto M_\infty(z)$ is related to the function $z \mapsto N_\infty(z)$ of Proposition 3.9, for which we aim to prove that it has almost surely no zero on some real interval $(z_-\,, z_+)$ which contains 0. We will prove a similar result for $z \mapsto M_\infty(z)$ in Lemma 3.19, and we start by proving the following weaker statement.

Lemma 3.17. *For all $z \in \mathscr{V} \cap \mathbb{R}$, we have almost surely $M_\infty(z) > 0$. As a consequence, the number of zeros on every compact of $\mathscr{V}$ is almost surely finite.*

Proof. This follows from an application of Kolmogorov's $0-1$ law. Indeed, fix $N \geq J$ and $z \in \mathscr{V} \cap \mathbb{R}$ and for all $n \geq N$, let

$$M_n^{(N)}(z) = \frac{1}{C_n(z)} \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{d}(u_i, \mathtt{T}_N)}.$$

Now remark the following:

(i) $(M_n^{(N)}(z))_{n \geq N}$ is a positive martingale which satisfies the same assumptions as $M_n(z)$ so it converges a.s. and in L^1 towards a non-negative limit, $M_\infty^{(N)}(z)$.

(ii) We have $(1 \wedge e^z)^N M_n^{(N)}(z) \leq M_n(z) \leq (1 \vee e^z)^N M_n^{(N)}(z)$.

(iii) The sequence $(M_n^{(N)}(z))_{n \geq N}$, hence its limit $M_\infty^{(N)}(z)$, is independent of the N first steps of the construction, coded by the vector $(K_2, \ldots, K_N)$.

Using all these observations we deduce that for any $N \geq J$ we have the equality of events $\{M_\infty(z) > 0\} = \{M_\infty^{(N)}(z) > 0\}$. This proves that $\{M_\infty(z) > 0\}$ is measurable with respect to the tail σ-algebra generated by the sequence $(K_2, K_3, \ldots)$, which are independent, and has hence probability 0 or 1. By L^1 convergence we have $\mathbb{E}[M_\infty(z)] = \mathbb{E}[M_J(z)] > 0$ and this proves our claim. It follows immediately that the number of zeros of the limit $z \mapsto M_\infty(z)$ on any compact $K \subset \mathscr{V}$ is almost surely finite, because otherwise the function would be identically 0 with positive probability. $\square$

Remark 3.18. *In fact, Lemma 3.17 is already sufficient to prove the almost-sure lower bound* $\liminf_{n\to\infty} \frac{\mathrm{ht}(\mathtt{T}_n)}{\log n} \geq \gamma e^{z_+}$ *using arguments that are standard for branching random walks. Indeed, under the probability measure $\mathbb{P}^z$ which has density $M_\infty(z)$ with respect to $\mathbb{P}$, there is a natural description of the tree $\mathtt{T}_n$ together with a distinguished vertex u_I chosen among $\{u_i \mid 1 \leq i \leq n\}$ with probability proportional to $w_i e^{z\,\mathrm{ht}(u_i)}$ conditionally on $\mathtt{T}_n$, and the height of u_I has the description as a sum of independent Bernoulli variables and is concentrated around $\gamma e^z \log n$. Since $\mathbb{P}^z$ and $\mathbb{P}$ are equivalent for any $z \in \mathscr{V} \cap \mathbb{R}$ thanks to Lemma 3.17, this also proves the existence in $\mathtt{T}_n$ of vertices of height $\gamma e^z \log n$ with high probability under $\mathbb{P}$.*

Let us now prove the stronger statement:

Lemma 3.19. *The function $M_\infty(z)$ almost surely has no zero on $\mathscr{V} \cap \mathbb{R}$.*

In order to prove this lemma, we use an argument of self-similarity: essentially, if we take two vertices u_i and u_j in the tree, then conditionally on the sequences of vertices that are grafted above u_i or above u_j, the subtrees above u_i and u_j evolve as two *independent* weighted recursive trees. Using Proposition 3.14 and Lemma 3.17, the normalized Laplace transform of the weighted profile of each of those two subtrees should converge to some random analytic function which is non-negative on $\mathscr{V} \cap \mathbb{R}$ and has at most countably many zeros. Since the two are independent, their zeros should not overlap and hence the sum of their contribution should result in a function that is positive on $\mathscr{V} \cap \mathbb{R}$.

Proof. Let us formalise this line of reasoning. Using Theorem 3.4, we know that the measure μ on $\partial\mathbb{U}$ is almost surely diffuse, hence we can define

$$I^{(1)} := \inf\{i \geq 1 \mid \mu(T(u_i)) \in (0,1)\} \quad \text{and} \quad I^{(2)} := \inf\{i \geq I_1 \mid u_{I_1} \npreceq u_i \text{ and } \mu(T(u_i)) \in (0,1)\},$$

and they are almost surely finite. Let us consider the sequences $\left(\mathbf{1}_{\{u_{I^{(j)}}\preceq u_n\}}\right)_{n\geq 1}$ for $j=1,2$, which record the times when a vertex is added to $T(u_{I^{(1)}})$ or $T(u_{I^{(2)}})$, and work conditionally on them. Thanks to our definition of $I^{(1)}$ and $I^{(2)}$, we know that the number of vertices in each of those subtrees will grow linearly in time (in particular, they go to infinity). We let

$$\forall n\geq 1,\quad N_n^{(j)} := \sum_{i=1}^{n}\mathbf{1}_{\{u_{I^{(j)}}\preceq u_i\}}\qquad\text{and}\qquad \forall k\geq 1,\quad \tau_k^{(j)} := \inf\left\{n\geq 1 \;\middle|\; N_n^{(j)}\geq k\right\},$$

which record respectively the number of vertices in $T(u_{I^{(j)}})$ at time n and conversely, the time when the k-th vertex is added. We let $w_k^{(j)} := w_{\tau^{(j)}(k)}$ and $W_k^{(j)} := \sum_{i=1}^{k} w_k^{(j)}$, and also $u_{\tau_k^{(j)}} = u_k^{(j)}$. We also define for $j=1,2$ and $k\geq 1$

$$\mathtt{T}_k^{(j)} := \left\{u\in\mathbb{U}\;\middle|\; u_{I^{(j)}}u\in\mathtt{T}_{\tau_k^{(j)}}\right\},$$

the subtree hanging above $u_{I^{(j)}}$ at the time where it contains exactly k vertices (translated to the origin in order to be considered as a plane tree). Let us state the following intermediate result, which we will prove at the end of the section.

Lemma 3.20. *The following holds.*

(i) *For $j=1,2$, we have $N_n^{(j)} \underset{n\to\infty}{\sim} \mu(T(u_{I^{(j)}}))\cdot n$ almost surely.*

(ii) *For $j=1,2$, the sequence $(w_k^{(j)})_{k\geq 1}$ satisfies $(\square_\gamma^p)$ almost surely.*

(iii) *Conditionally on the two sequences $\left(\mathbf{1}_{\{u_{I^{(1)}}\preceq u_n\}}\right)_{n\geq 1}$ and $\left(\mathbf{1}_{\{u_{I^{(2)}}\preceq u_n\}}\right)_{n\geq 1}$, the sequences of trees $(\mathtt{T}_k^{(1)})_{k\geq 1}$ and $(\mathtt{T}_k^{(2)})_{k\geq 1}$ are independent and have respective distributions $\mathrm{WRT}((w_k^{(1)})_{k\geq 1})$ and $\mathrm{WRT}((w_k^{(2)})_{k\geq 1})$.*

Recall the discussion before Lemma 3.10 and fix $J'\geq 1$ such that for $j=1,2$, for all $k\geq J'$ and for all $z\in\mathscr{E}$ we have $1+(e^z-1)\frac{w_k^{(j)}}{W_k^{(j)}}\neq 0$. Then we can define for $j=1,2$ and $k\geq J'$,

$$M_k^{(j)}(z) := \frac{1}{C_k^{(j)}(z)}\sum_{i=1}^{k}\frac{w_i^{(j)}}{W_k^{(j)}}e^{zd(u_{I^{(j)}},u_i^{(j)})}\quad\text{with}\quad C_k^{(j)}(z) := \prod_{i=J'}^{k}\left(1+(e^z-1)\frac{w_i^{(j)}}{W_i^{(j)}}\right).$$

These processes are the martingales associated to the weighted profile of the trees $(\mathtt{T}_k^{(j)})_{k\geq 1}$, to which we can apply the result of Proposition 3.14 thanks to Lemma 3.20. This entails that these two sequences of functions converge almost surely to analytic limits $z\mapsto M_\infty^{(j)}(z)$ on the domain $\mathscr{V}$. Now we can write, for n sufficiently large

$$\begin{aligned}M_n(z) &= \frac{1}{C_n(z)}\sum_{i=1}^{n}\frac{w_i}{W_n}e^{z\,\mathrm{ht}(u_i)}\\ &\geq \frac{C_{N_n^{(1)}}^{(1)}(z)\cdot W_{N_n^{(1)}}^{(1)}}{C_n(z)\cdot W_n}\cdot e^{z\,\mathrm{ht}(u_{I^{(1)}})}\cdot M_{N_n^{(1)}}^{(1)}(z) + \frac{C_{N_n^{(2)}}^{(2)}(z)\cdot W_{N_n^{(2)}}^{(2)}}{C_n(z)\cdot W_n}\cdot e^{z\,\mathrm{ht}(u_{I^{(2)}})}\cdot M_{N_n^{(2)}}^{(2)}(z).\end{aligned}\tag{3.32}$$

Using Lemma 3.11, we have almost surely for $j=1,2$,

$$C_k^{(j)}(z) = \exp\Big(\phi(z)\log k + c^{(j)}(z) + o_{\mathscr{E}}(1)\Big).$$

Using the asymptotics $N_n^{(j)} \underset{n\to\infty}{=} \mu(T(u_{I(j)})) \cdot n \cdot (1+o(1))$ from Lemma 3.20(i) we get

$$C^{(j)}_{N_n^{(j)}}(z) = \exp\Big(\phi(z)\log n + \phi(z)\log(\mu(T(u_{I(j)}))) + c^{(j)}(z) + o_{\mathscr{E}}(1)\Big).$$

From the a.s. convergence of the sequence of measures $(\mu_n)_{n\geq 1}$, see Theorem 3.4, we also get $\frac{W^{(j)}_{N_n^{(j)}}}{W_n} = \mu_n(T(u_{I(j)})) \underset{n\to\infty}{\to} \mu(T(u_{I(j)}))$, which entails that for $j=1,2$, uniformly on all compacts included in $\mathscr{E}$, we have the a.s. convergence

$$\frac{W^{(j)}_{N_n^{(j)}}}{W_n} \cdot \frac{C^{(j)}_{N_n^{(j)}}(z)}{C_n(z)} \underset{n\to\infty}{\to} A_j(z),$$

where the limiting function $z\mapsto A_j(z)$ is analytic (as the uniform limit of analytic functions) and only takes positive values on $\mathscr{E}\cap\mathbb{R}$ (as the exponential of some real-valued function). Then, for any $z\in\mathscr{V}\cap\mathbb{R}$, taking the limit $n\to\infty$ in (3.32) yields

$$M_\infty(z) \geq e^{z\,\mathrm{ht}(u_{I(1)})}\cdot A_1(z)\cdot M_\infty^{(1)}(z) + e^{z\,\mathrm{ht}(u_{I(2)})}\cdot A_2(z)\cdot M_\infty^{(2)}(z).$$

Now, thanks to Lemma 3.17, the function $z\mapsto M_\infty^{(1)}(z)$ can only have at most countably many zeros on $\mathscr{V}\cap\mathbb{R}$ and for all $z\in\mathscr{V}\cap\mathbb{R}$, we have $M_\infty^{(2)}(z)>0$ almost surely. Then if we condition on the location of the zeros $z_1,z_2\ldots$ of $M_\infty^{(1)}$ on $\mathscr{V}\cap\mathbb{R}$, since $M_\infty^{(2)}$ is independent of $z_1,z_2\ldots$, we have $M_\infty^{(2)}(z_i)>0$ for all $i\geq 1$ almost surely. Hence M_∞ has almost surely no zeros on $\mathscr{V}\cap\mathbb{R}$. $\square$

Now let us prove Lemma 3.20 which we used in the preceding proof.

Proof of Lemma 3.20. Point (i) follows just from Theorem 3.4 and the fact that for $j=1,2$ we have $N_n^{(j)} = n\nu_n(T(u_{I(j)}))$. Point (iii) is obvious from the attachment dynamics. We just need to prove (ii). In order to do that we are going to prove that for $j=1,2$, we have

$$\mu_n(T(u_{I(j)})) \underset{n\to\infty}{\bowtie} \mu(T(u_{I(j)})) \qquad \text{and} \qquad \tau_n^{(j)} \underset{n\to\infty}{\bowtie} \mu(T(u_{I(j)}))^{-1}\cdot n. \tag{3.33}$$

Let us conclude from here: using the fact that $\mathbf{w}$ satisfies ($\square_\gamma$), we get

$$W_n^{(j)} = W_{\tau_n^{(j)}}\cdot \mu_{\tau_n^{(j)}}(T(u_{I(j)})) \underset{n\to\infty}{\bowtie} \mathrm{cst}\cdot(\tau_n^{(j)})^\gamma\cdot\mu(T(u_{I(j)})) \underset{n\to\infty}{\bowtie} \mathrm{cst}\cdot n^\gamma,$$

with a positive constant. Then we also have

$$\sum_{k=n}^{2n}(w_k^{(j)})^p = \sum_{k=n}^{2n}(w_{\tau_k^{(j)}})^p \leq \sum_{i=\tau_n^{(j)}}^{\tau_{2n}^{(j)}} w_i^p \leq n^{1+(\gamma-1)p+o(1)},$$

where the last inequality is due to the linear growth of $\tau_n^{(j)}$ and the fact that $\mathbf{w}$ satisfies ($\square_\gamma^p$).

So it remains only to prove (3.33). Recall the proof of Theorem 3.4. For all $k\geq 1$ the process $(\mu_n(T(u_k)))_{n\geq k}$ is a martingale and almost surely we have $|\mu_{n+1}(T(u_k))-\mu_n(T(u_k))|\leq \frac{w_{n+1}}{W_{n+1}}$, hence using Lemma 3.33 in the appendix we get

$$\mathbb{E}\left[|\mu_{2n}(T(u_k))-\mu_n(T(u_k))|^p\right] \leq 2^p\cdot\sum_{i=n+1}^{2n}\left(\frac{w_i}{W_i}\right)^p = O\Big(n^{1-p+o(1)}\Big).$$

Using then Lemma 3.34 with $q=p$ and $\alpha=0$ and $\delta=(p-1)/2$, we get that $|\mu_n(T(u_k))-\mu(T(u_k))| = O(n^{-\epsilon})$ almost surely for some $\epsilon>0$. Since this is true almost surely for all $k \geq 1$, we use it with $k \in \{I^{(1)}, I^{(2)}\}$. As by definition for $j = 1, 2$ we have $\mu(T(u_{I^{(j)}})) > 0$, we conclude that $\mu_n(T(u_{I^{(j)}})) \underset{n\to\infty}{\bowtie} \mu(T(u_{I^{(j)}}))$.

Then, for any $k \geq 1$, consider the process $\big(n\nu_n(T(u_k)) - \sum_{i=k+1}^{n} \mu_i(T(u_k))\big)_{n\geq k}$. It is easy to verify that this process is a martingale in its own filtration and that its increments are bounded by 1. Using again Lemma 3.34 with $q = 2$ and $\alpha = 1$ and $\delta = 1$, we get $n^{-1}\big|n\nu_n(T(u_k)) - \sum_{i=k+1}^{n} \mu_i(T(u_k))\big| = O(n^{-\epsilon})$. Using again that for $j = 1, 2$ the limit $\mu(T(u_{I^{(j)}}))$ is almost surely positive, we can write $N_n^{(j)} = n\nu_n(T(u_{I^{(j)}})) \underset{n\to\infty}{\bowtie} \mu(T(u_{I^{(j)}})) \cdot n$. Using the definition of $\tau_n^{(j)}$, we can check that this entails that $\tau_n^{(j)} \underset{n\to\infty}{\bowtie} \mu(T(u_{I^{(j)}}))^{-1} \cdot n$ almost surely. This concludes the proof of (3.33) and so the lemma is proved.

□

3.3.2 From the weighted to the unweighted sum.

Now we want to transfer these results of convergence to the Laplace transform of the real profile. In this aim, we introduce the following quantity, for $n \geq J$,

$$\begin{aligned} X_n(z) &:= n^{1+\phi(z)} \cdot N_n(z) - e^z \sum_{k=J}^{n-1} C_k(z) M_k(z) \\ &= \sum_{i=1}^{n} e^{z\,\mathrm{ht}(u_i)} - e^z \sum_{k=J}^{n-1} \left(\sum_{i=1}^{k} \frac{w_i}{W_k} e^{z\,\mathrm{ht}(u_i)} \right) \end{aligned}$$

The goal of this subsection is to show that the quantity $X_n(z)$ is negligible as $n \to \infty$ compared to any of the two terms in the difference, for z contained in some domain. This way we will transfer the asymptotics that we have proved for $M_n(z)$ and $C_n(z)$ in the last section to asymptotics for $N_n(z)$, which is the quantity that we want to study in the end.

Recall the definition of z_+ and z_- in (3.6). Let us define the domain $\mathscr{D}$ to which we refer in the statement of Proposition 3.9 as

$$\mathscr{D} := \mathscr{V} \cap \{z \in \mathbb{C} \mid 1 + \mathrm{Re}(\phi(z)) > 0\}\,.$$

In this way $\mathscr{D}$ is a connected domain of $\mathbb{C}$ and $\mathscr{D} \cap \mathbb{R} = (z_-\,, z_+)$. Indeed, recall from Lemma 3.13 that $\mathscr{V} \cap \mathbb{R} = I_\gamma = \{x \in \mathbb{R} \mid 1 + \gamma(e^x - 1 - xe^x) > 0\}$ is an open interval which contains 0 and has z_+ as its right bound. Now just check that $\{z \in \mathbb{R} \mid 1 + \mathrm{Re}(\phi(z)) > 0\} = (z_-\,, \infty)$ and that $z_- \in I_\gamma$.

For technical reasons, we also introduce

$$\mathscr{D}' = (z_-\,, z_+) \times (0\,, 2\pi),$$

on which the process $(z \mapsto M_n(z))_{n\geq J}$, and hence also $(z \mapsto X_n(z))_{n\geq J}$, is well-defined. Let us further decompose $\mathscr{D}'$ into a union of open sets

$$\mathscr{D}' = \bigcup_{1<q\leq p} \mathscr{D}'_q \qquad \text{where} \qquad \mathscr{D}'_q = \big\{z \in \mathscr{D}' \;\big|\; g(\mathrm{Re}\, z, q) < 0\big\}\,,$$

and the function $g : \mathscr{E} \times (1\,, p] \to \mathbb{R}$ is the one defined in (3.24).

Lemma 3.21. *The process $(X_n(z))_{n\geq J}$ is a martingale with respect to the filtration generated by $(\mathtt{T}_n)_{n\geq 1}$. Furthermore, for all $q \in (1\,, p]$,*

$$\mathbb{E}\left[|X_{2n}(z) - X_n(z)|^q\right] = O_{\mathscr{E}}\Big(n^{1+(q\,\mathrm{Re}(\phi(z)) \vee \phi(q\,\mathrm{Re}\, z)) + o_{\mathscr{E}}(1)}\Big)\,.$$

Proof. This process is of course $(\sigma(\mathtt{T}_n))$-adapted and integrable. For the martingale property we compute

$$\begin{aligned}\mathbb{E}\left[X_{n+1}(z) \mid \mathtt{T}_n\right] &= \mathbb{E}\left[X_n(z) - e^z C_n(z) M_n(z) + e^{z\,\mathrm{ht}(u_{n+1})} \,\middle|\, \mathtt{T}_n\right] \\ &= X_n(z) - e^z C_n(z) M_n(z) + e^z \sum_{i=1}^{n} \frac{w_i}{W_n} e^{z\,\mathrm{ht}(u_i)} = X_n(z).\end{aligned}$$

For $z \in \mathscr{E}$ and $q \in (1\,, p]$, we make the following computation, using Lemma 3.11 and Lemma 3.15,

$$\begin{aligned}\mathbb{E}\left[|X_{n+1}(z) - X_n(z)|^q\right] &= \mathbb{E}\left[\left|-e^z C_n(z) M_n(z) + e^{z\,\mathrm{ht}(u_{n+1})}\right|^q\right] \\ &\leq 2^q \cdot \left(e^{qz}|C_n(z)|^q \mathbb{E}\left[|M_n(z)|^q\right] + e^{q\,\mathrm{Re}\,z}\mathbb{E}\left[\sum_{i=1}^{n} \frac{w_i}{W_n} e^{\mathrm{ht}(u_i) q\,\mathrm{Re}\,z}\right]\right), \\ &= O_{\mathscr{E}}\left(n^{q\,\mathrm{Re}\,\phi(z) + 0\vee g(z,q) + o_{\mathscr{E}}(1)}\right) + O_{\mathscr{E}}\left(n^{\phi(q\,\mathrm{Re}\,z)}\right) \\ &= O_{\mathscr{E}}\left(n^{q\,\mathrm{Re}\,\phi(z) \vee (q\,\mathrm{Re}\,\phi(z) + g(z,q)) \vee \phi(q\,\mathrm{Re}\,z) + o_{\mathscr{E}}(1)}\right),\end{aligned}$$

and the last exponent reduces to $q\,\mathrm{Re}\,\phi(z) \vee \phi(q\,\mathrm{Re}\,z)$ because $(q\,\mathrm{Re}\,\phi(z) + g(z,q)) = \phi(q\,\mathrm{Re}\,z) + 1 - q < \phi(q\,\mathrm{Re}\,z)$. Hence, using Lemma 3.33,

$$\mathbb{E}\left[|X_{2n}(z) - X_n(z)|^q\right] \leq 2^q \sum_{i=n}^{2n} \mathbb{E}\left[|X_{i+1}(z) - X_i(z)|^q\right] = O_{\mathscr{E}}\left(n^{1 + (q\,\mathrm{Re}(\phi(z)) \vee \phi(q\,\mathrm{Re}\,z)) + o_{\mathscr{E}}(1)}\right),$$

which finishes the proof of the lemma. □

Lemma 3.22. *The following holds:*

(i) *For all compact $K \subset \mathscr{D}$ there exists $\epsilon(K) > 0$ such that almost surely*

$$n^{-(1+\mathrm{Re}\,\phi(z))} \cdot |X_n(z)| = O_K\left(n^{-\epsilon(K)}\right).$$

(ii) *For all compact $K \subset \mathscr{D}'$, there exists $\epsilon(K) > 0$ such that*

$$n^{-(1+\phi(\mathrm{Re}\,z))} \cdot \left|\sum_{i=J}^{n-1} C_i(z) M_i(z)\right| = O_K\left(n^{-\epsilon(K)}\right).$$

(iii) *For all compact $K \subset \mathscr{D}'$, there exists $\epsilon(K) > 0$ such that almost surely*

$$n^{-(1+\phi(\mathrm{Re}\,z))} \cdot |X_n(z)| = O_K\left(n^{-\epsilon(K)}\right).$$

Proof. For (i), for any $q \in (1\,, p]$ we can apply Lemma 3.34 on the domain $\mathscr{V}_q \cap \{z \in \mathbb{C} \mid 1 + \mathrm{Re}(\phi(z)) > 0\}$ with $\alpha(z) = 1 + \mathrm{Re}(\phi(z)) > 0$ and $\delta(z) = \min(q-1, -g(z,q)) > 0$, thanks to Lemma 3.21. Then using the compactness property, (i) is true for every compact $K \subset \mathscr{D}$.

Let us prove point (ii). For any $q \in (1\,, p]$, on the domain $\mathscr{D}'_q$ we have $\mathbb{E}\left[|M_{2n}(z) - M_n(z)|^q\right] = O_{\mathscr{D}'}\left(n^{(1-q)\vee g(z,q) + o_{\mathscr{D}'_q}(1)}\right)$ and

$$g(z,q) = q\underbrace{(\phi(\mathrm{Re}\,z) - \mathrm{Re}\,\phi(z))}_{>0} + \underbrace{g(\mathrm{Re}\,z, q)}_{<0}.$$

Applying Lemma 3.34 for the martingale $(z \mapsto M_n(z))_{n\geq J}$ on any compact $K \subset \mathscr{D}'_q$ with $\alpha(z) = \phi(\operatorname{Re} z) - \operatorname{Re}\phi(z) > 0$ and $\delta(z) = \min(-1 + q + q(\phi(\operatorname{Re} z) - \operatorname{Re}(\phi(z))), -g(\operatorname{Re} z, q)) > 0$ yields:

$$n^{-\phi(\operatorname{Re} z)+\operatorname{Re}\phi(z)} \cdot M_n(z) = O_K\left(n^{-\epsilon(K)}\right).$$

Using the estimates of Lemma 3.11, we have $|C_n(z)| = O_K\left(n^{\operatorname{Re}\phi(z)}\right)$, and so $|C_i(z)M_i(z)| = O_K\left(i^{\phi(\operatorname{Re} z)-\epsilon(K)}\right)$. Hence $\sum_{i=J}^{n-1} |C_i(z)M_i(z)| = O_K\left(n^{0\vee(1+\phi(\operatorname{Re} z)-\epsilon(K))}\right)$.

For the last point, we use Lemma 3.34 on $\mathscr{D}'_q$ for the martingale $(z \mapsto X_n(z))_{n\geq J}$ with $\alpha(z) = 1 + \phi(\operatorname{Re} z) > 0$ and $\delta(z) = \min(-1 + q + q(\phi(\operatorname{Re} z) - \operatorname{Re}(\phi(z))), -g(\operatorname{Re} z, q))$. ☐

In order to conclude, we will also need the following lemma, which is a direct consequence of Lemma 3.11.

Lemma 3.23. *For any compact $K \subset \mathscr{E} \cap \{z \in \mathbb{C} \mid 1 + \operatorname{Re}(\phi(z)) > 0\}$, there exists $\epsilon(K)$ such that*

$$\left| n^{-(1+\phi(z))} \cdot \sum_{i=J}^{n-1} C_i(z) - \frac{e^{c(z)}}{1+\phi(z)} \right| = O_K\left(n^{-\epsilon(K)}\right)$$

Proof. On any compact $K \subset \mathscr{E} \cap \{z \in \mathbb{C} \mid 1 + \operatorname{Re}(\phi(z)) > 0\}$, using Lemma 3.11 we write

$$C_n(z) = e^{c(z)} \cdot n^{\phi(z)} \cdot (1 + O_K\left(n^{-\epsilon}\right)),$$

so that

$$\begin{aligned}\sum_{i=1}^{n-1} C_i(z) &= e^{c(z)} \cdot \sum_{i=1}^{n-1} i^{\phi(z)} + e^{c(z)} \cdot \sum_{i=1}^{n-1} i^{\phi(z)} \cdot O_K\left(i^{-\epsilon}\right) \\ &= \frac{e^{c(z)} n^{1+\phi(z)}}{1+\phi(z)} \cdot (1 + O_K\left(n^{-1}\right)) + O_K\left(n^{1+\phi(z)-\epsilon(K)}\right),\end{aligned}$$

where in the second line, we use the fact that $\inf_{z\in K}(1 + \operatorname{Re}\phi(z)) > 0$, and we define $\epsilon(K) := \epsilon \wedge \inf_{z\in K}(1 + \operatorname{Re}\phi(z))$. This proves the lemma. ☐

We can now prove Proposition 3.9.

Proof of Proposition 3.9. Let us start by proving simultaneously that $N_\infty(z) = \frac{e^{z+c(z)}}{1+\phi(z)} M_\infty(z)$ and both point (i) and (ii) of the proposition. For $K \subset \mathscr{D}$ compact and $z \in K$, we write

$$\begin{aligned}&\left| N_n(z) - \frac{e^{z+c(z)}}{1+\phi(z)} M_\infty(z) \right| \\ &\qquad \leq \left| n^{-(1+\phi(z))} X_n(z) \right| + \left| n^{-(1+\phi(z))} e^z \sum_{i=J}^{n} C_i(z) M_i(z) - \frac{e^{z+c(z)}}{1+\phi(z)} M_\infty(z) \right|.\end{aligned}$$

The first term is $O_K\left(n^{-\epsilon(K)}\right)$ thanks to Lemma 3.22(i). We bound the second one by the following quantity

$$|M_\infty(z)| \cdot |e^z| \cdot \underbrace{\left| n^{-(1+\phi(z))} \cdot \sum_{i=J}^{n} C_i(z) - \frac{e^{c(z)}}{(1+\phi(z))} \right|}_{O_K\left(n^{-\epsilon(K)}\right)} + n^{-(1+\operatorname{Re}\phi(z))} \cdot |e^z| \cdot \sum_{i=J}^{n-1} \underbrace{|C_i(z)| \cdot |M_i(z) - M_\infty(z)|}_{O_K\left(i^{\operatorname{Re}\phi(z)-\epsilon(K)}\right)}.$$

We then use Lemma 3.23 and Lemma 3.11 together with Proposition 3.14) to prove that the first and second terms of the last display are $O_K\big(n^{-\epsilon(K)}\big)$. The limiting function $N_\infty(z)$ is analytic as a uniform limit of analytic functions and has almost surely no zero on $(z_-\,,z_+)$ because of Lemma 3.19. For (iii), let us prove the stronger statement: for any compact set $K\subset(z_-\,,z_+)$ and $0<a<\pi$, there exists $\epsilon(K,a)>0$ such that almost surely,

$$\sup_{x\in K}\sup_{a\le\eta\le\pi} n^{-(1+\phi(x))}\left|\sum_{i=1}^{n} e^{(x+i\eta)\,\mathrm{ht}(u_i)}\right| = O\Big(n^{-\epsilon(K,a)}\Big)\,.$$

For this, we write

$$n^{-(1+\phi(x))}\left|\sum_{i=1}^{n} e^{(x+i\eta)\,\mathrm{ht}(u_i)}\right| \le n^{-(1+\phi(x))}|X_n(x+i\eta)| + n^{-(1+\phi(x))}\left|\sum_{i=J}^{n-1} C_i(x+i\eta)M_i(x+i\eta)\right|.$$

We apply points (ii) and (iii) of Lemma 3.22 to the compact $K\times[a\,,\pi]$ and get the desired bound. ☐

3.3.3 Height of the tree

In this section, we study the behaviour of the height $\mathrm{ht}(\mathtt{T}_n)$ of the tree $\mathtt{T}_n$, which is defined as the maximal height of the vertices of $\mathtt{T}_n$, i.e.,

$$\mathrm{ht}(\mathtt{T}_n) = \max_{1\le k\le n}\mathrm{ht}(u_k).$$

We start by showing that under the assumption ($\square_\gamma^p$) we have the convergence (3.9). Then, for the sake of completeness, we also study the simpler case where $\log n = o\Big(\sum_{i=1}^n \frac{w_i}{W_i}\Big)$.

One key argument in our proofs is the following equality for the *annealed* moment generating function of the height of u_k, for any fixed $k\ge1$, which can be seen as a corollary of Lemma 3.10,

$$\mathbb{E}\left[e^{z\,\mathrm{ht}(u_k)}\right] = e^z\cdot\prod_{j=2}^{k-1}\left(1+(e^z-1)\frac{w_j}{W_j}\right). \tag{3.34}$$

Some elementary computations using the Chernoff bound and the last display yield the following lemma:

Lemma 3.24. *Suppose that the sequence of weights* $\mathbf{w}$ *satisfies*

$$\limsup_{n\to\infty}\frac{1}{\log n}\sum_{i=2}^{n}\frac{w_i}{W_i}\le u\in\mathbb{R}_+^*.$$

Then almost surely we have

$$\limsup_{n\to\infty}\frac{\mathrm{ht}(\mathtt{T}_n)}{\log n}\le ue^{z_+(u)},$$

where $z_+(u)$ *is the unique positive root of* $u(ze^z-e^z+1)-1=0$.

Proof. Using the expression (3.34) for the moment generating function of $\mathrm{ht}(u_n)$ we get, for any $z>0$,

$$\begin{aligned}\mathbb{E}\left[e^{z\,\mathrm{ht}(u_n)}\right] = e^z\cdot\prod_{j=2}^{n-1}\left(1+(e^z-1)\frac{w_j}{W_j}\right) &\le \exp\left(1+(e^z-1)\sum_{j=2}^{n-1}\frac{w_j}{W_j}\right)\\ &\le \exp\left((\log n)\cdot(u(e^z-1)+o(1))\right),\end{aligned}$$

where we use the inequality $(1+x) \le e^x$ and the assumption on $\mathbf{w}$. Then, for any $z > 0$ and $n \ge 1$,

$$\mathbb{P}\left(\mathrm{ht}(u_n) \ge ue^z \log n\right) \le e^{-uze^z \log n}\mathbb{E}\left[e^{z\,\mathrm{ht}(u_n)}\right] \le \exp\left(-u\log n(ze^z - e^z + 1 + o(1))\right).$$

If we take $z > 0$ such that $u(ze^z - e^z + 1) > 1$ then the right-hand-side is summable and hence using the Borel-Cantelli lemma shows that for all n large enough, we have $\mathrm{ht}(u_n) \le ue^z \log n$. Letting $z \searrow z^+(u)$, we get the result. □

Let us prove the last claim of Theorem 3.3. Here we suppose that the weight sequence $\mathbf{w}$ satisfies ($\square_\gamma^p$) for some $\gamma > 0$ and some $p \in (1\,,2]$.

Proof of Theorem 3.3. Recall the asymptotics (3.8) in Theorem 3.3. It ensures that there almost surely exist vertices at height $\gamma e^z \log n$, for any $z \in (z_-\,, z_+)$. Hence the height of the tree $\mathtt{T}_n$ satisfies

$$\liminf_{n\to\infty} \frac{\mathrm{ht}(\mathtt{T}_n)}{\log n} \ge \gamma e^{z_+}.$$

For the limsup, we use Lemma 3.25 with $u = \gamma$ (this is justified by Lemma 3.12), which yields $\limsup_{n\to\infty} \frac{\mathrm{ht}(\mathtt{T}_n)}{\log n} \ge \gamma e^{z_+}$. □

To finish the section, we state a proposition.

Proposition 3.25. *Let $f(n) := \sum_{i=2}^n \frac{w_i}{W_i}$. If $f(n) \gg \log n$ as $n \to \infty$ then we have the almost sure convergence*

$$\lim_{n\to\infty} \frac{\mathrm{ht}(\mathtt{T}_n)}{f(n)} = 1.$$

Proof. For the upper-bound, we proceed as above. For any $\epsilon > 0$ and $z > 0$:

$$\begin{aligned}\mathbb{P}\left(\mathrm{ht}(u_n) \ge (1+\epsilon)f(n)\right) &\le \exp(-z(1+\epsilon)f(n))\mathbb{E}\left[e^{z\,\mathrm{ht}(u_n)}\right] \\ &\le \exp\left((e^z-1)f(n-1) - (1+\epsilon)zf(n)\right) \\ &\le \exp\left(f(n)\left[e^z - 1 - (1+\epsilon)z + o(1)\right]\right)\end{aligned}$$

If we choose $z > 0$ close enough to 0 then the last display is summable, due to our assumption on f. This implies using the Borel-Cantelli lemma that $\limsup_{n\to\infty} \frac{\mathrm{ht}(\mathtt{T}_n)}{f(n)} \le 1 + \epsilon$ almost surely, for any fixed $\epsilon > 0$. Then we let $\epsilon \searrow 0$.

For the lower-bound, we use the fact that we can construct jointly with $(\mathtt{T}_n)_{n\ge1}$ a sequence $(D_n)_{n\ge1}$ such that $\forall n \ge 1, D_n \in \mathtt{T}_n$, increasing for the genealogical order and such that, as a sequence, we have

$$(\mathrm{ht}(D_n))_{n\ge1} \overset{(\mathrm{d})}{=} \left(\sum_{i=2}^n \mathbf{1}_{\left\{U_i \le \frac{w_i}{W_i}\right\}}\right)_{n\ge1}$$

with $(U_i)_{i\ge2}$ i.i.d. uniform random variables. See for example [41, Section 2.2] or in our setting [91, Corollary 8]. Using the law of large numbers, we get that almost surely $\mathrm{ht}(D_n) \sim \sum_{i=2}^n \frac{w_i}{W_i} = f(n)$ as $n \to \infty$. Since $\mathrm{ht}(\mathtt{T}_n) \ge \mathrm{ht}(D_n)$, this proves the lower-bound and finishes the proof. □

3.4 Preferential attachment trees are weighted recursive trees

In this section, we study preferential attachment trees with initial fitnesses $\mathbf{a}$ as defined in the introduction. First, in Section 3.4.1, we prove Theorem 3.1 which allows us to see them as weighted random trees $\mathrm{WRT}(\mathbf{w^a})$ for some random weight sequence $\mathbf{w^a}$. Then in Section 3.4.2 we prove Proposition 3.2 which relates the asymptotic behaviour of $\mathbf{w^a}$ to the behaviour of $\mathbf{a}$. Finally, in Section 3.4.3 we prove Proposition 3.27, which ensures that the sequence $\mathbf{m^a}$ obtained as the scaling limit of the degrees can be expressed as the increments of a Markov chain.

3.4.1 Coupling with a sequence of Pólya urns

Here we fix an arbitrary sequence $\mathbf{a}$ such that $a_1 > -1$ and $\forall n \geq 2,\ a_n \geq 0$. Let us recall the notation, for $n \geq 0$,

$$A_n := \sum_{i=1}^{n} a_i,$$

with the convention that $A_0 = 0$. We consider a sequence of trees $(\mathtt{P}_n)_{n\geq 1}$ evolving according to the distribution $\mathrm{PA}(\mathbf{a})$ and we want to prove Theorem 3.1, namely that there exists a random sequence of weights $\mathbf{w^a}$ for which the sequence evolves as a $\mathrm{WRT}(\mathbf{w^a})$. The proof uses a decomposition of this process into an infinite number of Pólya urns. This is very close to what is used in the proofs of [17, Theorem 2.1] or [28, Section 1.2] in similar settings. The novelty of our approach is to express this result using weighted random trees, since it allows us to apply all the results developed in the preceding section.

Pólya urns. For us, a Pólya urn process $(\mathsf{Urn}(n))_{n\geq 0} = (X(n), \mathrm{Total}(n))_{n\geq 0}$ is a Markov chain on $E := \left\{(x,z) \in \mathbb{R}_+ \times \mathbb{R}_+^* \;\middle|\; x \leq z\right\}$ with transition probabilities given by the matrix P where for all $(x,z) \in E$,

$$P\left((x,z),(x+1,z+1)\right) = \frac{x}{z} \qquad \text{and} \qquad P\left((x,z),(x,z+1)\right) = \frac{z-x}{z}. \tag{3.35}$$

The quantities $X(n)$ and $\mathrm{Total}(n)$ represent respectively the number of red balls and the total number of balls at time n in a urn containing red and blacks balls, in which we add a ball at each time, the colour of which is chosen at random proportionally to the current proportion in the urn. Starting at time 0 from the state $(a, a+b)$, i.e. with a red balls and b black balls, it is well-known that the sequence $(\Delta X(n))_{n\geq 1} = (X(n) - X(n-1))_{n\geq 1}$ of random variables is *exchangeable*, and an application of de Finetti's representation theorem ensures that it has the same distribution as i.i.d. samples of Bernoulli random variables with a *random* parameter β, which has distribution $\mathrm{Beta}(a,b)$, where we use the convention that $\mathrm{Beta}(a,b) = \delta_1$ if $b = 0$.

Nested structure of urns in the tree. For all $k \geq 1$ we define the following process in $n \geq k$

$$W_k(n) := A_k + \sum_{i=1}^{k} \deg^+_{\mathtt{P}_n}(u_i),$$

the "total fitness" of the vertices $\{u_1, u_2, \dots, u_k\}$, for which we remark that for any $k \geq 1$ we have

$$W_k(k) = A_k + k - 1 \qquad \text{and} \qquad W_k(k+1) = A_k + k. \tag{3.36}$$

Imagine that $\mathtt{P}_n$ is constructed and we add a new vertex u_{n+1} to the tree. We choose its parent in a downward sequential way:

- we first determine whether the parent is u_n, this happens with probability

$$\frac{a_n + \deg^+_{\mathsf{P}_n}(u_n)}{W_n(n)} = 1 - \frac{W_{n-1}(n)}{W_n(n)},$$

- then with the complementary probability $\frac{W_{n-1}(n)}{W_n(n)}$ it is not, so conditionally on this we determine whether it is u_{n-1}, this happens with (conditional) probability

$$\frac{a_{n-1} + \deg^+_{\mathsf{P}_n}(u_{n-1})}{W_{n-1}(n)} = 1 - \frac{W_{n-2}(n)}{W_{n-1}(n)}.$$

- then with the complementary probability $\frac{W_{n-2}(n)}{W_{n-1}(n)}$ it is not, etc... We continue this process until we stop at some u_i.

Now let us fix $k \geq 1$ and introduce the following time-change: for all $N \geq 0$, we let

$$\theta_k(N) := \inf\left\{n \geq k+1 \mid W_{k+1}(n) = A_{k+1} + k + N\right\}, \tag{3.37}$$

be the N-th time that a vertex is attached to one of the vertices $\{u_1, \dots, u_{k+1}\}$. Remark that it can be the case that $\theta_k(N)$ is not defined for large N, if there is only a finite number of vertices attaching to $\{u_1, \dots, u_{k+1}\}$. Let us ignore this possible problem for the moment, and only consider bounded sequences $\mathbf{a}$, for which this will almost surely not happen. In this case for all $N \geq 0$ we set

$$\mathsf{Urn}_k(N) := (W_k(\theta_k(N)), W_{k+1}(\theta_k(N))) = (W_k(\theta_k(N)), A_{k+1} + k + N). \tag{3.38}$$

Now, the two following facts are the key observations in order to prove Theorem 3.1:

(i) for all $k \geq 1$, the process $\mathsf{Urn}_k = (\mathsf{Urn}_k(N))_{N \geq 0}$ has the distribution of a Pólya urn starting from the state $(A_k + k, A_{k+1} + k)$,

(ii) those process are jointly independent for $k \geq 1$.

Point (i) already follows from the discussion above. A moment of thought shows that (ii) holds as well: of course the processes $(W_k(n), W_{k+1}(n))_{n \geq k+1}$ for different k are not independent at all but the point is that they only interact through the time-changes $(\theta_k(\cdot), k \geq 1)$.

Reversing the construction and using the exchangeability. Using de Finetti's theorem and points (i) and (ii), each of the processes Urn_k can be produced by sampling $\beta_k \sim \mathrm{Beta}(A_k + k, a_{k+1})$ and adding a red ball at each step independently with probability β_k and a black ball with probability $1 - \beta_k$. This is of course done independently for different $k \geq 1$.

In terms of our downward sequential procedure defined above for finding the parent of each newcomer, it amounts to saying that each time that we have to choose between attaching to u_{k+1} or attach to a vertex among $\{u_1, \dots, u_k\}$, the former is chosen with probability $1 - \beta_k$ and the latter with probability β_k. Let us verify that the law of $(\mathsf{P}_n)_{n \geq 1}$ conditionally on the sequence $(\beta_k)_{k \geq 1}$ can indeed be expressed as WRT with the random sequence of weights $\mathsf{w}^{\mathbf{a}}$ defined in Theorem 3.1, which is defined from the sequence $(\beta_k)_{k \geq 1}$ as

$$\forall n \geq 1, \qquad \mathsf{W}^{\mathbf{a}}_n = \prod_{i=1}^{n-1} \beta_i^{-1} \qquad \text{and} \qquad \mathsf{w}^{\mathbf{a}}_n = \mathsf{W}^{\mathbf{a}}_n - \mathsf{W}^{\mathbf{a}}_{n-1},$$

with the convention that $\mathsf{W}_1^{\mathbf{a}} = 1$ and $\mathsf{W}_0^{\mathbf{a}} = 0$. Let us reason conditionally on the sequence $(\beta_k)_{k\geq 1}$ (or equivalently the sequence $(\mathsf{w}_n^{\mathbf{a}})_{n\geq 1}$). When determining the parent of u_{n+1}, we successively try to attach to $u_n, u_{n-1}, \dots$ until we stop at some u_k. Using the independence, we get

$$\mathbb{P}\left(J_{n+1} = k \mid \mathtt{P}_n, \beta_1, \beta_2, \dots\right) = \beta_{n-1}\beta_{n-2}\dots\beta_k(1-\beta_{k-1}) = \frac{\mathsf{W}_k^{\mathbf{a}} - \mathsf{W}_{k-1}^{\mathbf{a}}}{\mathsf{W}_n^{\mathbf{a}}} = \frac{\mathsf{w}_k^{\mathbf{a}}}{\mathsf{W}_n^{\mathbf{a}}}.$$

Remark that the above construction is still valid without the assumption that the sequence $\mathbf{a}$ is bounded, and hence Theorem 3.1 is proved.

3.4.2 Proof of Proposition 3.2

Let $(\mathsf{W}_n^{\mathbf{a}})_{n\geq 1}$ be the random sequence of cumulated weights defined in Theorem 3.1, whose distribution depends on a sequence $\mathbf{a}$ of initial fitnesses, and is expressed using a sequence of independent Beta-distributed random variables $(\beta_k)_{k\geq 1}$. We are going to prove Proposition 3.2, which relates the growth of $(\mathsf{W}_n^{\mathbf{a}})_{n\geq 1}$ to that of $(A_n)_{n\geq 1}$. In this proof, we omit the superscript $\mathbf{a}$ for readability.

Proof of Proposition 3.2. As in [63], we introduce

$$X_n := \prod_{i=1}^{n-1} \frac{\beta_i}{\mathbb{E}\left[\beta_i\right]}.$$

It is easy to see that X_n is a positive martingale, hence it almost surely converges to a limit X_∞ as $n \to \infty$. Now, using the fact that the $(\beta_n)_{n\geq 1}$ are independent and that the expectation of a random variable with Beta(a,b) distribution distribution has q-th moment, for $q \geq 0$,

$$\frac{\Gamma(a+q)\Gamma(a+b)}{\Gamma(a)\Gamma(a+b+q)} = \prod_{k=0}^{q-1} \frac{a+k}{a+b+k}, \tag{3.39}$$

we can compute

$$\begin{aligned}\prod_{i=1}^{n-1} \mathbb{E}\left[\beta_i^p\right] = \prod_{i=1}^{n-1}\left(\prod_{k=0}^{p-1} \frac{i+A_i+k}{i+A_{i+1}+k}\right) &= \prod_{k=0}^{p-1}\left(\frac{1+A_1+k}{n+A_n+k-1}\prod_{i=2}^{n-1}\frac{i+A_i+k}{i+A_i+k-1}\right)\\ &= \left(\prod_{k=0}^{p-1}\frac{1+A_1+k}{n+A_n+k-1}\right)\cdot\prod_{k=0}^{p-1}\prod_{i=2}^{n-1}\left(1+\frac{1}{i+A_i+k-1}\right)\end{aligned}$$

Now from our hypotheses on the sequence (A_n), we have for all $k \in [\![0\,, p-1]\!]$

$$n + A_n + k - 1 \underset{n\to\infty}{=} (c+1)n + O\big(n^{1-\epsilon}\big) \quad \text{and so} \quad \frac{1}{n+A_n+k-1} \underset{n\to\infty}{=} \frac{1}{(c+1)n} + O\big(n^{-1-\epsilon}\big).$$

Hence

$$\begin{aligned}\prod_{k=0}^{p-1}\prod_{i=2}^{n-1}\left(1+\frac{1}{i+A_i+k-1}\right) &= \prod_{k=0}^{p-1}\prod_{i=2}^{n-1}\left(1+\frac{1}{(c+1)i}+O\big(i^{-1-\epsilon}\big)\right)\\ &= \exp\left(\sum_{k=0}^{p-1}\sum_{i=2}^{n}\left(\frac{1}{(c+1)i}+O\big(i^{-1-\epsilon}\big)\right)\right)\\ &= \exp\left(\frac{p}{c+1}\log n + \mathrm{cst} + O\big(n^{-\epsilon}\big)\right)\\ &= \mathrm{cst}\cdot n^{\frac{p}{c+1}}\left(1+O\big(n^{-\epsilon}\big)\right).\end{aligned}$$

In the end, since $\prod_{k=0}^{p-1} \frac{1+A_1+k}{n+A_n+k-1} = \mathrm{cst} \cdot \prod_{k=0}^{p-1} \frac{1}{(c+1)n+O(n^{1-\epsilon})} = \mathrm{cst} \cdot n^{-p} \cdot (1+O(n^{-\epsilon}))$, we get

$$\prod_{i=1}^{n-1} \mathbb{E}\left[\beta_i^p\right] = C_p \cdot n^{-p+p/(c+1)} \cdot (1+O\big(n^{-\epsilon}\big)) \tag{3.40}$$

where C_p is a positive constant which depends on the sequence $\mathbf{a}$ and p. This entails that, under our assumptions, for any $p \geq 1$, we have $\mathbb{E}\left[X_n^p\right] \to C_p/C_1^p$ as $n \to \infty$, which shows that this martingale is bounded in L^p for all $p \geq 1$ and hence it is uniformly integrable. Consequently, it converges a.s. and in L^p to a limit random variable X_∞, with moments determined by

$$\forall p \geq 1, \qquad \mathbb{E}\left[X_\infty^p\right] = \frac{C_p}{C_1^p}. \tag{3.41}$$

Furthermore, we have

$$\mathbb{E}\left[(X_{n+1}-X_n)^2\right] = \mathbb{E}\left[X_n^2\left(\frac{\beta_n}{\mathbb{E}\left[\beta_n\right]}-1\right)^2\right] \leq C_2 \cdot \frac{\mathbb{V}\left(\beta_n\right)}{\mathbb{E}\left[\beta_n\right]^2}. \tag{3.42}$$

Since $\beta_n \sim \mathrm{Beta}(n+A_n, a_{n+1})$, we get

$$\mathbb{E}\left[\beta_n\right] = \frac{n+A_n}{n+A_{n+1}} \to 1 \qquad \text{and} \qquad \mathbb{V}\left(\beta_n\right) = \frac{a_{n+1}(n+A_n)}{(n+A_{n+1})^2(n+A_{n+1}+1)} = O\big(n^{-1-\epsilon}\big). \tag{3.43}$$

Using equation (3.42), equation (3.43), Lemma 3.33 and summing over $n \leq k \leq 2n-1$ we get that $\mathbb{E}\left[(X_{2n}-X_n)^2\right] = O(n^{-\epsilon})$. Using Lemma 3.34, we get, for some $\epsilon > 0$,

$$|X_n - X_\infty| = O\big(n^{-\epsilon}\big).$$

Since $\beta_i > 0$ almost surely for every $i \geq 1$, the event $\{X_\infty = 0\}$ is a tail event for the filtration generated by the β_i and has probability 0 or 1. In the end, it has probability 0 because $\mathbb{E}\left[X_\infty\right] = 1$. We deduce that

$$\begin{aligned}\mathsf{W}_n^{-1} = \prod_{i=1}^{n-1} \beta_i = X_n \cdot \prod_{i=1}^{n-1} \mathbb{E}\left[\beta_i\right] \\ &= X_\infty \cdot (1+O\big(n^{-\epsilon}\big)) \cdot C_1 \cdot n^{-1+\frac{1}{c+1}} \cdot (1+O\big(n^{-\epsilon}\big)) \\ &= C_1 \cdot X_\infty \cdot n^{-1+\frac{1}{c+1}} \cdot (1+O\big(n^{-\epsilon}\big)).\end{aligned}$$

Hence, we have,

$$\mathsf{W}_n = Z \cdot n^{\frac{c}{(c+1)}} \cdot (1+O\big(n^{-\epsilon}\big)) \qquad \text{with} \qquad Z := \frac{1}{X_\infty \cdot C_1}. \tag{3.44}$$

Whenever $a_n \leq n^{c'+o(1)}$ as $n \to \infty$, we can show the following (we postpone the proof to the end of the section):

Lemma 3.26. *For any $\delta > 0$, we have*

$$\mathbb{P}\left(1-\beta_k > k^{-1+c'+\delta}\right) \leq \exp\left(-k^{\delta+o(1)}\right).$$

Since the last quantity is summable in k we can use the Borel-Cantelli lemma (and a sequence of δ going to 0) to show that almost surely $1-\beta_k \leq k^{-1+c'+o(1)}$ as $k \to \infty$. This finishes to prove the proposition, because we can write

$$\mathsf{w}_k = \mathsf{W}_k - \mathsf{W}_{k-1} = \mathsf{W}_k \cdot (1-\beta_{k-1}) \leq k^{c'-1/(c+1)+o(1)}. \qquad \square$$

We finish by giving a proof of Lemma 3.26.

Proof of Lemma 3.26. Let $x > 0$ and $y > 1$ and let β be a random variable with $\mathrm{Beta}(x, y)$ distribution. Then for any $z \in [0,1]$ we have, using the explicit expression of the density of β:

$$\begin{aligned}\mathbb{P}(\beta > z) &= \frac{\Gamma(x+y)}{\Gamma(x)\Gamma(y)} \int_z^1 u^{x-1}(1-u)^{y-1}\,\mathrm{d}u \\ &\leq \frac{\Gamma(x+y)}{\Gamma(x)\Gamma(y)} \exp(-(y-1)z) \int_z^1 u^{x-1}\,\mathrm{d}u \\ &\leq \frac{\Gamma(x+y)}{\Gamma(x+1)\Gamma(y)} \cdot \exp(-(y-1)z).\end{aligned}$$

For any two sequences (x_n) and (y_n) simultaneously going to infinity with $x_n = o(y_n)$, we have the following bound using Stirling's approximation:

$$\log\left(\frac{\Gamma(x_n+y_n)}{\Gamma(x_n+1)\Gamma(y_n)}\right) \underset{n\to\infty}{\sim} x_n \log(y_n).$$

Applying the above computations for $(1-\beta_n) \sim \mathrm{Beta}(a_{n+1}, A_n + n)$, and using the assumptions on the sequence $\mathbf{a}$, we get

$$\log \mathbb{P}\left(1 - \beta_n > n^{-1+c'+\delta}\right) \leq -n^{\delta+o(1)},$$

which is what we wanted. □

3.4.3 The distribution of the limiting sequence

Recall the convergence of the degree sequence stated in Proposition 3.5. Thanks to what precedes, we know that if some sequence $\mathbf{a}$ satisfies (H_c) then the associated random sequence $(\mathsf{w}_n^{\mathbf{a}})_{n\geq 1}$ satisfies $\mathsf{W}_n^{\mathbf{a}} \underset{n\to\infty}{\sim} Z \cdot n^{c/(c+1)}$ and so in this setting the convergence of degrees can be stated as

$$n^{-\frac{1}{c+1}} \cdot (\deg^+_{\mathsf{P}_n}(u_1), \deg^+_{\mathsf{P}_n}(u_2), \dots) \underset{n\to\infty}{\longrightarrow} (\mathsf{m}_1^{\mathbf{a}}, \mathsf{m}_2^{\mathbf{a}}, \dots)$$

where $(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1} = \frac{c+1}{Z} \cdot (\mathsf{w}_n^{\mathbf{a}})_{n\geq 1}$. Remark that the random variable Z depends on the whole sequence $(\beta_n)_{n\geq 1}$ used in the definition of $(\mathsf{W}_n^{\mathbf{a}})_{n\geq 1}$, so the sequence $(\mathsf{M}_n^{\mathbf{a}})_{n\geq 1} = \frac{c+1}{Z} \cdot (\mathsf{W}_n^{\mathbf{a}})_{n\geq 1}$ can not be seen as an iterated product of independent random variables, which was the case for $(\mathsf{W}_n^{\mathbf{a}})_{n\geq 1}$. We will prove that this new process still has some nice properties.

Proposition 3.27. *For any sequence* $\boldsymbol{a}$ *that satisfies the condition* (H_c)*, the sequence* $(\mathsf{M}_k^{\boldsymbol{a}})_{k\geq 1}$ *is a (possibly time-inhomogeneous) Markov chain such that for all* $k \geq 1$*,* $\mathsf{M}_{k+1}^{\boldsymbol{a}}$ *is independent of* $\beta_1, \beta_2, \dots, \beta_k$*. The fact that for all* $k \geq 1$ *we have* $\mathsf{M}_k^{\boldsymbol{a}} = \beta_k \cdot \mathsf{M}_{k+1}^{\boldsymbol{a}}$ *with* $\beta_k \sim \mathrm{Beta}(A_k + k, a_{k+1})$ *independent of* $\mathsf{M}_{k+1}^{\boldsymbol{a}}$ *characterises the backward transitions of the chain.*

Proof. We follow the same steps as [63, Lemma 1.1]. Let us fix a sequence $\mathbf{a}$ that satisfies the hypotheses of the proposition and make the dependence on it implicit to ease notation. Recall from (3.40) the definition of C_1 and from (3.44) the definition of Z from X_∞. We have

$$\mathsf{M}_1 = (C_1 \cdot (c+1) \cdot X_\infty) \quad \text{and for } k \geq 2, \quad \mathsf{M}_k = \mathsf{M}_1 \cdot \left(\prod_{i=1}^{k-1} \beta_i\right)^{-1}. \tag{3.45}$$

It then follows that we can write, for $k \geq 1$,

$$\mathsf{M}_{k+1} = C_1 \cdot (c+1) \cdot X_\infty \cdot \left(\prod_{i=1}^{k} \beta_i\right)^{-1} = C_1 \cdot (c+1) \cdot \lim_{n\to\infty} \frac{\prod_{i=k+1}^{n-1} \beta_i}{\prod_{i=1}^{n-1} \mathbb{E}\left[\beta_i\right]},$$

which ensures that M_{k+1} is independent of $\beta_1, \beta_2, ..., \beta_k$. The limit in the last equality exists almost surely thanks to the results of the preceding section.

Now we prove the Markov property of the chain. Let $k \geq 1$. Because of the definition of the chain as a product, the distribution of M_{k+1} conditional on the past trajectory $\mathsf{M}_1, \mathsf{M}_2, \dots, \mathsf{M}_k$ is the same as the distribution of M_{k+1} conditional on $\mathsf{M}_k, \beta_1, \dots, \beta_{k-1}$. Since $\mathsf{M}_{k+1} = \beta_k^{-1} \cdot \mathsf{M}_k$ and that β_k and M_k are both independent of $\beta_1, \dots, \beta_{k-1}$, this conditional distribution corresponds to the one of M_{k+1} conditional on the present state of the chain M_k. □

Computing the moments. In some cases where the sequence $\mathbf{a}$ is sufficiently regular, we can compute explicitly every moment of the random variable $\mathsf{M}_k^{\mathbf{a}}$ for every $k \geq 1$. Indeed, using (3.41) and (3.45) and the independence, we get

$$\begin{aligned}\mathbb{E}\left[\mathsf{M}_k^p\right] = \mathbb{E}\left[\left(C_1 \cdot (c+1) \cdot \lim_{n\to\infty} \frac{\prod_{i=k}^{n-1} \beta_i}{\prod_{i=1}^{n-1} \mathbb{E}\left[\beta_i\right]}\right)^p\right] &= C_1^p \cdot (c+1)^p \cdot \lim_{n\to\infty} \frac{\prod_{i=k}^{n-1} \mathbb{E}\left[\beta_i^p\right]}{\left(\prod_{i=1}^{n-1} \mathbb{E}\left[\beta_i\right]\right)^p} \\ &= \frac{(c+1)^p \cdot C_p}{\prod_{i=1}^{k-1} \mathbb{E}\left[\beta_i^p\right]}. \end{aligned} \tag{3.46}$$

In general, if the collection $(\mu_p)_{p\geq 1}$ of p-th moments of some positive random variable satisfies the so-called Carleman's condition: $\sum_{p=1}^{\infty} \mu_p^{-1/(2p)} = \infty$, then its distribution is uniquely determined from those moments.

3.5 Examples and applications

In this section, we compute the explicit distribution of $(\mathsf{M}_n^{\mathbf{a}})$ for some particular sequences $\mathbf{a}$. We apply this result to another model of preferential attachment.

3.5.1 The limit chain for particular sequences a

As stated in the preceding section, we can compute the distribution of $\mathsf{M}_k^{\mathbf{a}}$ for some fixed k by the expression of its moments (3.46), provided that they satisfy Carleman's condition. Knowing these distributions and the backward transitions given in Proposition 3.27 then characterises the law of the whole process. For two particular examples, this law has a nice expression.

Proposition 3.28. *In the two following cases, the distribution of the chain $(\mathsf{M}_n^{\mathbf{a}})$ is explicit.*

(i) *If $\mathbf{a}$ is of the form $\mathbf{a} = (a, b, b, b, \dots)$ with $a > -1$ and $b > 0$, then the limiting sequence $(\mathsf{M}_n^{\mathbf{a}})_{n\geq 1}$ is a Mittag-Leffler Markov chain* MLMC $\left(\frac{1}{b+1}, \frac{a}{b+1}\right)$.

(ii) *If $\mathbf{a}$ is of the form $\mathbf{a} = (a, \underbrace{0, 0, \dots, 0}_{\ell-1}, m, \underbrace{0, 0, \dots, 0}_{\ell-1}, m, \dots)$ with $a > -1$ and $\ell, m \in \mathbb{N}$, then $(\mathsf{M}_n^{\mathbf{a}})_{n\geq 1}$ is constant on the interval of the form $[\![1+k\ell\,, (k+1)\ell]\!]$ and the sequence*

$$\frac{\ell^{\frac{\ell}{m+\ell}}}{m+\ell} \cdot (\mathsf{N}_k^{\mathbf{a}})_{k\geq 1} = \frac{\ell^{\frac{\ell}{m+\ell}}}{m+\ell} \cdot (\mathsf{M}_{(k-1)\ell+1}^{\mathbf{a}})_{k\geq 1}$$

has the Product Generalised Gamma *distribution* $\mathrm{PGG}(a, \ell, m)$.

We will prove the two points of this proposition in separate subsections. The proper definitions of the distributions to which we refer are given along the proof. For the rest of the section, we drop the superscript $\mathbf{a}$ and write $(\mathsf{M}_n)_{n\geq 1}$.

Mittag-Leffler Markov chains

Let us study the case where the underlying preferential attachment tree has a sequence of initial fitnesses $\mathbf{a}$ that are of the form $(a, b, b, b, \dots)$. We start by recalling the definitions of Mittag-Leffler distributions and Mittag-Leffler Markov chains introduced in [63], and also studied in [79].

Mittag-Leffler distributions. Let $0 < \alpha < 1$ and $\theta > -\alpha$. The generalized Mittag-Leffler $\mathrm{ML}(\alpha, \theta)$ distribution has pth moment

$$\frac{\Gamma(\theta)\Gamma(\theta/\alpha + p)}{\Gamma(\theta/\alpha)\Gamma(\theta + p\alpha)} = \frac{\Gamma(\theta+1)\Gamma(\theta/\alpha+p+1)}{\Gamma(\theta/\alpha+1)\Gamma(\theta+p\alpha+1)} \tag{3.47}$$

and the collection of p-th moments for $p \in \mathbb{N}$ uniquely characterizes this distribution.

Mittag-Leffler Markov Chains. For any $0 < \alpha < 1$ and $\theta > -\alpha$, we introduce the (a priori) inhomogenous Markov chain $(\mathsf{M}_n^{\alpha,\theta})_{n\geq 1}$, the distribution of which we call the Mittag-Leffler Markov chain of parameters (α, θ), or $\mathrm{MLMC}(\alpha, \theta)$. This type of Markov chain was already defined in [63], for some choice of parameters α and θ. It is a Markov chain such that for any $n \geq 1$,

$$\mathsf{M}_n^{\alpha,\theta} \sim \mathrm{ML}\left(\alpha, \theta + n - 1\right),$$

and the transition probabilities are characterised by the following equality in law:

$$\left(\mathsf{M}_n^{\alpha,\theta}, \mathsf{M}_{n+1}^{\alpha,\theta}\right) = \left(B_n \cdot \mathsf{M}_{n+1}^{\alpha,\theta}, \mathsf{M}_{n+1}^{\alpha,\theta}\right),$$

with $B_n \sim \mathrm{Beta}\left(\frac{\theta+k-1}{\alpha} + 1, \frac{1}{\alpha} - 1\right)$, independent of $\mathsf{M}_{n+1}^{\alpha,\theta}$. These chains are constructed (for a some values of θ depending on α) in [63]. In fact, our proof of Proposition 3.28(i) ensures that these chains exists for any choice of parameters $0 < \alpha < 1$ and $\theta > -\alpha$. Let us mention that the proof of [63, Lemma 1.1] is still valid for the whole range of parameters $0 < \alpha < 1$ and $\theta > -\alpha$, which proves that these Markov chains are in fact time-homogeneous.

The limiting Markov chain is a Mittag-Leffler. Recall the definition of the sequence $(\beta_k)_{k\geq 1}$ and their respective distributions $\beta_k \sim \mathrm{Beta}(A_k + k, a_{k+1})$. From our assumptions on the sequence $\mathbf{a}$ we have for all $k \geq 1$,

$$(A_k + k, a_{k+1}) = (1 + a + (k-1)b, b).$$

Proof of Proposition 3.28 (i). For $p \geq 1$, we can make the following computation, using (3.39), one change of indices and several times the property of the Gamma function that for any $z > 0$

we have $\Gamma(z+1) = z\Gamma(z)$:

$$\begin{aligned}
\prod_{i=1}^{n-1} \mathbb{E}\left[\beta_i^p\right] &= \prod_{i=1}^{n-1} \frac{\Gamma(1+a+p+(b+1)(i-1))\,\Gamma(a+(b+1)i)}{\Gamma(1+a+(b+1)(i-1))\,\Gamma(a+(b+1)i+p)} \\
&= \prod_{i=0}^{n-2} \frac{\Gamma(1+a+p+(b+1)i)}{\Gamma(1+a+(b+1)i)} \cdot \prod_{i=1}^{n-1} \frac{\Gamma(a+(b+1)i)}{\Gamma(a+(b+1)i+p)} \\
&= \frac{\Gamma(1+a+p)}{\Gamma(1+a)} \cdot \frac{\Gamma(a+(b+1)(n-1))}{\Gamma(a+(b+1)(n-1)+p)} \cdot \prod_{i=1}^{n-2} \frac{i(b+1)+a+p}{i(b+1)+a} \\
&= \frac{\Gamma(1+a+p)}{\Gamma(1+a)} \cdot \frac{\Gamma(a+(b+1)(n-1))}{\Gamma(a+(b+1)(n-1)+p)} \cdot \frac{\Gamma\left(\frac{a+p}{b+1}+n-1\right)}{\Gamma\left(\frac{a}{b+1}+n-1\right)} \cdot \frac{\Gamma\left(1+\frac{a}{b+1}\right)}{\Gamma\left(1+\frac{a+p}{b+1}\right)}. \qquad (3.48)
\end{aligned}$$

Using Stirling formula, we can then compute the numbers C_p introduced in (3.40),

$$C_p = (b+1)^{-p} \cdot \frac{\Gamma(1+a+p)\,\Gamma\left(1+\frac{a}{b+1}\right)}{\Gamma(1+a)\,\Gamma\left(1+\frac{a+p}{b+1}\right)}. \tag{3.49}$$

Using (3.46), the moments of M_k are given, for any $p \in \mathbb{N}$ by the formula

$$\mathbb{E}\left[\mathsf{M}_k^p\right] = \frac{(b+1)^p \cdot C_p}{\prod_{i=1}^{k-1} \mathbb{E}\left[\beta_i^p\right]} \underset{(3.49),(3.48)}{=} \frac{\Gamma\left(\frac{a}{b+1}+k-1\right)\Gamma(a+(b+1)(k-1)+p)}{\Gamma(a+(b+1)(k-1))\,\Gamma\left(\frac{a+p}{b+1}+k-1\right)}$$

These moments identify using (3.47) the distribution of M_k for all $k \geq 1$,

$$\mathsf{M}_k \sim \mathrm{ML}\left(\frac{1}{b+1}, \frac{a}{b+1}+k-1\right).$$

From this, and the form of the backward transitions, we can identify $(\mathsf{M}_k)_{k\geq 1}$ as having a MLMC $\left(\frac{1}{b+1}, \frac{a}{b+1}\right)$ distribution. □

Products of generalised Gamma.

The following paragraphs aim at proving Proposition 3.28(ii). In the first paragraph we define the family of distributions of PGG-process. In the second one we prove that the distribution of $(\mathsf{M}_k)_{k\geq 1}$ belongs to this family whenever the sequence $\mathbf{a}$ is of the form assumed in Proposition 3.28(ii).

Construction of a PGG(a,ℓ,m)-process. For $a > -1$ a real number and $\ell, m \geq 1$ integers, we define the following. Let $\left\{Z_i^{(q)} \,\middle|\, 0 \leq q \leq m-1,\ i \geq 1\right\}$ be a family of independent variables with the following distribution: for all $0 \leq q \leq m-1$,

$$Z_1^{(q)} \sim \mathrm{Gamma}\left(\frac{\ell+a+q}{\ell+m}, 1\right) \quad \text{and for } i \geq 2, \quad Z_i^{(q)} \sim \mathrm{Gamma}(1,1),$$

where, for any $k, \theta > 0$, the distribution $\mathrm{Gamma}(k,\theta)$ has density $x \mapsto \frac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} \mathbf{1}_{\{x>0\}}$. Then for all $k \geq 1$ we define $\mathcal{G}_k$ as,

$$\mathcal{G}_k := \prod_{q=0}^{m-1} \left(\sum_{i=1}^{k} Z_i^{(q)}\right)^{\frac{1}{m+\ell}}. \tag{3.50}$$

We say that the process $(\mathcal{G}_k)_{k\geq 1}$ has the distribution *Product of Generalised Gamma* with parameters (a,ℓ,m) which we denote PGG(a,ℓ,m).

The limiting chain is a PGG. Fix $\ell \geq 1$ and $m \geq 1$ some integers and suppose that the sequence $\mathbf{a}$ has the form

$$\mathbf{a} = a, \underbrace{0, 0, \ldots, 0}_{\ell-1}, m, \underbrace{0, 0, \ldots, 0}_{\ell-1}, m, \ldots,$$

meaning that for all $j \geq 0$ we have $a_{\ell \cdot j+1} = m$, and $a_n = 0$ whenever $n-1$ is not a multiple of ℓ, and $a_1 = a > -1$.

Proof of Proposition 3.28 (ii). For all $j \geq 1$ we let $\beta'_j := \beta_{\ell \cdot j}$ in our preceding notation. Of course $\beta_i = 1$ whenever i is not a multiple of ℓ, hence the sequence $(\mathsf{M}_n)_{n\geq 1}$ is constant on intervals of the type $[\![k\ell+1\,,(k+1)\ell]\!]$. Recall that for all $k \geq 1$ we denote $\mathsf{N}_k = \mathsf{M}_{\ell\cdot(k-1)+1}$. For any $j \geq 1$, we have

$$\beta'_j \sim \mathrm{Beta}\left(a + \ell + (j-1)\cdot(m+\ell), m\right).$$

For any $j \geq 1, p \geq 1$, we use the moments (3.39) of a Beta random variable and a telescoping argument to write

$$\mathbb{E}\left[(\beta'_j)^p\right] = \prod_{q=0}^{p-1} \frac{a+\ell+(j-1)(m+\ell)+q}{a+\ell+(j-1)(m+\ell)+m+q} = \prod_{q=0}^{m-1} \frac{a+\ell+(j-1)(m+\ell)+q}{a+\ell+(j-1)(m+\ell)+p+q}.$$

Then we compute, using the properties of the Gamma function,

$$\begin{aligned}\prod_{i=1}^{n-1} \mathbb{E}\left[(\beta_i)^p\right] = \prod_{j=1}^{\lfloor \frac{n-1}{\ell} \rfloor} \mathbb{E}\left[(\beta'_j)^p\right] &= \prod_{j=1}^{\lfloor \frac{n-1}{\ell} \rfloor} \prod_{q=0}^{m-1} \frac{a+\ell+(j-1)(m+\ell)+q}{a+\ell+(j-1)(m+\ell)+p+q} \\ &= \prod_{q=0}^{m-1} \frac{\Gamma\left(\lfloor \frac{n-1}{\ell} \rfloor + \frac{q+a+\ell}{m+\ell}\right)\Gamma\left(\frac{q+a+\ell+p}{m+\ell}\right)}{\Gamma\left(\lfloor \frac{n-1}{\ell} \rfloor + \frac{q+a+\ell+p}{m+\ell}\right)\Gamma\left(\frac{q+a+\ell}{m+\ell}\right)}.\end{aligned}$$

Using Stirling's approximation we get

$$\prod_{i=1}^{n-1} \mathbb{E}\left[(\beta_i)^p\right] \underset{n\to\infty}{\sim} n^{-\frac{pm}{m+\ell}} \cdot \ell^{\frac{pm}{m+\ell}} \cdot \prod_{q=0}^{m-1} \frac{\Gamma\left(\frac{q+a+\ell+p}{m+\ell}\right)}{\Gamma\left(\frac{q+a+\ell}{m+\ell}\right)}.$$

Hence, recalling the definition of C_p in (3.40), we get

$$C_p = \ell^{\frac{pm}{m+\ell}} \cdot \prod_{q=0}^{m-1} \frac{\Gamma\left(\frac{q+a+\ell+p}{m+\ell}\right)}{\Gamma\left(\frac{q+a+\ell}{m+\ell}\right)}.$$

Then using (3.46) with $c = m/\ell$,

$$\begin{aligned}\mathbb{E}\left[\mathsf{N}_k^p\right] = \mathbb{E}\left[\mathsf{M}^p_{1+(k-1)\ell}\right] = \frac{(c+1)^p \cdot C_p}{\prod_{i=1}^{(k-1)\ell} \mathbb{E}\left[\beta_i^p\right]} &= \left(\frac{m+\ell}{\ell}\right)^p \cdot \ell^{\frac{pm}{m+\ell}} \cdot \prod_{q=0}^{m-1} \frac{\Gamma\left(k-1+\frac{q+a+\ell+p}{m+\ell}\right)}{\Gamma\left(k-1+\frac{q+a+\ell}{m+\ell}\right)} \\ &= \left((m+\ell)\cdot(\ell^{\frac{-\ell}{m+\ell}})\right)^p \cdot \prod_{q=0}^{m-1} \frac{\Gamma\left(k-1+\frac{q+a+\ell}{m+\ell}+\frac{p}{m+\ell}\right)}{\Gamma\left(k-1+\frac{q+a+\ell}{m+\ell}\right)}.\end{aligned}$$

Using the last display and the fact that a random variable with distribution $\mathrm{Gamma}(x, 1)$ has p-th moment equal to $\frac{\Gamma(x+p)}{\Gamma(p)}$, we can identify the distribution of the marginals $\frac{\ell^{\frac{\ell}{m+\ell}}}{m+\ell} \cdot \mathsf{N}_k$ for any $k \geq 1$ with the ones of the process described in (3.50). The identification of the distribution of the whole process $\frac{\ell^{\frac{\ell}{m+\ell}}}{m+\ell} \cdot (\mathsf{N}_k)_{k\geq 1}$ with a $\mathrm{PGG}(a, \ell, m)$ is then obtained by checking that their backward transitions are the same. ◻

Remark 3.29. *For $m = a = 1$, the process $(\mathcal{G}_k)_{k\geq 1}$ has exactly the distribution of the points of a Poisson process on $\mathbb{R}_+$ with intensity $(\ell+1)t^\ell \mathrm{d}t$, listed in increasing order.*

Remark 3.30. *The distribution of $\mathcal{G}_1$ coincides with the one proved in [14] for the limiting proportion of some periodic Pólya urn, which is not a surprise because the degree of the first vertex in the tree follows exactly the urn dynamic that they study (with completely different tools).*

3.5.2 Applications to some other models of preferential attachment

Let us present here another model of preferential attachment which appears in the literature, for example in [100]. This model does not produce a tree as ours does, but we can couple them in such a way that some of their features coincide. We only focus on one particular model of graph here but the method presented here can be adapted to other similar models.

A model of (m,α)-preferential attachment Let $\mathtt{S}$ be a non-empty graph, with vertex-set $\{v_1^{(1)},\dots,v_1^{(k)}\}$ which have degrees $(d_1,\dots d_k)$, and $m\geq 2$ an integer and $\alpha > -m$ a real number such that $\alpha + d_i > 0$ for all $1\leq i\leq k$. The model is then the following: we let $\mathtt{G}_1 = \mathtt{S}$. Then, at any time $n\geq 1$, the graph $\mathtt{G}_{n+1}$ is constructed from the graph $\mathtt{G}_n$ by:

- adding a new vertex labelled v_{n+1} with m outgoing edges,
- choosing sequentially to which other vertex each of these edges are pointed, each vertex being chosen with probability proportional to α plus its degree (the degree of the vertices are updated after each edge-creation).

The degree of a vertex in a graph refers in this section to the number of edges incident to it. Here the growth procedure in fact produces *multigraphs*, in which it is possible for two vertices to be connected to each other by more than one edge. In this case, all those edges contribute to the count of their degree.

We can couple this model to a preferential attachment tree with sequence of initial fitnesses $\mathbf{a}$ defined as:

$$\mathbf{a} = (w(\mathtt{S}), \underbrace{0,0,\dots,0}_{m-1}, m+\alpha, \underbrace{0,0,\dots,0}_{m-1}, m+\alpha, 0, 0\dots),$$

where $w(\mathtt{S}) := d_1 + d_2 + \dots + d_k + k\alpha$.

Indeed, we can construct $(\mathtt{T}_n)$ with distribution $\mathrm{PA}(\mathbf{a})$. Then, for any $n\geq 1$, consider the tree $\mathtt{T}_{1+m(n-1)}$ and for all $2\leq i\leq n$, merge together each vertex with initial fitness $m+\alpha$ together with the $m-1$ vertices with fitness 0 that arrived just before it. If $\mathtt{G}_1$ only contains one vertex, it is immediate that the obtained sequence of graphs has exactly the same distribution as $(\mathtt{G}_n)_{n\geq 1}$. For general seed graphs $\mathtt{S}$, we can still use the same construction and the obtained sequence of graphs has the same evolution as some sequence $(\widetilde{\mathtt{G}}_n)_{n\geq 1}$ which would be obtained from $(\mathtt{G}_n)_{n\geq 1}$ by merging all the vertices $\{v_1^{(1)},\dots,v_1^{(k)}\}$ into a unique vertex v_1.

Note that a similar construction would also be possible if the initial degrees of the vertices $v_2, v_3, \dots$ were given by a sequence of integers $(m_2, m_3, \dots)$ instead of all being equal to some constant value m. This is for example the case in the model studied in [44], where the initial degrees are random.

We have the following convergence for degrees of vertices in the graph, as $n\to\infty$:

Proposition 3.31. *The following convergence holds almost surely in any ℓ^p with $p > 2+\frac{\alpha}{m}$:*

$$n^{-\frac{1}{2+\alpha/m}}(\deg_{\mathtt{G}_n}(v_1^{(1)}), \deg_{\mathtt{G}_n}(v_1^{(k)}), \dots, \deg_{\mathtt{G}_n}(v_1^{(k)}), \deg_{\mathtt{G}_n}(v_2), \deg_{\mathtt{G}_n}(v_3), \dots) \\ \underset{n\to\infty}{\longrightarrow} (\mathsf{N}_1 \cdot B^{(1)}, \mathsf{N}_1 \cdot B^{(2)}, \dots \mathsf{N}_1 \cdot B^{(k)}, \mathsf{N}_2 - \mathsf{N}_1, \mathsf{N}_3 - \mathsf{N}_2, \dots),$$

where

$$(B^{(1)}, B^{(2)}, \dots B^{(k)}) \sim \mathrm{Dir}(d_1+\alpha, d_2+\alpha, \dots, d_k+\alpha),$$

and the process $(\mathsf{N}_n)_{n\geq 1}$ is independent of $(B^{(1)}, B^{(2)}, \dots B^{(k)})$.

Furthermore, whenever $\alpha \in \mathbb{Z}$ or $m = 1$ then the distribution of $(\mathsf{N}_n)_{n\geq 1}$ is explicit and given by:

- $\frac{m^{\frac{-2m}{2m+\alpha}}}{2m+\alpha} \cdot (\mathsf{N}_n)_{n\geq 1} \sim \mathrm{PGG}(w(\mathsf{S}), m, m+\alpha)$ *if $\alpha \in \mathbb{Z}$,*
- $(\mathsf{N}_n)_{n\geq 1} \sim \mathrm{MLMC}\left(\frac{1}{2+\alpha}, \frac{w(\mathsf{S})}{2+\alpha}\right)$ *if $m = 1$.*

This result strengthens the one of [100, Theorem 1, Theorem 2 and Proposition 1] which corresponds (up to some definition convention) to the case $\alpha = 1 - m$. We emphasize that the convergence here is almost sure in an ℓ^p space.

Proof of Proposition 3.31. Using the coupling argument, we know that the sequence

$$((\deg_{\mathtt{G}_n}(v_1^{(1)}) - d_1) + (\deg_{\mathtt{G}_n}(v_1^{(2)}) - d_2) + \dots + (\deg_{\mathtt{G}_n}(v_1^{(k)}) - d_k), \\ (\deg_{\mathtt{G}_n}(v_2) - m), (\deg_{\mathtt{G}_n}(v_3) - m), \dots)$$

evolves as the out-degrees of the vertices in order of apparition in a preferential attachment tree $\mathrm{PA}(\mathbf{a})$ with sequence

$$\mathbf{a} = (w(S), \underbrace{0, 0, \dots, 0}_{m-1}, m+\alpha, \underbrace{0, 0, \dots, 0}_{m-1}, m+\alpha, \dots).$$

Using Theorem 3.1, Proposition 3.2 and Proposition 3.5 we get

$$n^{-\frac{1}{2+\alpha/m}}(\deg_{\mathtt{G}_n}(v_1) + \deg_{\mathtt{G}_n}(v_2) \dots + \deg_{\mathtt{G}_n}(v_k), \deg_{\mathtt{G}_n}(u_2), \deg_{\mathtt{G}_n}(u_3), \dots) \\ \underset{n\to\infty}{\longrightarrow} (\mathsf{N}_1, \mathsf{N}_2 - \mathsf{N}_1, \mathsf{N}_3 - \mathsf{N}_2, \dots),$$

almost surely in ℓ^p for all $p > 2+\frac{\alpha}{m}$, for some random sequence $(\mathsf{N}_k)_{k\geq 1}$. In the case $\alpha \in \mathbb{Z}$ or $m = 1$, Proposition 3.28 identifies the distribution of the limiting sequence. Last, the convergence of $\frac{1}{\deg_{\mathtt{G}_n}(v_1^{(1)})+\deg_{\mathtt{G}_n}(v_1^{(2)})+\dots+\deg_{\mathtt{G}_n}(v_1^{(k)})}(\deg_{\mathtt{G}_n}(v_1^{(1)}), \deg_{\mathtt{G}_n}(v_1^{(2)}), \dots, \deg_{\mathtt{G}_n}(v_1^{(k)}))$ just follows from the classical result of convergence for the proportion of balls in a Pólya urn. $\square$

3.A Technical proofs and results

This appendix contains the proofs of technical results that are used throughout this chapter. Let start by stating a useful conditional version of the Borel-Cantelli lemma.

Lemma 3.32. *Let $(\mathcal{F}_n)$ be a filtration and let $(B_n)_{n\geq 1}$ be a sequence of events adapted to this filtration. For all $n \geq 1$, let $\mathsf{p}_n := \mathbb{P}(B_n \mid \mathcal{F}_{n-1})$. We have*

$$\frac{\sum_{i=1}^n \mathbf{1}_{B_i}}{\sum_{i=1}^n \mathsf{p}_i} \underset{n\to\infty}{\to} 1 \qquad \text{a.s. on the event} \quad \left\{\sum_{i=1}^{\infty} \mathsf{p}_i = \infty\right\}$$

and also

$$\sum_{i=1}^n \mathbf{1}_{B_i} \quad \text{converges a.s. on the event} \quad \left\{\sum_{i=1}^{\infty} \mathsf{p}_i < \infty\right\}.$$

Proof. The first convergence is the content of Theorem 5.4.11 and the second one is an application of Theorem 5.4.9, both taken from [56]. □

The following lemma is a rewriting of [26, Lemma 1]. We provide the proof for completeness.

Lemma 3.33 ("Biggins' lemma"). *Let $(M_n)_{n\geq 1}$ be a complex-valued martingale with finite q-th moment for some $q \in [1\,,2]$. Then for every $n \geq 1$*

$$\mathbb{E}\left[|M_{n+1}|^q\right] \leq \mathbb{E}\left[|M_n|^q\right] + 2^q \cdot \mathbb{E}\left[|M_{n+1} - M_n|^q\right].$$

Proof. Let $X_{n+1} := M_{n+1} - M_n$ and let X'_{n+1} be a random variable such that conditionally on $(M_1, \dots, M_n)$ the random variable X'_{n+1} is independent of, and has the same distribution as, X_{n+1}. Then

$$\begin{aligned}
\mathbb{E}\left[|M_{n+1}|^q\right] &= \mathbb{E}\left[\left|\mathbb{E}\left[M_{n+1} - X'_{n+1} \mid M_1, \dots M_{n+1}\right]\right|^q\right] \\
&\leq \mathbb{E}\left[\left|M_{n+1} - X'_{n+1}\right|^q\right] \\
&= \mathbb{E}\left[\left|M_n + X_{n+1} - X'_{n+1}\right|^q\right] \\
&\leq \mathbb{E}\left[|M_n|^q\right] + \mathbb{E}\left[\left|X_{n+1} - X'_{n+1}\right|^q\right] \\
&\leq \mathbb{E}\left[|M_n|^q\right] + 2^q \cdot \mathbb{E}\left[|X_{n+1}|^q\right],
\end{aligned}$$

where the first equality comes from the fact that $\mathbb{E}\left[X'_{n+1} \mid M_1, \dots M_{n+1}\right] = 0$. The first inequality is the one of Jensen for conditional expectation, applied to the convex function $z \mapsto |z|^q$. The second inequality is due to Clarkson, see [13, Lemma 1], and can be applied because the distribution of $X_{n+1} - X'_{n+1}$ conditional on M_n is symmetric and $1 \leq q \leq 2$. The last inequality comes from the triangle inequality for the L^q-norm. □

Let us state another result about martingales, which we use numerous times throughout the chapter. Recall our uniform big-O and small-o notation, introduced in (3.22).

Lemma 3.34. *Suppose that $(z \mapsto Z_n(z))_{n\geq 1}$ is a sequence of analytic functions on some open domain $\mathscr{O} \subset \mathbb{C}$, adapted to some filtration $(\mathcal{G}_n)$. Suppose that for every $z \in \mathscr{O}$, the sequence $(Z_n(z))_{n\geq 1}$ is a martingale with respect to the filtration $(\mathcal{G}_n)$. If there exists a parameters $q > 1$ and continuous functions $\alpha : \mathscr{O} \to \mathbb{R}$ and $\delta : \mathscr{O} \to \mathbb{R}_+^*$ such that for all $n \geq 1$ we have*

$$\mathbb{E}\left[|Z_{2n}(z) - Z_n(z)|^q\right] = O_{\mathscr{O}}\left(n^{\alpha(z)q - \delta(z)}\right),$$

then for any compact $K \subset \mathscr{O}$, there exists $\epsilon(K) > 0$ such that

(i) *if $\alpha > 0$ on $\mathscr{O}$ we have $n^{-\alpha(z)} \cdot |Z_n(z)| = O_K\left(n^{-\epsilon(K)}\right)$ almost surely and also in expectation,*

(ii) *if $\alpha \leq 0$ on $\mathscr{O}$, the almost sure limit $Z_\infty(z)$ exists for $z \in \mathscr{O}$ and we have $n^{-\alpha(z)} \cdot |Z_n(z) - Z_\infty(z)| = O_K\left(n^{-\epsilon(K)}\right)$ almost surely and also in expectation.*

Proof of Lemma 3.34. By compactness, it is sufficient to prove the result for a small disk around each $x \in K$. Since $\mathscr{O}$ is an open set, let $\rho > 0$ be such that $\mathrm{D}(x, 2\rho) \subset \mathscr{O}$, where $\mathrm{D}(x, 2\rho)$ is the closed disk in the complex plane with centre x and radius 2ρ. We denote

$$\underline{\alpha} = \inf_{\mathrm{D}(x,2\rho)} \alpha, \quad \overline{\alpha} = \sup_{\mathrm{D}(x,2\rho)} \alpha, \quad \underline{\delta} = \inf_{\mathrm{D}(x,2\rho)} \delta,$$

and choose ρ small enough so that $\underline{\alpha} - \overline{\alpha} + \frac{1}{q}\underline{\delta} > 0$. Then if we let $\xi : [0\,, 2\pi] \to \mathbb{C}$ such that $\xi(t) = x + 2\rho e^{it}$, we have for any n and m, using Cauchy formula

$$\sup_{z \in \mathrm{D}(x,\rho)} |Z_n(z) - Z_m(z)| \leq \pi^{-1} \int_0^{2\pi} |Z_n(\xi(t)) - Z_m(\xi(t))| \mathrm{d}t.$$

Now,

$$\begin{aligned} \sup_{2^s \leq n \leq 2^{s+1}} \sup_{z \in \mathrm{D}(x,\rho)} |Z_n(z) - Z_{2^s}(z)| &\leq \pi^{-1} \sup_{2^s \leq n \leq 2^{s+1}} \int_0^{2\pi} |Z_n(\xi(t)) - Z_{2^s}(\xi(t))| \mathrm{d}t \\ &\leq \pi^{-1} \int_0^{2\pi} \sup_{2^s \leq n \leq 2^{s+1}} |Z_n(\xi(t)) - Z_{2^s}(\xi(t))| \mathrm{d}t. \end{aligned} \tag{3.51}$$

Sequentially using Jensen's inequality and Doob's maximal inequality in L^q, gives us for every $z \in \mathrm{D}(x, \rho)$:

$$\begin{aligned} \mathbb{E}\left[\sup_{2^s \leq n \leq 2^{s+1}} |Z_n(z) - Z_{2^s}(z)|\right] &\leq \mathbb{E}\left[\sup_{2^s \leq n \leq 2^{s+1}} |Z_n(z) - Z_{2^s}(z)|^q\right]^{\frac{1}{q}} \\ &\leq \frac{q}{q-1} \cdot \mathbb{E}\left[|Z_{2^{s+1}}(z) - Z_{2^s}(z)|^q\right]^{\frac{1}{q}} \\ &\underset{s \to \infty}{=} O_{\mathrm{D}(x,2\rho)}\left(2^{\left(\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}\right)s}\right). \end{aligned} \tag{3.52}$$

So using (3.51), Fubini's theorem and (3.52), we get

$$\begin{aligned} \mathbb{E}\left[\sup_{2^s \leq n \leq 2^{s+1}} \sup_{z \in \mathrm{D}(x,\rho)} |Z_n(z) - Z_{2^s}(z)|\right] &\leq \pi^{-1} \int_0^{2\pi} \mathbb{E}\left[\sup_{2^s \leq n \leq 2^{s+1}} |Z_n(\xi(t)) - Z_{2^s}(\xi(t))|\right] \mathrm{d}t \\ &\underset{s \to \infty}{=} O\left(2^{\left(\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}\right)s}\right). \end{aligned}$$

Now let us treat the two cases $\alpha > 0$ and $\alpha \leq 0$ separately. Remark that the quantity $\left(\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}\right)$ is negative when $\alpha \leq 0$, but can be of any sign in the case $\alpha > 0$.

- For $\alpha > 0$ and $n \geq 1$, let $r \in \mathbb{N}$ be such that $2^r \leq n \leq 2^{r+1}$ and write

$$\begin{aligned} \mathbb{E}\left[n^{-\alpha(z)} \sup_{1 \leq k \leq n} \sup_{z \in \mathrm{D}(x,\rho)} |Z_n(z) - Z_1(z)|\right] &\leq 2^{-\underline{\alpha}r} \cdot \sum_{s=0}^{r} \mathbb{E}\left[\sup_{2^s \leq n \leq 2^{s+1}} \sup_{z \in \mathrm{D}(x,\rho)} |Z_n(z) - Z_{2^s}(z)|\right] \\ &\leq \mathrm{cst} \cdot 2^{-\underline{\alpha}r} \cdot \sum_{s=0}^{r} 2^{\left(\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}\right)s} \\ &\leq \mathrm{cst} \cdot 2^{-(\underline{\alpha} - 0 \vee (\overline{\alpha} - \frac{1}{q} \cdot \underline{\delta}))r}. \end{aligned}$$

The expectation of the right-hand side tends to 0 exponentially fast in r hence also almost surely, which proves point (i).

• For $\alpha \leq 0$ and $n \geq 1$, let $r \in \mathbb{N}$ be such that $2^r \leq n \leq 2^{r+1}$ and write

$$\begin{aligned}
\mathbb{E}\left[n^{-\alpha(z)} \sup_{k\geq n} \sup_{z\in \mathrm{D}(x,\rho)} |Z_k(z) - Z_n(z)|\right] &\leq \mathrm{cst} \cdot 2^{-\underline{\alpha} r} \cdot \sum_{s=r}^{\infty} \mathbb{E}\left[\sup_{2^s\leq k\leq 2^{s+1}} \sup_{z\in \mathrm{D}(x,\rho)} |Z_k(z) - Z_{2^s}(z)|\right] \\
&\leq \mathrm{cst} \cdot 2^{-\underline{\alpha} r} \cdot \sum_{s=r}^{\infty} 2^{\left(\overline{\alpha} - \frac{1}{q}\cdot \underline{\delta}\right)s} \\
&\leq \mathrm{cst} \cdot 2^{-(\underline{\alpha} - \overline{\alpha} + \frac{1}{q}\cdot\underline{\delta})r}.
\end{aligned}$$

and the last display converges exponentially fast to 0. So the function $z \mapsto Z_n(z)$ converges almost surely to some $z \mapsto Z_\infty(z)$ uniformly on the disc, and point (ii) is satisfied. $\square$

Finally, let us give a proof of Lemma 3.12.

Proof of Lemma 3.12. Let $\epsilon > 0$ and suppose that $W_n = \mathrm{cst} \cdot n^\gamma + O(n^{\gamma-\epsilon})$ as $n \to \infty$. It is immediate that $w_n = W_{n+1} - W_n = O(n^{\gamma-\epsilon})$. Then

$$\sum_{i=n}^{2n} \left(\frac{w_i}{W_i}\right)^2 \leq \frac{1}{W_n^2} \cdot \max_{n\leq i\leq 2n} w_i \cdot \sum_{i=n}^{2n} w_i \leq \frac{W_{2n}}{W_n^2} \cdot \max_{n\leq i\leq 2n} w_i = O\big(n^{-\epsilon}\big),$$

and the first point follows by summing over intervals of the type $[\![n2^k , n2^{k+1}]\!]$.

Now write

$$\frac{W_1}{W_n} = \prod_{i=2}^{n} \frac{W_{i-1}}{W_i} = \prod_{i=2}^{n} \left(1 - \frac{w_i}{W_i}\right) = \exp\left(\sum_{i=2}^{n} \log\left(1 - \frac{w_i}{W_i}\right)\right).$$

Since $\frac{w_i}{W_i} \to 0$ as $n \to \infty$, we get

$$\log\left(1 - \frac{w_i}{W_i}\right) = -\frac{w_i}{W_i} + O\left(\left(\frac{w_i}{W_i}\right)^2\right)$$

Putting everything together, we get

$$\begin{aligned}
\sum_{i=2}^{n} \frac{w_i}{W_i} &= -\sum_{i=2}^{n} \log\left(1 - \frac{w_i}{W_i}\right) + \sum_{i=2}^{n} O\left(\left(\frac{w_i}{W_i}\right)^2\right) \\
&= \log W_n - \log W_1 + \sum_{i=2}^{\infty} O\left(\left(\frac{w_i}{W_i}\right)^2\right) - O\left(\sum_{i=n+1}^{\infty} \left(\frac{w_i}{W_i}\right)^2\right) \\
&= \log W_n + \mathrm{cst} + O\big(n^{-\epsilon}\big).
\end{aligned}$$

Last, just remark that $\log W_n = \log(\mathrm{cst} \cdot n^\gamma \cdot (1 + O(n^{-\epsilon}))) = \gamma \log n + \mathrm{cst} + O(n^{-\epsilon})$, which finishes the proof. $\square$

Chapter 4

Growing random graphs with a preferential attachment structure

This chapter is adapted from the work in progress [111].

The aim of this chapter is to develop a method for proving almost sure convergence in Gromov-Hausdorff-Prokhorov topology for models of growing random graphs that have some hidden preferential attachment structure. We describe the obtained limits using some iterative gluing construction that generalises the famous line-breaking construction of Aldous' Brownian tree. We develop a framework which allows us to handle metric spaces seen as the gluing of metric spaces along a tree structure. We prove the convergences using an argument of "finite-dimensional" convergence together with some relative compactness property. This approach strongly relies on results for preferential attachment and weighted recursive trees obtained in Chapter 3.

Contents

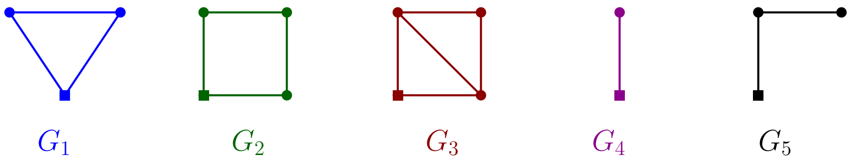


Figure 4.1 – An example of a sequence of graphs used to run the algorithm, the root of each graph is represented by a square vertex

4.1 Introduction

Let us introduce a generalised version of Rémy's algorithm [107], which should be considered as a particular example of the models that are handled by our method. Other models are discussed at the end of the introduction.

4.1.1 A generalised version of Rémy's algorithm

Consider $(G_n)_{n\geq 1}$ a sequence of finite connected rooted graphs and construct the sequence $(H_n)_{n\geq 1}$ recursively as follows. Let $H_1 = G_1$. Then, for any $n \geq 1$, conditionally on the structure H_n already constructed, take an edge in H_n uniformly at random, split it into two edges by adding a vertex "in the middle" of this edge, and glue a copy of G_{n+1} to the structure by identifying the root vertex of G_{n+1} with the newly created vertex. Call the obtained graph H_{n+1}.

When all the graphs $(G_n)_{n\geq 1}$ are equal to the single-edge graph, we obtain the so-called Rémy's algorithm, which produces for each n a uniform planted binary tree with n leaves (if the leaves are labelled for example). Remark that this kind of generalisation of the algorithm has already been studied for particular sequences $(G_n)_{n\geq 1}$, namely for $(G_n)_{n\geq 1}$ constant equal to the star-graph with $k-1$ branches in [71], where the authors show that the obtained trees converge in the scaling limit to some fragmentation tree; and in [109], for G_n being equal to the single-edge graph for every $n \equiv 1 \mod \ell$, for some $\ell \geq 2$, and equal to the single-vertex graph whenever $n \not\equiv 1 \mod \ell$.

We see the graphs $(H_n)_{n\geq 1}$ as measured metric spaces, by considering their set of vertices endowed with the usual graph distance and the uniform measure on vertices. It is well-known [42] that the sequence of trees created through the standard Rémy's algorithm with distances rescaled by $n^{-1/2}$ converges almost surely in the Gromov-Hausdorff-Prokhorov topology to a constant multiple of Aldous' Brownian tree. We give here an analogous result, under some conditions on the sequence $(G_n)_{n\geq 1}$, which ensures that the graphs $(H_n)_{n\geq 1}$ appropriately rescaled converge almost surely in the Gromov-Hausdorff-Prokhorov topology to a random compact metric space.

Proposition 4.1. *Call $(a_n)_{n\geq 1}$ the respective numbers of edges in the graphs $(G_n)_{n\geq 1}$. Suppose there exists $c > 0$ and $0 \leq c' < \frac{1}{c+1}$ and $\epsilon > 0$ such that*

$$\sum_{i=1}^{n} a_i = c \cdot n \cdot \left(1 + O\left(n^{-\epsilon}\right)\right), \qquad \text{and} \qquad a_n \leq n^{c'+o(1)},$$

then we have the following convergence, almost surely in the GHP topology

$$\left(H_n, n^{\frac{-1}{c+1}} \cdot \mathrm{d_{gr}}, \mu_{\mathrm{unif}}\right) \underset{n\to\infty}{\longrightarrow} (\mathcal{H}, \mathrm{d}, \mu). \tag{4.1}$$

As for Aldous' Brownian tree, the limiting random compact metric space $(\mathcal{H}, \mathrm{d}, \mu)$, which depends on the whole sequence $(G_n)_{n\geq 1}$, can be described as the result of an *iterative gluing construction*, as defined in Chapter 2. It is a natural extension of the famous *line-breaking construction* invented by Aldous [9], but with branches that are allowed to be more complex than just segments.

A line-breaking construction. The construction of $(\mathcal{H}, \mathrm{d}, \mu)$ is described as follows. We first run some increasing time-inhomogeneous Markov chain $(\mathsf{M}_n^{\mathbf{a}})_{n\geq 1}$ which takes values in $\mathbb{R}_+$, and whose law depends only on the sequence $\mathbf{a} = (a_n)_{n\geq 1}$.

Cut the semi-infinite line $\mathbb{R}_+$ at the values taken by the chain, this creates an ordered sequence of segments with length $\mathsf{M}_1^{\mathbf{a}}, (\mathsf{M}_2^{\mathbf{a}} - \mathsf{M}_1^{\mathbf{a}}), (\mathsf{M}_3^{\mathbf{a}} - \mathsf{M}_2^{\mathbf{a}}), \dots$. Now for any $n \geq 1$, we do the following:

(i) Cut the n-th segment into a_n sub-segments by throwing $a_n - 1$ uniform points on it, and call $(L_{n,1}, L_{n,2}, \dots, L_{n,a_n})$ the respective lengths of the obtained sub-segments.

(ii) Take the graph G_n and replace every edge $e_k \in \{e_1, \dots, e_{a_n}\}$, where the edges of G_n are labelled in an arbitrary order, with a segment of length $L_{n,k}$. Call the result $\mathcal{G}_n$.

Now, start from $\mathcal{H}_1 := \mathcal{G}_1$ and recursively when $\mathcal{H}_n$ is already constructed, sample a point according to the length measure on $\mathcal{H}_n$ and identify the root of $\mathcal{G}_{n+1}$ to the chosen point. The space $\mathcal{H}$ is obtained as the completion of the increasing union

$$\mathcal{H} = \overline{\bigcup_{n\geq 1} \mathcal{H}_n}.$$

The measure μ is the weak limit of the normalised length measure carried by the $\mathcal{H}_n$'s.

4.1.2 Metric spaces glued along a tree structure

We introduce a general framework that allows us to handle objects that are defined as the result of gluing together metric spaces along a discrete tree structure. Consider the Ulam tree with its usual representation as

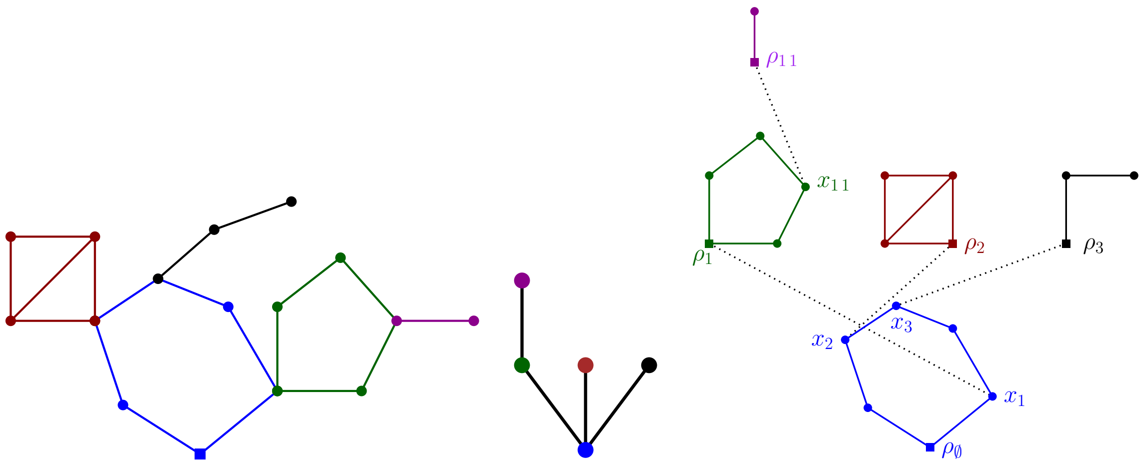
$$\mathbb{U} = \bigcup_{n\geq 0} \mathbb{N}^n. \tag{4.2}$$

We say that $\mathcal{D} = (\mathcal{D}(u))_{u\in\mathbb{U}}$ is a *decoration* on the Ulam tree if for any $u \in \mathbb{U}$,

$$\mathcal{D}(u) = (D_u, d_u, \rho_u, (x_{ui})_{i\geq 1}),$$

is a compact rooted metric space, with underlying set D_u, distance d_u, rooted at a point ρ_u and endowed with a sequence $(x_{ui})_{i\geq 1} \in D_u$. Then for any such decoration $\mathcal{D}$, we make sense of the following metric space $\mathscr{G}(\mathcal{D})$, which is informally what we get if we take the disjoint union $\bigsqcup_{u\in\mathbb{U}} D_u$ and identify every root $\rho_{ui} \in D_{ui}$ to the distinguished point $x_{ui} \in D_u$ for every $u \in \mathbb{U}$ and every $i \in \mathbb{N}$, and take the metric completion of the obtained metric space.

This setting also encompasses the case where we only glue a finite number of blocks along a plane tree. If τ is a plane tree, it can be natural to consider a decoration $\mathcal{D} = (\mathcal{D}(u))_{u\in\tau}$ which is only defined on the vertices of τ and is such that for all $u \in \tau$, the block $\mathcal{D}(u) = (D_u, d_u, \rho_u, (x_{ui})_{1\leq i\leq \deg_\tau^+(u)})$ is only endowed with a finite number of distinguished points that corresponds to the number $\deg_\tau^+(u)$ of children of u in τ. In this case we automatically extend $\mathcal{D}$ by letting $\mathcal{D}(u)$ be the one-point space $(\star, 0, \star, (\star)_{i\geq 1})$ for all $u \notin \mathbb{U}$ and by letting $x_{ui} = \rho_i$ for all $u \in \tau$ and $i \geq \deg_\tau^+(u)$. Thanks to this identification, we always consider decorations that are defined on the whole Ulam tree $\mathbb{U}$ and for which all the blocks have infinitely many distinguished points.



(a) A realisation of H_5

(b) The tree $\mathtt{P}_5$

(c) The graph H_5 seen as the gluing of a decoration

Figure 4.2 – Decomposition of H_5 as the gluing of some decoration. The graphs $(G_n)_{n\geq 1}$ used for the construction are the one appearing in Figure 4.1.

Convergence of metric spaces glued along $\mathbb{U}$. For a sequence $(\mathcal{D}_n)_{n\geq 1}$ of decorations, we have a sufficient condition for the convergence of the sequence $(\mathscr{G}(\mathcal{D}_n))_{n\geq 1}$ in the Gromov-Hausdorff topology. Indeed, we will see in Theorem 4.2 that it suffices that for every $u \in \mathbb{U}$, we have the convergence $\mathcal{D}_n(u) \to \mathcal{D}_\infty(u)$ for some decoration $\mathcal{D}_\infty$ in the some appropriate infinitely pointed Gromov-Hausdorff topology and that the sequence of decorations $(\mathcal{D}_n)_{n\geq 1}$ satisfies the relative compactness property

$$\inf_{\substack{\theta\subset\mathbb{U}\\ \theta \text{ plane tree}}} \sup_{u\in\mathbb{U}} \left(\sum_{\substack{v\prec u\\ v\notin\theta}} \sup_{n\geq 1} \operatorname{diam}(\mathcal{D}_n(v)) \right) = 0,$$

to get $\mathscr{G}(\mathcal{D}_n) \to \mathscr{G}(\mathcal{D}_\infty)$ in the Gromov-Hausdorff topology as $n \to \infty$. With some appropriate assumptions, we can also endow these metric spaces with measures and get a similar statement in Gromov-Hausdorff-Prokhorov topology. We recall the definition and some properties of those topologies in Section 4.2.3.

Scaling limit for the generalised Rémy algorithm. This framework will allow us to prove Proposition 4.1. The idea is to interpret $(H_n)_{n\geq 1}$ in this framework by constructing a sequence of decorations $(\mathcal{D}_n)_{n\geq 1}$, in such a way that for all $n \geq 1$ the graph H_n seen as a metric space coincides with $\mathscr{G}(\mathcal{D}_n)$. That way, the problem of understanding the whole structure of H_n is decomposed into the easier problem of understanding separately all the $\mathcal{D}_n(u)$ for all $u \in \mathbb{U}$.

In fact, the construction of such a sequence $(\mathcal{D}_n)_{n\geq 1}$ is naturally coupled with the construction of a sequence $(\mathtt{P}_n)_{n\geq 1}$ of preferential attachment trees with initial fitnesses $(a_n)_{n\geq 1}$, the definition of which is recalled in Section 4.3, which were studied in Chapter 3. The sequence $(\mathtt{P}_n)_{n\geq 1}$ is an increasing sequence of plane trees, so in particular they can be seen as subsets of $\mathbb{U}$.

With this particular construction, each process $(\mathcal{D}_n(u))_{n\geq 1}$ for a fixed $u \in \mathbb{U}$ only evolves at times n when the degree of u evolves in the tree $(\mathtt{P}_n)_{n\geq 1}$ and stays constant otherwise. Also, at times where the block $\mathcal{D}_n(u)$ evolves, it does so independently of all the other blocks and follows some simple dynamics. This allows us to study the evolution of the processes $(\mathcal{D}_n(u))_{n\geq 1}$, including their scaling limit, separately for every $u \in \mathbb{U}$.

The fact that the limiting metric space can be described using an iterative gluing construction depends crucially on the fact that the distribution of the trees $(\mathsf{P}_n)_{n\geq 1}$ can also be expressed as that of a weighted recursive tree (see Section 4.3 for a definition) using the random sequence $(\mathsf{M}^{\mathbf{a}}_n - \mathsf{M}^{\mathbf{a}}_{n-1})_{n\geq 1}$ to which we referred above in definition of the line-breaking construction.

Two families of continuous distributions on decorations. Aside from iterative gluing constructions, an example of which we already mentioned, we define the family of *self-similar* decorations. Under some assumptions, the distribution of the gluing $\mathscr{G}(\mathcal{D})$ of some random self-similar decoration $\mathcal{D}$ is the unique fixed point of some contraction in an appropriate space of distribution on metric spaces, in the same spirit as the self-similar random trees of Rembart and Winkel in [106]. Some distributions on decorations can belong to both of these families, which is often the case for distributions arising as scaling limits of some natural discrete models.

4.1.3 Scope of our results and their relation to previous work

Let us discuss the results proved in this chapter and how they are related to the existing literature.

Subcases of the generalised Rémy's algorithm. Proposition 4.1 already encompasses several models that were already studied using other methods, when specifying particular sequences of graphs $(G_n)_{n\geq 1}$.

- Of course, we recover the convergence for the standard Rémy's algorithm whenever $(G_n)_{n\geq 1}$ is constant and taken to be a single-edge graph.
- When $(G_n)_{n\geq 1}$ is the constant sequence equal to a vertex with a single loop, the model is equivalent to the looptree of the linear preferential attachment tree, and we recover the convergence proved in [40].
- In [71], Haas and Stephenson study the case where G_1 is the single-edge graph and the sequence $(G_n)_{n\geq 2}$ is constant equal to the star-graph with $k-1$ branches, for $k\geq 2$. They describe the scaling limit as a fragmentation tree, as introduced in [68]. In this case, we improve their convergence which was only in probability and give another construction of the limit.
- Let us also cite the work of Ross and Wen [109], whose model (depending on an integer-valued parameter $\ell\geq 2$) is obtained by setting G_n to be a single-edge graph if $n-1$ is a multiple of ℓ, and reduced to a single vertex otherwise. We recover their results.
- In an ongoing work of Haas and Stephenson [72] the authors also study the case where $(G_n)_{n\geq 1}$ is taken as an i.i.d. sequence of rooted trees taken from a finite set. They describe the limit as a multi-type fragmentation tree as introduced in [114]. Again, our result ensures that the convergence is almost sure in the Gromov-Hausdorff-Prokhorov topology and gives another construction of the limit.

Other models of growing random graphs. Our general method can be applied to various models of growth such as Ford's α-model [60], Marchal's algorithm [92] or their generalisation the $\alpha-\gamma$-growth [37], possibly started from an arbitrary graph. The same methods apply also for discrete *looptrees* associated to those models (using an appropriate planar embedding) or to planar preferential attachment trees. We note the following:

- We improve the convergence [69, 67] of Ford trees and α-γ-trees, from convergence in probability to almost sure convergence, and also prove the convergence of their respective discrete looptrees to continuous limits which can be described as the result of iterative gluing constructions.
- We provide a new iterative gluing construction for α-stable trees and α-stable components, different from the ones appearing in [63, 64].
- We prove a conjecture of Curien et al. [40] for the scaling limits of looptrees of planar preferential attachment trees with offset δ, and describe the limit as an iterative gluing construction with circles.

4.1.4 Organisation of the paper

This chapter is organised as follows.

We start in Section 4.2 by developing a framework that allows us to define the gluing of infinitely many metric spaces along the structure of the Ulam tree. We prove Theorem 4.2 which ensures that this procedure is continuous in some sense with respect to the blocks that we glue together, as soon as they satisfy some relative compactness property. Then in Section 4.3 we recall some properties of affine preferential attachment trees and weighted recursive trees that were proved in the companion paper [110], and on which our study of sequences of growing graphs in Section 4.5 strongly relies. In Section 4.4 we present two families of distributions on decorations, the iterative gluing constructions and the self-similar decorations, which can appear as continuous limits of the discrete distributions that we study. We derive some of their properties, in particular we give some sufficient condition for the associated metric space obtained under the gluing map $\mathscr{G}$ to be compact almost surely. We also provide examples of random decorations that belong to both families of distributions. Last, in Section 4.5, we apply the preceding result to obtain scaling limits of some families of growing random graphs. We first start by proving Theorem 4.10, which is the general case in which our scaling limit results apply. The rest of the section is devoted to applying this theorem to examples of growing random graphs.

Appendix 4.A contains some computations related to the application of our results to some specific models.

4.2 Gluing metric spaces along the Ulam tree

In this section, we introduce what we call *decorations* on the Ulam tree, which are families of infinitely pointed compact metric spaces, indexed by the vertices of the Ulam tree. This structure should be thought of as a plan that specifies how to construct a metric space by gluing together all those decorations onto one another, along the structure of the Ulam tree. We then provide sufficient conditions that ensures that the resulting metric space is compact and depends continuously on the decorations in a sense that we make precise.

4.2.1 The Ulam tree

The completed Ulam tree. Recall the definition of the Ulam tree $\mathbb{U} = \bigcup_{n\geq 0} \mathbb{N}^n$. We also introduce the set $\partial\mathbb{U} = \mathbb{N}^{\mathbb{N}}$ to which we refer as the *leaves* of the Ulam tree, which we see as the infinite rays joining the root to infinity and let $\overline{\mathbb{U}} := \mathbb{U} \cup \partial\mathbb{U}$. On this set, we have a natural genealogical order $\preceq$ defined such that $u \preceq v$ if and only if u is a prefix of v. From this order we can define for any $u \in \mathbb{U}$ the subtree descending from u as the set $T(u) := \{v \in \mathbb{U} \mid u \preceq v\}$. The collection of sets $\{T(u),\ u \in \mathbb{U}\}$ and $\{\{u\},\ u \in \mathbb{U}\}$ generate a topology on $\overline{\mathbb{U}}$, which can also be generated using the appropriate ultrametric distance. Endowed with this distance the set $\overline{\mathbb{U}}$ is then a separable and complete metric space.

Plane trees as subsets of $\mathbb{U}$. Classically, a plane tree τ is defined as a finite non-empty subset of $\mathbb{U}$ such that

(i) if $v \in \tau$ and $v = ui$ for some $i \in \mathbb{N}$, then $u \in \tau$,

(ii) for all $u \in \tau$, there exists $\deg_\tau^+(u) \in \mathbb{N} \cup \{0\}$ such that for all $i \in \mathbb{N}$, $ui \in \tau$ iff $i \leq \deg_\tau^+(u)$.

We denote $\mathbb{T}$ the set of planes trees.

Elements of notation. Let us define some pieces of notation.

- Elements of $\overline{\mathbb{U}}$ are defined as finite or infinite sequences of integers, which we handle as words on the alphabet $\mathbb{N}$, we usually use the symbols u or v to denote elements of this space.
- Sometimes we also use a bold letter $\mathbf{i}$ to denote a finite or infinite word $\mathbf{i} = i_1 i_2 \dots$. In this case, for any integer k smaller than the length of $\mathbf{i}$ we also write $\mathbf{i}_k = i_1 \dots i_k$ for the word truncated to its k first letters.
- For any two $u, v \in \overline{\mathbb{U}}$, we write $u \wedge v$ for the most recent common ancestor of u and v.
- For any $u \in \mathbb{U}$, the height of u is the unique number n such that $u \in \mathbb{N}^n$. We denote it by $\mathrm{ht}(u)$ or sometimes also $|u|$.

4.2.2 Decorations on the Ulam tree

In a general manner, we call any function $f : \mathbb{U} \to E$ from the Ulam tree to a space E an E-valued *decoration* on the Ulam tree.

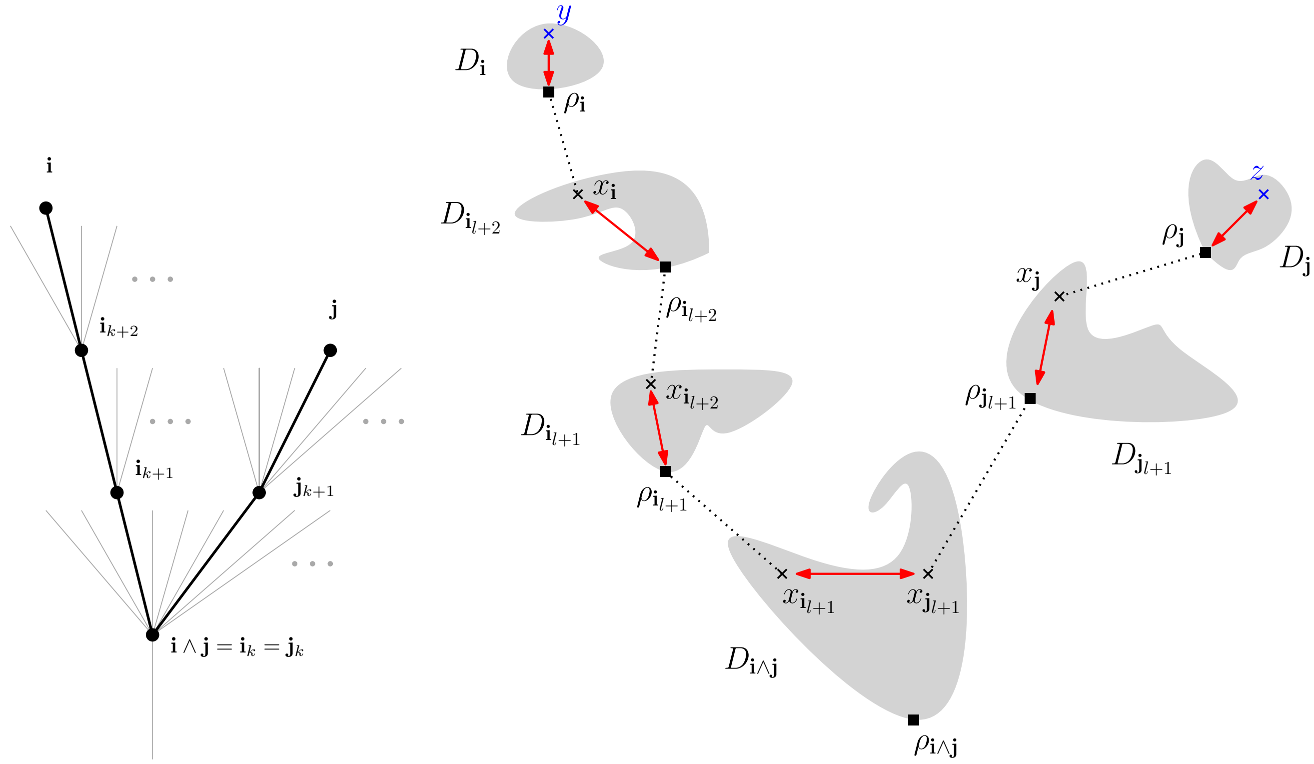
Real-valued decorations. As a first example, a function $\ell : \mathbb{U} \to \mathbb{R}_+$ is a real-valued decoration on the Ulam tree. We say that ℓ is *non-explosive* if

$$\inf_{\theta \in \mathbb{T}} \sup_{u \in \mathbb{U}} \left(\sum_{\substack{v \preceq u \\ v \notin \theta}} \ell(v) \right) = 0. \tag{4.3}$$

Metric space-valued decorations. One of the main objects studied in this chapter are decorations $\mathcal{D} : \mathbb{U} \to \mathbb{M}^{\infty\bullet}$, where the set $\mathbb{M}^{\infty\bullet}$ is the set of non-empty compact metric spaces endowed with an infinite sequence of distinguished points, up to isometry (see below for a proper definition). More precisely

$$\mathcal{D} : u \mapsto \mathcal{D}(u) = (D_u, d_u, \rho_u, (x_{ui})_{i\geq 1}),$$

where D_u is a set, d_u is a distance function on D_u, and ρ_u and the $(x_{ui})_{i\geq 1}$ are distinguished points of D_u. The point ρ_u is called the *root* of $\mathcal{D}(u)$, and we sometimes call $\mathcal{D}(u)$ a *block* of the decoration.



(a) The path between $\mathbf{i}$ and $\mathbf{j}$ in the Ulam tree

(b) The contribution from every decoration along the path

Figure 4.3 – The distance between two points is computed as the sum of the contributions denoted in red, computed using the distance in the corresponding decoration.

Let us define a particular element of $\mathbb{M}^{\infty\bullet}$, which we call the *trivial* or *one-point* space $(\{\star\}, 0, \star, (\star)_{i\geq 1})$. For any decoration $\mathcal{D}$, the subset $S \subset \mathbb{U}$ of elements u for which $\mathcal{D}(u)$ is not trivial is called the *support* of the decoration $\mathcal{D}$. In the rest of this chapter we will often consider decorations that are supported on finite plane trees.

For $a > 0$, we will use the notation $a \cdot \mathcal{D}$ to denote the decoration created from $\mathcal{D}$ by multiplying all the distances in all the blocks by a factor a.

The gluing operation. We define a *gluing operation* $\mathscr{G}$ on the set of metric-space-valued decorations $(\mathbb{M}^{\infty\bullet})^{\mathbb{U}}$. For any $\mathcal{D} = (\mathcal{D}(u))_{u\in\mathbb{U}}$ we first define the metric space $\mathscr{G}^*(\mathcal{D})$ as

$$\mathscr{G}^*(\mathcal{D}) = \left(\bigsqcup_{u\in\mathbb{U}} D_u \right) / \sim \tag{4.4}$$

where the equivalence relation $\sim$ is such that for every $u \in \mathbb{U}$ and $i \in \mathbb{N}$ the root ρ_{ui} of D_{ui} is in relation with the distinguished point $x_{ui} \in D_u$. This set $\mathscr{G}^*(\mathcal{D})$ is endowed with a distance d.

This distance is such that for all $\mathbf{i} = i_1 i_2 \dots i_n$ and $\mathbf{j} = j_1 j_2 \dots j_m$ and points $y \in D_{\mathbf{i}}$, $z \in D_{\mathbf{j}}$,

$$\begin{aligned} \mathrm{d}(y,z) = \mathrm{d}(z,y) = d_{\mathbf{i}}(y,z) \qquad &\text{if} \qquad \mathbf{i} = \mathbf{j}, \\ = d_{\mathbf{i}}(y, x_{\mathbf{j}_{n+1}}) + \sum_{k=n+1}^{m-1} d_{\mathbf{j}_k}(\rho_{\mathbf{j}_k}, x_{\mathbf{j}_{k+1}}) + d_{\mathbf{j}}(\rho_{\mathbf{j}}, z) \qquad &\text{if } \mathbf{j} \prec \mathbf{i}, \end{aligned}$$

and if $\mathbf{i} \wedge \mathbf{j} = \mathbf{i}_l = \mathbf{j}_l$ is different from $\mathbf{i}$ and $\mathbf{j}$ we let

$$\begin{aligned}\mathrm{d}(y,z) = \mathrm{d}(z,y) = d_{\mathbf{i}_k}(x_{\mathbf{i}_{k+1}}, x_{\mathbf{j}_{k+1}}) &+ \sum_{k=l+1}^{n-1} d_{\mathbf{i}_k}(\rho_{\mathbf{i}_k}, x_{\mathbf{i}_{k+1}}) + d_{\mathbf{i}}(\rho_{\mathbf{i}}, z) \\ &+ \sum_{k=l+1}^{m-1} d_{\mathbf{j}_k}(\rho_{\mathbf{j}_k}, x_{\mathbf{j}_{k+1}}) + d_{\mathbf{j}}(\rho_{\mathbf{j}}, z).\end{aligned}$$

This last configuration is illustrated in Figure 4.3. We then set

$$\mathscr{G}(\mathcal{D}) = \overline{\mathscr{G}^*(\mathcal{D})},$$

its metric completion. We also let $\mathscr{L}(\mathcal{D}) = \mathscr{G}(\mathcal{D}) \setminus \mathscr{G}^*(\mathcal{D})$ be its *set of leaves*.

Whenever the associated function $\ell : \mathbb{U} \to \mathbb{R}_+$ defined as $u \mapsto \ell(u) = \mathrm{diam}(D_u)$ is non-explosive, it is easy to see that the defined object $\mathscr{G}(\mathcal{D})$ is compact, and it can be approximated by gluing only finitely many blocks of the decoration.

Remark that if $\mathcal{D}$ is supported on a plane tree τ, then for any $u \in \tau$ the result of the gluing operation does not depend on the distinguished points $(x_{ui})_{i \geq \deg_\tau^+(u)+1}$ of $\mathcal{D}(u)$ with index greater than $\deg_\tau^+(u) + 1$.

Identification of the leaves. Suppose that $\mathcal{D}$ is such that $\mathscr{G}(\mathcal{D})$ is compact. Then there exists a natural map

$$\iota_{\mathcal{D}} : \partial\mathbb{U} \to \mathscr{G}(\mathcal{D}), \tag{4.5}$$

that maps every leaf of the Ulam-Harris tree to a point of $\mathscr{G}(\mathcal{D})$. Indeed, for any $\mathbf{i} = i_1 i_2 \cdots \in \partial\mathbb{U}$, we define

$$\iota_{\mathcal{D}}(\mathbf{i}) = \lim_{n\to\infty} x_{\mathbf{i}_n} \in \mathscr{G}(\mathcal{D}),$$

and the limit exists because of the compactness of the space. It is then straightforward to see that this map is continuous.

Adding measures. Let $\mathcal{D}$ be a metric-space-valued decoration. Suppose that we have a family $\boldsymbol{\nu} : u \mapsto (\nu_u)_{u\in\mathbb{U}}$ such that for all $u \in \mathbb{U}$, ν_u is a Borel measure on $\mathcal{D}(u)$. Then we can define a corresponding measure ν on $\mathbb{U}$, so that $\nu(\{u\}) = \nu_u(D_u)$ for all $u \in \mathbb{U}$. We define the *support* of $\boldsymbol{\nu}$ as the support of the corresponding measure ν on $\mathbb{U}$.

In this setting, we can in a natural way define a measure on $\mathscr{G}(\mathcal{D})$ by seeing $\sum_{u\in\mathbb{U}} \nu_u$ as a measure on $\mathscr{G}(\mathcal{D})$, identifying every decoration as a subspace. In this case we write

$$\mathscr{G}(\mathcal{D}, \boldsymbol{\nu}) \tag{4.6}$$

for the corresponding measured metric space. In the case where $\mathscr{G}(\mathcal{D})$ is compact, then the function $\iota_{\mathcal{D}} : \partial\mathbb{U} \to \mathscr{G}(\mathcal{D})$ is well-defined and continuous so that if μ denotes a measure on $\partial\mathbb{U}$, then we can consider the push-forward measure $(\iota_{\mathcal{D}})_*\mu$ on $\mathscr{G}(\mathcal{D})$. In this case we write

$$\mathscr{G}(\mathcal{D}, \mu) = (\mathscr{G}(\mathcal{D}), (\iota_{\mathcal{D}})_*\mu). \tag{4.7}$$

We can now state the main result of Section 4.2.

Theorem 4.2. *Suppose that $(\mathcal{D}_n)_{n\geq 1}$ is a sequence of decorations such that there exists a decoration $\mathcal{D}_\infty$ such that for every $u \in \mathbb{U}$,*

$$\mathcal{D}_n(u) \underset{n\to\infty}{\longrightarrow} \mathcal{D}_\infty(u),$$

for the infinitely pointed Gromov-Hausdorff-Prokhorov topology and such that the associated real-valued decoration $(u \mapsto \sup_{n\geq 1} \operatorname{diam}(\mathcal{D}_n(u)))$ is non-explosive.

(i) *Then, we have the convergence*

$$\mathscr{G}(\mathcal{D}_n) \longrightarrow \mathscr{G}(\mathcal{D}_\infty) \quad \text{as } n \to \infty \text{ for the Gromov-Hausdorff topology.}$$

(ii) *Furthermore, suppose that for all $n \geq 1$, we have $\boldsymbol{\nu}_n = (\nu_{u,n})_{u\in\mathbb{U}}$, measures over $\mathcal{D}_n$ such that the corresponding measures $(\nu_n)_{n\geq 1}$ are probabilities on $\mathbb{U}$ and converge weakly in $\overline{\mathbb{U}}$ as $n \to \infty$ to some probability measure ν_∞ that charges only $\partial\mathbb{U}$, then we have the convergence*

$$\mathscr{G}(\mathcal{D}_n, \boldsymbol{\nu}_n) \longrightarrow \mathscr{G}(\mathcal{D}_\infty, \nu_\infty) \quad \text{as } n \to \infty,$$

for the Gromov-Hausdorff-Prokhorov topology.

The first point of this theorem states that the convergence of a global structure defined as $\mathscr{G}(\mathcal{D}_n)$, for some sequence $\mathcal{D}_n$ of decorations, can be obtained by proving the convergence of every $\mathcal{D}(u)$, for all $u \in \mathbb{U}$ (convergence of finite-dimensional marginals) with the additional assumption that they satisfy some relative compactness property which is here expressed as the non-explosion condition. The second point ensures that if we add measures on our decorations and if these measures converge nicely then we can improve our convergence to Gromov-Hausdorff-Prokhorov topology on measured metric spaces. We only treat the case where the measure gets "pushed to the leaves" because only this case arises in our applications. A more general statement where ν is not carried on $\partial\mathbb{U}$ could be proven under the appropriate assumptions.

4.2.3 Some formal topological arguments

The aim of this section is to justify and properly define the construction described in the preceding section, in a way that can be adapted to random decorations without any measurability problem. This section is rather technical and can be skipped at first reading. We begin by recalling some topological facts about the Urysohn universal space, and the so-called Hausdorff/Gromov-Hausdorff/Gromov-Hausdorff-Prokhorov topologies.

Urysohn space and Gromov-Hausdorff-Prokhorov topology

Urysohn universal space. Let us consider $(\mathcal{U}, \delta)$ the Urysohn space, and fix a point $* \in \mathcal{U}$. The space $\mathcal{U}$ is defined as the only Polish metric space (up to isometry) which has the following extension property (see [78] for constructions and basic properties of $\mathcal{U}$): given any finite metric space X, and any point $x \in X$, any isometry from $X \setminus \{x\}$ to $\mathcal{U}$ can be extended to an isometry from X to $\mathcal{U}$. This property ensures in particular that any separable metric space can be isometrically embedded into $\mathcal{U}$. In what follows we will use the fact that if (K, d, ρ) is a rooted compact metric space, there exists an isometric embedding of K into $\mathcal{U}$ such that ρ is mapped to $*$. It has also a very useful property called compact homogeneity (see [94, Corollary 1.2]) which ensures that any isometry φ between two compact subsets K and L of $\mathcal{U}$ can be extended to the whole space $\mathcal{U}$, meaning that there exists a global isometry ϕ such that φ is just the restriction $\varphi = \phi_{|K}$.

Hausdorff distance, Lévy-Prokhorov distance. For any two compact subsets A and B of the same metric space (E, d), we can define their Hausdorff distance as

$$\mathrm{d}_{\mathrm{H}}^{E}(A,B) = \inf\left\{\epsilon > 0 \;\middle|\; A \subset B^{(\epsilon)}, \quad B \subset A^{(\epsilon)}\right\},$$

where $A^{(\epsilon)}$ and $B^{(\epsilon)}$ are the ϵ-fattening of the corresponding sets. We denote the set of Borel probability measures on E by $\mathcal{P}(E)$. For any two $\mu, \nu \in \mathcal{P}(E)$, we can define their Lévy-Prokhorov distance as

$$\mathrm{d}_{\mathrm{LP}}^{E}(\mu,\nu) = \inf\left\{\epsilon > 0 \;\middle|\; \forall F \in \mathcal{B}(E), \mu(F) \leq \nu(F^{(\epsilon)}) + \epsilon \text{ and } \nu(F) \leq \mu(F^{(\epsilon)}) + \epsilon\right\}.$$

Whenever the space E is the Urysohn space, we drop the index E in the notation for those distances.

Infinitely pointed Gromov-Hausdorff topology. We write $\mathbb{M}^{k\bullet}$ for the space of all equivalence classes of $(k+1)$-pointed measure metric spaces. We can define the Gromov-Hausdorff distance on $\mathbb{M}^{k\bullet}$ by

$$\begin{aligned}
&\mathrm{d_{GH}}^{(k)}((X,d,\rho_0,(\rho_1,\dots,\rho_k)),(X',d',\rho_0,(\rho_1',\dots,\rho_k')))\\
&= \inf_{\phi: X\to\mathcal{U}, \phi': X'\to\mathcal{U}} \left\{d_{\mathrm{H}}(\phi(X),\phi'(X)) \vee \max_{0\leq i\leq k} \delta(\phi(\rho_i),\phi'(\rho_i'))\right\},
\end{aligned}$$

where, as previously, the infimum is over all isometric embeddings ϕ and ϕ' of X and X' into the Urysohn space $\mathcal{U}$. We write $\mathbb{M}^{\infty\bullet}$ for the space of all (equivalence classes of) ∞-pointed measured metric spaces. We can define the infinitely pointed Gromov-Hausdorff distance on $\mathbb{M}^{\infty\bullet}$ by

$$\begin{aligned}
&\mathrm{d_{GH}}^{(\infty)}((X,d,\rho_0,(\rho_i)_{i\geq 1}),(X',d',\rho_0,(\rho_i')_{i\geq 1}))\\
&= \sum_{k=1}^{\infty} \frac{1}{2^k} \mathrm{d_{GH}}^{(k)}((X,d,\rho_0,(\rho_1,\dots,\rho_k)),(X',d',\rho_0,(\rho_1',\dots,\rho_k')))\\
&(\leq (\operatorname{diam} X + \operatorname{diam} X') < \infty)
\end{aligned}$$

By abuse of notation, we will also consider (equivalence classes of) finitely pointed compact metric spaces $(X, \mathrm{d}, \rho, (x_i)_{1\leq i\leq k})$ as elements of $\mathbb{M}^{\infty\bullet}$, by completing the sequence (x_i) by setting $x_i = \rho$ for all $i \geq k+1$.

Infinitely pointed Gromov-Hausdorff-Prokhorov topology. In some of our applications, we work on $\mathbb{K}^{\infty\bullet}$, which is the corresponding space for elements of $\mathbb{M}^{\infty\bullet}$ endowed with a Borel probability measure. In the same way as before, elements of $\mathbb{K}^{\infty\bullet}$ are 5-tuples $(X, \mathrm{d}, \rho, (x_i)_{i\geq 1}, \mu)$, where $(X, \mathrm{d}, \rho, (x_i)_{i\geq 1}) \in \mathbb{K}^{\infty\bullet}$ and μ is a finite Borel measure on X. Again we set

$$\begin{aligned}
&\mathrm{d_{GHP}}^{(k)}((X,d,\rho_0,(\rho_1,\dots,\rho_k),\mu),(X',d',\rho_0,(\rho_1',\dots,\rho_k'),\mu'))\\
&= \inf_{\phi: X\to\mathcal{U}, \phi': X'\to\mathcal{U}} \left\{d_{\mathrm{H}}(\phi(X),\phi'(X)) \vee d_{\mathrm{LP}}((\phi)_*\mu,(\phi')_*\mu') \vee \max_{0\leq i\leq k} d(\phi(\rho_i),\phi'(\rho_i'))\right\},
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{d_{GHP}}^{(\infty)}((X,d,\rho_0,(\rho_i)_{i\geq 1},\mu),(X',d',\rho_0,(\rho_i')_{i\geq 1},\mu'))\\
&= \sum_{k=1}^{\infty} \frac{1}{2^k} \mathrm{d_{GHP}}^{(k)}((X,d,\rho_0,(\rho_1,\dots,\rho_k),\mu),(X',d',\rho_0,(\rho_1',\dots,\rho_k'),\mu')).
\end{aligned}$$

Construction in the appropriate ambient space

In order to ease the definition of our objects and avoid some measurability issues that arise when working with abstract equivalence classes of metric spaces, we define a way of only dealing with some particular representatives of those equivalence classes that are compact subsets of the set $\mathcal{U}$. For that matter we define $\mathbb{K}^{\infty\bullet}(\mathcal{U})$ the counterpart of $\mathbb{K}^{\infty\bullet}$, as

$$\begin{aligned}\mathbb{K}^{\infty\bullet}(\mathcal{U}) := \{(K, \delta_{|_K}, *, (\rho_i)_{i\geq 1}, \mu) \ \mid\ * \in K \subset \mathcal{U},\ K \text{ compact},\\ \forall i \geq 1, \rho_i \in K,\ \mu \in \mathcal{P}(\mathcal{U}),\ \mathrm{supp}(\mu) \subset K\},\end{aligned}$$

where $\delta_{|_K}$ is the distance on $\mathcal{U}$ restricted to the subset K. We set accordingly,

$$\mathrm{d_{HP}}^{(k)}((K, (\rho_1, \dots, \rho_k), \mu), (K', (\rho'_1, \dots, \rho'_k), \mu')) = d_{\mathrm{H}}(K, K') \vee d_{\mathrm{LP}}(\mu, \mu') \vee \max_{1\leq i\leq k} d(\rho_i, \rho'_i),$$

and

$$\begin{aligned}&\mathrm{d_{HP}}^{(\infty)}(K, \delta_{|_K}, *, (\rho_i)_{i\geq 1}, \mu), (K', \delta_{|_{K'}}, *, (\rho'_i)_{i\geq 1}, \mu'))\\ &= \sum_{k=1}^{\infty} \frac{1}{2^k} \mathrm{d_{HP}}^{(k)}((K, (\rho_1, \dots, \rho_k), \mu), (K', (\rho'_1, \dots, \rho'_k), \mu')).\end{aligned}$$

We define the projection map $\pi : \mathbb{K}^{\infty\bullet}(\mathcal{U}) \longrightarrow \mathbb{K}^{\infty\bullet}$, such that

$$\pi((K, \delta_{|_K}, *, (\rho_i)_{i\geq 1}, \mu)) = \left[(K, \delta_{|_K}, *, (\rho_i)_{i\geq 1}, \mu)\right],$$

the corresponding equivalence class in $\mathbb{K}^{\infty\bullet}$. This map is surjective by the properties of Urysohn space and continuous because it is obviously 1-Lipschitz. Using the surjectivity, we know that we can lift any deterministic element of $\mathbb{K}^{\infty\bullet}$ to an element $\mathbb{K}^{\infty\bullet}(\mathcal{U})$.

Actually, we are going to deal with random variables with values in the space $\mathbb{K}^{\infty\bullet}$ and we want to ensure that we can consider versions of those random variables with values in $\mathbb{K}^{\infty\bullet}(\mathcal{U})$. In fact, remarking that both sets are Polish spaces, we can use a theorem of measure theory from [90] which ensures that every probability distribution τ on $\mathbb{K}^{\infty\bullet}$ can be lifted to a probability measure σ on $\mathbb{K}^{\infty\bullet}(\mathcal{U})$, such that $\pi_*\sigma = \tau$. Hence, whenever we consider a random variable with values in $\mathbb{K}^{\infty\bullet}$, we can always work with a version of our random variable that is embedded in the space $\mathcal{U}$, and whose root coincides with $*$. The same line of reasoning can be made with $\mathbb{M}^{\infty\bullet}$.

From now on, we work with decorations $\mathcal{D} \in (\mathbb{K}^{\infty\bullet}(\mathcal{U}))^{\mathbb{U}}$ by taking a representative for every one of the decorations.

Construction embedded in a space. We introduce the following space, in which we will be able to define a representative of the space $\mathscr{G}(\mathcal{D})$ for any family of decoration $\mathcal{D}$.

$$\ell^1(\mathcal{U}, \mathbb{U}, *) := \left\{ (y_u)_{u\in\mathbb{U}} \in \mathcal{U}^{\mathbb{U}} \ \middle|\ \sum_{u\in\mathbb{U}} \delta(y_u, *) < +\infty \right\}.$$

We endow $\ell^1(\mathcal{U}, \mathbb{U}, *)$ with the distance $\mathrm{d}((y_u)_{u\in\mathbb{U}}, (z_u)_{u\in\mathbb{U}}) = \sum_{u\in\mathbb{U}} \delta(y_u, z_u)$, which makes it a Polish space.

Remark 4.3. *If for each $u \in \mathbb{U}$, we are given an isometry $\phi_u : \mathcal{U} \to \mathcal{U}$ such that $\phi_u(*) = *$, then we can introduce*

$$\begin{aligned}\phi := \prod_{u\in\mathbb{U}} \phi_u :\ &\ell^1(\mathcal{U}, \mathbb{U}, *) \to \ell^1(\mathcal{U}, \mathbb{U}, *),\\ &(y_u)_{u\in\mathbb{U}} \mapsto (\phi_u(y_u))_{u\in\mathbb{U}},\end{aligned}$$

and ϕ is an isometry of the space $\ell^1(\mathcal{U}, \mathbb{U}, *)$.

For each $u \in \mathbb{U}$, we consider a representative of the decoration $(D_u, d_u, \rho_u, (x_{ui})_{i\geq 1})$ that belongs to $\mathbb{M}^{\infty\bullet}(\mathcal{U})$, meaning that we see D_u as a subset of $\mathcal{U}$ and

$$(D_u, d_u, \rho_u, (x_{ui})_{i\geq 1}) = \left(D_u, \delta_{|D_u}, *, (x_{ui})_{i\geq 1}\right)$$

Then the gluing operation is defined this way. Let $\mathbf{i} = i_1 i_2 \dots i_n \in \mathbb{U}$. For any such $\mathbf{i} \in \mathbb{U}$, we define

$$\tilde{D_{\mathbf{i}}} = \left\{ (y_u)_{u\in\mathbb{U}} \;\middle|\; y_\emptyset = x_{\mathbf{i}_1}, y_{\mathbf{i}_1} = x_{\mathbf{i}_2}, \dots, y_{\mathbf{i}_{n-1}} = x_{\mathbf{i}_n}, y_{\mathbf{i}} \in D_{\mathbf{i}}, \text{ and } \forall u \npreceq \mathbf{i}, y_u = * \right\}.$$

Remark that each of the subsets $\tilde{D_{\mathbf{i}}}$ is isometric to the corresponding decoration $D_{\mathbf{i}}$. Then we consider

$$\mathscr{G}^*(\mathcal{D}) = \bigcup_{\mathbf{i}\in\mathbb{U}} \tilde{D_{\mathbf{i}}}. \tag{4.8}$$

The structure $\mathscr{G}(\mathcal{D})$ is then defined as the closure of $\mathscr{G}^*(\mathcal{D})$ in the space $\ell^1(\mathcal{U}, \mathbb{U}, *)$. Thanks to Remark 4.3, the resulting space (up to isometry) does not depend on the choice of representative for the different decorations.

For convenience, for any plane tree θ we also introduce the metric space obtained by only gluing the decorations that are indexed by the vertices in θ, which we denote by $\mathscr{G}(\theta, \mathcal{D})$, i.e.,

$$\mathscr{G}(\theta, \mathcal{D}) := \bigcup_{\mathbf{i}\in\theta} \tilde{D_{\mathbf{i}}} \tag{4.9}$$

We do not need to complete it since it is already compact, as a union of a finite number of compact metric spaces.

Identification of the leaves. Suppose that $\mathcal{D}$ is such that $\mathscr{G}(\mathcal{D})$ is compact. Then, in this setting, the map $\iota_{\mathcal{D}} : \partial\mathbb{U} \to \mathscr{G}(\mathcal{D})$ defined in (4.5) has the following form: for any $\mathbf{i} = i_1 i_2 \dots \in \partial\mathbb{U}$,

$$\begin{aligned} \iota_{\mathcal{D}}(\mathbf{i}) = (y_u)_{u\in\mathbb{U}} \quad \text{with } y_{\mathbf{i}_n} &= x_{\mathbf{i}_{n+1}} \quad \text{for all } n \geq 0, \\ y_u &= * \quad \text{whenever} \quad u \nprec \mathbf{i}. \end{aligned}$$

4.2.4 Proof of Theorem 4.2

Before proving the theorem, let us state a lemma that ensures that the gluing operation is continuous when considering a finite number of decorations.

Lemma 4.4. *For any θ finite plane tree, and $\mathcal{D}$ and $\mathcal{D}'$ decorations, we have*

$$\mathrm{d_{GH}}\left(\mathscr{G}\left(\theta, \mathcal{D}\right), \mathscr{G}\left(\theta, \mathcal{D}'\right)\right) \leq 2 \cdot \sum_{u\in\theta} \mathrm{d_{GH}}^{(\deg_\theta^+(u))}\left(\mathcal{D}(u), \mathcal{D}'(u)\right).$$

Proof. For all $u \in \theta$, and thanks to the compact homogeneity of $\mathcal{U}$, we can find an isometry $\phi_u : \mathcal{U} \to \mathcal{U}$ such that $\phi_u(*) = *$ and

$$\mathrm{d_H}(\phi_u\left(D'_u\right)), D_u) \vee \max_{1\leq i\leq \deg_\theta^+(u)} \delta(\phi_u(x'_{ui}), x_{ui}) \leq 2\mathrm{d_{GH}}^{(\deg_\theta^+(u))}\left(\mathcal{D}(u), \mathcal{D}'(u)\right). \tag{4.10}$$

Then let $\phi_u = \mathrm{id}_{\mathcal{U}}$, for every $u \notin \theta$, and let $\phi = \prod_{u\in\mathbb{U}} \phi_u$ be the corresponding isometry of $\ell^1(\mathcal{U}, \mathbb{U}, *)$. Then let us show that we control the Hausdorff distance between

$$\begin{aligned} \mathscr{G}(\theta, \mathcal{D}) &= \bigcup_{\mathbf{i}\in\theta} \left\{ (y_u)_{u\in\mathbb{U}} \;\middle|\; y_\emptyset = x_{\mathbf{i}_1},\ y_{\mathbf{i}_1} = x_{\mathbf{i}_2},\ \dots,\ y_{\mathbf{i}_{n-2}} = x_{\mathbf{i}_{n-1}},\ y_{\mathbf{i}_{n-1}} = x_{\mathbf{i}},\ \text{and } \forall u \npreceq \mathbf{i}, y_u = * \right\}, \\ &= \bigcup_{\mathbf{i}\in\theta} \tilde{D_{\mathbf{i}}}, \end{aligned}$$

and

$$\begin{aligned}\phi(\mathscr{G}(\theta,\mathcal{D}')) &= \bigcup_{\mathbf{i}\in\theta}\left\{(y_u)_{u\in\mathbb{U}} \;\middle|\; y_\emptyset=\phi_\emptyset(x'_{\mathbf{i}_1}),\dots,y_{\mathbf{i}_{n-1}}=\phi_{\mathbf{i}_{n-1}}(x'_{\mathbf{i}}), y_{\mathbf{i}}\in\phi_{\mathbf{i}}(D'_{\mathbf{i}}), \text{ and } \forall u\not\preceq\mathbf{i}, y_u=*\right\}\\ &= \bigcup_{\mathbf{i}\in\theta}\phi\left(\tilde{D}'_{\mathbf{i}}\right).\end{aligned}$$

Now for any $\mathbf{i}=i_1i_2\dots i_n\in\theta$, any $y=(y_u)_{u\in\mathbb{U}}\in\tilde{D}_{\mathbf{i}}$ and $z=(z_u)_{u\in\mathbb{U}}\in\phi\left(\tilde{D}'_{\mathbf{i}}\right)$, we can write

$$\mathrm{d}(y,z)=\delta(y_{\mathbf{i}},z_{\mathbf{i}})+\sum_{\ell=1}^{n}\delta(x_{\mathbf{i}_\ell},\phi_{\mathbf{i}_{\ell-1}}(x'_{\mathbf{i}_\ell})),$$

with $y_{\mathbf{i}}\in D_{\mathbf{i}}$ and $z_{\mathbf{i}}\in\phi(D'_{\mathbf{i}})$. Now using equation (4.10), we get that

$$\begin{aligned}\mathrm{d_H}\left(\tilde{D}_{\mathbf{i}},\phi\left(\tilde{D}'_{\mathbf{i}}\right)\right) &\leq \mathrm{d_H}(D_{\mathbf{i}},\phi_{\mathbf{i}}(D'_{\mathbf{i}}))+\sum_{\ell=1}^{n}\delta(x_{\mathbf{i}_\ell},\phi_{\mathbf{i}_{\ell-1}}(x'_{\mathbf{i}_\ell}))\\ &\leq 2\cdot\sum_{u\in\theta}\mathrm{d_{GH}}^{(\deg_\theta^+(u))}\left(\mathcal{D}(u),\mathcal{D}'(u)\right).\end{aligned}\tag{4.11}$$

The last inequality is true for any $\mathbf{i}\in\theta$, hence taking a union yields

$$\mathrm{d_H}\left(\bigcup_{\mathbf{i}\in\theta}\tilde{D}_{\mathbf{i}},\bigcup_{\mathbf{i}\in\theta}\phi\left(\tilde{D}'_{\mathbf{i}}\right)\right)=\mathrm{d_H}\left(\mathscr{G}(\theta,\mathcal{D}),\phi(\mathscr{G}(\theta,\mathcal{D}'))\right)\leq 2\cdot\sum_{u\in\theta}\mathrm{d_{GH}}^{(\deg_\theta^+(u))}\left(\mathcal{D}(u),\mathcal{D}'(u)\right),$$

which finishes the proof of the lemma. ☐

Proof of Theorem 4.2. Let $n\in\mathbb{N}$ be an integer and θ a finite plane tree and $y=(y_u)_{u\in\mathbb{U}}\in\mathscr{G}(\mathcal{D}_n)$. From our construction of $\mathscr{G}(\mathcal{D}_n)$, we know that the indices v for which $y_v\neq *$ are all contained in an infinite ray in $\mathbb{U}$, meaning that there exists $u\in\partial\mathbb{U}$ such that $y_v=*$ for all $v\not\preceq u$. Now we can check that

$$\begin{aligned}\mathrm{d}(y,\mathscr{G}(\theta,\mathcal{D}_n)) = \sum_{\substack{v\prec u\\ v\notin\theta}}\delta(y_v,*) &\leq \sum_{\substack{v\prec u\\ v\notin\theta}}\sup_{n\geq 1}\operatorname{diam}(\mathcal{D}_n(v))\\ &\leq \sup_{u\in\mathbb{U}}\sum_{\substack{v\prec u\\ v\notin\theta}}\sup_{n\geq 1}\operatorname{diam}(\mathcal{D}_n(v)).\end{aligned}$$

Since it holds for any $y\in\mathscr{G}(\mathcal{D}_n)$ and the bound on the right-hand side is uniform for all such y, we have

$$\mathrm{d_H}\left(\mathscr{G}(\theta,\mathcal{D}_n),\mathscr{G}(\mathcal{D}_n)\right)\leq\sup_{u\in\mathbb{U}}\sum_{\substack{v\prec u\\ v\notin\theta}}\sup_{n\geq 1}\operatorname{diam}(\mathcal{D}_n(v)).$$

Now we can write

$$\begin{aligned}\mathrm{d_{GH}}\left(\mathscr{G}(\mathcal{D}_n),\mathscr{G}(\mathcal{D})\right) &\leq \mathrm{d_{GH}}\left(\mathscr{G}(\mathcal{D}_n),\mathscr{G}(\theta,\mathcal{D}_n)\right)+\mathrm{d_{GH}}\left(\mathscr{G}(\theta,\mathcal{D}_n),\mathscr{G}(\theta,\mathcal{D})\right)+\mathrm{d_{GH}}\left(\mathscr{G}(\theta,\mathcal{D}),\mathscr{G}(\mathcal{D})\right)\\ &\leq 2\sup_{u\in\mathbb{U}}\sum_{\substack{v\prec u\\ v\notin\theta}}\sup_{n\geq 1}\operatorname{diam}(\mathcal{D}_n(v))+\mathrm{d_{GH}}\left(\mathscr{G}(\theta,\mathcal{D}_n),\mathscr{G}(\theta,\mathcal{D})\right)\end{aligned}$$

Using the non-explosion of the function $(u\mapsto\sup_{n\geq 1}\operatorname{diam}(\mathcal{D}_n(u)))$ we can make the first term as small as we want by taking the appropriate θ, and when θ is fixed, the second term vanishes as $n\to\infty$ thanks to Lemma 4.4. This finishes the proof of (i).

Now let us prove point (ii). For simplicity, we write $\mu_n = \sum_{u\in\mathbb{U}} \nu_{u,n}$ and also $\mu_\infty = (\iota_{\mathcal{D}_\infty})_*\nu_\infty$. Let $\epsilon > 0$. From the non-explosion condition we know that we can find a plane tree θ such that

$$\sup_{u\in\mathbb{U}} \left(\sum_{\substack{v\prec u\\ v\notin\theta}} \sup_{n\geq 1} \operatorname{diam}(\mathcal{D}_n(v)) \right) < \epsilon. \tag{4.12}$$

Now, we construct another finite plane tree θ', such that $\theta \subset \theta'$, by adding only children of vertices of θ. We do so in such a way that

$$\sum_{v\in\theta'\setminus\theta} \nu_\infty\left(T(v)\right) \geq 1-\epsilon/2.$$

Remark that from (4.12), for any $v \in \theta'\setminus\theta$ and any $n \geq 1$, we have $\operatorname{diam}\left(\mathcal{D}_n(v)\right) < \epsilon$.

Introduce the projection $p_{\theta'} : \ell^1(\mathcal{U}, \mathbb{U}, *) \to \ell^1(\mathcal{U}, \mathbb{U}, *)$, such that for any $(y_u)_{u\in\mathbb{U}}$, the image $(z_u)_{u\in\mathbb{U}} = p_{\theta'}\left((y_u)_{u\in\mathbb{U}}\right)$ is such that $z_u = y_u$ for any $u \in \theta'$ and $z_u = *$ otherwise. Using (4.12), we can check that for any $n \geq 1$ and $y \in \mathscr{G}(\mathcal{D}_n)$, we have

$$\mathrm{d}\left(p_{\theta'}(y), y\right) < \epsilon.$$

This observation suffices to show that for any $n \in \mathbb{N}\cup\{\infty\}$,

$$\mathrm{d}_{\mathrm{GHP}}\left(\left(\mathscr{G}(\mathcal{D}_n), \mu_n\right), \left(p_{\theta'}\left(\mathscr{G}(\mathcal{D}_n)\right), (p_{\theta'})_*\mu_n\right)\right) < \epsilon.$$

Then,

$$\begin{aligned}
\mathrm{d}_{\mathrm{GHP}}\left(\left(\mathscr{G}(\mathcal{D}_n), \mu_n\right), \left(\mathscr{G}(\mathcal{D}_\infty), \mu_\infty\right)\right) \leq\; & \mathrm{d}_{\mathrm{GHP}}\left(\left(\mathscr{G}(\mathcal{D}_n), \mu_n\right), \left(p_{\theta'}\left(\mathscr{G}(\mathcal{D}_n)\right), (p_{\theta'})_*\mu_n\right)\right)\\
&+ \mathrm{d}_{\mathrm{GHP}}\left(\left(\mathscr{G}(\mathcal{D}_\infty), \mu_\infty\right), \left(p_{\theta'}\left(\mathscr{G}(\mathcal{D}_\infty)\right), (p_{\theta'})_*\mu_\infty\right)\right)\\
&+ \mathrm{d}_{\mathrm{GHP}}\left(\left(p_{\theta'}\left(\mathscr{G}(\mathcal{D}_n)\right), (p_{\theta'})_*\mu_n\right), \left(p_{\theta'}\left(\mathscr{G}(\mathcal{D}_\infty)\right), (p_{\theta'})_*\mu_\infty\right)\right).
\end{aligned}$$

The first two terms on the right-hand side are smaller than ϵ from what precedes so that we only have to prove that the last one is also small whenever n is large enough. Remark that for any $\mathcal{D}$, $p_{\theta'}\left(\mathscr{G}(\mathcal{D})\right) = \mathscr{G}(\theta', \mathcal{D})$. Let us fix $n \geq 1$ large enough such that

$$2\cdot \sum_{u\in\theta'} {\mathrm{d}_{\mathrm{GH}}}^{(\deg_{\theta'}(u))}\left(\mathcal{D}_n(u), \mathcal{D}_\infty(u)\right) < \epsilon,$$

and

$$\sum_{v\in\theta'\setminus\theta} \left|\nu_n(T(v)) - \nu_\infty(T(v))\right| < \epsilon. \tag{4.13}$$

From (4.11) in the proof of Lemma 4.4, we can find an isometry ϕ such that for all $\mathbf{i} \in \theta'$,

$$\mathrm{d}_{\mathrm{H}}\left(\tilde{D}_{\infty,\mathbf{i}}, \phi\left(\tilde{D}_{n,\mathbf{i}}\right)\right) < \epsilon. \tag{4.14}$$

Now because of (4.13), we know that we can find a coupling (X_n, X_∞) of random variables with values in $\overline{\mathbb{U}}$ having distribution ν_n and ν_∞, such that with probability $> 1-\epsilon$, they both fall in the same $T(v)$ for $v \in \theta'\setminus\theta$. From this coupling, we can construct another one between (Y_n, Y_∞) of random variables on respectively $\mathscr{G}(\theta', \mathcal{D}_n)$ and $\mathscr{G}(\theta', \mathcal{D}_\infty)$ such that one has distribution $(p_{\theta'})_*\mu_n$ and the other $(p_{\theta'})_*\mu_\infty$ and such that the probability that there exists $v \in \theta'\setminus\theta$ such that $Y_n \in \tilde{D}_{n,v}$ and $Y_\infty \in \tilde{D}_{\infty,v}$ is greater than $1-\epsilon$. Using this plus (4.14) shows that the couple $(Y_n, \phi(Y_\infty))$ is at distance at most ϵ with probability at least $1-\epsilon$. This shows that the Lévy-Prokhorov distance between $(p_{\theta'})_*\mu_\infty$ and $\phi_*((p_{\theta'})_*\mu_n)$ is smaller than ϵ. In this end, we just showed that

$$\mathrm{d}_{\mathrm{GHP}}\left(\left(p_{\theta'}\left(\mathscr{G}(\mathcal{D}_n)\right), (p_{\theta'})_*\mu_n\right), \left(p_{\theta'}\left(\mathscr{G}(\mathcal{D}_\infty)\right), (p_{\theta'})_*\mu_\infty\right)\right) < \epsilon,$$

which finishes the proof of the theorem. $\square$

4.2.5 Sufficient condition for non-explosion

Let us finish this section by proving a useful result which ensures non-explosion for some type of real-valued decorations. Let $(x_n)_{n\geq 1}$ be a sequence of non-negative real numbers. We define a real-valued decoration on the Ulam tree $\ell : \mathbb{U} \to \mathbb{R}_+$ using this sequence and a sequence $(u_n)_{n\geq 1}$ of distinct elements of $\mathbb{U}$ as

$$\ell(u_k) = x_k \quad \text{for all } k \geq 1,$$
$$\ell(u) = 0 \quad \text{for any } u \notin \{u_k \mid k \geq 1\}$$

The following lemma ensures the non-explosion of ℓ under some assumptions that are often met in our cases of application.

Lemma 4.5. *If there exists constant $\epsilon > 0$ and $K > 0$ such that for all $n \geq 1$*

$$x_n \leq (n+1)^{-\epsilon+o(1)} \qquad \text{and} \qquad \mathrm{ht}(u_n) \leq K \cdot \log n,$$

then the function ℓ defined above is non-explosive.

Proof. Let $i \in \mathbb{N}$. For any $u \in \mathbb{U}$ we have

$$\begin{aligned} \sum_{\substack{v \prec u \\ v \in \{u_k,\ 2^i < k \leq 2^{i+1}\}}} \ell(v) &\leq \# \left\{ k \in [\![2^i+1\,, 2^{i+1}]\!] \mid u_k \prec u \right\} \cdot \left(\max_{2^i < k \leq 2^{i+1}} \ell(u_k) \right) \\ &\leq K \cdot \log 2 \cdot (i+1) \cdot (2^i)^{-\epsilon + o(1)}, \end{aligned}$$

where the last display is independent of u. Now, if we consider any plane tree τ_i which contains all the vertices $\{u_1, u_2, \ldots, u_{2^i}\}$ then we have, for any $u \in \mathbb{U}$,

$$\sum_{\substack{v \prec u \\ v \notin \tau_i}} \ell(v) \leq \sum_{j=i}^{\infty} \sum_{\substack{v \prec u \\ v \in \{u_k,\ 2^i < k \leq 2^{i+1}\}}} \ell(v) \leq \sum_{j=i}^{\infty} K \cdot \log 2 \cdot (j+1) \cdot (2^j)^{-\epsilon + o(1)}$$

and the last display converges to 0 as $i \to \infty$, which proves the lemma. □

4.3 Preferential attachment and weighted recursive trees

In this section we recall some results about preferential attachment trees with initial fitnesses and weighted recursive trees that are proved in the previous chapter.

4.3.1 Definitions

Weighted recursive trees (WRT). For any sequence of non-negative real numbers $(w_n)_{n\geq 1}$ with $w_1 > 0$, the distribution $\mathrm{WRT}((w_n)_{n\geq 1})$ of the *weighted recursive tree with weights* $(w_n)_{n\geq 1}$ is defined on sequences of growing plane trees. A sequence $(\mathtt{T}_n)_{n\geq 1}$ having this distribution is constructed iteratively starting from $\mathtt{T}_1$ containing only one vertex $u_1 = \emptyset \in \mathbb{U}$, which has label 1, in the following manner: the tree $\mathtt{T}_{n+1}$ is obtained from $\mathtt{T}_n$ by adding a vertex u_{n+1} with label $n+1$. The parent of this new vertex is chosen to be any of the vertices $u_k \in \mathtt{T}_n$ with probability proportional to w_k, and u_{n+1} is added to the tree so that it is the rightmost child of its parent. Whenever we consider a random sequence of weight $(\mathsf{w}_n)_{n\geq 1}$, the distribution $\mathrm{WRT}((\mathsf{w}_n)_{n\geq 1})$ denotes the law of the random tree obtained by the above process conditionally on $(\mathsf{w}_n)_{n\geq 1}$.

Preferential attachment trees (PA). For any sequence $\mathbf{a} = (a_n)_{n\geq 1}$ of real numbers, with $a_1 > -1$ and $a_n \geq 0$ for $n \geq 2$, we define another distribution on growing sequences $(\mathtt{P}_n)_{n\geq 1}$ of plane trees called the *affine preferential attachment tree with initial fitnesses* $(a_n)_{n\geq 1}$ which is denoted by $\mathrm{PA}((a_n)_{n\geq 1})$. The construction goes on as before: $\mathtt{P}_1$ contains only one vertex u_1 labelled 1 and $\mathtt{P}_{n+1}$ is obtained from $\mathtt{P}_n$ by adding a vertex u_{n+1} with label $n+1$, whose parent is chosen to be any $u_k \in \mathtt{T}_n$ with probability proportional to $\deg^+_{\mathtt{P}_n}(u_k) + a_k$, where $\deg^+_{\mathtt{P}_n}(\cdot)$ denotes the number of children in the tree $\mathtt{P}_n$. By convention if $n = 1$, the second vertex u_2 is always defined as a child of u_1.

4.3.2 Properties of preferential attachment and weighed recursive trees

Let us state the properties proved in the previous chapter which will be needed in our analysis. Let us suppose here that we consider a sequence $\mathbf{a} = (a_n)$ such that

$$A_n := \sum_{i=1}^{n} a_i = c \cdot n + O\big(n^{1-\epsilon}\big) \quad \text{and} \quad a_n \leq n^{c'+o(1)} \tag{$H_{c,c'}$}$$

for some constants $c > 0$, some $0 \leq c' < \frac{1}{c+1}$ and some $\epsilon > 0$.

Convergence of degrees and representation theorem. A first result concerns the scaling limit of the degrees of the vertices in their order of creation and the distribution of the sequence of trees conditionally on the limit sequence; it can be read from [110, Theorem 1, Proposition 2 and Proposition 5]. We have the following convergence in the product topology to a random sequence

$$n^{-\frac{1}{c+1}} \cdot \left(\deg^+_{\mathtt{P}_n}(u_1), \deg^+_{\mathtt{P}_n}(u_2), \dots\right) \underset{n\to\infty}{\overset{\text{a.s.}}{\to}} (\mathsf{m}^{\mathbf{a}}_1, \mathsf{m}^{\mathbf{a}}_2, \dots), \tag{4.15}$$

and conditionally on the sequence $(\mathsf{m}^{\mathbf{a}}_k)_{k\geq 1}$ the sequence $(\mathtt{P}_n)$ has distribution $\mathrm{WRT}((\mathsf{m}^{\mathbf{a}}_k)_{k\geq 1})$. Also, the limiting sequence $(\mathsf{m}^{\mathbf{a}}_k)_{k\geq 1}$ has the following behaviour, which depends on the parameters c and c'

$$\mathsf{M}^{\mathbf{a}}_k := \sum_{i=1}^{k} \mathsf{m}^{\mathbf{a}}_i \underset{k\to\infty}{\sim} (c+1) \cdot k^{\frac{c}{c+1}} \qquad \text{and} \qquad \mathsf{m}^{\mathbf{a}}_k \leq (k+1)^{c' - \frac{1}{c+1} + o_\omega(1)}.$$

for a random function $o_\omega(1)$ which only depends on k and tends to 0 as $k \to \infty$. The convergence (4.15) is such that for all n large enough

$$\forall k \geq 1, \quad \deg^+_{\mathtt{P}_n}(u_k) \leq n^{\frac{1}{c+1}} \cdot (k+1)^{c' - \frac{1}{c+1} + o_\omega(1)}, \tag{4.16}$$

also for a random function $o_\omega(1)$ of k.

Distribution of $(\mathsf{M}^{\mathbf{a}}_k)_{k\geq 1}$ In some very specific cases for the sequence $\mathbf{a}$, the process $(\mathsf{M}^{\mathbf{a}}_k)_{k\geq 1}$ has an explicit distribution. In particular if $\mathbf{a} = a, b, b, b \dots$ then the sequence $(\mathsf{M}^{\mathbf{a}}_k)_{k\geq 1}$ has the Mittag-Leffer Markov chain distribution $\mathrm{MLMC}(\frac{1}{b+1}, \frac{a}{b+1})$, which we define below.

Let $0 < \alpha < 1$ and $\theta > -\alpha$. The generalized Mittag-Leffler $\mathrm{ML}(\alpha, \theta)$ distribution has pth moment

$$\frac{\Gamma(\theta)\Gamma(\theta/\alpha + p)}{\Gamma(\theta/\alpha)\Gamma(\theta + p\alpha)} = \frac{\Gamma(\theta+1)\Gamma(\theta/\alpha + p + 1)}{\Gamma(\theta/\alpha + 1)\Gamma(\theta + p\alpha + 1)} \tag{4.17}$$

and the collection of p-th moments for $p \in \mathbb{N}$ uniquely characterizes this distribution. Then, a Markov chain $(\mathsf{M}_n)_{n\geq 1}$ has the distribution $\mathrm{MLMC}(\alpha, \theta)$ if for all $n \geq 1$,

$$\mathsf{M}_n \sim \mathrm{ML}\left(\alpha, \theta + n - 1\right),$$

and its transition probabilities are characterised by the following equality in law:

$$\left(\mathsf{M}_n, \mathsf{M}_{n+1}\right) = \left(B_n \cdot \mathsf{M}_{n+1}, \mathsf{M}_{n+1}\right),$$

where $B_n \sim \mathrm{Beta}\left(\frac{\theta+k-1}{\alpha} + 1, \frac{1}{\alpha} - 1\right)$ is independent of $\mathsf{M}_{n+1}^{\alpha,\theta}$.

Height. In this setting, we know that the height of the tree P_n grows logarithmically in n using Theorem 3.3 of Chapter 3. We only need here the following weak version: there exists some constant K such that

$$\mathrm{ht}(\mathsf{P}_n) \leq K \cdot \log n, \tag{4.18}$$

almost surely for all n large enough. This estimate is also true for any sequence $(\mathtt{T}_n)_{n\geq 1}$ of weighted random trees with weights $(w_n)_{n\geq 1}$ as soon as $W_n := \sum_{i=1}^n w_i$ has at most a polynomial growth, which we always assume.

Measures. For a sequence of trees $(\mathtt{T}_n)_{n\geq 1}$ evolving under the distribution $\mathrm{WRT}((w_n)_{n\geq 1})$ for any weight sequence (w_n), [110, Theorem 4] ensures that the probability measures $(\mu_n)_{n\geq 1}$, defined in such a way that for all $k \in \{1, \dots n\}$ we have $\mu_n(u_k) = \frac{w_k}{W_n}$, converge almost surely weakly on $\overline{\mathbb{U}}$ towards a limiting measure μ.

Under the conditions $\sum_{n=1}^{\infty} w_n = \infty$ and $\sum_{n=1}^{\infty} \left(\frac{w_n}{W_n}\right)^2 < \infty$, which are almost surely satisfied by our sequence $(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1}$, the limiting measure μ is carried on $\partial\mathbb{U}$, and other sequences of probability measures $(\eta_n)_{n\geq 1}$ and $(\nu_n)_{n\geq 1}$, which we define below, also converge almost surely weakly towards μ.

For any $n \geq 1$, the measure ν_n is just defined as the uniform measure on the set $\{u_1, \dots, u_n\}$. The second sequence of measures $(\eta_n)_{n\geq 1}$ depends on a sequence $(b_n)_{n\geq 1}$ of real numbers which satisfies $b_1 > -1$ and $b_n \geq 0$ for all $n \geq 2$. We suppose that $b_n = O\left(n^{1-\epsilon}\right)$ for some $\epsilon > 0$ and that $B_n := \sum_{i=1}^n b_i = O(n)$. The measures are then defined in such a way that η_1 only charges the vertex u_1, and for every $n \geq 2$, the measure η_n charges only the vertices $\{u_1, u_2, \dots, u_n\}$ in such a way that for any $1 \leq k \leq n$,

$$\eta_n(u_k) = \frac{b_k + \deg^+_{\mathtt{T}_n}(u_k)}{B_n + n - 1}, \tag{4.19}$$

Other description of the measure μ. Suppose now that $\mathbf{a}$ is constant from the second term, say $a_1 = a > -1$ and $a_n = b > 0$ for all $n \geq 2$, so that it satisfies ($H_{c,c'}$) with $c = b$ and $c' = 0$. For all $u \in \mathbb{U}$, and all $i \geq 1$, using the limiting measure μ define the quantities

$$\mathsf{p}_{ui} = \frac{\mu(T(ui))}{\mu(T(u))}, \tag{4.20}$$

which describe how the mass above every vertex u is split into the subtrees above its children. In this case we can explicitly describe the law of the $(\mathsf{p}_u)_{u\in\mathbb{U}}$ and hence also the law of μ.

Moreover, let $\ell : \mathbb{U} \to \mathbb{R}_+$ be defined as

$$\ell(u_n) = \mathsf{m}_n^{\mathbf{a}} \qquad \forall n \geq 1.$$

Remark that almost surely this defines ℓ on all vertices of $\mathbb{U}$ and that, for $u \in \mathbb{U}$,

$$\ell(u) := \lim_{n \to \infty} n^{-1/(b+1)} \deg^+_{\mathsf{P}_n}(u),$$

In fact, thanks to [81], the values $(\ell(u))_{u \in \mathbb{U}}$ can be expressed from the ones of $(\mathsf{p}_u)_{u \in \mathbb{U}}$. The following proposition describes the joint distribution of those random variables, see e.g. [104] for a definition of the two parameter GEM distributions that appear in the statement.

Proposition 4.6. *In this setting we have*

$$(\mathsf{p}_i)_{i \geq 1} \sim \mathrm{GEM}\left(\frac{1}{b+1}, \frac{a}{b+1}\right) \quad \textit{and} \quad \forall u \in \mathbb{U} \setminus \{\emptyset\},\ (\mathsf{p}_{ui})_{i \geq 1} \sim \mathrm{GEM}\left(\frac{1}{b+1}, \frac{b}{b+1}\right),$$

and they are all independent. For all $u \in \mathbb{U}$, denote

$$S_u := \Gamma\left(\frac{b}{b+1}\right) \cdot \lim_{i \to \infty} i \cdot \mathsf{p}_{ui}^{\frac{1}{b+1}},$$

the $\frac{1}{b+1}$-diversity of the sequence $(\mathsf{p}_{ui})_{i \geq 1}$. Then for all $u \in \mathbb{U}$,

$$\ell(u) = \left(\prod_{v \preceq u} \mathsf{p}_v\right)^{\frac{1}{b+1}} \cdot S_u.$$

Proof. This result almost follows from [81, Theorem 1.5] and the adaptation to our case is left to the reader. □

4.4 Distributions on decorations

In this section, we define two families of distributions on decorations on the Ulam tree that will arise as limits of our discrete models.

4.4.1 The iterative gluing construction

Let $(\mathsf{B}_n, \mathsf{D}_n, \rho_n, (X_{n,i})_{i \geq 1})_{n \geq 1}$ be a sequence of independent random variables in $\mathbb{M}^{\infty \bullet}$, meaning compact pointed metric spaces endowed with a sequence of points. Let also $(w_n)_{n \geq 1}$ and $(\lambda_n)_{n \geq 1}$ be two sequences of non-negative real numbers, which we call the weights and scaling factors respectively. The model is the following: first sample $(\mathtt{T}_n)_{n \geq 1}$ with distribution $\mathrm{WRT}((w_n)_{n \geq 1})$. Then we define, for all $n \geq 1$, denoting u_n the vertex with label n in the trees $(\mathtt{T}_n)_{n \geq 1}$,

$$\mathcal{D}(u_n) = (D_{u_n}, d_{u_n}, \rho_{u_n}, (x_{u_n i})_{i \in \mathbb{N}}) := (\mathsf{B}_n, \lambda_n \cdot \mathsf{D}_n, \rho_n, (X_{n,i})_{i \geq 1}), \tag{4.21}$$

and for all $u \notin \{u_n \mid n \geq 1\}$, we set

$$\mathcal{D}(u) = (\{\star\}, 0, \star, (\star)_{i \geq 1}),$$

the one-point space. Let us assume that the cumulated sum W_n does not grow faster to infinity than polynomially, so that the height of the tree grows at most logarithmically. We also assume that there exists $\alpha > 0$ and $p > 1$ with $\alpha p > 1$ such that

$$\lambda_n \leq n^{-\alpha + o(1)} \quad \text{and} \quad \sup_{n \geq 1} \mathbb{E}\left[\mathrm{diam}(\mathsf{B}_n)^p\right] < \infty,$$

then we have $\operatorname{diam}(\mathcal{D}(u_n)) \leq n^{-\epsilon+o_\omega(1)}$ almost surely, with $\epsilon = \alpha - \frac{1}{p} > 0$. Using Lemma 4.5, the function $(u \mapsto \operatorname{diam}(\mathcal{D}_n(u)))$ is then almost surely non-explosive so $\mathscr{G}(\mathcal{D})$ is almost surely compact.

Assuming that the sequence (w_n) has an infinite sum, the limit μ of the weight measure associated to the trees $(\mathtt{T}_n)_{n\geq 1}$ is carried on $\partial\mathbb{U}$ and the random metric space $\mathscr{G}(\mathcal{D})$ can a.s. be endowed with a probability measure $(\iota_{\mathcal{D}})_*\mu$ and this yields a random measured metric space.

We call this procedure the *iterative gluing construction* with blocks $(\mathsf{B}_n, \mathsf{D}_n, \rho_n, (X_{n,i})_{i\geq 1})_{n\geq 1}$, scaling factors $(\lambda_n)_{n\geq 1}$ and weights $(w_n)_{n\geq 1}$. We allow the sequences $(\lambda_n)_{n\geq 1}$ and $(w_n)_{n\geq 1}$ to be random and in this case we assume that they are independent of the blocks $(\mathsf{B}_n, \mathsf{D}_n, \rho_n, (X_{n,i})_{i\geq 1})_{n\geq 1}$ and that we perform this procedure conditionally on those sequences.

The case of exchangeable distinguished points. A special case of the above construction is given when $(\mathsf{B}_n, \mathsf{D}_n, \rho_n, (X_{n,i})_{i\geq 1}, \nu_n)_{n\geq 1}$ is a sequence in $\mathbb{K}^{\infty\bullet}$ and that for all $n \geq 1$, conditionally on ν_n, the points $(X_{n,i})_{i\geq 1}$ are i.i.d. with law ν_n, independent of everything else. In this case we still call this distribution the iterative gluing construction with blocks $(\mathsf{B}_n, \mathsf{D}_n, \rho_n, \nu_n)_{n\geq 1}$, scaling factors $(\lambda_n)_{n\geq 1}$ and weights $(w_n)_{n\geq 1}$. This is the setting studied in [112].

4.4.2 The self-similar case

In this section, we investigate particular cases of models such that the limiting space is self-similar in distribution. This setting is adapted from the one studied by Rembart and Winkel in [106], which deals with several models of self-similar random trees. Let us define a model inspired from theirs. We fix $\beta > 0$ and the law of a couple $((\mathsf{B}, \mathsf{D}, \rho, (X_i)_{i\geq 1}), (P_i)_{i\geq 1})$, where the first coordinate is a random variable in $\mathbb{M}^{\infty\bullet}$ and $(P_i)_{i\geq 1}$ is a random variable in, say, $[0\,,1]^{\mathbb{N}}$. In order to use mimic the notation of [106], we let $\Xi = \mathbb{M}^{\bullet\infty} \times [0\,,1]^{\mathbb{N}}$. We consider a family

$$(\xi_u)_{u\in\mathbb{U}} = ((\mathsf{B}_u, \mathsf{D}_u, \rho_u, (X_{ui})_{i\geq 1}), (P_{ui})_{i\geq 1})_{u\in\mathbb{U}}$$

of random variables in Ξ which are i.i.d, with the same law as $\xi = ((\mathsf{B}, \mathsf{D}, \rho, (X_i)_{i\geq 1}), (P_i)_{i\geq 1})$. We set $P_\emptyset = 1$ and

$$\lambda_u = \left(\prod_{v \preceq u} P_v\right)^\beta.$$

We then define our random decorations as, for all $u \in \mathbb{U}$,

$$\mathcal{D}(u) := (\mathsf{B}_u, \lambda_u \cdot \mathsf{D}_u, \rho_u, (X_{ui})_{i\geq 1}). \tag{4.22}$$

We say that $\mathcal{D}$ is a self-similar decoration with exponent β and spine distribution given by ξ.

We want to show that, under suitable assumptions, the resulting $\mathscr{G}(\mathcal{D})$ is almost surely compact. In our examples, the distribution of $(P_i)_{i\geq 1}$ will always be $\mathrm{GEM}(\alpha, \theta)$ for some parameters $\alpha \in (0\,,1)$ and $\theta > -\alpha$, but the arguments presented here are still valid in greater generality.

The function ϕ_β. We first define, for any $n \geq 1$ the function

$$\phi_\beta^{(n)} : \Xi \times (\mathbb{M}^\bullet)^{\mathbb{N}} \to \mathbb{M}^\bullet,$$

as follows: $\phi_\beta^{(n)}((b, d, \rho, (x_i)_{i\geq 1}), (p_i)_{i\geq 1}, (b_i, d_i, \rho_i)_{i\geq 1})$ is the metric space obtained after gluing the n first b_i with distance scaled by p_i^β by identifying their root ρ_i with the point $x_i \in b$. For any

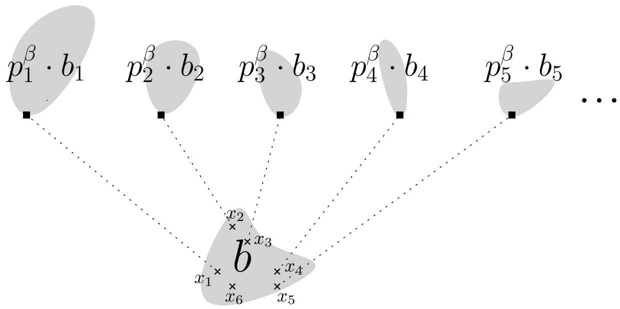


Figure 4.4 – The function ϕ_β

$n \geq 1$ this operation is continuous with respect to the product topology on the starting space, hence it is measurable.

Now we define ϕ_β as $\lim_{n\to\infty} \phi_\beta^{(n)}$ on the set where this limit exists, and constant equal to $(\{\star\}, 0, \star)$ on the complementary set. Since $\mathbb{M}^\bullet$ is Polish, this function is measurable. Remark that the condition for the limit to exists is

$$p_i \operatorname{diam}(b_i) \underset{i\to\infty}{\longrightarrow} 0.$$

The contraction Φ_β. Consider the set of probability measures $\mathcal{P}(\mathbb{M}^\bullet)$ on the space $\mathbb{M}^\bullet$ and for $p \geq 1$, the subset $\mathcal{P}_p \subset \mathcal{P}(\mathbb{M}^\bullet)$ given by

$$\mathcal{P}_p := \{\eta \in \mathcal{P}(\mathbb{M}^\bullet) \mid \mathbb{E}\left[\operatorname{diam}(\tau)^p\right] < \infty \text{ for } \tau \sim \eta\}. \tag{4.23}$$

We equip $\mathcal{P}_p$ with the *Wasserstein metric* of order $p \geq 1$, which is defined by

$$W_p(\eta, \eta') := \left(\inf \mathbb{E}\left[\left|d_{\mathrm{GH}}(\tau, \tau')\right|^p\right]\right)^{1/p}, \qquad \eta, \eta' \in \mathcal{P}_p, \tag{4.24}$$

where the infimum is taken over all joint distributions of (τ, τ') on $(\mathbb{M}^\bullet)^2$ with marginal distributions $\tau \sim \eta$ and $\tau' \sim \eta'$. The space $(\mathcal{P}_p, W_p)$ is complete since d_{GH} is a complete metric on $\mathbb{M}^\bullet$. Convergence in $(\mathcal{P}_p, W_p)$ implies weak convergence on $\mathbb{M}^\bullet$ and convergence of pth diameter moments.

Let $\Phi_\beta : \mathcal{P}_p \to \mathcal{P}_p$ be such that for any distribution η, the measure $\Phi_\beta(\eta)$ is the distribution of the space

$$\phi_\beta(\xi, (\tau_i)_{i\geq 1}),$$

where the $(\tau_i)_{i\geq 1}$ are i.i.d. random variables with law η, independent of ξ. Now let us state a result that was stated in the context of trees but remains valid in our case.

Lemma 4.7 (Lemma 3.4 of [106]). *Let $\beta > 0$, $p \geq 1$ and $((\mathsf{B}, \mathsf{D}, \rho, (X_i)_{i\geq 1}), (P_i)_{i\geq 1})$ with some distribution ν such that $\mathbb{E}\left[\operatorname{diam}(\mathsf{B})^p\right] < \infty$ and $\mathbb{E}\left[\sum_{j\geq 1} P_j^{p\beta}\right] < 1$. Then the map $\Phi_\beta \colon \mathcal{P}_p \to \mathcal{P}_p$ associated with ϕ_β is a strict contraction with respect to the Wasserstein metric of order p, i.e.*

$$\sup_{\eta,\eta' \in \mathcal{P}_p, \eta\neq\eta'} \frac{W_p(\Phi_\beta(\eta), \Phi_\beta(\eta'))}{W_p(\eta, \eta')} < 1. \tag{4.25}$$

Now using Banach fixed-point theorem, we know that there exists in $\mathcal{P}_p$ a unique fixed-point of the function Φ_β.

Compactness. Finally, the almost-sure compactness of our structure $\mathscr{G}(\mathcal{D})$ is ensured by [106, Prop. 3.7], and actually the distribution of $\mathscr{G}(\mathcal{D})$ is exactly the fixed-point of Φ_β, and this fixed-point is attractive.

Measure on the leaves. If we restrict ourselves to the case where the sequence $(P_i)_{i\geq 1}$ is such that $\sum_{i=1}^\infty P_i = 1$ almost surely, we can naturally define a measure μ on $\partial\mathbb{U}$ as follows:

$$\forall u \in \mathbb{U}, \quad \mu(T(u)) = \prod_{v \preceq u} P_v.$$

Then $\mathscr{G}(\mathcal{D})$ can be naturally endowed with the measure $\tilde{\mu} = (\iota_{\mathcal{D}})_* \mu$, and under the condition $\mathbb{P}(\exists i \geq 1,\ \mathsf{D}(\rho, X_i) > 0 \text{ and } P_i > 0) > 0$, one can check that this measure is carried on the set of leaves $\mathscr{L}(\mathcal{D})$.

Hausdorff dimension of the leaves. Under some mild hypotheses on the distribution of our blocks, we can compute the Hausdorff dimension of $\mathscr{L}(\mathcal{D})$ almost surely:

Proposition 4.8. *Let $\beta > 0$, $p \geq 1$ and $((\mathsf{B}, \mathsf{D}, \rho, (X_i)_{i\geq 1}), (P_i)_{i\geq 1})$ with some distribution ν such that $\mathbb{E}[\mathrm{diam}(\mathsf{B})^p] < \infty$ and $\mathbb{E}\left[\sum_{j\geq 1} P_j^{p\beta}\right] < 1$. Suppose furthermore that almost surely $\sum_{j\geq 1} P_j = 1$ and that $\mathbb{P}(\exists i \geq 1,\ \mathsf{D}(\rho, X_i) > 0$ and $P_i > 0) > 0$. Then the Hausdorff dimension of $\mathscr{L}(\mathcal{D})$ is almost surely*

$$\dim_{\mathrm{H}}(\mathscr{L}(\mathcal{D})) = \frac{1}{\beta}.$$

Proof. We prove this by providing an upper-bound and a lower-bound for the dimension. The upper-bound follows from the proof of [106, Lemma 4.6] which adapts to our new setting. For the lower-bound, we provide a direct argument, which crucially uses the assumption that $\sum_{j\geq 1} P_j = 1$ a.s. Indeed, in this case, the preceding paragraph ensures the existence of a measure $\tilde{\mu}$ on $\mathscr{L}(\mathcal{D})$. Let us show that for $\tilde{\mu}$-almost every point x, we have

$$\liminf_{r\to 0} \frac{\log \tilde{\mu}(B(x,r))}{-\log r} \leq -\frac{1}{\beta}, \tag{4.26}$$

which will prove the proposition, using the mass distribution principle (see [59] for example). Actually, it is easy to see that, for (4.26) to hold, it is enough to provide a sequence $(r_n)_{n\geq 1}$ tending to 0 such that $\frac{\log r_n}{\log r_{n+1}} \to 1$ and

$$\liminf_{n\to\infty} \frac{\log \tilde{\mu}(B(x,r_n))}{-\log r_n} \leq -\frac{1}{\beta}, \tag{4.27}$$

Let us prove that (4.27) holds almost surely for a point L taken under the measure $\tilde{\mu}$ and a random sequence R_n. Using the product definition of μ, it is straightforward to see that if $\mathbf{I} = I_1 I_2, \dots \in \partial\mathbb{U}$ is taken under the measure μ, then the sequence $I_1, I_2, \dots$ is i.i.d. with the same distribution as I given by $\mathbb{P}(I = i \mid (P_j)_{j\geq 1}) = P_i$, where the sequence $(P_j)_{j\geq 1}$ has the distribution of that in the theorem. We can compute $\mathbb{E}[\log P_I] = \mathbb{E}[\mathbb{E}[\log P_I \mid (P_j)_{j\geq 1}]] = \mathbb{E}[\sum_{i=1}^\infty P_i \log P_i]$.

Let $R_n := \mathrm{d}(\rho_{\mathbf{I}_n}, \iota_{\mathcal{D}}(\mathbf{I}))$ be the distance of the random leaf $L := \iota_{\mathcal{D}}(\mathbf{I})$ to the root $\rho_{\mathbf{I}_n}$ of the block $\mathcal{D}(\mathbf{I}_n)$. Remark that the open ball $\mathrm{B}(L, R_n)$ of centre L and radius R_n only contains points that come from decorations with indices $u \succeq \mathbf{I}_n$ so that $\mu(\mathrm{B}(L, R_n)) \leq \mu(T(\mathbf{I}_n))$.

Now write

$$\log \mu(\mathrm{B}(L, R_n)) \leq \log \mu(T(\mathbf{I}_n)) = \sum_{i=1}^{n} \log P_{\mathbf{I}_i} \underset{n\to\infty}{\sim} n \cdot \mathbb{E}\left[\sum_{i=1}^{\infty} P_i \log P_i\right],$$

almost surely, because of the law of large numbers. Now it suffices to prove that almost surely

$$\log R_n \underset{n\to\infty}{\sim} n\beta \cdot \mathbb{E}\left[\sum_{i=1}^{\infty} P_i \log P_i\right], \tag{4.28}$$

and (4.27) would follow for the random leaf L thanks to the two last displays. In order to prove (4.28) we write

$$R_n := \mathrm{d}(\rho_{\mathbf{I}_n}, \iota_{\mathcal{D}}(\mathbf{I})) = \sum_{k=n}^{\infty} \left(\prod_{i=1}^{k} P_{\mathbf{I}_i}\right)^{\beta} \mathsf{D}_{\mathbf{I}_k}(\rho, X_{\mathbf{I}_{k+1}}),$$

using the definition of the distances in $\mathscr{G}(\mathcal{D})$. Then let us fix $\delta > 0$ such that $\mathbb{P}\left(\mathsf{D}(\rho, X_I) > \delta\right) > \delta$, and let $\tau_n = \inf\left\{i \geq n \;\middle|\; \mathsf{D}_{\mathbf{I}_i}(\rho, X_{\mathbf{I}_{i+1}}) > \delta\right\}$. Then we have

$$\mathbb{P}\left(\tau_n \geq n + \sqrt{n}\right) \leq (1-\delta)^{\sqrt{n}},$$

which is summable in n, so that using Borel-Cantelli lemma, we have $n \leq \tau_n \leq n + \sqrt{n}$ almost surely. Then for all n large enough

$$R_n \geq \left(\prod_{i=1}^{n+\sqrt{n}} P_{\mathbf{I}_i}\right)^{\beta} \cdot \delta,$$

and this proves that $\log R_n \geq \beta \sum_{i=1}^{n+\sqrt{n}} \log P_{\mathbf{I}_i} + \log \delta$. For an upper bound, remark that

$$R_n = \left(\prod_{i=1}^{n} P_{\mathbf{I}_i}\right)^{\beta} \cdot \underbrace{\left(\mathsf{D}_{\mathbf{I}_n}(\rho, X_{\mathbf{I}_{n+1}}) + P_{\mathbf{I}_{n+1}} \sum_{k=n+1}^{\infty} \left(\prod_{i=n+1}^{k} P_{\mathbf{I}_i}\right)^{\beta} \mathsf{D}_{\mathbf{I}_i}(\rho, X_{\mathbf{I}_{i+1}})\right)}_{R'_n},$$

where R'_n has the same law as R_0, which admits a finite first moment. Using the Markov inequality and the Borel-Cantelli lemma, we get that almost surely for any n large enough, $R'_n \leq n^p$ for some $p > 1$. Then for all $n \geq 1$ large enough

$$\log R_n = \log \left(\prod_{i=1}^{n} P_{\mathbf{I}_i}\right)^{\beta} + \log R'_n \leq \beta \sum_{i=1}^{n} \log P_{\mathbf{I}_i} + p \log n.$$

In the end, using the upper and lower bound on R_n and the law of large numbers we get (4.28), which proves the proposition. □

Almost-self-similar decorations. For our needs, we define a slight variation of this model where we only suppose that the random variables $(\xi_u)_{u\in\mathbb{U}\setminus\{\emptyset\}}$ have the same law as ξ, and $\xi_\emptyset = ((\mathsf{B}_\emptyset, \mathsf{D}_\emptyset, \rho_\emptyset, (X_i)_{i\geq 1}), (P_i)_{i\geq 1})$ is independent of the other $(\xi_u)_{u\in\mathbb{U}\setminus\{\emptyset\}}$ but can possibly have a different law.

In this case, we say that the obtained $\mathcal{D}$ is almost-self-similar with exponent β and spine distributions $\xi_\emptyset$ and ξ. If $\xi_\emptyset$ satisfies the conditions of Lemma 4.7 as well as ξ, the above arguments still hold and the law of the obtained metric space is the law of $\phi_\beta(\xi_\emptyset, (\tau_i)_{i\geq 1})$ where the $(\tau_i)_{i\geq 1}$ are i.i.d. with distribution η which is the unique fixed point of Φ_β.

4.4.3 Structures constructed by iterative gluing can also be self-similar

Some random decorations that are described using an iterative gluing construction also belong to the family of almost-self-similar decorations. The following proposition ensures that this is the case for a particular family of iterative gluing constructions.

Proposition 4.9. *Suppose that $\mathcal{D}$ is defined as an iterative gluing construction using*

(i) *a sequence of weights $(\mathsf{m}_n)_{n\geq 1}$ defined as the increments of a Mittag-Leffler Markov chain $(\mathsf{M}_n)_{n\geq 1} \sim \mathrm{MLMC}(\frac{1}{b+1}, \frac{a}{b+1})$,*

(ii) *a sequence of scaling factors taken as $(\mathsf{m}_n^\gamma)_{n\geq 1}$ for some $\gamma > 0$,*

(iii) *a sequence of independent blocks $(\mathsf{B}_n, \mathsf{D}_n, \rho_n, (X_{n,i})_{i\geq 1})$, with the same distribution starting from $n \geq 2$ such that their diameter admits a p-th moment with $p > 1$.*

Then $\mathcal{D}$ is an almost-self-similar decoration with exponent $\frac{\gamma}{b+1}$ and spine distributions $\xi_\emptyset$ and ξ such that

- $\xi_\emptyset \overset{(d)}{=} \left((\mathsf{B}_1, S_\emptyset^\gamma \cdot \mathsf{D}_1, \rho_1, (X_{1,i})_{i\geq 1}), (P_i)_{i\geq 1}\right)$, *with $(P_i)_{i\geq 1} \sim \mathrm{GEM}\left(\frac{1}{b+1}, \frac{a}{b+1}\right)$ independent of B_1 and $S_\emptyset$ its $\frac{1}{b+1}$-diversity,*
- $\xi \overset{(d)}{=} \left((\mathsf{B}_2, S^\gamma \cdot \mathsf{D}_2, \rho_2, (X_{2,i})_{i\geq 1}), (\tilde{P}_i)_{i\geq 1}\right)$, *with $(\tilde{P}_i)_{i\geq 1} \sim \mathrm{GEM}\left(\frac{1}{b+1}, \frac{b}{b+1}\right)$ independent of B_2 and S its $\frac{1}{b+1}$-diversity.*

Proof. Recall the definition of μ the probability measure on $\partial\mathbb{U}$ obtained as the weak limit of the mass measure of the weighted recursive tree used for this iterative construction. If we denote for all $u \in \mathbb{U}$ and $i \in \mathbb{N}$

$$\mathsf{p}_{ui} = \frac{\mu(T(ui))}{\mu(T(u))},$$

then from Proposition 4.6 we have a complete description of the distribution of $(\mathsf{p}_u)_{u\in\mathbb{U}}$ using GEM distributions. For every $u \in \mathbb{U}$, we let S_u be the $\frac{1}{b+1}$-diversity of the sequence $(\mathsf{p}_{ui})_{i\geq 1}$. Hence, denoting

$$\xi_\emptyset := \left((\mathsf{B}_1, S_\emptyset^\gamma \cdot \mathsf{D}_1, \rho_1, (X_{1,i})_{i\geq 1}), (\mathsf{p}_i)_{i\geq 1}\right)$$
$$\xi_{u_k} := \left((\mathsf{B}_k, S_{u_k}^\gamma \cdot \mathsf{D}_k, \rho_k, (X_{k,i})_{i\geq 1}), (\mathsf{p}_{u_k i})_{i\geq 1}\right) \qquad \forall k \geq 2.$$

it is immediate from the previous section that the $(\xi_u)_{u\in\mathbb{U}}$ are independent and $(\xi_u)_{u\in\mathbb{U}\setminus\{\emptyset\}}$ are i.i.d. Hence the distribution of $\mathcal{D}$ coincides with that of an almost-self-similar decoration with scaling exponent $\frac{\gamma}{b+1}$ with these spine distributions. ☐

4.5 Application to models of growing random graphs

Let us use this framework of random decorations to prove scaling limits for random graph models. We first present a general proof that will apply to all our different applications. Every example that we treat is of the following form: we start with a model of objects $(H_n)_{n\geq 1}$ defined iteratively, that can be considered as measured metric spaces.

For every $n \geq 1$, we construct simultaneously a decoration $\mathcal{D}^{(n)}$ and measures $\boldsymbol{\nu}^{(n)}$, supported on a tree P_n, such that the distribution of the sequence $\left(\mathscr{G}(\mathcal{D}^{(n)}), \boldsymbol{\nu}^{(n)}\right)_{n\geq 1}$ coincides with that of $(H_n)_{n\geq 1}$, seen as measured metric spaces.

4.5.1 An abstract result that handles all our applications

Assume that we study a sequence $(\mathcal{D}^{(n)})_{n\geq 1}$ of decorations endowed with measures $\boldsymbol{\nu}^{(n)}$ which are at each time supported on the trees $(\mathsf{P}_n)_{n\geq 1}$, constructed using processes $(\mathcal{A}_k(m), m \geq 0)_{k\geq 1}$ with values in $\mathbb{M}^{\infty\bullet}$ that are jointly independent and independent of $(\mathsf{P}_n)_{n\geq 1}$, in such a way that

(i) the sequence $(\mathsf{P}_n)_{n\geq 1}$ evolves as a preferential attachment tree with some sequence of fitnesses $\mathbf{a} = (a_n)_{n\geq 1}$ which satisfies ($H_{c,c'}$) for some $c > 0$ and $0 \leq c' < \frac{1}{c+1}$;

(ii) for all $n \geq 1$, the decoration $\mathcal{D}^{(n)}$ is such that for all $k \in \{1, \ldots, n\}$,

$$\mathcal{D}^{(n)}(u_k) = \mathcal{A}_k(\deg^+_{\mathsf{P}_n}(u_k)),$$

and for all $u \notin \{u_1, \ldots, u_n\}$, the associated block is trivial i.e. $\mathcal{D}^{(n)}(u) = (\star, 0, \star, (\star)_{i\geq 1})$;

(iii) there exists $\gamma > 0$ such that for all $k \geq 1$,

$$m^{-\gamma} \cdot \mathcal{A}_k(m) \xrightarrow[m\to\infty]{\text{a.s.}} (\mathsf{B}_k, \mathsf{D}_k, \rho_k, (X_{k,i})_{i\geq 1}) \quad \text{in } \mathbb{M}^{\bullet\infty};$$

(iv) there exists a sequence $(x_k)_{k\geq 1}$ such that for all $p > 0$ we have

$$\sup_{k\geq 1} \mathbb{E}\left[\sup_{m\geq 1} \left(\frac{\operatorname{diam}(\mathcal{A}_k(m))}{(m + x_k)^{\gamma}}\right)^{p}\right] < \infty,$$

and the sequence $(x_k)_{k\geq 1}$ is such that $x_k \leq k^{s+o(1)}$ for some $s < \frac{1}{c+1}$;

(v) the associated measures $(\nu^{(n)})_{n\geq 1}$ are probability measures of the form (4.19) or are uniform on $\{u_1, \ldots, u_n\}$.

Under all those assumptions we have a convergence result for our decorations.

Theorem 4.10. *Suppose that the decorations $\mathcal{D}^{(n)}$ are constructed as above. Then we have the following almost sure convergence*

$$\mathscr{G}(n^{-\frac{\gamma}{c+1}} \cdot \mathcal{D}^{(n)}, \boldsymbol{\nu}^{(n)}) \underset{n\to\infty}{\longrightarrow} \mathscr{G}(\mathcal{D}, \nu) \qquad \textit{in the GHP topology,}$$

where the limit is described as an iterative construction with blocks $(\mathsf{B}_k, \mathsf{D}_k, \rho_k, (X_{k,i})_{i\geq 1})_{k\geq 1}$, scaling factors $((\mathsf{m}^{\mathbf{a}}_k)^{\gamma})_{k\geq 1}$ and weights $(\mathsf{m}^{\mathbf{a}}_k)_{k\geq 1}$.

Proof. First we check that for all $k \geq 1$ and all $n \geq k$ we have, almost surely,

$$\begin{aligned} n^{-\frac{\gamma}{c+1}} \cdot \mathcal{D}^{(n)}(u_k) &= n^{-\frac{\gamma}{c+1}} \cdot \mathcal{A}_k(\deg^+_{\mathsf{P}_n}(u_k)) \\ &= \left(n^{-\frac{1}{c+1}} \deg^+_{\mathsf{P}_n}(u_k)\right)^{\gamma} \cdot (\deg^+_{\mathsf{P}_n}(u_k))^{-\gamma} \cdot \mathcal{A}_k(\deg^+_{\mathsf{P}_n}(u_k)) \\ &\underset{n\to\infty}{\longrightarrow} (\mathsf{m}^{\mathbf{a}}_k)^{\gamma} \cdot (\mathsf{B}_k, \mathsf{D}_k, \rho_k, (X_{k,i})_{i\geq 1}), \end{aligned}$$

in the topology of $\mathbb{M}^{\infty\bullet}$. By (iii), for any $u \notin \{u_1, u_2, \ldots\}$ the block $\mathcal{D}^{(n)}(u)$ is constant and equal to the trivial block. So $n^{-\frac{\gamma}{c+1}} \cdot \mathcal{D}^{(n)}$ converges to some limiting decoration $\mathcal{D}$ in the product topology.

Now, for any $k \geq 1$ and $n \geq k$ we have

$$\operatorname{diam}\left(n^{-\frac{\gamma}{c+1}} \cdot \mathcal{D}^{(n)}(u_k)\right) = \left(n^{-\frac{1}{c+1}} \left(\deg^+_{\mathsf{P}_n}(u_k) + x_k\right)\right)^{\gamma} \cdot \frac{\operatorname{diam}(\mathcal{A}_k(\deg^+_{\mathsf{P}_n}(u_k)))}{(\deg^+_{\mathsf{P}_n}(u_k) + x_k)^{-\gamma}}.$$

The first term can be shown to be smaller than some random bound $(k+1)^{-\epsilon+o_\omega(1)}$ using (4.16). The second term is bounded above by some $k^{o_\omega(1)}$ thanks to (iv). In the end we almost surely have

$$\operatorname{diam}\left(n^{-\frac{\gamma}{c+1}} \cdot \mathcal{D}^{(n)}(u_k)\right) \leq k^{-\epsilon+o_\omega(1)}. \tag{4.29}$$

Using the logarithmic growth of the trees $(\mathtt{P}_n)_{n\geq1}$ and Lemma 4.5, we get the almost sure non-explosion of the function $\ell : u \mapsto \sup_{n\geq1} \operatorname{diam}\left(n^{-\frac{\gamma}{c+1}} \cdot \mathcal{D}^{(n)}(u)\right)$. Thanks to Theorem 4.2, this ensures the almost sure Gromov-Hausdorff convergence of the spaces $\mathscr{G}(n^{-\frac{\gamma}{c+1}} \cdot \mathcal{D}^{(n)})$ towards $\mathscr{G}(\mathcal{D})$. Now, thanks to (v) and (4.15) we can add the measures and the convergence takes place instead in the GHP topology, and hence the theorem is proved. $\square$

How to apply this theorem. This theorem may seem abstract at that point, but it encompasses all our specific examples of growing random graphs. Now for all our sequences of graphs, the goal will be to provide a sequence $(a_n)_{n\geq1}$ satisfying ($H_{c,c'}$) for some parameters c and c' and processes $(\mathcal{A}_k)_{k\geq1}$ so that the decorations $(\mathcal{D}^{(n)})_{n\geq1}$ satisfying (i) and (ii) indeed evolve in such a way that $\mathscr{G}(\mathcal{D}^{(n)})$ coincides with our process. Then we check that the other assumptions are also satisfied in order to get the scaling limit.

Particular form of processes $\mathcal{A}$. In general, remark that for any $k \geq 1$ and $m \geq 0$, all the distinguished points in $\mathcal{A}_k(m)$ that matter for the construction are the first m ones. All the others can be set equal to the root vertex without changing the distribution of $\left(\mathscr{G}(\mathcal{D}^{(n)})\right)_{n\geq1}$, so we can always suppose that at each step $m \geq 0$, the metric space $\mathcal{A}_k(m)$ is endowed with only m distinguished points in addition to the root and can hence be seen as an element of $\mathbb{M}^{m\bullet}$.

In all our examples, the different processes $\mathcal{A}_k$ for $k \geq 0$ all evolve under the same Markovian transitions, possibly starting from different states $\mathcal{A}_k(0)$ for different values of $k \geq 1$. These transitions are often more naturally defined on *weighted graphs*, in which each of the vertices and edges are given some weight. The dynamics involves taking a vertex or edge at random proportionally to its weight, do some local transformation of the graph at that point by possibly adding one or several vertices and edges to the graph. The list of distinguished points is then updated by appending some vertex to the end of the existing list.

Almost self-similar limits. Theorem 4.10 describes the limiting space as the result of an iterative construction. In our examples, it will be often the case that Proposition 4.9 applies to the limiting space and hence that it is almost-self-similar in the sense of Section 4.4.2. We will not detail further this type of description, this is left to the reader.

4.5.2 Generalised Rémy algorithm

Recall the construction described in the introduction. Consider $(G_n, o_n)_{n\geq1}$ a sequence of finite rooted graphs with respective number of edges $\mathbf{a} = (a_n)_{n\geq1}$ which satisfies ($H_{c,c'}$) for some $c > 0$ and $c' < \frac{1}{c+1}$. We construct the sequence $(H_n)_{n\geq1}$ recursively as follows. Let $H_1 = G_1$. Then, for any $n \geq 1$, conditionally on the structure H_n already constructed, take an edge in H_n uniformly at random, split it into two edges by adding a vertex "in the middle" of this edge, and glue a copy of G_{n+1} to the structure by identifying o_{n+1} the root vertex of G_{n+1} with the newly created vertex. Call the obtained graph H_{n+1}. See Figure 4.2 for a realisation of H_5 using the sequence $(G_n)_{n\geq1}$ of Figure 4.1.

Construction as a gluing of decorations. Let us provide a construction of some sequence $(\mathcal{D}^{(n)})_{n\geq 1}$ of decorations, endowed with measures $(\boldsymbol{\nu}^{(n)})_{n\geq 1}$, that satisfies the assumptions of Theorem 4.10 and for which the process $(\mathscr{G}(\mathcal{D}^{(n)},\boldsymbol{\nu}^{(n)}))_{n\geq 1}$ coincides with $(H_n)_{n\geq 1}$ endowed with its graph distance and the uniform measure on its vertices.

For this let $(\mathsf{P}_n)_{n\geq 1}$ be a preferential attachment tree with fitnesses $(a_n)_{n\geq 1}$ and for all $k\geq 1$, define the process $(\mathcal{A}_k(m))_{m\geq 0}$ as follows: $\mathcal{A}_k(0)$ is just (the set of vertices of) the graph G_k endowed with the corresponding graph distance, rooted at o_k, with an empty list of distinguished points. Then $\mathcal{A}_k(m+1)$ is obtained from $\mathcal{A}_k(m)$ by duplicating an edge uniformly at random by adding some point x_{m+1} in its centre. The vertex x_{m+1} is then appended at the end of the list of distinguished points, now becoming of length $m+1$.

Now, consider the measures $\boldsymbol{\nu}^{(n)}$ such that for all $u\in\mathbb{U}$, $\nu_u^{(n)}$ charges every point of $\mathcal{D}^{(n)}(u)$ except its root if $u\neq\emptyset$, with the same mass, normalised in such a way that the associated measure $\nu^{(n)}$ on the Ulam tree is a probability measure.

It is now an exercise to check that the sequence of graphs $(H_n)_{n\geq 1}$ seen as measured metric spaces has the same distribution as $(\mathscr{G}(\mathcal{D}^{(n)},\boldsymbol{\nu}^{(n)})_{n\geq 1}$.

Applying the theorem. Assumptions (i) and (ii) are satisfied by construction, so let us verify that the other ones hold as well.

The convergence (iii) is quite easy to prove for $\gamma=1$. The limiting block $(\mathsf{B}_k,\mathsf{D}_k,\rho_k,(X_{k,i})_{i\geq 1})$ can be described as a continuous version of the graph G_k, where each edge $e\in\{e_1,\dots,e_{a_k}\}$ has been replaced by a segment of length $L(e)$ where the lengths are such that

$$(L(e_1),L(e_2),\dots,L(e_{a_k}))\sim \mathrm{Dir}(1,1,\dots,1), \tag{4.30}$$

so that the total length is 1. The $(X_{k,i})_{i\geq 1}$ are then obtained conditionally on this construction as i.i.d. points taken under the length measure.

The control (iv) is immediate with $(x_k)_{k\geq 1}=(a_k)_{k\geq 1}$ because we have the deterministic upper-bound for all $k\geq 1$ and $m\geq 0$, $\mathrm{diam}(\mathcal{A}_k(m))\leq a_k+m$.

The last point (v) is obtained by using (4.19). Indeed, if we let $(b_n)_{n\geq 1}$ be defined such that b_1 is the number of vertices of G_1 and for $n\geq 2$, b_n is the number of vertices minus 1 of the graph G_n. Then the measures $\nu^{(n)}$ on the Ulam tree are probability measures of the form $\nu^{(n)}(u_k)\propto b_n+\deg^+_{\mathsf{P}_n}(u_k)$ for all $k\leq n$ and $\nu^{(n)}(u)=0$ on other vertices u. It is easy to check that the sequence $(b_n)_{n\geq 1}$ satisfies the appropriate condition so that it converges almost surely weakly to a limiting measure ν on $\partial\mathbb{U}$, which is also the limit of the weight measure.

In the end, this yields a proof of Proposition 4.1.

4.5.3 Generalised Rémy algorithm, version 2

Let us use a different decomposition of the preceding model, which will lead to another description of the limit for a particular sequence of graphs $(G_n)_{n\geq 1}$ with G_1 equal to the single-edge graph and constant starting from the second term, equal to a line with two edges, rooted at one end. The limiting space has a particularly nice description as a gluing of rescaled i.i.d Brownian trees.

In this case we can take $(a_n)_{n\geq 1}=(\frac{1}{2},\frac{1}{2},\dots)$ and the processes $(\mathcal{A}_k)_{k\geq 1}$ all have the same distribution as the standard Rémy algorithm started from a single edge, such that the distinguished points correspond to the added leaves in order of creation. We also add the measures $\boldsymbol{\nu}^{(n)}$ in the same way as before.

Remark that the last paragraph would describe the limit as an iterative gluing construction using blocks that are all equal to the $[0\,,1]$ interval rooted at 0 endowed with i.i.d. random points.

The associated sequence of scaling factors and weights would then be both equal to a sequence $(\mathsf{m}_n)_{n\geq 1}$ having the distribution of the increment of a $\mathrm{MLMC}(\frac{1}{3},\frac{1}{3})$ Mittag-Leffler Markov chain.

Evolution of decorations. Using [42, Theorem 5], for any $k\geq 1$,

$$m^{-\frac{1}{2}}\cdot \mathcal{A}_k(m) \to (\mathsf{B}_k,\mathsf{D}_k,\rho_k,(X_{k,i})_{i\geq 1}) \qquad \text{in } \mathbb{M}^{\infty\bullet}, \tag{4.31}$$

where the limiting metric space $(\mathsf{B}_k,\mathsf{D}_k,\rho_k,(X_{k,i})_{i\geq 1})$ has the distribution of (2 times) the Brownian tree, endowed with an i.i.d. sequence of points taken under its mass measure.

The condition (iv) is satisfied thanks to Lemma 4.16, proved in the appendix. It is easy to check that the measure have the form (4.19), so that (v) is satisfied and so Theorem 4.10 applies, which proves the following:

Proposition 4.11. *Under these conditions, the limiting space can be constructed by an iterative gluing construction using scaling factors* $((\mathsf{m}_n^{\mathbf{a}})^{\frac{1}{2}})_{n\geq 1}$ *and weights* $(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1}$*, where the sequence* $(\mathsf{m}_n^{\mathbf{a}})_{n\geq 1}$ *has the distribution of the increments of a* $\mathrm{MLMC}\left(\frac{2}{3},\frac{1}{3}\right)$ *process, and i.i.d. blocks with the distribution of 2 times the Brownian tree endowed with a sequence of i.i.d. leaves taken under their mass measure.*

This convergence proves in particular the non-trivial fact that the two iterative gluing constructions with segments or Brownian trees lead to the same object.

4.5.4 Marchal algorithm started from an arbitrary seed

Let us define Marchal's algorithm started from a rooted connected multigraph G, as introduced in [64], following the same idea as [92]. Fix $\alpha\in(1,2)$. We let $H_1^\alpha=G$ and for each $n\geq 1$ define H_{n+1}^α recursively. If H_n^α is defined then take a vertex or an edge with probability proportional to their weight, where their weights are defined as

- $\alpha-1$ for any edge,
- $\deg(v)-1-\alpha$ for a vertex v with degree 3 or more,
- 0 for a vertex of degree 2 or less.

Then if it is an edge, split this edge into 2 edges with a common endpoint, add an edge linking that newly created vertex to a new leaf. Otherwise, attach an edge linking the selected vertex to a new leaf. The obtained graph is then H_{n+1}^α.

This construction is also defined and studied in Chapter 5, where Proposition 5.2, Theorem 5.5 and Theorem 5.8 already ensure that these graphs appropriately rescaled converge in the GHP topology to some random object that is constructed using an iterative gluing construction. In this case, there are two natural interpretations of this graph process in terms of decorations on the Ulam tree, which give two different descriptions of the limiting object: one of them coincides with the one given in Chapter 5, but the other one is different. We only describe this one. Note that if we denote $w(G)$ the sum of the weights of all vertices and edges of a graph G with surplus s, with ℓ vertices of degree 1 and m vertices of degree 2 then $w(G)=(\ell-1)\alpha+m(\alpha-1)+s(\alpha+1)-1$.

Splitting the width

In this first decomposition, we take $(a_n) = (w(G), \alpha - 1, \alpha - 1, \dots)$ and $(\mathtt{P}_n)_{n\geq 1}$ taken as a preferential attachment tree $\mathrm{PA}(\mathbf{a})$. The processes $(\mathcal{A}_k)_{k\geq 1}$ all follow the same Markov transitions on weighted graphs, starting from the rooted graph G in the case of $\mathcal{A}_1$ and from the single-edge graph for $\mathcal{A}_k$ for any $k \geq 2$.

In this setting, the weight of an edge is always $\alpha - 1$ but the weight of vertices evolves with time. At time 0, for any seed graph G, every vertex with degree $d \geq 3$ is given weight $d - 1 - \alpha$ and other vertices have weight 0. The process evolves then in the following way: to obtain $\mathcal{A}(m+1)$ from $\mathcal{A}(m)$ we choose at random an edge or a vertex proportionally to their weights:

- if it is a vertex x then $\mathcal{A}(m+1)$ is obtained from $\mathcal{A}(m)$ by setting its $(m+1)$-st distinguished point to be x, and incrementing the weight of x by one,
- if it is an edge, then $\mathcal{A}(m+1)$ is obtained from $\mathcal{A}(m)$ by splitting this edge in 2 by adding a new vertex x, giving weight $2 - \alpha$ to this vertex, and setting its $(m+1)$-st distinguished point to x.

With this dynamics, we have the following convergence almost surely in $\mathbb{M}^{\infty\bullet}$, for $\mathcal{A}$ starting from any graph G with at least one edge,

$$m^{-(\alpha-1)} \cdot \mathcal{A}(m) \to \mathcal{M}_\alpha^{\mathrm{wid}}(G), \tag{4.32}$$

as $n \to \infty$ in $\mathbb{M}^{\infty\bullet}$, where the distribution of the limiting object is described below.

This convergence is obtained using a coupling with a Chinese restaurant process with parameters $(\alpha - 1, |E| \cdot (\alpha - 1))$, for which the number of tables is an upper bound on the diameter of $\mathcal{A}(m)$. We do not detail the construction here, but the condition (iv) is obtained using Lemma 4.15 stated in the Appendix of this chapter.

Limiting block. Let us describe the law of the random metric space $\mathcal{M}_\alpha^{\mathrm{wid}}(G) = (\mathcal{G}, \mathrm{d}, \rho, (X_i)_{i\geq 1})$. It is a continuous version of G meaning that we define it by replacing every edge of G by a segment of some length. We label its edges $e_1, e_2, \dots e_{|E|}$ in arbitrary order and replace each edge e with a segment of length $L(e)$, whose distribution is characterised by what follows. We let I be the set of vertices of G that have degree greater than 3 and write $I = \{v_1, v_2, ..., v_{|I|}\}$. All the random variables used in the construction are supposed to be independent.

- We let
$$\left(W_E, W_{v_1}, \dots, W_{v_{|I|}}\right) \sim \mathrm{Dir}\left(|E| \cdot (\alpha - 1), d_{v_1} - 1 - \alpha, \dots, d_{v_{|V|}} - 1 - \alpha\right)$$
- We let
$$(P_i)_{i\geq 1} \sim \mathrm{PD}\left(\alpha - 1, |E| \cdot (\alpha - 1)\right), \quad \text{and} \quad S \quad \text{its } (\alpha - 1)\text{-diversity},$$
- The length of the edges are defined as
$$(L(e_1), L(e_2), \dots, L(e_{|E|})) = W_e^{\alpha-1} \cdot S \cdot (B_1, B_2, \dots, B_{|E|}),$$
with $(B_1, B_2, \dots, B_{|E|}) \sim \mathrm{Dir}(1, 1, \dots 1)$.

- Conditionally on the lengths, let $(Z_i)_{i\geq 1}$ be independent and uniformly distributed on the total length of the graph.

- We set

$$\mu = \sum_{v\in I} W_v \delta_v + \sum_{i=1}^{\infty} P_i \delta_{Z_i},$$

and conditionally on all the rest, the points $(X_i)_{i\geq 1}$ are obtained as i.i.d. samples under the probability measure μ.

Convergence result. We get the following convergence:

Proposition 4.12. *The graphs H_n^α endowed with the uniform measure on their vertices converge almost surely in Gromov-Hausdorff-Prokhorov topology:*

$$n^{1-1/\alpha} \cdot H_n^\alpha \underset{n\to\infty}{\longrightarrow} (\mathcal{G}^\alpha(G), \mathrm{d}, \mu).$$

The distribution of the limiting space $(\mathcal{G}^\alpha(G), \mathrm{d}, \mu)$ is obtained as an iterative gluing construction with blocks

$$(\mathsf{B}_1, \mathsf{D}_1, \rho_1, \nu_1) \sim \mathcal{M}_\alpha^{\mathrm{wid}}(G) \quad \textit{and} \quad \forall n \geq 2,\ (\mathsf{B}_n, \mathsf{D}_n, \rho_n, \nu_n) \sim \mathcal{M}_\alpha^{\mathrm{wid}}(-)$$

and sequence of scaling factors $(\mathsf{m}_k^{(\alpha-1)})_{k\geq 1}$ and weights $(\mathsf{m}_k)_{k\geq 1}$, where $(\mathsf{m}_k)_{k\geq 1}$ is obtained as the increments of a $\mathrm{MLMC}(\frac{1}{\alpha}, \frac{w(G)}{\alpha})$.

Another decomposition

Using another decomposition into decorations we retrieve the other description of the limiting space that is given in Chapter 5. It is obtained as an iterative gluing construction with blocks (the distribution of which we define below)

$$(\mathsf{B}_1, \mathsf{D}_1, \rho_1, \nu_1) \sim \mathcal{M}_\alpha^{\mathrm{len}}(G) \quad \text{and} \quad \forall n \geq 2,\ (\mathsf{B}_n, \mathsf{D}_n, \rho_n, \nu_n) \sim \mathcal{M}_\alpha^{\mathrm{len}}(\bullet-),$$

and sequence of weights and scaling factors $(\mathsf{m}_k)_{k\geq 1}$, where $(\mathsf{m}_k)_{k\geq 1}$ is obtained as the increments of a $\mathrm{MLMC}(1-\frac{1}{\alpha}, \frac{w(G)}{\alpha})$.

Let us describe the random metric space $\mathcal{M}_\alpha^{\mathrm{len}}(G) = (\mathcal{G}, \mathrm{d}, \rho, (X_i)_{i\geq 1})$, for any rooted connected multigraph G. As before, we let I be the set of vertices of G that have degree greater than 3 and write $I = \{v_1, v_2, ..., v_{|I|}\}$ and we arbitrarily label its edges by $e_1, e_2, \dots e_{|E|}$. Then

- we define

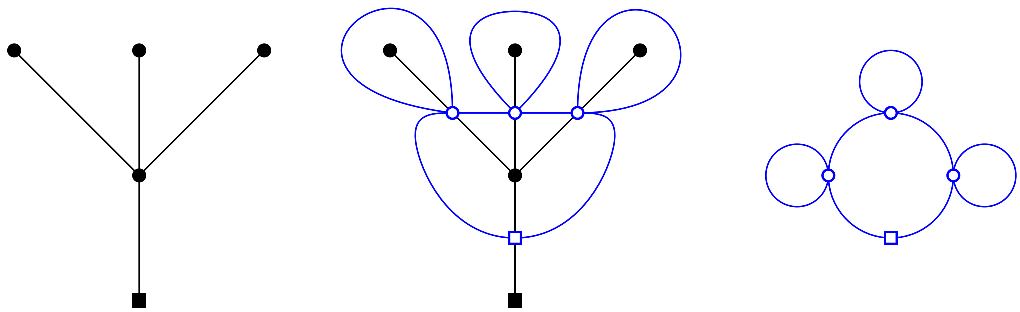
$$\big(L(e_1), \dots, L(e_{|E|}), L(v_1), \dots, L(v_{|I|})\big) \sim \mathrm{Dir}\left(1, \dots, 1, \frac{d_{v_1} - 1 - \alpha}{\alpha - 1}, \dots, \frac{d_{v_{|I|}} - 1 - \alpha}{\alpha - 1}\right),$$

- and set

$$\nu = \sum_{v\in I} L(v) \cdot \delta_v + \mu_{\mathrm{len}},$$

and conditionally on all the rest, the sequence $(X_i)_{i\geq 1}$ is i.i.d. with distribution ν.

For the single-edge graph, this yields only a segment of unit length endowed with the uniform measure, but we also introduce a variant of this one. We define $\mathcal{M}_\alpha^{\mathrm{len}}(\bullet-) = (\mathcal{G}, \mathrm{d}, \rho, \nu)$ as follows:



(a) A plane tree τ (b) Construction of the loops (c) The obtained $\mathrm{Loop}(\tau)$

Figure 4.5 – The construction of $\mathrm{Loop}(\tau)$ from the plane tree τ.

- We set
$$(L(e), L(\rho)) \sim \mathrm{Dir}\left(1, \frac{2-\alpha}{\alpha-1}\right),$$
- and also
$$\nu = L_\rho \delta_\rho + \mu_{\mathrm{len}},$$
and conditionally on this, the sequence $(X_i)_{i\geq 1}$ is i.i.d. with distribution ν.

4.5.5 Scaling limits for growing trees and/or their looptrees.

The looptree $\mathrm{Loop}(\tau)$ of a plane tree τ is a multigraph constructed from τ as follows: we first place a blue vertex in the middle of every edge of the tree τ. Then, we connect two blue vertices if they correspond to two consecutive edges according to the cyclic ordering around vertices that are not the root, as pictured in this is illustrated in Figure 4.5. Then $\mathrm{Loop}(\tau)$ is obtained by removing all the vertices and edges that belong to the tree τ.

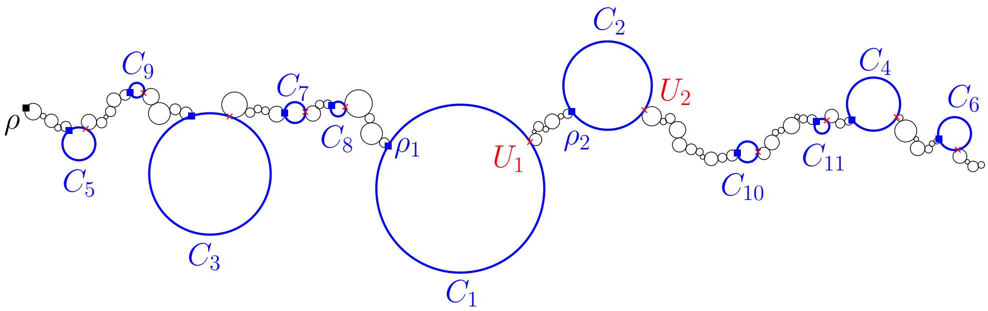
Whenever we work with a model of trees that have degrees that grow to infinity, studying the associated looptrees may allow to pass this information to the limit in terms of metric scaling limits. Among the two models that we present here, one of them admits scaling limits for both the tree itself and its associated looptree. For the other one, only the looptree behaves well in this sense.

The $\alpha-\gamma$−growth model

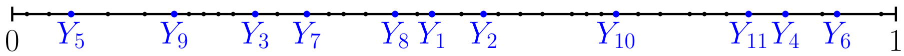
Fix $\alpha \in (0,1)$ and $\gamma \in (0,\alpha]$. The $\alpha-\gamma$-growth model is defined as follows: $T_1^{\alpha,\gamma}$ is a tree with a single edge. Then if $T_n^{\alpha,\gamma}$ is already constructed, take an edge or a vertex at random with probability proportional to

- $1-\alpha$ for edges that are adjacent to a leaf,
- γ for edges that are not adjacent to a leaf,
- $(d-1)\alpha-\gamma$ for every vertex of degree $d+1\geq 3$.

Then as in Marchal's algorithm, if an edge is chosen, it is split into two edges and a new edge leading to a new leaf is grafted at the newly created vertex. If a vertex is chosen, add an edge connecting it to a new leaf. We can also use a planar variation of this algorithm, where every time that we attach a new leaf to a vertex we attach it in a uniform corner around this vertex.



(a) The block $\mathcal{B}^{\alpha,\gamma}$ constructed from a countable number of circles $(C_n)_{n\geq 1}$



(b) The points $(Y_n)_{n\geq 1}$ along the segment $[0\,,1]$

Figure 4.6 – The block $\mathcal{B}^{\alpha,\gamma}$ is constructed by agglomerating countably many circles $(C_n)_{n\geq 1}$, in the order given by the relative position of points $(Y_n)_{n\geq 1}$. Since they are dense in $[0\,,1]$, no two circles are ever adjacent.

Proposition 4.13. *We have the following joint convergence almost surely in the Gromov–Hausdorff–Prokhorov topology*

$$(n^{-\gamma}\cdot T_n^{\alpha,\gamma}, n^{-\alpha}\cdot \mathrm{Loop}(T_n^{\alpha,\gamma})) \underset{n\to\infty}{\longrightarrow} (\mathcal{T}^{\alpha,\gamma},\mathcal{L}^{\alpha,\gamma}).$$

The limiting objects can be constructed using an iterative gluing construction with i.i.d. blocks using a weight sequence $(\mathsf{m}_n)_{n\geq 1}$ *obtained as the increment of a Mittag-Leffler Markov chain* $\mathrm{MLMC}(\alpha, 1-\alpha)$. *The scaling factors are taken as* $(\mathsf{m}_n^{\gamma/\alpha})_{n\geq 1}$ *for the first one and* $(\mathsf{m}_n)_{n\geq 1}$ *for the second one, using block i.i.d. blocks with the same joint distribution as* $(\mathcal{S}^{\alpha,\gamma},\mathcal{B}^{\alpha,\gamma})$ *which we define below.*

Joint construction of the limiting blocks. Let us define a random sequence $(Y_n)_{n\geq 1}$ on $[0\,,1]$ as follows:

- Let $Y_1 \sim \mathrm{Beta}(1, \frac{1-\alpha}{\gamma})$.
- Recursively, if $(Y_1,\dots,Y_n)$ are already defined then conditionally on them the point Y_{n+1} is distributed uniformly on $[0\,,\max_{1\leq i\leq n} Y_i]$ with probability $(\max_{1\leq i\leq n} Y_i)$ and as $1-R_n\cdot(1-\max_{1\leq i\leq n} Y_i)$ with complementary probability, with $R_n\sim\mathrm{Beta}(1,\frac{1-\alpha}{\gamma})$ independent of everything else.

Then, if $\gamma=\alpha$, the couple $(\mathcal{S}^{\alpha,\gamma},\mathcal{B}^{\alpha,\gamma})$ is such that $\mathcal{S}^{\alpha,\gamma}=\mathcal{B}^{\alpha,\gamma}$ which are just defined as the interval $[0\,,1]$, rooted at 0 and endowed with the points $(Y_n)_{n\geq 1}$.

If $\gamma\neq\alpha$, we define the following random variables, independently of the sequence $(Y_n)_{n\geq 1}$:

- We let $(P_i)_{i\geq 1} \sim \mathrm{GEM}(\frac{\gamma}{\alpha}, \frac{1-\alpha}{\alpha})$ with $\frac{\gamma}{\alpha}$-diversity S.
- Define recursively the sequence $(N_k)_{k\geq 1}$ starting from $N_1 = 1$. Conditionally on $(N_1, \dots, N_k)$ we have

$$\begin{aligned} N_{k+1} = i \quad &\text{with probability } P_i \quad \text{for any } i \in \{1, \dots, \max_{1\leq i\leq k} N_i\}, \\ = 1 + \max_{1\leq i\leq k} N_i \quad &\text{with complementary probability.} \end{aligned}$$

- The sequence $(X_k)_{k\geq 1}$ is then defined on the interval $[0\,, S]$ as $(S \cdot Y_{N_k})_{k\geq 1}$.

The block $\mathcal{S}^{\alpha,\gamma}$ is then defined as the interval $[0\,, S]$ rooted at 0 endowed with the sequence $(X_k)_{k\geq 1}$. In order to construct $\mathcal{B}^{\alpha,\gamma}$, we introduce a sequence $(C_i)_{i\geq 1}$ of circles such that for all $i \geq 1$, (C_i, d_i, ρ_i) is a circle with circumference P_i endowed with its path distance and rooted at some point ρ_i. Conditionally on that, we take on each C_i a point U_i and a sequence $(V_{i,j})_{j\geq 1}$ of i.i.d. uniform random points on C_i. Then we consider their disjoint union

$$\bigsqcup_{i=1}^{\infty} C_i, \tag{4.33}$$

which we endow with the distance d characterised by

$$\begin{aligned} \mathrm{d}(x, y) &= d_i(x, y) \qquad \text{if} \quad x, y \in C_i, \\ &= d_i(x, U_i) + \sum_{k:\ Y_i < Y_k < Y_j} d_k(\rho_k, U_k) + d_j(\rho_j, y) \qquad \text{if} \quad x \in C_i,\ y \in C_j,\ Y_i < Y_j. \end{aligned}$$

Then $\mathcal{B}^{\alpha,\gamma}$ is defined as the completion of $\bigsqcup_{i=1}^{\infty} C_i$ equipped with this distance, with distinguished points $(V_{N_k,k})_{k\geq 1}$. Its root ρ can be obtained as a limit $\rho = \lim_{i\to\infty} \rho_{\sigma_i}$ for any sequence $(\sigma_i)_{i\geq 1}$ for which $Y_{\sigma_i} \to 0$.

Several remarks on the limiting space. First, when $\gamma = 1 - \alpha$ then the limiting spaces $(\mathcal{T}^{\alpha,\gamma}, \mathcal{L}^{\alpha,\gamma})$ are respectively (a constant multiple of) the $\frac{1}{\gamma}$-stable trees and its associated $\frac{1}{\gamma}$-stable looptree, thanks to the convergence results in [43, 42].

Second, remark in the construction of the block $\mathcal{B}^{\alpha,\gamma}$ for $\gamma \neq \alpha$ that we glued circles onto one another along a line structure. This could not have been done in the framework of decorations on the Ulam tree, because in this particular case no two distinct circles are adjacent, they are always separated from each other by a countable number of other circles that accumulate in between, see Figure 4.6.

Looptrees constructed using affine preferential attachment

This last model is very similar to the case of the generalised Rémy's algorithm. We mention it separately because it was the object of a conjecture by Curien, Duquesne, Kortchemski and Manolescu [40], to which we provide a positive answer.

First, for any $\delta > -1$, let us define the model LPAM$^{\delta}$ which produces sequences $(T_n^{\delta})_{n\geq 1}$ of plane trees. Start with T_1^{δ} containing a unique vertex connected to a root by an edge (the root is not considered as a real vertex here). Then, if T_n^{δ} is already constructed, take a vertex in the tree at random (the root does not count) with probability proportional to its degree plus δ, then add an edge connected to a new vertex in a uniformly chosen corner around this vertex. This yields T_{n+1}^{δ}.

Proposition 4.14. *We have the almost sure convergence*

$$n^{-\frac{1}{2+\delta}} \cdot \mathrm{Loop}(T_n^{\delta}) \underset{n\to\infty}{\longrightarrow} \mathcal{L}^{\delta}$$

in the Gromov–Hausdorff–Prokhorov topology. The limiting object can be constructed using an iterative gluing construction with deterministic blocks equal to a circle with unit circumference, using a sequence of weights and scaling factors $(\mathsf{m}_n)_{n\geq 1}$ *obtained as the increments of a Mittag-Leffler Markov chain* $\mathrm{MLMC}\left(\frac{1}{2+\delta}, \frac{1+\delta}{2+\delta}\right)$.

The Hausdorff dimension of $\mathcal{L}^{\delta}$ is $2+\delta$ almost surely using [112, Theorem 1]. The proof of this is really close to the one used for the generalised Rémy algorithm, so we omit it.

4.A Computations

This section is devoted to proving some results that are used in some of our applications.

4.A.1 Number of tables in a Chinese Restaurant Process

Fix two parameters $\alpha \in (0,1)$ and $\theta > -\alpha$. Let us introduce the so-called Chinese restaurant process with parameters (α, θ). We refer to [104] for the definition and properties of this process. Under $\mathbb{P}_{\alpha,\theta}$, the process starts with one table occupied by one customer and then evolves in a Markovian way as follows: given that at stage n there are k occupied tables with n_i customers at table i, a new customer is placed at table i with probability $(n_i - \alpha)/(n+\theta)$ and placed at a new table with probability $(\theta + k\alpha)/(n+\theta)$. Let $N_n(i), i \geq 1$ be the number of customers at table i at stage n. Let also K_n denote the number of occupied tables at stage n. Then

$$\left(\frac{N_n(i), i \geq 1}{n}\right) \underset{n\longrightarrow\infty}{\xrightarrow{\text{a.s. in } \ell^1}} (Y_i, i \geq 1) \quad \text{and} \quad \frac{K_n}{n^{\alpha}} \underset{n\to\infty}{\xrightarrow{\text{a.s.}}} W,$$

where $(Y_i, i \geq 1)$ follows a $\mathrm{GEM}(\alpha,\theta)$-distribution and W a generalized ML (α,θ)-Mittag-Leffler distribution. Let us prove the following

Lemma 4.15. *For every $p \geq 1$, we have*

$$\mathbb{E}\left[\sup_{n\geq 1}\left(\frac{K_n}{n^{\alpha}}\right)^p\right] < \infty. \tag{4.34}$$

Proof. Let $f_{\alpha,\theta}(k) := \frac{\Gamma(\theta/\alpha+k)}{\Gamma(\theta/\alpha+1)\Gamma(k)}$, and $\mathcal{F}_n$ the σ-field generated by the n first steps of the process, then

$$\left(\frac{d\mathbb{P}_{\alpha,\theta}}{d\mathbb{P}_{\alpha,0}}\right)_{|\mathcal{F}_n} = \frac{f_{\alpha,\theta}(K_n)}{f_{1,\theta}(n)} = M_{\alpha,\theta,n},$$

which is a martingale under $\mathbb{P}_{\alpha,0}$ and bounded in L^p, for all $p > 0$. Actually, there exists a constant $c > 1$ such that for any $k, n \geq 1$:

$$\frac{1}{c}\left(\frac{k}{n^{\alpha}}\right)^{\theta} \leq \frac{f_{\alpha,\theta}(k)}{f_{1,\theta}(n)} \leq c\left(\frac{k}{n^{\alpha}}\right)^{\theta}.$$

Introduce $M^*_{\alpha,\theta} := \sup_{n\geq 1} M_{\alpha,\theta,n}$. Using Doob's maximal inequality, we get

$$\mathbb{E}_{\alpha,0}\left[(M^*_{\alpha,\theta})^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}_{\alpha,0}\left[M^p_{\alpha,\theta}\right]$$

Hence

$$\mathbb{E}_{\alpha,\theta}\left[\left(\sup_{n\geq 1}\frac{K_n}{n^\alpha}\right)^p\right] \leq c\mathbb{E}_{\alpha,\theta}\left[(M^*_{\alpha,1})^p\right] = \mathbb{E}_{\alpha,0}\left[(M^*_{\alpha,1})^p\cdot M_{\alpha,\theta}\right]$$
$$\leq \sqrt{\mathbb{E}_{\alpha,0}\left[\left(M^*_{\alpha,1}\right)^{2p}\right]\mathbb{E}_{\alpha,0}\left[M^2_{\alpha,\theta}\right]} < \infty.$$

which finishes the proof of the lemma. □

4.A.2 The supremum of the normalised height in Rémy's algorithm

Let $(T_n)_{n\geq 1}$ be a sequence of trees evolving using Rémy's algorithm. This sequence is a Markov chain in a state space of binary planted trees. Let us denote $H := \sup_{n\geq 1}(n^{-1/2}\operatorname{ht}(T_n))$. We prove the following:

Lemma 4.16. *There exists constants C_1 and C_2 such that for all $x>0$,*

$$\mathbb{P}(H > x) \leq C_1\exp\bigl(-C_2x^2\bigr).$$

In particular, H admits moments of all orders.

Proof. Let $\tau_x := \inf\left\{n\geq 1 \;\middle|\; \operatorname{ht}(T_n)\geq xn^{1/2}\right\}$. We write

$$\begin{aligned}\mathbb{P}(H>x) &= \mathbb{P}(\tau_x<+\infty)\\ &\leq \mathbb{P}\left(\lim_{n\to\infty} n^{-1/2}\operatorname{ht}(T_n)\geq \frac{x}{2}\right) + \mathbb{P}\left(\tau_x<+\infty, \lim_{n\to\infty} n^{-1/2}\operatorname{ht}(T_n)\geq\frac{x}{2}\right)\end{aligned}$$

We know thanks to [42] that the trees constructed using Rémy's algorithm converge almost surely in the Gromov-Hausdorff topology to Aldous' Brownian tree, so $n^{-1/2}\cdot T_n\to\mathcal{T}$ as $n\to\infty$. By continuity we have $\lim_{n\to\infty} n^{-1/2}\operatorname{ht}(T_n) = \operatorname{ht}(\mathcal{T})$. Some estimates on the height of the Brownian tree (expressed for the maximum of a Brownian excursion in [86]) show that the first term of the above sum is smaller than $C_1\exp\bigl(-C_2x^2\bigr)$, for some choice of constants C_1 and C_2. Fix some $N_0\geq 1$ that we will choose later. Then, compute

$$\begin{aligned}\mathbb{P}\left(\tau_x<+\infty, \lim_{n\to\infty} n^{-1/2}\operatorname{ht}(T_n)\geq\frac{x}{2}\right) &= \sum_{N=1}^{+\infty}\mathbb{P}(\tau_x=N)\,\mathbb{P}\left(\operatorname{ht}(\mathcal{T})\leq\frac{x}{2}\;\middle|\;\tau_x=N\right)\\ &\leq \sum_{N=1}^{N_0}\mathbb{P}(\tau_x=N) + \sup_{N\geq N_0}\mathbb{P}\left(\operatorname{ht}(\mathcal{T})\leq\frac{x}{2}\;\middle|\;\tau_x=N\right)\\ &\leq N_0\cdot C_1\exp\bigl(-C_2x^2\bigr) + \sup_{N\geq N_0}\mathbb{P}\left(\operatorname{ht}(\mathcal{T})\leq\frac{x}{2}\;\middle|\;\tau_x=N\right),\end{aligned}$$

where in the last inequality we use the fact that for all N,

$$\mathbb{P}(\tau_x=N)\leq\mathbb{P}\left(\operatorname{ht}(T_N)\geq xN^{1/2}\right)\leq C_1\exp\bigl(-C_2x^2\bigr),$$

using the results of Addario-Berry [2].

Now, let us reason conditionally on the event $\{\tau_x=N\}$. On that event, in the tree T_N there exists at least one path of length $\lfloor xN^{1/2}\rfloor$ starting from the root ending at a vertex v. The height $H_n(v)$ of the vertex v at time n evolves under the same dynamics as the number of balls in a triangular urn model with replacement matrix $\begin{pmatrix}1&1\\0&2\end{pmatrix}$, and starting proportion

$(\lfloor xN^{1/2} \rfloor, 2N+1-\lfloor xN^{1/2} \rfloor)$, see [80] for definition and results for those urns. Using a theorem of [80], as $n \to \infty$, we have the almost sure convergence

$$\frac{H_n(v)}{n^{1/2}} \longrightarrow W_N^x,$$

where $W_N^x = \beta_N^x \cdot M_N$ is the product of two independent variables, with

$$\beta_N^x \sim \mathrm{Beta}(\lfloor xN^{1/2} \rfloor, 2N+1-\lfloor xN^{1/2} \rfloor) \qquad \text{and} \qquad M_N \sim \mathrm{ML}\left(\frac{1}{2}, \frac{2N+1}{2}\right). \tag{4.35}$$

Then we write

$$\begin{aligned}
\mathbb{P}\left(W_N^x \leq \frac{x}{2}\right) &= \mathbb{P}\left(\beta_N^x \cdot M_N \leq \frac{x}{2}\right) \\
&\leq \mathbb{P}\left(\beta_N^x \cdot M_N \leq \frac{x}{2},\ M_N \geq \frac{3}{2}N^{1/2}\right) + \mathbb{P}\left(M_N \leq \frac{3}{2}N^{1/2}\right) \\
&\leq \mathbb{P}\left(\beta_N^x \leq \frac{1}{3} \cdot x \cdot N^{-1/2}\right) + \mathbb{P}\left(M_N \leq \frac{3}{2}N^{1/2}\right).
\end{aligned}$$

We bound the two terms in the last sum using the Chebychev inequality, using

$$\mathbb{E}\left[\beta_N^x\right] = \frac{\lfloor xN^{1/2} \rfloor}{2N+1} \sim \frac{x}{2}N^{-1/2}, \qquad \mathbb{V}\left(\beta_N^x\right) = \frac{\lfloor xN^{1/2} \rfloor (2N+1-\lfloor xN^{1/2} \rfloor)}{(2N+1)^2(2N+2)} \sim \frac{x}{4}N^{-3/2}.$$

We also have

$$\mathbb{E}\left[M_N\right] = \frac{\Gamma\left(N+1/2\right)\Gamma\left(2N+2\right)}{\Gamma\left(2N+1\right)\Gamma\left(N+1\right)} \sim 2N^{1/2},$$

and

$$\mathbb{V}\left(M_N\right) = \frac{\Gamma\left(N+1/2\right)\Gamma\left(2N+3\right)}{\Gamma\left(2N+1\right)\Gamma\left(N+3/2\right)} - \left(\frac{\Gamma\left(N+1/2\right)\Gamma\left(2N+2\right)}{\Gamma\left(2N+1\right)\Gamma\left(N+1\right)}\right)^2 \leq 1.$$

Hence,

$$\begin{aligned}
\mathbb{P}\left(\beta_N^x \leq \frac{1}{3} \cdot x \cdot N^{-1/2}\right) &\leq \mathbb{P}\left(\left|\beta_N^x - \mathbb{E}\left[\beta_N^x\right]\right| \geq \frac{1}{8} \cdot x \cdot N^{-1/2}\right) \\
&\leq 8x^2 N \mathbb{V}\left(\beta_N^x\right) \\
&\leq Cx^3 N^{-1/2},
\end{aligned}$$

with C a constant that is independent of x and N. We also have

$$\begin{aligned}
\mathbb{P}\left(M_N \leq \frac{3}{2}N^{1/2}\right) &\leq \mathbb{P}\left(\left|M_N - \mathbb{E}\left[M_N\right]\right| \geq \frac{1}{3}N^{1/2}\right) \\
&\leq 3N^{-1}\mathbb{V}\left(M_N\right) \\
&\leq C'N^{-1},
\end{aligned}$$

with C' another constant that does not depend on x or N. All this analysis was done conditionally on the event $\{\tau_x = N\}$, so in fact, we have for all $N \geq N_0$,

$$\mathbb{P}\left(\mathrm{ht}(\mathcal{T}) \leq \frac{x}{2} \,\middle|\, \tau_x = N\right) \leq Cx^3 N^{-1/2} + C'N^{-1} \leq Cx^3 N_0^{-1/2} + C'N_0^{-1}.$$

Now we just take $N_0 = \exp\left(\frac{C_2}{2}x^2\right)$ and the result follows. $\square$

Chapter 5

Stable graphs: distributions and line-breaking construction

This chapter is adapted from the preprint [64], which is joint work with Christina Goldschmidt and Bénédicte Haas.

For $\alpha \in (1, 2]$, the α-stable graph arises as the universal scaling limit of critical random graphs with i.i.d. degrees having a given α-dependent power-law tail behavior. It consists of a sequence of compact measured metric spaces (the limiting connected components), each of which is tree-like, in the sense that it consists of an $\mathbb{R}$-tree with finitely many vertex-identifications (which create cycles). Indeed, given their masses and numbers of vertex-identifications, these components are independent and may be constructed from a spanning $\mathbb{R}$-tree, which is a biased version of the α-stable tree, with a certain number of leaves glued along their paths to the root. In this chapter we investigate the geometric properties of such a component with given mass and number of vertex-identifications. We (1) obtain the distribution of its kernel and more generally of its discrete finite-dimensional marginals; we will observe that these distributions are related to the distributions of some configuration models (2) determine the distribution of the α-stable graph as a collection of α-stable trees glued onto its kernel and (3) present a line-breaking construction, in the same spirit as Aldous' line-breaking construction of the Brownian continuum random tree.

Contents

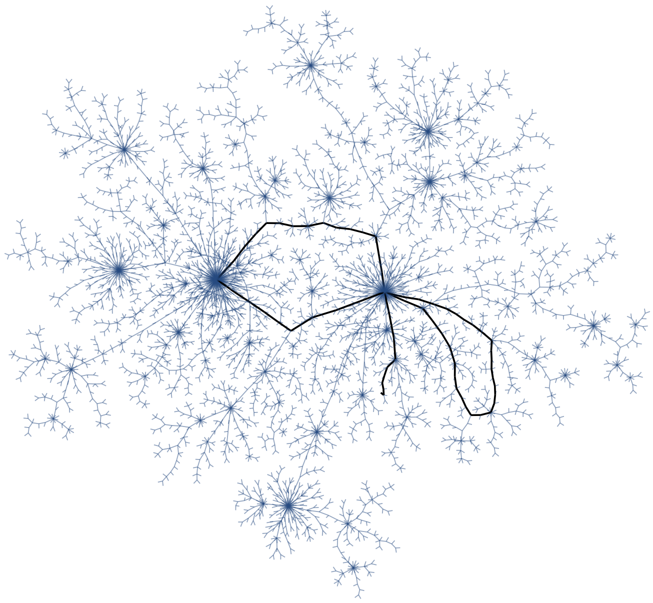


Figure 5.1 – A simulation of a connected component of the stable graph when $\alpha = 1.5$ and the surplus is 2. The cycle structure is shown in black.

5.1 Introduction and main results

5.1.1 Motivation

The purpose of this chapter is to understand the distributional properties of the scaling limit of a critical random graph with independent and identically distributed degrees having certain power-law tail behaviour. Let us first describe the random graph model precisely. Let $D_1, D_2, \ldots, D_n \in \mathbb{N}$ be independent and identically distributed random variables such that $\mathbb{E}D_1^2 < \infty$. We build a graph with vertices labelled by $1, 2, \ldots, n$. For $i = 1, \ldots, n-1$, let vertex i have degree D_i. If $\sum_{i=1}^n D_i$ is even, let vertex n have degree D_n; otherwise, let vertex n have degree $D_n + 1$. Now pick a simple graph G_n uniformly at random from among those with these given vertex degrees (at least one such graph exists with probability tending to 1 as $n \to \infty$).

Molloy and Reed [97] showed that there is a phase transition in the sizes of the connected components: if the parameter $\nu := \mathbb{E}[D_1(D_1-1)]/\mathbb{E}[D_1]$ is larger than 1 there exists a unique *giant component* of size proportional to n, while if ν is smaller than or equal to 1 there is no giant component. We will here tune the degree distribution so as to be exactly at the point of the phase transition, i.e. $\nu = 1$. The component-size behaviour is here at its most delicate: even after performing the correct rescaling and taking a limit, there is residual randomness in the sequence of component sizes. For the questions in which we are interested, the critical case with $\mathbb{E}D_1^3 < \infty$ has already been thoroughly investigated in previous work, which we summarise in Section 5.1.3. So we will rather assume that the degree distribution has infinite third moment and a specific power-law behaviour. Henceforth, fix $1 < \alpha < 2$ and assume that

$$\nu = 1 \quad \text{and} \quad \mathbb{P}(D_1 = k) \sim ck^{-2-\alpha} \quad \text{as } k \to \infty, \tag{5.1}$$

where $c > 0$ is constant. (Note that $\nu = 1$ is equivalent to $\mathbb{E}D_1^2 = 2\mathbb{E}D_1$.)

The analogous model of a random *tree* is a Galton-Watson tree with critical offspring distribution in the domain of attraction of an α-stable law. In that case, there is a well-known scaling limit, the α-stable tree [52]. We will explore the relationship between these two models, at the level of scaling limits, in the sequel.

It is now standard to formulate random graph scaling limits in terms of sequences of *measured metric spaces*, namely metric spaces endowed with a measure. Throughout this chapter we let $(\mathscr{C}, \mathrm{d}_{\mathrm{GHP}})$ denote the set of measured isometry-equivalence classes of compact measured metric spaces equipped with the Gromov-Hausdorff-Prokhorov topology (see, for example, Section 2.1 of [6] for the formulation we use here) and endow it with the associated Borel σ-algebra. (We will often elide the difference between a measured metric space and its equivalence class but it should be understood that we are really thinking about the equivalence class.) As we are dealing with graphs which have many components, we need a topology on sequences of (equivalence classes of) compact measured metric spaces. Let $\mathbf{Z}$ be an infinite sequence of "zero" measured metric spaces, each consisting of a single point endowed with measure 0. Consider a pair $\mathbf{M} = (M_i, d_i, \mu_i)_{i\geq 1}$ and $\mathbf{M}' = (M_i', d_i', \mu_i)_{i\geq 1}$ of sequences of compact measured metric spaces. For $p \geq 1$ define

$$\mathrm{dist}_p(\mathbf{M}, \mathbf{M}') = \left(\sum_{i=1}^{\infty} \mathrm{d}_{\mathrm{GHP}}\Big((M_i, d_i, \mu_i), (M_i', d_i', \mu_i') \Big)^p \right)^{1/p},$$

and let

$$\mathscr{L}_p = \Big\{ \mathbf{M} \in \mathscr{C}^{\mathbb{N}} : \mathrm{dist}_p(\mathbf{M}, \mathbf{Z}) < \infty \Big\}.$$

Then $(\mathscr{L}_p, \mathrm{dist}_p)$ is a Polish space [6].

Write $C_1^n, C_2^n, \ldots$ for the vertex-sets of the components of the graph G_n, listed in decreasing order of size (with ties broken arbitrarily). Set

$$A_\alpha = \left(\frac{c\Gamma(2-\alpha)}{\alpha(\alpha-1)} \right)^{1/(\alpha+1)}. \tag{5.2}$$

We think of the components as metric spaces by endowing each one with a scaled version of the usual graph distance, d_{gr}: let

$$d_i^n := \frac{A_\alpha^2}{\mathbb{E}[D_1] n^{(\alpha-1)/(\alpha+1)}} \mathrm{d}_{\mathrm{gr}}$$

be the distance in C_i^n. We also endow each of them with the scaled counting measure

$$\mu_i^n := \frac{A_\alpha}{\mathbb{E}[D_1] n^{\alpha/(\alpha+1)}} \sum_{v \in C_i^n} \delta_v.$$

Let $\mathrm{C}_i^n = (C_i^n, d_i^n, \mu_i^n)$ be the resulting measured metric space. We write $s(\mathrm{C}_i^n)$ for the number of *surplus edges* (i.e. edges more than a tree) possessed by the component C_i^n. Formally, for a connected graph $G = (V, E)$, the number of surplus edges or, more briefly, *surplus*, is defined to be

$$s(G) = |E| - |V| + 1.$$

The following theorem is proved in [38].

Theorem 5.1. *As $n \to \infty$,*

$$(\mathrm{C}_1^n, \mathrm{C}_2^n, \ldots) \xrightarrow{\mathrm{d}} (\mathrm{C}_1, \mathrm{C}_2, \ldots),$$

in $(\mathscr{L}_{2\alpha/(\alpha-1)}, \mathrm{dist}_{2\alpha/(\alpha-1)})$, for some random sequence $(\mathrm{C}_1, \mathrm{C}_2, \ldots)$ which we call the α-

stable graph.

(In Section 5.1.3 below we will describe the relationship of this theorem to earlier work.)

Theorem 5.1 also holds in the setting of a random multigraph (i.e. it may contain self-loops and multiple edges) sampled from the *configuration model* with i.i.d. degrees. Formally, a *multigraph* G is an ordered pair $G = (V, E)$ where V is the set of vertices and E the *multiset* of edges (i.e. elements of $\{\{u, v\},\ u \in V,\ v \in V\}$). Let $\mathrm{supp}(E)$ denote the support of E, i.e. the underlying set of distinct elements of E, and, for $e \in \mathrm{supp}(E)$, let $\mathrm{mult}(e)$ denote its multiplicity. Let $\mathrm{sl}(G)$ denote the cardinality of the multiset of self-loops. For a vertex $v \in V$, we write $\deg(v)$ for its degree, or $\deg_G(v)$ if there is potential ambiguity over which graph we are looking at. The surplus is still defined to be $s(G) = |E| - |V| + 1$, where we emphasise that $|E| = \sum_{e \in \mathrm{supp}(E)} \mathrm{mult}(e)$. Let us briefly explain the set-up of the configuration model for deterministic degrees $d_1, d_2, \ldots, d_n$ with even sum. (The configuration model was introduced in varying degrees of generality in [16, 29, 115]. We refer to Chapter 7 of the recent book of van der Hofstad [75] for the proofs of the claims made in this paragraph.) To vertex i we assign d_i half-edges, for $1 \leq i \leq n$. We give the half-edges an arbitrary labelling (so that we may distinguish them) and then choose a matching of the half-edges uniformly at random. Two matched half-edges form an edge of the resulting structure, which is a multigraph. Then for a particular multigraph G with degrees $d_1, d_2, \ldots, d_n$, the probability that the configuration model generates G is

$$\frac{\prod_{i=1}^n d_i!}{(\sum_{i=1}^n d_i - 1)!!\ 2^{\mathrm{sl}(G)} \prod_{e \in \mathrm{supp}(E)} \mathrm{mult}(e)!}, \tag{5.3}$$

where $a!!$ denotes the double factorial of a. From this expression, it is easy to see that if there exists at least one simple graph with degrees $d_1, d_2, \ldots, d_n$ then conditioning the multigraph to be simple yields a *uniform* graph with the given degree sequence. We are interested in the setting where the degrees are random variables $D_1, D_2, \ldots, D_n$ satisfying the conditions (5.1) (with the small modification mentioned above to make the sum of the degrees even). In this case, there exists a simple graph with these degrees with probability tending to 1 as $n \to \infty$, which enables us to convert results for the configuration model into results for the uniform random graph with given degree sequence; in the setting of Theorem 5.1 the conditioning turns out not to affect the result.

The α-stable graph is constructed using a spectrally positive α-stable Lévy process; we give the details, which are somewhat involved, in Section 5.2.2. For $i \geq 1$, write $\mathrm{C}_i = (C_i, d_{C_i}, \mu_{C_i})$, $i \geq 1$. These measured metric spaces are $\mathbb{R}$-*graphs* in the sense of [6] i.e. they are locally $\mathbb{R}$-trees, but may also possess cycles. It is possible to make sense of the surpluses of the limiting components, for which we write $s(\mathrm{C}_i)$, $i \geq 1$. It is a consequence of Theorem 5.1 that

$$\frac{A_\alpha}{\mathbb{E}[D_1] n^{\alpha/(\alpha+1)}} (|C_1^n|, |C_2^n|, \ldots) \xrightarrow{\mathrm{d}} (\mu_{C_1}(C_1), \mu_{C_2}(C_2), \ldots) \tag{5.4}$$

in $\ell^{2\alpha/(\alpha-1)}$, jointly with the convergence in the sense of the product topology

$$(s(\mathrm{C}_1^n), s(\mathrm{C}_2^n), \ldots) \xrightarrow{\mathrm{d}} (s(\mathrm{C}_1), s(\mathrm{C}_2), \ldots) \tag{5.5}$$

for the sequences of surplus edges. The joint law of $(\mu_{C_1}(C_1), \mu_{C_2}(C_2), \ldots)$ and $(s(\mathrm{C}_1), s(\mathrm{C}_2), \ldots)$ is explicit in terms of the underlying α-stable Lévy process; see Section 5.2.2. Moreover the limiting components $(\mathrm{C}_1, \mathrm{C}_2, \ldots)$ are conditionally independent given $(\mu_{C_1}(C_1), \mu_{C_2}(C_2), \ldots)$ and $(s(\mathrm{C}_1), s(\mathrm{C}_2), \ldots)$, with distributions coming from a collection of fundamental building-blocks:

there exist random measured metric spaces $(\mathcal{G}^s, d^s, \mu^s)$, $s \geq 0$, where μ^s is a probability measure, such that, for all i, given $\mu_{C_i}(C_i)$ and $s(\mathrm{C_i})$, we have

$$\left(C_i, d_{C_i}, \mu_{C_i}\right) \overset{\text{(d)}}{=} \left(\mathcal{G}^{s(\mathrm{C}_i)}, \mu_{C_i}(C_i)^{1-1/\alpha} \cdot d^{s(\mathrm{C}_i)}, \mu_{C_i}(C_i) \cdot \mu^{s(\mathrm{C}_i)}\right).$$

For $s = 0$, $(\mathcal{G}^s, d^s, \mu^s)$ is simply the standard rooted α-stable tree, the definition of which is recalled in Section 5.2.1. Informally, for $s \geq 1$, $(\mathcal{G}^s, d^s, \mu^s)$, is constructed by randomly choosing s leaves in an s-biased version of this α-stable tree, and then gluing them to randomly-chosen branch-points along their paths to the root, with probabilities proportional to the "local time to the right" of the branch-points. (We will define these quantities in the sequel.) We will often think of the resulting $\mathbb{R}$-graph $\mathcal{G}^s$ as being rooted; in this case, the root is simply inherited from that of the s-biased α-stable tree. The measure μ^s on $\mathcal{G}^s$ is then the probability measure inherited from the s-biased α-stable tree. We will often abuse notation and simply write $\mathcal{G}^s$ in place of $(\mathcal{G}^s, d^s, \mu^s)$. For $a > 0$, we will also write $a \cdot \mathcal{G}^s$ to denote the same measured metric space with all distances scaled by a, i.e. $(\mathcal{G}^s, ad^s, \mu^s)$.

In order to understand the geometric properties of the α-stable graph, it therefore suffices to consider the measured metric spaces

$$\mathcal{G}^s, s \geq 0.$$

We will call $\mathcal{G}^s$ the *connected α-stable graph with surplus s*. Let us note immediately that $\mathcal{G}^s$ naturally inherits the Hausdorff dimension of the α-stable tree and that, therefore,

$$\dim_{\mathrm{H}}(\mathcal{G}^s) = \frac{\alpha}{\alpha - 1} \quad \text{a.s.}$$

Like a connected combinatorial graph, the $\mathbb{R}$-graph $\mathcal{G}^s$ may be viewed as a cycle structure to which pendant subtrees are attached. Let $\mathcal{K}^s$ be the image after the gluing procedure of the subtree spanned by the s selected leaves and the root of the s-biased version of the α-stable tree. (When $s = 0$, we use the convention that $\mathcal{K}^s$ is the empty set.) The space $\mathcal{K}^s$ encodes the rooted cycle structure of $\mathcal{G}^s$. We refer to it as the *continuous kernel* because it is a continuous analogue of the usual graph-theoretic notion of a kernel (except that it is rooted at a vertex of degree 1). We will think of it as a rooted multigraph which is endowed with real-valued edge-lengths, and write K^s for the rooted multigraph *without* the edge-lengths, which we call the *discrete kernel*.

In order to better understand the structure of the $\mathbb{R}$-graph $\mathcal{G}^s$, we will approximate it by a sequence $(\mathcal{G}^s_n)_{n \geq 0}$ of multigraphs with edge-lengths, starting from the continuous kernel, $\mathcal{G}^s_0 = \mathcal{K}^s$. Consider an infinite sample of leaves from $\mathcal{G}^s$, labelled $1, 2, \ldots$. For each $n \in \mathbb{N}$, let $\mathcal{G}^s_n$ be the connected subgraph of $\mathcal{G}^s$ consisting of the union of the kernel $\mathcal{K}^s$ and the paths from the n first leaves to the root. These are the $\mathbb{R}$-graph analogues of Aldous' *random finite-dimensional marginals* for a continuum random tree. For brevity, we will call them the *marginals* of $\mathcal{G}^s$. In Lemma 5.24 below, we note that $\mathcal{G}^s$ can be recovered as the completion of $\cup_{n \geq 0} \mathcal{G}^s_n$. We will also make extensive use of the discrete counterparts of the $\mathcal{G}^s_n$. For $n \geq 0$, let G^s_n be the *combinatorial shape* of $\mathcal{G}^s_n$ (i.e. "forget the edge-lengths", so as to obtain a finite graph with surplus s and no vertices of degree 2 – see (5.14) for a formal definition in the framework of trees that adapts immediately to our graphs), so that $\mathsf{K}^s = \mathsf{G}^s_0$. Note that the root vertex has degree 1 in all of these graphs. When $s \geq 2$, we can *erase the root* in the discrete kernel (formally, we remove the root and the adjacent edge, and if this creates a vertex of degree 2 we erase it) to obtain a multigraph that we denote by G^s_{-1}.

5.1.2 Main results

Throughout this section, we fix the surplus $s \in \mathbb{Z}_+$.

Our first main results characterise the joint distributions of the discrete marginals $(\mathsf{G}_n^s)_{n\geq 0}$. This family of random multigraphs has particularly attractive properties: for fixed n, the graph G_n^s has the distribution of a certain conditioned configuration model with i.i.d. random degrees, with a particular canonical degree distribution. Moreover, as a process, $(\mathsf{G}_n^s)_{n\geq 0}$ evolves in a Markovian manner according to a simple recursive construction which is a version of Marchal's algorithm [92] for building the marginals of the stable tree, $(\mathsf{G}_n^0)_{n\geq 0}$. Although $\mathcal{G}^s$ is constructed from a biased version of the α-stable tree, we emphasise that it was not at all obvious to us *a priori* that Marchal's algorithm would generalise in this way.

An advantage of this recursive construction is that it has many urn models embedded in it, which enable us to easily get at different aspects of $\mathcal{G}^s$. We provide two different constructions of $\mathcal{G}^s$, which rely on relatively simple random building blocks. The distributions of these building blocks (Beta, generalised Mittag-Leffler, Dirichlet and Poisson-Dirichlet) are defined in Section 5.5, where we also recall various of their standard properties and discuss their relationships to urns. Our two constructions are as follows:

(i) The first takes a collection of i.i.d. α-stable trees which are randomly scaled and then glued onto K^s in such a way that each edge of K^s is replaced by a tree with two marked points, and such that every vertex of K^s acquires a (countable) collection of pendant subtrees.

(ii) The second starts by replacing the edges of the kernel by line-segments of lengths with a given joint distribution, and then proceeds by recursively gluing a countable sequence of segments of random lengths onto the structure. We call this a *line-breaking construction* and obtain the limit space in the end by completion.

These constructions generalise, in a natural way, the distributional properties and line-breaking construction proved in [4] for the components of the *Brownian graph*, a term we coin here to mean the common scaling limit of the critical Erdős-Rényi random graph [5] and the critical random graph with i.i.d. degrees having a finite third moment [24] as well as various other models (see Section 5.1.3). We emphasise, however, that the proofs in the stable setting are much harder, essentially due to the added complication of dealing with Lévy processes rather than just Brownian motion. Our line-breaking construction is the graph counterpart of the line-breaking construction of the stable trees given in [63].

The discrete marginals of $\mathcal{G}^s$

We can recover the measured metric space $\mathcal{G}^s$ from the discrete marginals G_n^s by equipping them with the graph distance and the uniform distribution on their leaves, as follows.

Proposition 5.2.

$$\frac{\mathsf{G}_n^s}{n^{1-1/\alpha}} \xrightarrow[n\to\infty]{\text{a.s.}} \alpha \cdot \mathcal{G}^s$$

for the Gromov-Hausdorff-Prokhorov topology.

This generalises a result which says that the α-stable tree is the (almost sure) scaling limit of its discrete marginals, see [92, 42]. See Section 5.4.1.

For any multigraph $G = (V, E)$, recall that we let $\mathrm{sl}(G)$ denote its number of self-loops, and for an element $e \in \mathrm{supp}(E)$, we let $\mathrm{mult}(e)$ denote its multiplicity. Let $I(G) \subset V$ denote the set of internal vertices of G. We say that a permutation τ of the set $I(G)$ is a *symmetry* of G if, after having extended τ to the identity function on the leaves, τ preserves the adjacency relations

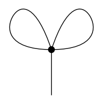
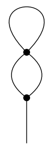
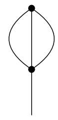
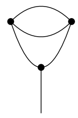
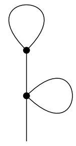
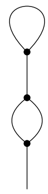
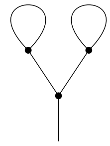
Graph $G \in \mathbb{M}_{2,0}$							
$\mathrm{sl}(G)$	2	1	0	0	2	1	2
$\prod_{v \in I(G)} w_{\deg(v)-1}$ $(\alpha = 5/4)$	$\frac{21}{64}$	$\frac{3}{64}$	$\frac{3}{64}$	$\frac{1}{64}$	$\frac{3}{64}$	$\frac{1}{64}$	$\frac{1}{64}$
$\prod_{e \in E(G)} \mathrm{mult}(e)!$	2	2	6	2	1	2	1
$\lvert\mathrm{Sym}(G)\rvert$	1	1	1	2	1	1	2
$\mathbb{P}\left(\mathsf{K}^2 = G\right)$ $(\alpha = 5/4)$	$\frac{1}{2}$	$\frac{1}{7}$	$\frac{2}{21}$	$\frac{1}{21}$	$\frac{1}{7}$	$\frac{1}{21}$	$\frac{1}{42}$
$\mathbb{P}\left(\mathsf{K}^2_{\mathrm{Br}} = G\right)$	0	0	0	$\frac{2}{5}$	0	$\frac{2}{5}$	$\frac{1}{5}$

Figure 5.2 – The possible kernels for $s = 2$ with their probabilities for $\alpha = 5/4$ (given in the penultimate line). For comparison, the last line gives the distribution of the kernel of the connected Brownian graph with surplus 2.

in the graph and for all $u, v \in V$, the edges $\{u, v\}$ and $\{\tau(u), \tau(v)\}$ have the same multiplicity. We let $\mathrm{Sym}(G)$ denote the set of symmetries of G. For $n \geq 0$, let $\mathbb{M}_{s,n}$ be the set of connected multigraphs with $n + 1$ labelled leaves, surplus s and no vertices of degree 2. (Observe that the internal vertices are not labelled.) When $s \geq 2$, let $\mathbb{M}_{s,-1}$ be the set of unlabelled connected multigraphs with surplus s and minimum degree at least 3. Finally, let us define a sequence of weights by

$$w_0 := 1, \quad w_1 := 0, \quad w_2 := \alpha - 1, \quad w_k := (k - 1 - \alpha) \dots (2 - \alpha)(\alpha - 1), \quad \text{for } k \geq 3. \tag{5.6}$$

Viewing the root as a leaf with label 0, we note that G^s_n is an element of $\mathbb{M}_{s,n}$. We can now describe the distributions of the random multigraphs G^s_n.

Theorem 5.3. *Let $n \geq 0$. For every connected multigraph $G = (V, E) \in \mathbb{M}_{s,n}$,*

$$\mathbb{P}\left(\mathsf{G}^s_n = G\right) \propto \frac{\prod_{v \in I(G)} w_{\deg(v)-1}}{\lvert\mathrm{Sym}(G)\rvert 2^{\mathrm{sl}(G)} \prod_{e \in \mathrm{supp}(E)} \mathrm{mult}(e)!}.$$

This, in particular, gives the distribution of the kernel K^s when $n = 0$. When $s \geq 2$, this expression also gives the distribution of G^s_{-1} on $\mathbb{M}_{s,-1}$.

This result is proved in Section 5.3. To illustrate it, in Figure 5.2 we give the distribution of the kernel explicitly in the case $s = 2$ and $\alpha = \frac{5}{4}$.

Comparing the form of the distribution of G^s_n with (5.3) suggests a connection with a conditioned configuration model. To make this precise, let $D^{(\alpha)}$ be a random variable on $\mathbb{N}$ with distribution

$$\mathbb{P}(D^{(\alpha)} = k) = \frac{2(1+\alpha)\alpha}{\alpha^2 + \alpha + 2} \cdot \frac{w_{k-1}}{k!}, \quad k \geq 2, \quad \text{and} \quad \mathbb{P}(D^{(\alpha)} = 1) = \frac{2(1+\alpha)}{\alpha^2 + \alpha + 2}. \tag{5.7}$$

Observe that $\mathbb{P}(D^{(\alpha)} = 2) = 0$. We will verify in Section 5.3.6 that this indeed defines a probability measure which, moreover, satisfies the conditions (5.1). Consider now the following particular instance of the configuration model. We fix $n \geq 0$ and $m \geq n+1$ (include the case $n = -1$ if $s \geq 2$), take vertices labelled $0, 1, \ldots, m-1$ to have i.i.d. degrees distributed according to $D^{(\alpha)}$ and write $\mathsf{C}^s_{n,m}$ for the resulting configuration multigraph conditioned to be in $\mathbb{M}_{n,s}$, after having forgotten the labels $n+1, n+2, \ldots, m-1$.

Corollary 5.4. *The random multigraph G^s_n conditioned to have m vertices has the same law as $\mathsf{C}^s_{n,m}$.*

This again generalises the analogous result for the α-stable tree: the combinatorial shape of the subtree obtained by sampling $n \geq 0$ leaves and the root is distributed as a planted (i.e. with a root of degree 1) non-ordered version of a Galton-Watson tree conditioned to have n leaves, whose offspring distribution η_α has probability generating function $z + \alpha^{-1}(1-z)^\alpha$. There is, of course, a connection between $D^{(\alpha)}$ and η_α: if we let $\hat{D}^{(\alpha)}$ denote the size-biased version

$$\mathbb{P}(\hat{D}^{(\alpha)} = k) := \frac{k\mathbb{P}(D^{(\alpha)} = k)}{\mathbb{E}\big[D^{(\alpha)}\big]}, \quad k \geq 1,$$

then $\hat{D}^{(\alpha)} - 1$ is distributed as η_α. See Section 5.3.6.

In fact, we may think of the configuration multigraph with i.i.d. degrees distributed as $D^{(\alpha)}$ as, in some sense, the *canonical* model in the universality class of the stable graph. For this model, the law of a component conditioned to have $n+1$ leaves and surplus s is exactly the same as the corresponding discrete marginal for its scaling limit, and there exists a coupling for different n which is such that we get almost sure (rather than just distributional) convergence, on rescaling, to the connected α-stable graph with surplus s.

We are also able to understand the joint distribution of the graphs $\mathsf{G}^s_n, n \geq 0$ (again, include the case $n = -1$ when $s \geq 2$): they evolve according to a multigraph version of Marchal's algorithm [92] for the discrete marginals of a α-stable tree. Let us define a step in the algorithm. Take a multigraph $G = (V, E) \in \mathbb{M}_{s,n}$. Declare every edge to have weight $\alpha - 1$, every internal vertex $u \in I(G)$ to have weight $\deg_G(u) - 1 - \alpha$ and every leaf to have weight 0. Then the total weight of G is

$$\sum_{u \in I(G)} (\deg_G(u) - 1 - \alpha) + (\alpha - 1) \cdot |E| = \alpha(s+n) + s - 1, \tag{5.8}$$

which depends only on the surplus and the number of leaves of the graph. We use the term *edge-leaf* to mean an edge with a leaf at one of its end-points. Choose an edge/vertex with probability proportional to its weight. Then

- if it is a vertex, attach a new edge-leaf where the leaf has label $n+1$ to this vertex,
- if it is an edge, attach a new edge-leaf where the leaf has label $n+1$ to a newly created vertex which splits the edge into two.

We say that a sequence of graphs evolves according to Marchal's algorithm if it is Markovian and the transitions are given by one step of Marchal's algorithm.

Theorem 5.5. *For $s \geq 0$, the sequence $(\mathsf{G}^s_n)_{n \geq 0}$ evolves according to Marchal's algorithm. For $s \geq 2$, more generally, the sequence $(\mathsf{G}^s_n)_{n \geq -1}$ evolves according to Marchal's algorithm.*

See Section 5.3.4 for a proof. We now turn to our constructions of the limit object $\mathcal{G}^s$.

Construction 1: from randomly scaled stable trees glued to the kernel

Given a connected multigraph $G \in \mathbb{M}_{s,0}$, with k edges and $k-s$ internal vertices having degrees $d_1, \dots, d_{k-s}$, consider independent random variables

$$(M_1, \dots, M_{2k-s}) \sim \mathrm{Dir}\Bigg(\underbrace{\frac{\alpha-1}{\alpha}, \dots, \frac{\alpha-1}{\alpha}}_{k}, \frac{d_1-1-\alpha}{\alpha}, \dots, \frac{d_{k-s}-1-\alpha}{\alpha}\Bigg) \tag{5.9}$$

and, for $1 \leq i \leq k-s$,

$$(\Delta_{i,j}, j \geq 1) \sim \mathrm{PD}\left(\frac{1}{\alpha}, \frac{d_i - 1 - \alpha}{\alpha}\right), \tag{5.10}$$

where $\mathrm{Dir}(a_1, \dots, a_n)$ denotes the Dirichlet distribution on the $(n-1)$-dimensional simplex, with parameters $a_1 > 0, a_2 > 0, \dots, a_n > 0$, and $\mathrm{PD}(a,b)$ denotes the Poisson-Dirichlet distribution on the set of positive decreasing sequences with sum 1, with parameters $a > 0, b > 0$.

Given all of these random variables, consider independent α-stable trees $\mathcal{T}_\ell$, $\mathcal{T}_{i,j}$, where $\mathcal{T}_\ell$ has mass M_ℓ and $\mathcal{T}_{i,j}$ has mass $M_{i+k} \cdot \Delta_{i,j}$, with $1 \leq \ell \leq k, 1 \leq i \leq k-s, j \geq 1$. For each ℓ let ρ_ℓ denote the root of $\mathcal{T}_\ell$ and L_ℓ be a uniform leaf. Similarly, let $\rho_{i,j}$ denote the root of the tree $\mathcal{T}_{i,j}$ for each i, j. Then denote by $e_1, \dots, e_k$ the edges of G in arbitrary order, with, say, $e_i = \{x_i, y_i\}$, and by $v_1, \dots, v_{k-s}$ the internal vertices of G, also in arbitrary order. Finally, let $\mathcal{G}(G)$ be the $\mathbb{R}$-graph obtained by:

- replacing the edge $\{x_\ell, y_\ell\}$ with the tree $\mathcal{T}_\ell$, identifying ρ_ℓ with x_ℓ and L_ℓ with y_ℓ, for each $1 \leq \ell \leq k$,
- gluing to the vertex v_i the collection of stable trees $\mathcal{T}_{i,j}, j \geq 0$, by identifying all the roots $\rho_{i,j}$ to v_i (this gluing a.s. gives a compact metric space, see Section 5.4.2), for each $1 \leq i \leq k-s$.

On an event of probability one the graph $\mathcal{G}(G)$ is therefore compact, and is naturally endowed with the probability measure induced by the rescaled probability measures on the α-stable trees $\mathcal{T}_\ell$, $\mathcal{T}_{i,j}$, $i, j, \ell \in \mathbb{N}$. We view it as a random variable in $(\mathscr{C}, \mathrm{d}_{\mathrm{GHP}})$.

Theorem 5.6. *Given the random kernel K^s, let $\mathcal{G}(\mathsf{K}^s)$ be the graph constructed above by gluing α-stable trees along the edges and vertices of K^s. Then*

$$\mathcal{G}^s \overset{\mathrm{d}}{=} \mathcal{G}(\mathsf{K}^s),$$

as random variables in $(\mathscr{C}, \mathrm{d}_{\mathrm{GHP}})$.

We prove Theorem 5.6 in Section 5.4.2 via the recursive construction of the discrete graphs $\mathsf{G}_n^s, n \geq 0$. As a byproduct of the proof, we obtain the distribution of the continuous marginals $\mathcal{G}_n^s$, which may be viewed as G_n^s with random edge-lengths. In particular, when $n = 0$, we obtain the distribution of the continuous kernel $\mathcal{K}^s$.

Proposition 5.7. *For $n \geq 0$, given $\mathsf{G}_n^s = (V, E)$, let $(L(e), e \in E)$ be the lengths of the corresponding edges in $\mathcal{G}_n^s$, in arbitrary order. Then,*

$$(\alpha \cdot L(e), e \in E)$$

is distributed as the product of three independent random variables:

$$\mathrm{Beta}\left(|E|, \frac{(n+s)\alpha+s-1}{\alpha-1} - |E|\right) \cdot \mathrm{ML}\left(1-\frac{1}{\alpha}, \frac{(n+s)\alpha+s-1}{\alpha}\right) \cdot \mathrm{Dir}(1,\dots,1). \quad (5.11)$$

Here, $\mathrm{ML}(\beta,\theta)$ denotes the generalised Mittag-Leffler distribution with parameters $0<\beta<1$ and $\theta > -\beta$.

Construction 2: line-breaking

Various prominent examples of random metric spaces may be obtained as the limit of a so-called line-breaking procedure that consists in gluing recursively segments of random lengths – or more complex measured metric structures – to obtain a growing structure. The most famous is the line-breaking construction of the Brownian continuum random tree discovered by Aldous in [9]. We refer to [4, 41, 63, 109, 112, 111] for other models studied since then.

The graph $\mathcal{G}^s$ may also be constructed in such a way, starting from its kernel. This construction makes use of an increasing $\mathbb{R}_+$-valued Markov chain $(R_n)_{n\geq 1}$ which is characterized by the following two properties for each $n \geq 1$:

$$R_n \sim \mathrm{ML}\left(1-\frac{1}{\alpha}, \frac{n\alpha+(s-1)}{\alpha}\right) \qquad \text{and} \qquad R_n = R_{n+1}\cdot B_n$$

where $B_n \sim \mathrm{Beta}\left(\frac{(n+1)\alpha+s-2}{\alpha-1}, \frac{1}{\alpha-1}\right)$ is a random variable independent of R_{n+1}. (An explicit construction of this Markov chain is given e.g. in [63, Section 1.2]. Note that similar Markov chains arise in the scaling limits of several stochastic models, see [79, 110].)

For the moment, assume that $s \geq 1$. Suppose we are given K^s with, say, k edges and internal vertices $v_1,\dots,v_{k-s}$ having degrees $d_1,\dots,d_{k-s}$ respectively (the order of labelling is unimportant). We first perform an initialisation step: independently of the Markov chain $(R_n)_{n\geq 1}$,

- sample
$$(\Theta_1,\dots,\Theta_{2k-s}) \sim \mathrm{Dir}\Big(\underbrace{1,\dots,1}_{k}, \frac{d_1-1-\alpha}{\alpha-1},\dots,\frac{d_{k-s}-1-\alpha}{\alpha-1}\Big);$$
- assign the lengths $R_s\cdot\Theta_1,\dots,R_s\cdot\Theta_k$ to the k edges of K^s (the order is again unimportant); viewing the edges as closed line-segments, this gives a metric space that we denote $\mathcal{H}_0^s$, with $k-s$ branch-points (i.e. vertices of degree at least 3) labelled $v_1,\dots,v_{k-s}$;
- let $\eta_0 := \lambda_{\mathcal{H}_0^s} + \sum_{i=1}^{k-s}(R_s\cdot\Theta_{k+i})\delta_{v_i}$, where $\lambda_{\mathcal{H}_0^s}$ denotes the Lebesgue measure on $\mathcal{H}_0^s$.

We now build a growing sequence of measured metric spaces $(\mathcal{H}_n^s,\eta_n)_{n\geq 0}$, starting from $(\mathcal{H}_0^s,\eta_0)$. Recursively,

- select a point v in $\mathcal{H}_n^s$ with probability proportional to η_n;
- attach to v a new closed line-segment σ of length $(R_{n+s+1}-R_{n+s})\cdot\beta_n$, where β_n has a $\mathrm{Beta}(1,(2-\alpha)/(\alpha-1))$-distribution and is independent of everything constructed until now; this gives $\mathcal{H}_{n+1}^s$;
- let $\eta_{n+1} := \eta_n + (R_{n+s+1}-R_{n+s})\cdot(1-\beta_n)\delta_v + \lambda_\sigma$, where λ_σ denotes the Lebesgue measure on σ.

When $s=0$ the construction works similarly except that the initialization starts at $n=1$ with $\mathcal{H}_1^0$ taken to be a closed segment of length R_1, equipped with the Lebesgue measure denoted by η_1. We have the following result, which is proved in Section 5.4.3.

Theorem 5.8. *The sequence* $(\mathcal{H}_n^s, n \geq 0)$ *is distributed as* $(\mathcal{G}_n^s, n \geq 0)$. *In consequence, the graph* $\mathcal{H}_n^s$, *endowed with the uniform probability on its set of leaves, converges almost surely for the Gromov-Hausdorff-Prokhorov topology to a random compact measured metric space distributed as* $\mathcal{G}^s$. *In particular,* $\overline{\cup_{n\geq 0}\mathcal{H}_n^s}$ *is a version of* $\mathcal{G}^s$.

Remark 5.9. *We adopt a "discrete" approach to proving Theorems 5.6 and 5.8; in other words, we make use of Marchal's algorithm and the fact that it gives us a sequence of approximations which, on rescaling, converges almost surely to the connected* α*-stable graph with surplus* s. *An alternative approach should be possible, whereby one would work directly in the continuum, but it is far from clear to us that it would be any simpler to implement.*

5.1.3 The finite third moment case, and other related work

The case where

$$\mathbb{E}D_1^2 = 2\mathbb{E}D_1 \quad \text{and} \quad \mathbb{E}D_1^3 < \infty$$

has already been well-studied. In particular, when $\mathbb{P}(D_1 = 2) < 1$, if we let $\beta = \mathbb{E}D_1(D_1 - 1)(D_1 - 2)$ then Theorem 5.1 holds with $\alpha = 2$ on rescaling the counting measure on each component by $\beta^{-1}n^{-2/3}\mathbb{E}D_1$ and the graph distances by $n^{-1/3}$. The limiting graphs are constructed similarly to ours but using a standard Brownian motion instead of a spectrally positive α-stable Lévy process (with the small variation that β appears in the change of measure). See [24, Theorem 2.4 and Construction 3.5] and also Section 3 of [62] for more details. This *Brownian graph* first appeared as the scaling limit of the critical Erdős-Rényi random graph [5] and is now known to be the universal scaling limit of various other critical random graph models. Precise analogues of our main results were already known in this Brownian case (except for Theorem 5.5).

It follows from the properties of Brownian motion that the branch-points in $\mathcal{G}_{\mathrm{Br}}^s$, the connected Brownian graph with surplus s, are then all of degree 3. Its discrete kernel $\mathsf{K}_{\mathrm{Br}}^s$ is therefore a 3-regular planted multigraph, whose distribution is given below.

Theorem 5.10 ([4, Figure (2)] and [82, Theorem 7]). *For a connected* 3*-regular planted multigraph* G *with surplus* s,

$$\mathbb{P}\left(\mathsf{K}_{\mathrm{Br}}^s = G\right) \propto \frac{1}{|\mathrm{Sym}(G)| 2^{\mathrm{sl}(G)} \prod_{e \in \mathrm{supp}(E(G))} \mathrm{mult}(e)!}.$$

(In the references given, the kernel is taken to be labelled and unrooted, but the labelling can be removed simply at the cost of the factor of $|\mathrm{Sym}(G)|^{-1}$ appearing in the above expression, and the root can be removed as detailed above.) See Figure 5.2 for numerical values when $s = 2$. Note that the formula above corresponds to that of Theorem 5.3 when $n = 0$ and $\alpha = 2$ since then

$$w_0 = w_2 = 1 \quad \text{and} \quad w_i = 0 \quad \text{for all other indices } i.$$

In fact, our proofs in Section 5.3 can be adapted to recover this case and more generally to obtain the joint distribution of the marginals $\mathsf{G}_{n,\mathrm{Br}}^s$ via a recursive construction which is particularly simple in this case: starting from the kernel $\mathsf{K}_{\mathrm{Br}}^s$, at each step a new edge-leaf is attached to an edge chosen uniformly at random from among the set of edges of the pre-existing structure. (For

$s = 0$, this is Rémy's algorithm [107] for generating a uniform binary leaf-labelled tree.) After n steps, this gives a version of $\mathsf{G}^s_{n,\mathrm{Br}}$, whose distribution is specified below.

Proposition 5.11. *For every multigraph $G \in \mathbb{M}_{s,n}$ with internal vertices all of degree 3,*

$$\mathbb{P}\left(\mathsf{G}^s_{n,\mathrm{Br}} = G\right) \propto \frac{1}{|\mathrm{Sym}(G)| 2^{\mathrm{sl}(G)} \prod_{e \in \mathrm{supp}(E(G))} \mathrm{mult}(e)!}.$$

As in the stable cases, these distributions are connected to configuration multigraphs. Indeed, let $D^{(\mathrm{Br})}$ denote a random variable with distribution

$$\mathbb{P}(D^{(\mathrm{Br})} = 1) = 3/4 \quad \text{and} \quad \mathbb{P}(D^{(\mathrm{Br})} = 3) = 1/4.$$

Consider then the following particular instance of the configuration model. We fix $n \geq 0$, $m \geq n+1$ and take vertices labelled $0, 1, \ldots, m-1$ to have i.i.d. degrees distributed according to $D^{(\mathrm{Br})}$. We then write $\mathsf{C}^s_{n,m}$ for the resulting configuration multigraph conditioned to be in $\mathbb{M}_{s,n}$, after having forgotten the labels $n+1, n+2, \ldots, m-1$.

Corollary 5.12. *The random multigraph $\mathsf{G}^s_{n,\mathrm{Br}}$ conditioned to have m vertices has the same law as $\mathsf{C}^s_{n,m}$.*

The paper [4] is devoted to the study of the distribution of $\mathcal{G}^s_{\mathrm{Br}}$ for $s \geq 0$. In particular, it is shown there that a version of $\mathcal{G}^s_{\mathrm{Br}}$ can be recovered by gluing appropriately rescaled Brownian continuum random trees along the edges of $\mathsf{K}^s_{\mathrm{Br}}$ ([4, Procedure 1]) or via a line-breaking construction ([4, Procedure 2 & Theorem 4]).

Let us turn now to other related work. The study of scaling limits for critical random graph models was initiated by Aldous in [8], where he proved in particular the convergence of the sizes and surpluses of the largest components of the Erdős-Rényi random graph in the critical window, as well as a similar result for the sizes of the largest components in an inhomogeneous random graph model. This was followed soon afterwards by Aldous and Limic [11], who explored the possible scaling limits for the sizes of the components in a "rank-one" inhomogeneous random graph, with the limiting sizes encoded as the lengths of excursions above past-minima of a so-called thinned Lévy process.

In [5], it was shown that Aldous' result for the sizes and surpluses of the largest components in a critical Erdős-Rényi random graph could be extended to include also the metric structure of the limiting components; the limiting object is what we refer to here as the *Brownian graph*. Since that paper, progress has been made in several directions. One direction has been to demonstrate the *universality* of the Brownian graph (first in terms of component sizes, and then in terms of the full metric structure). This has been done for the critical rank-one inhomogeneous random graph [22, 25], for critical Achlioptas processes with bounded size rules [19], for critical configuration models with finite third moment degrees [83, 108, 46, 24] and in great generality in [18].

Another line of enquiry, into which the present paper fits, is the investigation of other universality classes, generally those with power-law degree distributions. This has been pursued in the setting of rank-one inhomogeneous random graphs with power-law degrees in [76, 21, 23] and with very general weights by [31]. The configuration model with power-law degrees has been treated by [83, 47, 20]. The last three papers are the most directly related to the topic of the present paper, and so we will discuss them in a little more detail.

In [83], Joseph considers the configuration model with i.i.d. degrees satisfying the same conditions as us, and proves the convergence in distribution of the component sizes (5.4). (He leaves the equivalent convergence in the setting of the graph conditioned to be simple as a conjecture, but this is not hard to prove; see [38] for the details.) The results of [38] in Theorem 5.1 thus directly generalise those of Joseph. Dhara, van der Hofstad, van Leeuwaarden and Sen [47] and Bhamidi, Dhara, van der Hofstad, and Sen [20] consider the component sizes and metric structure respectively in configuration models with fixed degree sequences satisfying a certain power-law condition. The paper [20] proves a metric space scaling limit, where the limit components are derived from the thinned Lévy processes mentioned above. This scaling limit is proved in a somewhat weaker topology than that of [38] but is much more general in scope; in particular, it includes the case of i.i.d. degrees with the tail behaviour we assume. In principle, it should be possible to view the stable graph as an appropriately annealed version of the scaling limit of [20]. However, it is for the moment unclear how to prove independently that the two objects obtained must be the same. The limit spaces obtained in [20] are *a priori* much less easy to understand than ours; the advantage of our more restrictive setting is that we get very nice absolute continuity relations with the stable trees which are already well understood. Obtaining analogous results in the setting studied by [20] seems much more challenging.

5.1.4 Perspectives

As discussed above, the results of this paper provide heavy-tailed analogues of those in [4], which have been applied in other contexts. Firstly, the decomposition into a continuous kernel with explicit distribution plus pendant subtrees played a key role in the proof of the existence of a scaling limit for the minimum spanning tree of the complete graph on n vertices in [6]. More specifically, assign the edges of the complete graph i.i.d. random edge-weights with Exp(1) distribution. Now find the spanning tree M_n of the graph with minimum total edge-weight. (The law of M_n does not depend on the weight distribution as long as it is non-atomic.) Think of M_n as a measured metric space in the usual way by endowing it with the graph distance d_n and the uniform probability measure μ_n on its vertices. The main result of [6] is that

$$(M_n, n^{-1/3} d_n, \mu_n) \xrightarrow{\mathrm{d}} (\mathcal{M}, d, \mu)$$

as $n \to \infty$, in the Gromov-Hausdorff-Prokhorov sense, where the limit space $(\mathcal{M}, d, \mu)$ is a random measured $\mathbb{R}$-tree having Minkowski dimension 3 almost surely. This convergence has, up to a constant factor, recently been shown by Addario-Berry and Sen [7] to hold also for the MST of a uniform random 3-regular (simple) graph or for the MST of a 3-regular configuration model.

Following the scheme of proof developed in [6], it should be possible to use the results of the present paper together with those of [38] to prove an analogous scaling limit for the minimum spanning tree of the following model. First, generate a uniform random graph (or configuration model) with i.i.d. degrees $D_1, D_2, \ldots, D_n$ with the same power-law tail behaviour as discussed above, but now in the supercritical setting $\nu > 1$. For the purposes of this discussion, let us also assume that $\mathbb{P}(D_1 \geq 3) = 1$. Under this condition, the graph not only has a giant component, but that component contains all of the vertices with probability tending to 1 [33, Lemma 1.2]. As before, assign the edges of this graph i.i.d. random weights with Exp(1) distribution and find the minimum spanning tree M_n. Then we conjecture that in this setting we will have

$$(M_n, n^{-(\alpha-1)/(\alpha+1)} d_n, \mu_n) \xrightarrow{\mathrm{d}} (\mathcal{M}, d, \mu),$$

for some measured $\mathbb{R}$-tree $(\mathcal{M}, d, \mu)$. This conjecture will be the topic of future work.

Another application of the results of [4] has been in the context of random maps. The Brownian versions of the graphs $\mathcal{G}^s, s \geq 0$ arise as scaling limits of unicellular random maps on various compact surfaces. The results of [4] have, in particular, been used to study Voronoi cells in these objects. More specifically, for a surface S, let $(\mathcal{U}(S), d, \mu)$ be the continuum random unicellular map on S [3], endowed with its mass measure μ, and let $X_1, X_2, \ldots, X_k$ be independent random points sampled from μ. Let $V_1, V_2, \ldots, V_k$ be the Voronoi cells with centres $X_1, \ldots, X_k$. Then in [3] it is shown that

$$(\mu(V_1), \ldots, \mu(V_k)) \sim \mathrm{Dir}(1, 1, \ldots, 1).$$

In other words, the Voronoi cells of uniform points provide a way to split the mass of the space up uniformly. In principle, there should exist "stable" analogues of this result (in which the mass-split will no longer be uniform).

5.1.5 Organisation of the chapter

Section 5.2 is devoted to background on stable trees, and to the description of the distribution of the limiting sequence of metric spaces arising in Theorem 5.1 in terms of a spectrally positive α-stable Lévy process. In particular, we give a precise description of the elementary building-blocks $\mathcal{G}^s, s \geq 0$. We then enter the core of the paper with Section 5.3 which is dedicated to the proof of the joint distribution of the discrete marginals $\mathsf{G}_n^s, n \geq 0$ (Theorems 5.3 and 5.5), including the connection to a configuration model stated in Corollary 5.4. Section 5.4 is devoted to the proofs of the construction of the $\mathbb{R}$-graph $\mathcal{G}^s$ from randomly scaled trees glued to its kernel and of its line-breaking construction (Theorem 5.6, Proposition 5.7 and Theorem 5.8, as well as Proposition 5.2). Finally, in the appendix, Section 5.5, we recall the definitions and some properties of various distributions (generalized Mittag-Leffler, Beta, Dirichlet and Poisson-Dirichlet), as well as some classical urn model asymptotics, which are used at various points in the paper.

5.2 The stable graphs

We begin in Section 5.2.1 with some necessary background on stable trees. In particular, we recall Marchal's algorithm for constructing the discrete ordered marginals, and use it to obtain the joint distribution of various aspects (lengths, weights, local times) of the continuous marginals, which we will need later on. In Section 5.2.2, we turn to the distribution of the limiting sequence of metric spaces arising in Theorem 5.1 and in particular to the construction of the stable graphs.

Throughout this section, we fix $\alpha \in (1, 2)$.

5.2.1 Background on stable trees

Construction and properties

The α-stable tree was introduced by Duquesne and Le Gall [53], building on earlier work of Le Gall and Le Jan [89]. Our presentation of this material owes much to that of Curien and Kortchemski [43], which relies in turn on various key results from Miermont [95].

First, let ξ be a spectrally positive α-stable Lévy process with Laplace exponent

$$\mathbb{E}\mathrm{exp}(-\lambda \xi_t) = \exp(t\lambda^\alpha), \quad \lambda \geq 0, \quad t \geq 0.$$

Now consider a reflected version of this Lévy process, namely $(\xi_t - \inf_{0\le s\le t}\xi_s, t\ge 0)$. It is standard that this process has an associated excursion theory, and that one can make sense of an excursion conditioned to have length 1. We will write X for this excursion of length 1, and observe that, thanks to the scaling property of ξ we may obtain the law of an excursion conditioned to have length $x>0$ via $(x^{1/\alpha}X(t/x), 0\le t\le x)$. See Chaumont [34] for more details.

To a normalised excursion X we may associate an $\mathbb{R}$-tree. In order to do this, we first derive from X a *height function* H, defined as follows: for $t\in[0,1]$,

$$H(t) = \lim_{\varepsilon\to 0+}\frac{1}{\varepsilon}\int_0^t \mathbf{1}_{\{X(s)<\inf_{s\le r\le t}X(r)+\varepsilon\}}\mathrm{d}s.$$

The process H possesses a continuous modification such that $H(0)=H(1)=0$ and $H(t)>0$ for $t\in(0,1)$, which we consider in the sequel (see Duquesne and Le Gall [53] for more details). We then obtain an $\mathbb{R}$-tree in a standard way from H by first defining a pseudo-distance d on $\mathbb{R}_+$ via

$$d(s,t) = H(s)+H(t)-2\inf_{s\wedge t\le r\le s\vee t}H(r).$$

Now define an equivalence relation $\sim$ by declaring $s\sim t$ if $d(s,t)=0$. Then let $\mathcal{T}$ be the metric space obtained by endowing $[0,1]/\sim$ with the image of d under the quotienting operation. Let us write $\pi:[0,1]\to\mathcal{T}$ for the projection map. We additionally endow $\mathcal{T}$ with the push-forward of the Lebesgue measure on $[0,1]$ under π, which is denoted by μ. The point $\rho:=\pi(0)=\pi(1)$ is naturally interpreted as a *root* for the tree. We will refer to the random variable $(\mathcal{T},d,\mu)$ as the (standard) α*-stable tree*. In the usual notation, for points $x,y\in\mathcal{T}$, we will write $[\![x,y]\!]$ for the path between x and y in $\mathcal{T}$, and $]\!]x,y[\![$ for $[\![x,y]\!]\setminus\{x,y\}$. (These are isometric to closed and open line-segments of length $d(x,y)$, respectively.) We can use the root to endow the tree $\mathcal{T}$ with a *genealogical order*: we say $x\preceq y$ if $x\in[\![\rho,y]\!]$. We define the *degree*, $\deg(x)$, of a point $x\in\mathcal{T}$ to be the number of connected components into which its removal splits the space. If there is any potential ambiguity over which metric space we are working in, we will write $\deg_{\mathcal{T}}(x)$. The *branchpoints* are those with degree strictly greater than 2 and the leaves are those with degree 1; we write $\mathrm{Br}(\mathcal{T})=\{x\in\mathcal{T}:\deg(x)>2\}$ and $\mathrm{Leaf}(\mathcal{T})=\{x\in\mathcal{T}:\deg(x)=1\}$. We observe that the distance d induces a natural length measure on the tree $\mathcal{T}$, for which we write λ.

We also define a partial order $\preceq$ on $[0,1]$ by declaring

$$s\preceq t \quad\text{if}\quad s\le t \text{ and } X(s-)\le \inf_{s\le r\le t}X(r). \tag{5.12}$$

(We take as a convention that $X(0-)=0$.) This partial order is compatible with the genealogical order on $\mathcal{T}$ in the sense that for $x,y\in\mathcal{T}$, $x\preceq y$ if and only if there exist $s,t\in[0,1]$ such that $x=\pi(s)$ and $y=\pi(t)$ and $s\preceq t$.

We will require various properties of $\mathcal{T}$ in the sequel. We will make use of the fact that the root ρ acts as a uniform sample from the measure μ and so we will sometimes think of the tree as unrooted and regenerate a root from μ when necessary. Another key feature of $\mathcal{T}$ is that its branchpoints are all of infinite degree, almost surely. By Proposition 2 of Miermont [95], $x\in\mathrm{Br}(\mathcal{T})$ if and only if there exists a unique $s\in[0,1]$ such that $x=\pi(s)$ and $\Delta X(s) = X(s)-X(s-)>0$. For all other values $r\in[0,1]$ such that $\pi(r)=\pi(s)=x$, we have $\inf_{s\le u\le r}X(u)=X(r)\ge X(s-)$. For such s associated to a branchpoint $x=\pi(s)$, we will define $N(x):=\Delta X(s)$. By Miermont's equation [95, Eq. (1)], for all $x\in\mathrm{Br}(\mathcal{T})$ this quantity may be almost surely recovered as

$$N(x)=\lim_{\varepsilon\to 0+}\frac{1}{\varepsilon}\mu(\{y\in\mathcal{T}:x\in[\![\rho,y]\!], d(x,y)<\varepsilon\}),$$

and so $N(x)$ gives a renormalised notion of the degree of x. We will refer to this quantity as the *local time* of x, since it plays that role with respect to H.

For any $s, t \in [0,1]$ such that $\pi(s) \in \mathrm{Br}(\mathcal{T})$ and $s \preceq t$, we also define the *local time of* $\pi(s)$ *to the right of* $\pi(t)$ to be

$$N^{\mathrm{right}}(\pi(s), \pi(t)) = \inf_{s \leq u \leq t} X(u) - X(s-).$$

Then $N^{\mathrm{right}}(\pi(s), \pi(t)) \in [0, N(\pi(s))]$ is a measure of how far through the descendants of $\pi(s)$ we are when we visit $\pi(t)$. (Indeed, since $\pi(s) \in \mathrm{Br}(\mathcal{T})$, if $s \preceq t$ and $s \preceq u$ with $N^{\mathrm{right}}(\pi(s), \pi(t)) > N^{\mathrm{right}}(\pi(s), \pi(u))$ then necessarily $t < u$.) By Corollary 3.4 of [43], we can express $X(t)$ as the sum of the atoms of local time along the path from the root to $\pi(t)$:

$$X(t) = \sum_{0 \preceq s \preceq t} N^{\mathrm{right}}(\pi(s), \pi(t)), \tag{5.13}$$

almost surely for all $t \in [0,1]$. For any $s \preceq t$, we define the local time along the path $]\!]\pi(s), \pi(t)[\![$ by

$$N(\,]\!]\pi(s), \pi(t)[\![\,) := \sum_{b \in \mathrm{Br}(\mathcal{T}) \cap]\!]\pi(s), \pi(t)[\![} N(b),$$

and the local time to the right along the path $]\!]\pi(s), \pi(t)[\![$ by

$$N^{\mathrm{right}}(\,]\!]\pi(s), \pi(t)[\![\,) := \sum_{b \in \mathrm{Br}(\mathcal{T}) \cap]\!]\pi(s), \pi(t)[\![} N^{\mathrm{right}}(b, \pi(t)) = X(t-) - X(s),$$

where we observe that all of these sums are over countable sets.

Marchal's algorithm for ordered trees

Consider an infinite sample of leaves from $(\mathcal{T}, d, \mu)$ obtained as the images of i.i.d. uniform random variables $U_1, U_2, \ldots$ on $[0,1]$ under the quotienting. These leaves, which we label $1, 2, \ldots$, inherit an order from $[0,1]$. For $n \in \mathbb{N}$, let $\mathcal{T}_n^{\mathrm{ord}}$ be an ordered leaf-labelled version of the subtree of $\mathcal{T}$ spanned by the root and the first n leaves (the order being inherited from the leaves) and $\mathsf{T}_n^{\mathrm{ord}}$ its combinatorial shape, also with leaf-labels. Formally,

$$\mathsf{T}_n^{\mathrm{ord}} = \mathrm{shape}(\mathcal{T}_n^{\mathrm{ord}})$$

where, for any compact rooted (say at ρ) real tree τ (possibly ordered), shape(τ) is the (possibly ordered) rooted discrete tree (V, E) with no vertex of degree 2 except possibly the root, where

$$V = \{\rho\} \cup \{v \in \tau \backslash \{\rho\} : \deg_\tau(v) \neq 2\} \text{ and } E = \Big\{ \{u, v\} : \ u, v \in V, \ \deg_\tau(w) = 2, \\ \forall w \in]]u, v[[\text{ and } \rho \notin]]u, v[[\Big\}. \tag{5.14}$$

We define the shape of a discrete tree similarly. Note that all of the trees we shall consider have a root of degree 1: they are *planted.*

For any $n \geq 1$, we denote by $\mathbb{A}_n$ the set of planted ordered finite trees with n labelled leaves, with labels from 1 to n, and no vertex of degree 2. The root is thought of as a leaf with label 0. In [53, Section 3], Duquesne and Le Gall show that for each tree $T \in \mathbb{A}_n$ with set of internal vertices $I(T)$,

$$\mathbb{P}\big(\mathsf{T}_n^{\mathrm{ord}} = T\big) \propto \prod_{u \in I(T)} \frac{w_{\deg_T(u)-1}}{(\deg_T(u) - 1)!}, \tag{5.15}$$

where the weights $(w_k, k \geq 0)$ were defined in (5.6). In other words, $\mathsf{T}_n^{\text{ord}}$ is distributed as a planted version of a Galton-Watson tree with offspring distribution η_α as defined in Section 5.1.2 (below Corollary 5.4), conditioned on having n leaves uniformly labelled from 1 to n.

Building on this result, in [92] Marchal proposed a recursive construction of a sequence with the same law as $(\mathsf{T}_n^{\text{ord}}, n \geq 1)$. (In fact, Marchal gave a construction of the non-ordered versions of the trees $\mathsf{T}_n^{\text{ord}}, n \geq 1$ but combined with [92, Section 2.3] we easily obtain an ordered version.) For any $n \geq 1$ and any $T \in \mathbb{A}_n$, we construct randomly a tree in $\mathbb{A}_{n+1}$ as follows.

(1) Assign to every edge of T a weight $\alpha-1$ and every internal vertex u a weight $\deg_T(u)-1-\alpha$; the other vertices have weight 0;

(2) Choose an edge/vertex with probability proportional to its weight and then

- if it is a vertex, choose a uniform corner around this vertex, attach a new edge-leaf in this corner and give the leaf the label $n+1$,
- if it is an edge, create a new vertex which splits the edge into two edges, and attach an edge-leaf with leaf labelled $n+1$ pointing to the left/right with probability $1/2$.

If we start with the unique element of $\mathbb{A}_1$ and apply this procedure recursively, we obtain a sequence of trees distributed as $(\mathsf{T}_n^{\text{ord}}, n \geq 1)$.

Asymptotic behaviour. Consider now the discrete trees as metric spaces, endowed with the graph distance. Fix k and for each $k \leq n$ let $\mathsf{T}_k^{\text{ord}}(n)$ be the subtree of $\mathsf{T}_n^{\text{ord}}$ spanned by the k first leaves and the root. Hence, $\mathsf{T}_k^{\text{ord}} = \text{shape}(\mathsf{T}_k^{\text{ord}}(n))$ but the distances in $\mathsf{T}_k^{\text{ord}}(n)$ are inherited from those in $\mathsf{T}_n^{\text{ord}}$. We may therefore view $\mathsf{T}_k^{\text{ord}}(n)$ as a discrete tree having the same vertex- and edge-sets as $\mathsf{T}_k^{\text{ord}}$, but where the edges now have lengths. Similarly for $\mathcal{T}_k^{\text{ord}}$. Again from Marchal [92], we have

$$\frac{\mathsf{T}_k^{\text{ord}}(n)}{n^{1-1/\alpha}} \xrightarrow[n\to\infty]{\text{a.s.}} \alpha \cdot \mathcal{T}_k^{\text{ord}}, \tag{5.16}$$

as $n \to \infty$, where the convergence means that the rescaled lengths of the edges of $\mathsf{T}_k^{\text{ord}}(n)$ converge to the lengths, multiplied by α, of the corresponding edges in $\mathcal{T}_k^{\text{ord}}$. This convergence of random finite-dimensional marginals can be improved when considering trees as metric spaces (i.e. we forget the order) equipped with probability measures. Indeed, if T_n denotes the unordered version of $\mathsf{T}_n^{\text{ord}}$, with leaves still labelled $0, 1, 2, \ldots$ (0 is the root), μ_n the uniform probability measure on these leaves, then we have that

$$\left(\frac{\mathsf{T}_n}{n^{1-1/\alpha}}, \mu_n, 0, \ldots, k\right) \xrightarrow[n\to\infty]{\text{a.s.}} \alpha \cdot (\mathcal{T}, \mu, 0, \ldots, k) \tag{5.17}$$

for the $(k+1)$-pointed Gromov-Hausdorff-Prokhorov topology on the set of measured $(k+1)$-pointed compact trees, for each integer k. (See e.g. [96, Section 6.4] for a definition of this topology.) The convergence (5.17) was first proved in probability in [69, Corollary 24] and then improved to an almost sure convergence in [42, Section 2.4].

Suppose now that $\mathsf{T}_k^{\text{ord}}$ has edge-set $E(\mathsf{T}_k^{\text{ord}})$, labelled arbitrarily as $e_i, 1 \leq i \leq \left|E(\mathsf{T}_k^{\text{ord}})\right|$, and internal vertices $I(\mathsf{T}_k^{\text{ord}})$, labelled arbitrarily as v_j, $1 \leq j \leq \left|I(\mathsf{T}_k^{\text{ord}})\right|$. As discussed above, for $k \leq n$, the internal vertices $I(\mathsf{T}_k^{\text{ord}})$ all have counterparts in $\mathsf{T}_k^{\text{ord}}(n)$, which we will also call v_j, $1 \leq j \leq \left|I(\mathsf{T}_k^{\text{ord}})\right|$. To each edge $e_i \in E(\mathsf{T}_k^{\text{ord}})$ there corresponds a path γ_i in $\mathsf{T}_k^{\text{ord}}(n)$ whose endpoints are elements of $\{v_j, 1 \leq j \leq |I(\mathsf{T}_k^{\text{ord}})|\} \cup \{0, 1, \ldots, k\}$. Write γ_i° for the same path with its endpoints removed (γ_i° may be empty). Since $\mathsf{T}_k^{\text{ord}}(n) \subset \mathsf{T}_n^{\text{ord}}$, we refer to the corresponding vertices and paths in $\mathsf{T}_n^{\text{ord}}$ by the same names.

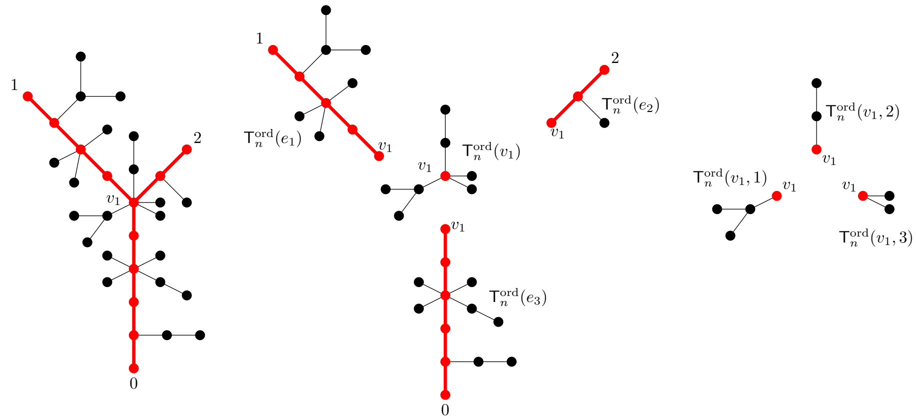


Figure 5.3 – Left: the tree $\mathsf{T}_n^{\mathrm{ord}}$ for $n = 18$ (leaf-labels $3, \ldots, 18$ are suppressed for purposes of readability). $\mathsf{T}_2^{\mathrm{ord}}(n)$ is emphasised in red and bold. The tree $\mathsf{T}_2^{\mathrm{ord}}$ has a single internal vertex called v_1 and edges $e_1 = \{v_1, 1\}$, $e_2 = \{v_1, 2\}$ and $e_3 = \{v_1, 0\}$. The corresponding paths in $\mathsf{T}_2^{\mathrm{ord}}(n)$ have lengths 4, 2 and 5 respectively. Middle: the subtrees $\mathsf{T}_n^{\mathrm{ord}}(e_1)$, $\mathsf{T}_n^{\mathrm{ord}}(e_2)$, $\mathsf{T}_n^{\mathrm{ord}}(e_3)$ and $\mathsf{T}_n^{\mathrm{ord}}(v_1)$. Right: the subtrees $\mathsf{T}_n^{\mathrm{ord}}(v_1, 1)$, $\mathsf{T}_n^{\mathrm{ord}}(v_1, 2)$ and $\mathsf{T}_n^{\mathrm{ord}}(v_1, 3)$.

We will now give names to certain important subtrees of $\mathsf{T}_n^{\mathrm{ord}}$ and refer the reader to Figure 5.3 for an illustration. For each vertex $v \in V(\mathsf{T}_n^{\mathrm{ord}})$, the unique directed path from v to 0 has a first point $\mathrm{int}(v)$ of intersection with $\mathsf{T}_k^{\mathrm{ord}}(n)$. For $1 \leq j \leq |I(\mathsf{T}_k^{\mathrm{ord}})|$, let $\mathsf{T}_n^{\mathrm{ord}}(v_j)$ be the subtree induced by the set of vertices $\{v : \mathrm{int}(v) = v_j\}$ and rooted at v_j. If $\mathrm{int}(v) \notin \{v_j : 1 \leq j \leq |I(\mathsf{T}_k^{\mathrm{ord}})|\}$ then $\mathrm{int}(v)$ belongs to γ_i° for some $1 \leq i \leq |E(\mathsf{T}_k^{\mathrm{ord}})|$. Let $\mathsf{T}_n^{\mathrm{ord}}(e_i)$ be the subtree of $\mathsf{T}_n^{\mathrm{ord}}$ induced by the vertices $\{v \in V(\mathsf{T}_n^{\mathrm{ord}}) : \mathrm{int}(v) \in \gamma_i^\circ\} \cup \gamma_i$ and rooted at the endpoint of γ_i closest to the root of $\mathsf{T}_n^{\mathrm{ord}}$.

If $\deg_{\mathsf{T}_k^{\mathrm{ord}}}(v_j) = d_j$ then $\mathsf{T}_n^{\mathrm{ord}}(v_j)$ can be split up into separate subtrees descending from the d_j different corners of v_j. We list these subtrees in clockwise order from the root as $\mathsf{T}_n^{\mathrm{ord}}(v_j, \ell)$, $1 \leq \ell \leq d_j$.

For each $e_i, 1 \leq i \leq |E(\mathsf{T}_k^{\mathrm{ord}})|$ then denote by

- $L_n(e_i)$ the length of γ_i in $\mathsf{T}_k^{\mathrm{ord}}(n)$,
- $M_n(e_i)$ the number of leaves in the subtree $\mathsf{T}_n^{\mathrm{ord}}(e_i)$,
- $N_n(e_i)$ the number of edges of $\mathsf{T}_n^{\mathrm{ord}}(e_i)$ adjacent to γ_i,
- $N_n^{\mathrm{right}}(e_i)$ the number of edges of $\mathsf{T}_n^{\mathrm{ord}}(e_i)$ attached to the right of γ_i,
- $N_n(e_i, \ell)$ the degree -2 of the ℓth largest branchpoint along the path γ_i in $\mathsf{T}_n^{\mathrm{ord}}(e_i)$, for $\ell \geq 1$, with ties broken arbitrarily,
- $N_n^{\mathrm{right}}(e_i, \ell)$ the degree to the right of the ℓth largest branchpoint along the path γ_i in $\mathsf{T}_n^{\mathrm{ord}}(e_i)$, for $\ell \geq 1$ (with the same labelling as in the previous point).
- $L_n(e_i, \ell)$ the distance from the ℓth largest branchpoint of γ_i to the root (endpoint nearest 0 in $\mathsf{T}_n^{\mathrm{ord}}$) of $\mathsf{T}_n^{\mathrm{ord}}(e_i)$, $\ell \geq 1$, again with the same labelling.

Observe that $N_n(e_i) = \sum_{\ell \geq 1} N_n(e_i, \ell)$ and $N_n^{\mathrm{right}}(e_i) = \sum_{\ell \geq 1} N_n^{\mathrm{right}}(e_i, \ell)$.

Similarly, for each vertex v_j, $1 \leq j \leq \left|I(\mathsf{T}_n^{\mathrm{ord}})\right|$, denote by

- $N_n(v_j)$ the degree of v_j in $\mathsf{T}_n^{\mathrm{ord}}$ (i.e. $\deg_{\mathsf{T}_n^{\mathrm{ord}}}(v_j)$),

- $N_n(v_j, \ell)$ the degree of v_j in $\mathsf{T}_n^{\mathrm{ord}}$ in the ℓth corner counting clockwise from the root, for $1 \leq \ell \leq \deg_{\mathsf{T}_k^{\mathrm{ord}}}(v_j)$,
- $M_n(v_j)$ the number of leaves in $\mathsf{T}_n^{\mathrm{ord}}(v_j)$,
- $M_n(v_j, \ell)$ the number of leaves in $\mathsf{T}_n^{\mathrm{ord}}(v_j, \ell)$, for $1 \leq \ell \leq \deg_{\mathsf{T}_k^{\mathrm{ord}}}(v_j)$.

We use the same edge- and vertex-labels for the corresponding parts of $\mathcal{T}_k^{\mathrm{ord}}$. Since $\mathcal{T}_k^{\mathrm{ord}}$ is (an ordered version of) a subset of $\mathcal{T}$, we have that e_i corresponds to an open path $]\!] x_{i,1}, x_{i,2} [\![$ for some pair of points $x_{i,1}, x_{i,2} \in \mathcal{T}$ such that $x_{i,1} \preceq x_{i,2}$. Let $L(e_i) = d(x_{i,1}, x_{i,2})$ be the length of this path. We will abuse notation somewhat by writing $N(e_i)$ and $N^{\mathrm{right}}(e_i)$ instead of $N(]\!] x_{i,1}, x_{i,2} [\![)$ and $N^{\mathrm{right}}(]\!] x_{i,1}, x_{i,2} [\![)$ for the local time of the edge and the local time to the right of the edge respectively. For $\ell \geq 1$, we will write $N(e_i, \ell)$ for the local time of the ℓth largest branchpoint along $]\!] x_{i,1}, x_{i,2} [\![$ (with ties broken arbitrarily), $N^{\mathrm{right}}(e_i, \ell)$ for the local time to the right at the same branchpoint, and $L(e_i, \ell)$ for the distance from that branchpoint to the lower endpoint $x_{i,1}$ of e_i. Each vertex v_j corresponds to some point of $\mathcal{T}$, which by abuse of notation we will also call v_j. (Note that, of course, we must have $\{v_j : 1 \leq j \leq |I(\mathsf{T}_k^{\mathrm{ord}})|\} \cup \{0, 1, \ldots, k\} = \{x_{i,p} : 1 \leq i \leq |E(\mathsf{T}_k^{\mathrm{ord}})|, p = 1, 2\}$.)

Let $\mathcal{T}(e_i)$ be the subtree of $\mathcal{T}$ containing $[\![x_{i,1}, x_{i,2}]\!]$, formally defined by

$$\mathcal{T}(e_i) = \{x \in \mathcal{T} : [\![\rho, x]\!] \cap]\!] x_{i,1}, x_{i,2} [\![\neq \emptyset, x_{i,2} \notin [\![\rho, x]\!]\} \cup \{x_{i,1}, x_{i,2}\}.$$

Let $M(e_i) = \mu(\mathcal{T}(e_i))$. Let $\mathcal{T}(v_j)$ be the subtree of $\mathcal{T}$ attached to v_j, namely

$$\mathcal{T}(v_j) = \{x \in \mathcal{T} : v_j \in [\![\rho, x]\!] \,,]\!] v_j, x [\![\cap [\![x_{i,1}, x_{i,2}]\!] = \emptyset \text{ for all } 1 \leq i \leq \left| E(\mathsf{T}_n^{\mathrm{ord}}) \right| \}.$$

Let $M(v_j) = \mu(\mathcal{T}(v_j))$. As in the discrete case, we can split up $\mathcal{T}(v_j)$ into subtrees sitting in the $\deg_{\mathsf{T}_k^{\mathrm{ord}}}(v_j)$ corners of v_j. We call these $\mathcal{T}(v_j, \ell)$ for $1 \leq \ell \leq \deg_{\mathsf{T}_k^{\mathrm{ord}}}(v_j)$. Let

$$N(v_j, \ell) = \lim_{\epsilon \to 0+} \frac{1}{\epsilon} \mu(\{y \in \mathcal{T}(v_j, \ell) : d(x_j, y) < \epsilon\}).$$

Lemma 5.13. *We have the almost sure joint convergence, for $1 \leq i \leq |E(\mathsf{T}_k^{\mathrm{ord}})|$ and $\ell \geq 1$,*

$$\begin{aligned} \frac{L_n(e_i)}{n^{1-1/\alpha}} &\underset{n\to\infty}{\longrightarrow} \alpha \cdot L(e_i), & \frac{M_n(e_i)}{n} &\underset{n\to\infty}{\longrightarrow} M(e_i), \\ \frac{N_n(e_i)}{n^{1/\alpha}} &\underset{n\to\infty}{\longrightarrow} N(e_i), & \frac{N_n^{\mathrm{right}}(e_i)}{n^{1/\alpha}} &\underset{n\to\infty}{\longrightarrow} N^{\mathrm{right}}(e_i), \\ \frac{N_n(e_i, \ell)}{n^{1/\alpha}} &\underset{n\to\infty}{\longrightarrow} N(e_i, \ell), & \frac{N_n^{\mathrm{right}}(e_i, \ell)}{n^{1/\alpha}} &\underset{n\to\infty}{\longrightarrow} N^{\mathrm{right}}(e_i, \ell), \end{aligned}$$

and for $1 \leq j \leq |I(\mathsf{T}_k^{\mathrm{ord}})|$, $1 \leq \ell \leq \deg_{\mathsf{T}_k^{\mathrm{ord}}}(v_j)$,

$$\begin{aligned} \frac{M_n(v_j)}{n} &\underset{n\to\infty}{\longrightarrow} M(v_j), \\ \frac{N_n(v_j)}{n^{1/\alpha}} &\underset{n\to\infty}{\longrightarrow} N(v_j), & \frac{N_n(v_j, \ell)}{n^{1/\alpha}} &\underset{n\to\infty}{\longrightarrow} N(v_j, \ell). \end{aligned}$$

Proof. The convergence of the lengths is Marchal's result (5.16). The convergence of the local times is proved in Dieuleveut [48, Lemma 2.7 & Lemma 2.8]. Finally, the convergences of the subtree masses are an immediate consequence of the strong law of large numbers. Note that since we are dealing with a countable collection of random variables, these convergences indeed hold simultaneously almost surely. □

Marginals of the stable tree

We now state explicitly the joint distributions of all of the limit quantities in Lemma 5.13.

Proposition 5.14. *Conditionally on $\mathsf{T}_k^{\mathrm{ord}}$ with $\left|E(\mathsf{T}_k^{\mathrm{ord}})\right| = m$ and $\left|I(\mathsf{T}_k^{\mathrm{ord}})\right| = r$, with $\deg_{\mathsf{T}_k^{\mathrm{ord}}}(v_j) = d_j$ for $1 \le j \le r$, we have jointly*

$$\begin{aligned}(M(e_1), \dots, M(e_m), M(v_1), \dots, M(v_r)) &\overset{(\mathrm{d})}{=} (D_1, D_2, \dots, D_{m+r})\\ (N(e_1), \dots, N(e_m), N(v_1), \dots, N(v_r)) &\overset{(\mathrm{d})}{=} (D_1^{1/\alpha} R_1, \dots, D_{m+r}^{1/\alpha} R_{m+r})\\ \alpha \cdot (L(e_1), \dots, L(e_m)) &\overset{(\mathrm{d})}{=} (D_1^{1-1/\alpha} R_1^{\alpha-1} \bar{R}_1, D_2^{1-1/\alpha} R_2^{\alpha-1} \bar{R}_2, \dots, D_m^{1-1/\alpha} R_m^{\alpha-1} \bar{R}_m),\end{aligned}$$

where the following elements are independent:

- $(D_1, \dots, D_m, D_{m+1}, \dots, D_{m+r}) \sim \mathrm{Dir}(1 - 1/\alpha, \dots, 1 - 1/\alpha, (d_1 - 1 - \alpha)/\alpha, \dots, (d_r - 1 - \alpha)/\alpha)$*;*
- $R_1, R_2, \dots, R_{m+r}$ *are mutually independent with* $R_1, \dots, R_m \sim \mathrm{ML}(1/\alpha, 1 - 1/\alpha)$ *and* $R_{m+i} \sim \mathrm{ML}(1/\alpha, (d_i - 1 - \alpha)/\alpha)$ *for* $1 \le i \le r$*;*
- $\bar{R}_1, \bar{R}_2, \dots, \bar{R}_m$ *are i.i.d.* $\mathrm{ML}(\alpha - 1, \alpha - 1)$.

Moreover, we have $R_i^{\alpha-1} \bar{R}_i \sim \mathrm{ML}(1 - 1/\alpha, 1 - 1/\alpha)$ *for* $1 \le i \le m$.

The random variables $N^{\mathrm{right}}(e_i, \ell)/N(e_i, \ell)$ *and* $L(e_i, \ell)/L(e_i)$ *for* $1 \le i \le m$, $\ell \ge 1$, *the random sequences* $(N(e_i, \ell)/N(e_i), \ell \ge 1)$ *for* $1 \le i \le m$, *and the random vectors* $(N(v_j, \ell)/N(v_j), 1 \le \ell \le d_j)$ *for* $1 \le j \le r$ *are mutually independent, and are also independent of* $N(e_i), 1 \le i \le m$ *and* $N(v_j), 1 \le j \le r$. *Moreover, we have*

$$\left(\frac{N(e_i, \ell)}{N(e_i)}, \ell \ge 1 \right) \sim \mathrm{PD}(\alpha - 1, \alpha - 1), \quad 1 \le i \le m,$$

$$\frac{N^{\mathrm{right}}(e_i, \ell)}{N(e_i, \ell)} \sim \mathrm{U}[0, 1], \quad 1 \le i \le m, \quad \ell \ge 1,$$

$$\frac{L(e_i, \ell)}{L(e_i)} \sim \mathrm{U}[0, 1], \quad 1 \le i \le m, \quad \ell \ge 1,$$

and

$$\left(\frac{N(v_j, \ell)}{N(v_j)}, 1 \le \ell \le d_j \right) \sim \mathrm{Dir}(1, 1, \dots, 1), \quad 1 \le j \le r.$$

The distributional results for the masses, lengths and total local times may be read off from [63], although the precise dependence between lengths and local times is left somewhat implicit there. Related results appeared earlier in [70]. We give a complete proof of Proposition 5.14 via an urn model which we now introduce.

Suppose we have k colours such that each colour has three *types*: a, b and c. Let $X_i^a(n)$, $X_i^b(n)$ and $X_i^c(n)$ be the weights of the three types of colour i in the urn at step n, respectively, for $1 \le i \le k$. At each step we draw a colour with probability proportional to its weight in the urn. If we pick the colour i type a, we add weight $\alpha - 1$ to colour i type a, $2 - \alpha$ to colour i type b and $\alpha - 1$ to colour i type c (recall that $\alpha \in (1, 2)$). If we pick colour i type b, we add 1 to colour i type b and $\alpha - 1$ to colour i type c. If we pick colour i type c, we simply add weight

α to colour i type c. We start with

$$X_i^a(0) = \gamma_i, \quad X_i^b(0) = 0, \quad X_i^c(0) = 0, \quad 1 \le i \le k.$$

Proposition 5.15. *As $n \to \infty$, we have the following almost sure limits:*

$$\frac{1}{(\alpha-1)n^{1-1/\alpha}}(X_1^a(n), \dots, X_k^a(n)) \to (D_1^{1-1/\alpha} R_1^{\alpha-1}\bar{R}_1, \dots, D_k^{1-1/\alpha} R_k^{\alpha-1}\bar{R}_k)$$
$$\frac{1}{n^{1/\alpha}}(X_1^b(n), \dots, X_k^b(n)) \to (D_1^{1/\alpha} R_1, \dots, D_k^{1/\alpha} R_k)$$
$$\frac{1}{\alpha n}(X_1^c(n), \dots, X_k^c(n)) \to (D_1, D_2, \dots, D_k),$$

where the sequences $(D_1, \dots, D_k)$, $(R_1, \dots, R_k)$ and $(\bar{R}_1, \dots, \bar{R}_k)$ are independent; we have $(D_1, \dots, D_k) \sim \mathrm{Dir}(\gamma_1/\alpha, \dots, \gamma_k/\alpha)$; the random variables $R_1, \dots, R_k$ are mutually independent with $R_i \sim \mathrm{ML}(1/\alpha, \gamma_i/\alpha)$; and the random variables $\bar{R}_1, \dots, \bar{R}_k$ are mutually independent with $\bar{R}_i \sim \mathrm{ML}(\alpha-1, \gamma_i)$.

The proof of Proposition 5.15 appears in Section 5.5.2.

Proof of Proposition 5.14. We make use of Marchal's algorithm. Recall that we are given an ordered tree $\mathsf{T}_k^{\mathrm{ord}}$ with k leaves labelled $1, 2, \dots, k$, m edges and r internal vertices with degrees $d_1, \dots, d_r$. Let us set

$$\gamma_1 = \cdots = \gamma_m = \alpha - 1$$

and

$$\gamma_{m+1} = d_1 - 1 - \alpha, \dots, \gamma_{m+r} = d_r - 1 - \alpha.$$

We then have $\sum_{i=1}^{m+r} \gamma_i = \alpha n - 1$.

We now show that the the urn process from Proposition 5.15 naturally occurs within our tree evolving according to Marchal's algorithm. Colours $1, 2, \dots, m$ represent the different edges of $\mathsf{T}_k^{\mathrm{ord}}$ and colours $m+1, \dots, m+r$ represent the different vertices. For edge e_i of $\mathsf{T}_k^{\mathrm{ord}}$, type a corresponds to the weight of edges inserted along e_i; type b corresponds to the weight at vertices along e_i and type c corresponds to the weight in vertices and edges in pendant subtrees hanging off e_i (excluding their roots along e_i). So $X_i^a(n) = (\alpha-1)L_n(e_i)$, $X_i^b(n) = N_n(e_i) + (1-\alpha)(L_n(e_i)-1)$ and $X^c(n) = \alpha M_n(e_i) - N_n(e_i)$. For vertex v_j of $\mathsf{T}_k^{\mathrm{ord}}$, types a and b together correspond to the weight at v_j and type c corresponds to the weight in edges and vertices in subtrees hanging from v_j. So $X_{m+j}^a(n) + X_{m+j}^b(n) = N_n(v_j) - 1 - \alpha$ and $X^c(n) = \alpha M_n(v_j) - N_n(v_j) + d_j$. Applying Proposition 5.15 and Lemma 5.13 then yields the claimed distributions for the $L(e_i)$, $N(e_i)$, $M(e_i)$, $N(v_j)$ and $M(v_j)$.

We now turn to $N_n(e_i, \ell), \ell \ge 1$, the ordered numbers of edges attached to the branchpoints along e_i. Independently for $1 \le i \le m$, let $(C_{i,\ell}(n), \ell \ge 1)$ be a Chinese restaurant process with $\beta = \theta = \alpha - 1$. This evolves in exactly the same way as Marchal's algorithm adds new edges along e_i. In particular, we have

$$(N_n(e_i, \ell), \ell \ge 1) = (C_{i,\ell}^{\downarrow}(N_n(e_i)), \ell \ge 1).$$

By again composing limits, it follows that

$$\left(\frac{N(e_i, \ell)}{N(e_i)}, \ell \ge 1\right) \sim \mathrm{PD}(\alpha-1, \alpha-1),$$

independently for $1 \le i \le m$ and independently of everything else.

Let us now consider how the local time is distributed among the corners of the vertices v_j. This again follows from an urn argument: for the vertex v_j which has degree d_j, consider an urn with d_j colours, one corresponding to each corner, $(A_{m+j,1}(n), \dots A_{m+j,d_j}(n))_{n\ge 0}$. Start the urn from a single ball of each colour. Then whenever we insert an edge into the corresponding corner, we increase the number of positions into which we can insert new edges by 1. Hence, we have precisely Pólya's urn (see Section 5.5 for a definition) and so by Theorem 5.30,

$$\frac{1}{n}(A_{m+j,1}(n), \dots, A_{m+j,d_j}(n)) \to (\Delta_1, \dots, \Delta_{d_j})$$

almost surely, where $(\Delta_1, \dots, \Delta_d) \sim \mathrm{Dir}(1, 1, \dots, 1)$. We have

$$(N_n(v_j, \ell), 1 \le \ell \le d_j) = (A_{m+j,\ell}(N_n(v_j)) - 1, 1 \le \ell \le d_j)$$

and it follows that

$$\left(\frac{N(v_j, \ell)}{N(v_j)}, 1 \le \ell \le d_j\right) \sim \mathrm{Dir}(1, 1, \dots, 1),$$

independently for $1 \le j \le r$ and independently of everything else.

A similar argument works for the local time to the left and right of the ℓth largest vertex along an edge e_i: start a two-colour urn $(A_{i,\ell,1}(n), A_{i,\ell,2}(n))_{n\ge 0}$ from one ball of each colour and at each step add a single ball of the picked colour. Then, again by Theorem 5.30,

$$\frac{1}{n}(A_{i,1}(n), A_{i,2}(n)) \to (\Delta, 1 - \Delta)$$

almost surely, where $\Delta \sim \mathrm{U}[0, 1]$. We get

$$N_n^{\text{right}}(e_i, \ell) = A_{i,2}(N^n(e_i, \ell)) - 1$$

and so it follows that

$$\frac{N^{\text{right}}(e_i, \ell)}{N(e_i, \ell)} \sim \mathrm{U}[0, 1],$$

independently for $1 \le i \le m$ and $\ell \ge 1$. □

Remark 5.16. *Let $N(T) := N(e_1) + \dots + N(e_m) + N(v_1) + \dots + N(v_r)$. Using Remark 5.33 below, we observe the following distributional relation: we have $N(T) \sim \mathrm{ML}(1/\alpha, k - 1/\alpha)$ and, independently,*

$$\left(\frac{N(e_1)}{N(T)}, \dots, \frac{N(e_m)}{N(T)}, \frac{N(v_1)}{N(T)}, \dots, \frac{N(v_r)}{N(T)}\right) \sim \mathrm{Dir}(\alpha - 1, \dots, \alpha - 1, d_1 - 1 - \alpha, \dots, d_r - 1 - \alpha).$$

5.2.2 Construction of the stable graphs

Construction from [38]. Returning now to the setting of our graphs, we wish to specify the distribution of the limiting sequence $\mathrm{C}_i = (C_i, d_{C_i}, \mu_{C_i})$, $i \ge 1$ arising in Theorem 5.1. The details of the following can be found in the paper [38]. Our graph notation was introduced in Section 5.1.1 and the processes ξ, X, H were introduced in Section 5.2.1.

We define a real-valued process $\tilde{\xi}$ via a change of measure from the Lévy process ξ. To this end, we observe first that $\left(\exp\left(\int_0^t s\mathrm{d}\xi_s - \frac{t^{\alpha+1}}{(\alpha+1)}\right), t \ge 0\right)$ is a martingale. Now for each $t \ge 0$ and any suitable test-function $f : \mathbb{D}([0, t], \mathbb{R}) \to \mathbb{R}$, define $\tilde{\xi}$ by

$$\mathbb{E}f(\tilde{\xi}_s, 0 \le s \le t) = \mathbb{E}\exp\left(\int_0^t s\mathrm{d}\xi_s - \frac{t^{\alpha+1}}{\alpha + 1}\right) f(\xi_s, 0 \le s \le t).$$

Superimpose a Poisson point process of rate A_α^{-1} (as defined in (5.2)) in the region $\{(t,y) \in \mathbb{R}_+ \times \mathbb{R}_+ : y \leq \tilde{\xi}_t - \inf_{0\leq s\leq t} \tilde{\xi}_s\}$. Then the limiting components $\mathrm{C}_i, i \geq 1$ are encoded by the excursions of the reflected process $(\tilde{\xi}_t - \inf_{0\leq s\leq t} \tilde{\xi}_s, t \geq 0)$ above 0 and the Poisson points falling under each such excursion. The total masses of the measures $\mu_{C_1}(C_1), \mu_{C_2}(C_2), \ldots$ are given by the lengths of the excursions of $\tilde{\xi}$ above its running infimum. The surpluses $s(\mathrm{C}_1), s(\mathrm{C}_2), \ldots$ are given by the the number of Poisson points falling under corresponding excursions. Then, the limiting components $(\mathrm{C}_1, \mathrm{C}_2, \ldots)$ are conditionally independent given the sequences $(\mu_{C_1}(C_1), \mu_{C_2}(C_2), \ldots)$ and $(s(\mathrm{C}_1), s(\mathrm{C}_2), \ldots)$, with

$$\left(C_i, d_{C_i}, \mu_{C_i}\right) \overset{(\mathrm{d})}{=} \left(\mathcal{G}^{s(\mathrm{C}_i)}, \mu_{C_i}(C_i)^{1-1/\alpha} \cdot d^{s(\mathrm{C}_i)}, \mu_{C_i}(C_i) \cdot \mu^{s(\mathrm{C}_i)}\right).$$

Construction of the connected α-stable graph with surplus s. For $s \geq 0$, it remains to describe the connected stable graph, $\mathcal{G}^s$. First sample excursions X^s and H^s with joint law specified by

$$\mathbb{E} f(X^s(t), H^s(t), 0 \leq t \leq 1) = \frac{\mathbb{E}\left(\int_0^1 X(u)\mathrm{d}u\right)^s f(X(t), H(t), 0 \leq t \leq 1)}{\mathbb{E}\left(\int_0^1 X(u)\mathrm{d}u\right)^s}.$$

Let $\mathcal{T}^s$ be the $\mathbb{R}$-tree encoded by H^s and let $\pi^s : [0,1] \to \mathcal{T}^s$ be its canonical projection. If $s = 0$, then X^s is a standard stable excursion and H^s is its corresponding height process i.e. $\mathcal{T}^0 \overset{(\mathrm{d})}{=} \mathcal{T}$. In this case, we simply set $\mathcal{G}^0 = \mathcal{T}^0$. If, on the other hand, $s \geq 1$, conditionally on X^s and H^s, sample conditionally independent points $V_1^s, V_2^s, \ldots, V_s^s$ from $[0,1]$, each having density

$$\frac{X^s(u)}{\int_0^1 X^s(t)\mathrm{d}t}, \quad u \in [0,1].$$

Then, for $1 \leq k \leq s$, let Y_k^s be uniformly distributed on the interval $[0\,, X^s(V_k^s)]$, independently for all k, and let $B_k^s = \inf\{t \geq V_k^s : X^s(t) = Y_k^s\}$. We obtain $\mathcal{G}^s$ from $\mathcal{T}^s$ by identifying the pairs of points $(\pi^s(V_k^s), \pi^s(B_k^s))$ for $1 \leq k \leq s$. (This is achieved formally by a further straightforward quotienting operation which we do not detail here.)

In fact, using the notation of Section 5.2.1 for the tree $\mathcal{T}^s$ which is absolutely continuous with respect to $\mathcal{T}$, this last operation corresponds to identifying the leaf $\pi^s(V_k^s)$ with a branchpoint on its ancestral line $]\!]\rho, \pi^s(V_k^s)[\![$, independently for $1 \leq k \leq s$. As a consequence of the discussion in Section 5.2.1, the point $\pi^s(B_k)$ is such that

$$\pi^s(B_k^s) = \pi^s(A_k^s), \quad \text{where} \quad A_k^s = \sup\left\{t \leq V_k^s : X^s(t) \leq \inf\{X^s(u) : t \leq u \leq Y_k^s\}\right\}.$$

Along with equation (5.13), this ensures that each branchpoint $b \in\,]\!]\rho, \pi^s(V_k^s)[\![$ is chosen with probability equal to

$$\frac{N^{\text{right}}(b, \pi^s(V_k^s))}{N^{\text{right}}(]\!]\rho, \pi^s(V_k^s)[\![)} = \frac{N^{\text{right}}(b, \pi^s(V_k^s))}{X(V_k^s)},$$

as claimed in the introduction. We view $\mathcal{G}^s$ as a measured metric space by endowing it with the image of the Lebesgue measure on $[0,1]$ by the projection π^s.

Continuous and discrete marginals. Recall the definition for any $n \geq 0$ of the continuous marginals $\mathcal{G}_n^s$ from the introduction: $\mathcal{G}_n^s$ is the union of the kernel $\mathcal{K}^s$ and the paths from n leaves to the root, where the leaves are taken i.i.d under the measure carried by $\mathcal{G}^s$. Indeed, the kernel is the image of the subtree of $\mathcal{T}^s$ spanned by the s selected leaves after the gluing procedure.

Let $(U_i)_{i\geq 1}$ be a sequence of i.i.d. U$[0,1]$ random variables independent of X^s, and let $n \geq 0$. In the construction described above, let $\mathcal{T}^s_{s,n}$ be the ordered subtree of $\mathcal{T}^s$ spanned by the root and the leaves corresponding to the real numbers $V^s_1, \dots, V^s_s, U_1, \dots, U_n$, and $\mathcal{T}^{s,\mathrm{ord}}_{s,n}$ its ordered version. Since $\pi^s(U_1), \dots, \pi^s(U_n)$ are (by definition) distributed according to the probability measure carried by $\mathcal{G}^s$, the image of $\mathcal{T}^s_{s,n}$ after the gluing procedure is a version of the continuous marginal $\mathcal{G}^s_n$ (and the discrete marginal G^s_n is then the combinatorial shape of the continuous marginal $\mathcal{G}^s_n$).

For future purposes, we also define $\mathsf{T}^{s,\mathrm{ord}}_{s,n}$ the discrete counterpart of $\mathcal{T}^{s,\mathrm{ord}}_{s,n}$. By convention, we consider that the s first leaves are unlabelled and the n leaves corresponding to $U_1, \dots, U_n$ inherit the label of their uniform variable.

Unbiasing. Let $(X; V_1, V_2, \dots, V_s, Y_1, \dots, Y_s)$ be the unbiased excursion endowed with

- $V_1, \dots, V_s$ i.i.d. U$[0,1]$ random variables
- $Y_1, \dots, Y_s$ which are conditionally independent given $(X; V_1, V_2, \dots V_s)$, with $Y_k \sim$ U$[0, X(V_k)]$.

We call $(X; V_1, V_2, \dots V_s, Y_1, \dots, Y_s)$ the *unbiased counterpart* of $(X^s; V^s_1, \dots, V^s_s, Y^s_1, \dots Y^s_s)$. Any random object defined as a measurable function $f(X^s; (V^s_k)_{1\leq k\leq s}, (Y^s_k)_{1\leq k\leq s}, (U_i)_{i\geq 1})$ then also has an unbiased counterpart, $f(X; (V_k)_{1\leq k\leq s}, (Y_k)_{1\leq k\leq s}, (U_i)_{i\geq 1})$ and vice versa. Using the fact that, conditionally on $(X; V_1, V_2, \dots V_s)$, the random variables $Y_1, \dots, Y_s$ have the same distribution as $Y^s_1, \dots, Y^s_s$ conditionally on $(X^s; V^s_1, V^s_2, \dots V^s_s)$, we observe that

$$\begin{aligned}
&\mathbb{E} f(X^s; (V^s_k)_{1\leq k\leq s}, (Y^s_k)_{1\leq k\leq s}, (U_i)_{i\geq 1}) \\
&= \frac{\mathbb{E}\int_{[0,1]^s} \mathrm{d}v_1 \dots \mathrm{d}v_s \frac{X(v_1)\dots X(v_s)}{\left(\int_0^1 X(t)\mathrm{d}t\right)^s} \int_0^{X(v_1)} \frac{\mathrm{d}y_1}{X(v_1)} \dots \int_0^{X(v_s)} \frac{\mathrm{d}y_s}{X(v_s)} f(X; (v_k), (y_k), (U_i)) \left(\int_0^1 X(t)\mathrm{d}t\right)^s}{\mathbb{E}\left(\int_0^1 X(t)\mathrm{d}t\right)^s} \\
&= \frac{\mathbb{E} f(X; (V_k)_{1\leq k\leq s}, (Y_k)_{1\leq k\leq s}, (U_i)_{i\geq 1}) X(V_1)X(V_2)\dots X(V_s)}{\mathbb{E} X(V_1)X(V_2)\dots X(V_s)}. \qquad (5.18)
\end{aligned}$$

In particular, this allows us to compute quantities in the unbiased setting in order to understand the biased one. We define $\widehat{\mathcal{G}}^s$ to be the unbiased counterpart of $\mathcal{G}^s$ and $\widehat{\mathcal{G}}^s_n$ to be the unbiased counterpart of $\mathcal{G}^s_n$ and $\widehat{\mathsf{G}}^s_n$ to be the unbiased counterpart of G^s_n. Similarly, $\widehat{\mathcal{T}}^{s,\mathrm{ord}}_{s,n}$ is the unbiased counterpart of $\mathcal{T}^{s,\mathrm{ord}}_{s,n}$ which, modulo the labelling of the leaves, has the same distribution as $\mathcal{T}^{\mathrm{ord}}_{s+n}$.

5.3 Distribution of the marginals G^s_n

Let $s \geq 0$. The goal of this section is to identify the joint distribution of the marginals G^s_n, for $n \geq 0$ (and for $n \geq -1$ if $s \geq 2$). By definition, for any $n \geq 0$, the random graph G^s_n is an element of $\mathbb{M}_{s,n}$, the set of connected multigraphs with surplus s, with $n+1$ labelled leaves, unlabelled internal vertices and no vertex of degree 2. To perform our calculations, it will be convenient to consider versions of this multigraph with some additional structure, namely cyclic orderings of the half-edges around each vertex. We denote by $\mathbb{M}^{\mathrm{ord}}_{s,n}$ the set of such graphs and we emphasise here that the orderings around different vertices need not be compatible with one another: the elements of $\mathbb{M}^{\text{ord}}_{s,n}$ are not necessarily planar. The advantage is that this additional structure breaks the symmetries present in elements of $\mathbb{M}_{s,n}$. (For $n=-1$ the cyclic ordering is insufficient to break all the symmetries and we will rather label the internal vertices.)

We will begin in Section 5.3.1 by computing the number of possible cyclic orderings of the half-edges around the different vertices of a graph $G \in \mathbb{M}_{s,n}$. Then, in Section 5.3.2, we will describe the elements of $\mathbb{M}^{\text{ord}}_{s,n}$ as ordered trees with n labelled and s unlabelled leaves together with a "gluing plan", that specifies how to glue each unlabelled leaf "to the right" of the ancestral path of that leaf. This description corresponds to the one we have for G^s_n, and we compute in Section 5.3.3 the distribution of the tree and the corresponding gluing plan, which then yields the distribution of G^s_n claimed in Theorem 5.3. In Section 5.3.4, we show that the sequence $(\mathsf{G}^s_n)_{n\geq 0}$ evolves according to Marchal's algorithm (Theorem 5.5). In Section 5.3.5, we extend this to $(\mathsf{G}^s_n)_{n\geq -1}$ for $s\geq 2$. Finally, Section 5.3.6 is devoted to the proof of Corollary 5.4, which identifies the distribution of G^s_n with that of a specific configuration model with i.i.d. random degrees.

We recall the following notation from the introduction. For each $G=(V(G),E(G))\in\mathbb{M}_{s,n}$, we denote $I(G)\subset V(G)$ the set of internal vertices of G (vertices of degree 3 or more), $\deg(v)=\deg_G(v)$ the degree of a vertex $v\in V(G)$, $\mathrm{sl}(G)$ the number of self-loops, $\mathrm{mult}(e)$ the multiplicity of the element $e\in\mathrm{supp}(E)$ and $\mathrm{Sym}(G)$ the set of permutations of vertices of G that are the identity on the leaves and that preserve the adjacency relations (with multiplicity).

5.3.1 Cyclic orderings of half-edges

Let $n\geq 0$. In this section we compute the number of possible cyclic orderings of the half-edges around each vertex of G, for each $G\in\mathbb{M}_{s,n}$ (we emphasise that Lemma 5.17 is false when $n=-1$ and $s\geq 2$). Let $\psi:\mathbb{M}^{\text{ord}}_{s,n}\to\mathbb{M}_{s,n}$ be the map that forgets the cyclic ordering around the vertices.

Lemma 5.17. *For each $G\in\mathbb{M}_{s,n}$,*

$$\left|\psi^{-1}(G)\right| = \frac{\prod_{v\in I(G)}(\deg(v)-1)!}{|\mathrm{Sym}(G)|2^{\mathrm{sl}(G)}\prod_{e\in\mathrm{supp}(E(G))}\mathrm{mult}(e)!}.$$

Proof. It is convenient to consider versions of G with labelled internal vertices. The number of possible labellings is

$$\frac{|I(G)|!}{|\mathrm{Sym}(G)|}. \tag{5.19}$$

Indeed, let $\tilde{G}$ denote an arbitrarily labelled version of G. The symmetric group $\mathfrak{S}_{|I(G)|}$ acts on the set of multigraphs with $|I(G)|$ internal labels by permuting those labels. The number of labellings we seek is thus the number of elements of the orbit of $\tilde{G}$ under this action. This is just $|I(G)|!$ divided by the cardinality of the stabilizer of $\tilde{G}$. Any permutation $\sigma\in\mathfrak{S}_{|I(G)|}$ that fixes $\tilde{G}$ corresponds to a permutation $\tau\in\mathrm{Sym}(G)$, hence the result.

Now, to compute $\left|\psi^{-1}(G)\right|$, we first label everything then forget the labels we do not need.

- Consider a version of G with labelled internal vertices: from the preceding paragraph, there are $\frac{|I(G)|!}{|\mathrm{Sym}(G)|}$ possible labellings.
- For each $e=\{u,v\}\in\mathrm{supp}(E(G))$, in order to distinguish between the $\mathrm{mult}(e)$ edges joining u and v, number them from 1 to $\mathrm{mult}(e)$.
- Give every self-loop an orientation.

- Endow the multigraph with a cyclic ordering around each vertex. For each $v \in I(G)$ we have $(\deg(v) - 1)!$ possibilities for an ordering of the half-edges adjacent to v. (The half-edges are distinguishable because the self-loops are oriented.)

- Forget the orientation on the self-loops. This transformation is $2^{\mathrm{sl}(G)}$-to-1 since with the ordering around the vertices, every orientation is distinguishable.

- Forget the labelling of the edges. This transformation is $\left(\prod_{e \in \mathrm{supp}(E(G))} \mathrm{mult}(e)!\right)$-to-1.

- Forget the labelling of the internal vertices. With the cyclic ordering around the vertices every vertex is distinguishable, and so this map is $|I(G)|!$-to-1.

(We emphasise here the importance of the fact that our multigraphs are planted in distinguishing edges and vertices.) We obtain a multigraph in $\mathbb{M}_{s,n}^{\mathrm{ord}}$ whose image by ψ is G. By the previous considerations, the number of such multigraphs is indeed given by the claimed formula. $\square$

5.3.2 Ordered multigraphs and the depth-first tree

We still consider integers $n \geq 0$.

Ordered trees with paired leaves. Let $\mathbb{A}_{s,n}$ be the set of planted ordered trees with no vertices of degree 2 that have s unlabelled leaves and n labelled leaves, with labels from 1 to n. Let $\mathbb{A}_{s,n}^{\mathrm{pair}}$ be the set of ordered trees with no vertices of degree 2 that have n labelled uncoloured leaves, s red leaves labelled 1 to s in clockwise order from the root, and s blue leaves also labelled from 1 to s. We think of the red and blue leaves labelled i as forming a pair, and impose the condition that the blue leaf labelled i must lie to the right of the ancestral line of the red leaf labelled i, for $1 \leq i \leq s$.

We first describe how every ordered multigraph $G \in \mathbb{M}_{s,n}^{\mathrm{ord}}$ is equivalent to an element of $\mathbb{A}_{s,n}^{\mathrm{pair}}$. We define two natural maps on $\mathbb{A}_{s,n}^{\mathrm{pair}}$. Let

$$\textsc{Glue} : \mathbb{A}_{s,n}^{\mathrm{pair}} \to \mathbb{M}_{s,n}^{\mathrm{ord}}$$

be the map that, for each red leaf i identifies i with its blue pair and then contracts the resulting path containing a vertex of degree 2 into a single edge. Let

$$\textsc{Erase} : \mathbb{A}_{s,n}^{\mathrm{pair}} \to \mathbb{A}_{s,n}$$

be the map that erases the blue leaves and their adjacent edges, then contracts any path of degree 2 vertices into a single edge, and finally forgets the labelling and colour of the red leaves.

Reverse construction: the depth-first tree. Let $G \in \mathbb{M}_{s,n}^{\mathrm{ord}}$. We imagine that each edge of G is made up of two half-edges, one attached to each end-point. We say that two half-edges are *adjacent* if they have a common end-point. We describe a procedure that explores all the half-edges of the graph in a deterministic manner and disconnects exactly s edges in order to transform G into a tree. At each step i of the algorithm, we will have an ordered stack of *active* half-edges A_i and a current surplus s_i. We write h_0 for the unique half-edge connected to the leaf with label 0.

Initialization $A_0 = (h_0)$, $s_0 = 0$.

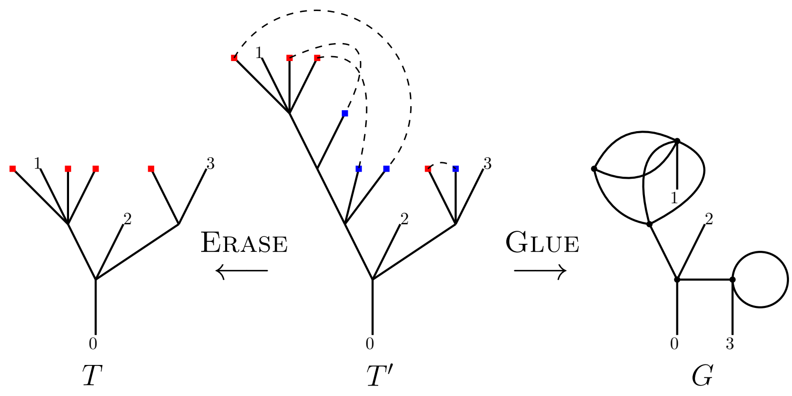


Figure 5.4 – The operations GLUE and ERASE applied to a tree T'. Here, T' is the depth-first tree of G, and T is the base tree.

STEP i $(0 \le i \le |E(G)| - 1)$: Let h_i be the half-edge at the top of the stack A_i. Let $\hat{h}_i$ be the half-edge to which it is attached. If $\hat{h}_i \notin A_i$, remove h_i from the stack and put the half-edges adjacent to $\hat{h}_i$ on the top of the stack, in clockwise order top to bottom. If $\hat{h}_i \in A_i$, first increment s_i, then remove both h_i and $\hat{h}_i$ from the stack, disconnect them, attach a red leaf labelled s_i to h_i and attach a blue leaf labelled s_i to $\hat{h}_i$.

It is straightforward to check that this algorithm produces a tree in $\mathbb{A}_{s,n}^{\mathrm{pair}}$, which we call the *depth-first tree*, and denote by $\mathrm{DEP}(G)$. (Note that this is a variant of the notion of depth-first tree introduced in [5].) We have $\mathrm{DEP}(G) = G$ if and only if G is a tree i.e. $s = 0$. The following lemma is then straightforward.

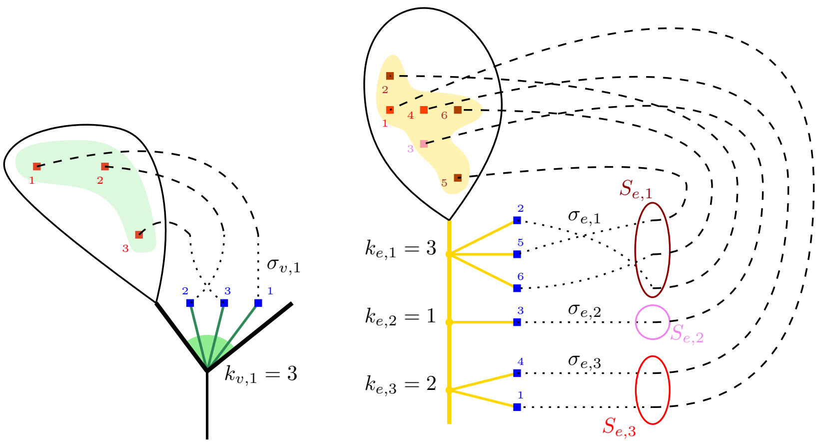
Lemma 5.18. *The maps* $\mathrm{GLUE} : \mathbb{A}_{s,n}^{\mathrm{pair}} \to \mathbb{M}_{s,n}^{\mathrm{ord}}$ *and* $\mathrm{DEP} : \mathbb{M}_{s,n}^{\mathrm{ord}} \to \mathbb{A}_{s,n}^{\mathrm{pair}}$ *are reciprocal bijections.*

For a multigraph G, call $\mathrm{ERASE}(\mathrm{DEP}(G))$ the *base tree.*

Gluing plans. Consider $T \in \mathbb{A}_{s,n}$. We now aim to describe the set $\mathrm{ERASE}^{-1}(\{T\})$. This is the set of possible depth-first trees T' obtainable from a fixed base tree T. As usual, we write $I(T)$ for the internal vertices of T and $E(T)$ for its edges. A vertex $v \in I(T)$ of degree $d = \deg_T(v)$ possesses d *corners*, which we call $c_{v,1}, \dots, c_{v,d}$ in clockwise order from the root. We write $C(T)$ for the set of corners of T. The *ancestral path* of a vertex is its unique path to the root. For the kth unlabelled leaf of T in clockwise order, let $\mathcal{A}(k)$ be the set of edges and corners that lie immediately to the right of its ancestral path, for $1 \le k \le s$.

Now let $T' \in \mathrm{ERASE}^{-1}(\{T\})$. The internal vertices of T each have a counterpart in T', for which we use the same name. The red leaves of T' correspond to the unlabelled leaves of T. A blue leaf is attached by its incident edge either into one of the corners of an internal vertex of T, or to an internal vertex of T' which disappears when the blue leaves are removed and paths of internal vertices of degree 2 are contracted into a single edge. For each $e \in E(T)$ let a_e be the number of additional vertices along the path in T' which get contracted to yield the edge e by ERASE. If $a_e \neq 0$, we will list these additional vertices as $v_{e,1}, \dots, v_{e,a_e}$ in decreasing order of distance from the root.

For each $v \in I(T)$, let $S_{v,\ell}$ be the set of labels of blue leaves attached to corner $c_{v,\ell}$, for $1 \le \ell \le \deg_T(v)$. (Any or all of these sets may be empty; in particular, $S_{v,1}$ is always empty because a blue leaf must lie to the right of the ancestral line of the corresponding red leaf.) If $S_{v,\ell}$ is non-empty, let $\sigma_{v,\ell}$ be the permutation of its elements which gives the clockwise ordering



(a) In a corner (b) On an edge

Figure 5.5 – Definition of a gluing plan

of the blue leaves in corner $c_{v,\ell}$; if it is empty, let $\sigma_{v,\ell}$ be the unique permutation of the empty set. For each $e \in E(T)$ such that $a_e \neq 0$, we let $S_{e,i}$ be the set of labels of blue leaves attached to vertex $v_{e,i}$ in T', for $1 \leq i \leq a_e$. These sets can not be empty. Let $\sigma_{e,i}$ be the permutation of the elements of $S_{e,i}$ giving the clockwise ordering of the blue leaves attached to $v_{e,i}$ (note that these are necessarily attached to the right of e). Observe that the collection of sets

$$\{S_{v,\ell} : v \in I(T), 1 \leq \ell \leq \deg_T(v), S_{v,\ell} \neq \emptyset\} \cup \{S_{e,i} : e \in E(T), 1 \leq i \leq a_e\}$$

partitions $\{1, 2, \ldots, s\}$. This induces a *gluing function* $g : \{1, 2, \ldots, s\} \to (I(T) \cup E(T)) \times \mathbb{N}$ as follows. For $1 \leq k \leq s$, if $k \in S_{v,\ell}$ set $g(k) = (v, \ell)$; if $k \in S_{e,i}$ set $g(k) = (e, i)$.

See Figure 5.5 for an illustration. This leads us to the formal definition of a gluing plan.

Definition 5.19. *We say that* $\Delta = \Big(\big((S_{v,\ell}, \sigma_{v,\ell})_{1 \leq \ell \leq \deg_T(v)} \big)_{v \in I(T)}, \big((S_{e,i}, \sigma_{e,i})_{1 \leq i \leq a_e} \big)_{e \in E(T)} \Big)$ *is a* gluing plan *for* T *if the following properties are satisfied.*

(i) *For all* $v \in I(T)$ *and all* $1 \leq \ell \leq \deg_T(v)$, *we have* $S_{v,\ell} \subseteq \{1, 2, \ldots, s\}$ *and* $\sigma_{v,\ell}$ *is a permutation of* $S_{v,\ell}$.

(ii) *For all* $e \in E$ *and all* $1 \leq i \leq a_e$, *the set* $S_{e,i} \subseteq \{1, 2, \ldots, s\}$ *is non-empty and* $\sigma_{e,i}$ *is a permutation of* $S_{e,i}$.

(iii) *The sets* $\{S_{v,\ell} : v \in I(T), 1 \leq \ell \leq \deg_T(v), S_{v,i} \neq \emptyset\}$ *and* $\{S_{e,i} : e \in E(T), 1 \leq i \leq a_e\}$ *partition* $\{1, 2, \ldots, s\}$.

(iv) *The induced gluing function* $g : \{1, 2, \ldots, s\} \to (I(T) \cup E(T)) \times \mathbb{N}$ *is such that if* $g(k) = (v, \ell)$ *then* $c_{v,\ell} \in \mathcal{A}(k)$ *and if* $g(k) = (e, i)$ *then* $e \in \mathcal{A}(k)$, *for all* $1 \leq k \leq s$.

It is straightforward to see that we can completely encode a tree $T' \in \text{ERASE}^{-1}(\{T\})$ by its gluing plan, and that conversely, every gluing plan for T encodes a tree $T' \in \text{ERASE}^{-1}(\{T\})$.

Lemma 5.20.

$$\mathbb{M}^{\text{ord}}_{s,n} \simeq \mathbb{A}^{\text{pair}}_{s,n} \simeq \{(T, \Delta) \mid T \in \mathbb{A}_{s,n} \textit{ and } \Delta \textit{ is a gluing plan for } T\}.$$

Suppose $T \in \mathbb{A}_{s,n}$ and that

$$\Delta = \Big(\big((S_{v,\ell}, \sigma_{v,\ell})_{1 \leq \ell \leq \deg_T(v)} \big)_{v \in I(T)}, \big((S_{e,i}, \sigma_{e,i})_{1 \leq i \leq a_e} \big)_{e \in E(T)} \Big)$$

is a gluing plan for the base tree T. We let $k_{v,\ell} = |S_{v,\ell}|$ be the number of blue leaves attached into corner $c_{v,\ell}$ and $k_v = \sum_{\ell=1}^{\deg_T(v)} k_{v,\ell}$ be the total number of blue leaves attached to v. We let $k_{e,i} = |S_{e,i}|$ be the number of blue leaves attached to the ith vertex inserted along e and let $k_e = \sum_{i=1}^{a_e} k_{e,i}$ be the total number of blue leaves attached to vertices along e. We call the family of numbers

$$\big((k_v, (k_{v,\ell})_{1\le \ell\le \deg_T(v)})_{v\in I(T)}, (k_e, a_e, (k_{e,i})_{1\le i\le a_e})_{e\in E(T)}\big)$$

the *type* of the gluing plan Δ.

Remark 5.21. *Suppose that $G \in \mathbb{M}_{s,n}^{\mathrm{ord}}$ corresponds to (T, Δ). The degrees in G depend only on T and the type of the gluing plan Δ. For an internal vertex v of G that was already present in $I(T)$, its degree in G is $\deg_G(v) = \deg_T(v) + k_v$. The internal vertices of G that do not correspond to internal vertices of T are the ones that were created along the edges of T during the gluing procedure. For each $e \in E(T)$, there are a_e newly-created vertices along the edge e, having degrees $2 + k_{e,1}, 2 + k_{e,2}, \dots, 2 + k_{e,a_e}$.*

5.3.3 The distribution of G_n^s

The goal of this section is to prove Theorem 5.3 for $n \ge 0$, which states that for every connected multigraph $G \in \mathbb{M}_{s,n}$,

$$\mathbb{P}\left(\mathsf{G}_n^s = G\right) \propto \frac{\prod_{v\in I(G)} w_{\deg(v)-1}}{|\mathrm{Sym}(G)| 2^{\mathrm{sl}(G)} \prod_{e\in \mathrm{supp}(E(G))} \mathrm{mult}(e)!},$$

where the weights $(w_k)_{k\ge 0}$ are defined in (5.6).

Recall the construction of the random graph $\mathcal{G}^s$ using a tilted excursion and biased chosen points $(X^s; V_1^s, \dots, V_s^s)$ from Section 5.2.2. Recall also the definitions of $\mathcal{T}_{s,n}^{s,\mathrm{ord}}$ (and its discrete version $\mathsf{T}_{s,n}^{s,\mathrm{ord}}$) and $\mathcal{G}_{s,n}^s$ (and its discrete version G_n^s), using an additional sequence of i.i.d. uniform random variables $(U_i)_{i\ge 1}$. In order to apply the results of the previous section, we want to work with ordered versions of our graphs. In particular, we will get an ordered version $\mathsf{G}_n^{s,\mathrm{ord}}$ of G_n^s by applying a gluing plan to the base tree $\mathsf{T}_{s,n}^{s,\mathrm{ord}}$. The change of measure (5.18) allows us to make calculations using the unbiased excursion with uniform points $(X; V_1, \dots, V_s, U_1, \dots, U_n)$. So we will define and work instead with an unbiased version $\widehat{\mathsf{G}}_n^{s,\mathrm{ord}}$, derived from the unbiased version $\widehat{\mathsf{T}}_{s,n}^{s,\mathrm{ord}}$ of $\mathsf{T}_{s,n}^{s,\mathrm{ord}}$.

Construction of $\widehat{\mathsf{G}}_n^{s,\mathrm{ord}}$. We define $\widehat{\mathsf{G}}_n^{s,\mathrm{ord}}$ via a random gluing plan Δ for $\widehat{\mathsf{T}}_{s,n}^{s,\mathrm{ord}}$. Conditionally on $\widehat{\mathsf{T}}_{s,n}^{s,\mathrm{ord}} = T \in \mathbb{T}_{s,n}$, let

$$W(T) := \{(v,\ell) : v \in I(T),\ 1 \le \ell \le \deg_T(v)\} \cup \{(e,j) : e \in E(T),\ j \ge 1\} \subset (I(T) \cup E(T)) \times \mathbb{N}.$$

This indexes all the atoms of local time in the corners (as usual, ordered clockwise around each internal vertex) and along the edges (ordered by decreasing local time in this instance) of the ordered tree $\mathcal{T}_{s,n}^{s,\mathrm{ord}}$. We will often abuse notation and think of the elements of $W(T)$ as the atoms themselves. In fact, the tree $\mathcal{T}_{s,n}^{s,\mathrm{ord}}$ has, up to the labelling of the leaves, the same distribution as $\mathcal{T}_{s+n}^{\mathrm{ord}}$, so using the discussion just before Lemma 5.13, we can decompose the whole (unbiased) stable tree as

$$\mathcal{T}_{s,n}^{\mathrm{ord}} \cup \bigcup_{w\in W(T)} \mathcal{T}(w).$$

In order to define our gluing plan, we need to be a little careful about labelling. For $1 \le k \le s$, let $l_k \in \{1, 2, \dots, s\}$ be the position of V_k in the increasing ordering of $V_1, \dots, V_s$ i.e. $l_k = \#\{1 \le j \le s : V_j \le V_k\}$. This gives the relative planar position of the (unlabelled) leaf in T corresponding to V_k. Almost surely, the value $B_k = \inf\{t \ge V_k : X(t) = Y_k\}$ is such that there exists an element $w_k \in W(T)$ along the ancestral line of l_k, such for ϵ small enough, the canonical projection of an ϵ-neighbourhood around B_k lies completely within some subtree hanging off $\mathcal{T}_{s,n}^{\text{ord}}$ i.e.

$$\pi\left((B_k - \epsilon\,, B_k + \epsilon)\right) \subset \mathcal{T}(w_k).$$

For $1 \le k \le s$, for the jth largest atom of local time along an edge $e \in \mathcal{A}(l_k)$ and every corner $(v, \ell) \in \mathcal{A}(l_k)$ on the right of the ancestral path of the root to l_k, conditionally on $(X; V_1, V_2, \dots V_s, U_1, U_2, \dots U_n)$ we have

$$w_k = \begin{cases} (v, \ell) & \text{with probability } \frac{N(v,\ell)}{X(V_k)}, \text{ for } v \in I(T), 1 \le \ell \le \deg_T(v), \\ (e, j) & \text{with probability } \frac{N^{\text{right}}(e,j)}{X(V_k)}, \text{ for } e \in E(T), j \ge 1, \end{cases}$$

independently for all k. For each edge $e \in E(T)$, let a_e be the number of distinct atoms of local time which appear among $w_1, \dots, w_s$. If $a_e \ge 1$, we denote by $j_1, j_2, \dots j_{a_e}$ the values in the set $\{j \ge 1 : (e, j) \in \{w_1, \dots, w_s\}\}$ (that is, the indices of the atoms along e that receive at least one gluing) listed now in decreasing order of height i.e. such that $L(e, j_1) > L(e, j_2) > \cdots > L(e, j_{a_e})$. The probability that for any fixed set $\{j_1, \dots, j_{a_e}\}$ of distinct indices we have $L(e, j_1) > L(e, j_2) > \cdots > L(e, j_{a_e})$ is $1/a_e!$, since the random variables $L(e, j_1), \dots, L(e, j_{a_e})$ are exchangeable and distinct with probability 1, by Proposition 5.14. Moreover, again by Proposition 5.14, these random variables are independent of the local times. For $1 \le k \le s$, let

$$g(l_k) = \begin{cases} (v, \ell) & \text{if } w_k = (v, \ell) \text{ for some } v \in I(T) \text{ and some } 1 \le \ell \le \deg_T(v), \\ (e, i) & \text{if } w_k = (e, j_i) \text{ for some } e \in E(T) \text{ and some } 1 \le i \le a_e. \end{cases}$$

This is the required gluing function for T. We now derive the full gluing plan. For $e \in E(T)$ such that $a_e \ge 1$ and $1 \le i \le a_e$, let $S_{e,i} = g^{-1}(\{(e, i)\})$ be the set of leaves mapped to the ith atom in decreasing order of height along the edge e. Define a permutation $\sigma_{e,i}$ of $S_{e,i}$ by

$$\sigma_{e,i}(l_k) = \#\{1 \le j \le s : l_j \in S_{e,i}, Y_j \ge Y_k\}.$$

Similarly, for any $(v, \ell) \in C(T)$, we define $S_{v,\ell} = g^{-1}(\{(v, \ell)\})$ and a permutation $\sigma_{v,\ell}$ of $S_{v,\ell}$ by

$$\sigma_{v,\ell}(l_k) = \#\{1 \le j \le s : l_j \in S_{v,\ell}, Y_j \ge Y_k\}.$$

Since $Y_1, \dots, Y_k$ are conditionally independent given $(X; V_1, \dots, V_s, U_1, \dots, U_n)$, we see that the permutations are conditionally independent. Conditionally on corresponding to the same atom of local time, the relative ordering of the associated Y_k's is uniform, so that the permutations are all uniform on their label-sets. By construction,

$$\Delta = \left(\left((S_{v,\ell}, \sigma_{v,\ell})_{1 \le \ell \le \deg_T(v)} \right)_{v \in I(T)}, \left((S_{e,i}, \sigma_{e,i})_{1 \le i \le a_e} \right)_{e \in E(T)} \right)$$

is a gluing plan for T. We call $\widehat{\mathsf{G}}_n^{s,\text{ord}}$ the corresponding (random) multigraph in $\mathbb{M}_{s,n}^{\text{ord}}$, obtained via the bijection of Lemma 5.20.

For $n \ge 1$, let $\mathbb{N}^{n,\neq} = \{(j_1, \dots, j_n) \in \mathbb{N}^n : \ j_1, j_2, \dots, j_n \text{ are distinct}\}$.

Proposition 5.22. *Fix $T \in \mathbb{A}_{s,n}$ and suppose that $G \in \mathbb{M}^{\mathrm{ord}}_{s,n}$ is obtained from T by a gluing plan Δ. Conditionally on $(X; V_1, \ldots, V_s, U_1, \ldots, U_n)$ such that $\widehat{\mathsf{T}}^{s,\mathrm{ord}}_{s,n} = T$, the probability that $\widehat{\mathsf{G}}^{s,\mathrm{ord}}_n$ is equal to G depends only on the type of the gluing plan Δ. Indeed, for any gluing plan of type*

$$\Big(\big(k_v, (k_{v,\ell})_{1\le \ell \le \deg_T(v)}\big)_{v \in V(T)}, (k_e, a_e, (k_{e,i})_{1\le i\le a_e})_{e\in E(T)} \Big),$$

this conditional probability is

$$\frac{1}{X(V_1)X(V_2)\ldots X(V_s)} \left(\prod_{v\in I(T)} \prod_{\ell=1}^{\deg_T(v)} \frac{N(v,\ell)^{k_{v,\ell}}}{k_{v,\ell}!} \right) \cdot \left(\prod_{e\in E(T)} \sum_{(j_1,\ldots,j_{a_e})\in \mathbb{N}^{a_e,\neq}} \frac{1}{a_e!} \prod_{i=1}^{a_e} \frac{N^{\mathrm{right}}(e,j_i)^{k_{e,i}}}{k_{e,i}!} \right). \tag{5.20}$$

Proof. We reason conditionally on $(X; V_1, \ldots, V_s, U_1, \ldots, U_n)$. Observe that the tree $\widehat{\mathsf{T}}^{s,\mathrm{ord}}_{s,n}$ and random variables $\big(N^{\mathrm{right}}(e,j) : e \in E(T),\ j \ge 1\big)$ and $(N(v,\ell) : v \in I(T), 1 \le \ell \le \deg_T(v))$ are measurable functions of these quantities, as are the relative orderings of the atoms of local time along an edge. The remaining randomness lies in the random variables $Y_1, \ldots, Y_s$. Consider first a vertex $v \in I(T)$ and $1 \le \ell \le \deg_T(v)$. The probability that the leaves among $l_1, \ldots, l_s$ with indices in $S_{v,\ell}$ (where $|S_{v,\ell}| = k_{v,\ell}$) are glued into corner $c_{v,\ell}$ is

$$\frac{N(v,\ell)^{k_{v,\ell}}}{\prod_{l_j \in S_{v,\ell}} X(V_j)}.$$

Now consider an edge $e \in E(T)$ and fixed $a_e \ge 1$. The probability that the leaves among $l_1, \ldots, l_s$ with indices in the sets $S_{e,1}, \ldots, S_{e,a_e}$ (with $|S_{e,i}| = k_{e,i}$) are grouped together in the gluing, in that top-to-bottom order, is given by summing over $(j_1, \ldots, j_{a_e}) \in \mathbb{N}^{a_e,\neq}$, corresponding to different ordered collections of atoms of local time along the edge e, and multiplying by the probability $1/a_e!$ that this vector is such that $L(e,j_1) > L(e,j_2) > \cdots > L(e,j_{a_e})$:

$$\sum_{(j_1,\ldots,j_{a_e})\in \mathbb{N}^{a_e,\neq}} \frac{1}{a_e!} \prod_{i=1}^{a_e} \frac{N^{\mathrm{right}}(e,j_i)^{k_{e,i}}}{\prod_{l_j\in S_{e,i}} X(V_j)}.$$

The corners and edges all behave independently, and so multiplying everything together, we obtain that the probability of seeing the particular sets $((S_{v,\ell})_{1\le\ell\le\deg_T(v)})_{v\in I(T)}$, $((S_{e,i})_{1\le i\le a_e})_{e\in E(T)}$ in the random gluing plan is

$$\frac{1}{X(V_1)X(V_2)\ldots X(V_s)} \cdot \left(\prod_{v\in I(T)} \prod_{\ell=1}^{\deg_T(v)} N(v,\ell)^{k_{v,\ell}} \right) \cdot \left(\prod_{e\in E(T)} \sum_{(j_1,\ldots,j_{a_e})\in \mathbb{N}^{a_e,\neq}} \frac{1}{a_e!} \prod_{i=1}^{a_e} N^{\mathrm{right}}(e,j_i)^{k_{e,i}} \right). \tag{5.21}$$

Since the permutations $(\sigma_{v,\ell})_{v\in I(T), 1\le\ell\le\deg_T(v)}$ and $(\sigma_{e,i})_{e\in E(T), 1\le i\le a_e}$ are uniform and independent given the sets $((S_{v,\ell})_{1\le\ell\le\deg_T(v)})_{v\in I(T)}$ and $((S_{e,\ell})_{1\le\ell\le a_e})_{e\in E(T)}$, we see that each particular collection of permutations arises with conditional probability

$$\left(\prod_{v\in I(T)} \prod_{\ell=1}^{\deg_T(v)} \frac{1}{k_{v,\ell}!} \right) \cdot \left(\prod_{e\in E(T)} \frac{1}{k_{e,1}! \ldots k_{e,a_e}!} \right).$$

Multiplying (5.21) by this quantity gives the desired result. □

Recall that $\widehat{\mathsf{G}}_n^{s,\mathrm{ord}}$ is an ordered version of $\widehat{\mathsf{G}}_n^s$. We denote by $\mathsf{G}_n^{s,\mathrm{ord}}$ the corresponding ordered version in the s-biased case.

The distribution of $\mathsf{G}_n^{s,\mathrm{ord}}$. We will show that for all ordered multigraph $G \in \mathbb{M}_{s,n}^{\mathrm{ord}}$

$$\mathbb{P}\left(\mathsf{G}_n^{s,\mathrm{ord}} = G\right) \propto \prod_{v \in I(G)} \frac{w_{\deg_G(v)-1}}{(\deg_G(v)-1)!}. \tag{5.22}$$

Fix $G \in \mathbb{M}_{s,n}^{\mathrm{ord}}$. As previously mentioned, the only way to obtain G by gluing the s unlabelled leaves of a tree $T \in \mathbb{A}_{s,n}$ onto their ancestral paths is that the tree T is the base-tree of G, i.e. that $T = \textsc{Erase}(\textsc{Dep}(G))$. Let $C_s := \mathbb{E}\left[X(V_1)\dots X(V_s)\right]^{-1}$. Then using the change of measure formula (5.18), we have

$$\begin{aligned}\mathbb{P}\left(\mathsf{G}_n^{s,\mathrm{ord}} = G\right) &= C_s \cdot \mathbb{E}\left[\mathbf{1}_{\left\{\widehat{\mathsf{G}}_n^{s,\mathrm{ord}}=G\right\}} X(V_1)X(V_2)\dots X(V_s)\right] \\ &= C_s \cdot \mathbb{P}\left(\widehat{\mathsf{T}}_{s,n}^{s,\mathrm{ord}} = T\right) \mathbb{E}\left[\mathbf{1}_{\left\{\widehat{\mathsf{G}}_n^{s,\mathrm{ord}}=G\right\}} X(V_1)X(V_2)\dots X(V_s) \,\middle|\, \widehat{\mathsf{T}}_{s,n}^{s,\mathrm{ord}} = T\right].\end{aligned} \tag{5.23}$$

Observe here again that, apart from the labels on the leaves, the tree $\widehat{\mathsf{T}}_{s,n}^{s,\mathrm{ord}}$ has exactly the same distribution as $\mathsf{T}_{s+n}^{\mathrm{ord}}$ defined at the beginning of Section 5.2.1. So by (5.15), we have

$$\mathbb{P}\left(\widehat{\mathsf{T}}_{s,n}^{s,\mathrm{ord}} = T\right) \propto \prod_{v \in I(T)} \frac{w_{\deg_T(v)-1}}{(\deg_T(v)-1)!}. \tag{5.24}$$

We then calculate

$$\mathbb{E}\left[\mathbf{1}_{\left\{\widehat{\mathsf{G}}_n^{s,\mathrm{ord}}=G\right\}} X(V_1)X(V_2)\dots X(V_s) \,\middle|\, \widehat{\mathsf{T}}_{s,n}^{s,\mathrm{ord}} = T\right]$$

by taking expectations in the formula of Proposition 5.22 conditionally on the event $\{\widehat{\mathsf{T}}_{s,n}^{s,\mathrm{ord}} = T\}$. Recall that we fixed $T = \textsc{Erase}(\textsc{Dep}(G))$. Using Proposition 5.14 and Remark 5.16, we know explicitly the (conditional) distributions of each of the terms in (5.20). Using the independence stated there, we get

$$\begin{aligned}&\mathbb{E}\left[\mathbf{1}_{\left\{\widehat{\mathsf{G}}_n^{s,\mathrm{ord}}=G\right\}} X(V_1)X(V_2)\dots X(V_s) \,\middle|\, \widehat{\mathsf{T}}_{s,n}^{s,\mathrm{ord}} = T\right] \\ &= \mathbb{E}\left(\prod_{v\in I(T)} \prod_{\ell=1}^{\deg_T(v)} \frac{N(v,\ell)^{k_{v,\ell}}}{k_{v,\ell}!}\right) \cdot \left(\prod_{e\in E(T)} \sum_{(j_1,\dots,j_{a_e})\in\mathbb{N}^{a_e,\neq}} \frac{1}{a_e!} \prod_{i=1}^{a_e} \frac{N^{\mathrm{right}}(e,j_i)^{k_{e,i}}}{k_{e,i}!}\right) \\ &= \mathbb{E}N(T)^s \mathbb{E} \prod_{v\in I(T)} \left(\frac{N(v)}{N(T)}\right)^{k_v} \prod_{e\in E(T)} \left(\frac{N(e)}{N(T)}\right)^{k_e} \prod_{v\in I(T)} \mathbb{E} \prod_{\ell=1}^{\deg_T(v)} \frac{1}{k_{v,\ell}!} \left(\frac{N(v,\ell)}{N(v)}\right)^{k_{v,\ell}} \\ &\quad \times \prod_{e\in E(T)} \mathbb{E} \sum_{(j_1,\dots,j_{a_e})\in\mathbb{N}^{a_e,\neq}} \frac{1}{a_e!} \prod_{i=1}^{a_e} \left(\frac{N^{\mathrm{right}}(e,j_i)}{N(e)}\right)^{k_{e,i}} \frac{1}{k_{e,i}!}\end{aligned}$$

We now compute the different terms in this product separately.

Using Remark 5.16, we have

$$N(T) = N(e_1) + N(e_2) + \dots N(e_{|E(T)|}) + N(v_1) + \dots N(v_{|I(T)|}) \sim \mathrm{ML}\left(1/\alpha; n+s-1/\alpha\right),$$

so we get

$$\mathbb{E}\left[N(T)^s\right] = \frac{\Gamma(n+s-1/\alpha)\Gamma((n+s)\alpha+s-1)}{\Gamma((n+s)\alpha-1)\Gamma(n+s+(s-1)/\alpha)}.$$

Using Remark 5.16 again,

$$\left(\frac{N(e_1)}{N(T)},\ldots,\frac{N(e_{|E(T)|})}{N(T)},\frac{N(v_1)}{N(T)},\ldots,\frac{N(v_{|I(T)|})}{N(T)}\right)\sim \mathrm{Dir}(\alpha-1,\ldots,\alpha-1,d_1-1-\alpha,\ldots,d_r-1-\alpha).$$

Note that $|I(T)|=|E(T)|-n-s$ and $\sum_{v\in I(T)}\deg_T(v)=2|E(T)|-n-s-1$, which yield that

$$(\alpha-1)|E(T)|+\sum_{v\in I(T)}(\deg_T(v)-1-\alpha)=(n+s)\alpha-1.$$

So (5.35) gives

$$\mathbb{E}\left[\prod_{v\in I(T)}\left(\frac{N(v)}{N(T)}\right)^{k_v}\prod_{e\in E(T)}\left(\frac{N(e)}{N(T)}\right)^{k_e}\right]$$
$$=\frac{\Gamma((n+s)\alpha-1)}{\Gamma((n+s)\alpha+s-1)}\cdot\prod_{v\in I(T)}\frac{\Gamma(\deg_T(v)+k_v-1-\alpha)}{\Gamma(\deg_T(v)-1-\alpha)}\cdot\prod_{e\in E(T)}\frac{\Gamma(\alpha-1+k_e)}{\Gamma(\alpha-1)}.$$

Let $v\in I(T)$. Proposition 5.14 gives

$$\left(\frac{N(v,1)}{N(v)},\ldots,\frac{N(v,\deg_T(v))}{N(v)}\right)\sim \mathrm{Dir}(1,\ldots,1),$$

and then (5.35) yields

$$\mathbb{E}\left[\prod_{\ell=1}^{\deg_T(v)}\frac{1}{k_{v,\ell}!}\left(\frac{N(v,\ell)}{N(v)}\right)^{k_{v,\ell}}\right]=\frac{\Gamma(\deg_T(v))}{\Gamma(\deg_T(v)+k_v)}\cdot\left(\prod_{\ell=1}^{\deg_T(v)}\frac{\Gamma(k_{v,\ell}+1)}{\Gamma(1)}\right)\cdot\left(\prod_{\ell=1}^{\deg_T(v)}\frac{1}{k_{v,\ell}!}\right)$$
$$=\frac{(\deg_T(v)-1)!}{(\deg_T(v)+k_v-1)!}.$$

Let $e\in E(T)$. Using Proposition 5.14, we have

$$\left(\frac{N(e,j)}{N(e)}\right)_{j\geq1}\sim \mathrm{PD}(\alpha-1,\alpha-1),\quad\text{and}\quad\left(\frac{N^{\mathrm{right}}(e,j)}{N(e,j)}\right)_{j\geq1}\ \text{are i.i.d. U}[0,1],$$

so using Lemma 5.29, and the fact that $\mathbb{E}[U^p]=1/(p+1)$ for $U\sim\mathrm{U}[0,1]$, we get

$$\mathbb{E}\left[\sum_{(j_1,\ldots,j_{a_e})\in\mathbb{N}^{a_e,\neq}}\left(\frac{N^{\mathrm{right}}(e,j_1)}{N(e)}\right)^{k_{e,1}}\cdots\left(\frac{N^{\mathrm{right}}(e,j_{a_e})}{N(e)}\right)^{k_{e,a_e}}\right]=\left(\prod_{i=1}^{a_e}\frac{w_{k_{e,i}+1}}{k_{e,i}+1}\right)\cdot\frac{\Gamma(\alpha-1)}{\Gamma(k_e+\alpha-1)}\cdot a_e!.$$

Multiplying this by the combinatorial factor $\dfrac{1}{a_e!k_{e,1}!\ldots k_{e,a_e}!}$, we get

$$\prod_{i=1}^{a_e}\frac{w_{k_{e,i}+1}}{(k_{e,i}+1)!}\cdot\frac{\Gamma(\alpha-1)}{\Gamma(k_e+\alpha-1)}.$$

So, multiplying everything together, we get

$$\mathbb{E}\left[\mathbf{1}_{\left\{\widehat{\mathsf{G}}_n^{s,\mathrm{ord}}=G\right\}}X(V_1)X(V_2)\ldots X(V_s)\,\middle|\,\mathsf{T}_{s,n}^{\mathrm{ord}}=T\right]$$
$$=\frac{\Gamma(n+s-1/\alpha)}{\Gamma(n+s+(s-1)/\alpha)}\cdot\left(\prod_{e\in E(T)}\prod_{i=1}^{a_e}\frac{w_{k_{e,i}+1}}{(k_{e,i}+1)!}\right)\cdot\prod_{v\in I(T)}\frac{\Gamma(\deg_T(v)+k_v-1-\alpha)}{(\deg_T(v)+k_v-1)!}\frac{(\deg_T(v)-1)!}{\Gamma(\deg_T(v)-1-\alpha)}.\tag{5.25}$$

Now, if we fix an ordered multigraph $G \in \mathbb{M}_{s,n}^{\mathrm{ord}}$, from (5.23) and (5.24) we get

$$\mathbb{P}\left(\mathsf{G}_n^{s,\mathrm{ord}} = G\right) \propto \prod_{v \in I(T)} \frac{w_{\deg_T(v)-1}\,\Gamma(\deg_T(v)+k_v-1-\alpha)}{(\deg_T(v)+k_v-1)!\,\Gamma(\deg_T(v)-1-\alpha)} \cdot \left(\prod_{e \in E(T)} \prod_{i=1}^{a_e} \frac{w_{k_{e,i}+1}}{(k_{e,i}+1)!}\right).$$

Observe finally that every *new* internal vertex in G corresponds to some $e \in E(T)$ and some $1 \leq i \leq a_e$, and has degree $k_{e,i}+2$. For a vertex $v \in I(T)$, its degree in G is $\deg_G(v) = \deg_T(v) + k_v$. Moreover,

$$w_{\deg_G(v)-1} = w_{\deg_T(v)+k_v-1} = w_{\deg_T(v)-1} \cdot \frac{\Gamma(\deg_T(v)+k_v-1-\alpha)}{\Gamma(\deg_T(v)-1-\alpha)}.$$

Putting everything together, we indeed get (5.22).

We have now assembled all of the ingredients needed for the proof of Theorem 5.3.

Proof of Theorem 5.3. Take a multigraph $G \in \mathbb{M}_{s,n}$ with internal vertices $I(G)$, edge multiset $E(G)$ and a number $\mathrm{sl}(G)$ of self-loops. From Lemma 5.17, the number of corresponding ordered multigraphs is

$$\frac{\prod_{v \in I(G)} (\deg(v)-1)!}{|\mathrm{Sym}(G)| 2^{\mathrm{sl}(G)} \prod_{e \in \mathrm{supp}(E(G))} \mathrm{mult}(e)!}.$$

Combining this with (5.22), we get that for any multigraph $G \in \mathbb{M}_{s,n}$,

$$\mathbb{P}\left(\mathsf{G}_n^s = G\right) \propto \frac{\prod_{v \in I(G)} w_{\deg(v)-1}}{|\mathrm{Sym}(G)| 2^{\mathrm{sl}(G)} \prod_{e \in \mathrm{supp}(E(G))} \mathrm{mult}(e)!},$$

as claimed. □

5.3.4 The distribution of $(\mathsf{G}_n^s, n \geq 0)$ as a process

We now turn to the proof of Theorem 5.5, which says that the sequence $(\mathsf{G}_n^s, n \geq 0)$ evolves according to the multigraph version of Marchal's algorithm given in Section 5.1.2. Again, it is easier to work with multigraphs having cyclic orderings of the half-edges around each vertex in order to break symmetries. Recall from Section 5.3.3 that $\mathsf{G}_n^{s,\mathrm{ord}}$ denotes a version of G_n^s with cyclic orderings around the vertices built from the trees $\mathsf{T}_{s,n}^{s,\mathrm{ord}}$. We observe that there is a natural coupling of $\mathsf{T}_{s,n}^{s,\mathrm{ord}}$ for $n \geq 0$ obtained by repeatedly sampling new uniform leaves. Let $(\mathsf{G}_n^s, n \geq 0)$ and $(\mathsf{G}_n^{s,\mathrm{ord}}, n \geq 0)$ be built from this coupled version of the base trees. Note that, for all n, $\mathsf{G}_n^{s,\mathrm{ord}}$ is obtained from $\mathsf{G}_{n+1}^{s,\mathrm{ord}}$ by erasing the leaf labelled $n+1$ together with the edge to which it is connected. Recall also from (5.22) that the distribution of $\mathsf{G}_n^{s,\mathrm{ord}}$ is

$$\mathbb{P}\left(\mathsf{G}_n^{s,\mathrm{ord}} = G\right) = c_{s,n} \cdot \prod_{v \in I(G)} \frac{w_{\deg_G(v)-1}}{(\deg_G(v)-1)!}, \quad \forall G \in \mathbb{M}_{s,n}^{\mathrm{ord}},$$

where $c_{s,n}$ is the normalizing constant. We need an ordered counterpart of Marchal's algorithm for graphs with cyclic orderings around vertices. Starting from a graph $G \in \mathbb{M}_{s,n}^{\mathrm{ord}}$ and assigning to its edges and vertices the weights of Marchal's algorithm, we decide that (1) if a vertex is selected, then we glue the new edge-leaf in a corner chosen uniformly around this vertex, while (2) if an edge is selected, then we place the new edge-leaf on the right or on the left of the selected edge each with probability 1/2.

We will prove Theorem 5.5 together with the following result.

Proposition 5.23. *The sequence* $(\mathsf{G}_n^{s,\mathrm{ord}}, n \geq 0)$ *is Markovian, with transitions given by the ordered version of Marchal's algorithm.*

Proof of Proposition 5.23 and Theorem 5.5. The Markov property of $(\mathsf{G}_n^s, n \geq 0)$ and $(\mathsf{G}_n^{s,\mathrm{ord}}, n \geq 0)$ is immediate since the backward transitions are deterministic. Now fix n and let $G^{\mathrm{ord}} \in \mathbb{M}_{s,n}^{\mathrm{ord}}$ and $H^{\mathrm{ord}} \in \mathbb{M}_{s,n+1}^{\mathrm{ord}}$ be such that G^{ord} is obtained from H^{ord} by erasing the leaf labelled $n+1$ and the adjacent edge. Note that the internal vertices of our graphs are mutually distinguishable since the graphs are planted, with cyclic orderings around internal vertices. Then,

$$\mathbb{P}\left(\mathsf{G}_{n+1}^{s,\mathrm{ord}} = H^{\mathrm{ord}} \mid \mathsf{G}_n^{s,\mathrm{ord}} = G^{\mathrm{ord}}\right) = \frac{\mathbb{P}\left(\mathsf{G}_{n+1}^{s,\mathrm{ord}} = H^{\mathrm{ord}}\right)}{\mathbb{P}\left(\mathsf{G}_n^{s,\mathrm{ord}} = G^{\mathrm{ord}}\right)} = \frac{c_{s,n+1}}{c_{s,n}} \cdot \frac{\prod_{v \in I(H^{\mathrm{ord}})} \frac{w_{\deg_{H^{\mathrm{ord}}}(v)-1}}{(\deg_{H^{\mathrm{ord}}}(v)-1)!}}{\prod_{v \in I(G^{\mathrm{ord}})} \frac{w_{\deg_{G^{\mathrm{ord}}}(v)-1}}{(\deg_{G^{\mathrm{ord}}}(v)-1)!}}.$$

Now there are two different cases, (a) and (b) below:

(a) The leaf $n+1$ of H^{ord} is attached to a vertex v of H^{ord} that has a degree greater than or equal to 4. In this case, v corresponds to a vertex of G^{ord}, still denoted by v, and $I(H^{\mathrm{ord}}) = I(G^{\mathrm{ord}})$, $\deg_{G^{\mathrm{ord}}}(v) = \deg_{H^{\mathrm{ord}}}(v) - 1$ and the degree of any other internal vertex is identical in G^{ord} and H^{ord}. Since

$$w_{\deg_{H^{\mathrm{ord}}}(v)-1} = w_{\deg_{G^{\mathrm{ord}}}(v)} = (\deg_{G^{\mathrm{ord}}}(v) - 1 - \alpha) w_{\deg_{G^{\mathrm{ord}}}(v)-1},$$

together with the above expression for $\mathbb{P}\left(\mathsf{G}_{n+1}^{s,\mathrm{ord}} = H^{\mathrm{ord}} \mid \mathsf{G}_n^{s,\mathrm{ord}} = G^{\mathrm{ord}}\right)$ this implies that

$$\mathbb{P}\left(\mathsf{G}_{n+1}^{s,\mathrm{ord}} = H^{\mathrm{ord}} \mid \mathsf{G}_n^{s,\mathrm{ord}} = G^{\mathrm{ord}}\right) = \frac{c_{s,n+1}}{c_{s,n}} \cdot \frac{\deg_{G^{\mathrm{ord}}}(v) - 1 - \alpha}{\deg_{G^{\mathrm{ord}}}(v)}. \tag{5.26}$$

(b) The vertex v has degree 3 in H^{ord} and is erased when erasing the leaf $n+1$ and the adjacent edge. In this case $I(H^{\mathrm{ord}}) = I(G^{\mathrm{ord}}) \cup \{v\}$ and

$$\mathbb{P}\left(\mathsf{G}_{n+1}^{s,\mathrm{ord}} = H^{\mathrm{ord}} \mid \mathsf{G}_n^{s,\mathrm{ord}} = G^{\mathrm{ord}}\right) = \frac{c_{s,n+1}}{c_{s,n}} \cdot \frac{\alpha - 1}{2}. \tag{5.27}$$

Proposition 5.23 follows immediately.

This argument also gives the transition probabilities of the process $(\mathsf{G}_n^s, n \geq 0)$. Recall the function $\psi : \mathbb{M}_{s,n}^{\mathrm{ord}} \to \mathbb{M}_{s,n}$ that forgets the cyclic ordering around vertices. We have that

$$\mathbb{P}\left(\mathsf{G}_{n+1}^s = H \mid \mathsf{G}_n^s = G\right) = \sum_{G^{\mathrm{ord}} \in \psi^{-1}(G)} \mathbb{P}\left(\mathsf{G}_{n+1}^s = H \mid \mathsf{G}_n^{s,\mathrm{ord}} = G^{\mathrm{ord}}\right) \mathbb{P}\left(\mathsf{G}_n^{s,\mathrm{ord}} = G^{\mathrm{ord}} \mid \mathsf{G}_n^s = G\right). \tag{5.28}$$

If H is obtained from G by attaching a leaf-edge to a vertex v of G, then, from (5.26), we get

$$\mathbb{P}\left(\mathsf{G}_{n+1}^s = H \mid \mathsf{G}_n^{s,\mathrm{ord}} = G^{\mathrm{ord}}\right) = \deg_G(v) \cdot \frac{c_{s,n+1}}{c_{s,n}} \cdot \frac{\deg_G(v) - 1 - \alpha}{\deg_G(v)} \quad \text{for all } G^{\mathrm{ord}} \in \psi^{-1}(G).$$

With (5.28), this gives

$$\mathbb{P}\left(\mathsf{G}_{n+1}^s = H \mid \mathsf{G}_n^s = G\right) = \frac{c_{s,n+1}}{c_{s,n}} \cdot (\deg_G(v) - 1 - \alpha).$$

Similarly, from (5.27) and (5.28), we get that when G' is obtained from G by attaching a leaf-edge to the middle of an edge of G, we have

$$\mathbb{P}\left(\mathsf{G}^s_{n+1} = H \mid \mathsf{G}^s_n = G\right) = \frac{c_{s,n+1}}{c_{s,n}} \cdot (\alpha - 1).$$

Theorem 5.5 follows. □

5.3.5 The unrooted kernel G^s_{-1}

In this section, we fix $s \geq 2$. Our goal is to prove that the distribution of G^s_{-1} is that given in Theorem 5.3, and that the conditional probability of G^s_0 given G^s_{-1} is given by a step in Marchal's algorithm. We cannot proceed as before since the use of cyclic orderings around vertices is not sufficient to break all the symmetries in the unrooted graph G^s_{-1}. We instead label the internal vertices: let $\mathsf{G}^{s,\mathrm{lab}}_0$ denote a version of G^s_0 with internal vertices labelled uniformly from 1 to $|V(\mathsf{G}^s_0)|$.

For any connected multigraph G (labelled or not) we write

$$w(G) := \frac{\prod_{v \in I(G)} w_{\deg(v)-1}}{|I(G)|!\ 2^{\mathrm{sl}(G)} \prod_{e \in \mathrm{supp}(E(G))} \mathrm{mult}(e)!},$$

with the usual notation. From Theorem 5.3 and (5.19), we know that the distribution of the labelled graph $\mathsf{G}^{s,\mathrm{lab}}_0$ is

$$\mathbb{P}\left(\mathsf{G}^{s,\mathrm{lab}}_0 = G\right) = \tilde{c}_{s,0} \cdot w(G), \tag{5.29}$$

where $\tilde{c}_{s,0}$ is the normalising constant.

Let H^{lab} and G^{lab} be labelled versions of multigraphs in $\mathbb{M}_{s,0}$ and $\mathbb{M}_{s,-1}$ respectively that are *compatible* in the sense that removing the root and the adjacent edge (in the following, we will use the word *root-edge*) in H^{lab} gives a graph which, after an increasing mapping of the labelling to $\{1, \dots, |V(G^{\mathrm{lab}})|\}$, is G^{lab}. We then distinguish 2 cases, precisely one of which occurs:

(a) The root-edge in H^{lab} is attached to a vertex v of degree $\deg_{H^{\mathrm{lab}}}(v) \geq 4$, in which case

$$w(H^{\mathrm{lab}}) = \frac{w_{\deg_{H^{\mathrm{lab}}}(v)-1}}{w_{\deg_{G^{\mathrm{lab}}}(v)-1}} \cdot w(G^{\mathrm{lab}}) = (\deg_{G^{\mathrm{lab}}}(v) - 1 - \alpha) \cdot w(G^{\mathrm{lab}}).$$

Note that, given G^{lab} and a vertex v of G^{lab}, there is a unique graph H^{lab} which has its root-edge attached to v and is compatible with G^{lab}.

(b) The root-edge is attached to a vertex v of degree $\deg_{H^{\mathrm{lab}}}(v) = 3$. Its deletion either "creates" an edge e of G^{lab} (possibly a self-loop, erasing then at the same time an edge of multiplicity 2 in H^{lab}) or increases by 1 the multiplicity of an edge $e \in \mathrm{supp}(H^{\mathrm{lab}})$ (possibly a multiple self-loop, erasing, again, at the same time an edge of multiplicity 2 in H^{lab}). In all cases,

$$w(H^{\mathrm{lab}}) = \frac{w_{\deg_{H^{\mathrm{lab}}}(v)-1} \cdot \mathrm{mult}(e)}{|I(G^{\mathrm{lab}})| + 1} \cdot w(G^{\mathrm{lab}}) = \frac{(\alpha - 1) \cdot \mathrm{mult}(e)}{|I(G^{\mathrm{lab}})| + 1} \cdot w(G^{\mathrm{lab}}),$$

where $\mathrm{mult}(e)$ refers here to the multiplicity of e seen as an element of $\mathrm{supp}(G^{\mathrm{lab}})$. Note that given an edge e of G^{lab}, there are exactly $|I(G^{\mathrm{lab}})| + 1$ graphs H^{lab} with the root-edge attached in the middle of (a copy of) e that are compatible with G^{lab}.

From this, (5.29) and the fact that the sum of the Marchal weights is $(s-1)(\alpha+1)$ for any graph in $\mathbb{M}_{s,-1}$ (see (5.8)), we obtain the distribution of $G_{-1}^{s,\mathrm{lab}}$ as

$$\begin{aligned}
\mathbb{P}\left(\mathsf{G}_{-1}^{s,\mathrm{lab}} = G^{\mathrm{lab}}\right) &= \sum_{\substack{H^{\mathrm{lab}} \text{ compatible} \\ \text{with } G^{\mathrm{lab}}}} \mathbb{P}\left(\mathsf{G}_{0}^{s,\mathrm{lab}} = H^{\mathrm{lab}}\right) \\
&= \tilde{c}_{s,0} \sum_{\substack{H^{\mathrm{lab}} \text{ compatible} \\ \text{with } G^{\mathrm{lab}}}} w(H^{\mathrm{lab}}) \\
&= \tilde{c}_{s,0} \cdot \left(\sum_{v \in I(G^{\mathrm{lab}})} (\deg_{G^{\mathrm{lab}}}(v) - 1 - \alpha) + \sum_{e \in \mathrm{supp}(E(G^{\mathrm{lab}}))} \mathrm{mult}(e)(\alpha - 1) \right) \cdot w(G^{\mathrm{lab}}) \\
&= \tilde{c}_{s,-1} \cdot w(G^{\mathrm{lab}}),
\end{aligned}$$

where $\tilde{c}_{s,-1} = \tilde{c}_{s,0}(s-1)(\alpha+1)$. Together with (5.19), which holds for graphs of $\mathbb{M}_{s,-1}$, this implies that G_{-1}^{s} has the required distribution. Next, to get the conditional distribution of G_{0}^{s} given G_{-1}^{s} we write, for $H \in \mathbb{M}_{s,0}$ and $G \in \mathbb{M}_{s,-1}$,

$$\mathbb{P}\left(\mathsf{G}_0^s = H \mid \mathsf{G}_{-1}^s = G\right) = \sum_{\substack{G^{\mathrm{lab}} \text{ a labelled} \\ \text{version of } G}} \frac{\mathbb{P}\left(\mathsf{G}_0^s = H, \mathsf{G}_{-1}^{s,\mathrm{lab}} = G^{\mathrm{lab}}\right)}{\mathbb{P}\left(\mathsf{G}_{-1}^{s,\mathrm{lab}} = G^{\mathrm{lab}}\right)} \mathbb{P}\left(\mathsf{G}_{-1}^{s,\mathrm{lab}} = G^{\mathrm{lab}} \mid \mathsf{G}_{-1}^s = G\right).$$

From the remarks above, we see that when H is obtained from G by gluing the root-edge to a vertex v of G, we get

$$\frac{\mathbb{P}\left(\mathsf{G}_0^s = H, \mathsf{G}_{-1}^{s,\mathrm{lab}} = G^{\mathrm{lab}}\right)}{\mathbb{P}\left(\mathsf{G}_{-1}^{s,\mathrm{lab}} = G^{\mathrm{lab}}\right)} = \frac{\tilde{c}_{s,0}}{\tilde{c}_{s,-1}} \cdot \frac{w(H)}{w(G)} = \frac{\tilde{c}_{s,0}}{\tilde{c}_{s,-1}} \cdot (\deg_G(v) - 1 - \alpha),$$

for all labelled versions $\mathsf{G}^{\mathrm{lab}}$. If, on the other hand, H is obtained from G by gluing the root-edge to (a copy of) an edge $e \in \mathrm{supp}(G)$, then

$$\frac{\mathbb{P}\left(\mathsf{G}_0^s = H, \mathsf{G}_{-1}^{s,\mathrm{lab}} = G^{\mathrm{lab}}\right)}{\mathbb{P}\left(\mathsf{G}_{-1}^{s,\mathrm{lab}} = G^{\mathrm{lab}}\right)} = (|I(G)|+1) \cdot \frac{\tilde{c}_{s,0}}{\tilde{c}_{s,-1}} \cdot \frac{w(H)}{w(G)} = \frac{\tilde{c}_{s,0}}{\tilde{c}_{s,-1}} \cdot (\alpha - 1) \cdot \mathrm{mult}(e).$$

Putting everything together, we see that we do indeed obtain the transition probabilities corresponding to a step of Marchal's algorithm.

5.3.6 The configuration model embedded in a limit component

The goal of this subsection is to prove Corollary 5.4 where we identify for each $n \geq 0$ (and $n = -1$ if $s \geq 2$) the distribution of G_n^s with that of a specific configuration model.

Two probability distributions. In Section 3 of Duquesne and Le Gall [53], it is shown that the rooted subtree obtained by sampling $n \geq 0$ leaves in the α-stable tree is distributed as a planted (non-ordered version of a) Galton-Watson tree conditioned to have n leaves, with critical offspring distribution η_α satisfying

$$\eta_\alpha(k) = \frac{w_k}{k!}, \quad k \geq 2, \qquad \eta_\alpha(1) = 0, \qquad \eta_\alpha(0) = \frac{1}{\alpha},$$

or, equivalently, with probability generating function $z + \alpha^{-1}(1-z)^{\alpha}, z \in (0,1]$, as already mentioned in Section 5.1.2. Note that $\eta_\alpha(k) \sim_{k\to\infty} ck^{-1-\alpha}$ for some constant $c > 0$, by Stirling's approximation. Now consider the random variable $D^{(\alpha)}$ with distribution introduced in (5.7), and note that it is indeed a probability distribution since

$$\sum_{k\geq 2} \frac{w_k}{k!} = \frac{(\alpha-1)}{2} + \sum_{k\geq 3} \frac{(k-1-\alpha)w_{k-1}}{k!} = \frac{(\alpha-1)}{2} + \sum_{k\geq 3} \frac{w_{k-1}}{(k-1)!} - (1+\alpha) \sum_{k\geq 3} \frac{w_{k-1}}{k!},$$

which implies that

$$\sum_{k\geq 2} \frac{w_{k-1}}{k!} + \frac{1}{\alpha} = \frac{(\alpha-1)}{2(1+\alpha)} + \frac{1}{\alpha} = \frac{\alpha^2+\alpha+2}{2(1+\alpha)\alpha}.$$

It is straightforward to see that $\mathbb{E}[D^{(\alpha)}] = 2$. Moreover, if we consider the biased version

$$\mathbb{P}(\hat{D}^{(\alpha)} = k) := \frac{k\mathbb{P}(D^{(\alpha)} = k)}{\mathbb{E}\big[D^{(\alpha)}\big]}, \quad k \geq 1$$

we immediately get that $\hat{D}^{(\alpha)} - 1$ has the same distribution as η_α. This in particular implies that $D^{(\alpha)}$ satisfies the conditions (5.1).

The stable configuration model. Fix $n \geq 0$ if $s \in \{0,1\}$ or $n \geq -1$ if $s \geq 2$. Then fix $m \geq n+1$ and consider the multigraph C_m sampled from the configuration model with i.i.d. degrees $D_0^{(\alpha)}, \dots, D_{m-1}^{(\alpha)}$ distributed as $D^{(\alpha)}$. From Proposition 7.7 in [75], we have that

$$\mathbb{P}\left(\mathsf{C}_m = G \;\middle|\; D_i^{(\alpha)} = d_i, 0 \leq i \leq m-1\right) = \frac{1}{(\sum_{0\leq i\leq m-1} d_i - 1)!!} \cdot \frac{\prod_{0\leq i\leq m-1} d_i!}{2^{\mathrm{sl}(G)} \prod_{e\in\mathrm{supp}(E)} \mathrm{mult}(e)!},$$

for every multigraph $G = (V,E)$ with m *labelled* vertices of respective degrees $d_0, \dots, d_{m-1}$ such that $\sum_{0\leq i\leq m-1} d_i$ is even. Hence, the distribution of C_m is given for each such multigraph by

$$\mathbb{P}(\mathsf{C}_m = G) = \left(\frac{2(1+\alpha)\alpha}{\alpha^2+\alpha+2}\right)^m \cdot \frac{1}{(\sum_{0\leq i\leq m-1} d_i - 1)!!} \cdot \frac{\prod_{0\leq i\leq m-1} d_i!}{2^{\mathrm{sl}(G)} \prod_{e\in\mathrm{supp}(E)} \mathrm{mult}(e)!} \cdot \frac{1}{\alpha^{\#\{i:d_i=1\}}} \cdot \prod_{i=0}^{m-1} \frac{w_{d_i-1}}{d_i!}.$$

On the event $\{\mathsf{C}_m \text{ is connected}, s(\mathsf{C}_m) = s\}$, the sum $\sum_{0\leq i\leq m-1} d_i$ depends only on m and s. Conditioning additionally on $\{D_0^{(\alpha)} = \dots = D_n^{(\alpha)} = 1, D_i^{(\alpha)} \neq 1, n+1 \leq i \leq m-1\}$, we have $\#\{i : D_i^{(\alpha)} = 1\} = n+1$. Forgetting the labels $n+1, \dots, m-1$ (which we now know belong to internal vertices), we obtain a factor of $(m-n-1)!/|\mathrm{Sym}(G)|$. (See (5.19) for further discussion.) Together with Theorem 5.3 this implies Corollary 5.4.

5.4 Two simple constructions of the graph $\mathcal{G}^s$

Let $s \geq 1$. We start in Section 5.4.1 by proving that the (measured) $\mathbb{R}$-graph $\mathcal{G}^s$ is the almost-sure limit of rescaled versions of its combinatorial shapes $\mathsf{G}_n^s, n \geq 0$ equipped with the uniform distribution on their leaves. Together with the algorithmic construction of the graphs G_n^s for $n \geq 0$ (Theorem 5.5) and some urn model asymptotics recalled in the Appendix, this will lead us to the two alternative constructions of $\mathcal{G}^s$ presented in the introduction: in Section 5.4.2, we prove Theorem 5.6 and Proposition 5.7, giving the distribution of $\mathcal{G}^s$ as a collection of rescaled α-stable trees appropriately glued onto the kernel K^s; Section 5.4.3 is then devoted to the line-breaking construction of Theorem 5.8.

5.4.1 The graph as the scaling limit of its marginals

Recall from Section 5.2.2 that $\mathcal{G}^s$ is constructed from $\mathcal{T}^s$, a biased version of the α-stable tree, by appropriately gluing s marked leaves onto randomly selected branch-points. Recall also that X^s denotes the s-biased stable excursion from which $\mathcal{T}^s$ is built, that $\pi^s(V_1^s),\dots,\pi^s(V_s^s)$ are the s leaves to be glued and that $\pi^s(U_i), i\geq 1$ are i.i.d. uniform leaves. For all $n\geq 1$, $\mathcal{T}_{s,n}^s$ then denotes the subtree of $\mathcal{T}^s$ spanned by the root and the leaves $\pi^s(V_1^s),\dots,\pi^s(V_s^s),\pi^s(U_1),\dots,\pi^s(U_n)$ and we let $\mathsf{T}_{s,n}^s$ be its combinatorial shape. Finally, recall that $\mathcal{G}_n^s$ is the connected subgraph of $\mathcal{G}^s$ consisting of the union of the kernel and the paths from the leaves $\pi^s(U_1),\dots,\pi^s(U_n)$ to the root, for all $n\geq 0$, and that the finite graph G_n^s denotes the combinatorial shape of $\mathcal{G}_n^s$. We will use the following observation: for all n larger than some finite random variable, $\mathcal{G}_n^s$ is obtained from $\mathcal{T}_{s,n}^s$ by an appropriate gluing of the s leaves $\pi^s(V_1^s),\dots,\pi^s(V_s^s)$ to some of its *internal* vertices (for small n, it may be that we instead glue some leaves along edges of $\mathcal{T}_{s,n}^s$).

The goal of this section is to prove Proposition 5.2: when the graph G_n^s is equipped with the uniform distribution on its leaves,

$$\frac{\mathsf{G}_n^s}{n^{1-1/\alpha}} \xrightarrow[n\to\infty]{\text{a.s.}} \alpha\cdot\mathcal{G}^s \tag{5.30}$$

for the Gromov-Hausdorff-Prokhorov topology. With this aim in mind, we first observe that $\mathcal{G}^s$ can be recovered from the completion of the union of its continuous marginals.

Lemma 5.24. *With probability one,*

$$\mathcal{G}^s=\overline{\cup_{n\geq 0}\mathcal{G}_n^s},$$

and consequently $\mathcal{G}^s$ is the a.s. limit of $\mathcal{G}_n^s$ in $(\mathscr{C},\mathrm{d}_{\mathrm{GHP}})$, when the graph $\mathcal{G}_n^s$ is endowed with the uniform distribution on its leaves for $n\geq 1$.

Indeed, it is well-known that the α-stable tree is almost surely the completion of the union of its continuous marginals, which entails a similar result for the biased version $\mathcal{T}^s$ and then for the graph $\mathcal{G}^s$, using its construction from $\mathcal{T}^s$. The measures can then be incorporated by using the strong law of large numbers.

Proof of Proposition 5.2. We make use of the fact (5.17) that the α-stable tree is the almost-sure scaling limit of its discrete marginals. We refer the reader to the book of Burago, Burago and Ivanov [32] for background on the notions of a *correspondence* and its *distortion*, which are used here for the proof.

By Lemma 5.24, it suffices to prove that almost surely

$$\mathrm{d}_{\mathrm{GHP}}\left(\frac{\mathsf{G}_n^s}{n^{1-1/\alpha}},\alpha\cdot\mathcal{G}_n^s\right)\xrightarrow[n\to\infty]{}0.$$

We observe first that almost surely

$$\mathrm{d}_{\mathrm{GHP}}\left(n^{1/\alpha-1}\mathsf{T}_{s,n}^s,\alpha\cdot\mathcal{T}_{s,n}^s\right)\xrightarrow[n\to\infty]{}0.$$

This is proved for $s=0$ in [42, Section 2.4] and may be transferred to $s\geq 1$ by absolute continuity. The $s=0$ case is proved in [42] by using a natural correspondence which we introduce here for general s and call $\mathcal{R}_n^s$. It is a correspondence between $n^{\frac{1}{\alpha}-1}\mathsf{T}_{s,n}^s$ and $\alpha\cdot\mathcal{T}_{s,n}^s$. The leaves with the same labels correspond to one another, and the internal vertices of $\mathsf{T}_{s,n}^s$ are put in correspondence with the branch-points of $\mathcal{T}_{s,n}^s$ in the obvious way. Finally, the edges of $\mathcal{T}_{s,n}^s$ (which have real-valued lengths and which we think of as line-segments) are put in correspondence with the vertex or vertices of $\mathsf{T}^s_{s,n}$ corresponding to their end-points. From [42] we obtain that the distortion $\mathrm{dist}(\mathcal{R}^s_n)$ of the correspondence $\mathcal{R}^s_n$ tends to 0 almost surely as $n \to \infty$. To deal with the gluing, we use the fact already observed above that for n sufficiently large, $\mathcal{G}^s_n$ is obtained from $\mathcal{T}^s_{s,n}$ by an appropriate gluing of the s leaves $\pi^s(V^s_1), \dots, \pi^s(V^s_s)$ to its internal vertices; similarly G^s_n is obtained by the gluing of the corresponding leaves of $\mathsf{T}^s_{s,n}$ to the corresponding internal vertices of this tree. It then follows from Lemma 4.2 of [6] that

$$\mathrm{d}_{\mathrm{GHP}}\left(\frac{\mathsf{G}^s_n}{n^{1-1/\alpha}}, \alpha \cdot \mathcal{G}^s_n\right) \leq \frac{(s+1)}{2}\mathrm{dist}(\mathcal{R}^s_n)$$

and the claimed almost-sure convergence follows easily. □

5.4.2 Construction from randomly scaled stable trees glued to the kernel

We now turn to the proof of Theorem 5.6: in $(\mathscr{C}, \mathrm{d}_{\mathrm{GHP}})$, we have the identity in distribution of the measured compact metric spaces

$$\mathcal{G}^s \overset{\mathrm{d}}{=} \mathcal{G}(\mathsf{K}^s) \tag{5.31}$$

(with the notation used in Section 5.1.2). We will also prove Proposition 5.7 in this section.

Proof of Theorem 5.6. Using (5.30), we just need to prove that

$$\frac{\mathsf{G}^s_n}{n^{1-1/\alpha}} \xrightarrow[n\to\infty]{\mathrm{d}} \alpha \cdot \mathcal{G}(\mathsf{K}^s)$$

for the Gromov-Hausdorff-Prokhorov topology, when the graph G^s_n is equipped with the uniform distribution on its leaves. (We will prove the compactness of the object on the right-hand side below.) As discussed earlier, the graph G^s_n may be viewed as a collection of trees glued to the kernel K^s. We will show that each of these tree-blocks converges after rescaling to its continuous counterpart used in the construction of $\mathcal{G}(\mathsf{K}^s)$. Our argument and notation are similar to those used in the proof of Proposition 5.14 concerning the stable tree.

We work conditionally on K^s. Let m denote the number of edges of K^s, which are arbitrarily labelled as $e_1, \dots, e_m$. Let $v_1, \dots, v_{m-s}$ denote the internal vertices of K^s, again in arbitrary order, and $d_1, \dots, d_{m-s}$ their respective degrees. For each $n \geq 0$, we interpret these edges (resp. vertices) as edges of G^s_n with edge-lengths (resp. vertices). For each k, we write $T_n(e_k)$ for the subtree of G^s_n induced by the vertices closer to e_k than to any other edge $e_i, i \neq k$, including the two end-points of e_k. These end-vertices are interpreted as leaves of $T_n(e_k)$ and count as distinct leaves even if e_k is a loop. (These formulation may seem arbitrary but it is the one needed to initiate properly the urn model we will use below.) The number of leaves of $T_n(e_k)$ is then denoted by $M_n(e_k)$. Similarly we let $T_n(v_i)$ denote the subtree of G^s_n induced by the set of all vertices closer to v_i than to any edge $e_k, 1 \leq k \leq m$, including v_i which is considered as its root. Then $M_n(v_i)$ denotes its number of leaves (here v_i is *not* considered to be a leaf so that, in particular, $M_n(v_i) = 0$ if $T_n(v_i)$ has vertex-set $\{v_i\}$). Next, for each $1 \leq i \leq m-s$, let $T_n(v_i, j), j \geq 1$ denote the connected components of $T_n(v_i)\backslash\{v_i\}$. We think of these subtrees as planted (and we again call the root of each v_i), so that if we identify their roots we recover $T_n(v_i)$. The number of such subtrees is finite (possibly zero) for each n but tends to infinity as $n \to \infty$. We label them $T_n(v_i, 1), T_n(v_i, 2), \dots$ in order of appearance, with the convention that $T_n(v_i, j)$ is the empty set if there are strictly fewer than j subtrees at step n. Let $M_n(v_i, j)$ be the number of leaves of $T_n(v_i, j), j \geq 1$.

• *Scaling limits of the numbers of leaves.* It is easy to see using the algorithmic construction of the sequence $(\mathsf{G}_n^s, n \geq 0)$ from Theorem 5.5 that the process

$$(\alpha M_n(e_1) - \alpha - 1, \dots, \alpha M_n(e_m) - \alpha - 1, \\ \alpha M_n(v_1) + d_1 - 1 - \alpha, \dots, \alpha M_n(v_{m-s}) + d_{m-s} - 1 - \alpha)_{n\geq 0} \quad (5.32)$$

evolves according to Pólya's urn (see Theorem 5.30) with $2m - s$ colours of initial weights

$$(\alpha - 1, \dots, \alpha - 1, d_1 - 1 - \alpha, \dots, d_{m-s} - 1 - \alpha),$$

respectively, and weight parameter α. Hence, there exists a random vector $(M_1, \dots, M_{2m-s})$ with the Dirichlet distribution of parameters specified in (5.9) such that

$$\left(\frac{M_n(e_1)}{n}, \dots, \frac{M_n(e_m)}{n}, \frac{M_n(v_1)}{n}, \dots, \frac{M_n(v_{m-s})}{n}\right) \xrightarrow[n\to\infty]{\text{a.s.}} (M_1, \dots, M_{2m-s}).$$

Next we observe that for all i the jumps of $((M_n(v_i, j))_{j\geq 1}, n \geq 0)$ follow the same dynamics as a Chinese restaurant process with parameters $1/\alpha$ and $(d_i - 1 - \alpha)/\alpha$, independently of everything else. Since the total number of jumps at step n is $M_n(v_i)$, Theorem 5.31 yields

$$\left(\frac{M_n^{\downarrow}(v_i, j)}{M_n(v_i)}, j \geq 1\right) \xrightarrow[n\to\infty]{\text{a.s.}} (\Delta_{i,j}, j \geq 1),$$

where $(M_n^{\downarrow}(v_i, j), j \geq 1)$ denotes the decreasing reordering of $(M_n(v_i, j), j \geq 1)$ and the limit $(\Delta_{i,j}, j \geq 1)$ follows a Poisson-Dirichlet $\mathrm{PD}(1/\alpha, (d_i - 1 - \alpha)/\alpha)$ distribution, independent of the random variables $(M_1, \dots, M_{2m-s})$. (The convergence holds in ℓ^1 equipped with its usual metric.)

• *Scaling limits of the trees $T_n(e_k), T_n(v_i, j)$.* Given the processes $(M_n(e_k), n \geq 0)$, $(M_n(v_i, j), n \geq 0)$, for all k, i, j, the jump evolutions of the trees $T_n(e_k)$, $T_n(v_i, j)$, $n \geq 0$ are independent and all follow Marchal's algorithm. Then writing $e_k = \{x_k, y_k\}$ for $1 \leq k \leq m$, we know by (5.17) that there exist *rescaled* (measured) α-stable trees $\mathcal{T}_k, \mathcal{T}_{i,j}$, k, i, j such that, given $(M_1, \dots, M_{2m-s})$ and $(\Delta_{i,j}, j \geq 1)$, the trees are independent, $\mathcal{T}_k$ has total mass M_k, $\mathcal{T}_{i,j}$ total mass $M_{i+m} \cdot \Delta_{i,j}$ and, furthermore,

(a) for all k,

$$\left(\frac{T_n(e_k)}{n^{1-1/\alpha}}, x_k, y_k\right) = \left(\left(\frac{M_n(e_k)}{n}\right)^{1-1/\alpha} \cdot \frac{T_n(e_k)}{M_n(e_k)^{1-1/\alpha}}, x_k, y_k\right) \xrightarrow[n\to\infty]{\text{a.s.}} (\alpha \cdot \mathcal{T}_k, \rho_k, L_k)$$

for the 2-pointed Gromov-Hausdorff-Prokhorov topology, the tree $T_n(e_k)$ being implicitly endowed with the measure that assigns weight $1/n$ to each of its leaves (here, ρ_k denotes the root of $\mathcal{T}_k$ and L_k a uniform leaf);

(b) for all i, j,

$$\left(\frac{T_n(v_i, j)}{n^{1-1/\alpha}}, v_i\right) = \left(\left(\frac{M_n(v_i, j)}{n}\right)^{1-1/\alpha} \cdot \frac{T_n(v_i, j)}{M_n(v_i, j)^{1-1/\alpha}}, v_i\right) \xrightarrow[n\to\infty]{\text{a.s.}} (\alpha \cdot \mathcal{T}_{i,j}, \rho_{i,j})$$

for the pointed Gromov-Hausdorff-Prokhorov topology, where again $T_n(v_i, j)$ is endowed with the measure that assigns weight $1/n$ to each of its leaves, and $\rho_{i,j}$ is the root of $\mathcal{T}_{i,j}$.

• *Scaling limits of the trees $T_n(v_i)$, and the compactness of the limit.* Fix $i \geq 1$ and recall that $T_n(v_i)$ is obtained by identifying the roots of the trees $T_n(v_i, j), j \geq 1$. We now show that $n^{-(1-1/\alpha)} T_n(v_i)$ converges in probability for the pointed GHP-topology to the measured $\mathbb{R}$-tree $\mathcal{T}_{(i)}$ obtained by identifying the roots of the trees $\alpha \cdot \mathcal{T}_{i,j}$.

Let us first show that $\mathcal{T}_{(i)}$ is compact and is the almost sure GHP-limit as $j_0 \to \infty$ of the $\mathbb{R}$-tree $\mathcal{T}_{(i)}^{j_0}$ obtained by gluing the first j_0 trees $\mathcal{T}_{i,j}, j \leq j_0$ together at their roots. (For different values of j_0 we think of the underlying spaces as being nested and all contained within $\mathcal{T}_{(i)}$.) For a rooted $\mathbb{R}$-tree $\mathtt{t}$, we write $\mathrm{ht}(\mathtt{t})$ for its height. Let $\mathcal{T}$ denote a standard α-stable tree (of total mass 1). Then by the scaling property of the stable tree we have

$$\mathbb{E}\left(\sup_{j>j_0} \mathrm{ht}(\mathcal{T}_{i,j})\right)^{\alpha/(\alpha-1)} \leq \sum_{j>j_0} \mathbb{E}\left[\mathrm{ht}(\mathcal{T}_{i,j})^{\alpha/(\alpha-1)}\right] = \mathbb{E}\left[\mathrm{ht}(\mathcal{T})^{\alpha/(\alpha-1)}\right] \mathbb{E}\left[M_{i+m}\right] \sum_{j>j_0} \mathbb{E}\left[\Delta_{i,j}\right].$$

Since $\mathrm{ht}(\mathcal{T})$ has finite exponential moments (see, for example, equation (2) of [87] for a convenient statement) the right-hand side is finite, and clearly tends to 0 as $j_0 \to \infty$. Hence the decreasing sequence $\sup_{j>j_0} \mathrm{ht}(\mathcal{T}_{i,j})$ converges a.s. to 0 as $j_0 \to \infty$. This implies in particular that $\mathcal{T}_{(i)}$ is a.s. compact. Then, note that

$$\mathrm{d_{GHP}}\left(\mathcal{T}_{(i)}, \mathcal{T}_{(i)}^{j_0}\right) \leq \max\left(\sup_{j>j_0} \mathrm{ht}(\mathcal{T}_{i,j}), M_{i+m} \cdot \sum_{j>j_0} \Delta_{i,j}\right),$$

since $M_{i+m} \cdot \sum_{j>j_0} \Delta_{i,j}$ is the total mass of $\mathcal{T}_{(i)} \setminus \mathcal{T}_{(i)}^{j_0}$. This total mass also converges to 0. Hence, $\mathcal{T}_{(i)}^{j_0} \to \mathcal{T}_{(i)}$ almost surely as $j_0 \to \infty$ with respect to the GHP-topology.

Next, note that for $j_0 \in \mathbb{N}$,

$$\begin{aligned}\mathrm{d_{GHP}}\left(\frac{T_n(v_i)}{n^{1-1/\alpha}}, \alpha \cdot \mathcal{T}_{(i)}\right) \leq \sum_{j=1}^{j_0} \mathrm{d_{GHP}}\left(\frac{T_n(v_i, j)}{n^{1-1/\alpha}}, \alpha \cdot \mathcal{T}_{i,j}\right) + \alpha \cdot \mathrm{d_{GHP}}\left(\mathcal{T}_{(i)}^{j_0}, \mathcal{T}_{(i)}\right) \\ + \sup_{j>j_0} \mathrm{ht}\left(\frac{T_n(v_i, j)}{n^{1-1/\alpha}}\right) + \sum_{j>j_0} \frac{M_n(v_i, j)}{n}.\end{aligned}$$

We already know that the first term on the right-hand side converges a.s. to 0 as $n \to \infty$ (for j_0 fixed) and that the second term converges a.s. to 0 as $j_0 \to \infty$. Moreover, since $M_n(v_i) \leq n$, by dominated convergence we have

$$\mathbb{E}\left[\sum_{j>j_0} \frac{M_n(v_i, j)}{n}\right] = \mathbb{E}\left[\frac{M_n(v_i)}{n} - \sum_{j \leq j_0} \frac{M_n(v_i, j)}{n}\right] \underset{n\to\infty}{\longrightarrow} \mathbb{E}\left[M_{i+m}\left(1 - \sum_{j \leq j_0} \Delta_{i,j}\right)\right],$$

and then

$$\lim_{j_0\to\infty} \lim_{n\to\infty} \mathbb{E}\left[\sum_{j>j_0} \frac{M_n(v_i, j)}{n}\right] = 0.$$

Now note that

$$\begin{aligned}\limsup_{n\to\infty} \sum_{j>j_0} \mathbb{E}\left[\frac{(\mathrm{ht}(T_n(v_i, j))^{\alpha/(\alpha-1)}}{n}\right] &\leq \limsup_{n\to\infty} \sum_{j>j_0} \mathbb{E}\left[\frac{(\mathrm{ht}(T_{n,j}(v_i, j))^{\alpha/(\alpha-1)}}{M_n(v_i, j)} \cdot \frac{M_n(v_i, j)}{n}\right] \\ &\leq C_\alpha \limsup_{n\to\infty} \sum_{j>j_0} \mathbb{E}\left[\frac{M_n(v_i, j)}{n}\right],\end{aligned}$$

by [67, Lemma 33], where C_α is a finite constant depending only on α. So by Markov's inequality, we get

$$\lim_{j_0\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\sup_{j>j_0} \mathrm{ht}\left(\frac{T_n(v_i,j)}{n^{1-1/\alpha}}\right) > \varepsilon \right) = 0$$

for all $\varepsilon > 0$. Putting everything together, we obtain the convergence in probability

$$\mathrm{d_{GHP}}\left(\frac{T_n(v_i)}{n^{1-1/\alpha}}, \alpha \cdot \mathcal{T}_{(i)}\right) \xrightarrow{p} 0.$$

- *Final gluing.* Finally, the graph G_n^s is obtained by gluing appropriately the $2m-s$ trees $T_n(e_k), T_n(v_i), 1 \le k \le m, 1 \le i \le m-s$ along the kernel K^s. Using the results above, it therefore converges in probability, after multiplication of distances by $n^{-(1-1/\alpha)}$, to a version of $\alpha \cdot \mathcal{G}(\mathsf{K}^s)$. □

From this we immediately obtain the joint distribution of the edge-lengths of the continuous kernel $\mathcal{K}^s$. Given that the number of edges of $\mathcal{K}^s$ is m and keeping the notation of the proofs, we see that the lengths of the m edges are given by $M_i^{1-1/\alpha} \cdot \Lambda_i, 1 \le i \le m$ where the Λ_i are i.i.d. $\mathrm{ML}(1-1/\alpha, 1-1/\alpha)$ random variables (this is the distribution of the distance between a uniform leaf and the root in a standard α-stable tree) and independent of $(M_1, \dots, M_{2m-s})$. We may combine Remark 5.33 and Lemma 5.26 to check that the distribution of this m-tuple of random variables coincides with the one of Proposition 5.7 when $n = 0$. More generally, we could deduce from (5.31) the joint distribution of the edge-lengths of the continuous marginals $\mathcal{G}_n^s$, $n \ge 0$. However, it is simpler to prove this directly using urn arguments similar to those above.

Proof Proposition 5.7. Fix $n_0 \ge 0$. We work conditionally on $\mathsf{G}_{n_0}^s = (V, E)$. For each edge $e \in E$ and each $n \ge n_0$, let $L_n(e)$ denote the length of e in G_n^s and let $L_n^{\mathrm{tot}} := \sum_{e\in E} L_n(e)$. From the algorithmic construction of $(\mathsf{G}_n^s, n \ge n_0)$ we get that

(a) the process

$$(L_n^{\mathrm{tot}}, n \ge n_0)$$

is a triangular urn scheme as defined in Theorem 5.32 with initial weights

$$a = |E|, \quad b = \frac{(n_0+s)\alpha + s - 1}{\alpha - 1} - |E|$$

(b is the initial total weight of the vertices of $\mathsf{G}_{n_0}^s$, divided by $\alpha - 1$) and additional weight parameters $\gamma = 1$ and $\beta = \alpha/(\alpha-1)$;

(b) the jumps of the process $((L_n(e), e \in E), n \ge n_0)$ evolve according to Pólya's urn with initial weights $a_i = 1, 1 \le i \le |E|$, and additional weight parameter $\beta = 1$, independently of L_n^{tot}.

Theorem 5.32 and Theorem 5.30 therefore imply that $\left(L_n(e)/n^{1-1/\alpha}, e \in E\right)$ converges almost surely to a random vector with distribution (5.11). The conclusion then follows from the convergence (5.30). □

5.4.3 The line-breaking construction

The proof of Theorem 5.8 for $s \geq 1$ is inspired by the approach used in [63] to obtain a line-breaking construction of the stable trees. As we have already mentioned, we rely again on the algorithmic construction of the sequence $(\mathsf{G}_n^s, n \geq 0)$. The notation below coincides with that of Section 5.1.2. Moreover, for each n, we let H_n^s denote the combinatorial shape of $\mathcal{H}_n^s$. The metric space $\mathcal{H}_n^s$ is then interpreted as a finite graph (the graph H_n^s) with edge-lengths. We let L_n denote this sequence of edge-lengths, ordered arbitrarily, and let W_n denote the sequence of weights at internal vertices of $\mathcal{H}_n^s$ (i.e. the weights attributed by the measure η_n to each of these vertices), also ordered arbitrarily. We start with a preliminary lemma.

Lemma 5.25. *Given H_k^s, $0 \leq k \leq n$, and in particular that H_n^s has m edges and $m-(n+s)$ internal vertices with degrees $d_1, \dots, d_{m-(n+s)}$, we have*

$$(L_n, W_n) \overset{(\mathrm{d})}{=} \mathrm{ML}\left(1-\frac{1}{\alpha}, \frac{(n+s)\alpha+(s-1)}{\alpha}\right) \cdot \mathrm{Dir}\left(\underbrace{1,\dots,1}_{m}, \frac{d_1-1-\alpha}{\alpha-1}, \dots, \frac{d_{m-(n+s)}-1-\alpha}{\alpha-1}\right),$$

the random variables on the right-hand side being independent. In particular,

$$L_n \overset{(\mathrm{d})}{=} \mathrm{ML}\left(1-\frac{1}{\alpha}, \frac{(n+s)\alpha+(s-1)}{\alpha}\right) \cdot \mathrm{Beta}\left(m, \frac{(n+s)\alpha+s-1}{\alpha-1}-m\right) \cdot \mathrm{Dir}\,(1,\dots,1)\,.$$

Proof. For $n=0$, the first identity in distribution holds by definition of $(\mathcal{H}_0^s, \eta_0)$ in the line-breaking construction. The rest of the proof proceeds by induction on n, and is based essentially on manipulations of Dirichlet distributions. The steps are exactly the same as those of Proposition 3.2 in [63], to which we refer the interested reader. The only slight change to highlight is that here the degrees $d_1, \dots, d_{m-(n+s)}$ of the internal vertices of a graph in $\mathbb{M}_{s,n}$ with m edges necessarily satisfy

$$\sum_{i=1}^{m-(n+s)} \frac{d_i-1-\alpha}{\alpha-1} = \frac{(n+s)\alpha+s-1}{\alpha-1} - m,$$

as already observed in (5.8). This fact is also used, together with Lemma 5.26, to deduce the distribution of L_n from that of the pair (L_n, W_n). ☐

Proof of Theorem 5.8. Note that the metric spaces $\mathcal{H}_n^s, n \geq 0$ have implicit leaf-labels, given by their order of appearance in the construction. The metric spaces $\mathcal{G}_n^s, n \geq 0$ are also leaf-labelled by construction. Both models are sampling consistent: the metric space indexed by n is obtained from the metric space indexed by $n+1$ by removing the leaf labelled $n+1$ and the adjacent line-segment (this description is a little informal but hopefully clear). Hence, we only need to prove that, for all $n \geq 0$,

$$\mathcal{H}_n^s \overset{\mathrm{d}}{=} \mathcal{G}_n^s, \tag{5.33}$$

these compact metric spaces being implicitly endowed with the uniform distribution on their leaves, and still leaf-labelled. Together with the sampling consistency, this will imply that the processes of compact measured metric spaces $(\mathcal{H}_n^s, n \geq 0)$ and $(\mathcal{G}_n^s, n \geq 0)$ have the same distribution. Since $\mathcal{G}^s$ is the almost sure GHP-scaling limit of $\mathcal{G}_n^s$ (Lemma 5.24) and since $(\mathscr{C}, \mathrm{d_{GHP}})$ is complete, this will in turn entail that $\mathcal{H}_n^s$ converges a.s. to a random compact measured metric space distributed as $\mathcal{G}^s$.

To prove (5.33), we first check that the sequence of finite graphs $(\mathsf{H}_n^s, n \geq 0)$ evolves according to Marchal's algorithm, as does $(\mathsf{G}_n^s, n \geq 0)$. This relies on Lemmas 5.25 and 5.27 which imply that for each n, given $(\mathsf{H}_k^s, 0 \leq k \leq n)$, the probability that the new segment in the line-breaking construction is attached to a given edge of H_n^s is proportional to 1, whereas the probability that it is attached to a given vertex with degree $d_i \geq 3$ is proportional to $(d_i - 1 - \alpha)/(\alpha - 1)$. Hence, the sequences of graphs $(\mathsf{H}_n^s, n \geq 0)$ and $(\mathsf{G}_n^s, n \geq 0)$ have the same distribution since $\mathsf{G}_0^s = \mathsf{H}_0^s = \mathsf{K}^s$, including leaf-labels. Then we get (5.33) by simply noticing that the distribution of the edge-lengths of $\mathcal{H}_n^s$ given $(\mathsf{H}_k^s, 0 \leq k \leq n)$ is the same as that of the edge-lengths of $\mathcal{G}_n^s$ given $(\mathsf{G}_k^s, 0 \leq k \leq n)$, by Lemma 5.25 and Proposition 5.7. □

5.5 Appendix: distributions, urn models and applications

We detail in this appendix some classical asymptotic results on urn models that are needed at various points in the paper. We first recall the definitions and some properties of several distributions that are related to these asymptotics.

5.5.1 Some probability distributions

For more detail on the material in this section, we refer to Pitman [104].

Definitions and moments

Beta distributions. For parameters $a, b > 0$, the Beta(a, b) distribution has density

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$$

with respect to the Lebesgue measure on $(0, 1)$. If $B \sim \text{Beta}(a, b)$, then for $p, q \in \mathbb{R}_+$,

$$\mathbb{E}\left[B^p(1-B)^q\right] = \frac{\Gamma(a+b)}{\Gamma(a+b+p+q)} \frac{\Gamma(a+p)}{\Gamma(a)} \frac{\Gamma(b+q)}{\Gamma(b)}. \tag{5.34}$$

Dirichlet distributions. For parameters $a_1, a_2, \dots, a_n > 0$, the Dirichlet distribution $\text{Dir}(a_1, a_2, \dots, a_n)$ has density

$$\frac{\Gamma(\sum_{i=1}^n a_i)}{\prod_{i=1}^n \Gamma(a_i)} \prod_{j=1}^n x_i^{a_j - 1}$$

with respect to the Lebesgue measure on the simplex $\{(x_1, \dots, x_n) \in [0, 1]^n : \sum_{i=1}^n x_i = 1\}$. When $(X_1, \dots, X_n) \sim \text{Dir}(a_1, \dots, a_n)$, for $k_1, \dots, k_n \in \mathbb{R}_+$,

$$\mathbb{E}\left[X_1^{k_1} X_2^{k_2} \dots X_n^{k_n}\right] = \frac{\Gamma\left(\sum_{i=1}^n a_i\right)}{\Gamma(\sum_{i=1}^n (a_i + k_i))} \cdot \prod_{i=1}^n \frac{\Gamma(a_i + k_i)}{\Gamma(a_i)}. \tag{5.35}$$

Generalized Mittag-Leffler distributions. Let $0 < \beta < 1$, $\theta > -\beta$. An $\mathbb{R}_+$-valued random variable M has the generalized Mittag-Leffler distribution ML(β, θ) if, for all suitable test functions f, we have

$$\mathbb{E}\left[f(M)\right] = \frac{\mathbb{E}\left[\sigma_\beta^{-\theta} f\left(\sigma_\beta^{-\beta}\right)\right]}{\mathbb{E}\left[\sigma_\beta^{-\theta}\right]}, \tag{5.36}$$

where σ_β is a stable random variable with Laplace transform $\mathbb{E}[e^{-\lambda \sigma_\beta}] = \exp\left(-\lambda^\beta\right), \lambda \geq 0$. For $p \in \mathbb{R}_+$,

$$\mathbb{E}\left[M^p\right] = \frac{\Gamma(\theta)\Gamma(\theta/\beta + p)}{\Gamma(\theta/\beta)\Gamma(\theta + p\beta)} = \frac{\Gamma(\theta+1)\Gamma(\theta/\beta + p + 1)}{\Gamma(\theta/\beta + 1)\Gamma(\theta + p\beta + 1)}.$$

Poisson-Dirichlet distributions. Let $0 < \beta < 1$, $\theta > -\beta$ and for $i \geq 1$, let $B_i \sim \mathrm{Beta}(1-\beta, \theta + i\beta)$ independently. Then the decreasing sequence $(P_i)_{i\geq 1} = (Q_i^{\downarrow})_{i\geq 1}$ where $Q_j = B_j \prod_{i=1}^{j-1}(1-B_i)$ has the $\mathrm{PD}(\beta,\theta)$ distribution. The almost sure limit $W := \Gamma(1-\beta)\lim_{i\to\infty} i(P_i^{\downarrow})^{\beta}$ has the $\mathrm{ML}(\beta,\theta)$ distribution.

Distributional properties

Lemma 5.26. *If $(X_1,\ldots,X_n) \sim \mathrm{Dir}(a_1,\ldots,a_n)$ then for all $1 \leq m \leq n-1$, $(X_1,\ldots,X_m)$ is distributed as the product of two independent random variables:*

$$\mathrm{Beta}\left(\sum_{i=1}^{m} a_i, \sum_{i=m+1}^{n} a_i\right) \cdot \mathrm{Dir}(a_1,\ldots,a_m).$$

Lemma 5.27. *Suppose that $(X_1,X_2,\ldots,X_n) \sim \mathrm{Dir}(a_1,a_2,\ldots,a_n)$. Let I be the index of a size-biased pick from amongst the co-ordinates i.e. $\mathbb{P}(I=i|X_1,X_2,\ldots,X_n) = X_i$, for $1 \leq i \leq n$. Then*

$$\mathbb{P}(I=i) = \frac{a_i}{a_1+a_2+\ldots+a_n}$$

for $1 \leq i \leq n$ and, conditionally on $I=i$,

$$(X_1,X_2,\ldots,X_n) \sim \mathrm{Dir}(a_1,\ldots,a_{i-1},a_i+1,a_{i+1},\ldots,a_n).$$

Lemma 5.28. *Let $0<\beta<1$, $\theta>-\beta$, and let $(P_i)_{i\geq1}$ have distribution $\mathrm{PD}(\beta,\theta)$. Let J be the index of a size-biased pick from this sequence, i.e. $\mathbb{P}(J=j \mid (P_i)_{i\geq 1}) = P_j$, for $j \geq 1$. We let $(P_i')_{i\geq1}$ be the decreasing sequence $(1-P_J)^{-1}\cdot(P_i)_{i\geq1, i\neq J}$, reindexed by $\mathbb{N}$. Then*

$$P_J \sim \mathrm{Beta}(1-\beta,\theta+\beta) \quad \textit{and} \quad (P_i')_{i\geq1} \sim \mathrm{PD}(\beta,\theta+\beta),$$

and these two random variables are independent.

Let $\mathbb{N}^{n,\neq} := \{(i_1,\ldots,i_n) \in \mathbb{N}^n \mid i_1,\ldots,i_n \text{ are distinct}\}$.

Lemma 5.29. *Let $(P_i)_{i\geq1} \sim \mathrm{PD}(\beta,\theta)$ with $0<\beta<1$ and $\theta>-\beta$. Then for all $k_1,k_2,\ldots,k_n \in [1,\infty)$,*

$$\mathbb{E}\left[\sum_{(i_1,\ldots,i_n)\in\mathbb{N}^{n,\neq}} P_{i_1}^{k_1}\ldots P_{i_n}^{k_n}\right] = \left(\prod_{i=1}^{n}\beta\frac{\Gamma(k_i-\beta)}{\Gamma(1-\beta)}\right)\frac{\Gamma(\theta)}{\Gamma(\theta+\sum_{j=1}^{n}k_j)}\frac{\Gamma(\theta/\beta+n)}{\Gamma(\theta/\beta)}. \tag{5.37}$$

In particular, for $(P_i)_{i\geq1} \sim \mathrm{PD}(\alpha-1,\alpha-1)$ with $\alpha \in (1,2)$, and $k_1,\ldots k_n \in \mathbb{N}$, we have

$$\begin{aligned}\mathbb{E}\left[\sum_{(i_1,\ldots,i_n)\in\mathbb{N}^{n,\neq}} P_{i_1}^{k_1}\ldots P_{i_n}^{k_n}\right] &= \left(\prod_{i=1}^{n}(\alpha-1)\frac{\Gamma(k_i+1-\alpha)}{\Gamma(2-\alpha)}\right)\frac{\Gamma(\alpha-1)\; n!}{\Gamma(\alpha-1+\sum_{j=1}^{n}k_j)} \\ &= \left(\prod_{i=1}^{n} w_{k_i+1}\right)\frac{\Gamma(\alpha-1)\; n!}{\Gamma(\alpha-1+\sum_{j=1}^{n}k_j)},\end{aligned}$$

where the weights $w_1,w_2,\ldots$ are defined in (5.6).

Proof. We proceed by induction on n. For $n=0$ we use the convention that the left-hand side of (5.37) is 1 and so the identity is true. Let $n \geq 1$ and suppose that the identity is true for $n-1$. Then letting J be such that $\mathbb{P}\left(J=j \mid (P_i)_{i\geq 1}\right) = P_j$, we have

$$\begin{aligned}
&\mathbb{E}\left[\sum_{(i_1,\ldots,i_n)\in\mathbb{N}^{n,\neq}} P_{i_1}^{k_1}\ldots P_{i_n}^{k_n}\right] \\
&\quad = \mathbb{E}\left[P_J^{k_n-1}(1-P_J)^{k_1+\cdots+k_{n-1}} \sum_{\substack{(i_1,\ldots,i_{n-1})\\ \in(\mathbb{N}\setminus\{J\})^{n-1,\neq}}} \left(\frac{P_{i_1}}{1-P_J}\right)^{k_1}\cdots\left(\frac{P_{i_{n-1}}}{1-P_J}\right)^{k_{n-1}}\right] \\
&\quad = \mathbb{E}\left[P_J^{k_n-1}(1-P_J)^{k_1+\cdots+k_{n-1}}\right]\cdot\mathbb{E}\left[\sum_{(i_1,\cdots,i_{n-1})\in\mathbb{N}^{n-1,\neq}} (P'_{i_1})^{k_1}\cdots(P'_{i_{n-1}})^{k_{n-1}}\right],
\end{aligned}$$

by Lemma 5.28, where $(P'_i)_{i\geq 1} \sim \mathrm{PD}(\beta,\beta+\theta)$ and $P_J \sim \mathrm{Beta}(1-\beta,\theta+\beta)$. Using (5.34), we have

$$\begin{aligned}
\mathbb{E}\left[P_J^{k_n-1}(1-P_J)^{k_1+\cdots+k_{n-1}}\right] &= \frac{\Gamma(1+\theta)\Gamma(1-\beta+k_n-1)\Gamma(\theta+\beta+\sum_{i=1}^{n-1}k_i)}{\Gamma(\theta+\beta)\Gamma(1-\beta)\Gamma(1+\theta+\sum_{i=1}^{n}k_i-1)} \\
&= \left(\beta\frac{\Gamma(k_n-\beta)}{\Gamma(1-\beta)}\right)\frac{\Gamma(\theta)\Gamma(\theta+\beta+\sum_{i=1}^{n-1}k_i)}{\Gamma(\theta+\beta)\Gamma(\theta+\sum_{j=1}^{n}k_j)}\frac{\theta}{\beta}.
\end{aligned}$$

The induction hypothesis applied to the sequence $(P'_i)_{i\geq 1}$, which has distribution $\mathrm{PD}(\beta,\beta+\theta)$, then yields

$$\begin{aligned}
&\mathbb{E}\left[\sum_{(i_1,\ldots,i_{n-1})\in\mathbb{N}^{n-1,\neq}} (P'_{i_1})^{k_1}\ldots(P'_{i_{n-1}})^{k_{n-1}}\right] \\
&\quad = \left(\prod_{i=1}^{n-1}\beta\frac{\Gamma(k_i-\beta)}{\Gamma(1-\beta)}\right)\frac{\Gamma(\theta+\beta)}{\Gamma(\theta+\beta+\sum_{j=1}^{n-1}k_j)}\frac{\Gamma((\theta+\beta)/\beta+n-1)}{\Gamma((\theta+\beta)/\beta)} \\
&\quad = \left(\prod_{i=1}^{n-1}\beta\frac{\Gamma(k_i-\beta)}{\Gamma(1-\beta)}\right)\frac{\Gamma(\theta+\beta)}{\Gamma(\theta+\beta+\sum_{j=1}^{n-1}k_j)}\frac{\Gamma(\theta/\beta+n)}{(\theta/\beta)\Gamma(\theta/\beta)},
\end{aligned}$$

and the result for n follows by multiplying the last display with the preceding one. □

5.5.2 Pólya's urn, Chinese restaurant processes and triangular urn schemes

We gather here some classical results for urn models.

Theorem 5.30 (Pólya's urn)**.** *Consider an urn model with k colours, with initial weights $a_1,\ldots,a_k>0$ respectively. At each step, draw a colour with a probability proportional to its weight and add an extra weight $\beta>0$ to this colour. Let $W_n^{(1)},\ldots,W_n^{(k)}$ denote the weights of the k colours after n steps. Then*

$$\left(\frac{W_n^{(1)}}{\beta n},\ldots,\frac{W_n^{(k)}}{\beta n}\right) \xrightarrow[n\to\infty]{\text{a.s.}} (W^{(1)},\ldots,W^{(k)})$$

where $(W^{(1)}, \ldots, W^{(k)}) \sim \mathrm{Dir}(a_1/\beta, \ldots, a_k/\beta)$.

Theorem 5.31 (The Chinese restaurant process)**.** *Fix two parameters* $\beta \in (0,1)$ *and* $\theta > -\beta$. *The process starts with one table occupied by a single customer and then evolves in a Markovian way as follows: given that at step* n *there are* k *occupied tables with* n_i *customers at table* i*, a new customer is placed at table* i *with probability* $(n_i - \beta)/(n+\theta)$ *and placed at a new table with probability* $(\theta + k\beta)/(n+\theta)$. *Let* $N_i(n), i \geq 1$ *be the number of customers at table* i *at step* n *and let* $(N_i^{\downarrow}(n), i \geq 1)$ *be the decreasing rearrangement of these terms. Let* $K(n)$ *denote the number of occupied tables at step* n. *Then*

$$\left(\frac{N_i^{\downarrow}(n), i \geq 1}{n}\right) \underset{n\to\infty}{\xrightarrow{\text{a.s. in } \ell^1}} (Y_i, i \geq 1) \quad \textit{and} \quad \frac{K(n)}{n^{\beta}} \underset{n\to\infty}{\xrightarrow{\text{a.s.}}} W$$

where $(Y_i, i \geq 1) \sim \mathrm{PD}(\beta, \theta)$ *and* $W \sim \mathrm{ML}\,(\beta, \theta)$.

We refer to Pitman's book [104, Chapter 3] for more detail on these first two theorems.

Theorem 5.32 (Triangular urn schemes)**.** *Consider an urn model with two colours, red and black. Suppose that initially red has weight* $a > 0$ *and black has weight* $b \geq 0$. *At each step, we sample a colour with probability proportional to its current weight in the urn. Let* $\beta > \gamma > 0$ *and assume that when red is drawn then weight* γ *is added to red and weight* $\beta - \gamma$ *to black, whereas when black is drawn then weight* β *is added to black (and nothing to red). Let* R_n *denote the red weight after* n *steps. Then,*

$$\frac{R_n}{n^{\gamma/\beta}} \underset{n\to\infty}{\xrightarrow{\text{a.s.}}} R$$

where the random variable R *is such that* $R \sim \gamma \cdot \mathrm{Beta}(\frac{a}{\gamma}, \frac{b}{\gamma}) \cdot \mathrm{ML}\left(\frac{\gamma}{\beta}, \frac{(a+b)}{\beta}\right)$ *with the Beta and Mittag-Leffler random variables being independent, and the convention that* $\mathrm{Beta}(a, 0) = 1$ *a.s.*

(Note that, since the total weight in the urn at step n is $a+b+n\beta$, we trivially deduce that the black weight $B_n = a+b+n\beta - R_n$ satisfies $B_n/n \to \beta$ almost surely.) There is a vast literature on triangular urn schemes, which give rise to profoundly different asymptotic behaviour. We refer to Janson [80] for an overview, and in particular to Theorems 1.3 and 1.7 therein which together imply the convergence of Theorem 5.32 (but only in distribution). The almost sure convergence can, in fact, be deduced from Theorems 5.30 and 5.31. Observe first that we may reduce to the case $\gamma = 1$ by scaling. Now note that in the context of Theorem 5.32 when $\gamma = 1$ and $b = 0$, the red weight evolves as a plus the number of occupied tables in a Chinese restaurant process with parameters $(1/\beta, a/\beta)$, and so the almost sure limit has $\mathrm{ML}(1/\beta, a/\beta)$ distribution. To treat the case $b > 0$, consider a refinement of the urn model in which the red colour comes in two variants, light and dark. Start with a light red weight, b dark red weight and 0 black weight. Sample a colour with probability proportional to its current weight in the urn. When black is drawn, add weight β to black. When red is drawn in either of its variants, add weight 1 to that variant and weight $\beta - 1$ to black. Clearly, light red and dark red + black taken together follow the β-triangular urn scheme with respective initial weights a and b. Moreover, (1) the proportion of the total red weight which is light red converges almost surely to a random variable with

Beta(a, b) distribution by Theorem 5.30, and (2) this evolution holds independently of that of the total proportion of red weight in the urn, which converges to a ML$(1/\beta, (a+b)/\beta)$-distributed random variable, by the Chinese restaurant process as noted above.

We finally turn to the proof of Proposition 5.15. The notation is introduced in the vicinity of its statement in Section 5.2.1.

Proof of Proposition 5.15. Imagine first not distinguishing between the different types of a colour, i.e. consider the evolution of

$$X_i^{a,b,c}(n) = X_i^a(n) + X_i^b(n) + X_i^c(n), \quad 1 \le i \le k.$$

Then $(X_1^{a,b,c}(n), \dots, X_k^{a,b,c}(n))_{n\ge 0}$ performs a classical Pólya's urn in which we always add weight α of the colour picked, and which is started from

$$(X_1^{a,b,c}(0), \dots, X_k^{a,b,c}(0)) = (\gamma_1, \dots, \gamma_k).$$

So we have

$$\frac{1}{\alpha n}(X_1^{a,b,c}(n), \dots, X_k^{a,b,c}(n)) \to (D_1, \dots, D_k) \tag{5.38}$$

almost surely as $n \to \infty$, where $(D_1, \dots, D_k) \sim \mathrm{Dir}(\gamma_1/\alpha, \dots, \gamma_k/\alpha)$. Observe that $(X_i^{a,b,c}(n) - \gamma_i)/\alpha$ is the number of times by step n that colour i has been picked.

Now consider the triangular sub-urn which just watches the evolution of colour i, which doesn't distinguish between types a and b, but does distinguish type c. In particular, at each step we pick either type $\{a, b\}$ or type c with probability proportional to its current weight. If we pick $\{a, b\}$, we add 1 to its weight and $\alpha - 1$ to the weight of c; if we pick c, we simply add weight α to c. Write $Y_i^{a,b}(n)$ and $Y_i^c(n)$ for the weights after n steps within this urn, with $Y_i^{a,b}(0) = \gamma_i$ and $Y_i^c(0) = 0$. Then by Theorem 5.32, we have

$$\frac{1}{n^{1/\alpha}} Y_i^{a,b}(n) \to R_i, \quad \frac{1}{\alpha n} Y_i^c(n) \to 1, \tag{5.39}$$

almost surely as $n \to \infty$, where $R_i \sim \mathrm{ML}(1/\alpha, \gamma_i/\alpha)$. Moreover, the number of times we add to type a or b is $Y_i^{a,b}(n) - \gamma_i$.

Now consider the sub-urn which just watches the evolution of types a and b of colour i. So if we pick a, we add weight $\alpha - 1$ to a and $2 - \alpha$ to b, whereas if we pick b we just add weight 1 to b. Write $Z_i^a(n)$ and $Z_i^b(n)$ for the weights of types a and b after n steps of this sub-urn, with $Z_i^a(0) = \gamma_i$ and $Z_i^b(0) = 0$. Then again by Theorem 5.32 we have

$$\frac{1}{(\alpha - 1)n^{\alpha-1}} Z_i^a(n) \to \bar{R}_i, \quad \frac{1}{n} Z_i^b(n) \to 1 \tag{5.40}$$

almost surely, where $\bar{R}_i \sim \mathrm{ML}(\alpha - 1, \gamma_i)$. Finally, observe that the full urn process may be decomposed as follows:

$$\begin{aligned}
X_i^a(n) &= Z_i^a\left(Y_i^{a,b}\left(\frac{X_i^{a,b,c}(n) - \gamma_i}{\alpha}\right) - \gamma_i\right)\\
X_i^b(n) &= Z_i^b\left(Y_i^{a,b}\left(\frac{X_i^{a,b,c}(n) - \gamma_i}{\alpha}\right) - \gamma_i\right)\\
X_i^c(n) &= Y_i^c\left(\frac{X_i^{a,b,c}(n) - \gamma_i}{\alpha}\right),
\end{aligned}$$

where the processes $(X_1^{a,b,c}(n), \dots, X_k^{a,b,c}(n))_{n\geq 0}$, $(Y_i^{a,b}(n), Y_i^c(n))_{n\geq 0}$ for $1 \leq i \leq k$, and $(Z_i^a(n), Z_i^b(n))_{n\geq 0}$ for $1 \leq i \leq k$, are all independent. The claimed results then follow by composing the limits (5.38), (5.39) and (5.40). $\square$

Remark 5.33. *The following statements follow using similar arguments:*

$$(D_1^{1/\alpha} R_1, \dots, D_k^{1/\alpha} R_k) \overset{(\mathrm{d})}{=} R \cdot (\tilde{D}_1, \dots, \tilde{D}_k),$$

where $R \sim \mathrm{ML}(1/\alpha, \gamma/\alpha)$ *is independent of* $(\tilde{D}_1, \dots, \tilde{D}_k) \sim \mathrm{Dir}(\gamma_1, \dots, \gamma_k)$, *and*

$$R_i^{\alpha-1} \bar{R}_i \sim \mathrm{ML}(1 - 1/\alpha, \gamma_i/\alpha)$$

for $1 \leq i \leq k$.

Acknowledgements

B.H. and D.S.'s research is partially supported by the ANR GRAAL ANR-14-CE25-0014. C.G.'s research is supported by EPSRC Fellowship EP/N004833/1. She is grateful to LIX, École polytechnique for two professeur invité positions in November 2016 and November 2017, and for funding from Université Paris 13 as part of the Projet MathStic to pay for a week's visit in March 2018 and for a professeur invité position in September 2018, all of which facilitated work on this project.

Bibliographie

[1] Romain Abraham, Jean-François Delmas, and Patrick Hoscheit. A note on the Gromov–Hausdorff–Prokhorov distance between (locally) compact metric measure spaces. *Electronic Journal of Probability.* 18(0), 2013.

[2] Louigi Addario-Berry. Most trees are short and fat. *Probability Theory and Related Fields*, 2019.

[3] Louigi Addario-Berry, Omer Angel, Guillaume Chapuy, Éric Fusy, and Christina Goldschmidt. Voronoi tessellations in the CRT and continuum random maps of finite excess. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 933–946. SIAM, Philadelphia, PA, 2018.

[4] Louigi Addario-Berry, Nicolas Broutin, and Christina Goldschmidt. Critical random graphs: limiting constructions and distributional properties. *Electronic Journal of Probability.* 15(0):741–775, 2010.

[5] Louigi Addario-Berry, Nicolas Broutin, and Christina Goldschmidt. The continuum limit of critical random graphs. *Probability Theory and Related Fields.* 152(3):367–406, 2012.

[6] Louigi Addario-Berry, Nicolas Broutin, Christina Goldschmidt, and Grégory Miermont. The scaling limit of the minimum spanning tree of the complete graph. *The Annals of Probability.* 45(5):3075–3144, 2017.

[7] Louigi Addario-Berry and Sanchayan Sen. Geometry of the minimal spanning tree of a random 3-regular graph. 2018. arXiv: 1810.03802 [math.PR].

[8] David Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability.* 25(2):812–854, 1997.

[9] David Aldous. The continuum random tree. I. *The Annals of Probability.* 19(1):1–28, 1991.

[10] David Aldous. The continuum random tree. III. *The Annals of Probability.* 21(1):248–289, 1993.

[11] David Aldous and Vlada Limic. The entrance boundary of the multiplicative coalescent. *Electronic Journal of Probability.* 3:No. 3, 59 pages, 1998.

[12] Omid Amini, Luc Devroye, Simon Griffiths, and Neil Olver. Explosion and linear transit times in infinite trees. *Probability Theory and Related Fields.* 167(1):325–347, 2017.

[13] Bengt von Bahr and Carl-Gustav Esseen. Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Annals of Mathematical Statistics.* 36:299–303, 1965.

[14] Cyril Banderier, Philippe Marchal, and Michael Wallner. Periodic Pólya urns and an application to Young tableaux. In *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms.* Volume 110, LIPIcs. Leibniz Int. Proc. Inform. Art. No. 11, 13. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.

[15] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*. 286(5439):509–512, 1999.

[16] Edward A. Bender and E. Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory. Series A*. 24(3):296–307, 1978.

[17] Noam Berger, Christian Borgs, Jennifer T. Chayes, and Amin Saberi. Asymptotic behavior and distributional limits of preferential attachment graphs. *The Annals of Probability*. 42(1):1–40, 2014.

[18] Shankar Bhamidi, Nicolas Broutin, Sanchayan Sen, and Xuan Wang. Scaling limits of random graph models at criticality: universality and the basin of attraction of the Erdős-Rényi random graph. 2014. arXiv: 1411.3417 [math.PR].

[19] Shankar Bhamidi, Amarjit Budhiraja, and Xuan Wang. The augmented multiplicative coalescent, bounded size rules and critical dynamics of random graphs. *Probability Theory and Related Fields*. 160(3-4):733–796, 2014.

[20] Shankar Bhamidi, Souvik Dhara, Remco van der Hofstad, and Sanchayan Sen. Universality for critical heavy-tailed network models: metric structure of maximal components. 2017. arXiv: 1703.07145 [math.PR].

[21] Shankar Bhamidi, Remco van der Hofstad, and Johan S. H. van Leeuwaarden. Novel scaling limits for critical inhomogeneous random graphs. *The Annals of Probability*. 40(6):2299–2361, 2012.

[22] Shankar Bhamidi, Remco van der Hofstad, and Johan S. H. van Leeuwaarden. Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electronic Journal of Probability*. 15:no. 54, 1682–1703, 2010.

[23] Shankar Bhamidi, Remco van der Hofstad, and Sanchayan Sen. The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs. *Probability Theory and Related Fields*. 170(1-2):387–474, 2018.

[24] Shankar Bhamidi and Sanchayan Sen. Geometry of the vacant set left by random walks on random graphs, Wright's contants, and critical random graphs with prescribed degrees. 2016. arXiv: 1608.07153 [math.PR].

[25] Shankar Bhamidi, Sanchayan Sen, and Xuan Wang. Continuum limit of critical inhomogeneous random graphs. *Probability Theory and Related Fields*. 169(1-2):565–641, 2017.

[26] John D. Biggins. Uniform convergence of martingales in the branching random walk. *The Annals of Probability*. 20(1):137–151, 1992.

[27] Patrick Billingsley. *Convergence of probability measures*. Wiley series in probability and statistics. Probability and statistics section. Wiley, New York, 2nd ed edition, 1999.

[28] Benjamin Bloem-Reddy and Peter Orbanz. Preferential attachment and vertex arrival times. 2017. arXiv: 1710.02159 [math.PR].

[29] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*. 1(4):311–316, 1980.

[30] Konstantin A. Borovkov and Vladimir A. Vatutin. On the asymptotic behaviour of random recursive trees in random environments. *Advances in Applied Probability*. 38(4):1047–1070, 2006.

[31] Nicolas Broutin, Thomas Duquesne, and Minmin Wang. Limits of multiplicative inhomogeneous random graphs and Lévy trees. 2018. arXiv: 1804.05871 [math.PR].

[32] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001, pages xiv+415.

[33] Shirshendu Chatterjee and Rick Durrett. Contact processes on random graphs with power law degree distributions have critical value 0. *The Annals of Probability*. 37(6):2332–2356, 2009.

[34] Loïc Chaumont. Excursion normalisée, méandre et pont pour les processus de Lévy stables. *Bulletin des Sciences Mathematiques*. 121(5):377–404, 1997.

[35] Brigitte Chauvin, Michael Drmota, and Jean Jabbour-Hattab. The profile of binary search trees. *Annals of Applied Probability*:1042–1062, 2001.

[36] Brigitte Chauvin, Thierry Klein, Jean-François Marckert, and Alain Rouault. Martingales and profile of binary search trees. *Electronic Journal of Probability*. 10(0):420–435, 2005.

[37] Bo Chen, Daniel Ford, and Matthias Winkel. A new family of Markov branching trees: the alpha-gamma model. *Electronic Journal of Probability*. 14:400–430, 2009.

[38] Guillaume Conchon-Kerjan and Christina Goldschmidt. Stable graphs: the metric space scaling limits of critical random graphs with i.i.d. power-law degrees. In preparation, 2019.

[39] Miklós Csörg. On the strong law of large numbers and the central limit theorem for martingales. *Transactions of the American Mathematical Society*. 131(1):259–275, 1968.

[40] Nicolas Curien, Thomas Duquesne, Igor Kortchemski, and Ioan Manolescu. Scaling limits and influence of the seed graph in preferential attachment trees. *Journal de l'École Polytechnique — Mathématiques*. 2:1–34, 2015.

[41] Nicolas Curien and Bénédicte Haas. Random trees constructed by aggregation. *Annales de l'institut Fourier*. 67(5):1963–2001, 2017.

[42] Nicolas Curien and Bénédicte Haas. The stable trees are nested. *Probability Theory and Related Fields*. 157(3):847–883, 2013.

[43] Nicolas Curien and Igor Kortchemski. Random stable looptrees. *Electronic Journal of Probability*. 19(0), 2014.

[44] Maria Deijfen, Henri van den Esker, Remco van der Hofstad, and Gerard Hooghiemstra. A preferential attachment model with random initial degrees. *Arkiv för Matematik*. 47(1):41–72, 2009.

[45] Luc Devroye. Branching processes in the analysis of the heights of trees. *Acta Informatica*. 24(3):277–298, 1987.

[46] Souvik Dhara, Remco van der Hofstad, Johan SH van Leeuwaarden, and Sanchayan Sen. Critical window for the configuration model: finite third moment degrees. *Electronic Journal of Probability*. 22, 2017.

[47] Souvik Dhara, Remco van der Hofstad, Johan van Leeuwaarden, and Sanchayan Sen. Heavy-tailed configuration models at criticality. 2016. arXiv: 1612.00650 [math.PR].

[48] Daphné Dieuleveut. The vertex-cut-tree of Galton–Watson trees converging to a stable tree. *The Annals of Applied Probability*. 25(4):2215–2262, 2015.

[49] Marian Dondajewski, Jerzy Szymańlskt, and Peter Kirschenhofer. Vertex-degrees in strata of a random recursive tree. In *North-Holland Mathematics Studies*. Volume 118, pages 99–105. Elsevier, 1985.

[50] Michael Drmota. *Random trees: an interplay between combinatorics and probability*. Springer Science & Business Media, 2009.

[51] Michael Drmota and Hsien-Kuei Hwang. Profiles of random trees: correlation and width of random recursive trees and binary search trees. *Advances in Applied Probability*. 37(2):321–341, 2005.

[52] Thomas Duquesne. A limit theorem for the contour process of conditioned Galton–Watson trees. *The Annals of Probability*. 31(2):996–1027, 2003.

[53] Thomas Duquesne and Jean-François Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*. 281:vi+147, 2002.

[54] Thomas Duquesne and Guanying Wang. Exceptionally small balls in stable trees. *Bulletin de la Société Mathématique de France*. 142(2):223–254, 2014.

[55] Thomas Duquesne and Minmin Wang. Decomposition of Lévy trees along their diameter. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*. 53(2):539–593, 2017.

[56] Rick Durrett. *Probability: Theory and Examples*. Cambridge University Press, Cambridge, 4th edition, 2010.

[57] David A Edwards. The structure of superspace. In *Studies in topology*, pages 121–133. Elsevier, 1975.

[58] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 5:17–61, 1960.

[59] Kenneth J. Falconer. *Fractal geometry: mathematical foundations and applications*. Wiley, Chichester, West Sussex, third edition edition, 2014.

[60] Daniel J. Ford. Probabilities on cladograms: introduction to the alpha model. 2005. arXiv: 0511246 [math.PR].

[61] Michael Fuchs, Hsien-Kuei Hwang, and Ralph Neininger. Profiles of random trees: limit theorems for random recursive trees and binary search trees. *Algorithmica*. 46(3):367–407, 2006.

[62] Christina Goldschmidt. Scaling limits of random trees and random graphs. To appear in Random Graphs, Phase Transitions and the Gaussian Free Field: Proceedings of the 2017 PIMS-CRM Summer School in Probability, Martin Barlow and Gordon Slade (eds.), available at `http://www.stats.ox.ac.uk/~goldschm/PIMSNotes.pdf`, 2018.

[63] Christina Goldschmidt and Bénédicte Haas. A line-breaking construction of the stable trees. *Electronic Journal of Probability*. 20(0), 2015.

[64] Christina Goldschmidt, Bénédicte Haas, and Delphin Sénizergues. Stable graphs: distributions and line-breaking construction, 2018. arXiv: 1811.06940 [math.PR].

[65] Mikhael Gromov. Groups of polynomial growth and expanding maps. *Institut des Hautes Études Scientifiques. Publications Mathématiques*. (53):53–73, 1981.

[66] Bénédicte Haas. Asymptotics of heights in random trees constructed by aggregation. *Electronic Journal of Probability*. 22(0), 2017.

[67] Bénédicte Haas and Grégory Miermont. Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees. *The Annals of Probability.* 40(6):2589–2666, 2012.

[68] Bénédicte Haas and Grégory Miermont. The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electronic Journal of Probability.* 9:57–97, 2004.

[69] Bénédicte Haas, Grégory Miermont, Jim Pitman, and Matthias Winkel. Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *The Annals of Probability.* 36(5):1790–1837, 2008.

[70] Bénédicte Haas, Jim Pitman, and Matthias Winkel. Spinal partitions and invariance under re-rooting of continuum random trees. *The Annals of Probability.* 37(4):1381–1411, 2009.

[71] Bénédicte Haas and Robin Stephenson. Scaling limits of k-ary growing trees. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques.* 51(4):1314–1341, 2015.

[72] Bénédicte Haas and Robin Stephenson. Scaling limits of multi-type Markov-branching trees. In preparation, 2019.

[73] Felix Hausdorff. *Grundzüge der Mengenlehre.* Veit and Company, Leipzig, 1914. Das Hauptwerk von Felix Hausdorff.

[74] Ella Hiesmayr and Ümit Işlak. Asymptotic results on Hoppe trees and its variations. 2017. arXiv: `1712.03572 [math.PR]`.

[75] Remco van der Hofstad. *Random graphs and complex networks*, number 43 in Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, Cambridge, 2017.

[76] Remco van der Hofstad. Critical behavior in inhomogeneous random graphs. *Random Structures & Algorithms.* 42(4):480–508, 2013.

[77] Remco van der Hofstad, Gerard Hooghiemstra, and Piet Van Mieghem. On the covariance of the level sizes in random recursive trees. *Random Structures & Algorithms.* 20(4):519–539, 2002.

[78] Miroslav Hušek. Urysohn universal space, its development and Hausdorff's approach. *Topology and its Applications.* 155(14):1493–1501, 2008.

[79] Lancelot F. James. Generalized Mittag-Leffler distributions arising as limits in preferential attachment models. 2015. arXiv: `1509.07150v4 [math.PR]`.

[80] Svante Janson. Limit theorems for triangular urn schemes. *Probability Theory and Related Fields.* 134(3):417–452, 2006.

[81] Svante Janson. Random recursive trees and preferential attachment trees are random split trees. 2017. arXiv: `1706.05487 [math.PR]`.

[82] Svante Janson, Donald E. Knuth, Tomasz Łuczak, and Boris Pittel. The birth of the giant component. *Random Structures & Algorithms.* 4(3):233–358, 1993.

[83] Adrien Joseph. The component sizes of a critical random graph with given degree sequence. *The Annals of Applied Probability.* 24(6):2560–2594, 2014.

[84] Zakhar Kabluchko, Alexander Marynych, and Henning Sulzbach. General Edgeworth expansions with applications to profiles of random trees. *The Annals of Applied Probability.* 27(6):3478–3524, 2017.

[85] Zsolt Katona. Width of a scale-free tree. *Journal of Applied Probability.* 42(3):839–850, 2005.

[86] Douglas P Kennedy. The distribution of the maximum Brownian excursion. *Journal of Applied Probability.* 13(2):371–376, 1976.

[87] Igor Kortchemski. Sub-exponential tail bounds for conditioned stable Bienaymé–Galton–Watson trees. *Probability Theory and Related Fields.* 168(1):1–40, 2017.

[88] Jean-François Le Gall. Geodesics in large planar maps and in the Brownian map. *Acta Mathematica.* 205(2):287–360, 2010.

[89] Jean-François Le Gall and Yves Le Jan. Branching processes in Lévy processes: the exploration process. *The Annals of Probability.* 26(1):213–252, 1998.

[90] Arthur Lubin. Extensions of measures and the von Neumann selection theorem. *Proceedings of the American Mathematical Society.* 43:118–122, 1974.

[91] Cécile Mailler and Gerónimo Uribe Bravo. Random walks with preferential relocations and fading memory: a study through random recursive trees. 2018. arXiv: 1810.02735 [math.PR].

[92] Philippe Marchal. A note on the fragmentation of a stable tree. *Discrete Mathematics & Theoretical Computer Science*, 2008.

[93] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canadian Journal of Mathematics. Journal Canadien de Mathématiques.* 30(5):997–1015, 1978.

[94] Julien Melleray. On the geometry of Urysohn's universal metric space. *Topology and its Applications.* 154(2):384–403, 2007.

[95] Grégory Miermont. Self-similar fragmentations derived from the stable tree II: splitting at nodes. *Probability Theory and Related Fields.* 131(3):341–375, 2005.

[96] Grégory Miermont. Tessellations of random maps of arbitrary genus. *Annales Scientifiques de l'École Normale Supérieure. Quatrième Série.* 42(5):725–781, 2009.

[97] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random structures & algorithms.* 6(2):161–180, 1995.

[98] J. W. Moon. The distance between nodes in recursive trees. In T. P. McDonough and V. C. Mavron, editors, *Combinatorics*, pages 125–132. Cambridge University Press, Cambridge, 1974.

[99] Erol A. Peköz, Adrian Röllin, and Nathan Ross. Generalized gamma approximation with rates for urns, walks and trees. *The Annals of Probability.* 44(3):1776–1816, 2016.

[100] Erol Peköz, Adrian Röllin, and Nathan Ross. Joint degree distributions of preferential attachment random graphs. *Advances in Applied Probability.* 49(2):368–387, 2017.

[101] Erol Peköz, Adrian Röllin, and Nathan Ross. Pólya urns with immigration at random times. *Official Journal of the Bernoulli Society for Mathematical Statistics and Probability.* 25(1):189–220, 2019.

[102] Robin Pemantle. A survey of random processes with reinforcement. *Probability Surveys.* 4(0):1–79, 2007.

[103] Robin Pemantle. A time-dependent version of Pólya's urn. *Journal of Theoretical Probability.* 3(4):627–637, 1990.

[104] Jim Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006, pages x+256. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002.

[105] Boris Pittel. Note on the heights of random recursive trees and random m-ary search trees. *Random Structures & Algorithms*. 5(2):337–347, 1994.

[106] Franz Rembart and Matthias Winkel. Recursive construction of continuum random trees. *The Annals of Probability*. 46(5):2715–2748, 2018.

[107] Jean-Luc Rémy. Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire. *RAIRO. Informatique théorique*. 19(2):179–195, 1985.

[108] Oliver Riordan. The phase transition in the configuration model. *Combinatorics, Probability and Computing*. 21(1-2):265–299, 2012.

[109] Nathan Ross and Yuting Wen. Scaling limits for some random trees constructed inhomogeneously. *Electronic Journal of Probability*. 23(0), 2018.

[110] Delphin Sénizergues. Geometry of preferential attachment and weighted recursive trees. 2019. arXiv: `1904.07115 [math.PR]`.

[111] Delphin Sénizergues. Growing random graphs with a preferential attachment structure. In preparation, 2019.

[112] Delphin Sénizergues. Random gluing of metric spaces. *The Annals of Probability*, 2019. To appear.

[113] Laurent Serlet. A large deviation principle for the Brownian snake. *Stochastic Processes and their Applications*. 67(1):101–115, 1997.

[114] Robin Stephenson. On the exponential functional of Markov additive processes, and applications to multi-type self-similar fragmentation processes and trees. *ALEA. Latin American Journal of Probability and Mathematical Statistics*. 15(2):1257–1292, 2018.

[115] Nicholas Wormald. *Some problems in the enumeration of labelled graphs*. PhD thesis, University of Newcastle, Australia, 1978.