



HAL
open science

Metric coalescence of homogeneous and inhomogeneous random graphs

Nicolas Frilet

► **To cite this version:**

Nicolas Frilet. Metric coalescence of homogeneous and inhomogeneous random graphs. Algebraic Topology [math.AT]. Université Grenoble Alpes [2020-..], 2021. English. NNT : 2021GRALM061 . tel-03667362

HAL Id: tel-03667362

<https://theses.hal.science/tel-03667362>

Submitted on 13 May 2022

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Mathématiques

Arrêté ministériel : 25 mai 2016

Présentée par

Nicolas FRILET

Thèse dirigée par **Vincent BEFFARA**
et co-encadrée par **Raphael ROSSIGNOL**

préparée au sein du **Laboratoire Institut Fourier**
dans l'**École Doctorale Mathématiques, Sciences et technologies de l'information, Informatique**

Coalescence métrique des graphes homogènes et inhomogènes

Metric coalescence of homogeneous and inhomogeneous random graphs

Thèse soutenue publiquement le **6 décembre 2021**,
devant le jury composé de :

Monsieur NICOLAS CURIEN

PROFESSEUR DES UNIVERSITES, UNIVERSITE PARIS-SACLAY,
Rapporteur

Monsieur JEAN-FRANÇOIS MARCKERT

DIRECTEUR DE RECHERCHE, CNRS DELEGATION AQUITAINE,
Rapporteur

Madame BENEDICTE HAAS

PROFESSEUR DES UNIVERSITES, UNIVERSITE SORBONNE PARIS
NORD, Examinatrice

Madame MARIE ALBENQUE

CHARGÉE DE RECHERCHE, CNRS DELEGATION ILE-DE-FRANCE
SUD, Examinatrice

Monsieur GREGORY MIERMONT

PROFESSEUR DES UNIVERSITES, ENS DE LYON, Président

Monsieur EMMANUEL JACOB

MAITRE DE CONFERENCES, ENS DE LYON, Examineur

Monsieur DIDIER PIAU

PROFESSEUR DES UNIVERSITES, UNIVERSITE GRENOBLE ALPES,
Examineur

Madame CHRISTINA GOLDSCHMIDT

PROFESSEUR, University of Oxford, Examinatrice



Metric coalescence of homogeneous and inhomogeneous random
graphs

Nicolas Fiolet

Remerciements

Je souhaite remercier Raphaël Rossignol et Vincent Beffara, qui ont encadré ma thèse. Ils m'ont tous les deux fournis de nombreuses perspectives et intuitions, et je leur dois la grande majorité de mes connaissances en probabilités. Merci en particulier d'avoir su encourager et soutenir mes expérimentations mathématiques, tout en me maintenant sur les fondations nécessaires de la rigueur.

Je souhaite également remercier Nicolas Curien et Jean-François Marckert d'avoir accepté de rapporter ma thèse. De même, je remercie Bénédicte Haas, Marie Albenque, Grégory Miermont, Emmanuel Jacob, Didier Piau et Christina Goldschmidt d'avoir accepté de participer à mon jury de thèse.

Naturellement, cette thèse a largement bénéficiée de la qualité excellente de la vie scientifique de l'Institut Fourier, et plus particulièrement de l'équipe de probabilité, que je remercie.

Finalement, je souhaite remercier ma famille, mes parents, mon frère, mes grands-parents, ma belle-famille ainsi que tous mes amis et mes proches qui ont dû subir l'explication laborieuse, répétée et modérément fructueuse de mon sujet de thèse. D'une manière générale, merci pour tous ces instants passés, à la maison, à l'extérieur, sur discord ou sur ow, à parler d'autres choses. Par-dessus tout, je remercie ma partenaire, Clotilde, pour son soutien constant et inconditionnel, et pour nos projets inspirants.

Enfin je souhaite dédier ce texte à mon grand-père qui a assisté à ses débuts enthousiastes (et incompréhensible) et qui n'est malheureusement plus là pour en apprécier le résultat (incompréhensible cependant, toujours.)

Table of contents

- I Introductions 9**
- 1 Introduction en français 13**
 - 1.1 Contexte: graphes aléatoires discrets au point critique 13
 - 1.2 Coalescent multiplicatif d’Aldous 15
 - 1.3 Espaces d’espaces métriques 19
 - 1.4 Encodage d’espaces métriques aléatoires 21
 - 1.5 Extension métrique de la coalescence 24
 - 1.6 Résultats 30
- 2 English introduction 33**
 - 2.1 Context: discrete random graphs at criticality 33
 - 2.2 Aldous’ multiplicative coalescent 34
 - 2.3 Spaces of metric spaces 39
 - 2.4 Encoding of random metric spaces 42
 - 2.5 Metric extension of the Coalescence 45
 - 2.6 Results 52
- II Tools 55**
- 3 Link between global statistics and component-wise statistics 57**
 - 3.1 Results 57
 - 3.2 Consequence and discussion 58
 - 3.3 Proof of Theorem 3.1.1 62
 - 3.4 Proof of Theorem 3.1.3 64
- 4 Truncation and Gromov-weak topology 67**
 - 4.1 Motivations 67
 - 4.2 Structural result 67
 - 4.3 Bounding the differences between truncated and non truncated coalescent 71
 - 4.4 Existence of a suitable truncation 72
- 5 From simple to complete coalescence 75**
 - 5.1 Bounding the difference between simple and complete coalescences 75
 - 5.2 Reformulation of universality theorems in terms of complete coalescence 77
- 6 Eternal multiplicative coalescent, mixing and entrance boundary 81**
 - 6.1 The lower mass bound is free for an eternal metric multiplicative coalescent 81
 - 6.2 The boundary conditions uniquely determine eternal multiplicative coalescents 83
- III Main results 85**
- 7 Palm measure of the point process associated with $\mathcal{E}(\alpha, \lambda, c, \gamma)$ 87**
 - 7.1 Result 87
 - 7.2 Proof 87

8	Mixing property for the coalescence of homogeneous multiplicative spaces	91
8.1	Results	91
8.2	Prerequisite	91
8.3	First moment of metric statistics	92
8.4	Second moment of general statistics	94
8.5	Proof of Theorem 8.2.2	95
9	Noise sensitivity for critical Erdős–Rényi random graphs	99
9.1	Asymptotic Independence	99
9.2	Some random graphs background	100
9.3	Some noise sensitive events	103
IV	Perspectives	107
10	Toward a characterization of all coalescence processes of compact measured metric spaces	109
10.1	What we have, what we want	109
10.2	Toward a generalized universality theorem	111
10.3	Completion of the discrete process	111
10.4	Coupling with an eternal spanning forest	115
A	Lower mass bound and randomness	119
B	Feller property for metric coalescence	121
	List of notation and terminology	123

Part I

Introductions

Les chapitres 1 et 2 sont tous deux autonomes et contiennent essentiellement les mêmes informations, on suggère donc au lecteur de n'en lire qu'un des deux. Le chapitre 2 contient quelques résultats techniques supplémentaires qui n'interviennent que dans les preuves, et il est tout à fait envisageable de ne s'y référer qu'au moment venu.

Chapter 1 et Chapter 2 are both self-contained and provide essentially the same information, so we suggest to the reader to read only one of them. Chapter 2 contains some additional technical results but it is perfectly conceivable to read them only when necessary.

Chapter 1

Introduction en français

1.1 Contexte: graphes aléatoires discrets au point critique

Bien que l'objet principal de cette thèse soit un processus d'évolution d'espaces métriques continus, son étude est motivée d'abord par celle des graphes aléatoires discrets. Commençons d'abord par préciser ce que l'on entend par là. Un graphe désigne ici la donnée d'un ensemble de sommets, usuellement $V = \llbracket 0, n \rrbracket = [n]$, et la donnée d'un ensemble d'arêtes non orientées $E \subset \{\{i, j\} \in V^2\}$. De plus, la donnée d'un graphe (V, E) s'accompagne de celle d'une distance sur V , celle du plus court chemin.

Sans rentrer outre mesure dans des questions de modélisations, généralement l'objectif d'un modèle est de capturer la notion de graphe typique exhibant tel ou tel comportement : est-ce que le graphe s'injecte isométriquement dans \mathbb{R}^d ? Est-ce que le degré des sommets est contraint à une valeur précise ? Quelle est la distance moyenne entre deux sommets ?

On peut imaginer ainsi définir une infinité de modèles, correspondant chacun à un comportement particulier, et on peut ainsi voir l'intérêt de faire apparaître les similarités entre ces modèles. L'étude de ces similarités est l'objectif sous-jacent de cette thèse, dans le cas d'une classe particulière de modèles qui informellement satisfont les critères suivants :

- Ce sont des grands graphes, c'est à dire que l'on est intéressé par les asymptotiques en $n \rightarrow +\infty$.
- Le degré moyen des sommets est fini.
- Il n'y a pas de contrainte géométrique, c'est à dire de contrainte à grande échelle.

Introduisons un modèle qui répond effectivement à ces critères et qui est historiquement l'un des premiers modèles historiques de graphes aléatoires. Pour $n \in \mathbb{N}$, $p \in [0, 1]$, on considère le graphe $G^n = (V, E)$ sur $V = \llbracket 0, n \rrbracket = [n]$ avec

$$\forall i, j \in [n], i \neq j, \mathbb{P}(\{i, j\} \in E) = p$$

indépendamment pour toutes les paires $\{i, j\}$. On appelle ces graphes des graphes aléatoires d'Erdős–Rényi, et on note leur loi $G(n, p)$. Pour que la contrainte de degré moyen des sommets soit respectée, on s'autorise à faire varier $p = (p_n)_{n \in \mathbb{N}}$ avec n , et on prendra p de sorte que $p_n n \xrightarrow[n \rightarrow \infty]{} c > 0$.

1.1.1 La fenêtre critique

Un résultat emblématique du travail de Erdős et Rényi [19] sur $G(n, p)$ est le phénomène de seuil pour les tailles des grandes composantes connexes. Dans ce contexte, une composante connexe est un ensemble de sommets $V' \subset V$ tel que pour chaque couple de sommets $i, j \in V'$, il existe un chemin d'arêtes qui lie i à j . De plus, la taille d'une composante est simplement le cardinal de celle-ci. On dit qu'une composante V_1 est plus grande qu'une autre composante V_2 si $|V_1| \geq |V_2|$, et lorsque $V_1 = V_2$, si le plus petit $i \in V_1$ est plus petit que le plus petit $j \in V_2$. Enfin, par commodité, on considère qu'un graphe possède toujours une infinité de composantes vides, de telle sorte que l'on peut associer à G^n la suite $(C_i^n)_{i \in \mathbb{N}}$ de ces composantes en ordre décroissant. En particulier, si pour $i \in \mathbb{N}$, $C_i^n = \emptyset$, cela signifie que le graphe contient au plus i composantes connexes non vides.

La loi $G(n, p)$ donne lieu à la transition de phase suivante [19]:

$$\text{Si } c < 1, \exists \delta > 0 \quad \mathbb{P}(|C_0^n| \leq \delta \log(n)) \xrightarrow[n \rightarrow +\infty]{} 1, \quad (1.1)$$

$$\text{Si } c > 1, \exists \delta > 0, \forall k \geq 1, \mathbb{P}(|C_0^n| \geq \delta n, |C_k^n| \leq \delta \log(n)) \xrightarrow[n \rightarrow +\infty]{} 1. \quad (1.2)$$

Au point critique, $c = 1$, on a de plus le phénomène suivant [13] : pour $\lambda \in \mathbb{R}$,

$$\text{si } c = 1 + \lambda n^{-1/3}, \forall a, b > 0, \text{ tel que } a < b, \forall k \in \mathbb{N}, \lim_{n \rightarrow +\infty} \mathbb{P}(an^{2/3} \leq |C_k^n| \leq bn^{2/3}) \in (0, 1) \quad (1.3)$$

1.1.2 Approche dynamique et fusion multiplicative

La dépendance en λ de (1.3) sous-entend l'existence non pas d'un point mais d'une fenêtre critique, paramétrée par λ , durant laquelle naît la composante géante, celle qui est composée de la plupart des sommets lorsque $c > 1$. La formulation de cette remarque suggère d'adopter un point de vue dynamique sur ces graphes : pour chaque $i, j \in [n], i < j$, on échantillonne $E_{i,j} \sim \mathcal{E}(1)$ indépendamment de tout le reste. D'autre part, pour $t \geq -n^{1/3}$, on définit $G^n(t) = (V, E(t))$ avec

$$E(t) = \left\{ \{i, j\} : i, j \in [n], i < j, E_{i,j} \leq \frac{1 + tn^{-1/3}}{n} \right\}.$$

En particulier, la probabilité d'existence d'une arête $\{i, j\}$ à un temps t est, lorsque n tend vers l'infini,

$$\mathbb{P}(\{i, j\} \in E(t)) = 1 - e^{-\frac{1+tn^{-1/3}}{n}} = \frac{1 + tn^{-1/3}}{n} + o(n^{-4/3})$$

ainsi $G^n(t) \sim G(n, \frac{1+tn^{-1/3}}{n} + o(n^{-4/3}))$. De plus, pour $V_1, V_2 \in V$, en notant $E(t, V_1, V_2)$ l'ensemble des arêtes entre V_1 et V_2 qui existent au temps t , du fait de l'indépendance entre les arêtes la probabilité de fusion des ensembles disjoints V_1 et V_2 durant la période $[s, t]$ est

$$\mathbb{P}(E(t, V_1, V_2) \neq \emptyset | E(s, V_1, V_2) = \emptyset) = 1 - e^{-\frac{t-s}{n^{4/3}} |V_1| |V_2|}.$$

Ainsi, deux ensembles disjoints fusionnent à un taux $|V_1|n^{-2/3}|V_2|n^{-2/3}$ proportionnel au produit des tailles des ensembles V_1 et V_2 . Ce comportement, que l'on qualifiera de *multiplicatif*, est au cœur de cette thèse. On dit qu'un graphe G sur $[n]$ est *multiplicatif* s'il existe une suite de poids $(x_i)_{i \in [n]} \in \mathbb{R}_+^{n+1}$ telle que G est le résultat d'un processus où deux sous-ensembles disjoints V_1 et V_2 de $[n]$ fusionnent à taux $\sum_{i \in V_1} x_i \sum_{i \in V_2} x_i$, indépendamment du reste du graphe. Ces graphes sont aussi appelés graphes aléatoires inhomogènes et sont définis dans la Définition 1.2.1. Dans [14], les auteurs ont mis en évidence une famille de distributions vers laquelle les graphes aléatoires inhomogènes convergent, généralisant ainsi des résultats partiels de [1] et [11]. De plus, dans [9] et [10], les auteurs ont montré que sous certaines conditions, les modèles de configuration surcritiques percolés convergent vers certaines des distributions décrites dans [14]. Cette convergence vient entre autre du fait que la percolation sur les modèle de configurations surcritiques peut être couplée avec un processus de graphe multiplicatif de manière asymptotiquement transparente [17]. Ces modèles, les modèles de configuration surcritiques percolés, sont les principaux exemples de modèles *presque* multiplicatif, au sens qu'ils sont asymptotiquement multiplicatif. Comme le fait d'être multiplicatif est essentiellement une propriété d'indépendance, on peut s'attendre à ce que d'autres modèles possèdent également cette propriété d'être presque multiplicatif. De plus le fait que tous ces graphes soient issues de processus, on peut également s'attendre à ce que ces processus eux-même convergent en tant que processus stochastiques.

Dans ce cadre, une question naturelle se pose :

Question directrice 1. *La famille d'espaces métriques aléatoires décrite dans [14] et [15] contient-elle toutes les lois marginales possibles des limites du processus de graphe presque multiplicatif ?*

Avant de finir cette section, nous pouvons développer à propos de ce qui serait intuitivement un bon candidat pour la Question directrice 1 : bien sûr, il doit y avoir une notion de comportement multiplicatif dans sa définition. Il sera d'ailleurs appelé coalescent multiplicatif. De plus, comme la fenêtre critique décrite dans (1.3) est de taille infinie ($\lambda \in (-\infty, +\infty)$), le processus doit être éternel à la fois dans le passé et dans le futur. Comme nous le verrons, sonder la Question directrice 1 nous ramène à trouver tous les coalescents multiplicatifs *éternels*.

Il apparaît en fait plusieurs sortes de coalescents multiplicatifs dans ce texte, et on utilisera l'expression "coalescent multiplicatif" lorsque le contexte est clair, et on précisera exactement de quel coalescent multiplicatif on parle lorsque c'est nécessaire.

1.2 Coalescent multiplicatif d'Aldous

1.2.1 Remarques préliminaires et définitions

Pour commencer à étudier la Question directrice 1, nous présentons d'abord le coalescent multiplicatif d'Aldous [5]. Ici, on oublie la structure de graphe et on se concentre uniquement sur la taille des composantes. On souhaite à terme prendre en compte le contenu des composantes mais dans [6], Aldous et Limic parviennent à une réponse à une forme affaiblie de Question directrice 1, et leur article constituera notre point de départ méthodologique. Nous sommes à la recherche d'un loi limite, sous forme de processus stochastique, pour les tailles de composantes à l'intérieur de la fenêtre critique. En particulier les conditions initiales d'un tel processus sont sous la forme de poids $x = (x_i)_{i \in \mathbb{N}}$. Par souci de commodité dans la manipulation des suites de poids, on suppose toujours que $\forall i < j, x_i \geq x_j$.

Définition 1.2.1 (Coalescent multiplicatif d'Aldous). *Soit $x = (x_i)_{i \in \mathbb{N}}$ une suite de réels positifs. On définit un processus qui prend ses valeurs dans les suites décroissantes de $\overline{\mathbb{R}}^+$ de la manière suivante: soit $gCoal(x, q)$ le graphe d'ensemble de sommets \mathbb{N} tel que pour $i, j \in \mathbb{N}$ tel que $i < j$,*

$$\{i, j\} \text{ est une arête de } gCoal(x, q) \text{ si et seulement si } E_{i,j} \leq q$$

avec $(E_{i,j})$ une suite de variables aléatoires exponentielles indépendantes telles que :

$$\forall i, j \in \mathbb{N}, \mathbb{E}(E_{i,j}) = \frac{1}{x_i x_j}.$$

On appelle ce modèle un graphe aléatoire inhomogène de rang 1, ou modèle de Norros-Reittu. On peut alors définir le coalescent multiplicatif d'Aldous $mCoal(x, q)$ partant de x au temps q comme la suite ordonnée des tailles de composantes de $gCoal(x, q)$. De plus, on note $E(x, q)$ l'ensemble des arêtes de $gCoal(x, q)$.

Cette construction, appelée construction "graphique" est due à Aldous dans [5]. Remarquons que dans la Définition 1.2.1 on n'impose pas de contrainte particulière sur la suite de poids initiale. Cependant, rien n'empêche a priori que le poids total de certaines composantes ne soit pas borné. Retournons à notre Question directrice 1 et déterminons quels suites de poids initiales prendre. Nous sommes à la recherche de l'objet limite et du fait de (1.3), il doit posséder une définition valide pour n'importe quel $\lambda \in \mathbb{R}$, à tout temps plusieurs grandes composantes doivent coexister et il ne doit jamais y avoir de composante de taille infinie. On pourrait à première vue être tenté de se restreindre aux x tels que $\sum_i x_i < +\infty$ ce qui garantirait que la plus grande composante soit toujours de taille finie. Cependant, on remarque qu'avec probabilité $(1 - e^{-q x_1 \sum_{i \geq 2} x_i}) > 0$, toutes les composantes de $gCoal(x, q)$ ont fusionné au temps q . Ainsi, on doit forcément avoir $\sigma_1(x) = \sum_i x_i = +\infty$ et trouver un autre moyen de garantir que les tailles des composantes restent finies.

1.2.2 Analyse du coalescent multiplicatif d'Aldous

Dans le but, entre autre, de contrôler la taille des composantes, Aldous introduit les espaces suivants:

Définition 1.2.2 (Espace des suites de poids). *Pour tout $r > 0$, x une suite décroissante de \mathbb{R}^+ , soit*

$$\sigma_r(x) = \sum_i x_i^r = \|x\|_r^r,$$

on définit

$$l_{\searrow}^r = \{(x_i)_{i \in \mathbb{N}} : 0 \leq x_{i+1} \leq x_i, \sigma_r(x) < +\infty\}.$$

En particulier, si $r = 2$ on équipe cet espace de la distance

$$\forall x, y \in l_{\searrow}^2, d_2(x, y) = \sqrt{\sum_i (x_i - y_i)^2}.$$

Dans la suite, s'il n'y a pas d'ambiguïté sur x , nous écrivons σ_r pour $\sigma_r(x)$.

Voyons quel rôle jouent ces espaces pour le coalescent multiplicatif, pour ce faire introduisons :

Définition 1.2.3. Soit $x \in l_{\searrow}^2$, si $x(q) = mCoal(x, q)$, on note

$$S(x, q) = \sigma_2(x(q)).$$

On peut alors écrire le résultat peut être le plus fondamental à propos du coalescent multiplicatif d'Aldous:

Lemme 1.2.4 (Lemme 20 de [5]). Pour $x \in l_{\searrow}^2$, $s > S(x, 0) = \sigma_2$,

$$\mathbb{P}(S(x, q) > s) \leq \frac{qs\sigma_2}{s - \sigma_2}$$

La théorie du coalescent multiplicatif nous apprend que l_{\searrow}^2 est la bonne contrainte sur x , en particulier à travers le théorème suivant:

Théorème 1.2.1 (Propriété de Feller, Proposition 5 de [5]). Soit $(x^n)_{n \in \mathbb{N}}$ une suite de l_{\searrow}^2 et $x^\infty \in l_{\searrow}^2$ tel que

$$x^n \xrightarrow[n \rightarrow +\infty]{} x^\infty$$

dans (l_{\searrow}^2, d^2) . Alors

$$\forall q \in \mathbb{R}^+, mCoal(x^n, q) \xrightarrow[n \rightarrow +\infty]{(d)} mCoal(x^\infty, q)$$

dans (l_{\searrow}^2, d^2) .

À la lumière du Théorème 1.2.1, on notera sl_2 l'espace des suites de l_{\searrow}^2 . De plus, si $x = (x^n)_{n \in \mathbb{N}} \in l_{\searrow}^2$, on note σ_r la suite $(\sigma_r(x^n))_{n \in \mathbb{N}}$.

1.2.3 Problèmes d'Aldous-Limic et conditions d'Aldous-Limic

Rappelons encore une fois que les coalescents multiplicatifs qui conviendraient comme limites de graphes multiplicatifs doivent être tels qu'ils soient bien définis aussi loin que l'on veut dans le passé. Plus précisément, il faut qu'ils satisfassent la définition suivante :

Définition 1.2.5 (Coalescent multiplicatif éternel). On dit qu'un processus de Markov $(x(t))_{t \in \mathbb{R}}$ dans l_{\searrow}^2 est un coalescent multiplicatif éternel si

$$\forall s, t \in \mathbb{R} \text{ t.q. } s \leq t, (x(t), x(s)) \stackrel{(d)}{=} (mCoal(x(s), t - s), x(s)).$$

Par la suite on appellera le problème consistant à trouver un coalescent multiplicatif éternel un problème d'Aldous-Limic. Avant d'étudier plus précisément ce problème, en particulier son caractère bien posé, on introduit quelques définitions :

Définition 1.2.6.

$$\mathcal{I} = \{(\alpha, \lambda, c) \in \mathbb{R}_+ \times \mathbb{R} \times l_{\searrow}^3 : \alpha > 0 \text{ ou } c \notin l_{\searrow}^2\}.$$

$$\mathcal{I}^\circ = \{(\alpha, \lambda, 0) \in \mathcal{I} : \alpha > 0\}$$

$$\mathcal{I}^* = \{(0, \lambda, c) \in \mathcal{I} : c \in l_{\searrow}^3 \setminus l_{\searrow}^2\}$$

De plus, pour $s \in \mathbb{R}$ et $(\alpha, \lambda, c) \in \mathcal{I}$, on définit

$$\mathcal{Z}^{(\alpha, \lambda, c)}(s) = \alpha^{1/2} \mathcal{W}(s) + (\lambda + t)s - \frac{1}{2} \alpha s^2 + \sum_i (c_i 1_{\zeta_i \leq s} - c_i^2 s),$$

avec \mathcal{W} un mouvement Brownien et $(\zeta_i)_{i \in \mathbb{N}}$ une suite de variables aléatoires exponentielles indépendantes telles que pour $i \in \mathbb{N}$, $\mathbb{E}(\zeta_i) = 1/c_i$. Quand il n'y a pas d'ambiguïté sur les paramètres, on écrira simplement \mathcal{Z} pour $\mathcal{Z}^{(\alpha, \lambda, c)}$. Remarquons que, en toute rigueur, \mathcal{Z} est définie comme un processus de Lévy spectralement positif caractérisé par une certaine transformée de Laplace dans [14]. Les auteurs de [14] montrent (équation (143) de la section 4.2), que la notation que l'on a introduit ici fait sens : la somme infinie doit être interprétée comme la limite de semi-martingales L^2 . Enfin, on note

$$\mathcal{B}^{(\alpha, \lambda, c)}(s) = \mathcal{Z}^{(\alpha, \lambda, c)}(s) - \inf_{[0, s]} \mathcal{Z}^{(\alpha, \lambda, c)}.$$

Revenons au problème d'Aldous-Limic. La première question qui se pose est celle de l'existence de solutions. On peut trouver la réponse dans [5] où l'on a :

Théorème 1.2.2 (Convergence des tailles des grandes composantes de graphes aléatoires d'Erdős-Rényi, Proposition 4 de [5]). *Pour $(1, \lambda, 0) \in \mathcal{I}^\circ$, on peut définir $(|e_j|)_{j \in \mathbb{N}}$ la suite ordonnée des longueurs d'excursion de $\mathcal{B}^{(1, \lambda, 0)}$ au dessus de 0. Pour $n \in \mathbb{N}$, soit $(|C_j^n|)_{j \in \mathbb{N}}$ la suite ordonnée des tailles de composantes de $G^n \sim G(n, \frac{1+\lambda n^{-1/3}}{n})$. On a*

$$(n^{-2/3}|C_j^n|)_{j \in \mathbb{N}} \xrightarrow[n \rightarrow +\infty]{(d)} (|e_j|)_{j \in \mathbb{N}}$$

dans (l_{\searrow}^2, d^2) .

Le Théorème 1.2.2, combiné au Théorème 1.2.1 donne, comme dit dans le Corollaire 24 de [5], une réponse positive à la question de l'existence de solutions au problème d'Aldous-Limic. En effet soit $(t_0, \dots, t_k) \in \mathbb{R}^{k+1}$ une famille de réels croissante et $x^n \in sl_2$, et pour $t \in \mathbb{R}$, $x^n(t) = \text{mCoal}(x^n, \sigma_2^{-1} + t)$. On considère la suite de vecteurs $(x^n(t_i))_{i \in [k], n \in \mathbb{N}}$. Pour f mesurable bornée, $x \in l_{\searrow}^2$ et $t \in \mathbb{R}$, soit $P_t f(x) = \mathbb{E}(f(\text{mCoal}(x, t)))$. Soit (f_1, \dots, f_k) une famille de fonctions mesurables bornées. Comme le processus est de Markov, on a

$$\mathbb{E}\left(\prod_{i \in [k]} f_i(x^n(t_i))\right) = \mathbb{E}(f_0(x^n(t_0))P_{t_1-t_0}[f_1 P_{t_2-t_1}[\dots]](x^n(t_0))).$$

Or, grâce au Théorème 1.2.2,

$$x^n(t_0) \xrightarrow[n \rightarrow \infty]{(d)} x^\infty(t_0)$$

et grace au Théorème 1.2.1, $f_0(\cdot)P_{t_1-t_0}[f_1 P_{t_2-t_1}[\dots]](\cdot)$ est continue bornée. Ainsi,

$$\mathbb{E}(f_0(x^n(t_0))P_{t_1-t_0}[f_1 P_{t_2-t_1}[\dots]](x^n(t_0))) \xrightarrow[n \rightarrow \infty]{} \mathbb{E}(f_0(x^\infty(t_0))P_{t_1-t_0}[f_1 P_{t_2-t_1}[\dots]](x^\infty(t_0)))$$

et alors

$$\mathbb{E}\left(\prod_{i \in [k]} f_i(x^n(t_i))\right) \xrightarrow[n \rightarrow \infty]{} \mathbb{E}\left(\prod_{i \in [k]} f_i(x^\infty(t_i))\right).$$

Cette convergence nous donne une famille de distributions telle que pour tout (t_0, \dots, t_k) croissant, $(x^\infty(t_0), \dots, x^\infty(t_k))$ est compatible, Markovien et de plus

$$(\text{mCoal}(x^\infty(t_0), t_1 - t_0), x^\infty(t_0)) \stackrel{(d)}{=} (x^\infty(t_1), x^\infty(t_0)).$$

Le théorème d'extension de Kolmogorov nous donne alors un processus solution du problème d'Aldous-Limic. Une construction alternative basée sur l'utilisation d'un même mouvement Brownien pour tout $t \in \mathbb{R}$ peut être trouvée dans [16]. Le coalescent ainsi obtenu est appelé le coalescent multiplicatif standard.

Naturellement, la question qui se pose alors est celle de l'unicité d'un tel processus. Cependant, tel quel, le problème est mal posé: d'une part la solution triviale toujours égale à $x = (1, 0, \dots) \in l_{\searrow}^2$ existe, et d'autre part si l'on a un coalescent multiplicatif éternel $(x(t))_{t \in \mathbb{R}}$, on peut changer l'origine des temps ou le multiplier par une constante et il restera un coalescent multiplicatif éternel. De plus, si un processus est un mélange de coalescents multiplicatifs éternels il est lui-même un coalescent multiplicatif éternel. Cela rend convexe l'ensemble des coalescents multiplicatifs éternels. Plus précisément, cet ensemble est un simplexe [18], c'est à dire un ensemble convexe dont tout élément s'écrit de manière unique comme combinaison convexe d'éléments extrémaux. De plus, les distributions extrêmes d'un processus de Markov éternel sont exactement celles pour lesquelles la tribu asymptotique en $-\infty$ est triviale [18]. On peut donc affiner la définition d'un problème d'Aldous-Limic en ajoutant que la distribution doit être extrême. De manière surprenante, on ne couvre pas tous les coalescents multiplicatifs éternels en considérant l'enveloppe convexe des solutions considérées au début de ce paragraphe (solutions triviales, coalescent multiplicatif standard, ses décalages en temps etc.). Dans [6], Aldous et Limic caractérisent tous les coalescents multiplicatifs éternels extrêmes, on présente leur travail ici dans le cadre spécifique de la question directrice 1. On définit d'abord, pour $x \in l_{\searrow}^2$,

$$\Gamma(x, q) = \left(\sigma_2, \left(\frac{x_j}{\sigma_2}\right)_{j \in \mathbb{N}}, \frac{\sigma_3}{\sigma_2^3}, q - \frac{1}{\sigma_2}\right).$$

Ceci-nous permet de définir proprement ce qui s'avérera par la suite être un problème d'Aldous-Limic bien posé :

Définition 1.2.7 (Problème d'Aldous-Limic avec paramètres $(\alpha, \lambda, c) \in \mathcal{I}$). *On dit qu'un processus de Markov $x = (x(t))_{t \in \mathbb{R}}$ dans l_{\searrow}^2 est une solution du problème d'Aldous-Limic avec paramètre $(\alpha, \lambda, c) \in \mathcal{I}$ si sa distribution est extrême et*

$$\left\{ \begin{array}{l} \forall s, t \in \mathbb{R}, s \leq t, (x(s), x(t)) \stackrel{(d)}{=} (x(s), m\text{Coal}(x(s), t-s)) \quad (\text{c'est un coalescent multiplicatif éternel}) \\ \lim_{t \rightarrow -\infty} \Gamma(x(t), t) = (0, c, \alpha + \sum_j c_j^3, \lambda) \text{ a.s.} \quad (\text{condition de frontière d'entrée}) \end{array} \right.$$

On peut maintenant énoncer le principal théorème à propos des coalescent multiplicatifs éternels d'Aldous :

Théorème 1.2.3 (Théorèmes 2 et 3 de [6]). *Pour $(\alpha, \lambda, c) \in \mathcal{I}$, le problème d'Aldous-Limic de la Définition 1.2.7 est bien posé. De plus, pour tout coalescent multiplicatif éternel extrême $(x(t))_{t \in \mathbb{R}}$, soit x est constant, soit il existe $(\alpha, \lambda, c) \in \mathcal{I}$, tel que pour tout $t \in \mathbb{R}$, si $(|e_j|)_{j \in \mathbb{N}}$ est la suite ordonnée des longueurs d'excursion de $\mathcal{B}^{(\alpha, \lambda+t, c)}$ au dessus de 0, alors,*

$$x(t) \stackrel{(d)}{=} (|e_j|)_{j \in \mathbb{N}}.$$

On notera la distribution d'un tel processus $\mathcal{M}(\alpha, \lambda, c)$ et $\mathcal{M}_t(\alpha, \lambda, c)$ désignera la loi marginale au temps t .

On remarque que, contrairement aux paramètres α et λ , qui décrivent la mise à l'échelle et le décalage en temps, le nouveau paramètre c ne possède pas d'interprétation immédiate. Il représente deux phénomènes liés:

- (i) Si $c \neq 0$, cela signifie que lorsque $t \rightarrow -\infty$, il reste des composantes dont la taille est comparable à σ_2 , ce qui n'est pas le cas des graphes d'Erdős-Rényisous-critiques. C'est la toute première fois que l'on observe que la classe des objets multiplicatifs est plus large que celle des variantes des graphes d'Erdős-Rényi.
- (ii) Si $c \neq 0$, le processus \mathcal{B} possède des saut. Comme pour l'instant on ignore la structure interne des composantes, on ne peut pas interpréter rigoureusement ces sauts, mais intuitivement ils révèlent une inhomogénéité dans la structure métrique des composantes.

Dans l'esprit du Théorème 1.2.3, on définit:

Hypothèse 1 (Conditions d'Aldous-limic pour des suites de suites de poids). *On dit que le couple (x, q) satisfait les conditions d'Aldous-Limic avec paramètre $(\alpha, \lambda, c) \in \mathcal{I}$ si $x = (x^n)_{n \in \mathbb{N}}$ est une suite de l_{\searrow}^2 , $q = (q_n)_{n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$ et*

$$\left\{ \begin{array}{ll} \sigma_2(x^n) \xrightarrow{n \rightarrow +\infty} 0 & (\text{sous-criticalité initiale}) \\ \frac{x_j^n}{\sigma_2(x^n)} \xrightarrow{n \rightarrow +\infty} c_j & (\text{distribution des sauts}) \\ \frac{\sigma_3(x^n)}{\sigma_2(x^n)^3} \xrightarrow{n \rightarrow +\infty} \alpha + \sum_i c_i^3 & (\text{diffusion}) \\ q_n - \frac{1}{\sigma_2(x^n)} \xrightarrow{n \rightarrow +\infty} \lambda. & (\text{criticalité}) \end{array} \right.$$

En d'autres termes,

$$\lim_{n \rightarrow +\infty} \Gamma(x^n, q_n) = (0, c, \alpha + \sum_i c_i^3, \lambda).$$

Comme notre objectif est d'obtenir un théorème similaire mais en tenant compte de la structure interne des composantes, disons quelques mots de la méthode qu'employèrent Aldous et Limic pour obtenir Théorème 1.2.3. L'idée repose sur l'usage du théorème suivant:

Théorème 1.2.4 (Proposition 18 de [6]). *Soit $(\alpha, \lambda, c) \in \mathcal{I}$, $x = (x^n)_{n \in \mathbb{N}} \in sl_2$ et $q = (q_n)_{n \in \mathbb{N}}$ tels que x et q satisfont les conditions d'Aldous-Limic (1) avec paramètres (α, λ, c) et pour tout $i > n$, $x_i^n = 0$. Alors*

$$m\text{Coal}(x^n, q) \xrightarrow[n \rightarrow +\infty]{d} (|e_j|)_{j \in \mathbb{N}}$$

où $(|e_j|)_{j \in \mathbb{N}}$ est la suite ordonnée des longueurs d'excursions de $\mathcal{B}^{(\alpha, \lambda, c)}$.

Ce théorème donne non seulement l'existence de solution, mais aussi, après une quantité notable de travail, l'unicité: si, étant donnée un coalescent multiplicatif éternel, on parvient à montrer qu'il satisfait les conditions d'Aldous-Limic 1 avec paramètre $(\alpha, \lambda, c) \in \mathcal{I}$, on peut espérer utiliser le Théorème 1.2.4 pour caractériser la loi de ce coalescent en tant que $\mathcal{M}(\alpha, \lambda, c)$. C'est pourquoi, tout au long de cette thèse, la quête d'un théorème analogue au Théorème 1.2.4 prenant en compte la structure interne des composantes sera l'une de nos principales préoccupations.

1.2.4 Mesure de Palm des coalescents multiplicatifs

Dans [22], Janson et Spencer adopte le point de vu des processus ponctuels pour décrire $\mathcal{M}_0(1, \lambda, 0)$. En particulier, ils en calculent la mesure de Palm:

Théorème 1.2.5 (Théorème 8.1 de [22]). *Soit f continue bornée à support compact de $(0, +\infty)$ vers \mathbb{R}_+ , et g continue bornée de \mathbb{R}_+^2 vers \mathbb{R}_+ . Soit $\lambda \in \mathbb{R}$ et $x \sim \mathcal{M}_0(1, \lambda, 0)$ On a:*

$$\mathbb{E}\left(\sum_i f(x_i)g(x)\right) = \mathbb{E}\left(\sum_i f(x_i)h(x_i)\right)$$

où la fonction h est définie de la manière suivante : pour $u \in \mathbb{R}_+$,

$$h(u) = \mathbb{E}(g(y^u))$$

où y^u est une variable aléatoire définie comme la suite ordonnée des poids de $y \sim \mathcal{M}_0(1, \lambda - u, 0)$ à laquelle on a ajouté u .

Ce théorème est particulièrement utile pour le calcul des moments d'ordre 2 puisqu'il ramène le calcul à celui d'un moment d'ordre 1. On peut donc espérer s'en servir pour contrôler des variances et donc des probabilités quand λ tend vers $-\infty$.

1.3 Espaces d'espaces métriques

1.3.1 Convergence d'espaces métriques

Cette section est essentiellement technique et pose le contexte dans lequel on peut étendre l'analyse d'Aldous et Limic du coalescent multiplicatif afin de prendre en compte la structure interne des composantes. Plus précisément, nous nous intéresserons aux structures métriques des composantes des graphes multiplicatifs.

Definition 1.3.1 (Espace des espaces métriques mesurés compacts). *On dit qu'un triplet $m = (\mathcal{M}, d, \mu)$ est un espace métrique mesuré si d est une distance sur \mathcal{M} et μ est une mesure de Borel sur \mathcal{M} . De plus, nous écrirons $|m| = \mu(\mathcal{M})$. On note \mathbb{M} l'espace de tous les espaces métriques mesurés compacts et $\tilde{\mathbb{M}}$ l'espace de tous les espaces métriques complets, séparables, à mesure finie, de support plein. De plus, pour $m = (\mathcal{M}, d, \mu)$ un espace métrique mesuré et $\alpha \in \mathbb{R}_+$, on note αm l'espace $(\mathcal{M}, \alpha d, \mu)$ et pour $\beta \in \mathbb{R}_+$, on note $scl(\alpha, \beta)m$ l'espace $(\mathcal{M}, \alpha d, \beta \mu)$. On équipe \mathbb{M} et $\tilde{\mathbb{M}}$ de la distance d_{GHP} qui est définie pour m et $m' \in \tilde{\mathbb{M}}$, comme suit: on dit que $\mathcal{R} \subset m \times m'$ est une correspondance entre (m, s, μ) et (m', d', μ') si*

$$\forall u \in m, \exists u' \in m' \text{ tel que } (u, u') \in \mathcal{R}$$

$$\forall u' \in m', \exists u \in m \text{ tel que } (u, u') \in \mathcal{R}.$$

On note $\mathcal{C}(m, m')$ pour l'ensemble de toutes les correspondances mesurables entre m et m' . On peut définir la distorsion d'une correspondance \mathcal{R} comme suit

$$dis(\mathcal{R}) = \inf\{\epsilon > 0 : \forall (u, u'), (v, v') \in \mathcal{R}, |d(u, v) - d'(u', v')| \leq \epsilon\}.$$

La distorsion offre une bonne comparaison des deux métriques mais nous aimerions aussi comparer les deux mesures. Soit $\Pi(m, m')$ l'ensemble des mesures boréliennes finies sur $m \times m'$. Pour $\pi \in \Pi(m, m')$, on note (π_1, π_2) la première et la deuxième marginales de π . Soit

$$D(\pi; \mu, \mu') = \|\pi_1 - \mu\| + \|\pi_2 - \mu'\|$$

avec $\|\cdot\|$ la distance en variation totale. On définit alors :

$$d_{GHP}(m, m') = \inf_{\mathcal{R} \in \mathcal{C}(m, m'), \pi \in \Pi(m, m')} \left\{ D(\pi; \mu, \mu') \wedge \frac{1}{2} dis(\mathcal{R}) \wedge \pi(\mathcal{R}^c) \right\}.$$

Bien sûr, telle quelle d_{GHP} n'est pas une distance sur \mathbb{M} puisqu'elle ne respecte pas l'axiome de séparation. Pour remédier à cela, on considère \mathbb{M}' l'ensemble des classes d'équivalences dans \mathbb{M} pour la relation d'équivalence

$$m \sim m' \Leftrightarrow d_{GHP}(m, m') = 0,$$

et on munit \mathbb{M}' de la distance engendrée par d_{GHP} . Dans le reste de ce travail, on abusera des notations en utilisant \mathbb{M} à la place de \mathbb{M}' .

1.3.2 Une topologie sur les suites d'espaces métriques issue de la topologie Gromov-Hausdorff-Prokhorov

Nous voulons traiter des espaces constitués d'une infinité de composantes connexes compactes, d'où la définition suivante:

Définition 1.3.2 (Espace des suites d'espaces métriques mesurés compacts). *On note $\mathbb{M}_2^{\mathbb{N}}$ l'espace des suites $X = (X_j)_{j \in \mathbb{N}}$ de \mathbb{M} telles que $\sum_j |X_j|^2 < +\infty$ et telles que*

$$\forall i, j \in \mathbb{N}, i < j, |X_i| \geq |X_j|.$$

On remarque qu'on considérera X comme un espace métrique mesuré en lui-même, dont les différents X_j sont les composantes. On note $\text{comp}(X)$ l'ensemble des composantes X_j pour $j \in \mathbb{N}$.

On remarque que l_{\searrow}^2 s'injecte dans $\mathbb{M}_2^{\mathbb{N}}$ en prenant la distance triviale sur chaque $\{i\}$ et la mesure de Dirac sur $\{i\}$ avec le poids x_i . Cela nous permet d'appliquer canoniquement à un élément de l_{\searrow}^2 une fonction qui prend un élément de $\mathbb{M}_2^{\mathbb{N}}$ comme argument.

Cependant, contrairement au cas du processus $\text{mCoal}(\cdot, \cdot)$ sur l_{\searrow}^2 il n'y a pas de topologie claire sur $\mathbb{M}_2^{\mathbb{N}}$ qui conviendrait à tous nos objectifs. Il existe principalement deux types de topologies provenant de d_{GHP} :

Définition 1.3.3. *Si $X, X' \in \mathbb{M}_2^{\mathbb{N}}$, on définit*

$$\delta_{GHP}^2(X, X') = \left(\sum_j (d_{GHP}(X_j, X'_j) \wedge 1) 2^{-j} \right) \vee \| |X| - |X'| \|_2.$$

La topologie associée sur $\mathbb{M}_2^{\mathbb{N}}$ sera notée T_{GHP}^2 .

Cependant, dans notre contexte, nous souhaiterons souvent travailler avec la topologie la plus faible possible. La première topologie qui vient à l'esprit qui satisfait cela tout en restant dans l'esprit de d_{GHP} est la topologie produit. La topologie produit sur $\mathbb{M}_2^{\mathbb{N}}$ héritée de d_{GHP} sera notée T_{GHP} .

1.3.3 Affaiblir d_{GHP} : la topologie Gromov-faible

Récemment, il y a eu une augmentation de l'utilisation de la topologie Gromov-faible pour étudier la convergence de graphes aléatoires. Cette topologie est définie comme suit:

Définition 1.3.4 (Topologie Gromov-faible). *Soit (X, d, μ) un espace métrique mesuré tel que $|X| < +\infty$. Pour $u = (u_1, \dots, u_l) \in X^l$, on définit $D(u) = (d(u_i, u_j))_{i,j \in [l]}$. On dit que $\Phi : \mathbb{M} \rightarrow \mathbb{R}$ est un polynôme s'il existe $\phi : \mathbb{R}^{l^2} \rightarrow \mathbb{R}$ continue et bornée telle que*

$$\forall X \in \mathbb{M}_2^{\mathbb{N}}, \Phi(X) = \int_{X^l} \phi(D(x)) d\mu^{\otimes l}(x).$$

On dit qu'une suite d'espaces métriques mesurés de mesure finie $(X^n)_{n \in \mathbb{N}}$ converge vers un espace métrique mesuré de mesure finie X^∞ si pour tout polynôme Φ ,

$$\Phi(X^n) \xrightarrow{n \rightarrow +\infty} \Phi(X^\infty).$$

Au cours de ce travail, nous rencontrerons de multiples situations où cette topologie est la mieux adaptée à l'étude de la coalescence. Cependant, là où nous gagnons en facilité, nous perdons en qualité de l'information issue de ces convergences. Par exemple, cette convergence ne permet pas de parler des diamètres de composantes. Heureusement, la différence entre ces deux topologies est bien connue:

Lemme 1.3.5 (Caractérisation de la borne inférieure des masses, Théorème 6.1 dans [7]). *Soit $(X^n, d_n, \mu_n)_{n \in \mathbb{N}}$ une suite de \mathbb{M} et $X \in \mathbb{M}$. Les assertions suivantes sont équivalentes :*

(i) $(X^n)_{n \in \mathbb{N}}$ converge vers X pour la topologie Gromov-faible et

$$\forall \delta > 0, \liminf_{n \rightarrow +\infty} m_\delta(X^n) > 0 \quad (1.4)$$

où $m_\delta(X^n) = \inf\{\mu_n(\{v : d_n(u, v) \leq \delta\}) : u \in X^n\}$.

(ii) X est compact et $(X^n)_{n \in \mathbb{N}}$ converge vers X pour la topologie de Gromov-Hausdorff-Prokhorov.

Notons que la borne inférieure des masses fonctionne de la même manière avec les espaces aléatoires : dans ce cas, nous avons besoin de la tension de $(1/m_\delta(X^n))_{n \in \mathbb{N}}$.

Lemme 1.3.6 (Caractérisation de la borne inférieure des masses, version aléatoire). *Soit $(X^n)_{n \in \mathbb{N}}$ une suite de variables aléatoire dans \mathbb{M} et X une variable aléatoire dans $\mathbb{M} \cap \tilde{\mathbb{M}}$. Les assertions suivantes sont équivalentes*

(i) $(X^n)_{n \in \mathbb{N}}$ converge en distribution vers X pour la topologie Gromov-faible et

$$\forall \delta > 0, \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow +\infty} \mathbb{P}(m_\delta(X^n) < \epsilon) = 0. \quad (1.5)$$

(ii) $(X^n)_{n \in \mathbb{N}}$ converge en distribution vers X pour la topologie de Gromov-Hausdorff-Prokhorov.

Proof. Voir Appendix A. □

De manière cruciale, comme nous le découvrirons plus tard, la propriété de borne inférieure des masses (1.5) sera souvent garantie dans notre contexte. Comme nous aimerions également étendre la topologie Gromov-faible à $\mathbb{M}_2^{\mathbb{N}}$, nous introduisons, pour $X \in \mathbb{M}_2^{\mathbb{N}}$

$$m_\delta^k(X) = \inf_{j \leq k} m_\delta(X_j).$$

1.3.4 Arbre réel, graphe réel et surplus

Le surplus d'un graphe connexe $G = (V, E)$ est égal à $|E| - |V| + 1$ et indique à quel point il diffère d'un arbre. Un arbre est un graphe connexe sans cycles et a toujours un surplus de 0. Puisque nous sommes dans le cadre général des convergences de graphes, on pourrait éventuellement vouloir savoir ce qu'il advient du surplus. Pour un graphe discret G donné, on peut associer à la suite ordonnée de ses composantes $(C_j)_{j \in \mathbb{N}}$ la suites de ses surplus $(s_j)_{j \in \mathbb{N}}$. Nous verrons plus loin dans le cas particulier des graphes aléatoires d'Erdős-Rényicritiques que les surplus des plus grandes composantes sont tendus.

De plus les composantes limites seront des graphes réels, ce qui signifie que pour tout point, il existe une boule telle que la restriction de la composante à cette boule soit un arbre réel : un espace métrique tel que deux points quelconques soient reliés par un unique chemin continu injectif. Ceci est défini précisément dans [3]. La définition du surplus s'étend naturellement aux graphes réels et on peut donc légitimement se demander si les surplus convergent. Remarquons que tout graphe discret G est canoniquement associé à un unique graphe réel \mathcal{G} où les arêtes sont remplacées par des copies de $[0, 1]$. De plus G , en tant qu'espace métrique, est isométriquement plongé dans \mathcal{G} . En conséquence, lorsque cela est nécessaire, tout graphe discret qui apparaît dans ce travail peut être considéré comme un graphe réel. Enfin, supposons que $(X^n)_{n \in \mathbb{N}} \in \mathbb{M}_2^{\mathbb{N}}$ soit une suite de graphes réels qui converge vers un graphe réel $X^\infty \in \mathbb{M}_2^{\mathbb{N}}$ pour T_{GHP} . De plus, supposons que le surplus de chaque X_i^n converge, pour $i \in \mathbb{N}$, vers le surplus de X_i^∞ . On dit alors que X^n converge vers X^∞ pour la *topologie de Gromov-Hausdorff-Prokhorov plus surplus produit*, notée T_{GHPS} .

1.4 Encodage d'espaces métriques aléatoires

Comme le suggère l'exposition des résultats sur le coalescent multiplicatif d'Aldous dans section 1.1.2, le processus Z peut encoder bien plus que la masse des composantes connexes. Il y a différentes façon d'encoder les propriétés des espaces métriques dans les fonctions réelles, ici nous donnons l'approche utilisée par Broutin, Duquesne et Wang dans [14]. C'est une méthode très puissante, car elle leur permet de décrire les espaces métriques limites de $\text{gCoal}(\cdot, \cdot)$ critiques, quels que soient les paramètres $(\alpha, \lambda, c) \in \mathcal{I}$.

L'encodage présenté dans ce chapitre est une variante de celui dont se servent Aldous et Limic dans [6]. Cependant, à la place d'effectuer un parcours en largeur sur les sommets, c'est à dire, explorer tous les enfants d'un sommet avant de passer au suivant, il effectue un parcours en profondeur des sommets. Les deux manières de faire sont intéressantes, mais la dernière nous permet de contrôler la fonction de hauteur, ce qui nous donne une vue plus précise sur la structure métrique des composantes. Tout d'abord, nous allons introduire le processus discret qui correspond au graphe discret aléatoire.

1.4.1 La file d'attente client-serveur

Soit $x \in \overline{l_{\searrow}}$ tel que $x_i > 0$ si et seulement si $i \leq n$. Le schéma proposé par Broutin, Duquesne et Wang pour la construction d'un graphe aléatoire discret multiplicatif $\text{gCoal}(x, q)$ est le suivant: chaque sommet $i \in [n]$ du graphe est un client qui arrive dans la file d'attente d'un serveur au temps aléatoire exponentiel ξ_i tel que

$$\mathbb{E}(\xi_i) = \frac{\sigma_1}{x_i}.$$

Si le serveur est occupé, une stratégie LIFO (last in first out) s'applique au temps ξ_i : le serveur interrompt le service de tous les précédents client pour servir le client i pendant une durée x_i . Introduisons nos deux premiers processus:

Définition 1.4.1 (Charge du serveur). *On définit la charge algébrique du serveur au temps t :*

$$Z(t) = -\frac{t}{\sigma_2} + \sum_{i \in [n]} \frac{x_i}{\sigma_2} 1_{\xi_i \leq t}$$

et la charge du serveur

$$B(t) = Z(t) - \inf_{[0, t]} Z.$$

L'évolution de la situation produit une représentation arborescente où, quand un client i arrive, si un client j est entrain d'être servi, alors j sera un ancêtre de i . Plus précisément:

Définition 1.4.2 (Processus de taille). *Soit Z une charge algébrique, on définit*

$$\alpha_i(t) = \begin{cases} 1 & \text{si } \xi_i \leq t, \text{ et } \inf_{[\xi_i, t]} Z > Z(\xi_i^-) \\ 0 & \text{sinon.} \end{cases}$$

De plus, on définit

$$H(t) = \sigma_2 \sum_i \alpha_i(t).$$

Finalement, on peut définir la forêt $T(x, q)$ comme le graphe sur $[n]$ tel que i est un descendant de j si et seulement si $\alpha_i(\xi_j) = 1$.

Examinons ce que signifie cette description. Quand, pour $t \in \mathbb{R}$ et $i \leq n$, $\alpha_i(t) = 1$, cela signifie que le client qui est arrivé au temps ξ_i n'a pas été entièrement servi au temps t . Dans la représentation arborescente nous disons que i est un ancêtre du client servi au temps t . Par conséquent, $H(t)$ est proportionnel au nombre d'ancêtres du client servi au temps t , c'est donc un processus de hauteur associé à $T(x, q)$. Bien sur $\text{gCoal}(x, q)$ n'est pas un arbre donc il nous faut ajouter une nouvelle construction afin d'obtenir $\text{gCoal}(x, q)$ partir de l'arbre encodé par H .

1.4.2 Le graphe multiplicatif discret

Ici, nous allons exposer la méthode qui nous permet de construire $\text{gCoal}(x, q)$ depuis B :

Définition 1.4.3 (Points de recollement). *Soit B une charge de serveur pour $x \in \overline{l_{\searrow}}$, soit \mathcal{P} qui est, conditionnellement à B , un processus de Poisson sur $[0, +\infty)^2$ avec pour intensité $1_{\{0 < s < B(t)\}} ds dt$. Considérons l'espace métrique $G(x, q)$ sur $[n]$ construit partir de $T(x, q)$ comme suit: on ajoute, pour chaque $(t, s) \in \mathcal{P}$, une arête entre le sommet servi au temps t et le sommet servi au temps $\inf\{u \leq t : \inf_{[u, t]} B > s\}$.*

Alors, nous avons

Théorème 1.4.1 (Théorème 2.1 dans [14]). *Soit $x \in \overline{l_{\downarrow}}$ et $q \in \mathbb{R}$. Alors le graphe $G(x, q)$ issu de la Définition 1.4.2 a la même loi que $g\text{Coal}(x, q)$.*

Finalement, comme le point de vue classique pour étudier la structure asymptotique de ces graphes est un point de vue métrique, on introduit:

Définition 1.4.4 (Espace discret multiplicatif). *Dans le même contexte que la Définition 1.4.3, on définit $G^n \in \mathbb{M}_2^{\mathbb{N}}$ l'espace métrique où l'ensemble de points est $[n]$, muni de la distance de graphe sur $G(x, q)$, multipliée par σ_2 et de la mesure de comptage sur $[n]$ multipliée par x . De plus, pour chaque j , on associe l_j l'instant ξ_i de l'arrivée du premier client i de la composante G_j^n , r_j le dernier instant auquel un sommet i de la composante G_j^n est servi, et s_j^n le surplus du graphe G_j^n .*

1.4.3 Processus continu

Ici nous décrivons comment le processus de limite d'Aldous et Limic dans [6] peut être utilisé pour définir un élément de $\mathbb{M}_2^{\mathbb{N}}$. Cela se fait d'une façon similaire au processus discret. Tout d'abord, on introduit l'équivalent continu de la charge du serveur:

Définition 1.4.5 (Processus de hauteur continu). *Pour $(\alpha, \lambda, c) \in \mathcal{I}$, soit $\mathcal{Z}^{(\alpha, \lambda, c)}$ tel que défini dans la Définition 1.4.2. Soit $\mathcal{H}^{(\alpha, \lambda, c)}$ le processus de hauteur associé à $\mathcal{Z}^{(\alpha, \lambda, c)}$ défini comme dans la Proposition 2.6 dans [15].*

Grace à la Proposition 14 dans Aldous et Limic [6], on sait que les excursions de \mathcal{B} peuvent être ordonnées par ordre décroissant de taille $(r_j - l_j)_{j \in \mathbb{N}}$. Nous allons montrer comment encoder les arbres réels à partir de fonctions réelles. Ce sujet originalement issu du travail d'Aldous [4] dans les années quatre-vingt-dix, est devenu maintenant standard. Soit $h : [l, r] \rightarrow \mathbb{R}_+$ une fonction continue telle que $h(r) = 0$, $h(s) \geq 0$ pour tout $s \in (l, r)$, on peut définir l'arbre [20] en posant

$$\forall u, v \in [l, r], d(u, v) = h(u) + h(v) - 2 \inf_{[u \wedge v, u \vee v]} h$$

et puis en prenant $[l, r] / \sim$ où $u \sim v \Leftrightarrow d(u, v) = 0$. Cela nous permet de définir la forêt $(\mathcal{T}_j)_{j \in \mathbb{N}}$ de la manière suivante:

$$\forall j \in \mathbb{N}, \forall u, v \in [l_j, r_j], d_j^{\mathcal{T}}(u, v) = \mathcal{H}(u) + \mathcal{H}(v) - 2 \inf_{[u \wedge v, u \vee v]} \mathcal{H},$$

$$\forall j \in \mathbb{N}, \forall u, v \in [l_j, r_j], u \sim_j^{\mathcal{T}} v \text{ si et seulement si } d_j^{\mathcal{T}}(u, v) = 0$$

$$\forall j \in \mathbb{N}, \mathcal{T}_j = [l_j, r_j] / \sim_j^{\mathcal{T}}.$$

Définition 1.4.6 (Espaces continus multiplicatifs). *Soit $(\alpha, \lambda, c) \in \mathcal{I}$, $\gamma \in \mathbb{R}_+$ et $t \in \mathbb{R}_+$, on définit la distribution $\mathcal{E}_0(\alpha, \lambda, c, \gamma)$ d'un espace multiplicatif au temps 0 comme suit. Soient $\mathcal{Z}^{(\alpha, \lambda, c)}$, $\mathcal{B}^{(\alpha, \lambda, c)}$, $\mathcal{H}^{(\alpha, \lambda, c)}$ définis comme dans la Définition 1.4.5 et soit \mathcal{P} un point du processus de Poisson sur \mathbb{R}_+^2 avec intensité $1_{v \leq q} \mathcal{B}^{(\alpha, \lambda, c)}(u) du dv$, conditionnellement sur $\mathcal{B}^{(\alpha, \lambda, c)}$. Pour tout $j \in \mathbb{N}$, on définit une distance d_j sur $[l_j, r_j]$ telle que:*

$$\forall u, v \in [l_j, r_j], d_j(u, v) = \inf_{k \geq 1} \left\{ \gamma \sum_{i=1}^k d_j^{\mathcal{T}}(u_i, v_i) : \begin{cases} (u_i)_{i \in [k]} \in [l_j, r_j]^k, (v_i)_{i \in [k]} \in [l_j, r_j]^k, \\ \forall i \in [2, k], (u_i, v_{i-1}) \in \mathcal{P} \text{ ou } (v_{i-1}, u_i) \in \mathcal{P}, \\ u_1 = u, v_k = v \end{cases} \right\}.$$

Alors si,

$$\forall j \in \mathbb{N}, \forall u, v \in [l_j, r_j], u \sim_j v \Leftrightarrow d_j(u, v) = 0,$$

on définit

$$\mathcal{X}_j = [l_j, r_j] / \sim_j.$$

Finalement, on dit que $\mathcal{X} = (\mathcal{X}_j)_{j \in \mathbb{N}}$ est un espace multiplicatif et on définit $\mathcal{E}_0(\alpha, \lambda, c, \gamma)$ comme la distribution de \mathcal{X} . De plus, pour $t \in \mathbb{R}$, on définit $\mathcal{E}_t(\alpha, \lambda, c, \gamma) = \mathcal{E}_0(\alpha, \lambda - t, c, \gamma)$ et on note $p_j^{\mathcal{X}}$ la projection de $[l_j, r_j]$ sur \mathcal{X}_j .

Par construction, un espace multiplicatif est une suite de graphes réels et par conséquent nous pouvons associer à chaque composante \mathcal{X}_j un surplus s_j .

1.4.4 Inhomogénéité dans une composante

Pour une raison qui s'expliquera par la suite, nous devons souvent quantifier l'importance de l'homogénéité à l'intérieur d'une composante d'un espace multiplicatif. Pour ce faire, nous allons définir les outils suivants:

Définition 1.4.7 (Quantité d'homogénéité dans une composante). *Soit $\mathcal{X} \sim \mathcal{E}_t(\alpha, \lambda, c, \gamma)$ construit à partir de $\mathcal{Z}^{(\alpha, \lambda-t, c, \gamma)}$. Soit $(\zeta_i)_{i \in \mathbb{N}}$ la suite de temps aléatoires exponentiels utilisée dans*

$$\mathcal{Z}^{(\alpha, \lambda-t, c, \gamma)}(s) = \alpha^{1/2} \mathcal{W}(s) + (\lambda - t)s - \frac{1}{2} \alpha s^2 + \sum_i (c_i 1_{\zeta_i \leq s} - c_i^2 s).$$

Pour tout $j > 0$, on pose

$$\Delta_j = \{i \in \mathbb{N} : \zeta_i \in (l_j, r_j]\},$$

$$\forall r > 0, \delta_r^{(j)} = \sum_{i \in \Delta_j} c_i^r.$$

et $c^{(j)}$ la suite telle que

$$c_i^{(j)} = \begin{cases} c_i & \text{si } i \in \Delta_j \\ 0 & \text{sinon.} \end{cases}.$$

1.4.5 Convergence des espaces métriques

Au cours de la dernière décennie, des efforts ont été fournis pour obtenir de cet encodage plus que la seule convergence des poids des composantes: la convergence de toute la structure métrique. Dans [1], Addario-Berry, Broutin et Goldschmidt ont été capables de montrer cette convergence quand tous les poids initiaux sont égaux, ce qui correspond au paramètre $(\alpha, \lambda, 0)$. Plus tard, dans [3], Addario-Berry, Broutin, Goldschmidt et Miermont affinent cette convergence en prenant en compte la distribution de la masse, obtenant ainsi une convergence pour la topologie de Gromov-Hausdorff-Prokhorov. D'autre part, dans [11], Bhamidi, van der Hofstad et Sen montrent la convergence pour les paramètres $(0, \lambda, c)$ où $c \in l_{\searrow}^3 \setminus l_{\searrow}^2$. Finalement, dans [14], Broutin, Duquesne et Wang développent une théorie unificatrice qui fonctionne pour tous les paramètres possibles dans les conditions d'Aldous-Limic 1. En effet, ils obtiennent le théorème suivant:

Théorème 1.4.2 (Convergence des espaces multiplicatifs, Théorème 2.5 and 2.8 of [14]). *Soient $(x, q) \in \overline{sl} \times \mathbb{R}^{\mathbb{N}}$ qui satisfont les conditions d'Aldous-Limic 1 avec les paramètres $(\alpha, \lambda, c) \in \mathcal{I}$. Soient \mathcal{Z}, H les processus associés et soit $G^n = (G_j^n) \in \mathbb{M}_2^{\mathbb{N}}$ l'espace multiplicatif discret de la Définition 1.4.4 et comme dans cette définition, soient $(s_j^n)_{j \in \mathbb{N}}, (l_j^n)_{j \in \mathbb{N}}, (r_j^n)_{j \in \mathbb{N}}$ les surplus, instants de début et de fin des excursions. Soient \mathcal{Z} et \mathcal{H} les processus d'exploration de paramètre (α, λ, c) de la Définition 1.4.6 et $(s_j)_{j \in \mathbb{N}}, (l_j)_{j \in \mathbb{N}}$ et $(r_j)_{j \in \mathbb{N}}$ les surplus, instants de début et de fin des excursions. Alors,*

$$(\mathcal{Z}, H, (s_j^n)_{j \in \mathbb{N}}, (l_j^n)_{j \in \mathbb{N}}, (r_j^n)_{j \in \mathbb{N}}) \xrightarrow[n \rightarrow +\infty]{(d)} (\mathcal{Z}, \mathcal{H}, (s_j)_{j \in \mathbb{N}}, (l_j)_{j \in \mathbb{N}}, (r_j)_{j \in \mathbb{N}}).$$

La convergence du processus est relative à la topologie de Skorokhod et la convergence des suites à la topologie produit. De plus,

$$G^n \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_0(\alpha, \lambda, c, 1)$$

pour T_{GHPS} .

1.5 Extension métrique de la coalescence

1.5.1 Différentes extensions possibles du coalescent multiplicatif d'Aldous

Rappelons que d'après notre Question directrice 1, nous sommes à la recherche d'un objet potentiellement limite des graphes multiplicatifs dans la fenêtre critique. L'approche d'Aldous et Limic qui consiste à ignorer la structure interne des composantes est un premier pas concluant et grâce à la section 1.4.5 on peut deviner qu'un bon point de vue pour étudier la structure interne des composantes est celui des

espaces métriques. Ainsi l'objectif de cette section est de définir un coalescent multiplicatif qui prend en entrée une suite d'espaces métriques mesurés et dont l'évolution des masses correspond à celle d'un coalescent multiplicatif d'Aldous. Ces contraintes peuvent être réalisées par diverses constructions, mais nous allons d'abord introduire la plus naturelle, qui provient originellement de Rossignol dans [28]. Tout d'abord, il nous faut la définition technique suivante:

Définition 1.5.1 (Collage d'espaces métriques). *Soit $X = (X, d, \mu) \in \mathbb{M}_2^{\mathbb{N}}$, soit d' une semi-métrique sur X , c'est à dire une fonction qui vérifie toutes les propriétés usuelles d'une métrique excepté que pour $u, v \in X$ tels que $d'(u, v) = 0$ on peut avoir $u \neq v$. Soit \mathcal{R} la relation d'équivalence définie par*

$$u\mathcal{R}v \Leftrightarrow d'(u, v) = 0$$

et soit $p_{\mathcal{R}} : X \rightarrow X/\mathcal{R}$ la projection canonique associée à cette relation d'équivalence. On définit $Glue(X, d')$ comme la complétion de $(X/\mathcal{R}, d' \circ p^{-1}, \mu \circ p^{-1})$.

On peut alors définir l'objet le plus important de cette thèse: la coalescence complète d'espaces métriques mesurés. Le terme "complet" vient du fait que l'on verra par la suite d'autres extensions du coalescent multiplicatif d'Aldous qui effectueront, en quelque sorte, moins de coalescence.

Définition 1.5.2 (Coalescence complète). *Soit $X = (X, d, \mu) \in \mathbb{M}_2^{\mathbb{N}}$, soit \mathcal{P} un processus ponctuel de Poisson sur $X^2 \times [0, +\infty)$ d'intensité $\frac{1}{2}\mu^2 \times Leb_{\mathbb{R}_+}$ et pour tout $t \geq 0$, soit*

$$\mathcal{P}_t = \{(u, v) \in X^2 : \exists s \leq t, (u, v, s) \in \mathcal{P}\}.$$

Pour tout $t > 0$, on définit la semi-métrique d_t^{Coal} sur X comme:

$$\forall u, v \in X, d_t^{Coal}(u, v) = \inf\left\{\sum_{i=1}^k d(u_i, v_i) : u_1 = u, v_k = v, \forall i \leq k-1, (v_i, u_{i+1}) \in \mathcal{P}_t\right\}.$$

On définit

$$Coal(X, t) = Glue(X, d_t^{Coal}).$$

Remarquons qu'il est montré dans [28] que les composantes de $Coal(X, q)$ sont dans $\tilde{\mathbb{M}}$ mais qu'il n'est pas évident qu'elles soient dans \mathbb{M} . D'autre part, les points de recollement qui se trouvent à l'intérieur de composantes déjà formées lorsqu'ils apparaissent ne modifient pas la suite des tailles de composantes. On peut ainsi définir un processus, que l'on peut espérer plus simple, qui n'ajoute pas ces points:

Définition 1.5.3 (Coalescence simple). *Soit $X = (X, d, \mu) \in \mathbb{M}_2^{\mathbb{N}}$, soit \mathcal{P} un processus ponctuel de Poisson sur $X^2 \times [0, +\infty)$ d'intensité $\frac{1}{2}\mu^2 \times Leb_{\mathbb{R}_+}$ et pour tout $t \geq 0$, soit*

$$\mathcal{SP}_t = \left\{ (u, v) \in X^2 : \begin{cases} \exists i, j \in \mathbb{N}, i \neq j \text{ s.t. } (u, v) \in X_i \times X_j \text{ and } \mathcal{P} \cap X_i \times X_j \times [0, t) = \emptyset \\ \exists s \leq t, (u, v, s) \in \mathcal{P} \end{cases} \right\}.$$

Pour tout $a \geq 0, t > 0$, on définit la semi-métrique $d_t^{SCoal_a}$ sur X comme:

$$\forall u, v \in X, d_t^{SCoal_a}(u, v) = \inf\left\{ka + \sum_{i=1}^k d(u_i, v_i) : u_1 = u, v_k = v, \forall i \leq k-1, (v_i, u_{i+1}) \in \mathcal{SP}_t\right\}.$$

On définit

$$SCoal_a(X, t) = Glue(X, d_t^{SCoal_a}).$$

De plus, on note $SCoal(X, t) = SCoal_0(X, t)$.

On remarque que, partant du même espace initial $X \in \mathbb{M}_2^{\mathbb{N}}$, il y a un couplage évident entre $Coal(X, \cdot)$ et $SCoal(X, \cdot)$ qui consiste à prendre le même processus de Poisson \mathcal{P} pour définir chaque processus.

Enfin, on définit:

Définition 1.5.4 (Coalescent (métrique) multiplicatif éternel). *On dit que $(X(t))_{t \in \mathbb{R}}$ est un coalescent multiplicatif éternel si c'est un processus de Markov dans \mathbb{M} tels que pour s et $t \in \mathbb{R}$ tel que $s \leq t$,*

$$(Coal(X(s), t-s), X(s)) \stackrel{(d)}{=} (X(t), X(s)).$$

1.5.2 Peut-on appliquer la méthode d'Aldous et Limic à la coalescence d'espaces métriques ?

La première question est celle de l'existence de coalescents multiplicatifs éternels. Rappelons que la réponse d'Aldous à cette question repose sur deux résultats: d'abord un résultat de convergence de graphes aléatoires discrets, qui serait le Théorème 1.4.2 dans notre cas, et une propriété de Feller. En effet, on a:

Théorème 1.5.1 (Propriété de Feller aléatoire pour la coalescence complète, conséquence du Théorème 3.1 de [28]). *Soit $X = (X^n)_{n \in \mathbb{N}}$ une suite de $\mathbb{M}_2^{\mathbb{N}}$, et pour $t \in \mathbb{R}_+$, $\mathcal{X}^n(t) = \text{Coal}(X^n, t)$. Supposons qu'il existe $X^\infty \in \mathbb{M}_2^{\mathbb{N}}$ et $t \in \mathbb{R}_+$ tels que*

$$X^n \xrightarrow[n \rightarrow +\infty]{(d)} X^\infty,$$

pour T_{GHP}^2 ,

$$\forall t \in \mathbb{R}, \forall \delta > 0, \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(m_\delta^k(\mathcal{X}^n(t)) < \epsilon) = 0.$$

et

$$\forall t \in \mathbb{R}_+, \mathbb{P}(\exists i, j \in \mathbb{N}, i \neq j \text{ s.t. } |\text{Coal}(X^\infty, t)_i| = |\text{Coal}(X^\infty, t)_j|) = 0.$$

Alors pour tout $t \in \mathbb{R}_+$, $\text{Coal}(X^\infty, t)$ est presque sûrement dans $\mathbb{M}_2^{\mathbb{N}}$ et pour toute famille $(t_0, \dots, t_k) \in \mathbb{R}_+^{k+1}$,

$$(\text{Coal}(X^n, t_i))_{i \in [k]} \xrightarrow[n \rightarrow \infty]{(d)} (\text{Coal}(X^\infty, t_i))_{i \in [k]}$$

pour T_{GHP}^2 .

Proof. Voir Appendix B. □

Dans le même esprit que la discussion qui suit le Théorème 1.2.2, le Théorème 1.5.1 donne, avec le Théorème 1.4.2, l'existence d'un coalescent multiplicatif éternel limite dont les marginales sont données par $\mathcal{E}_t(\alpha, \lambda, c, \gamma)_{t \in \mathbb{R}}$ (voir Appendix B).

La distribution d'un tel processus sera notée $\mathcal{E}(\alpha, \lambda, c, \gamma)$. Remarquons que rien ne garantit que ces coalescents soient extrêmes. Passons en revue les étapes qu'il faudrait franchir pour implémenter la méthode d'Aldous et Limic pour la coalescence métrique. Dans cette optique on introduit quelques points de théorie des processus de Markov: soit f mesurable bornée pour T_{GHP} , et $X = (X(t))_{t \in \mathbb{R}}$ un coalescent multiplicatif éternel. Pour $t \in \mathbb{R}$, soit \mathcal{F}_t la tribu engendrée par $(X(t-s))_{s \geq 0}$. Comme X est un processus de Markov, pour $t \in \mathbb{R}$ et $s > 0$,

$$\mathbb{E}(f(\text{Coal}(X(t-s), s)) | X(t-s)) = \mathbb{E}(f(\text{Coal}(X(t-s), s)) | \mathcal{F}_{t-s}).$$

De plus comme $(\text{Coal}(X(t-s), s), X(s)) \stackrel{(d)}{=} (X(t), X(s))$, d'après le théorème de convergence pour les martingales inverses, on a que

$$\mathbb{E}(f(X(t)) | X(t-s)) \xrightarrow[s \rightarrow \infty]{} \mathbb{E}(X(t) | \mathcal{F}_{-\infty}). \quad (1.6)$$

D'autre part, d'après [18], X est extrême si et seulement si pour toute fonction f mesurable bornée,

$$\mathbb{E}(f(X(t)) | \mathcal{F}_{-\infty}) = \mathbb{E}(f(X(t))). \quad (1.7)$$

On voudrait prouver les deux assertions suivantes:

Assertion 1. *Si $X \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$, pour $(\alpha, \lambda, c, \gamma) \in \mathcal{I} \times \mathbb{R}_+$, X est extrême.*

Idée de preuve possible. cette assertion serait une conséquence des assertions suivantes, non démontrées:

- (i) Il existe un espace de paramètres H de dimension dénombrable, une fonction $\Gamma' : \mathbb{M}_2^{\mathbb{N}} \times \mathbb{R}_+ \rightarrow H$ et une fonction $P : \mathcal{I} \times \mathbb{R}_+ \rightarrow H$ tel que

$$\Gamma'(X(t), -t) \xrightarrow[t \rightarrow -\infty]{\mathbb{P}} P(\alpha, \lambda, c, \gamma),$$

ce qui est un équivalent des conditions d'Aldous-Limic 1.

(ii) Si $X = (X^n)_{n \in \mathbb{N}}$ est une suite d'espaces métriques $\mathbb{M}_2^{\mathbb{N}}$ et $q = (q_n)_{n \in \mathbb{N}}$ est une suite de \mathbb{R}_+ telle que

$$\Gamma'(X^n, q_n) \xrightarrow[n \rightarrow \infty]{} P(\alpha, \lambda, c, \gamma),$$

alors

$$\text{Coal}(X^n, q_n) \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{E}_0(\alpha, \lambda, c, \gamma).$$

En effet, on aurait

$$\mathbb{E}(f(X(t)) | X(t-s)) \xrightarrow[s \rightarrow \infty]{\mathbb{P}} \mathbb{E}(f(X(t))),$$

donc d'après (1.6),

$$\mathbb{E}(f(X(t))) = \mathbb{E}(f(X(t)) | \mathcal{F}_{-\infty})$$

et donc d'après (1.7), X est extrême. \square

Assertion 2. Si X est un coalescent multiplicatif éternel, il existe $(\alpha, \lambda, c, \gamma) \in \mathcal{I} \times \mathbb{R}_+$ tel que $X \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$.

Idée de preuve possible. cette assertion serait une conséquence des assertions suivantes, non démontrées:

(i) Il existe un espace de paramètres H de dimension dénombrable, une fonction $\Gamma' : \mathbb{M}_2^{\mathbb{N}} \times \mathbb{R}_+ \rightarrow H$ et une fonction $P : \mathcal{I} \times \mathbb{R}_+ \rightarrow H$ tel que

$$\Gamma'(X(t), -t) \xrightarrow[t \rightarrow -\infty]{a.s.} P(\alpha, \lambda, c, \gamma).$$

(ii) Si $X = (X^n)_{n \in \mathbb{N}}$ est une suite de $\mathbb{M}_2^{\mathbb{N}}$ et $q = (q_n)_{n \in \mathbb{N}}$ est une suite de \mathbb{R}_+ telle que

$$\Gamma'(X^n, q_n) \xrightarrow[n \rightarrow \infty]{} P(\alpha, \lambda, c, \gamma),$$

alors

$$\text{Coal}(X^n, q_n) \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{E}_0(\alpha, \lambda, c, \gamma).$$

En effet, on aurait

$$\mathbb{E}(f(X(t)) | X(t-s)) \xrightarrow[s \rightarrow \infty]{\mathbb{P}} \mathbb{E}(f(\mathcal{X}(t))),$$

avec $\mathcal{X}(t) \sim \mathcal{E}_t(\alpha, \lambda, c, \gamma)$, et comme d'après (1.6),

$$\mathbb{E}(f(X(t)) | X(t-s)) \xrightarrow[s \rightarrow \infty]{\mathbb{P}} \mathbb{E}(f(X(t)) | \mathcal{F}_{-\infty}),$$

cela donnerait que

$$\mathbb{E}(f(X(t))) = \mathbb{E}(\mathbb{E}(f(X(t)) | \mathcal{F}_{-\infty})) = \mathbb{E}(f(\mathcal{X}(t)))$$

donc $X \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$. \square

On remarque que, bien que l'Assertion 1 et l'Assertion 2 soient d'apparence opposées, les plans de preuves hypothétiques sont assez similaires. Dans les deux cas, l'étape (i) est une question complexe, que l'on va laisser de côté pour l'instant. Pour l'Assertion 1, le cas où $X \sim \mathcal{E}(1, \lambda, 0, 1)$ sera le principal objectif du chapitre 8. Examinons quels résultats nous pouvons trouver dans la littérature qui satisferaient à l'exigence de (ii). On aura besoin des définitions suivantes:

Définition 1.5.5. Soit $X = (X, d, \mu) \in \mathbb{M}_2^{\mathbb{N}}$, pour tout $r \in \mathbb{N}$, on définit

$$\rho_r = \rho_r(X) = \int_{X^2} d(u, v)^r \mathbf{1}_{d(u, v) < +\infty} \mu(du) \mu(dv).$$

En particulier, ρ_1 peut aussi s'écrire

$$\rho_1 = \sum_i x_i^2 u_i$$

où, pour tout $i \in \mathbb{N}$,

$$x_i = |X_i|$$

et

$$u_i = x_i^{-2} \int_{X_i^2} d_i(u, v) \mu_i(du) \mu_i(dv)$$

est l'espérance de la distance entre deux points aléatoires dans X_i . De plus, le diamètre des composantes X_i sera noté D_i . Enfin, si $X \in \overline{\mathbb{M}}^{\mathbb{N}}$, $x_{\min} = x_n$ avec n le dernier indice tel que $x_i > 0$. On remarque que, comme d'habitude, si $X = (X^n)_{n \in \mathbb{N}}$, respectivement $X = (X(t))_{t \in \mathbb{R}}$, est une famille d'éléments de $\mathbb{M}_2^{\mathbb{N}}$, ρ_1 sera sous-entendu comme la suite $\rho_1 = (\rho_1^n)_{n \in \mathbb{N}}$, respectivement $\rho_1 = (\rho_1(t))_{t \in \mathbb{R}}$, avec ρ_1^n associé à X^n , respectivement $\rho_1(t)$ associé à $X(t)$. On fait de même pour $u = (u_i)_{i \in \mathbb{N}}$, $x = (x_i)_{i \in \mathbb{N}}$, $D = (D_i)_{i \in \mathbb{N}}$ et x_{\min} .

Théorème 1.5.2 (Convergence de la coalescence simple dans le cas homogène, Théorème 3.4 de [9]). *Soit $X = (X^n)_{n \in \mathbb{N}}$ une suite de $\overline{\mathbb{M}}^{\mathbb{N}}$, $x = (x^n)_{n \in \mathbb{N}} = (|X_i^n|)_{n \in \mathbb{N}, i \in \mathbb{N}}$ et $q = (q_n)_{n \in \mathbb{N}}$ tels que (x, q) satisfait les conditions d'Aldous-Limic 1 avec paramètres $(1, \lambda, 0)$. De plus on suppose les conditions supplémentaires suivantes: il existe $\eta > 0$ et $r > 0$ tel que*

$$\begin{cases} \frac{x_0}{\sigma_2} = o(\sigma_2^{1/2+\eta}) \\ \frac{\sigma_2^r}{x_{\min}} \rightarrow 0, \end{cases}$$

et la condition métrique:

$$\begin{cases} \frac{\sup_i D_i}{\rho_1 + \sigma_2} = o(\sigma_2^{1/2+\eta}) \\ \frac{x_0 \sigma_2 \sup_i D_i}{\rho_1} \rightarrow 0, \end{cases}$$

Alors,

$$\frac{\sigma_2^2}{\rho_1 + \sigma_2} S\text{Coal}_{\sigma_2}(X, q) \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_0(1, \lambda, 0, 1)$$

pour T_{GHP} .

Ce théorème porte sur des graphes qui sont fortement homogènes: le paramètre c est nul. Il existe un résultat, le Théorème 5.4 de [10] qui rend compte similairement de la situation où $(\alpha, \lambda, c) \in \mathcal{I}^*$. Ce résultat sera énoncé précisément en Section 2.5.3 mais il s'accompagne de beaucoup de détails techniques qui ne sont pas absolument indispensables pour le moment.

1.5.3 La question plus faible du mélange de la percolation dynamique et sa relation à la sensibilité au bruit

Que ce soit pour le plan de preuve de l'Assertion 1 où de l'Assertion 2, la première étape consiste à montrer un résultat sur des asymptotiques d'un coalescent multiplicatif éternel X en $-\infty$. Seulement, dans un cas, Assertion 1, la distribution de X est donnée et dans l'autre, Assertion 2, on sait que X est extrême. On peut s'attendre à ce que le fait de connaître la distribution de X soit un peu plus facile, étant donné que les marginales $\mathcal{E}_t(\alpha, \lambda, c, \gamma)$ sont bien comprises. Il se trouve par ailleurs que ce serait un résultat intéressant en lui-même, grâce à la théorie développée dans [28]:

Théorème 1.5.3 (Théorème 3.7 de [28]). *Soit $X \sim \mathcal{E}_0(1, \lambda, 0, 1)$, il existe un processus $\text{Perc}(\cdot, \cdot)$ qu'on appelle la percolation dynamique tel que, si $G(0) \sim G(n, \frac{1+\lambda n^{-1/3}}{n})$ et $(G(t))_{t \in \mathbb{R}}$ est le processus qui partant de $G(0)$ met à jour indépendamment les arêtes à taux $n^{-1/3}$,*

$$(G(t))_{t \geq 0} \xrightarrow[n \rightarrow +\infty]{(d)} (\text{Perc}(X, t))_{\geq 0}$$

pour T_{GHP} . De plus, pour tout $t \geq 0$, $\text{Perc}(X, t) \sim \mathcal{E}_0(1, \lambda, 0, 1)$. Enfin, si $(X(t))_{t \in \mathbb{R}} \sim \mathcal{E}(1, \lambda, 0, 1)$, pour tout f positive mesurable pour T_{GHP} ,

$$(X(t), \text{Perc}(X(t), s)) \stackrel{(d)}{=} (\text{Coal}(X(s), t-s), \widetilde{\text{Coal}}(X(s), t-s)). \quad (1.8)$$

où les processus de Poisson utilisés dans $\text{Coal}(X(s), \cdot)$ et $\widetilde{\text{Coal}}(X(s), \cdot)$ sont indépendants.

En particulier, du fait de (1.8), l'Assertion 1 donnerait la propriété de mélange de Perc(\cdot, t) suivante: pour toutes f, g positives et mesurables,

$$\mathbb{E}(f(\text{Perc}(X(s), t-s))g(X(s))) = \mathbb{E}(f(\text{Coal}(X(s), t-s))g(X(s))) \rightarrow \mathbb{E}(f(X(t)))\mathbb{E}(g(X(t))).$$

Cette propriété de mélange est intéressante en particulier grâce à sa relation avec la sensibilité au bruit. Donnont un peu de contexte:

Définition 1.5.6 (Sensibilité au bruit [8]). *Pour tout $n \in \mathbb{N}$, soit Λ_n un ensemble fini tel que $|\Lambda_n| \xrightarrow{n \rightarrow +\infty} +\infty$, soit \mathbb{P}_n une mesure de probabilité sur $\Omega_n = \{0, 1\}^{\Lambda_n}$. Soit ω_n une variable aléatoire de loi \mathbb{P}_n . Pour $\epsilon > 0$, on définit ω_n^ϵ de la manière suivante: soit $\omega'_n \sim \mathbb{P}_n$ indépendant de ω_n , on définit pour $i \in \Lambda_n$, $\omega_n^\epsilon(i) = \omega_n(i)$ avec probabilité $1 - \epsilon$ et $\omega_n^\epsilon(i) = \omega'_n(i)$ avec probabilité ϵ indépendamment pour chaque coordonnée. On dit que l'évènement $E_n \subset \Omega_n$ est sensible au bruit si*

$$\begin{cases} \mathbb{P}(E_n) \text{ est bornée loin de } 0 \text{ et de } 1, \\ \forall \epsilon > 0, \mathbb{P}(\omega_n \in E_n, \omega_n^\epsilon \in E_n) - \mathbb{P}(\omega_n \in E_n)\mathbb{P}(\omega_n^\epsilon \in E_n) \xrightarrow{n \rightarrow +\infty} 0. \end{cases}$$

Le domaine des graphes aléatoires est un bon laboratoire de sensibilité au bruit: on prend $\Lambda_n = \{\{i, j\} \in [n]^2 : i < j\}$ et pour $e \in \Lambda_n$ on associe $\omega(e) = 1$ à l'existence de l'arrête e dans le graphe G^n . Ici on se concentre sur le cas où \mathbb{P}_n est tel que pour $u \in \{0, 1\}^n$,

$$\mathbb{P}(\forall i \in [n], \omega_n(i) = u_i) = p^{\sum_i u_i} (1-p)^{n-\sum_i u_i}$$

pour un certain $p = (p_n)_{n \in \mathbb{N}}$. Le graphe que l'on construit alors est bien sûr un graphe aléatoire d'Erdős-Rényi $G(n, p)$. Le graphe construit à partir de ω_n^ϵ , c'est à dire après application du bruit à G^n sera noté $G^{n, \epsilon}$. On considère toujours G^n et $G^{n, \epsilon}$ comme des espaces métriques éléments de $\mathbb{M}_2^{\mathbb{N}}$, en prenant la distance de graphe divisée par $n^{1/3}$ et la mesure de comptage divisée par $n^{2/3}$.

Pour X et X' dans $\mathbb{M}_2^{\mathbb{N}}$ on définit

$$d_{GHP}^4(X, X') = \sum_j (d_{GHP}(X_j, X'_j))^4.$$

La remarque fondamentale est qu'appliquer le bruit à G^n est la même chose qu'effectuer une percolation dynamique dessus. Grâce à ca, Lubetzky et Peled ont pu obtenir dans [23]:

Théorème 1.5.4 (Théorème 2 de [23]). *Soit $G^n \sim G(n, p)$ équipé de la distance de graphe divisé par $n^{1/3}$ et de la mesure de comptage divisée par $n^{2/3}$, et $G^{n, \epsilon}$ le résultat de l'application d'un bruit de niveau ϵ à G^n . On suppose que $p = \frac{1+tn^{-1/3}}{n}$. Soit $\delta > 0$, on suppose que $\epsilon \geq n^{-1/3+\delta}$. Alors*

$$(G^n, G^{n, \epsilon}) \xrightarrow[n \rightarrow +\infty]{(d)} (X, Y)$$

où $X \sim \mathcal{E}_0(1, t, 0, 1)$ et $Y \stackrel{(d)}{=} X$ est indépendant de X . Cette convergence a lieu pour d_{GHP}^4 . D'autre part, si $\epsilon = o(n^{-1/3})$,

$$d_{GHP}^4(G^n, G^{n, \epsilon}) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Remarquons que la convergence pour d_{GHP}^4 est une conséquence de la convergence pour T_{GHP}^2 et du fait que G^n et $G^{n, \epsilon}$ convergent séparément pour d_{GHP}^4 . La seconde partie du Théorème 1.5.4 est en fait une conséquence de:

Théorème 1.5.5 (Conséquence directe de [28], Théorème 3.7). *Soit $G^n \sim G(n, p)$ équipé de la distance de graphe divisée par $n^{1/3}$ et de la mesure de comptage divisée par $n^{2/3}$, et $G^{n, \epsilon}$ le résultat de l'application d'un bruit de niveau ϵ à G^n . On suppose que $p = \frac{1+tn^{-1/3}}{n}$, $\epsilon_n n^{-1/3} \xrightarrow{n \rightarrow +\infty} s \in \mathbb{R}^+$. Alors,*

$$(G^n, G^{n, \epsilon}) \xrightarrow[n \rightarrow +\infty]{(d)} (X, \text{Perc}(X, s))$$

où $X \sim \mathcal{E}_0(1, t, 0, 1)$.

Ce théorème a également l'avantage de fournir un indice sur la façon d'obtenir le résultat de sensibilité au bruit à partir de la propriété de mélange de $\text{Perc}(\cdot, \cdot)$: à mesure que le bruit augmente, il efface de plus en plus d'informations sur le graphe d'origine. Si $\epsilon n^{1/3} \rightarrow +\infty$, pour n assez grand, $\epsilon \geq \epsilon'$ pour tout ϵ' tel que $\epsilon' n^{1/3} \rightarrow s$ dans \mathbb{R}_+ . Ainsi, grâce au* Théorème 1.5.5, pour n assez grand, le couple $(G^n, G^{n, \epsilon'})$ se comporte comme $(X, \text{Perc}(X, s))$, et comme $(G^n, G^{n, \epsilon})$ est "plus indépendant" que $(G^n, G^{n, \epsilon'})$ pour tout s , il est "plus indépendant" que $(X, \text{Perc}(X, s))$ pour tout s . Si la propriété de mélange est vraie, le couple $(X, \text{Perc}(X, s))$ devient lui-même asymptotiquement indépendant donc c'est également le cas de $(G^n, G^{n, \epsilon})$.

1.6 Résultats

Tous les résultats énoncés dans cette section sont originaux. L'analyse de la section précédente laisse entendre que la Question directrice 1 se décompose naturellement en deux sous-questions:

Question directrice 2 (À propos des asymptotique en $-\infty$). *Soit $(X(t))_{t \in \mathbb{R}}$ un coalescent multiplicatif éternel. Que peut-on dire de son comportement en $-\infty$?*

Question directrice 3 (À propos de la convergence de la coalescence). *Soit $X = (X^n)_{n \in \mathbb{N}}$ une suite de $\mathbb{M}_2^{\mathbb{N}}$ et $q = (q_n)_{n \in \mathbb{N}} \in \mathbb{N}$. Sous quelles conditions sur (X, q) a-t-on que*

$$\text{Coal}(X^n, q_n) \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_t(\alpha, \lambda, c, \gamma)$$

pour des paramètres $(\alpha, \lambda, c, \gamma)$?

Tous nos résultats s'articulent autour de ces deux axes, l'objectif étant l'exploration complète de la frontière d'entrée du coalescent métrique multiplicatif éternel, ce qui est une autre façon de dire que nous voulons caractériser tous les coalescents métriques extrêmes. Nous n'avons pas de théorème qui réponde entièrement à cette question, mais plusieurs réponses partielles à la Question directrice 2 et la Question directrice 3.

Avant d'énoncer nos résultats, nous souhaitons souligner un point en particulier concernant la Question directrice 2. Dans [6], l'équivalent de cette question pour le processus des masses d'un coalescent multiplicatif extrême repose sur des méthodes de martingales. Cependant, comme nous l'avons vu dans la Section 1.5.3, il est ici nécessaire de contrôler des distances moyennes sur les composantes. Cela rend la mise en œuvre des méthodes de martingales plus difficile car il n'y a pas de moyen évident de voir ce que devient la distance moyenne sur une composante après que deux de ses points aient coalescé. Cependant, si l'on ne s'intéresse qu'à la propriété de mélange de $\text{Perc}(\cdot, \cdot)$ à partir de $X \sim \mathcal{E}_t(\alpha, \lambda, c, \gamma)$, les composantes ne sont plus arbitraires et l'analyse devient plus facile.

1.6.1 Outils

La première partie de cette thèse est consacrée à une série d'outils permettant de manipuler les coalescents multiplicatifs.

Des statistiques globales aux statistiques par composantes Il est fréquent de voir dans la littérature des résultats de convergence portant sur des composantes conditionnées à être connexes plutôt que sur toute la suite d'espaces métriques. L'objectif du Chapitre 3 est de mettre en évidence dans quelles conditions on peut inférer des résultats globaux, c'est à dire sur toute la suite d'espaces métriques, à partir de résultats composante par composante.

Troncature et topologie Gromov-faible De même, on a souvent des résultats sur la coalescence d'un nombre fini de composantes, alors que nous souhaitons coalescer un nombre infini de composantes. Dans le Chapitre 4 on verra que dans le cadre de la topologie Gromov-faible, la différence entre les deux approches n'est pas significative sous de bonnes conditions.

De la coalescence simple à la coalescence complète Comme la littérature se concentre sur la coalescence simple et que nous nous intéressons principalement à la coalescence complète, le Chapitre 5 fera le lien entre les deux.

Coalescent multiplicatif éternel, mélange et frontière d'entrée Dans le Chapitre 6, nous montrons que la tension de la borne de plus petite masse est garantie pour un coalescent éternel. Cela permet d'établir le Théorème 6.2.1 qui dit que la distribution d'un coalescent est partiellement caractérisée par la convergence de certaines quantités en $-\infty$.

1.6.2 Résultats principaux

La deuxième partie de ce travail apporte des réponses partielles à la question de la caractérisation complète de la frontière d'entrée, qui peut également être vue comme un moyen de mieux comprendre les coalescents multiplicatifs et la percolation dynamique.

Mesure de Palm du processus ponctuel associé à $\mathcal{E}_t(\alpha, \beta, c, \gamma)$ Dans le Théorème 1.2.5, Janson et Spencer ont obtenu une formule explicite de la mesure de Palm de $\mathcal{M}_0(1, \lambda, 0)$. Puisqu'il permet entre autres de transformer des moments d'ordre deux en moments d'ordre un, c'est un outil utile pour montrer la concentration des quantités liées à $x = (x(t))_{t \in \mathbb{R}} \sim \mathcal{M}(1, \lambda, 0)$. À la lumière du problème plus simple de mélange de la percolation dynamique à partir de $X \sim \mathcal{E}_t(\alpha, \lambda, c, \gamma)$, il serait utile d'avoir un résultat similaire mais avec des espaces métriques à la place des poids et de généraliser également à tous les paramètres dans \mathcal{I} . Dans le Chapitre 7, on obtient:

Theorem 7.1.1. *Soit $(\alpha, \lambda, c, \gamma) \in \mathcal{I} \times \mathbb{R}_+$ et $(\Delta_j)_{j \in \mathbb{N}}, (c^{(j)})_{j \in \mathbb{N}}$ et $(\delta_2^{(j)})_{j \in \mathbb{N}}$ associés à $\mathcal{X} \sim \mathcal{E}_0(\alpha, \lambda, c, \gamma)$ comme dans la Définition 1.4.7. Soit $f : \mathbb{M} \times \mathbb{M}_2^{\mathbb{N}} \rightarrow \mathbb{R}^+$ mesurable pour T_{GHP} et telle que*

$$\mathbb{E}\left(\sum_j f(\mathcal{X}_j, \mathcal{X})\right) < +\infty.$$

Pour $j \in \mathbb{N}$, conditionnellement à $(\mathcal{X}_j, \Delta_j)$, soit $\mathcal{Y}^{(j)}$ de loi $\mathcal{E}_0(\alpha, \lambda - |\mathcal{X}_j| - \delta_2^{(j)}, c - c^{(j)}, \gamma)$. On suppose que

$$\text{soit } c = 0, \text{ soit } \lambda < 0. \quad (7.1)$$

On a:

$$\mathbb{E}\left(\sum_j f(\mathcal{X}_j, \mathcal{X})\right) = \mathbb{E}\left(\sum_j \mathbb{E}(f(\mathcal{X}_j, \mathcal{Y}^{(j)} \sqcup \mathcal{X}_j | (\mathcal{X}_j, \Delta_j)))\right). \quad (7.2)$$

Propriété de mélange pour le coalescent multiplicatif métrique homogène Comme déjà expliqué dans la Section 1.5.3, la question du mélange de la percolation dynamique est une question liée à l'application de la méthode d'Aldous et Limic à la coalescence complète des espaces métriques. Comme on part d'une distribution connue $\mathcal{E}_t(1, \lambda, 0, 1)$, la question de la frontière d'entrée Question directrice 2 est plus facile. Nous nous appuyons ici sur la formule de Janson et Spencer du Théorème 7.1.1, qui dans le Chapitre 8 nous permet d'obtenir un contrôle sur la variance des statistiques métriques.

Theorem 8.2.2. *Soit $(X(t))_{t \in \mathbb{R}} \sim \mathcal{E}(1, \lambda, 0, 1)$. Soient ρ_1, u, x, D définis comme dans la Définition 1.5.5. Il existe une constante K_r telle que, lorsque $t \rightarrow -\infty$,*

$$\forall p \geq 2, r \geq 0, \sum_j x_j^p u_i^r \sim_{\mathbb{P}} K_r \sigma_{p+r/2}, \quad (8.1)$$

$$\forall p \geq 2, \sigma_p \sim \mathbb{E}(\sigma_p), \quad (8.2)$$

$$\forall \eta \in (0, 1), \frac{x_0}{\sigma_2} = o_{\mathbb{P}}(\sigma_2^\eta) \quad (8.3)$$

$$\forall \eta \in (0, 1), \sup_i D_i = o_{\mathbb{P}}(\sigma_2^\eta) \quad (8.4)$$

Ceci nous permet de montrer:

Corollary 8.2.3. *Soit $X \sim \mathcal{E}(1, \lambda, 0, 1)$, alors pour tout $t \in \mathbb{R}$, pour toutes les fonctions f et g mesurables et bornées pour T_{GHPS} ,*

$$\mathbb{E}(f(\text{Perc}(X(t), s))g(X(t))) \xrightarrow{s \rightarrow +\infty} \mathbb{E}(f(X(t)))\mathbb{E}(g(X(t))).$$

Sensibilité au bruit Comme déjà expliqué dans la Section 2.5.3, la propriété de mélange est fortement liée à la sensibilité au bruit. Dans le Chapitre 9, nous montrons que l'indépendance asymptotique du

Théorème 1.5.4 reste vraie pour un bruit de niveau ϵ tel que $\epsilon n^{1/3} \rightarrow \infty$, ce qui est plus fort. De plus, nous donnons divers exemples de propriétés sensibles au bruit: la propriété de contenir une grande structure topologique donnée (comme un grand cycle par exemple) et la propriété qu'il existe un plongement d'un graphe dans une surface donnée.

1.6.3 Perspective

La dernière partie de ce travail est consacrée à l'esquisse d'un programme qui, s'il était entièrement réalisé, donnerait une caractérisation complète de la frontière d'entrée du coalescent multiplicatif métrique. D'une part, nous proposons d'étudier un encodage de $\text{SCoal}(\cdot, \cdot)$ à la Broutin, Duquesne et Wang [14] et nous montrons que ce processus est tendu dans le Lemme 10.3.4. D'autre part, nous montrons dans le Théorème 10.4.2 qu'il est possible de coupler un coalescent donné $(X(t))_{t \in \mathbb{N}}$ avec un processus $(T(t))_{t \in \mathbb{N}}$ qui n'ajoute pas de complexité à l'intérieur des composantes, de telle sorte que leurs asymptotiques $-\infty$ soient les mêmes. Cela annule la difficulté déjà évoquée de contrôler les statistiques de distance lors de l'ajout de points de coalescence internes.

Chapter 2

English introduction

2.1 Context: discrete random graphs at criticality

Although our work does not exclusively deal with discrete random graphs, our goal is to better our understanding of a certain class of discrete random graphs who, loosely put, lack geometrical constraints. Let us first introduce the main historical example of such graph.

2.1.1 First meeting with random graphs

One of the simplest examples of random graphs that one might think of is as follows: for $n \in \mathbb{N}$, $p \in [0, 1]$, we consider an undirected graph $G^n = (V, E)$ on $V = \llbracket 0, n \rrbracket = [n]$ where

$$\forall i, j \in [n], i \neq j, \mathbb{P}(\{i, j\} \in E) = p$$

independently across all pairs $\{i, j\}$. The law of such a graph will be denoted by $G(n, p)$ and we say that it is an Erdős–Rényi random graph.

We are especially interested in large sparse graphs, which means that we expect a typical vertex to have a degree of order 1 when n tends to infinity. This can be realised by taking $p = \frac{c}{n}$ with $c > 0$.

2.1.2 The critical window

A striking finding in the original work of Erdős and Rényi [19] is the critical behaviour of the size of the largest connected components. To state this result, we give some definitions: two vertices are in the same connected component if one can be reached from the other via a path of edges in E . The size of a given component refers simply to its number of vertices. We say that a component is larger than another if it has more vertices, and when the numbers of vertices are equal, if its smallest index is smaller than the smallest index of the other one. Moreover as a matter of convenience, we always consider that a graph contains a countable infinite amount of empty components. This allows us to associate to G^n the sequence $(C_i^n)_{i \in \mathbb{N}}$ of its components listed in decreasing order.

We have the following phase transition [19]:

$$\text{If } c < 1, \exists \delta > 0 \text{ s.t. } \mathbb{P}(|C_0^n| \leq \delta \log(n)) \xrightarrow[n \rightarrow +\infty]{} 1, \quad (2.1)$$

$$\text{If } c > 1, \exists \delta > 0 \text{ s.t. } \forall k \geq 1, \mathbb{P}(|C_0^n| \geq \delta n, |C_k^n| \leq \delta \log(n)) \xrightarrow[n \rightarrow +\infty]{} 1. \quad (2.2)$$

The critical point, $c = 1$ of this phase transition exhibits the following behaviour [13]: let $\lambda \in \mathbb{R}$,

$$\text{If } c = 1 + \lambda n^{-1/3}, \forall a, b > 0 \text{ s.t. } a < b, \forall k \in \mathbb{N}, \lim_{n \rightarrow +\infty} \mathbb{P}(an^{2/3} \leq |C_k^n| \leq bn^{2/3}) \in (0, 1). \quad (2.3)$$

2.1.3 Dynamical approach and almost multiplicative merging

This last assertion (2.3) highlights the existence of a critical window, parameterized by λ , during which the giant component, that occupy most of the vertices when $c > 1$, is formed. As this fact suggests, it is classical to examine these graphs with a dynamical point of view: for each $i, j \in [n]$, $i \leq j$, let $E_{i,j} \sim \mathcal{E}(1)$

independently across all i, j such that $i \leq j$ and let $E_{i,j} = E_{j,i}$. Then for $t \geq -n^{1/3}$, let $G^n(t) = (V, E(t))$ where

$$E(t) = \left\{ \{i, j\} : i, j \in [n], i < j, E_{i,j} \leq \frac{1 + tn^{-1/3}}{n} \right\}.$$

In particular the probability of existence of a particular edge $\{i, j\}$ at a fixed time t is, as n goes to infinity,

$$\mathbb{P}(\{i, j\} \in E(t)) = 1 - e^{-\frac{1+tn^{-1/3}}{n}} = \frac{1 + tn^{-1/3}}{n} + o(n^{-4/3})$$

so $G^n(t) \sim G(n, \frac{1+tn^{-1/3}}{n} + o(n^{-4/3}))$. Moreover, if for $t > 0$ and $V_1, V_2 \in V$, $E(t, V_1, V_2)$ is the set of all edges between V_1 and V_2 existing at time t , as the edges are independent, the merging probability between the two disjoint sets V_1 and V_2 during the time period $[s, t]$ is

$$\mathbb{P}(E(t, V_1, V_2) \neq \emptyset | E(s, V_1, V_2) = \emptyset) = 1 - e^{-\frac{t-s}{n^{4/3}} |V_1| |V_2|}.$$

As a consequence, the rate of merging of the two disjoint sets is the product of the scaled sizes of the disjoint sets $|V_1|n^{-2/3}|V_2|n^{-2/3}$. This behaviour, that we will describe as *multiplicative*, is at the heart of this work. We say that a graph G on $[n]$ is *multiplicative* if there exists a weight sequence $(x_i)_{i \in [n]} \in \mathbb{R}_+^{n+1}$ and it is the result of a graph process where two subsets V_1 and V_2 of $[n]$ are joined at rate $\sum_{i \in V_1} x_i \sum_{i \in V_2} x_i$, independently from what happens on the rest of the graph. These graphs are inhomogeneous random graphs, defined in Definition 2.2.1. In [14], the authors exhibited a family of distributions toward which scaled inhomogeneous random graphs converge, a result that generalizes the previous partial results of [1] and [11]. Moreover, in [9] and [10], the authors showed that, under certain conditions, percolated supercritical configuration models converge toward some of the distributions described in [14]. This fact must be put in parallel with the fact that a percolated configuration model can be coupled with a multiplicative graph with asymptotically no difference [17]. This model, the percolated supercritical configuration model is the main example of a graph that is *almost* multiplicative, in the sense that it is asymptotically multiplicative. As the multiplicative behaviour is in essence, an independence property, we can expect other models presenting this almost multiplicative behaviour. Moreover, as we saw, the fact that these graphs are all constructed from processes on graphs, we can also expect that these almost multiplicative graphs converge as graph processes.

In this framework, a natural question emerges:

Guiding question 1. *Does the family of random metric spaces described in [14] and [15] encompasses all possible marginals of limits of almost multiplicative graph process ?*

For now we can expand a bit further on what would intuitively be a good candidate for Guiding question 1: of course there has to be the multiplicative behaviour at the core of the process. Such a process will then be called a multiplicative coalescent. Also, as the critical window described in (2.3) is infinite in size ($\lambda \in (-\infty, +\infty)$), the process must be eternal both in the past and in the future. As we will see, the path toward an answer for Guiding question 1 is actually a quest to find all *eternal* multiplicative coalescents.

Note that depending on the context, we might use the denomination "multiplicative coalescent" in a pretty generic way, and we will precise what kind of multiplicative coalescent we are dealing with when necessary.

2.2 Aldous' multiplicative coalescent

2.2.1 Preliminary remarks and definitions

To start investigating the guiding question 1, we take a look at Aldous' multiplicative coalescent [5]. In his work, Aldous ignores the adjacency structure of the graph and only keeps track of the sizes of the different components. This point of view is more restricted than what we ultimately would like to have, but the method he used will be our general framework to answer Guiding question 1. We are looking for a limit of the component size process along the critical window, where components are already formed. As a consequence, we want to be able to talk about the evolution of the process since an initial time that is not only 0, so we must allow the process to start from an initial sequence of component weights $x = (x_i)_{i \in \mathbb{N}}$ that are not all equal. For an easier time manipulating weight sequences, we always suppose that $\forall i < j, x_i \geq x_j$.

Definition 2.2.1 (Aldous' multiplicative coalescent). Let $x = (x_i)_{i \in \mathbb{N}}$ be a sequence of numbers in \mathbb{R}^+ . We define a process that takes values in decreasing sequences of \mathbb{R}^+ as follow: let $g\text{Coal}(x, q)$ be a graph on \mathbb{N} where for $i, j \in \mathbb{N}$ such that $i < j$,

$$\{i, j\} \text{ is an edge of } g\text{Coal}(x, q) \text{ if and only if } E_{i,j} \leq q$$

where $(E_{i,j})$ is a sequence of independent exponential random variables such that:

$$\forall i, j \in \mathbb{N}, \mathbb{E}(E_{i,j}) = \frac{1}{x_i x_j}.$$

This model is known as rank-1 inhomogeneous random graph, or Norros-Reittu model. Then, the multiplicative coalescent $m\text{Coal}(x, q)$ starting from x at time q is the ordered sequence of size of components of $g\text{Coal}(x, q)$. Moreover, let the set of edges of $g\text{Coal}(x, q)$ be denoted by $E(x, q)$.

Note that in Definition 2.2.1 we do not specify any constraint on the initial weight sequence, because as long as one accepts the existence of components of infinite weight, this construction makes perfect sense, thanks to the use of the so-called "graphical" construction, originated from [5]. Let us return to our Guiding question 1 to further examine what initial weight sequence we should take. Recall that we are looking for a limiting object, and as per (2.3), for any $\lambda \in \mathbb{R}$, the size of the largest component of G^n must be of the same order but there also must coexist components comparable in size to the largest one. This constraint means that our limiting object must almost surely preserve multiple large components and not allow for infinite components. Hence, one might be tempted to restrict x to be such that $\sum_i x_i < +\infty$ as it would in fact guarantee that components stay finite in size. However, if one focuses on the first weight x_1 , we remark that with probability $(1 - e^{-q x_1 \sum_{i \geq 2} x_i})$, all components of $g\text{Coal}(x, q)$ have merged at time q . Hence, we have to constraint $\sigma_1(x) = \sum_i x_i$ to be infinite and find another way to ensure that component sizes stay finite.

2.2.2 Analysis of Aldous' multiplicative coalescent

As a way to deal, among other things, with the problem of infinite components, Aldous introduces the following spaces:

Definition 2.2.2 (Spaces of weight sequences). For all $r > 0$, x a decreasing sequence of \mathbb{R}^+ , let

$$\sigma_r(x) = \sum_i x_i^r = \|x\|_r^r$$

and define

$$l_{\searrow}^r = \{(x_i)_{i \in \mathbb{N}} : 0 \leq x_{i+1} \leq x_i, \sigma_r(x) < +\infty\}.$$

In particular if $r = 2$ we equip this space with

$$\forall x, y \in l_{\searrow}^2, d_2(x, y) = \sqrt{\sum_i (x_i - y_i)^2}.$$

Note that when x is obvious from the context, we will often write σ_r for $\sigma_r(x)$.

Let us see how these spaces play a role.

Definition 2.2.3. Let $x \in l_{\searrow}^2$, $x(q) = m\text{Coal}(x, q)$ we denote by

$$S(x, q) = \sigma_2(x(q)).$$

Then we can state maybe the most fundamental fact about Aldous' multiplicative coalescent:

Lemma 2.2.4 (Lemma 20 of [5]). For $x \in l_{\searrow}^2$, $s > S(x, 0) = \sigma_2$,

$$\mathbb{P}(S(x, q) > s) \leq \frac{qs\sigma_2}{s - \sigma_2}$$

Aldous' theory of the multiplicative coalescent teaches us that l_{\searrow}^2 is the right constraint on x , in particular in the following theorem:

Theorem 2.2.5 (Feller property, Proposition 5 of [5]). *Let $(x^n)_{n \in \mathbb{N}}$ be a sequence of l_{\searrow}^2 and $x^\infty \in l_{\searrow}^2$ be such that*

$$x^n \xrightarrow[n \rightarrow +\infty]{} x^\infty$$

in (l_{\searrow}^2, d^2) . Then

$$\forall q \in \mathbb{R}^+, m\text{Coal}(x^n, q) \xrightarrow[n \rightarrow +\infty]{(d)} m\text{Coal}(x^\infty, q)$$

in (l_{\searrow}^2, d^2) .

In light of theorem 2.2.5, we will denote by sl_2 the space of sequences of l_{\searrow}^2 . Moreover, if $x = (x^n)_{n \in \mathbb{N}} \in l_{\searrow}^2$, we denote by σ_r the sequence $(\sigma_r(x^n))_{n \in \mathbb{N}}$.

2.2.3 The Aldous-Limic problem and the Aldous-Limic hypothesis

We already said that a multiplicative coalescent process suited to describe the limit of critical almost multiplicative random graphs must be such that $m\text{Coal}(x, q)$ has almost surely only finite weights and such that there is always at least two components. Observing that (2.3) is also true for any $\lambda < 0$, we must have that our process must be the result of a coalescence of smaller weights as far as we want in the past. More precisely, what we are looking for is encapsulated in the following definition:

Definition 2.2.6 (Eternal multiplicative coalescent). *We say that a Markov process $(x(t))_{t \in \mathbb{R}}$ in l_{\searrow}^2 is an eternal multiplicative coalescent if*

$$\forall s, t \in \mathbb{R} \text{ s.t. } s \leq t, (x(t), x(s)) \stackrel{(d)}{=} (m\text{Coal}(x(s), t-s), x(s)).$$

We will refer to the problem of finding an eternal multiplicative coalescent the Aldous-Limic problem. Before examining further this problem, especially its well-posedness, let us introduce some notations:

Definition 2.2.7.

$$\mathcal{I} = \{(\alpha, \lambda, c) \in \mathbb{R}_+ \times \mathbb{R} \times l_{\searrow}^3 : \text{either } \alpha > 0 \text{ or } c \notin l_{\searrow}^2\}.$$

$$\mathcal{I}^\circ = \{(\alpha, \lambda, 0) \in \mathcal{I} : \alpha > 0\}$$

$$\mathcal{I}^* = \{(0, \lambda, c) \in \mathcal{I} : c \in l_{\searrow}^3 \setminus l_{\searrow}^2\}$$

Moreover, for $s \in \mathbb{R}$, $(\alpha, \lambda, c) \in \mathcal{I}$, let

$$\mathcal{Z}^{(\alpha, \lambda, c)}(s) = \alpha^{1/2} \mathcal{W}(s) + (\lambda + t)s - \frac{1}{2} \alpha s^2 + \sum_i (c_i 1_{\zeta_i \leq s} - c_j^2 s),$$

with \mathcal{W} a Brownian motion and $(\zeta_i)_{i \in \mathbb{N}}$ a sequence of independent exponential random variables such that for $i \in \mathbb{N}$, $\mathbb{E}(\zeta_i) = 1/c_i$. When the parameters are obvious from the context, we will usually write \mathcal{Z} for $\mathcal{Z}^{(\alpha, \lambda, c)}$. Note that rigorously, \mathcal{Z} is defined as a spectrally positive Lévy process characterized by its Laplace in exponent in [14]. They also show in this article (equation (14.3) of section 4.2), that the former notation makes sense : the infinite sum has to be interpreted as the limit of L^2 semi-martingales. Finally, let

$$\mathcal{B}^{(\alpha, \lambda, c)}(s) = \mathcal{Z}^{(\alpha, \lambda, c)}(s) - \inf_{[0, s]} \mathcal{Z}^{(\alpha, \lambda, c)}.$$

The first natural question that arises is the question of the existence of a solution to the Aldous-Limic problem. The answer can be found in [5] where we have:

Theorem 2.2.8 (Convergence of the component sizes of Erdős–Rényi random graphs, Proposition 4 [5]). *For $(1, \lambda, 0) \in \mathcal{I}^\circ$, we can define $(|e_j|)_{j \in \mathbb{N}}$ the ordered sequence of lengths of excursions of $\mathcal{B}^{(1, \lambda, 0)}$ above 0. For $n \in \mathbb{N}$, let $(|C_j^n|)_{j \in \mathbb{N}}$ be the ordered sequence of component weights of $G^n \sim G(n, \frac{1+\lambda n^{-1/3}}{n})$. Then*

$$(n^{-2/3} |C_j^n|)_{j \in \mathbb{N}} \xrightarrow[n \rightarrow +\infty]{(d)} (|e_j|)_{j \in \mathbb{N}}$$

in (l_{\searrow}^2, d^2) .

We then have, as stated in Corollary 24 of [5], that Theorem 2.2.8 together with Theorem 2.2.5 gives a positive answer to the question of existence of solution to the Aldous-Limic problem :

Let $(t_0, \dots, t_k) \in \mathbb{R}^{k+1}$ be family of real numbers and $x^n \in \text{sl}_2$, and for $t \in \mathbb{R}$, $x^n(t) = \text{mCoal}(x^n, \sigma_2^{-1} + t)$. Consider the sequence of vectors $(x^n(t_i))_{i \in [k], n \in \mathbb{N}}$. For f bounded and measurable, $x \in l_{\searrow}^2$ and $t \in \mathbb{R}$, let $P_t f(x) = \mathbb{E}(f(\text{mCoal}(x, t)))$. Let (f_1, \dots, f_k) be a family of bounded measurable functions. As the process is Markov, we have

$$\mathbb{E}\left(\prod_{i \in [k]} f_i(x^n(t_i))\right) = \mathbb{E}(f_0(x^n(t_0))P_{t_1-t_0}[f_1 P_{t_2-t_1}[\dots]](x^n(t_0))).$$

Finally, thanks to Theorem 2.2.8,

$$x^n(t_0) \xrightarrow[n \rightarrow \infty]{(d)} x^\infty(t_0)$$

and thanks to Theorem 2.2.5, $f_0(\cdot)P_{t_1-t_0}[f_1 P_{t_2-t_1}[\dots]](\cdot)$ is bounded continuous. As a consequence,

$$\mathbb{E}(f_0(x^n(t_0))P_{t_1-t_0}[f_1 P_{t_2-t_1}[\dots]](x^n(t_0))) \xrightarrow[n \rightarrow \infty]{} \mathbb{E}(f_0(x^\infty(t_0))P_{t_1-t_0}[f_1 P_{t_2-t_1}[\dots]](x^\infty(t_0)))$$

and then,

$$\mathbb{E}\left(\prod_{i \in [k]} f_i(x^n(t_i))\right) \xrightarrow[n \rightarrow \infty]{} \mathbb{E}\left(\prod_{i \in [k]} f_i(x^\infty(t_i))\right).$$

This convergence give us a family of distributions such that for all increasing family (t_0, \dots, t_k) , $(x^\infty(t_0), \dots, x^\infty(t_k))$ is compatible, Markov and such that

$$(\text{mCoal}(x^\infty(t_0), t_1 - t_0), x^\infty(t_0)) \stackrel{(d)}{=} (x^\infty(t_1), x^\infty(t_0)).$$

This, Kolmogorov's extension theorem gives us a process that is a solution to the Aldous-Limic problem. An alternative construction based on a single Brownian motion for all $t \in \mathbb{R}$ ca be found in [16].

Naturally one would then want to ask about the uniqueness of such process. However, as such, the problem is obviously ill-posed because, other than the trivial solution always equal to $x = (1, 0, \dots) \in l_{\searrow}^2$, if we have an eternal multiplicative coalescent $(x(t))_{t \in \mathbb{R}}$, we can shift it in time or multiply it by a constant and it will stay an eternal multiplicative coalescent. Moreover any mixture of eternal multiplicative coalescents is itself an eternal multiplicative coalescent. This makes the set of eternal multiplicative coalescents convex. More precisely, this set is a simplex [18], i.e. a convex set for which every element can be written in a unique way as a convex combination of extreme elements. Moreover, extreme distributions of eternal Markov process are exactly those for which the asymptotic sigma field in $-\infty$ is trivial [18]. We can therefore refine the definition of an Aldous-Limic problem by adding that the distribution must be extreme. Surprisingly, we do not account for all the eternal multiplicative coalescents by considering the convex hull of the solutions considered at the beginning of this paragraph (trivial solutions, standard multiplicative coalescent, its time shifts etc.). In [6], Aldous and Limic find all extreme eternal multiplicative coalescents, let us present their result in the framework given by the Guiding question 1. First we define, for $x \in l_{\searrow}^2$,

$$\Gamma(x, q) = (\sigma_2, (\frac{x_j}{\sigma_2})_{j \in \mathbb{N}}, \frac{\sigma_3}{\sigma_2}, q - \frac{1}{\sigma_2}).$$

This allows us to state a clean definition of what will be a well-posed Aldous-Limic problem:

Definition 2.2.9 (Aldous-Limic problem with parameter $(\alpha, \lambda, c) \in \mathcal{I}$). *We say that a Markov process $x = (x(t))_{t \in \mathbb{R}}$ in state space l_{\searrow}^2 is a solution to the Aldous-Limic problem if its distribution is extreme and*

$$\left\{ \begin{array}{l} \forall s, t \in \mathbb{R}, s \leq t, (x(s), x(t)) \stackrel{(d)}{=} (x(s), \text{mCoal}(x(s), t - s)) \quad (\text{it is an eternal multiplicative coalescent}) \\ \lim_{t \rightarrow -\infty} \Gamma(x(t), t) = (0, c, \alpha + \sum_j c_j^3, \lambda) \text{ a.s.} \quad (\text{entrance boundary condition}) \end{array} \right.$$

We can now state the main theorem about the extreme eternal multiplicative coalescents:

Theorem 2.2.10 (Theorem 2 and 3 of [6]). *For any $(\alpha, \lambda, c) \in \mathcal{I}$, the associated Aldous-Limic problem from Definition 2.2.9 is well-posed: First there exists a solution $x = (x(t))_{t \in \mathbb{R}}$. Moreover, for all extreme*

eternal multiplicative coalescent $(x(t))_{t \in \mathbb{R}}$, either x is constant, either there exists $(\alpha, \lambda, c) \in \mathcal{I}$, such that for $t \in \mathbb{R}$, if $(|e_j|)_{j \in \mathbb{N}}$ the ordered sequence of lengths of excursions of $\mathcal{B}^{(\alpha, \lambda+t, c)}$ above 0, then,

$$x(t) \stackrel{(d)}{=} (|e_j|)_{j \in \mathbb{N}}.$$

The distribution of such a multiplicative coalescent will be denoted by $\mathcal{M}(\alpha, \lambda, c)$ and $\mathcal{M}_t(\alpha, \lambda, c)$ will denote the distribution of the process at time t .

Notice that, contrary to the parameters α and λ , that describe the possible scaling and shifting in time of a given eternal multiplicative coalescent, the new parameter c doesn't lend itself to an immediate interpretation. It is in fact the manifestation of two linked phenomena:

- (i) If $c \neq 0$, it means that when $t \rightarrow -\infty$, some component sizes stay comparable to σ_2 , which is not the case for subcritical Erdős–Rényirandom graphs. It is the very first time we see that the multiplicative behaviour may encompass more than graphs very similar to Erdős–Rényirandom graphs.
- (ii) If $c \neq 0$, this means that the process \mathcal{B} can jump during its excursion above 0. As we are for now prohibited from looking inside those component due to the mass-centered point of view, we cannot interpret this fact, but we will see later that these jumps encode a certain kind of inhomogeneity inside the structure of the components.

In light of Theorem 2.2.10, we define:

Hypothesis 1 (Aldous-limic conditions for sequences of weight sequence). *We say that the couple (x, q) satisfies the Aldous-Limic hypothesis with parameters $(\alpha, \lambda, c) \in \mathcal{I}$ if $x = (x^n)_{n \in \mathbb{N}}$ is a sequence in l_{\searrow}^2 , $q = (q_n)_{n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$ and for $n \rightarrow +\infty$,*

$$\left\{ \begin{array}{ll} \sigma_2(x^n) \xrightarrow[n \rightarrow +\infty]{} 0 & \text{(initially subcritical)} \\ \frac{x_j^n}{\sigma_2(x^n)} \xrightarrow[n \rightarrow +\infty]{} c_j & \text{(jumps distribution)} \\ \frac{\sigma_3(x^n)}{\sigma_2(x^n)^3} \xrightarrow[n \rightarrow +\infty]{} \alpha + \sum_i c_i^3 & \text{(diffusion part)} \\ q_n - \frac{1}{\sigma_2(x^n)} \xrightarrow[n \rightarrow +\infty]{} \lambda. & \text{(criticality)} \end{array} \right.$$

In other words,

$$\lim_{n \rightarrow +\infty} \Gamma(x^n, q_n) = (0, c, \alpha + \sum_i c_i^3, \lambda).$$

2.2.4 The method of Aldous and Limic

Considering the success described in the precedent subsection, let us state some key points of the method used by Aldous and Limic to obtain Theorem 2.2.10, as it would be a perfect starting point for studying multiplicative objects with more focus on their internal structure. First, for all extreme eternal multiplicative coalescent, it makes sense to talk about the boundary condition in $-\infty$:

Theorem 2.2.11 ([6]). *Let $(x(t))_{t \in \mathbb{R}}$ be an extreme eternal multiplicative coalescent. Then either $x(t)$ is constant, or else there exists $(\alpha, \lambda, c) \in \mathcal{I}$ such that*

$$\lim_{t \rightarrow -\infty} \Gamma(x(t), -t) = (0, c, \alpha + \sum_j c_j^3, \lambda) \quad a.s.$$

Now remark that $x(t) \stackrel{(d)}{=} \text{mCoal}(x(s), t-s)$. When s goes to infinity, we know that $\Gamma(x(s), q)$ tends to $(0, c, \alpha + \sum_j c_j^3, \lambda)$ almost surely. As a consequence, $x(t)$ is the result of the coalescence of a well-understood initial sequence so we can hope to have a theorem similar to Theorem 2.2.8 for more general parameters and obtain the convergence of $\text{mCoal}(x(s), t-s)$ toward the law $\mathcal{M}_t(\alpha, \lambda, c)$. However recall that Theorem 2.2.8 is about discrete finite graphs, and the theorem that Aldous and Limic obtained is similarly limited to finite weight sequences:

Theorem 2.2.12 (Proposition 18 of [6]). *Let $(\alpha, \lambda, c) \in \mathcal{I}$ and $x = (x^n)_{n \in \mathbb{N}} \in sl_2$ and $q = (q_n)_{n \in \mathbb{N}}$ be such that x and q satisfy the Aldous-Limic conditions hypothesis 1 with parameters (α, λ, c) and for all $i > n$, $x_i^n = 0$. Then*

$$m\text{Coal}(x^n, q) \xrightarrow[n \rightarrow +\infty]{d} (|e_j|)_{j \in \mathbb{N}}$$

where $(|e_j|)_{j \in \mathbb{N}}$ are the lengths of the excursions of $\mathcal{B}^{(\alpha, \lambda, c)}$.

To fill the gap between Theorem 2.2.11 and Theorem 2.2.12, we need an argument of truncation: given a sequence $x = (x^n)_{n \in \mathbb{N}}$ of l_{\searrow}^2 , we would like a sequence l_{\searrow}^2 that would be finite and that would well emulate the asymptotic behaviour of x . For this purpose, let state some more definitions:

Definition 2.2.13 (Finite weight sequences). *Let*

$$\overline{l_{\searrow}} = \{(x_i)_{i \in \mathbb{N}} \in l_{\searrow}^2 : \exists n \in \mathbb{N}, x_{n+1} = 0\} \subset l_{\searrow}^2,$$

$$\overline{sl} = \{x = (x^n)_{n \in \mathbb{N}} : \forall n \in \mathbb{N}, x^n \in \overline{l_{\searrow}}\} \subset sl_2.$$

Moreover, if $x \in sl_2$, we say that \overline{x} is a truncation of x if $\overline{x} \in \overline{sl}$ and

$$\forall i, n \in \mathbb{N}, \overline{x}_i^n > 0 \Rightarrow \overline{x}_i^n = x_i^n$$

Moreover, the rest of the truncation $\hat{x} = (\hat{x}^n)_{n \in \mathbb{N}}$ is such that for $n \in \mathbb{N}$, \hat{x}_n is the sequence of weights $x_i^n \overline{x}_i^n = 0$. Note that there is an obvious coupling between $g\text{Coal}(x, q)$, $g\text{Coal}(\overline{x}, q)$ and $g\text{Coal}(\hat{x}, q)$ where $g\text{Coal}(\overline{x}, q)$ and $g\text{Coal}(\hat{x}, q)$ are the subgraphs spanning the vertices with positive weight.

Then the main tool to handle the difference between $m\text{Coal}(x, q)$ and $m\text{Coal}(\overline{x}, q)$ is the following lemma:

Lemma 2.2.14 (Lemma 23 of [5]). *Let $x \in l_{\searrow}^2$, $\overline{x} \in \overline{l_{\searrow}}$ be a truncation of x and let \hat{x} be the rest of the weights. Then*

$$\forall \epsilon, \mathbb{P}(S(x, q) - S(\overline{x}, q) \geq \epsilon | S(\overline{x}, q), S(\hat{x}, q)) \leq \epsilon^{-1} (1 + q(S(\overline{x}, q) + \epsilon))^2 S(\hat{x}, q).$$

This lemma will also be one of our main tools to handle the effect of truncation on coalescence.

2.2.5 Palm measure of the multiplicative coalescent

In [22], Janson and Spencer adopt a point process point of view to describe $\mathcal{M}_0(1, \lambda, 0)$. In particular, they compute the Palm measure of such a process:

Theorem 2.2.15 (Theorem 8.1 of [22]). *Let f be a bounded continuous function from $(0, +\infty)$ to \mathbb{R}_+ with compact support, and g a bounded continuous function from l_{\searrow}^2 to \mathbb{R}_+ . Let $\lambda \in \mathbb{R}$ and $x \sim \mathcal{M}_0(1, \lambda, 0)$, we have*

$$\mathbb{E}\left(\sum_i f(x_i)g(x)\right) = \mathbb{E}\left(\sum_i f(x_i)h(X_i)\right)$$

with, for $u \in \mathbb{R}_+$,

$$h(u) = \mathbb{E}(g(y^u))$$

where y^u is the ordered sequence of the weights $y \sim \mathcal{M}_0(1, \lambda - u, 0)$ to which we have added u .

In words, Theorem 2.2.15 says that the law of $x \sim \mathcal{M}(1, \lambda, 0)$, conditionally to the fact that one of the weights of x is u is the same as taking $y \sim \mathcal{M}(1, \lambda - u, 0)$ and adding the weight u to the sequence. It is especially useful when one wants to compute moments of order 2, because it becomes a computation of a first order moment.

2.3 Spaces of metric spaces

2.3.1 Convergence of metric spaces

This section is essentially technical and sets the context in which we can extend Aldous and Limic's analysis of the multiplicative coalescent in order to take into account the internal structure of the components. More precisely, we will be interested in the metric structures of the components of multiplicative graphs.

Definition 2.3.1 (Space of compact measured metric spaces). *We say that a triplet $m = (\mathcal{M}, d, \mu)$ is a measured metric space if d is a distance on \mathcal{M} and μ is a Borel measure on \mathcal{M} . Moreover, we will write $|m| = \mu(\mathcal{M})$. We denote by \mathbb{M} the space of all compact measured metric spaces and $\tilde{\mathbb{M}}$ the space of all complete, separable, with finite measure, full support, metric spaces. Moreover, for $m = (\mathcal{M}, d, \mu)$ a measured metric space, and for $\alpha \in \mathbb{R}_+$, αm denotes $(\mathcal{M}, \alpha d, \mu)$ and for $\beta \in \mathbb{R}_+$, $\text{scl}(\alpha, \beta)m$ denotes $(\mathcal{M}, \alpha d, \beta \mu)$. We equip \mathbb{M} and $\tilde{\mathbb{M}}$ with the distance d_{GHP} that is defined for m and $m' \in \tilde{\mathbb{M}}$, as follows: we say that $\mathcal{R} \subset m \times m'$ is a correspondence between (m, s, μ) and (m', d', μ') if*

$$\begin{aligned} \forall u \in m, \exists u' \in m' \text{ such that } (u, u') \in \mathcal{R} \\ \forall u' \in m', \exists u \in m \text{ such that } (u, u') \in \mathcal{R}. \end{aligned}$$

We write $\mathcal{C}(m, m')$ for the set of all measurable correspondences between m and m' . We can define the distortion of a correspondence \mathcal{R} as follows:

$$\text{dis}(\mathcal{R}) = \inf\{\epsilon > 0 : \forall (u, u'), (v, v') \in \mathcal{R}, |d(u, v) - d'(u', v')| \leq \epsilon\}.$$

The distortion gives a good idea of the difference between the two metrics but we also would like to compare the two measures. Let $\Pi(m, m')$ be the set of finite Borel measures on $m \times m'$. For $\pi \in \Pi(m, m')$, we denote by (π_1, π_2) the first and second marginal of π . Then let

$$D(\pi; \mu, \mu') = \|\pi_1 - \mu\| + \|\pi_2 - \mu'\|$$

with $\|\cdot\|$ the total variation of a measure. We then set:

$$d_{GHP}(m, m') = \inf_{\mathcal{R} \in \mathcal{C}(m, m'), \pi \in \Pi(m, m')} \{D(\pi; \mu, \mu') \wedge \frac{1}{2} \text{dis}(\mathcal{R}) \wedge \pi(\mathcal{R}^c)\}.$$

Of course, as such d_{GHP} is not a distance on \mathbb{M} since it does not verify the identity of indiscernibles. To remedy this, we consider \mathbb{M}' the set of equivalence classes in \mathbb{M} for the following equivalence relation:

$$m \sim m' \Leftrightarrow d_{GHP}(m, m') = 0,$$

and we equipped \mathbb{M}' with the distance issued from d_{GHP} . In the rest of this work, we will abuse the notations by using \mathbb{M} in place of \mathbb{M}' .

2.3.2 A topology on sequences of metric spaces inherited from the Gromov-Hausdorff-Prokhorov topology

We want to deal with spaces that consist in an infinite amount of compact connected component, hence the following definition:

Definition 2.3.2 (Space of sequences of compact measured metric spaces). *We denote by $\mathbb{M}_2^{\mathbb{N}}$ the space of sequences $X = (X_j)_{j \in \mathbb{N}}$ of \mathbb{M} such that $\sum_j |X_j| < +\infty$ and such that*

$$\forall i, j \in \mathbb{N}, i < j, |X_i| \geq |X_j|.$$

Remark that we usually view X as a measured metric space itself, for which the different X_j are the components. We write $\text{comp}(X)$ for the set of components (X_j) for $j \in \mathbb{N}$.

Note that l_{\searrow}^2 is embedded in $\mathbb{M}_2^{\mathbb{N}}$ by taking the trivial distance on each $\{i\}$ and the Dirac measure on $\{i\}$ with weight x_i . This allows us to canonically apply to an element of l_{\searrow}^2 a function that takes an element of $\mathbb{M}_2^{\mathbb{N}}$ as an argument.

However, contrary to the case of the process $\text{mCoal}(\cdot, \cdot)$ on l_{\searrow}^2 there is no clear topology that we can put on $\mathbb{M}_2^{\mathbb{N}}$ that will fit all our purposes. There are mainly two kind of topologies that come from d_{GHP} :

Definition 2.3.3. *If $X, X' \in \mathbb{M}_2^{\mathbb{N}}$, we set*

$$\delta_{GHP}^2(X, X') = \left(\sum_j (d_{GHP}(X_j, X'_j)) \right) \wedge 1 \vee \| |X| - |X'| \|_2.$$

The associated topology on $\mathbb{M}_2^{\mathbb{N}}$ will be denoted by T_{GHP}^2 .

However, for our purpose, we sometimes prefer the weakest topology as possible, as it would mean less work when translating the Aldous and Limic method. The first topology that comes to mind which would satisfy this while staying under the spirit of d_{GHP} is the product topology. The product topology on $\mathbb{M}_2^{\mathbb{N}}$ inherited from d_{GHP} will be denoted by T_{GHP} .

2.3.3 Weakening d_{GHP} : the Gromov-weak topology

Recently, there has been a surge in the use of the Gromov-weak topology to handle convergence of random graphs. It is defined as follows:

Definition 2.3.4 (Gromov-weak topology). *Let (X, d, μ) be a measured metric space such that $|X| < +\infty$. For $u = (u_1, \dots, u_l) \in X^l$, we define $D(u) = (d(u_i, u_j))_{i,j \in [l]}$. We say that $\Phi : \mathbb{M} \rightarrow \mathbb{R}$ is a polynomial if there exists $\phi : \mathbb{R}^{l^2} \rightarrow \mathbb{R}$ continuous and bounded such that*

$$\forall X \in \mathbb{M}_2^{\mathbb{N}}, \Phi(X) = \int_{X^l} \phi(D(x)) d\mu^{\otimes l}(x).$$

We say that a sequence of measured metric spaces of finite measure $(X^n)_{n \in \mathbb{N}}$ converges toward a measured metric space of finite measure X^∞ if for all polynomial Φ ,

$$\Phi(X^n) \xrightarrow{n \rightarrow +\infty} \Phi(X^\infty).$$

Note that the Gromov-weak topology on \mathbb{M} can be metrized :

Theorem 2.3.5 (Theorem 5 in [21]). *The Gromov-weak topology on \mathbb{M} can be metrized by the Gromov-Prokhorov metric d_{GP} , defined by: for $X, Y \in \mathbb{M}$,*

$$d_{GP}(X, Y) = \inf_{(\phi_X, \phi_Y, Z)} \inf \left\{ \epsilon > 0 : \mu_X(\phi_X^{-1}(F)) \leq \mu_Y(\phi_Y^{-1}(F^\epsilon)) + \epsilon, \forall F \text{ closed} \right\}$$

where the infimum is taken over all isometric embeddings ϕ_X and ϕ_Y from X and Y , respectively, into some common metric space (Z, d_Z) and

$$F^\epsilon = \{u \in Z : \exists v \in F \text{ s.t. } d_Z(u, v) \leq \epsilon\}.$$

In the course of this work, we will encounter multiples situations where this topology is the best suited for the study of the coalescence. However, what we gain in ease of use, we lose in what information can be retrieved from these convergences. For instance, this convergence doesn't allow to talk about diameters of components. Luckily it is well known how the Gromov-weak topology differs from the Gromov-Hausdorff-Prokhorov topology.

Lemma 2.3.6 (Lower mass bound characterization, Theorem 6.1 in [7]). *Let $(X^n, d_n, \mu_n)_{n \in \mathbb{N}}$ be a sequence in $\tilde{\mathbb{M}}$ and $X \in \tilde{\mathbb{M}}$. The following are equivalent:*

(i) $(X^n)_{n \in \mathbb{N}}$ converges toward X for the Gromov-weak topology and

$$\forall \delta > 0, \liminf_{n \rightarrow +\infty} m_\delta(X^n) > 0 \tag{2.4}$$

where $m_\delta(X^n) = \inf\{\mu_n(\{v : d_n(u, v) \leq \delta\}) : u \in X^n\}$.

(ii) X is compact and $(X^n)_{n \in \mathbb{N}}$ converges toward X for the Gromov-Hausdorff-Prokhorov topology.

Note that the lower-mass bound property works similarly with random objects: in this case we require the tightness of $1/m_\delta(X)$.

Lemma 2.3.7 (Random lower mass bound characterisation). *Let $(X^n)_{n \in \mathbb{N}}$ be a random sequence of elements in $\tilde{\mathbb{M}}$ and X a random element in $\tilde{\mathbb{M}} \cap \tilde{\mathbb{M}}$. The following are equivalent:*

(i) $(X^n)_{n \in \mathbb{N}}$ converges in distribution to X for the Gromov-weak topology and

$$\forall \delta > 0, \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow +\infty} \mathbb{P}(m_\delta(X^n) < \epsilon) = 0. \tag{2.5}$$

(ii) $(X^n)_{n \in \mathbb{N}}$ converges in distribution toward X for the Gromov-Hausdorff-Prokhorov topology.

Proof. See in Appendix A. □

Crucially, as we will find out later, the lower mass bound property (2.5) will be often guaranteed in our setting. As we would also like to extend the Gromov-weak topology to $\mathbb{M}_2^{\mathbb{N}}$, we introduce, for $X \in \mathbb{M}_2^{\mathbb{N}}$

$$m_\delta^k(X) = \inf_{j \leq k} m_\delta(X_j).$$

2.3.4 Real tree, real graph and surplus

The surplus of a connected graph $G = (V, E)$ is equal to $|E| - |V| + 1$ and indicates how much it differs from a tree. A tree is a connected graph without cycles. A tree always has surplus 0. Since we are in the general setting of convergences of graphs, one could possibly want to know what happens to the surplus.

For a given discrete graph G , we can associate to the ordering of its components $(C_j)_{j \in \mathbb{N}}$ the sequence of its surplus $(s_j)_{j \in \mathbb{N}}$. We will see later in the particular setting of critical random graphs that the surpluses of the largest components are tight.

Moreover the limiting components will be real graphs, which means that for any point, one can find a ball such that the restriction of the component to this ball is a real tree: a metric space such that any two points are connected through a unique continuous injective path. We refer to [3] for more background on real graphs. The definition of surplus can be naturally extended to real graphs and so we can legitimately ask if the surplus converges. Remark that any discrete graph G is canonically associated to a unique real graph \mathcal{G} where the edges are replaced by copies of $[0, 1]$. Moreover G , as a metric space, is isometrically embedded in \mathcal{G} . As a consequence, when necessary any discrete graph that appears in this work can be considered as a real graph. Lastly, suppose that $(X^n)_{n \in \mathbb{N}} \in \mathbb{M}_2^{\mathbb{N}}$ is a sequence of real graphs that converges towards a real graph $X^\infty \in \mathbb{M}_2^{\mathbb{N}}$ for T_{GHP} . Moreover, suppose that the surplus of each X_i^n converges, for $i \in \mathbb{N}$, toward the surplus of X_i^∞ . We then say that X^n converges toward X^∞ for the *Gromov-Hausdorff-Prokhorov plus surplus product topology*, denoted by T_{GHPS} .

2.4 Encoding of random metric spaces

As we already hinted at when exposing the results on Aldous' multiplicative coalescent in section 2.1.3, the process Z can encode way more than the mass of the connected components. There are different ways to encode properties of metric spaces into real functions, here we present the approach of Broutin, Duquesne and Wang in [14]. It is a really powerful setting, as it allows them to describe the limiting metric spaces of critical $\text{gCoal}(\cdot, \cdot)$ for any parameters $(\alpha, \lambda, c) \in \mathcal{I}$. The encoding presented in this chapter is a variation of the one used by Aldous and Limic in [6]. However, instead of exploring the nodes in a breadth-first fashion, which means, exploring all the children of a node before moving on, it explores the nodes in depth-first fashion. Both cases are interesting, but the depth-first approach allows to talk about the height function, which gives a precise view of the metric structure of the components. First, let us introduce the discrete processes that correspond to discrete random graphs.

2.4.1 The client-server queue

Let $x \in \overline{\mathbb{L}}_{\downarrow}$ be such that $x_i > 0$ if and only if $i \leq n$. The picture drawn by Broutin, Duquesne and Wang of the construction of a discrete multiplicative random graph $\text{gCoal}(x, q)$ is as follows: let each node $i \in [n]$ of the graph be a client of a server that arrives at an exponential random time ξ_i such that

$$\mathbb{E}(\xi_i) = \frac{\sigma_1}{x_i}.$$

If the server is busy, a LIFO (last in first out) policy applies at time ξ_i , the server interrupts the services of any previous client to serve the client i during a time x_i . Let us introduce our first two processes:

Definition 2.4.1 (Server load). *We define the algebraic load of the server at time t :*

$$Z(t) = -\frac{t}{\sigma_2} + \sum_{i \in [n]} \frac{x_i}{\sigma_2} \mathbf{1}_{\xi_i \leq t}$$

and the load of the server

$$B(t) = Z(t) - \inf_{[0, t]} Z.$$

The evolution of the situation yields a tree representation where when a client i arrives, if a client j is being served, then j is an ancestor of i . More precisely:

Definition 2.4.2 (Height process). *Let Z be an algebraic load, we define*

$$\alpha_i(t) = \begin{cases} 1 & \text{if } \xi_i \leq t, \text{ and } \inf_{[\xi_i, t]} Z > Z(\xi_i^-) \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, we define

$$H(t) = \sigma_2 \sum_i \alpha_i(t).$$

Finally, we define the forest $T(x, q)$ to be the graph on $[n]$ such that i is a descendant of j if and only if $\alpha_i(\xi_j) = 1$.

Let us say a few words about this description. When, for some $t \in \mathbb{R}$ and $i \leq n$, $\alpha_i(t) = 1$, this means that the client that arrived at time ξ_i has not completely be served by time t . In the tree representation, we say that i is an ancestor of the client serve at time t . As a consequence, $H(t)$ is the scaled count of the ancestors of the client served at time t , so it is the scaled height process associated with $T(x, q)$. Of course $\text{gCoal}(x, q)$ is not a tree so there needs to be some additional construction to obtain $\text{gCoal}(x, q)$ from the tree encoded by H .

2.4.2 The discrete multiplicative graph

Here we expose the method that allows one to reconstruct $\text{gCoal}(x, q)$ from B :

Definition 2.4.3 (Pinching points). *Let B be server load for $x \in \overline{l_{\downarrow}}$, let \mathcal{P} be, conditionally to B , a Poisson point process on $[0, +\infty)^2$ with intensity $1_{\{0 < s < B(t)\}} ds dt$. Consider the metric space $G(x, q)$ on $[n]$ constructed from $T(x, q)$ as follows: add, for each $(t, s) \in \mathcal{P}$, an edge between the node served at time t and the node served at time $\inf\{u \leq t : \inf_{[u, t]} B > s\}$.*

Then we have:

Theorem 2.4.4 (Theorem 2.1 in [14]). *Let $x \in \overline{l_{\downarrow}}$ and $q \in \mathbb{R}$. Then the graph $G(x, q)$ from definition 2.4.2 has the same distribution as $\text{gCoal}(x, q)$.*

Finally, we said that the classical point of view to see the asymptotic structure of these graphs, is the metric point of view, so we introduce:

Definition 2.4.5 (Discrete multiplicative space). *In the same context as Definition 2.4.3, we define $G^n \in \mathbb{M}_2^{\mathbb{N}}$ to be a metric space such that the set of points is $[n]$, equipped with the graph distance on $G(x, q)$ multiplied by σ_2 and with the counting measure on $[n]$ weighted by x . Moreover, for each j , we associate l_j the instant of arrival ξ_i of the first client i of the component G_j^n , r_j the last time at which a vertex i of the component G_j^n is served, and s_j^n the surplus of the underlying graph of G_j^n .*

2.4.3 Continuous process

Here we describe how the limiting process from Aldous and Limic in [6] can be used to define an element of $\mathbb{M}_2^{\mathbb{N}}$. This is done in a similar way to its discrete counterpart. First, let us introduce the continuous equivalent of the server load:

Definition 2.4.6 (Continuous height process). *For $(\alpha, \lambda, c) \in \mathcal{I}$, let $\mathcal{Z}^{(\alpha, \lambda, c)}$ be as in Definition 2.2.7. Let $\mathcal{H}^{(\alpha, \lambda, c)}$ be the height process associated to $\mathcal{Z}^{(\alpha, \lambda, c)}$ as defined in Proposition 2.6 in [15].*

From Proposition 14 in Aldous and Limic [6], we know that the excursions of \mathcal{B} can be listed by decreasing order of size $(r_j - l_j)_{j \in \mathbb{N}}$. Let us show how to encode real trees from real functions. This topic originated in the nineties with the work of Aldous [4], is now quite standard. Given a continuous function $h : [l, r] \rightarrow \mathbb{R}_+$ such that $h(r) = 0$, $h(s) \geq 0$ for all $s \in (l, r)$, we can define a tree [20] by setting

$$\forall u, v \in [l, r], d(u, v) = h(u) + h(v) - 2 \inf_{[u \wedge v, u \vee v]} h$$

and then taking $[l, r] / \sim$ where $u \sim v \Leftrightarrow d(u, v) = 0$. This allows us to define a forest $(\mathcal{T}_j)_{j \in \mathbb{N}}$ as follows:

$$\forall j \in \mathbb{N}, \forall u, v \in [l_j, r_j], d_j^{\mathcal{T}}(u, v) = \mathcal{H}(u) + \mathcal{H}(v) - 2 \inf_{[u \wedge v, u \vee v]} \mathcal{H},$$

$$\forall j \in \mathbb{N}, \forall u, v \in [l_j, r_j], u \sim_j^{\mathcal{T}} v \text{ if and only if } d_j^{\mathcal{T}}(u, v) = 0$$

$$\forall j \in \mathbb{N}, \mathcal{T}_j = [l_j, r_j] / \sim_j^{\mathcal{T}}.$$

Definition 2.4.7 (Continuous multiplicative space). *Let $(\alpha, \lambda, c) \in \mathcal{I}$, $\gamma \in \mathbb{R}_+$ and $t \in \mathbb{R}_+$, we define the distribution $\mathcal{E}_0(\alpha, \lambda, c, \gamma)$ of a multiplicative space at time t as follows. Let $\mathcal{Z}^{(\alpha, \lambda, c)}$, $\mathcal{B}^{(\alpha, \lambda, c)}$, $\mathcal{H}^{(\alpha, \lambda, c)}$ be defined as in Definition 2.4.6 and let \mathcal{P} be a Poisson point process on \mathbb{R}_+^2 with intensity $1_{v \leq \mathcal{B}^{(\alpha, \lambda, c)}(u)} du dv$, conditionally on $\mathcal{B}^{(\alpha, \lambda, c)}$. For any $j \in \mathbb{N}$, we define a distance on $[l_j, r_j]$ as follows:*

$$\forall u, v \in [l_j, r_j], d_j(u, v) = \inf_{k \in \mathbb{N}} \left\{ \gamma \sum_{i=1}^k d_j^{\mathcal{T}}(u_i, v_i) : \begin{cases} (u_i)_{i \in [k]} \in [l_j, r_j]^k, (v_i)_{i \in [k]} \in [l_j, r_j]^k, \\ \forall i \in [2, k], (u_i, v_{i-1}) \in \mathcal{P} \text{ or } (v_{i-1}, u_i) \in \mathcal{P}, \\ u_1 = u, v_k = v \end{cases} \right\}.$$

Then if

$$\forall j \in \mathbb{N}, \forall u, v \in [l_j, r_j], u \sim_j v \Leftrightarrow d_j(u, v) = 0,$$

we define

$$\mathcal{X}_j = [l_j, r_j] / \sim_j.$$

Finally, we say that $\mathcal{X} = (\mathcal{X}_j)_{j \in \mathbb{N}}$ is a multiplicative space and we define $\mathcal{E}_0(\alpha, \lambda, c, \gamma)$ as the distribution of \mathcal{X} . Moreover, for $t \in \mathbb{R}$, we define $\mathcal{E}_t(\alpha, \lambda, c, \gamma) = \mathcal{E}_0(\alpha, \lambda - t, c, \gamma)$ and we denote by $p_j^{\mathcal{X}}$ the projection from $[l_j, r_j]$ onto \mathcal{X}_j .

By construction, a multiplicative space is a sequence of real graphs, as a consequence one can associate, to each component \mathcal{X}_j , a surplus s_j .

2.4.4 Inhomogeneity inside a component

For purposes that will be clear later, we often have to quantify the amount of inhomogeneity inside a component of a multiplicative space. As a consequence, we define the following tool:

Definition 2.4.8 (Amount of inhomogeneity in a component). *Let $\mathcal{X} \sim \mathcal{E}_t(\alpha, \lambda, c, \gamma)$ be constructed from $\mathcal{Z}^{(\alpha, \lambda - t, c, \gamma)}$. Let $(\zeta_i)_{i \in \mathbb{N}}$ be the exponential random time used in*

$$\mathcal{Z}^{(\alpha, \lambda - t, c, \gamma)}(s) = \alpha^{1/2} \mathcal{W}(s) + (\lambda - t)s - \frac{1}{2} \alpha s^2 + \sum_i (c_i 1_{\zeta_i \leq s} - c_i^2 s).$$

For any $j > 0$, we write

$$\Delta_j = \{i : \zeta_i \in (l_j, r_j]\},$$

$$\forall r > 0, \delta_r^{(j)} = \sum_{i \in \Delta_j} c_i^r.$$

and $c^{(j)}$ the sequence such that

$$c_i^{(j)} = \begin{cases} c_i & \text{if } i \in \Delta_j \\ 0 & \text{otherwise.} \end{cases}.$$

2.4.5 Convergence of metric spaces

In the last decade, efforts have been made to obtain from this encoding more than just the convergence of the weights of the components: the convergence of the whole metric structure. In [1], Addario-Berry, Broutin and Goldschmidt were able to show this convergence when all the initial weights are equal, which corresponds to a parameter $(\alpha, \lambda, 0)$. Later in [3], Addario-Berry, Broutin, Goldschmidt and Miermont further refined this convergence to take into account the distribution of the mass, obtaining a convergence in the Gromov-Hausdorff-Prokhorov topology. On the other hand, in [11], Bhamidi, van der Hofstad and Sen show the convergence with a parameter $(0, \lambda, c)$ where $c \in l_{\searrow}^3 \setminus l_{\searrow}^2$. Finally, in [14], Broutin, Duquesne and Wang developed a unifying theory that works with all admissible parameters for the Aldous-Limic condition 1. In fact they have:

Theorem 2.4.9 (Convergence of multiplicative spaces, Theorem 2.5 and 2.8 of [14]). *Let $(x, q) \in \overline{s\ell} \times \mathbb{R}^{\mathbb{N}}$ that satisfies the Aldous-Limic conditions 1 with parameters $(\alpha, \lambda, c) \in \mathcal{I}$. Let Z, H be its associated processes and let $G^n = (G_j^n) \in \mathbb{M}_2^{\mathbb{N}}$ be a discrete multiplicative space from Definition 2.4.5 and as in this definition, let $(s_j^n)_{j \in \mathbb{N}}$, $(l_j^n)_{j \in \mathbb{N}}$, $(r_j^n)_{j \in \mathbb{N}}$ be the surpluses, the starting and ending points of the excursions.*

Let \mathcal{Z}, \mathcal{H} be the exploration process with parameter (α, λ, c) of definitions Definition 2.4.7 and $(s_j)_{j \in \mathbb{N}}, (l_j)_{j \in \mathbb{N}}$ and $(r_j)_{j \in \mathbb{N}}$ be the surpluses, the starting and ending points of the excursions. Then

$$(Z, H, (s_j^n)_{j \in \mathbb{N}}, (l_j^n)_{j \in \mathbb{N}}, (r_j^n)_{j \in \mathbb{N}}) \xrightarrow[n \rightarrow +\infty]{(d)} (Z, \mathcal{H}, (s_j)_{j \in \mathbb{N}}, (l_j)_{j \in \mathbb{N}}, (r_j)_{j \in \mathbb{N}}).$$

The convergence of the processes is relative to the Skorokhod topology and the convergence of the sequences to the product topology. Moreover,

$$G^n \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_0(\alpha, \lambda, c, 1)$$

for T_{GHPS} .

2.5 Metric extension of the Coalescence

2.5.1 Different possible extensions of the multiplicative coalescent

Recall that according to our guiding question 1, we are looking for an object that could describe random graphs with multiplicative behaviour in their critical window. The mass-focused approach of Aldous and Limic was a first successful attempt at this and from section 2.4.5 we can guess that a good point of view to gain more information on what happens inside the connected components is a metric space point of view. Hence, the goal of this subsection is to define an extension of the multiplicative coalescent for metric spaces. This construction originated in Rossignol [28]. First, we state the following useful definition:

Definition 2.5.1 (Gluing of spaces). *Let $X = (X, d, \mu) \in \mathbb{M}_2^{\mathbb{N}}$, let d' be a semi-metric on X , which means that it verifies all the properties of a usual metric except the fact that for $u, v \in X$, $d'(u, v) = 0$ implies $u = v$. Let \mathcal{R} be the equivalence relation defined by*

$$u\mathcal{R}v \Leftrightarrow d'(u, v) = 0$$

and let $p_{\mathcal{R}} : X \rightarrow X/\mathcal{R}$ be the projection under this equivalence relation. We define $Glue(X, d')$ as the completion of $(X/\mathcal{R}, d' \circ p^{-1}, \mu \circ p^{-1})$.

Remark that it is showed in [28] that the components of $Coal(X, q)$ are in $\tilde{\mathbb{M}}$ but it is not obvious that they are in \mathbb{M} . Now we can define our main object of interest, the complete coalescence of measured metric spaces. The word "complete" comes from the fact that there is a lot of ways one could extend Aldous' multiplicative coalescent to measured metric spaces but this maybe the most natural one and the more inline with our guiding question 1.

Definition 2.5.2 (Complete coalescence). *Let $X = (X, d, \mu) \in \mathbb{M}_2^{\mathbb{N}}$, let \mathcal{P} be a Poisson point process on $X^2 \times [0, +\infty)$ of intensity $\frac{1}{2}\mu^2 \times Leb_{\mathbb{R}_+}$ and for all $t \geq 0$, let*

$$\mathcal{P}_t = \{(u, v) \in X^2 : \exists s \leq t, (u, v, s) \in \mathcal{P}\}.$$

For all $t > 0$, we define the semi metric d_t^{Coal} on X as:

$$\forall u, v \in X, d_t^{Coal}(u, v) = \inf \left\{ \sum_{i=1}^k d(u_i, v_i) : u_1 = u, v_k = v, \forall i \leq k-1, (v_i, u_{i+1}) \in \mathcal{P}_t \right\}.$$

We define

$$Coal(X, t) = Glue(X, d_t^{Coal}).$$

Notice that the gluing points added inside the initial components does not modify the weight sequence. Hence we could also not add them while keeping a process described by the multiplicative coalescent at the component size level:

Definition 2.5.3 (Simple coalescence). *Let $X = (X, d, \mu) \in \mathbb{M}_2^{\mathbb{N}}$, let \mathcal{P} be a Poisson point process on $X^2 \times [0, +\infty)$ of intensity $\frac{1}{2}\mu^2 \times Leb_{\mathbb{R}_+}$ and for all $t \geq 0$, let*

$$\mathcal{SP}_t = \left\{ (u, v) \in X^2 : \left\{ \begin{array}{l} \exists i, j \in \mathbb{N}, i \neq j \text{ s.t. } (u, v) \in X_i \times X_j \text{ and } \mathcal{P} \cap X_i \times X_j \times [0, t) = \emptyset \\ \exists s \leq t, (u, v, s) \in \mathcal{P} \end{array} \right. \right\}.$$

For all $a \geq 0$, $t > 0$, we define the semi metric $d_t^{SCoal_a}$ on X as:

$$\forall u, v \in X, d_t^{SCoal_a}(u, v) = \inf \left\{ ka + \sum_{i=1}^k d(u_i, v_i) : u_1 = u, v_k = v, \forall i \leq k-1, (v_i, u_{i+1}) \in \mathcal{SP}_t \right\}.$$

We denote

$$SCoal_a(X, t) = Glue(X, d_t^{SCoal_a}).$$

Moreover, we denote $SCoal(X, t) = SCoal_0(X, t)$.

Remark that, starting from the same $X \in \mathbb{M}_2^{\mathbb{N}}$, there is an obvious coupling between $Coal(X, \cdot)$ and $SCoal(X, \cdot)$ that is taking the same Poisson point process \mathcal{P} to define each process.

Finally, let us introduce the following natural notion

Definition 2.5.4 (Eternal metric coalescent). *We say that $(X(t))_{t \in \mathbb{R}}$ is an eternal metric coalescent if it is a Markov process in \mathbb{M} such that for all s and $t \in \mathbb{R}$ such that $s \leq t$,*

$$(Coal(X(s), t-s), X(s)) \stackrel{(d)}{=} (X(t), X(s)).$$

2.5.2 Can we implement the method of Aldous and Limic from Section 2.2.4 to the coalescence of metric spaces ?

The first question is the question of existence of eternal multiplicative coalescent. Recall that Aldous' answer from to this question in regard to the coalescence of masses relies on two results, first a result of convergence of discrete multiplicative random graphs, that would be Theorem 2.4.9 in our case, and a Feller property. In fact we have:

Theorem 2.5.5 (Feller random property for the complete coalescence, consequence of Theorem 3.1 of [28]). *Let $X = (X^n)_{n \in \mathbb{N}}$ be a sequence of $\mathbb{M}_2^{\mathbb{N}}$, and for $t \in \mathbb{R}_+$, $\mathcal{X}^n(t) = Coal(X^n, t)$. Suppose that there exists $X^\infty \in \mathbb{M}_2^{\mathbb{N}}$ and $t \in \mathbb{R}_+$ such that*

$$X^n \xrightarrow[n \rightarrow +\infty]{(d)} X^\infty,$$

for T_{GHP}^2 ,

$$\forall t \in \mathbb{R}, \forall \delta > 0, \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(m_\delta^k(\mathcal{X}^n(t)) < \epsilon) = 0.$$

and

$$\forall t \in \mathbb{R}_+, \mathbb{P}(\exists i, j \in \mathbb{N}, i \neq j \text{ s.t. } |Coal(X^\infty, t)_i| = |Coal(X^\infty, t)_j|) = 0.$$

Then for all $t \in \mathbb{R}_+$, $Coal(X^\infty, t)$ is a.s. in $\mathbb{M}_2^{\mathbb{N}}$ and for any family $(t_0, \dots, t_k) \in \mathbb{R}_+^{k+1}$,

$$(Coal(X^n, t_i))_{i \in [k]} \xrightarrow[n \rightarrow \infty]{(d)} (Coal(X^\infty, t_i))_{i \in [k]}$$

for T_{GHP}^2 .

Proof. See Appendix B. □

In the same spirit as the discussion following Theorem 2.2.8, Theorem 2.5.5 gives, combined with Theorem 2.4.9, the existence of a limit eternal metric coalescent such that its marginals are given by $\mathcal{E}_t(\alpha, \lambda, c, \gamma)_{t \in \mathbb{R}}$ (see Appendix B).

The distribution of such a process will be denoted by $\mathcal{E}(\alpha, \lambda, c, \gamma)$. Remark that we do not know if those eternal metric coalescent are extreme. Let us review what steps should be taken to carry out the method from Aldous and Limic with metric coalescence. Before that, we highlight some general Markov theory: let f be bounded measurable for T_{GHP} , and $X = (X(t))_{t \in \mathbb{R}}$ an eternal metric coalescent. For $t \in \mathbb{R}$, let \mathcal{F}_t be the sigma field generated by $(X(t-s))_{s \geq 0}$. Because X is a Markov process, for $t \in \mathbb{R}$ and $s > 0$,

$$\mathbb{E}(f(Coal(X(t-s), s)) | X(t-s)) = \mathbb{E}(f(Coal(X(t-s), s)) | \mathcal{F}_{t-s}).$$

Moreover as $(Coal(X(t-s), s), X(s)) \stackrel{(d)}{=} (X(t), X(s))$, by Lévy's downward theorem, we have

$$\mathbb{E}(f(X(t)) | X(t-s)) \xrightarrow[s \rightarrow \infty]{} \mathbb{E}(X(t) | \mathcal{F}_{-\infty}). \quad (2.6)$$

On the other hand, by [18], X is extreme if and only if for any bounded measurable f ,

$$\mathbb{E}(f(X(t)) | \mathcal{F}_{-\infty}) = \mathbb{E}(f(X(t))). \quad (2.7)$$

We would like to prove the two following assertions :

Assertion 1. *If $X \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$, for $(\alpha, \lambda, c, \gamma) \in \mathcal{I} \times \mathbb{R}_+$, X is extreme.*

Possible proof layout. (i) There exists a space H of countable dimension, a function $\Gamma' : \mathbb{M}_2^{\mathbb{N}} \times \mathbb{R}_+ \rightarrow H$ and a function $P : \mathcal{I} \times \mathbb{R}_+ \rightarrow H$ such that

$$\Gamma'(X(t), -t) \xrightarrow[t \rightarrow -\infty]{\mathbb{P}} P(\alpha, \lambda, c, \gamma),$$

which is an equivalent of the Aldous-Limic conditions Hypothesis 1.

(ii) If $X = (X^n)_{n \in \mathbb{N}}$ is a sequence in $\mathbb{M}_2^{\mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}}$ is a sequence in \mathbb{R}_+ such that

$$\Gamma'(X^n, q_n) \xrightarrow[n \rightarrow \infty]{} P(\alpha, \lambda, c, \gamma),$$

then

$$\text{Coal}(X^n, q_n) \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{E}_0(\alpha, \lambda, c, \gamma).$$

Indeed, we would have

$$\mathbb{E}(f(X(t)) | X(t-s)) \xrightarrow[s \rightarrow \infty]{\mathbb{P}} \mathbb{E}(f(X(t))),$$

so from (2.6),

$$\mathbb{E}(f(X(t))) = \mathbb{E}(f(X(t)) | \mathcal{F}_{-\infty})$$

so from (2.7), X is extreme. □

Assertion 2. *If X is an eternal extreme metric coalescent, there exists $(\alpha, \lambda, c, \gamma) \in \mathcal{I} \times \mathbb{R}_+$ such that $X \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$.*

Possible proof layout. This assertions would be the consequence of the unproven following assertions :

(i) There exists a space H of countable dimension, a function $\Gamma' : \mathbb{M}_2^{\mathbb{N}} \times \mathbb{R}_+ \rightarrow H$ and a function $P : \mathcal{I} \times \mathbb{R}_+ \rightarrow H$ such that

$$\Gamma'(X(t), -t) \xrightarrow[t \rightarrow -\infty]{a.s.} P(\alpha, \lambda, c, \gamma).$$

(ii) If $X = (X^n)_{n \in \mathbb{N}}$ is a sequence in $\mathbb{M}_2^{\mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}}$ is a sequence in \mathbb{R}_+ such that

$$\Gamma'(X^n, q_n) \xrightarrow[n \rightarrow \infty]{} P(\alpha, \lambda, c, \gamma),$$

then

$$\text{Coal}(X^n, q_n) \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{E}_0(\alpha, \lambda, c, \gamma).$$

Indeed we would have

$$\mathbb{E}(f(X(t)) | X(t-s)) \xrightarrow[s \rightarrow \infty]{\mathbb{P}} \mathbb{E}(f(\mathcal{X}(t))),$$

with $\mathcal{X}(t) \sim \mathcal{E}_t(\alpha, \lambda, c, \gamma)$, and as by Equation (2.6),

$$\mathbb{E}(f(X(t)) | X(t-s)) \xrightarrow[s \rightarrow \infty]{\mathbb{P}} \mathbb{E}(f(X(t)) | \mathcal{F}_{-\infty}),$$

this would gives that

$$\mathbb{E}(f(X(t))) = \mathbb{E}(\mathbb{E}(f(X(t)) | \mathcal{F}_{-\infty})) = \mathbb{E}(f(\mathcal{X}(t)))$$

so $X \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$. □

Remark that, despite being quite opposite in appearance, the hypothetical proof layout of Assertion 1 and Assertion 2 is very similar. In both cases, step (i) is a complex question, that we will leave aside for now. For Assertion 1, the case where $X \sim \mathcal{E}(1, \lambda, 0, 1)$ will be the main focus of Chapter 8. Let us examine what results we can find in the literature that would fulfill the requirement of (ii). We will need the following notation:

Definition 2.5.6. Let $X = (X, d, \mu) \in \mathbb{M}_2^{\mathbb{N}}$, for all $r \in \mathbb{N}$, we define

$$\rho_r = \rho_r(X) = \int_{X^2} d(u, v)^r 1_{d(u, v) < +\infty} \mu(du) \mu(dv).$$

In particular, ρ_1 can also be written as

$$\rho_1 = \sum_i x_i^2 u_i$$

where for all $i \in \mathbb{N}$,

$$x_i = |X_i|$$

and

$$u_i = x_i^{-2} \int_{X_i^2} d_i(u, v) \mu_i(du) \mu_i(dv)$$

is the expectation of the distance between two random points in X_i . Moreover, the diameter of a component X_i will be denoted by D_i . Lastly, if $X \in \overline{\mathbb{M}}^{\mathbb{N}}$, $x_{\min} = x_n$ with n being the last index such that $x_i > 0$. Remark that, as usual, if $X = (X^n)_{n \in \mathbb{N}}$, respectively $X = (X(t))_{t \in \mathbb{R}}$ is a family of elements of $\mathbb{M}_2^{\mathbb{N}}$, the quantity ρ_1 will be understood as the sequence $\rho_1 = (\rho_1^n)_{n \in \mathbb{N}}$, respectively $\rho_1 = (\rho_1(t))_{t \in \mathbb{R}}$, where ρ_1^n is associated with X^n , respectively $\rho_1(t)$ is associated with $X(t)$. The same goes for $u = (u_i)_{i \in \mathbb{N}}$, $x = (x_i)_{i \in \mathbb{N}}$, $D = (D_i)_{i \in \mathbb{N}}$ and x_{\min} .

Theorem 2.5.7 (Convergence of the simple coalescent of homogeneous multiplicative random graphs, Theorem 3.4 of [9]). Let $X = (X^n)_{n \in \mathbb{N}}$ be a sequence of $\overline{\mathbb{M}}^{\mathbb{N}}$, $x = (x^n)_{n \in \mathbb{N}} = (|X_i^n|)_{n \in \mathbb{N}, i \in \mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}}$ such that (x, q) satisfies the Aldous-Limic conditions 1 with parameters $(1, \lambda, 0)$. Moreover, assume the additional weight conditions: there exists $\eta > 0$ and $r > 0$ such that

$$\begin{cases} \frac{x_0}{\sigma_2} = o(\sigma_2^{1/2+\eta}) \\ \frac{\sigma_2^r}{x_{\min}} \rightarrow 0, \end{cases}$$

and the metric conditions:

$$\begin{cases} \frac{\sup_i D_i}{\rho_1 + \sigma_2} = o(\sigma_2^{1/2+\eta}) \\ \frac{x_0 \sigma_2 \sup_i D_i}{\rho_1} \rightarrow 0, \end{cases}$$

Then we have

$$\frac{\sigma_2^2}{\rho_1 + \sigma_2} S\text{Coal}_{\sigma_2}(X, q) \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_0(1, \lambda, 0, 1)$$

for T_{GHP} .

Notice how this result deals with graphs that are strongly homogeneous, in the sense that there is no positive inhomogeneity c that subsists in the limit. There exists a result for other settings but we need to introduce a new definition before:

Definition 2.5.8 (Connected components of the simple coalescence). Let $V \subset \mathbb{N}$ finite, $r > 0$, a sequence and probability mass function $p = (p_i)_{i \in V}$ and $(X_i, \mu_i, d_i)_{i \in V} \in \mathbb{M}^V$ such that

$$\forall i \in V, p_i = \frac{|X_i|}{\sum_{i \in V} |X_j|} > 0.$$

For $q \in \mathbb{R}$, consider the following distribution on connected graphs on V :

$$\mathbb{P}_{\text{con}}^{p, q}(G) \propto \prod_{\{i, j\} \in E(G)} (1 - e^{-q p_i p_j}) \prod_{\{i, j\} \notin E(G)} e^{-q p_i p_j}.$$

Let E be a set of edges generated from this distribution and let $(U_{i, j}, V_{i, j})_{i, j \in V}$ be a family of independent random variables such that:

$$\forall i, j \in V, U_{i, j} \sim \frac{\mu_i}{|X_i|}, V_{i, j} \sim \frac{\mu_j}{|X_j|}.$$

Then, let d_{con} be such that

$$\forall u, v \in X, d_{con}(u, v) = \inf \left\{ ka + \sum_{l=1}^k d(u_l, v_l) : u_1 = u, v_k = v, \forall l \leq k-1, \right. \\ \left. \exists i, j \in V \text{ such that } (u_l, v_l) = (U_{i,j}, V_{i,j}) \text{ and } \{i, j\} \in E \right\}.$$

Then we define

$$SCoal_a^V(X, q) = Glue(X, d_{con}),$$

equipped with the measure $\sum_i \mu_i$.

This definition may seem a bit convoluted but it is an alternative and powerful way of constructing $SCoal(X, t)$, as we can see in the following result:

Proposition 2.5.9 (Proposition 6.1 of [12]). *Let $(V^j)_{j \geq 1}$ be a partition of \mathbb{N} , Let $X \in \overline{\mathbb{M}}^{\mathbb{N}}$ and x be its associated weight sequence, let \mathcal{V} be the event:*

$$\mathcal{V} = \{\forall j \in \mathbb{N}, V^j \text{ is a connected component of } gCoal(x, t)\}.$$

For all $j \in \mathbb{N}$, let $X^j = (X_i)_{i \in V^j}$. Consider \mathcal{X} the element of $\overline{\mathbb{M}}_2^{\mathbb{N}}$ associated with the independent components $(SCoal_a^{V^j}(X^j, q_j))_{j \in \mathbb{N}}$, where $q_j = q(\sum_{i \in V^j} x_i)^2$. Then the law of $SCoal_a(X, t)$, conditionally on the partition \mathcal{V} , is the same as the one of \mathcal{X} .

This theorem means that we can generate $SCoal_a(X, q)$ by first generating a partition of \mathbb{N} via Aldous' multiplicative coalescent, which is well understood, and then generate connected components through Definition 2.5.8. This is why in the literature one can find results only stated for $SCoal_a^V(X, q)$. It will be on us to extend it to $SCoal_a(X, q)$ via Proposition 2.5.9.

Moreover, the same phenomenon occurs in the continuous realm:

Proposition 2.5.10. *Let $(\alpha, \lambda, c) \in \mathcal{I}^* \cup \mathcal{I}^\circ$, $\gamma \in \mathbb{R}_+$, $t \in \mathbb{R}$, let $X \sim \mathcal{E}_t(\alpha, \lambda, c, \gamma)$. Recall the definition of Δ_j , for $j \in \mathbb{N}$ from Definition 2.4.8. There exists a family of distributions denoted by $(\mathcal{E}_t^{con}(x, \beta))_{x, \beta}$ indexed by $x \in \mathbb{R}_+$ and $\beta = (\beta_i)$ a sequence of real numbers that satisfies $\beta \in \ell_1 \setminus \ell_2$ when $(\alpha, \lambda, c) \in \mathcal{I}^*$ and $\beta = 0$ when $(\alpha, \lambda, c) \in \mathcal{I}^\circ$, such that conditionally on $(|X_i|, c^{(i)})_{i \in \mathbb{N}}$, the components X_i for $i \in \mathbb{N}$ are independent and $X_i \sim \mathcal{E}_t^{con}(|X_i|, c^{(i)})$.*

Proof. Theorem 3.3 in [12] and Theorem 1.8 in [11] both give a construction of the limit as described in the proposition, and as these limits are particular cases of the convergence in Theorem 2.4.9, the conditional distribution of X knowing $(|X_i|, c^{(i)})_{i \in \mathbb{N}}$ is indeed given by the construction of [11] when $(\alpha, \lambda, c) \in \mathcal{I}^*$ and of [12] when $(\alpha, \lambda, c) \in \mathcal{I}^\circ$. \square

Theorem 2.5.11 (Convergence of connected components of strongly inhomogeneous multiplicative random graphs, Theorem 5.4 of [10]). *Let $X = (X^m)_{m \in \mathbb{N}}$ be a sequence of $\overline{\mathbb{M}}^{\mathbb{N}}$ and $p = (p_i^m)_{m \in \mathbb{N}, i \in [m]}$ be such that*

$$\forall i \in [m], p_i^m = \frac{|X_i^m|}{\sum_j |X_j^m|} > 0.$$

Let $q = (q^m)_{m \in \mathbb{N}}$ and $\beta = (\beta_i)_{i \in \mathbb{N}} \in \ell_2 \setminus \ell_1$ be such that $\sum_i \beta_i^2 = 1$. Suppose that there exists $x > 0$ such that

$$\begin{cases} \sigma_2(p) \xrightarrow{m \rightarrow +\infty} 0, \\ \frac{p_i}{\sqrt{\sigma_2(p)}} \xrightarrow{m \rightarrow +\infty} \beta_i, \\ q\sqrt{\sigma_2(p)} \xrightarrow{m \rightarrow +\infty} x \end{cases}$$

and

$$\left\{ \frac{\sup_i D_i \sqrt{\sigma_2(p)}}{\sum_i p_i u_i + 1} \xrightarrow{m \rightarrow +\infty} 0. \right.$$

Then there exists a distribution $\mathcal{G}^\infty(\beta, x)$ of random real graphs such that

$$scl \left(\frac{\sqrt{\sigma_2(p)}}{\sum_i p_i u_i + 1}, \frac{1}{\sum_j |X_j^m|} \right) SCoal_1^{[m]}(X, q) \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{G}^\infty(\beta, x).$$

for the Gromov-Hausdorff-Prokhorov topology. Furthermore, for any $c \in l_{\searrow}^2 \setminus l_{\searrow}^1$,

$$\mathcal{E}_t^{\text{con}}(x, c) \stackrel{(d)}{=} \text{scl} \left(\frac{x}{\|c\|_2}, x \right) \mathcal{G}^\infty \left(\frac{c}{\|c\|_2}, x \|c\|_2 \right)$$

As a way to compare Theorem 2.5.7 and Theorem 2.5.11, we give the equivalent result of Theorem 2.5.11 for the components, as it is stated in [9]:

Theorem 2.5.12 (Convergence of connected components of strongly homogeneous multiplicative random graphs, Theorem 6.4 of [9]). *Let $X = (X^m)_{m \in \mathbb{N}}$ be a sequence of $\overline{\mathbb{M}}^{\mathbb{N}}$ and $p = (p_i^m)_{m \in \mathbb{N}, i \in [m]}$ be such that*

$$\forall i \in [m], p_i^m = \frac{|X_i^m|}{\sum_j |X_j^m|} > 0.$$

Let $q = (q^m)_{m \in \mathbb{N}}$. Suppose that there exists $x > 0$ and $r > 0$ such that

$$\left\{ \begin{array}{l} \sigma_2(p) \xrightarrow{m \rightarrow +\infty} 0, \\ \frac{\max_i p_i}{\sqrt{\sigma_2(p)}} = o((\sigma_2(p))^{\frac{1+n}{4}}), \\ q \sqrt{\sigma_2(p)} \xrightarrow{m \rightarrow +\infty} x \\ \frac{(\sigma_2(p))^r}{\min_i p_i} \xrightarrow{m \rightarrow +\infty} 0 \end{array} \right.$$

and

$$\left\{ \frac{\sup_i D_i \sqrt{\sigma_2(p)}}{\sum_i p_i u_i + 1} = o((\sigma_2(p))^{\frac{1+n}{4}}). \right.$$

Then there exists a distribution $\mathcal{G}^\infty(0, x)$ of random real graphs such that

$$\text{scl} \left(\frac{\sqrt{\sigma_2(p)}}{\sum_i p_i u_i + 1}, \frac{1}{\sum_j |X_j^m|} \right) \text{SCoal}_1^{[m]}(X, q) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{G}^\infty(0, x)$$

for the Gromov-Hausdorff-Prokhorov topology. Furthermore,

$$\mathcal{E}_t^{\text{con}}(x, 0) \stackrel{(d)}{=} \text{scl}(\sqrt{x}, x) \mathcal{G}^\infty(0, x^{3/2})$$

2.5.3 The easier question of mixing of the dynamical percolation and its relation to noise sensitivity

Recall the proof layout of Assertion 1 and Assertion 2. In particular, each times the first step is a result about asymptotics in $-\infty$, but with different hypothesis: either, one assumes that the distribution of X is given (Assertion 1), either, one assumes that X is extreme (Assertion 2). One could hope that showing the convergence given the distribution of X could be slightly easier, as a lot of things are known on the marginals $\mathcal{E}_t(\alpha, \lambda, c, \gamma)$. However, this apparently weaker result, showing the convergence for $X \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$, would be an interesting result beyond its importance in the Aldous-Limic problem thanks to the theory developed in [28]:

Theorem 2.5.13 (Theorem 3.7 of [28]). *Let $X \sim \mathcal{E}_0(1, \lambda, 0, 1)$, there exists a process $\text{Perc}(\cdot, \cdot)$ called the dynamical percolation such that, if $G(0) \sim G(n, \frac{1+\lambda n^{-1/3}}{n})$ and $(G(t))_{t \in \mathbb{R}}$ is the process that, starting from $G(0)$ updates independently the edges at rate $n^{-1/3}$,*

$$(G(t))_{t \geq 0} \xrightarrow[n \rightarrow +\infty]{(d)} (\text{Perc}(X, t))_{\geq 0}$$

for the Gromov-Hausdorff-Prokhorov product topology. Moreover, for all $t \geq 0$, $\text{Perc}(X, t) \sim \mathcal{E}_0(1, \lambda, 0, 1)$. Lastly, if $(X(t))_{t \in \mathbb{R}} \sim \mathcal{E}(1, \lambda, 0, 1)$, for all f positive measurable for T_{GHP} ,

$$(X(t), \text{Perc}(X(t), s)) \stackrel{(d)}{=} (\text{Coal}(X(s), t-s), \widetilde{\text{Coal}}(X(s), t-s)). \quad (2.8)$$

where the Poisson point process used in $\text{Coal}(X(s), \cdot)$ and $\widetilde{\text{Coal}}(X(s), \cdot)$ is independent.

In particular, as a consequence of (2.8), Assertion 1 would give the following mixing property of $\text{Perc}(\cdot, t)$: for all f, g positive and measurable,

$$\mathbb{E}(f(\text{Perc}(X(s), t-s))g(X(s))) = \mathbb{E}(f(\text{Coal}(X(s), t-s))g(X(s))) \rightarrow \mathbb{E}(f(X(t)))\mathbb{E}(g(X(t))).$$

This mixing property is worth investigating by itself, as it is a close (but weaker) task to implementing the full Aldous-Limic method, but it is also interesting through its relation to the question of noise sensitivity. Let us introduce some background:

Definition 2.5.14 (Noise sensitivity [8]). *For all $n \in \mathbb{N}$, let Λ_n be a finite set such that $|\Lambda_n| \xrightarrow[n \rightarrow +\infty]{} +\infty$, let \mathbb{P}_n be a probability distribution on $\Omega_n = \{0, 1\}^{\Lambda_n}$. Let ω_n be distributed according to \mathbb{P}_n . Let $\epsilon > 0$, we define ω_n^ϵ as follow: let $\omega'_n \sim \mathbb{P}_n$ be independent of ω_n , then for $i \in \Lambda_n$, $\omega_n^\epsilon(i) = \omega_n(i)$ with probability $1 - \epsilon$ and $\omega_n^\epsilon(i) = \omega'_n(i)$ with probability ϵ independently for each coordinate. We then say that an event $E_n \subset \Omega_n$ is noise sensitive if*

$$\begin{cases} \mathbb{P}(E_n) \text{ is bounded away from } 0 \text{ and } 1, \\ \forall \epsilon > 0, \mathbb{P}(\omega_n \in E_n, \omega_n^\epsilon \in E_n) - \mathbb{P}(\omega_n \in E_n)\mathbb{P}(\omega_n^\epsilon \in E_n) \xrightarrow[n \rightarrow +\infty]{} 0. \end{cases}$$

The field of random graphs is a fertile ground to study noise sensitivity: one takes $\Lambda_n = \{\{i, j\} \in [n]^2 : i < j\}$ and for $e \in \Lambda_n$ associates $\omega(e) = 1$ with the existence of the edge e in the graph G^n . Here we will focus on the case where \mathbb{P}_n is such that for all $u \in \{0, 1\}^n$,

$$\mathbb{P}(\forall i \in [n], \omega_n(i) = u_i) = p^{\sum_i u_i} (1-p)^{n-\sum_i u_i}$$

for some $p = (p_n)_{n \in \mathbb{N}}$. This of course makes the aforementioned associated graph an Erdős–Rényirandom graph $G(n, p)$. The version of the graph after the noise has been applied, i.e. constructed from ω_n^ϵ will be denoted by $G^{n, \epsilon}$. In the spirit of the measured metric space point of view, G^n and $G^{n, \epsilon}$ will always be considered as elements of $\mathbb{M}_2^{\mathbb{N}}$, by setting the distance to be the graph distance divided by $n^{1/3}$ and the measure to be the counting measure divided by $n^{2/3}$.

For X and X' in $\mathbb{M}_2^{\mathbb{N}}$ we define

$$d_{GHP}^4(X, X') = \sum_j (d_{GHP}(X_j, X'_j))^4.$$

The fundamental thing to remark is that applying the noise on G^n is equivalent to performing a dynamical percolation on it. As a consequence, Lubetzky and Peled were able to obtain in [23]:

Theorem 2.5.15 (Theorem 2 of [23]). *Let $G^n \sim G(n, p)$ be equipped with the graph distance divided by $n^{1/3}$ and the counting measure divided by $n^{2/3}$, and $G^{n, \epsilon}$ the same object with a noise of level ϵ applied. Supposed that $p = \frac{1+tn^{-1/3}}{n}$. Let $\delta > 0$, suppose that $\epsilon \geq n^{-1/3+\delta}$. Then*

$$(G^n, G^{n, \epsilon}) \xrightarrow[n \rightarrow +\infty]{(d)} (X, Y)$$

where $X \sim \mathcal{E}_0(1, t, 0, 1)$ and $Y \stackrel{(d)}{=} X$ is independent from X . This convergence occurs for d_{GHP}^4 . On the other hand, if $\epsilon = o(n^{-1/3})$,

$$d_{GHP}^4(G^n, G^{n, \epsilon}) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Notice that the convergence for d_{GHP}^4 is a consequence of the convergence for T_{GHP}^2 and the fact that G^n and $G^{n, \epsilon}$ converge separately for d_{GHP}^4 . The second part of Theorem 2.5.15 can be recovered from:

Theorem 2.5.16 (Direct consequence of [28], Theorem 3.7). *Let $G^n \sim G(n, p)$ be equipped with the graph distance divided by $n^{1/3}$ and the counting measure divided by $n^{2/3}$, and $G^{n, \epsilon}$ the same object with a noise of level ϵ applied. Supposed that $p = \frac{1+tn^{-1/3}}{n}$, $\epsilon_n n^{-1/3} \xrightarrow[n \rightarrow +\infty]{} s \in \mathbb{R}^+$. Then,*

$$(G^n, G^{n, \epsilon}) \xrightarrow[n \rightarrow +\infty]{(d)} (X, \text{Perc}(X, s))$$

where $X \sim \mathcal{E}_0(1, t, 0, 1)$.

This theorem also has the advantage of providing a hint on how to derive noise sensibility result from the mixing property of $\text{Perc}(\cdot, \cdot)$: as the noise increases, it erases more and more information about the original graph. If $\epsilon n^{1/3} \rightarrow +\infty$, the noise is larger than any noise ϵ' that would be such that $\epsilon' n^{1/3} \rightarrow s \in \mathbb{R}_+$. From Theorem 2.5.16, for n large enough, the couple $(G^n, G^{n, \epsilon'})$ behaves like $(X, \text{Perc}(X, s))$, and as $(G^n, G^{n, \epsilon})$ is "more independent" than $(G^n, G^{n, \epsilon'})$ for any s , it is "more independent" than $(X, \text{Perc}(X, s))$ for any s . If the mixing property is true, the couple $(X, \text{Perc}(X, s))$ itself becomes asymptotically independent so as a consequence, so does $(G^n, G^{n, \epsilon})$.

2.6 Results

All the results given in this section are original. Recall the Guiding question 1. The analysis from the previous section hints that in order to perform Aldous and Limic's method from section 2.2.4, one needs to further break down Guiding question 1 into two sub-questions:

Guiding question 2 (About asymptotics in $-\infty$). *Let $(X(t))_{t \in \mathbb{R}}$ be an eternal multiplicative metric coalescent. What can we say about its behaviour at $-\infty$?*

Guiding question 3 (About the convergence of the coalescence). *Let $X = (X^n)_{n \in \mathbb{N}}$ be a sequence of $\mathbb{M}_2^{\mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}} \in \mathbb{N}$. Under what conditions on (X, q) can we say that*

$$\text{Coal}(X^n, q_n) \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_t(\alpha, \lambda, c, \gamma)$$

for certain parameters $(\alpha, \lambda, c, \gamma)$?

All our findings revolve around these two axes, the objective being the complete exploration of the entrance boundary of the eternal multiplicative metric coalescent, which is an other way of saying that we want to characterize all extreme metric coalescent. We do not have a definitive theorem that answers this question, but several partial answers to Guiding question 2 and Guiding question 3.

Before stating the result, we want to stress a particular issue about guiding question 2. In [6], the equivalent of guiding question 2 for the weight process for extreme multiplicative coalescent relies on martingales methods. However, as we saw in section 2.5.3, here we have to deal with average distances on components. This makes the implementation of martingales methods harder because there is no obvious way to see what becomes the average distance on a component after two of its points have coalesced. However, if one is only interested in the mixing property of $\text{Perc}(\cdot, \cdot)$ started from $X \sim \mathcal{E}_t(\alpha, \lambda, c, \gamma)$, the components are no longer arbitrary and the analysis becomes easier.

2.6.1 Tools

The first part of this work will be devoted to tools to handle multiplicative coalescent.

From global statistics to component-wise statistics For the Guiding question 3, we of course want to rely on Theorem 2.5.12 and Theorem 2.5.11 but there is a lot of obstacles when one wants to use these theorem with the complete coalescence on $\mathbb{M}_2^{\mathbb{N}}$. First, Theorem 2.5.11 is only about connected components conditionally on their vertex set. To overcome this, we need a tool to infer the content of a component, in terms of vertices, from global statistics. This is the subject of Chapter 3.

Truncation, Gromov-weak topology At this point, thanks to the work done in Chapter 3, we can reformulate Theorem 2.5.12 and Theorem 2.5.11 as theorems on the simple coalescence of a finite number of components. As we would like to obtain convergence results starting from an infinite sequence of metric spaces we need tools that allow to approximate the coalescence of this infinite sequence from the coalescence of a finite sequence. We will see in Chapter 4 that the Gromov-weak topology is really the most natural setting for this kind of results.

From simple to complete coalescence As the literature focuses on the simple coalescence and we are mainly interested in the complete coalescence, Chapter 5 will bridge the gap.

Eternal multiplicative coalescent, mixing and entrance boundary In Chapter 6, we deal specifically with eternal multiplicative coalescents, and we show that the tightness of the lower mass bound is guaranteed. This leads to Theorem 6.2.1 that says that the distribution of the coalescent is characterized, at least partially, by the convergence of certain quantities in $-\infty$.

2.6.2 Main results

The second part of this work gives partial answers to the question of the complete characterization of the entrance boundary, which can also be seen as a way to better our understanding of the multiplicative coalescent and dynamical percolation.

Palm measure of the point process associated with $\mathcal{E}_t(\alpha, \beta, c, \gamma)$ Recall Theorem 2.2.15 where Janson and Spencer were able to compute an explicit formula for the Palm measure of the distribution $\mathcal{M}_0(1, \lambda, 0)$. As it allows to transform higher order moments into first order moments, it is a useful tool to show the concentration of quantities related to $x = (x(t))_{t \in \mathbb{R}} \sim \mathcal{M}(1, \lambda, 0)$. In the light of the simpler problem of mixing of the dynamical percolation started from $X \sim \mathcal{E}_t(\alpha, \lambda, c, \gamma)$, it would be useful to have a similar result but with metric spaces in place of weights and for all admissible parameters. In Chapter 7, we obtain the following:

Theorem 7.1.1. *Let $(\alpha, \lambda, c, \gamma) \in \mathcal{I} \times \mathbb{R}_+$ and recall from Definition 2.4.8 the definition of $(\Delta_j)_{j \in \mathbb{N}}$, $(c^{(j)})_{j \in \mathbb{N}}$ and $(\delta_2^{(j)})_{j \in \mathbb{N}}$ associated with $\mathcal{X} \sim \mathcal{E}_0(\alpha, \lambda, c, \gamma)$. Let $f : \mathbb{M} \times \mathbb{M}_2^{\mathbb{N}} \rightarrow \mathbb{R}^+$ measurable for T_{GHP} and such that*

$$\mathbb{E}\left(\sum_j f(\mathcal{X}_j, \mathcal{X})\right) < +\infty.$$

For $j \in \mathbb{N}$, conditionally on $(\mathcal{X}_j, \Delta_j)$, let $\mathcal{Y}^{(j)}$ be distributed as $\mathcal{E}_0(\alpha, \lambda - |\mathcal{X}_j| - \delta_2^{(j)}, c - c^{(j)}, \gamma)$. Assume that

$$\text{either } c = 0, \text{ either } \lambda < 0. \quad (7.1)$$

We have:

$$\mathbb{E}\left(\sum_j f(\mathcal{X}_j, \mathcal{X})\right) = \mathbb{E}\left(\sum_j \mathbb{E}(f(\mathcal{X}_j, \mathcal{Y}^{(j)} \sqcup \mathcal{X}_j | (\mathcal{X}_j, \Delta_j)))\right). \quad (7.2)$$

Mixing property for homogeneous multiplicative spaces As already explained in Section 2.5.3, the question of mixing of the dynamical percolation is a simpler yet related question to the implementation of the Aldous-Limic method from Section 2.2.4 for the complete coalescence of metric spaces. Here we test our reformulation of Theorem 2.5.12 to the complete coalescence of element of $\mathbb{M}_2^{\mathbb{N}}$. As we start from a known distribution $\mathcal{E}_t(1, \lambda, 0, 1)$ and not a generic extreme eternal multiplicative coalescent, the question of the entrance boundary Guiding question 2 is easier. We rely here on the Janson and Spencer formula of Theorem 7.1.1, which in Chapter 8 allows us to obtain a control on the variance of the metric estimates required in Theorem 2.5.12.

Theorem 8.2.2. *Let $(X(t))_{t \in \mathbb{R}} \sim \mathcal{E}(1, \lambda, 0, 1)$. Recall from Definition 2.5.6 the definition of ρ_1, u, x, D . There exists constants K_r such that, as $t \rightarrow -\infty$,*

$$\forall p \geq 2, r \geq 0, \sum_j x_j^p u_i^r \sim_{\mathbb{P}} K_r \sigma_{p+r/2}, \quad (8.1)$$

$$\forall p \geq 2, \sigma_p \sim \mathbb{E}(\sigma_p), \quad (8.2)$$

$$\forall \eta \in (0, 1), \frac{x_0}{\sigma_2} = o_{\mathbb{P}}(\sigma_2^\eta) \quad (8.3)$$

$$\forall \eta \in (0, 1), \sup_i D_i = o_{\mathbb{P}}(\sigma_2^\eta) \quad (8.4)$$

This allows us to show:

Corollary 8.2.3. *Let $X \sim \mathcal{E}(1, \lambda, 0, 1)$, then for all $t \in \mathbb{R}$, for all f and g bounded and measurable for T_{GHPS} ,*

$$\mathbb{E}(f(\text{Perc}(X(t), s))g(X(t))) \xrightarrow{s \rightarrow +\infty} \mathbb{E}(f(X(t)))\mathbb{E}(g(X(t))).$$

Noise sensitivity As already explained in Section 2.5.3, the mixing property is strongly linked to the noise sensitivity. In Chapter 9, we show that the asymptotic independence of Theorem 2.5.15 stay true for a noise of level ϵ such that $\epsilon n^{1/3} \rightarrow \infty$, which is stronger. Moreover, we give various examples of noise sensitive properties: the property of containing a given large topological structure (as a large cycle for instance) and the property that there exists an embedding of a graph in a given surface.

2.6.3 Perspective

The last part of this work is devoted to a sketch of a program that would, once entirely carried out, possibly give a complete characterization of the entrance boundary of the metric multiplicative coalescent. On one hand, we propose to study an encoding of $\text{SCoal}(\cdot, \cdot)$ à la Broutin, Duquesne and Wang [14] and we show that this process is tight in Lemma 10.3.4. On the other hand, we show in Theorem 10.4.2 that it is possible to couple a given coalescent $(X(t))_{t \in \mathbb{N}}$ with a process $(T(t))_{t \in \mathbb{N}}$ that doesn't add complexity inside the components, in such way that their $-\infty$ asymptotics are the same. This negates the already alluded difficulty of controlling distance statistics when adding internal coalescent points.

Part II

Tools

Chapter 3

Link between global statistics and component-wise statistics

3.1 Results

Let $x \in \overline{sl}$ and $q = (q_n)_{n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$ be such that (x, q) satisfies the Aldous-Limic conditions 1. Let $y = (y_i^n)_{i \in \mathbb{N}, n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N} \times \mathbb{N}}$ be a sequence that attributes a certain non-negative statistic to each vertex. These statistics can be their weights, but also, in light of the complete coalescence on metric spaces, some metric statistics like the average distances. Our question here is, what can be said about the distribution of these statistics on the components of $\text{gCoal}(x, q)$, let us say on the largest component for instance? Recall from Definition 2.4.2 the exploration of the component of $\text{gCoal}(x, q)$ encoded in Z . This exploration has the particularity that it sees the largest component at a time that is tight. This hints that we could track the statistics y along this exploration. To this end, we define

$$Y(t) = \sum_{i \in \mathbb{N}} y_i 1_{\xi_i \leq t}.$$

Moreover, for two sequences $x = (x_i)_{i \in \mathbb{N}}$ and $y = (y_i)_{i \in \mathbb{N}}$, for $p \in \mathbb{N}$, $r \in \mathbb{N}$, we define $\sigma_{p,r}(x, y) = \sum_i x_i^p y_i^r$. If x and y are obvious from context, we will usually only write $\sigma_{p,r}$.

Hypothesis 2 (Statistics conditions). *Let $x \in \overline{sl}$, $y = (y_i^n)_{i \in \mathbb{N}, n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N} \times \mathbb{N}}$, $q \in \mathbb{R}^{\mathbb{N}}$, there exists (α, λ, c) and $(\beta, e) \in \mathbb{R}_+ \times \mathbb{R}^{\mathbb{N}}$, such that $\sum_{i \in \mathbb{N}} e_i c_i < \infty$, (x, q) satisfies Hypothesis 1 with parameter $(\alpha, \lambda, c) \in \mathcal{I}$ and x, q and y together satisfy:*

$$\begin{cases} \forall i \in \mathbb{N}, y_i \rightarrow e_i \\ q\sigma_{1,1} - \sum_{i \in \mathbb{N}} c_i e_i \rightarrow \beta \\ \lim_m \limsup_n q \sum_{i > m} x_i y_i^2 = 0. \end{cases}$$

Then the main result of this chapter is the following:

Theorem 3.1.1. *Let $x \in \overline{sl}$, $y = (y_i^n)_{i \in \mathbb{N}, n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N} \times \mathbb{N}}$, $q \in \mathbb{R}^{\mathbb{N}}$, that satisfy Hypothesis 2 with parameters $(\alpha, \lambda, c) \in \mathcal{I}$ and $(\beta, e) \in \mathbb{R}_+ \times \mathbb{R}^{\mathbb{N}}$. Recall for $n \in \mathbb{N}$, $(r_j^n)_{n \in \mathbb{N}}$ and $(r_j^n)_{j \in \mathbb{N}}$ from Definition 2.4.5, and $(r_j)_{n \in \mathbb{N}}$ and $(r_j)_{j \in \mathbb{N}}$ from Definition 2.4.7. Then:*

$$(Y, Z, (r_j^n)_{j \in \mathbb{N}}, (l_j^n)_{j \in \mathbb{N}}) \xrightarrow{\mathcal{L}} (\mathcal{Y}, \mathcal{Z}, (r_j)_{j \in \mathbb{N}}, (l_j)_{j \in \mathbb{N}})$$

for the Skorokhod topology and the product topology, with

$$\mathcal{Y}(s) = \sum_{i \in \mathbb{N}} e_i 1_{\zeta_i \leq s} + \beta s,$$

where (ζ_i) is the sequence of exponential random variables with parameters (c_i) that appears in

$$\mathcal{Z}(s) = \alpha^{1/2} \mathcal{W}(s) + \lambda s - \frac{1}{2} \alpha s^2 + \sum_j (c_j 1_{\zeta_j \leq s} - c_j^2 s).$$

Proof. See Section 3.3 □

Notice that \mathcal{Y} is almost surely finite since, for $s \in \mathbb{R}$, $\mathbb{E}(\mathcal{Y}(s)) = \sum_{i \in \mathbb{N}} c_i e_i + \beta s < \infty$. To completely fulfill the goal of this chapter, we must show that useful information on the individual components can be effectively inferred from the convergence of Y . For this purpose, we introduce the following notation

Definition 3.1.2. Let $x \in \mathcal{L}_{\downarrow}^2$, $q \in \mathbb{R}$, $\mathcal{X} = \text{SCoal}(x, q)$, $r > 0$, $j \in \mathbb{N}$, we define

$$\theta_r(\mathcal{X}_j) = \sum_{i:\{i\} \subset \mathcal{X}_j^n} \left(\frac{x_i}{\sigma_2} \right)^r.$$

Theorem 3.1.3 (Components statistics from global statistics). Let $x = (x^n)_{n \in \mathbb{N}} \in \overline{sl}$, $y = (y^n)_{n \in \mathbb{N}}$ a sequence in $\mathbb{R}_+^{\mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}}$ that satisfy Hypothesis 2 with parameters $(\alpha, \lambda, c) \in \mathcal{I}$ and $(\beta, e) \in \mathbb{R}_+ \times \mathbb{R}^{\mathbb{N}}$. For $n \in \mathbb{N}$, let $\mathcal{X}^n = \text{SCoal}(x^n, q_n)$. Recall from Definition 2.4.8 the definition of $(\Delta_j)_{j \in \mathbb{N}}$ and $(\delta_r^{(j)})_{j \in \mathbb{N}}$ associated with \mathcal{X} . For $j \in \mathbb{N}$, we have that

$$\sum_{i:\{i\} \subset \mathcal{X}_j^n} y_i \xrightarrow[n \rightarrow +\infty]{(d)} \sum_{i \in \Delta_j} e_i + \beta |\mathcal{X}_j^\infty|.$$

In particular,

$$\begin{aligned} \theta_2(\mathcal{X}_j^n) &\xrightarrow[n \rightarrow +\infty]{(d)} \delta_2^{(j)} + \alpha |\mathcal{X}_j^\infty|, \\ \forall r > 2, \theta_r(\mathcal{X}_j^n) &\xrightarrow[n \rightarrow +\infty]{(d)} \delta_r^{(j)}. \end{aligned}$$

Proof. See Section 3.4 □

3.2 Consequence and discussion

3.2.1 Special cases

Even if it seems that we need a precise control on the limit of each $y_i = (y_i^n)_{n \in \mathbb{N}}$, in the purely continuous case, the hypothesis $q\sigma_{1,2} \rightarrow 0$ is strong enough to justify the convergence without such a control. In fact:

Lemma 3.2.1 (The special case of a purely continuous limit). Let $x = (x^n)_{n \in \mathbb{N}} \in \overline{sl}$, $y = (y^n)_{n \in \mathbb{N}}$ a sequence in $\mathbb{R}_+^{\mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}}$. Assume that (x, q) satisfies the Aldous-Limic conditions 1. Assume moreover that there exists $\beta \geq 0$ such that:

$$\begin{cases} q\sigma_{1,1}(x, y) \rightarrow \beta, \\ q\sigma_{1,2}(x, y) \rightarrow 0. \end{cases}$$

Then there exists another sequence of $\mathbb{R}^{\mathbb{N}}$ denoted by y' such that, together with x , it satisfies Hypothesis 2 with parameters $(\beta, 0)$, in other words:

$$\begin{cases} \forall i \in \mathbb{N}, y'_i \rightarrow 0, \\ q\sigma_{1,1}(x, y') \rightarrow \beta, \\ q\sigma_{1,2}(x, y') \rightarrow 0. \end{cases}$$

And

$$\forall T \geq 0, \forall \eta > 0, \mathbb{P} \left(\sup_{t \leq T} \left| \sum_i y_i 1_{\xi_i \leq t} - \sum_i y'_i 1_{\xi_i \leq t} \right| \geq \eta \right) \xrightarrow[n \rightarrow +\infty]{} 0$$

Proof. Let (x, y, q) satisfy the hypothesis of the lemma. Let $\epsilon > 0$. We have:

$$\begin{aligned} q \sum_i x_i y_i - q \sum_i x_i y_i 1_{y_i \leq \epsilon} &= q \sum_i x_i y_i 1_{y_i > \epsilon} \\ &\leq q \frac{\sum_i x_i y_i^2}{\epsilon^2} \\ &= \frac{q\sigma_{1,2}(x, y)}{\epsilon^2}. \end{aligned} \tag{3.1}$$

and

$$\begin{aligned} \mathbb{P}\left(\sup_{t \leq T} \left| \sum_i y_i 1_{\xi \leq t} - \sum_i y_i 1_{y_i \leq \epsilon} 1_{\xi_i \leq t} \right| \geq \eta\right) &\leq \mathbb{P}(\exists i \in \mathbb{N}, \xi_i \leq T, y_i > \epsilon) \\ &= 1 - e^{-q \sum_i x_i 1_{y_i > \epsilon}} \\ &\leq \frac{q\sigma_{1,2}(x, y)}{\epsilon^2}. \end{aligned} \tag{3.2}$$

Let $\epsilon = (\epsilon_n)_{n \in \mathbb{N}}$ be defined by

$$\forall n \in \mathbb{N}, \epsilon_n = (q_n \sigma_{1,2}(x^n, y^n))^{1/3}.$$

Then according to (3.1) and (3.2), the sequence y' such that

$$\forall i \in \mathbb{N}, n \in \mathbb{N}, y'_i = y_i^n 1_{y_i^n \leq \epsilon_n}$$

fits the goal of the lemma. Indeed,

$$\begin{aligned} \forall i \in \mathbb{N}, y_i^n &\leq \epsilon^n \xrightarrow{n \rightarrow +\infty} 0, \\ 0 \leq q(\sigma_{1,1}(x, y) - \sigma_{1,1}(x, y')) &\leq (q\sigma_{1,2}(x, y))^{2/3} \xrightarrow{n \rightarrow +\infty} 0, \\ q\sigma_{1,2}(x, y') &\leq q\sigma_{1,2}(x, y) \xrightarrow{n \rightarrow +\infty} 0 \end{aligned}$$

and

$$\mathbb{P}\left(\sup_{t \leq T} \left| \sum_i y_i 1_{\xi \leq t} - \sum_i y'_i 1_{\xi_i \leq t} \right| \geq \eta\right) \leq (q\sigma_{1,2}(x, y))^{1/3} \xrightarrow{n \rightarrow +\infty} 0.$$

□

The meaning of Lemma 3.2.1 is that one can apply Theorem 3.1.3 without checking the pointwise convergence part of Hypothesis 2, as long as we are in the purely continuous regime.

One important problem to which we would like to apply Theorem 3.1.3 is the problem of the expectation of a statistic of a random point on a given component. In other words, if $u = (u_i^n)_{i \in \mathbb{N}, n \in \mathbb{N}}$, we want to compute $\sum_{i \in X_j^n} \frac{x_i}{|X_j^n|} u_i$.

Corollary 3.2.2 (Components statistics from global statistics). *Let $x = (x^n)_{n \in \mathbb{N}} \in \overline{sl}$ and $q = (q_n) \in \mathbb{R}_+^{\mathbb{N}}$. Assume that (x, q) satisfies the Aldous-Limic conditions Hypothesis 1. Let $u = (u^n)_{n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$ be such that, together with x , it satisfies:*

$$\frac{\sigma_2(x)\sigma_{3,2}(x, u)}{\sigma_{2,1}(x, u)^2} \rightarrow 0$$

For $n \in \mathbb{N}$, let $\mathcal{X}_j^n = S\text{Coal}(x^n, q)$. Then we have

$$\sum_{i: \{i\} \in \mathcal{X}_j^n} \frac{x_i u_i}{|\mathcal{X}_j^n|} \sim \frac{\sigma_{2,1}(x, u)}{\sigma_2}$$

in probability.

Proof. We simply define $y_i = \frac{\sigma_2 x_i u_i}{\sigma_{2,1}(x, u)}$. Then

$$q\sigma_{1,1} = q\sigma_2 \rightarrow 1$$

and

$$q\sigma_{1,2} = \frac{q\sigma_2^2}{\sigma_{2,1}(x, u)^2} \sum_i x_i^3 u_i^2 \sim \frac{\sigma_2 \sigma_{3,2}}{\sigma_{2,1}^2} \rightarrow 0$$

by hypothesis. Then Lemma 3.2.1 tells us that it is all we need to apply Theorem 3.1.3. □

3.2.2 Global reformulation for Theorem 2.5.12 and Theorem 2.5.11

We want to apply Theorem 3.1.3 to infer a theorem on the simple coalescence of metric spaces from Theorem 2.5.12 and Theorem 2.5.11. Note that it has already been done for Theorem 2.5.12 in [9]. Here we do the same with slightly weaker hypotheses.

Theorem 3.2.3. *Let X be a sequence of $\overline{\mathbb{M}}^{\mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$ such that if x is its weight sequence, (q, x) satisfies the Aldous-Limic conditions 1 with parameters $(1, \lambda, 0) \in \mathcal{I}^\circ$. Recall from Definition 2.5.6 the definition of ρ_1, u, x, D and x_{\min} . Assume moreover that if $\mathcal{X}^n = \text{SCoal}_1(X^n, q_n)$, there exists $\eta > 0$ and $r > 0$ such that:*

$$\begin{cases} \frac{x_0}{\sigma_2} = o(\sigma_2^{1/2+\eta}) \\ \frac{\sigma_2^r}{x_{\min}} \rightarrow 0, \end{cases}$$

and if $\mathcal{X}^n = \text{SCoal}_1(X^n, q_n)$,

$$\begin{cases} \forall j \in \mathbb{N}, \max_{i: X_i \subset \mathcal{X}_j^n} D_i \frac{\sigma_2^2}{\rho_1 + \sigma_2} = o_{\mathbb{P}}(\sigma_2^{1/2+\eta}) \\ \frac{\sigma_2^2}{\rho_1} \sigma_{3,2}(x, u) \rightarrow 0. \end{cases}$$

Then we have

$$\frac{\sigma_2^2}{\rho_1 + \sigma_2} \mathcal{X} \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_0(1, \lambda, 0, 1)$$

for the Gromov-Hausdorff-Prokhorov product topology.

Proof. In the same spirit as Proposition 6.7 in [9], let us show that

$$\begin{aligned} \frac{\sum_{i: X_i \subset \mathcal{X}_j} x_i^2 \sigma_2}{|\mathcal{X}_j| \sigma_3} &\xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 1 \\ \frac{\sum_{i: X_i \subset \mathcal{X}_j} x_i u_i \sigma_2}{|\mathcal{X}_j| \rho_1} &\xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 1. \end{aligned}$$

The first one is obtained by applying Theorem 3.1.3 and the second one is obtained by applying Corollary 3.2.2, considering that thanks to Lemma 3.2.1, one can skip the control of the maximums. Then, if for $i \in \mathbb{N}$, j is such that $X_i \subset \mathcal{X}_j$, we define $p_i^j = \frac{x_i}{|\mathcal{X}_j|}$ and $q^j = q|\mathcal{X}_j|^2$. Recall that, under the Aldous-Limic conditions Hypothesis 1 with parameter $(1, \lambda, 0)$, $\sigma_3 \sim \sigma_2^3$.

We have, for $j \in \mathbb{N}$,

$$\begin{aligned} \sigma_2(p^j) &= \frac{\sum_{i: X_i \subset \mathcal{X}_j} x_i^2}{|\mathcal{X}_j|^2} \sim_{\mathbb{P}} \frac{\sigma_3}{\sigma_2 |\mathcal{X}_j|} \sim_{\mathbb{P}} \frac{\sigma_2^2}{|\mathcal{X}_j|} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0 \\ \frac{\max_i p_i^j}{\sqrt{\sigma_2(p^j)}} &= \frac{|\mathcal{X}_j| \max_i p_i}{\sqrt{\sum_{i: X_i \subset \mathcal{X}_j} x_i^2}} \leq \frac{x_0}{\sqrt{\sum_{i: X_i \subset \mathcal{X}_j} x_i^2}} \sim_{\mathbb{P}} \frac{x_0 \sqrt{\sigma_2 |\mathcal{X}_j|}}{\sqrt{\sigma_3}} \sim_{\mathbb{P}} \sqrt{|\mathcal{X}_j|} \frac{x_0}{\sigma_2} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0, \\ \frac{q^j \sqrt{\sigma_2(p^j)}}{|\mathcal{X}_j|^{3/2}} &= q |\mathcal{X}_j|^{-1/2} \sqrt{\sum_{i: X_i \subset \mathcal{X}_j} x_i^2} \sim_{\mathbb{P}} q \sqrt{\frac{\sigma_3}{\sigma_2}} \sim_{\mathbb{P}} q \sigma_2 \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 1 \\ \frac{(\sigma_2(p^j))^{r/2}}{\min_i p_i} &\leq \frac{(\sigma_2(p^j))^{r/2}}{x_{\min}} = \frac{(\sum_{i: X_i \subset \mathcal{X}_j} x_i^2)^{r/2}}{x_{\min} |\mathcal{X}_j|^r} \sim_{\mathbb{P}} \frac{\sigma_3^{r/2} |\mathcal{X}_j|^{r/2}}{x_{\min} |\mathcal{X}_j|^r \sigma_2^{r/2}} \sim_{\mathbb{P}} \frac{\sigma_2^r}{x_{\min} |\mathcal{X}_j|^{r/2}} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0, \\ \sqrt{|\mathcal{X}_j|} \frac{\sqrt{\sigma_2(p^j)}}{\sum_{i: X_i \subset \mathcal{X}_j} p_i u_i + 1} &= \sqrt{|\mathcal{X}_j|} \frac{\sqrt{\sum_{i: X_i \subset \mathcal{X}_j} x_i^2}}{\sum_{i: X_i \subset \mathcal{X}_j} x_i u_i + |\mathcal{X}_j|} \sim_{\mathbb{P}} \frac{|\mathcal{X}_j| \sqrt{\sigma_2^2}}{|\mathcal{X}_j| \frac{\rho_1}{\sigma_2} + o_{\mathbb{P}}(\frac{\rho_1}{\sigma_2}) + |\mathcal{X}_j|} \sim_{\mathbb{P}} \frac{\sigma_2^2}{\rho_1 + \sigma_2} \end{aligned}$$

so

$$\frac{\sup_{i: X_i \subset \mathcal{X}_j} D_i \sqrt{\sigma_2(p^j)}}{\sum_{i: X_i \subset \mathcal{X}_j} p_i u_i + 1} \sim_{\mathbb{P}} \frac{\sup_{i: X_i \subset \mathcal{X}_j} D_i \sigma_2^2}{\rho_1 + \sigma_2} = o_{\mathbb{P}}(\sigma_2^{1/2+\eta}).$$

All these convergences can be rephrased as convergences in probability toward constants. Using Skorokhod's representation theorem and Theorem 2.2.12, one may suppose that almost surely, the sizes

$(|\mathcal{X}_j^n|)_{j \in \mathbb{N}}$ converge toward $(|\mathcal{X}_j^\infty|)_{j \in \mathbb{N}}$ and the convergences above hold. Thus, conditionally on $(\mathcal{V}_j^n)_{j \in \mathbb{N}, n \in \mathbb{N}}$, the sequence of set of vertices spanned by the components $(\mathcal{X}_j^n)_{n \in \mathbb{N}}$, we can apply Theorem 2.5.12 for each $j \in \mathbb{N}$, which gives

$$\text{scl} \left(\frac{\sqrt{\sigma_2(p^j)}}{\sum_i p_i^j u_i + 1}, \frac{1}{|\mathcal{X}_j|} \right) \text{SCoal}_1^{\mathcal{V}_j}(X, q^j) \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{G}^\infty(0, |\mathcal{X}_j^\infty|^{3/2})$$

as

$$|\mathcal{X}_j^n| \xrightarrow[n \rightarrow \infty]{a.s.} |\mathcal{X}_j^\infty|.$$

Moreover, as

$$\frac{\sqrt{\sigma_2(p^j)}}{\sum_i p_i^j u_i + 1} \sim_{\mathbb{P}} \frac{1}{\sqrt{|\mathcal{X}_j|}} \frac{\sigma_2^2}{\rho_1 + \sigma_2},$$

we have

$$\text{scl} \left(\frac{1}{\sqrt{|\mathcal{X}_j|}} \frac{\sigma_2^2}{\rho_1 + \sigma_2}, \frac{1}{|\mathcal{X}_j|} \right) \text{SCoal}_1^{\mathcal{V}_j}(X, q^j) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{G}^\infty(0, |\mathcal{X}_j^\infty|^{3/2}).$$

Finally, as $\mathcal{E}_t^{\text{con}}(x, 0) \stackrel{(d)}{=} \text{scl}(\sqrt{x}, x) \mathcal{G}^\infty(0, x^{3/2})$,

$$\frac{\sigma_2^2}{\rho_1 + \sigma_2} \text{SCoal}_1^{\mathcal{V}_j}(X, q^j) \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_t^{\text{con}}(|\mathcal{X}_j^\infty|, 0).$$

□

Remark that in Theorem 3.4 of [9], it is required that $\max_{i \in \mathbb{N}} D_i \frac{\sigma_2 x_0}{\rho_1} \rightarrow 0$ which is stronger than $\frac{\sigma_2}{\rho_1^2} \sigma_{3,2}(x, u) \rightarrow 0$ and it is required $\max_{i \in \mathbb{N}} D_i \frac{\sigma_2^2}{\rho_1 + \sigma_2} = o_{\mathbb{P}}(\sigma_2^{1/2+\eta})$ which is also stronger than for any $j \in \mathbb{N}$, $\max_{i: X_i \subset \mathcal{X}_j^n} D_i \frac{\sigma_2^2}{\rho_1 + \sigma_2} = o_{\mathbb{P}}(\sigma_2^{1/2+\eta})$. Our improvement lies in the fact that in Theorem 3.2.3, one doesn't need to know anything about the maximal diameters in all the components, which might as well go to infinity, as long as large diameters concern components that are small in weight. However, we need to know that there are only small diameters in large components. We can do the same for Theorem 2.5.11, and this has not been done before:

Theorem 3.2.4. *Let X be a sequence of $\overline{\mathbb{M}}^{\mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$ such that if x is its weight sequence, (q, x) satisfies the Aldous-Limic conditions 1 with parameters $(0, \lambda, c) \in \mathcal{I}^*$. Recall from Definition 2.5.6 the definition of ρ_1 , u , x , D and x_{\min} and recall from Definition 2.4.8 the definition of $(\Delta_j)_{j \in \mathbb{N}}$, $(\delta_2^{(j)})_{j \in \mathbb{N}}$ and $(c^j)_{j \in \mathbb{N}}$. Assume moreover that if $\mathcal{X}^n = \text{SCoal}_1(X^n, q_n)$,*

$$\begin{cases} \forall j \in \mathbb{N}, \max_{i: X_i \subset \mathcal{X}_j^n} D_i \frac{\sigma_2^2}{\rho_1 + \sigma_2} = o_{\mathbb{P}}(1) \\ \frac{\sigma_2^2}{\rho_1^2} \sum_i x_i^3 u_i^2 \rightarrow 0. \end{cases}$$

Then we have

$$\frac{\sigma_2^2}{\rho_1 + \sigma_2} \mathcal{X}^n \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_0(0, \lambda, c, 1)$$

for the Gromov-weak product topology.

Proof. It is the same proof than theorem 3.2.3, except that we have

$$\begin{aligned} \frac{\sum_{i: X_i \subset \mathcal{X}_j} x_i^2}{\sigma_2^2} &\xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \delta_2^{(j)} \\ \frac{\sum_{i: X_i \subset \mathcal{X}_j} x_i u_i}{|\mathcal{X}_j|} \frac{\sigma_2}{\rho_1} &\xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 1. \end{aligned}$$

As a consequence, we get

$$\begin{aligned} q^j \sqrt{\sigma_2(p^j)} &\sim_{\mathbb{P}} |\mathcal{X}_j| \sqrt{\delta_2^{(j)}}, \\ \frac{\sqrt{\sigma_2(p^j)}}{\sum_{i: X_i \subset \mathcal{X}_j} p_i u_i + 1} &\sim_{\mathbb{P}} \frac{\sqrt{\delta_2^{(j)}}}{\sqrt{|\mathcal{X}_j|}} \frac{\sigma_2^2}{\rho_1 + \sigma_2} \end{aligned}$$

and the other asymptotics are unchanged. Thus, conditionally on $(\mathcal{V}_j^n)_{j \in \mathbb{N}, n \in \mathbb{N}}$, the sequence of set of vertices spanned by the components $(\mathcal{X}_j^n)_{n \in \mathbb{N}}$, we can apply Theorem 2.5.11 for each $j \in \mathbb{N}$, which gives

$$\text{scl} \left(\frac{\sqrt{\sigma_2(p^j)}}{\sum_i p_i^j u_i + 1}, \frac{1}{|\mathcal{X}_j|} \right) \text{SCoal}_1^{\mathcal{V}_j}(X, q^j) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{G}^\infty \left(\frac{c^j}{\sqrt{\delta_2^{(j)}}}, \sqrt{\delta_2^{(j)}} |\mathcal{X}_j^\infty| \right)$$

as

$$|\mathcal{X}_j^n| \xrightarrow[n \rightarrow \infty]{a.s.} |\mathcal{X}_j^\infty|$$

and

$$\theta_2(\mathcal{X}_j^n) \xrightarrow[n \rightarrow \infty]{} \delta_2^{(j)}.$$

Moreover, as

$$\frac{\sqrt{\sigma_2(p^j)}}{\sum_{i: X_i \subset \mathcal{X}_j} p_i u_i + 1} \sim_{\mathbb{P}} \frac{\sqrt{\delta_2^{(j)}}}{|\mathcal{X}_j|} \frac{\sigma_2}{\rho_1 + \sigma_2},$$

we have

$$\text{scl} \left(\frac{\sqrt{\delta_2^{(j)}}}{|\mathcal{X}_j|} \frac{\sigma_2}{\rho_1 + \sigma_2}, \frac{1}{|\mathcal{X}_j|} \right) \text{SCoal}_1^{\mathcal{V}_j}(X, q^j) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{G}^\infty \left(\frac{c^j}{\sqrt{\delta_2^{(j)}}}, \sqrt{\delta_2^{(j)}} |\mathcal{X}_j^\infty| \right).$$

Finally, as $\mathcal{E}_t^{\text{con}}(x, c) \stackrel{(d)}{=} \text{scl} \left(\frac{x}{\|c\|_2}, x \right) \mathcal{G}^\infty \left(\frac{c}{\|c\|_2}, x \|c\|_2 \right)$,

$$\frac{\sigma_2}{\rho_1 + \sigma_2} \text{SCoal}_1^{\mathcal{V}_j}(X, q^j) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{E}_t^{\text{con}}(|\mathcal{X}_j^\infty|, c^{(j)}).$$

□

Remark that in the case of Theorem 3.2.4, as the hypothesis were weaker, our Corollary 3.2.2 become even more interesting since the fact for all $j \in \mathbb{N}$, $\max_{i: X_i \subset \mathcal{X}_j^n} D_i \frac{\sigma_2}{\rho_1 + \sigma_2} = o_{\mathbb{P}}(1)$, considering that $\frac{\sigma_2}{\rho_1 + \sigma_2}$ is the scaling factor, is only saying that there is no macroscopic subcomponent X_i of a large component \mathcal{X}_j .

3.3 Proof of Theorem 3.1.1

The trick here to disentangle the two regimes (continuous part and jump part) originated in Aldous and Limic [6] and consists in a dichotomy between two kinds of weight, small and large. Let $m \in \mathbb{N}$. From the proof of Lemma 5.6 in [14] and Lemma 5.7 from [14] we have the unsurprising fact that,

$$(Z, (\xi_i)_{i \in \mathbb{N}}, (r_j^n)_{j \in \mathbb{N}}, (l_j^m)_{j \in \mathbb{N}}) \xrightarrow[n \rightarrow +\infty]{(d)} (Z, (\zeta_i)_{i \in \mathbb{N}}, (r_j)_{j \in \mathbb{N}}, (l_j)_{j \in \mathbb{N}}).$$

For the product of the Skorokhod topology and the product topology. Suppose, thanks to Skorokhod's representation theorem, that this convergence occurs almost surely. This allows us to define \mathcal{Y} on the same probability space than Y . Let $m \in \mathbb{N}$ and $\epsilon > 0$. For $T > 0$, let d_S^T be the Skorokhod distance on $[0, T]$. We have, for any $n \geq m$,

$$\mathbb{P}(d_S^T(Y, \mathcal{Y}) \geq \epsilon) \leq \mathbb{P}(d_S^T(Y^\circ, \mathcal{Y}^\circ) \geq \epsilon/2) + \mathbb{P}(d_S^T(Y^*, \mathcal{Y}^*) \geq \epsilon/2) \quad (3.3)$$

where

$$\begin{aligned} Y^*(t) &= \sum_{i=0}^m y_i 1_{\xi_i \leq t}, \\ \mathcal{Y}^*(t) &= \sum_{i=0}^m e_i 1_{\zeta_i \leq t}, \\ Y^\circ(t) &= \sum_{i=m+1}^n y_i 1_{\xi_i \leq t}, \end{aligned}$$

$$\mathcal{Y}^\circ(t) = \beta t.$$

It is clear that, for m fixed,

$$\lim_{n \rightarrow \infty} \mathbb{P}(d_S^T(Y^*, \mathcal{Y}^*) \geq \epsilon/2) = 0 \quad (3.4)$$

On the other hand we can bound $\mathbb{P}(d_S^T(Y^\circ, \mathcal{Y}^\circ) \geq \epsilon/2)$ as follows:

$$\mathbb{P}(d_S^T(Y^\circ, \mathcal{Y}^\circ) \geq \epsilon/2) \leq \mathbb{P}\left(\sup_{t \leq T} \left| Y^\circ(t) - q \sum_{i=m+1}^n x_i y_i \right| \geq \epsilon/4\right) + 1_{T|q \sum_{i=m+1}^n x_i y_i - \beta| \geq \epsilon/4}.$$

We have that $Y^\circ(t) - q \sum_{i=m+1}^n x_i y_i t$ is a supermartingale: for $s \leq t$,

$$\begin{aligned} \mathbb{E}(Y^\circ(t) | \mathcal{F}_s) &= \sum_{i=m+1}^n y_i 1_{\xi_i \leq s} + 1_{\xi_i > s} (1 - e^{-q x_i (t-s)}) \\ &\leq Y^\circ(s) + \sum_{i=m+1}^n y_i (1 - e^{-q(t-s)x_i}) \\ &\leq Y^\circ(s) + q(t-s) \sum_{i=m+1}^n x_i y_i \end{aligned}$$

Moreover, we can bound the expectation of this supermartingale as follows:

$$\begin{aligned} \left| \mathbb{E}(Y^\circ(T)) - q \sum_{i=m+1}^n x_i y_i T \right| &= \left| \sum_{i=m+1}^n y_i [(1 - e^{-qT x_i}) - qT x_i] \right| \\ &\leq \sum_{i=m+1}^n y_i \frac{q^2 T^2 x_i^2}{2} \end{aligned}$$

by a simple Taylor bound and so:

$$\left| \mathbb{E}(Y^\circ(T)) - q \sum_{i=m+1}^n x_i y_i T \right| \leq q x_{m+1} T^2 \frac{q \sigma_{1,1}}{2}.$$

Finally, we can also bound its variance:

$$\begin{aligned} \text{Var}(Y^\circ(T)) &= \sum_{i=m+1}^n y_i^2 (1 - e^{-qT x_i}) e^{-qT x_i} \\ &\leq Tq \sum_{i=m+1}^n x_i y_i^2. \end{aligned}$$

Then, using Doob's supermartingale inequality from Lemma 2.54.5 in [27], we have

$$\mathbb{P}\left(\sup_{t \leq T} \left| Z^\circ(t) - q \sum_{i=m+1}^n x_i y_i \right| \geq \epsilon/4\right) = O(q x_{m+1}) + O\left(q \sum_{i=m+1}^n x_i y_i^2\right).$$

As for $m \in \mathbb{N}$ $\lim_{n \rightarrow +\infty} q x_m = c_m$, and as c is in l_{\searrow}^3 , $\lim_{m \rightarrow \infty} \lim_{n \rightarrow +\infty} q x_m = 0$. Hence

$$\lim_{m \rightarrow +\infty} \limsup_{n \rightarrow +\infty} \mathbb{P}\left(\sup_{t \leq T} \left| Z^\circ(t) - q \sum_{i=m+1}^n x_i y_i \right| \geq \epsilon/4\right) = 0. \quad (3.5)$$

Moreover, as $q \sigma_{1,1} \rightarrow \beta + \sum_i c_i e_i$ by hypothesis, $q \sum_{i=m+1}^n x_i y_i \rightarrow \sum_{i=m+1}^n c_i e_i + \beta$ for m fixed, so

$$\lim_{m \rightarrow +\infty} \limsup_{n \rightarrow +\infty} 1_{T|q \sum_{i=m+1}^n x_i y_i - \beta| \geq \epsilon/4} = 0. \quad (3.6)$$

As a consequence of (3.3), (3.4) (3.5) and (3.6), we get that

$$\lim_n \mathbb{P}(d_S^T(Y, \mathcal{Y}) \geq \epsilon) = 0$$

and we have the convergence in probability of $(Y, Z, (r_j^n)_{n \in \mathbb{N}}, (l_j^n)_{n \in \mathbb{N}})$, which yields the joint convergence of the Theorem.

3.4 Proof of Theorem 3.1.3

The idea to retrieve $(\sum_{i: X_i \subset \mathcal{X}_j^n} y_i)_{j \in \mathbb{N}}$ from $(\sum_i y_i 1_{\xi_i \leq t})_{t \in \mathbb{R}_+}$ is pretty straightforward once one realizes that the components are jointly explored: if i and j are in the same component, all k such that $\xi_i \leq \xi_k \leq \xi_j$ are in this component. Hence, we get that for each $j \in \mathbb{N}$, we have:

$$\sum_{i: X_i \subset \mathcal{X}_j^n} y_i = Y(r_j^n) - Y(l_j^n).$$

Remark that, as $\mathcal{Y}(s) = \sum_i e_i 1_{\zeta_i \leq s} + \beta s$, any point s that is a discontinuity for \mathcal{Y} is also a discontinuity for \mathcal{Z} . Indeed, if $c_i = 0$, $\zeta_i = +\infty$ a.s., so it is a jump neither of \mathcal{Y} nor of \mathcal{Z} . If there is a jump at time s for \mathcal{Y} , there is $c_i > 0$ such that $\zeta_i = s$ and then there is a jump at time s for \mathcal{Z} . On the other hand, according to Lemma 4.4 in [14], \mathcal{Z} is continuous at the points l_j and r_j for $j \in \mathbb{N}$. As a consequence, \mathcal{Y} is continuous at those points, and, as we have the joint convergence, for all $j \in \mathbb{N}$,

$$\begin{aligned} Y(r_j^n) - Y(l_j^n) &\xrightarrow[n \rightarrow \infty]{(d)} \mathcal{Y}(r_j) - \mathcal{Y}(l_j) \\ &= \sum_{i \in \mathbb{N}} e_i 1_{\zeta_i \in (l_j, r_j]} + \beta(l_j - r_j) \\ &= \sum_{i \in \mathbb{N}} e_i 1_{\zeta_i \in (l_j, r_j]} + \beta |\mathcal{X}_j| \end{aligned}$$

which is the first convergence. In particular, by taking $y_i = (\frac{x_i}{\sigma_2})^r$, we have

$$\forall i \in \mathbb{N}, \frac{x_i}{\sigma_2} \rightarrow c_i$$

so $e_i = c_i^r$. Moreover, as $q\sigma_2 \rightarrow 1$,

$$\begin{aligned} q\sigma_{1,1} &= \frac{q}{\sigma_2^r} \sigma_{r+1} \\ &\sim \frac{\sigma_{r+1}}{\sigma_2^{r+1}} \end{aligned}$$

and for $m \in \mathbb{N}$,

$$\begin{aligned} q \sum_{i>m} x_i y_i^2 &= q \sum_{i>m} \frac{x_i^{2r+1}}{\sigma_2^{2r}} \\ &\sim \sum_{i>m} \frac{x_i^{2r+1}}{\sigma_2^{2r+1}} \\ &\leq \frac{x_m^{2r-2}}{\sigma_2^{2r-2}} \sum_{i>m} \frac{x_i^3}{\sigma_2^3} \\ &\leq \frac{x_m^{2r-2}}{\sigma_2^{2r-2}} \frac{\sigma_3}{\sigma_2^3} \\ &= O\left(\frac{x_m^{2r-2}}{\sigma_2^{2r-2}}\right). \end{aligned}$$

Because $r > 0$, as $\lim_{n \rightarrow \infty} \frac{x_m^{2r-2}}{\sigma_2^{2r-2}} = c_m^{2r-2}$, $\lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} q \sum_{i>m} x_i y_i^2 = 0$. If $r = 2$, $q\sigma_{1,1} \sim \frac{\sigma_3}{\sigma_2^3} \rightarrow \alpha + \sum_i c_i^3$ so $\beta = \alpha$ which gives the second convergence of the theorem.

If $r > 2$, we have

$$\left| q\sigma_{1,1} - \sum_i c_i^{r+1} \right| \leq \left| q\sigma_{1,1} - \frac{\sigma_{r+1}}{\sigma_2^{r+1}} \right| + \sum_i \left| \frac{x_i^{r+1}}{\sigma_2^{r+1}} - c_i^{r+1} \right|$$

and we already saw that $q\sigma_{1,1} \sim \frac{\sigma_{r+1}}{\sigma_2^{r+1}}$, furthermore, $\frac{\sigma_{r+1}}{\sigma_2^{r+1}} \leq \frac{x_0^{r-2}}{\sigma_2^{r-2}} \frac{\sigma_3}{\sigma_2^3} \rightarrow c_0(\alpha + \sum_i c_i^3)$. So $|q\sigma_{1,1} - \frac{\sigma_{r+1}}{\sigma_2^{r+1}}| = o(1)$. Moreover, for $m \in \mathbb{N}$,

$$\sum_i \left| \frac{x_i^{r+1}}{\sigma_2^{r+1}} - c_i^{r+1} \right| = \sum_{i=0}^m \left| \frac{x_i^{r+1}}{\sigma_2^{r+1}} - c_i^{r+1} \right| + \frac{x_m^{r-2}}{\sigma_2^{r-2}} \sum_{i>m} \frac{x_i^3}{\sigma_2^3} + \sum_{i>m} c_i^{r+1}$$

so

$$\limsup_{n \rightarrow \infty} \sum_i \left| \frac{x_i^{r+1}}{\sigma_2^{r+1}} - c_i^{r+1} \right| \leq c_m^{r-2} (\alpha + \sum_i c_i^3) + \sum_{i>m} c_i^{r+1}.$$

As $r > 2$ and $c \in \ell^3$ as it is true for all $m \in \mathbb{N}$, we get

$$\limsup_{n \rightarrow \infty} \sum_i \left| \frac{x_i^{r+1}}{\sigma_2^{r+1}} - c_i^{r+1} \right| = 0.$$

As a consequence, if $r > 2$, $\beta = 0$.

Chapter 4

Truncation and Gromov-weak topology

4.1 Motivations

Recall that non trivial eternal multiplicative coalescents are defined on $l_{\searrow}^2 \setminus l_{\searrow}^1$. In particular, for $x \in l_{\searrow}^2 \setminus l_{\searrow}^1$, for any time interval $[t, t']$, $t < t'$, the set of edges added by the coalescence on a bounded time interval $[t, t']$ is almost surely infinite. For this reason, it is often more convenient to work with $x \in l_{\searrow}^1$. The question which arises then is: can we infer properties of the coalescence on l_{\searrow}^2 from the coalescence on l_{\searrow}^1 ? In his proof of the Feller property, Aldous had to deal with this very issue in [5] and developed a method to approximate coalescence on l_{\searrow}^2 by coalescence on finite weight sequences. For the same reasons, in [28], Rossignol had to improve on this work and establish what conditions make $\text{Coal}(\bar{X}, q)$ a good approximation of $\text{Coal}(X, q)$, when \bar{X} is a truncation of X . In the following section, we expose his result through the lens of truncated weight sequences and in the subsequent sections, we will see how to find good truncations when $q_n \xrightarrow{n \rightarrow +\infty} -\infty$.

4.2 Structural result

All the results in this section come from Rossignol in [28]. More specifically, we pick apart Lemma 5.5 in [28] to emphasize the following heuristic: let $x \in l_{\searrow}^2$ and \bar{x} be a truncation of x , $q \in \mathbb{R}$, then with some conditions on \bar{x} , all the complexity of the largest component of $\text{gCoal}(x, q)$ comes from the coalescence of \bar{x} .

In the course of formalizing and proving this heuristic, we will extensively use the following obvious result:

Lemma 4.2.1 (Staged coalescence principle). *Let I and J be two disjoint subset of \mathbb{N} . Let $x_I = (x_i)_{i \in I}$ and $x_J = (x_j)_{j \in J}$ be two sequences of non negatives weights. For each $t \geq 0$, let $g_I(t) = \text{gCoal}(x_I, t)$ and $g_J(t) = \text{gCoal}(x_J, t)$. Let $(E_{i,j})_{i \in I, j \in J}$ be a sequence of independent exponential random variables such that for each $(i, j) \in I \times J$,*

$$\mathbb{E}(E_{i,j}) = \frac{1}{x_i x_j}.$$

For all $t \geq 0$ let $g_{IJ}(t)$ be the graph on $I \sqcup J$ where there is an edge between i and j for each one of the following cases:

- *If i and $j \in I$, there is an edge between i and j in $\text{gCoal}((x)_{k \in I}, t)$,*
- *If i and $j \in J$, there is an edge between i and j in $\text{gCoal}((x)_{k \in J}, t)$,*
- *If $i \in I$ and $j \in J$, there is an edge between i and j if $E_{i,j} \leq t$.*

Then for all $t \in \mathbb{R}$,

$$g_{IJ}(t) \stackrel{(d)}{=} \text{Coal}(x_{IJ}, t)$$

where $x_{IJ} = (x_i)_{i \in I \sqcup J}$.

Uniformly this lemma says that if one wants to coalesce a given input, one can proceed as follows:

- (i) Divide the input into smaller parts,

- (ii) Coalesce each part independently,
- (iii) Independently coalesce the results of step (ii).

For $x \in \mathcal{L}_{\searrow}^2$, recall from Definition 2.2.13 the definitions of a truncation \bar{x} of x , and the rest \hat{x} associated with this truncation. First we want to make sure, as previously stated, that when coalescing the rest of a truncation, the process doesn't add complexity to already existing large components.

Definition 4.2.2. Let $x \in \mathcal{L}_{\searrow}^2$, $q \in \mathbb{R}_+$ and $j \in \mathbb{N}$, we denote by $\text{comp}_j(x, q)$ the connected component of j in $\text{gCoal}(x, q)$ and $\text{comp}(x, q)$ the set of all the connected components of $\text{gCoal}(x, q)$.

Lemma 4.2.3 (Part of Lemma 5.5 in [28]). Let $x \in \mathcal{L}_{\searrow}^2$, \bar{x} a truncation of x and \hat{x} the rest associated with this truncation. Let A be the event that there exists a vertex i in the rest which has two neighbours in $\text{gCoal}(x, q)$ that stay in the same component after removing i :

$$A = \left\{ \exists i, j, k \in \mathbb{N}, \bar{x}_i = 0, j \neq k : \left\{ \begin{array}{l} (i, j), (i, k) \in E(x, q) \\ \text{comp}_j(x \setminus \{x_i\}, q) = \text{comp}_k(x \setminus \{x_i\}, q) \end{array} \right\} \right\}.$$

We have

$$\mathbb{P}(A \cap \{S(x, q) \leq K\}) \leq q^2 \hat{\sigma}_2 K.$$

Proof. We define for each $i \in \mathbb{N}$:

$$A_i = \left\{ \exists j, k \in \mathbb{N} : \left\{ \begin{array}{l} (i, j), (i, k) \in E(x, q) \\ \text{comp}_j(x \setminus \{x_i\}, q) = \text{comp}_k(x \setminus \{x_i\}, q) \end{array} \right\} \right\}.$$

so that

$$A = \bigcup_{i: \bar{x}_i = 0} A_i.$$

Here we remember Lemma 4.2.1 and start by coalescing only $x \setminus \{x_i\}$:

$$\begin{aligned} \mathbb{P}(A_i | \text{gCoal}(x \setminus \{x_i\}, q)) &\leq q^2 \sum_{C \in \text{comp}(x \setminus \{x_i\}, q)} x_i^2 |C|^2 \\ &= q^2 x_i^2 S(x \setminus \{x_i\}, q). \end{aligned}$$

so

$$\mathbb{P}(A_i | S(x \setminus \{x_i\}, q)) \leq q^2 x_i^2 S(x \setminus \{x_i\}, q)$$

As a consequence for $K > 0$,

$$\begin{aligned} \mathbb{P}(A \cap \{S(x, q) \leq K\}) &\leq \sum_{i: \bar{x}_i = 0} \mathbb{P}(A_i \cap \{S(x \setminus \{x_i\}, q) \leq K\}) \\ &\leq \sum_{i: \bar{x}_i = 0} \mathbb{E}(\mathbb{P}(A_i | S(x \setminus \{x_i\}, q)) \mathbf{1}_{\{S(x \setminus \{x_i\}, q) \leq K\}}) \\ &\leq \sum_{i: \bar{x}_i = 0} \mathbb{E}(q^2 x_i^2 S(x \setminus \{x_i\}, q) \mathbf{1}_{\{S(x \setminus \{x_i\}, q) \leq K\}}) \\ &= \sum_{i: \bar{x}_i = 0} q^2 x_i^2 K \\ &= q^2 \hat{\sigma}_2 K. \end{aligned}$$

□

Now if $\hat{\sigma}_2$ is small enough, we know that large components of the coalescence of the truncation are with large probability not impacted in their complexity by coalescing the whole sequence, but we want to be sure that coalescing the rest will not create large components on its own.

Lemma 4.2.4 (Part of Lemma 5.5 in [28]). Let $x \in \mathcal{L}_{\searrow}^2$, \bar{x} be a truncation of x and \hat{x} the rest associated with this truncation. Let $\epsilon > 0$,

$$\mathbb{P}(\exists C \in \text{comp}(\hat{x}, q), |C| \geq \epsilon) \leq \frac{q\epsilon^2 \hat{\sigma}_2}{\epsilon^2 - \hat{\sigma}_2}.$$

Proof. On the event

$$\{\exists C \in \text{comp}(\hat{x}, q), |C| \geq \epsilon\}$$

we have in particular that $S(\hat{x}, q) \geq \epsilon^2$. As a consequence and according to Lemma 2.2.4,

$$\mathbb{P}(\exists C \in \text{comp}(\text{gCoal}(\hat{x}, q)), |C| \geq \epsilon) \leq \mathbb{P}(S(\hat{x}, q) \geq \epsilon^2) \leq \frac{q\epsilon^2\hat{\sigma}_2}{\epsilon^2 - \hat{\sigma}_2}$$

□

This lemma has one major shortcoming: it says that coalescing only the rest will not create large components, however, it might be the case that coalescing small component of $\text{gCoal}(\bar{x}, q)$ with components of $\text{gCoal}(\hat{x}, q)$ creates large components. Recall from Lemma 2.2.14 that we can control $S(x, q) - S(\bar{x}, q)$. Then as $|\bar{C}|(|C| - |\bar{C}|) \leq S(x, q) - S(\bar{x}, q)$ for $\bar{C} \in \text{gCoal}(\bar{x}, q)$ and $\bar{C} \subset C \in \text{comp}(\text{gCoal}(x, q))$, it might be a way to address this shortcoming. However, this inequality is only useful if we already know that \bar{C} is not too small. We deal with this issue by introducing a subtruncation $\bar{\bar{x}}$. Of course we will write $\bar{\bar{\sigma}}_2 = \sigma_2(\bar{\bar{x}})$ and $\hat{\bar{\sigma}}_2 = \sigma_2(\hat{\bar{x}})$ with $\hat{\bar{x}}$ the rest associated with the truncation $\bar{\bar{x}}$.

Lemma 4.2.5 (Part of Lemma 5.5 from [28]). *Let $x \in \mathcal{l}_{\leq}^2$, \bar{x} be a truncation of x and $\bar{\bar{x}}$ a truncation of \bar{x} . For $\epsilon, \eta > 0$, let $B_{\epsilon, \eta}$ be such that:*

$$B_{\epsilon, \eta} = \left\{ \exists C \in \text{comp}(x, q) : \begin{cases} |C| \geq \eta, \\ \forall \bar{C} \in \text{comp}(\bar{x}, q), \bar{C} \subset C \Rightarrow |\bar{C}| \leq |C| - \epsilon. \end{cases} \right\}$$

Then

$$\mathbb{P}(B_{\epsilon, \eta} \cap \{S(x, q) \leq K\}) \leq \frac{q\eta^2\hat{\bar{\sigma}}_2}{\eta^2 - \hat{\bar{\sigma}}_2} + (\epsilon\bar{\bar{x}}_{\min})^{-1}(1 + q(K + \epsilon\bar{\bar{x}}_{\min}))^2 2\hat{\bar{\sigma}}_2 + 2q\hat{\bar{\sigma}}_2.$$

Proof. We have that:

$$B_{\epsilon, \eta} \subset B_{\eta}^1 \cup B_{\epsilon, \eta}^2$$

where

$$B_{\eta}^1 = \left\{ \exists C \in \text{comp}(x, q) : \begin{cases} |C| \geq \eta, \\ \forall \bar{C} \in \text{comp}(\bar{x}, q), \bar{C} \subset C, \Rightarrow |\bar{C}| < \bar{\bar{x}}_{\min}. \end{cases} \right\}$$

and

$$B_{\epsilon, \eta}^2 = \left\{ \exists C \in \text{comp}(x, q) : \begin{cases} |C| \geq \eta, \\ \forall \bar{C} \in \text{comp}(\bar{x}, q), \bar{C} \subset C, \Rightarrow |\bar{C}| \leq |C| - \epsilon, \\ \exists \bar{C} \subset C, \bar{\bar{x}}_{\min} \leq |\bar{C}|. \end{cases} \right\}$$

On B_{η}^1 , there is a component of $\text{gCoal}(x, q)$ larger than η that contains no vertex of $\bar{\bar{x}}$. Thus, on this event there is a component of $\text{gCoal}(\bar{x}, q)$ larger than η . This, from Lemma 4.2.4, has a probability controlled as follows:

$$\mathbb{P}(B_{\eta}^1) \leq \frac{q\eta^2\hat{\bar{\sigma}}_2}{\eta^2 - \hat{\bar{\sigma}}_2}.$$

On the other hand, we have, for $C \in \text{comp}(x, q)$ and $\bar{C} \in \text{comp}(\bar{x}, q)$ such that $C \subset \bar{C}$, $|\bar{C}|(|C| - |\bar{C}|) \leq S(x, q) - S(\bar{x}, q)$ for $\bar{C} \in \text{gCoal}(\bar{x}, q)$. On $B_{\epsilon, \eta}^2$ we have thus that $S(x, q) - S(\bar{x}, q) \geq \bar{\bar{x}}_{\min}\epsilon$. Thus, according to Lemma 2.2.14:

$$\mathbb{P}(B_{\epsilon, \eta}^2 | S(\bar{x}, q), S(\hat{x}, q)) \leq (\epsilon\bar{\bar{x}}_{\min})^{-1}(1 + q(S(\bar{x}, q) + \epsilon\bar{\bar{x}}_{\min}))^2 S(\hat{x}, q).$$

As a consequence, using Lemma 2.2.4 with \hat{x} ,

$$\begin{aligned} \mathbb{P}(B_{\epsilon, \eta}^2 \cap \{S(x, q) \leq K\}) &\leq \mathbb{P}(B_{\epsilon, \eta}^2 \cap \{S(x, q) \leq K\} \cap \{S(\hat{x}, q) \leq 2\hat{\bar{\sigma}}_2\}) + \mathbb{P}(S(\hat{x}, q) > 2\hat{\bar{\sigma}}_2) \\ &\leq \mathbb{P}(B_{\epsilon, \eta}^2 \cap \{S(\bar{x}, q) \leq K\} \cap \{S(\hat{x}, q) \leq 2\hat{\bar{\sigma}}_2\}) + 2q\hat{\bar{\sigma}}_2 \\ &= \mathbb{E}(\mathbb{P}(B_{\epsilon, \eta}^2 | S(\bar{x}, q), S(\hat{x}, q)) 1_{S(\hat{x}, q) \leq 2\hat{\bar{\sigma}}_2} 1_{S(\bar{x}, q) \leq K}) + 2q\hat{\bar{\sigma}}_2 \\ &\leq (\epsilon\bar{\bar{x}}_{\min})^{-1}(1 + q(K + \epsilon\bar{\bar{x}}_{\min}))^2 2\hat{\bar{\sigma}}_2 + 2q\hat{\bar{\sigma}}_2. \end{aligned}$$

□

Let us take a step back and analyze the inequality of Lemma 4.2.5. It appears that in order to make the event improbable we need that both $q\widehat{\sigma}_2$ and, granted that $S(x, q)$ were tight, $\frac{q^2\widehat{\sigma}_2}{x_{\min}}$ are small. This leads to the following definition:

Definition 4.2.6. Let $x = (x^n)_{n \in \mathbb{N}} \in sl_2$, $q = (q_n)_{n \in \mathbb{N}}$ and $\bar{x} = (\bar{x}^n)_{n \in \mathbb{N}}$ be a truncation of X . We say that \bar{x} is suitable if there exists $\overline{\bar{x}}$ a truncation of \bar{x} such that:

$$\widehat{\sigma}_2 = o(q^{-1}) \text{ and } \frac{\widehat{\sigma}_2}{\overline{\bar{x}}_{\min}} = o(q^{-2}).$$

Lemma 4.2.7 (Part of Lemma 5.5 from [28]). Let $X = (X^n)_{n \in \mathbb{N}}$ be a sequence of $\mathbb{M}_2^{\mathbb{N}}$ and $\overline{X} = (\overline{X}^n)_{n \in \mathbb{N}}$ be a truncation of X . Suppose that \bar{x} , the sequence of weight sequences of \overline{X} is a suitable truncation of X , the sequence of weight sequences of x . Suppose moreover that $S(x, q)$ is tight.

For $l \in \mathbb{N}$, Let (C_0, \dots, C_l) be the $l+1$ largest components of $\text{Coal}(X, q)$ and (d^0, \dots, d^l) the associated distances. For $\epsilon, \eta > 0$ Let $\Omega_{\epsilon, \eta}^l$ be the following event:

$$\Omega_{\epsilon, \eta}^l = \bigcap_{j \in [l]} \left\{ |C_j| \leq \eta \text{ or } \exists (\overline{C}_j, \overline{d}^j) \in \text{comp}(\text{Coal}(\overline{X}, q)) : \overline{C}_j \subset C_j, d_{|\overline{C}_j}^j = \overline{d}^j \text{ and } |C_j| - |\overline{C}_j| \leq \epsilon \right\}.$$

Then

$$\mathbb{P}(\Omega_{\epsilon, \eta}^l) \xrightarrow{n \rightarrow +\infty} 1.$$

Proof. We have that, with A the event of Lemma 4.2.3 and $B_{\epsilon, \eta}$ the event of Lemma 4.2.5, because \bar{x} is a suitable truncation of x ,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(A \cup B_{\epsilon, \eta}) \leq \lim_{K \rightarrow +\infty} \limsup_{n \rightarrow +\infty} (\mathbb{P}((A \cup B_{\epsilon, \eta}) \cap \{S(x, q) \leq K\}) + \mathbb{P}(S(x, q) > K)) = 0.$$

Let (C, d) be one of the (C_0, \dots, C_l) . On the event $\neg A \cap \neg B_{\epsilon, \eta}$, we have that either $|C| \leq \eta$, either there exists $\overline{C} \subset C$ that is such that $(\overline{C}, \overline{d}) \in \text{comp}(\text{Coal}(\overline{X}, q))$. Moreover, we have that $d_{|\overline{C}} \leq \overline{d}$. Suppose that $d_{|\overline{C}} \neq \overline{d}$. Then there exists a simple path of components (i_1, \dots, i_k) , a sequence of entry points (u_1, \dots, u_k) and a sequence of exit points (v_1, \dots, v_k) such that

$$u_j, v_j \in X_{i_j}$$

$$d(v_i, u_{i+1}) = 0$$

$$u_1 = u, v_k = v$$

$$\sum_{j=1}^k d_{X_{i_j}}(u_j, v_j) < \overline{d}(u, v).$$

If this is the case, it is necessary that there exists $j \in [k]$ such that X_{i_j} is not a subset of \overline{C} and therefore not in \overline{X} . This implies that i_j has two neighbours in a component of $\text{gCoal}(x \setminus i_j, q)$. This cannot be the case on $\neg A$. \square

4.3 Bounding the differences between truncated and non truncated coalescent

Lemma 4.3.1. *Under the hypothesis of Lemma 4.2.7, if we suppose that*

$$\text{Coal}(\bar{X}^n, q_n) \xrightarrow[n \rightarrow +\infty]{(d)} Y$$

and if almost surely, no two components of Y have the same weight, then

$$\text{Coal}(X^n, q_n) \xrightarrow[n \rightarrow +\infty]{(d)} Y$$

for the Gromov-weak product topology. Moreover, recall from Section 2.3.3 the definition of $m_\delta^k(\cdot)$ and assume in addition that

$$\forall \delta > 0, \forall k \in \mathbb{N}, \liminf_{n \rightarrow +\infty} \mathbb{P}(m_\delta^k(\text{Coal}(X^n, q_n)) \geq \epsilon) \xrightarrow{\epsilon \rightarrow 0} 1.$$

Then

$$\text{Coal}(X^n, q_n) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} Y$$

occurs for the Gromov-Hausdorff-Prokhorov product topology.

Proof. We suppose that, thanks to Skorokhod's representation theorem, the convergence

$$\text{Coal}(\bar{X}^n, q_n) \xrightarrow[n \rightarrow +\infty]{} Y$$

occurs almost surely for the Gromov-weak topology. Let $l \in \mathbb{N}$, (C_0, \dots, C_l) be the $l+1$ largest components of $\text{Coal}(X, q)$. Let $\epsilon, \delta > 0$. In the rest of the proof, we consider that the event $\Omega_{\epsilon, \eta}^l$ occurs. Notice that either $|C_l| < \eta$, either for all $i \in [l]$, $|C_l| \geq \eta$.

At this point, we want to know whether or not $(\bar{C}_i)_{i \in [l]}$ are the ordered $l+1$ largest components of $\text{Coal}(\bar{X}, q)$. First if there exists \bar{C}_{l+1} such that $|\bar{C}_{l+1}| \geq |\bar{C}_l|$, as $|\bar{C}_l| \geq |C_l| - \epsilon$, $|\bar{C}_{l+1}| \geq |C_l| - \epsilon$, and as C_{l+1} is a subset of a component of $\text{Coal}(X, q)$ smaller than $|C_l|$, $|\bar{C}_l| - |\bar{C}_{l+1}| \leq \epsilon$. Second if $(\bar{C}_i)_{i \in [l]}$ are in fact the largest components of $\text{Coal}(\bar{X}, q)$, but there exists $i < j \leq l$ such that $|\bar{C}_i| \leq |\bar{C}_j|$, then as $|C_i| \geq |C_j| \geq |\bar{C}_j|$, and $|\bar{C}_i| \geq |C_i| - \epsilon$, this implies that $||\bar{C}_i| - |\bar{C}_j|| \leq \epsilon$. As a consequence, on the event where

$$\left\{ \forall i, j, 0 \leq i < j \leq l+1, ||\bar{C}_i| - |\bar{C}_j|| > \epsilon \right\},$$

we in fact have that $(\bar{C}_i)_{i \in [l]}$ are the ordered $l+1$ largest components of $\text{Coal}(\bar{X}, q)$.

Notice that for all $\Phi \in \mathbb{F}$, if for $i \in [l]$, $C = C_i$ and $\bar{C} = \bar{C}_i$,

$$\Phi(C) - \Phi(\bar{C}) = \int_{C^k \setminus \bar{C}^k} \phi((d(x_i, x_j))_{i, j \leq k}) \prod_{i=1}^k \mu_i(dx_i) \leq \|\phi\|_\infty k \epsilon$$

due to the fact that on $\Omega_{\epsilon, \eta}^l$, $d_{|\bar{C}|} = \bar{d}$ and $|C| - |\bar{C}| \leq \epsilon$.

As a consequence,

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i \in [l]} \sup_{\Phi \in \mathbb{F}: k \|\phi\|_\infty \leq 1} |\Phi(C_i) - \Phi(\bar{C}_i)| > \epsilon\right) &\leq \mathbb{P}(\neg \Omega_{\epsilon, \eta}^l) \\ &+ \mathbb{P}(\exists i, j, \text{ s.t. } 0 \leq i < j \leq l+1, ||\bar{C}_i| - |\bar{C}_j|| \leq \epsilon) \\ &+ \mathbb{P}(|\bar{C}_l| \leq \eta). \end{aligned}$$

Finally, as $\text{Coal}(\bar{X}^n, q_n) \xrightarrow[n \rightarrow +\infty]{} Y$ for the Gromov-weak product topology, in particular, the masses $(|\bar{C}_0|, \dots, |\bar{C}_l|)$ converges. As a consequence,

$$\lim_{\epsilon \rightarrow +\infty} \lim_{n \rightarrow \infty} \mathbb{P}(\exists i, j, \text{ s.t. } 0 \leq i < j \leq l+1, ||\bar{C}_i| - |\bar{C}_j|| \leq \epsilon) = 0$$

since almost surely no two components of Y have the same weight. As $\mathbb{P}\left(\bigcup_{i \in [l]} \sup_{\Phi \in \mathbb{F}: k \|\phi\|_\infty \leq 1} |\Phi(C_i) - \Phi(\bar{C}_i)| > \epsilon\right)$ is decreasing in ϵ , by taking the limit in $n \rightarrow \infty$, then in $\eta \rightarrow 0$ and then $\epsilon \rightarrow 0$, we get for all $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\bigcup_{i \in [l]} \sup_{\Phi \in \mathbb{F}: k \|\phi\|_\infty \leq 1} |\Phi(C_i) - \Phi(\bar{C}_i)| > \epsilon\right) = 0.$$

□

4.4 Existence of a suitable truncation

Now we know that the definition of a suitable truncation is enough to guarantee the convergence of the coalescence from the convergence of the coalescence of the truncation. Next we analyze how restrictive this definition is. First remark that for a given sequence $x = (x^n)_{n \in \mathbb{N}} \in sl_2$ and $q = (q_n)_{n \in \mathbb{N}}$, there always exists a suitable truncation. In fact, it suffices to choose first \bar{x} such that $q^2 \hat{\sigma}_2 = o(1)$ and then augment the truncation into \bar{x} that is such that $\frac{\hat{\sigma}_2}{\bar{x}_{\min}} = o(q^{-2})$. However, when adding constraints on the truncation, there is a bit more work needed to guarantee its existence. The following lemma gives a suitable truncation that satisfies a constraint that will be useful later.

Lemma 4.4.1. *Let $x = (x^n)_{n \in \mathbb{N}} \in sl_2$ and $q = (q_n)_{n \in \mathbb{N}}$ such that $q\sigma_2 = O(1)$. If*

$$\exists \epsilon \in (0, 2), \lim_{n \rightarrow +\infty} \sigma_{2-\epsilon} = 0,$$

then

there exists a suitable truncation \bar{x} and $r > 0$, such that $\frac{\bar{\sigma}_2^r}{\bar{x}_{\min}} = o(1)$.

Proof. Let $\epsilon \in (0, 2)$ be such that $\sigma_{2-\epsilon} = o(1)$. In particular $\sigma_2 \leq (\sigma_{2-\epsilon})^{\frac{2}{2-\epsilon}}$. We define \bar{x} as the truncation such that

$$\forall i \in \mathbb{N}, \bar{x}_i > 0 \Leftrightarrow x_i \geq \sigma_2^{2/\epsilon^2 + 2/\epsilon}.$$

Then

$$\hat{\sigma}_2 = \sum_i x_i^2 1_{x_i < \sigma_2^{2/\epsilon^2 + 2/\epsilon}} \leq \sigma_2^{2/\epsilon^2 + 2/\epsilon} \sigma_{2-\epsilon}.$$

As a consequence, $\sigma_2 \sim \bar{\sigma}_2$. We then define \bar{x} as the truncation such that

$$\forall i \in \mathbb{N}, \bar{x}_i > 0 \Leftrightarrow x_i \geq \sigma_2^{2/\epsilon}.$$

We get that

$$q^2 \hat{\sigma}_2 \leq q^2 \sigma_2^2 \sigma_{2-\epsilon}.$$

On the other hand, for n large enough as $\sigma_2 \leq x_0^\epsilon \sigma_{2-\epsilon} \leq x_0^\epsilon$ we have that $x_0 \geq \sigma_2^{1/\epsilon} \geq \sigma_2^{2/\epsilon}$. As a consequence, for n large enough \bar{x}_{\min} exists and

$$\frac{q^2(\sigma_2 - \bar{\sigma}_2)}{\bar{x}_{\min}} \leq q^2 \sigma_2^{2/\epsilon^2 + 2/\epsilon} \sigma_{2-\epsilon} \sigma_2^{-2/\epsilon} = q^2 \sigma_2^2 \sigma_{2-\epsilon}.$$

Finally, as \bar{x}_{\min} exists, \bar{x}_{\min} exists too and

$$\frac{\bar{\sigma}_2^{2/\epsilon^2 + 2/\epsilon}}{\bar{x}_{\min}} \leq 1$$

so we can in fact find $r > 2/\epsilon^2 + 2/\epsilon$ such that $\frac{\bar{\sigma}_2^r}{\bar{x}_{\min}} = o(1)$. □

Of course we also need to know that the condition to be a suitable truncation also preserves the Aldous-Limic conditions 1, which is the case:

Lemma 4.4.2. *Let $(x, q) = (x^n, q_n)_{n \in \mathbb{N}}$ satisfy the Aldous-Limic conditions 1 with parameters $(\alpha, \lambda, c) \in \mathcal{I}$. Let $\bar{x} = (\bar{x}^n)_{n \in \mathbb{N}}$ be a suitable truncation of x . Then (\bar{x}, q) satisfies the Aldous-Limic conditions 1 with the same parameters (α, λ, c) .*

Proof. First, of course $\bar{\sigma}_2 \leq \sigma_2 \xrightarrow{n \rightarrow +\infty} 0$. Moreover, as $\bar{\sigma}_2 \sim \sigma_2$, for $i \in \mathbb{N}$,

$$c_i = \lim_{n \rightarrow +\infty} \frac{x_i}{\sigma_2} = \lim_{n \rightarrow +\infty} \frac{\bar{x}_i}{\bar{\sigma}_2}.$$

In fact, if that was not the case, at least along a subsequence $(\phi(n))_{n \in \mathbb{N}}$, we would have the existence of $m \in \mathbb{N}$ such that for all $n \geq m$, for all $j \geq i$, $\bar{x}_j^{\phi(n)} = 0$ and then $\frac{\bar{\sigma}_2}{\sigma_2} = \sum_{k=0}^{i-1} \frac{x_k^2}{\sigma_2} \sim \sigma_2 \sum_{k=0}^{i-1} c_k = o(1)$ which would contradict $\bar{\sigma}_2 \sim \sigma_2$. Now, the more difficult ones. On one hand, $q - \frac{1}{\bar{\sigma}_2} \leq q - \frac{1}{\sigma_2}$ so

$$\limsup_{n \rightarrow +\infty} q - \frac{1}{\bar{\sigma}_2} \leq \lambda$$

On the other hand,

$$\begin{aligned} q - \frac{1}{\bar{\sigma}_2} &= q - \frac{1}{\sigma_2} - \left(\frac{1}{\bar{\sigma}_2} - \frac{1}{\sigma_2} \right) \\ &= q - \frac{1}{\sigma_2} - \frac{1}{\sigma_2 \bar{\sigma}_2} (\sigma_2 - \bar{\sigma}_2) \\ &= q - \frac{1}{\sigma_2} - \frac{\hat{\sigma}_2}{\sigma_2 \bar{\sigma}_2}. \end{aligned}$$

As $q - \frac{1}{\sigma_2} \xrightarrow{n \rightarrow +\infty} \lambda$ and $\frac{\hat{\sigma}_2}{\sigma_2 \bar{\sigma}_2} \sim \frac{\hat{\sigma}_2}{\sigma_2^2} \xrightarrow{n \rightarrow +\infty} 0$, this shows that $q - \frac{1}{\bar{\sigma}_2} \xrightarrow{n \rightarrow +\infty} \lambda$. Finally, we want to show that

$$\frac{\bar{\sigma}_3}{\bar{\sigma}_2^3} \sim \frac{\sigma_3}{\sigma_2^3}.$$

As $\bar{\sigma}_2 \sim \sigma_2$, we only need to show that

$$\frac{\bar{\sigma}_3}{\sigma_3} \rightarrow 1.$$

We have that

$$\frac{\bar{\sigma}_3}{\sigma_3} = 1 - \frac{\hat{\sigma}_3}{\sigma_3}$$

and

$$\begin{aligned} \frac{\hat{\sigma}_3}{\sigma_3} &\leq x_0 \frac{\hat{\sigma}_2}{\sigma_3} \\ &\sim O\left(\frac{x_0 \hat{\sigma}_2}{\sigma_2 \sigma_2^2}\right). \end{aligned}$$

By hypothesis, $\frac{\hat{\sigma}_2}{\sigma_2} \rightarrow 0$ and $\frac{x_0}{\sigma_2} = O(1)$, which shows that

$$\bar{\sigma}_3 \sim \sigma_3.$$

□

Chapter 5

From simple to complete coalescence

5.1 Bounding the difference between simple and complete coalescences

Recall that Theorem 2.5.12 and Theorem 2.5.11 are stated specifically for the simple coalescence. We show that in our setting, the difference between the two is negligible.

Lemma 5.1.1. *Let $X = (X^n)_{n \in \mathbb{N}}$ be a sequence of $\mathbb{M}_2^{\mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$. For $n \in \mathbb{N}$, let $\mathcal{X}^n = \text{Coal}(X^n, q_n)$. Recall from Definition 2.5.6 the definition of $D = (D_i^n)_{i, n \in \mathbb{N}}$ associated with X . Suppose that for all $j \in \mathbb{N}$,*

$$\sup_{i: X_i^n \subset \mathcal{X}_j^n} D_i^n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0.$$

For $n \in \mathbb{N}$, let $\mathcal{Y}^n = \text{SCoal}(X^n, q_n)$. Then for all $j \in \mathbb{N}$,

$$d_{GHP}(\mathcal{X}^n, \mathcal{Y}^n) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0$$

Proof. Let \mathcal{N} be the number of pairs of vertices that are joint by two edges. Then

$$\mathbb{E}(\mathcal{N}) \leq \sum_{i \neq j \in \mathbb{N}} \frac{(qx_i x_j)^2}{2} = q^2 \sigma_2^2 \rightarrow 1.$$

As a consequence, for $j \in \mathbb{N}$, $\epsilon > 0$, $K \in \mathbb{N}$,

$$\begin{aligned} \mathbb{P}(d_{GHP}(\mathcal{X}_j^n, \mathcal{Y}_j^n) \geq \epsilon) &\leq \mathbb{P}(\mathcal{N} \sup_{i: X_i^n \subset \mathcal{X}_j^n} 2D_i^n \geq \epsilon) \\ &\leq \mathbb{P}(\sup_{i: X_i^n \subset \mathcal{X}_j^n} 2D_i^n \geq \epsilon/K) + \mathbb{P}(\mathcal{N} \geq K) \end{aligned}$$

so the convergence is verified by taking first $n \rightarrow \infty$ and then $K \rightarrow \infty$. \square

Recall from Definition 2.5.6 the definition of ρ_1 . Recall that Theorem 3.2.3 and Theorem 3.2.4 are stated specifically for the simple coalescence with positive edge length $\frac{\sigma_2^2}{\rho_1 + \sigma_2} \text{SCoal}_1(\cdot, \cdot)$. For $X \in \mathbb{M}_2^{\mathbb{N}}$ and $q \in \mathbb{R}_+$,

$$\frac{\sigma_2^2}{\rho_1 + \sigma_2} \text{SCoal}_1(X, q) \stackrel{(d)}{=} \text{SCoal}_{\frac{\sigma_2^2}{\rho_1 + \sigma_2}} \left(\frac{\sigma_2^2}{\rho_1 + \sigma_2} X, q \right)$$

We would like to have a result on $\text{SCoal}(\cdot, \cdot)$ without any scaling inside the coalescence. We can try to pre-scale X by a factor γ , noting that:

$$\frac{\sigma_2^2}{\gamma \rho_1 + \sigma_2} \text{SCoal}_1(\gamma X, q) \stackrel{(d)}{=} \text{SCoal}_{\frac{\sigma_2^2}{\rho_1 + \sigma_2}} \left(\frac{\sigma_2^2}{\rho_1 + \sigma_2} \gamma X, q \right)$$

hoping that $\frac{\sigma_2^2}{\gamma \rho_1 + \sigma_2} \gamma \sim 1$. One way to do this would be to take $\gamma \sim \sigma_2^{-1}$ and $\rho_1 \sim \sigma_2^2$. However, in this case, $\gamma \rho_1 \sim \sigma_2$ would be of the same order than σ_2 , which means that the average distances on the components are of the same order as the length of the edges, which would make the

task of comparing it to $\text{SCoal}_0(X, q)$ hopeless. The trick here is to inflate the input components a bit:

Lemma 5.1.2. *Let $X = (X^n)_{n \in \mathbb{N}}$ be a sequence of $\overline{\mathbb{M}}^{\mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}}$. Recall from Definition 2.5.6 the definition of the associated x , u and ρ_1 . Suppose that*

$$\frac{\rho_1}{\sigma_2^2} \rightarrow 1.$$

Let $\gamma = (\gamma_n)_{n \in \mathbb{N}}$ be such that

$$\gamma \sigma_2 \xrightarrow{n \rightarrow +\infty} +\infty$$

For $i \in \mathbb{N}$, let $Y = \gamma X$. For all $n \in \mathbb{N}$, let $\mathcal{X}^n = \text{SCoal}_0(X^n, q_n)$, $\mathcal{Y}^n = \frac{\sigma_2^2}{\gamma \rho_1 + \sigma_2} \text{SCoal}_1(Y^n, q_n)$ and $G^n = g\text{Coal}(x^n, q_n)$ with the obvious coupling between the three of them. Suppose that for each $j \in \mathbb{N}$, $(\text{Diam}(\mathcal{Y}_j^n))_{n \in \mathbb{N}}$ is tight. Moreover, let

$$L(G_j^n) = \sup\{k \geq 1 : \text{there exists a simple path of length } k \text{ in } G_j^n\}.$$

We suppose that for each $j \in \mathbb{N}$, $\sigma_2 L(G_j^n)$ is tight. Then for all $j \in \mathbb{N}$,

$$d_{\text{GHP}}(\mathcal{X}_j^n, \mathcal{Y}_j^n) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Proof. Let n and $j \in \mathbb{N}$. Consider the obvious correspondence where if $(u_1, u_2) \in \mathcal{X}_j^n \times \mathcal{Y}_j^n$, $u_1 \sim u_2$ if they are the projection of the same point $w \in X^n$ onto \mathcal{X}_j^n and \mathcal{Y}_j^n . To lighten notations, we drop the n in the sequel.

Let $u_1, v_1 \in \mathcal{X}_j$ and $u_2, v_2 \in \mathcal{Y}_j$ be such that $u_1 \sim u_2$ and $v_1 \sim v_2$.

Recall that by definition of simple coalescence, for any $\eta > 0$, there exists $k \in \mathbb{N}$ and two sequences $(a_i)_{1 \leq i \leq k}$ and $(b_i)_{1 \leq i \leq k}$ of X such that:

$$\begin{cases} p^{\mathcal{X}}(a_1) = u_1, p^{\mathcal{Y}}(b_k) = v_1, \\ \forall i \in \{1, \dots, k-1\}, (b_i, a_{i+1}) \in \mathcal{SP}_q \\ \sum_{1 \leq i \leq k} d_X(a_i, b_i) - d_X(u_1, v_1) \leq \eta \end{cases}$$

where $p^{\mathcal{X}}$ and $p^{\mathcal{Y}}$ are the projectors from X respectively Y onto \mathcal{X} and \mathcal{Y} and \mathcal{SP}_q the set of points of X^2 added by the simple coalescence until q . As a consequence,

$$d_{\mathcal{Y}}(u_2, v_2) - d_{\mathcal{X}}(u_1, v_1) \leq \eta + d_{\mathcal{Y}}(u_2, v_2) - \sum_{1 \leq i \leq k} d_X(a_i, b_i).$$

On the other hand, because of the coupling between \mathcal{X} and \mathcal{Y} ,

$$d_{\mathcal{Y}}(u_2, v_2) \leq \frac{\gamma \sigma_2^2}{\gamma \rho_1 + \sigma_2} \sum_{i \leq k} d_X(a_i, b_i) + k \frac{\sigma_2^2}{\gamma \rho_1 + \sigma_2},$$

and as

$$d_{\mathcal{Y}}(u_2, v_2) = d_{\mathcal{Y}}(u_2, v_2) \left(1 - \frac{\gamma \rho_1 + \sigma_2}{\gamma \sigma_2^2}\right) + \frac{\gamma \rho_1 + \sigma_2}{\gamma \sigma_2^2} d_{\mathcal{Y}}(u_2, v_2),$$

we get that

$$d_{\mathcal{Y}}(u_2, v_2) - \sum_{i=1}^k d_X(a_i, b_i) \leq d_{\mathcal{Y}}(u_2, v_2) \left(1 - \frac{\gamma \rho_1 + \sigma_2}{\gamma \sigma_2^2}\right)_+ + k/\gamma.$$

This allows us to compare this quantity with $d_{\mathcal{X}}(u_1, v_1)$:

$$d_{\mathcal{Y}}(u_2, v_2) - d_{\mathcal{X}}(u_1, v_1) \leq \eta + \text{Diam}(\mathcal{Y}_j) \left(1 - \frac{\gamma \rho_1 + \sigma_2}{\gamma \sigma_2^2}\right)_+ + k/\gamma.$$

Of course,

$$\frac{\gamma \rho_1 + \sigma_2}{\gamma \sigma_2^2} = \frac{\rho_1}{\sigma_2^2} + \frac{1}{\gamma \sigma_2} \xrightarrow{n \rightarrow +\infty} 1$$

so the tightness of the diameter assumption guarantees that $\text{Diam}(\mathcal{Y}_j)(1 - \frac{\gamma\rho_1 + \sigma_2}{\gamma\sigma_2^2})$ goes to 0 as n goes to infinity. The conclusion is the same regarding k/γ considering that $k \leq L(G_j^n)$, $\sigma_2 L(G_j^n)$ is tight and $\sigma_2\gamma \rightarrow +\infty$.

Now if we start with a sequence approximating $d(u_2, v_2)$ and not $d(u_1, v_1)$: To obtain a lower bound, we perform the same kind of work except that we start from a sequence approximating $d_{\mathcal{Y}}(u_2, v_2)$ and not $d_{\mathcal{X}}(u_1, v_1)$: for any $\eta > 0$, there exists $k \in \mathbb{N}$ and two sequences $(a_i)_{1 \leq i \leq k}$ and $(b_i)_{1 \leq i \leq k}$ of X such that:

$$\begin{cases} p^{\mathcal{X}}(a_1) = u_1, b_k = p^{\mathcal{Y}}(v_k), \\ \forall i \in \{1, \dots, k-1\}, (b_i, a_{i+1}) \in \mathcal{SP}_q \\ \frac{\gamma\sigma_2^2}{\gamma\rho_1 + \sigma_2} \sum_{1 \leq i \leq k} d_{\mathcal{X}}(a_i, b_i) + \frac{\sigma_2^2}{\gamma\rho_1 + \sigma_2} k - d_{\mathcal{Y}}(u_2, v_2) \leq \eta. \end{cases}$$

Again, because of the coupling between \mathcal{X} and \mathcal{Y} ,

$$d_{\mathcal{X}}(u_1, v_1) \leq \sum_{1 \leq i \leq k} d_{\mathcal{X}}(a_i, b_i).$$

As a consequence,

$$\begin{aligned} d_{\mathcal{X}}(u_1, v_1) - d_{\mathcal{Y}}(u_2, v_2) &= d_{\mathcal{X}}(u_1, v_1) - \frac{\gamma\rho_1 + \sigma_2}{\gamma\sigma_2^2} d_{\mathcal{Y}}(u_2, v_2) + (1 - \frac{\gamma\rho_1 + \sigma_2}{\gamma\sigma_2^2}) d_{\mathcal{Y}}(u_2, v_2) \\ &\leq \frac{\gamma\rho_1 + \sigma_2}{\gamma\sigma_2^2} \eta + \text{Diam}(\mathcal{Y}_j) \left(1 - \frac{\gamma\rho_1 + \sigma_2}{\gamma\sigma_2^2}\right)_+ \end{aligned}$$

which goes to 0 in probability too with the same arguments. \square

5.2 Reformulation of universality theorems in terms of complete coalescence

Theorem 5.2.1. *Let $X = (X^n)_{n \in \mathbb{N}}$ be a sequence of $\mathbb{M}_2^{\mathbb{N}}$ such that its weight sequence $x \in l_{\downarrow}^2$ satisfies the Aldous-Limic conditions 1 with parameters $(1, \lambda, 0)$. Recall from Definition 2.5.6 the definition of ρ_1 . Assume that there exists $\eta > 0$ such that X satisfies*

$$\begin{cases} \frac{x_0}{\sigma_2} = o(\sigma_2^{1/2+\eta}) \\ \sigma_{2-\eta} \rightarrow 0, \end{cases}$$

and if $\mathcal{X}^n = \text{Coal}(X^n, q_n)$,

$$\begin{cases} \frac{\rho_1}{\sigma_2^2} \rightarrow 1 \\ \forall j \in \mathbb{N}, \sup_{i: X_i \subset \mathcal{X}_j^n} D_i = o_{\mathbb{P}}(\sigma_2^{1/2+\eta}), \\ \sum_i \frac{x_i^2 u_i^2}{\sigma_2^2} \rightarrow 0. \end{cases}$$

Then we have

$$\mathcal{X}^n \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_0(1, \lambda, 0, 1)$$

for the Gromov-weak product topology. If moreover

$$\forall \delta > 0, \forall k \in \mathbb{N}, \liminf_{n \rightarrow +\infty} \mathbb{P}(m_{\delta}^k(\text{Coal}(X^n, q_n)) \geq \epsilon) \xrightarrow{\epsilon \rightarrow 0} 1$$

then it converges with respect to the Gromov-Hausdorff-Prokhorov product topology.

Proof. Of course the extension to the convergence for the Gromov-Hausdorff-Prokhorov product topology comes from Lemma 2.3.7. We only show the convergence for the Gromov-Weak product topology. Our goal is to apply Theorem 3.2.3 to a truncation of $\sigma_2^{-1-\eta} X$, where $\eta \in (0, 2)$ is chosen such that $\sigma_{2-\eta} \rightarrow 0$. First we show that there is a truncation of $\sigma_2^{-1-\eta} X$ that satisfies the hypothesis of Theorem 3.2.3. From Lemma 4.4.1, we have that X always possesses a suitable truncation \bar{X} and there exists $r > 0$ such that

$$\frac{\bar{\sigma}_2^r}{\bar{x}_{\min}} \rightarrow 0.$$

Moreover, from Lemma 4.4.2 and the fact that x satisfies the Aldous-Limic conditions 1, we get that this truncation \bar{x} satisfies the Aldous-Limic conditions 1. Finally, by applying Lemma 4.3.1 to X , we get that the limit of $\text{Coal}(X, q)$ and $\text{Coal}(\bar{X}, q)$ are the same. By applying Lemma 5.1.1 to \bar{X} , we get that the limit of $\text{Coal}(\bar{X}, q)$ and $\text{SCoal}(\bar{X}, q)$ are the same. Let $\mathcal{Y} = \frac{\sigma_2^2}{\sigma_2^{-1-\eta}\rho_1 + \sigma_2} \text{SCoal}_1(\sigma_2^{-1-\eta}\bar{X}, q)$. We want to apply Theorem 3.2.3 to $\sigma_2^{-1-\eta}\bar{X}$. Of course the condition

$$\frac{x_0}{\sigma_2} = o(\sigma_2^{1/2+\eta})$$

translates directly to the same condition on the truncation since for a suitable truncation, $\bar{x}_0 \sim x$ and $\bar{\sigma}_2 \sim \sigma_2$. Let us now turn our attention on the metric aspects of the process. Let us show that, under the conditions of the theorem,

$$\rho_1 \sim \bar{\rho}_1,$$

and

$$\frac{\sigma_2 \sum_i \bar{x}_i^3 u_i^2}{\bar{\rho}_1^2} = o(1).$$

For all $\delta > 0$,

$$\sum_i \hat{x}_i^2 u_i 1_{u_i \geq \delta} \leq \sum_i x_i^2 u_i^2 / \delta$$

so

$$\begin{aligned} \rho_1 - \bar{\rho}_1 &\leq \sum_i x_i^2 u_i^2 / \delta + \sum_i \hat{x}_i^2 u_i 1_{u_i \leq \delta} \\ &\leq \sum_i x_i^2 u_i^2 / \delta + \delta \hat{\sigma}_2. \end{aligned}$$

As a consequence,

$$\begin{aligned} 1 - \frac{\bar{\rho}_1}{\rho_1} &= \frac{\rho_1 - \bar{\rho}_1}{\rho_1} \frac{\sigma_2^2}{\sigma_2^2} \\ &\leq \left(\frac{\sum_i x_i^2 u_i^2}{\delta \sigma_2^2} + \delta \hat{\sigma}_2 \right) \frac{\sigma_2^2}{\rho_1}. \end{aligned}$$

Moreover, $\frac{\sigma_2^2}{\rho_1} \rightarrow 1$ by hypothesis and $\frac{\hat{\sigma}_2}{\sigma_2^2} \sim q^2 \hat{\sigma}_2 = o(1)$ so $\rho_1 \sim \bar{\rho}_1$. Now,

$$\begin{aligned} \frac{\sigma_2 \sum_i \bar{x}_i^3 u_i^2}{\bar{\rho}_1^2} &\sim \sum_i \frac{\bar{x}_i^3 u_i^2}{\sigma_2^3} \\ &\leq \frac{x_1}{\sigma_2} \sum_i \frac{x_i^2 u_i^2}{\sigma_2^2} \\ &\rightarrow 0. \end{aligned}$$

Finally, if $X' = \sigma_2^{-1-\eta}\bar{X}$, $u'_i = \frac{u_i}{\sigma_2^{1+\eta}}$, $\rho'_1 = \rho_1 \sigma_2^{-1-\eta}$, we have

$$\frac{\sigma_2^2}{\rho'_1 + \sigma_2} = \frac{\frac{\sigma_2^2}{\rho_1} \sigma_2^{1+\eta}}{1 + \frac{\sigma_2^2}{\rho_1} \sigma_2^\xi} \sim \sigma_2^{1+\eta}$$

so for $j \in \mathbb{N}$,

$$\sup_{i: X'_i \subset \mathcal{Y}_j} D'_i \frac{\sigma_2^2}{\rho'_1 + \sigma_2} \sim \sup_{i: X_i \subset \mathcal{X}_j} D_i = o_{\mathbb{P}}(\sigma_2^{1/2+\eta}).$$

We get that

$$\mathcal{Y} \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{E}_0(1, \lambda, 0, 1).$$

We can apply Lemma 5.1.2 to \bar{X} with $\gamma = \sigma_2^{-1-\eta}$. Indeed,

$$\gamma \sigma_2 = \sigma_2^{-\eta} \xrightarrow[n \rightarrow \infty]{} +\infty.$$

Moreover, the diameters $\text{Diam}(\mathcal{Y}_j)$ are tight for $j \in \mathbb{N}$ because \mathcal{Y} converges. Last, $\sigma_2 L(G_j^n)$ is tight thanks to Theorem 2.4.9.

Thus we get that the limit of $\text{SCoal}(\bar{X}, q)$ and \mathcal{Y} are the same. \square

Theorem 5.2.2. Let $X = (X^n)_{n \in \mathbb{N}}$ be a sequence of $\mathbb{M}_2^{\mathbb{N}}$ such that its weight sequence $x \in l_{\searrow}^2$ satisfy the Aldous-Limic conditions 1 with parameters $(0, \lambda, c)$. Assume moreover that if $\mathcal{X}^n = \text{Coal}(X^n, q_n)$,

$$\begin{cases} \frac{\rho_1}{\sigma_2^2} & \rightarrow 1 \\ \forall j \in \mathbb{N}, \sup_{i: X_i \subset \mathcal{X}_j^n} D_i & \xrightarrow{\mathbb{P}} 0, \\ \sum_i \frac{x_i^2 u_i^2}{\sigma_2^2} & \rightarrow 0. \end{cases}$$

Then we have that

$$\mathcal{X}^n \xrightarrow[n \rightarrow +\infty]{} \mathcal{E}_0(1, \lambda, 0, 1)$$

for the Gromov-weak product topology. If moreover

$$\forall k \in \mathbb{N}, \liminf_{n \rightarrow +\infty} \mathbb{P}(m_{\delta}^k(\text{Coal}(X^n, q_n)) \geq \epsilon) \xrightarrow[\epsilon \rightarrow 0]{} 1$$

then it converges with respect to the Gromov-Hausdorff-Prokhorov product topology.

Proof. It is basically the same proof that of theorem 5.2.1. □

Chapter 6

Eternal multiplicative coalescent, mixing and entrance boundary

6.1 The lower mass bound is free for an eternal metric multiplicative coalescent

The purpose of this section is to show why using the Gromov-weak topology is enough when dealing with eternal multiplicative coalescent. Basically, the fact that the law of $\text{Coal}(X(t-s), s)$ is stationary makes the tightness of the lower-mass bound close to trivial.

Lemma 6.1.1. *Let $(X(t))_{t \in \mathbb{R}}$ be an eternal multiplicative coalescent. Let $\epsilon > 0$, $s = (s_n)_{n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$ be such that $s_n \rightarrow \infty$ and*

$$L_\epsilon^\infty = \limsup_{n \rightarrow \infty} \mathbb{P}(m_\delta^k(\text{Coal}(X(t-s_n), s_n)) \leq \epsilon | X(t-s_n)).$$

Then

$$L_\epsilon^\infty \xrightarrow{\epsilon \rightarrow 0} 0,$$

in other words, the lower mass bound property for $\text{Coal}(X(t-s), s)$ is satisfied in probability with respect to $X(t-s)$.

Proof. By Levy's downwards theorem, $L_\epsilon^n = \mathbb{P}(m_\delta^k(\text{Coal}(X(t-s_n), s_n)) \leq \epsilon | X(t-s_n)) \mathbb{P}(m_\delta^k(\text{Coal}(X(t-s_n), s_n)) \leq \epsilon | \mathcal{F}_{t-s_n})$ converges almost surely as n goes to infinity. Since this random variable is bounded by one, we also get that

$$\lim_{n \rightarrow +\infty} \mathbb{E}(L_\epsilon^n) = \mathbb{E}(\limsup_{n \rightarrow +\infty} L_\epsilon^n) = \mathbb{E}(L_\epsilon^\infty).$$

However, the left hand side of this equation simplifies to:

$$\begin{aligned} \mathbb{E}(L_\epsilon^n) &= \mathbb{E}(\mathbb{P}(m_\delta^k(\text{Coal}(X(t-s_n), s_n)) \leq \epsilon | X(t-s_n))) \\ &= \mathbb{P}(m_\delta^k(\text{Coal}(X(t-s_n), s_n)) \leq \epsilon) \\ &= \mathbb{P}(m_\delta^k(X(t)) \leq \epsilon). \end{aligned}$$

As a consequence, by Markov's inequality,

$$\mathbb{P}(\limsup_{n \rightarrow +\infty} \mathbb{P}(m_\delta^k(\text{Coal}(X(t-s_n), s_n)) \leq \epsilon | X(t-s_n)) \geq \eta) \leq \frac{\mathbb{P}(m_\delta^k(X(t)) \leq \epsilon)}{\eta}$$

which converges to 0 as ϵ goes to 0 by hypothesis, since $X(t)$ belongs to $\mathbb{M}_2^{\mathbb{N}}$ almost surely. \square

Whereas the precedent lemma gives us a way to extend Gromov-weak results into Gromov-Hausdorff-Prokhorov results, the following one will even allow us to improve on this and add the surplus part of the convergence.

Lemma 6.1.2. *Let $(X^n, Y^n)_{n \in \mathbb{N}}$ be a sequence of $\mathbb{M}_2^{\mathbb{N}} \times \mathbb{M}_2^{\mathbb{N}}$ such that*

$$(X^n, Y^n) \xrightarrow[n \rightarrow +\infty]{d} (\mathcal{X}, \mathcal{Y})$$

for the Gromov-Hausdorff-Prokhorov product topology. Suppose that

$$X^n \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{X},$$

$$Y^n \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{Y}$$

for the Gromov-Hausdorff-Prokhorov plus surplus product topology. Then

$$(X^n, Y^n) \xrightarrow[n \rightarrow +\infty]{(d)} (\mathcal{X}, \mathcal{Y})$$

for the Gromov-Hausdorff-Prokhorov plus surplus product topology.

Proof. Suppose that (X^n, Y^n) doesn't converge in distribution toward $(\mathcal{X}, \mathcal{Y})$ with respect to T_{GHPS} , in particular there exists f and g continuous and bounded for T_{GHPS} such that

$$\exists \epsilon > 0, \forall N \in \mathbb{N}, \exists n \geq N \text{ such that } |\mathbb{E}(f(X^n)g(Y^n)) - \mathbb{E}(f(\mathcal{X})g(\mathcal{Y}))| \geq \epsilon.$$

This allows one to construct a subsequence $(X^{\phi(n)}, Y^{\phi(n)})$ such that

$$\exists \epsilon > 0, \forall n \in \mathbb{N}, |\mathbb{E}(f(G_{\phi(n)}^\epsilon)g(G_{\phi(n)})) - \mathbb{E}(f(\text{Perc}(\mathcal{X}, q))g(\mathcal{X}))| \geq \epsilon. \quad (6.1)$$

On the other hand, we have that $(X^n)_{n \in \mathbb{N}}$ and $(Y^n)_{n \in \mathbb{N}}$ both individually converge in distribution toward \mathcal{X} and \mathcal{Y} with respect to T_{GHPS} thus in particular for all $\epsilon > 0$, there exists K compact for T_{GHPS} such that

$$\mathbb{P}(X^n \in K) \geq 1 - \epsilon,$$

$$\mathbb{P}(Y^n \in K) \geq 1 - \epsilon,$$

As a consequence,

$$\mathbb{P}((X^n, Y^n) \in K^2) \geq 1 - 2\epsilon$$

so from Prokhorov's theorem, we get that (X^n, Y^n) is relatively compact for T_{GHPS} and then one can extract from $(X^{\phi(n)}, Y^{\phi(n)})_{n \in \mathbb{N}}$ a converging subsequence. This subsequence must converge toward $(\mathcal{X}, \mathcal{Y})$ because the addition of surplus in the topology makes it strictly stronger but this would contradict (6.1). \square

6.2 The boundary conditions uniquely determine eternal multiplicative coalescents

Theorem 6.2.1. *Let $(X(t))_{t \in \mathbb{R}}$ be an eternal metric multiplicative coalescent. Recall from Definition 2.5.6 the definition of the associated u, x, ρ_1 and D . Let $t \in \mathbb{R}$. Suppose that there exists $(s_n)_{n \in \mathbb{N}}$, $s_n \xrightarrow[n \rightarrow \infty]{} +\infty$ such that either one of those sets of convergences is satisfied in probability as $n \rightarrow \infty$:*

(i) $(X(t - s_n), s_n)$ verifies Hypothesis 1 with parameters $(1, 0, 0, \gamma) \in \mathcal{I}^\circ$ and $\exists \eta > 0$ such that

$$\begin{cases} \frac{\rho_1}{\sigma_2^2} \rightarrow \gamma, \\ \forall j \in \mathbb{N}, \sup_{i: X_i(t-s_n) \subset X_j(t)} D_i = o_{\mathbb{P}}(\sigma_2^{1/2+\eta}), \\ \frac{\sum_i x_i^3 u_i^2}{\rho_1} \rightarrow 0. \end{cases}$$

(ii) $(X(t - s_n), s_n)$ verifies Hypothesis 1 with parameters $(0, 0, c, \gamma) \in \mathcal{I}^*$, and

$$\begin{cases} \frac{\rho_1}{\sigma_2^2} \rightarrow \gamma \\ \forall j \in \mathbb{N}, \sup_{i: X_i(t-s_n) \subset X_j(t)} D_i = o_{\mathbb{P}}(1), \\ \frac{\sum_i x_i^3 u_i^2}{\rho_1} \rightarrow 0. \end{cases}$$

Then for $t \in \mathbb{N}$, $X(t)$ is real graph almost surely and for any f and g measurable for T_{GHPS} ,

$$\mathbb{E}(f(\text{Coal}(X(t-s), s))g(X(t))) \xrightarrow[s \rightarrow +\infty]{} \mathbb{E}(f(\mathcal{X}(t)))\mathbb{E}(g(\mathcal{X}(t)))$$

where $\mathcal{X}(t) \sim \mathcal{E}_t(1, 0, 0, \gamma)$ in the case of (i) and $\mathcal{X}(t) \sim \mathcal{E}_t(0, 0, c, \gamma)$ in the case of (ii). In particular,

$$(\text{Coal}(X(t-s), s), X(t)) \xrightarrow[s \rightarrow +\infty]{(d)} \mathcal{E}_t(1, 0, 0, \gamma) \times \mathcal{E}_t(1, 0, 0, \gamma)$$

in the case of (i) and

$$(\text{Coal}(X(t-s), s), X(t)) \xrightarrow[s \rightarrow +\infty]{(d)} \mathcal{E}_t(1, 0, 0, \gamma) \times \mathcal{E}_t(0, 0, c, \gamma)$$

in the case of (ii).

Proof. We have that for $t \in \mathbb{R}$, $s > 0$,

$$\mathbb{E}(f(\text{Coal}(X(t-s), s))g(X(t))) = \mathbb{E}(\mathbb{E}(f(\text{Coal}(X(t-s), s))|X(t-s))\mathbb{E}(g(X(t))|X(t-s)))$$

because by hypothesis the coalescence performed to obtain $\text{Coal}(X(t-s), s)$ from $X(t-s)$ is independent from the one performed to obtain $X(t)$. For $t \in \mathbb{R}$, let \mathcal{F}_t be the sigma field generated by $(X(t-s))_{s>0}$. Because X is a Markov process,

$$\mathbb{E}(f(\text{Coal}(X(t-s), s))|X(t-s)) = \mathbb{E}(f(\text{Coal}(X(t-s), s))|\mathcal{F}_{t-s})$$

and

$$\mathbb{E}(g(\text{Coal}(X(t-s), s))|X(t-s)) = \mathbb{E}(g(\text{Coal}(X(t-s), s))|\mathcal{F}_{t-s}).$$

Moreover as $(\text{Coal}(X(t-s), s), X(s)) \stackrel{(d)}{=} (X(t), X(s))$, by Lévy's downward theorem, we have

$$\mathbb{E}(f(\text{Coal}(X(t-s), s))|\mathcal{F}_{t-s}) \xrightarrow[s \rightarrow +\infty]{a.s.} \mathbb{E}(f(X(t))|\mathcal{F}_{-\infty})$$

$$\mathbb{E}(g(\text{Coal}(X(t-s), s))|\mathcal{F}_{t-s}) \xrightarrow[s \rightarrow +\infty]{a.s.} \mathbb{E}(g(X(t))|\mathcal{F}_{-\infty})$$

where $\mathcal{F}_{-\infty} = \bigcap_{s < 0} \mathcal{F}_s$. This fact shifts the matter at hand to identifying the limit. For this, we can examine these convergence but by replacing $s \rightarrow \infty$ by a sequence $(s_n)_{n \in \mathbb{N}}$ such that $s_n \xrightarrow[n \rightarrow \infty]{} +\infty$. By hypothesis, the convergences of either Theorem 5.2.1 or Theorem 5.2.2 are satisfied in probability as well as the lower mass bound property by Lemma 6.1.1. All these convergences can be strengthened

into almost sure convergence, at least along a subsequence of $(s_n)_{n \in \mathbb{N}}$. On the event of probability 1 where this occurs, 5.2.1 and theorem 5.2.2 give us that if f and g are bounded continuous for the Gromov-Hausdorff-Prokhorov product topology,

$$\mathbb{E}(f(\text{Coal}(X(t - s_n), s)) | X(t - s_n)) \xrightarrow[n \rightarrow +\infty]{a.s.} \mathbb{E}(f(X(t))),$$

$$\mathbb{E}(g(\text{Coal}(X(t - s_n), s)) | X(t - s_n)) \xrightarrow[n \rightarrow +\infty]{a.s.} \mathbb{E}(g(X(t)))$$

which then gives that

$$\mathbb{E}(f(\text{Coal}(X(t - s_n), s_n))g(X(t))) \rightarrow \mathbb{E}(f(\mathcal{X}(t)))\mathbb{E}(g(\mathcal{X}(t))).$$

In particular, by applying this convergence with $g = 1$ we get

$$\mathbb{E}(f(X(t))) = \mathbb{E}(f(\mathcal{X}(t)))$$

so $X(t)$ is a real graph almost surely. Of course, the marginals of $(\text{Coal}(X(t - s_n), s_n), X(t))$ converge for the Gromov-Hausdorff-Prokhorov plus surplus topology because $\text{Coal}(X(t - s_n), s_n) \stackrel{(d)}{=} X(t)$. As a consequence, thanks to Lemma 6.1.2, the convergence $\mathbb{E}(f(\text{Coal}(X(t - s_n), s_n))g(X(t))) \rightarrow \mathbb{E}(f(\mathcal{X}(t)))\mathbb{E}(g(\mathcal{X}(t)))$ can be extended to f and g bounded continuous for T_{GHPS} . Finally, as for f and g measurable for T_{GHPS} , it is already established that

$$\mathbb{E}(f(\text{Coal}(X(t - s), s))g(X(t))) \xrightarrow[s \rightarrow +\infty]{} \mathbb{E}(\mathbb{E}(f(X(t)|\mathcal{F}_{-\infty}))g(X(t)))$$

approximating f and g by bounded continuous functions allows one to conclude that for any f and g measurable for T_{GHPS} ,

$$\mathbb{E}(f(\text{Coal}(X(t - s), s))g(X(t))) \rightarrow \mathbb{E}(f(\mathcal{X}(t)))\mathbb{E}(g(\mathcal{X}(t))).$$

□

Part III

Main results

Chapter 7

Palm measure of the point process associated with $\mathcal{E}(\alpha, \lambda, c, \gamma)$

7.1 Result

For the discrete multiplicative coalescent, due to independence properties of the edges distribution, it is immediate to see that if one removes a connected component C from a realization of $\text{gCoal}(x, q)$, one gets a realization of a multiplicative coalescent on the truncated weight sequence $(x_i)_{i \notin C}$. It does not directly translate to an intuition on the law of the continuous object $\mathcal{X} \setminus \mathcal{C}$ where $\mathcal{X} \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$ and $\mathcal{C} \in \text{comp}(\mathcal{X})$, notably because only the largest components of $\text{gCoal}(x, q)$ are captured in the limit object \mathcal{X} . However it is still true, as stated in our main result:

Theorem 7.1.1. *Let $(\alpha, \lambda, c, \gamma) \in \mathcal{I} \times \mathbb{R}_+$ and recall from Definition 2.4.8 the definition of $(\Delta_j)_{j \in \mathbb{N}}$, $(c^{(j)})_{j \in \mathbb{N}}$ and $(\delta_2^{(j)})_{j \in \mathbb{N}}$ associated with $\mathcal{X} \sim \mathcal{E}_0(\alpha, \lambda, c, \gamma)$. Let $f : \mathbb{M} \times \mathbb{M}_2^{\mathbb{N}} \rightarrow \mathbb{R}^+$ measurable for T_{GHP} and such that*

$$\mathbb{E}\left(\sum_j f(\mathcal{X}_j, \mathcal{X})\right) < +\infty.$$

For $j \in \mathbb{N}$, conditionally on $(\mathcal{X}_j, \Delta_j)$, let $\mathcal{Y}^{(j)}$ be distributed as $\mathcal{E}_0(\alpha, \lambda - |\mathcal{X}_j| - \delta_2^{(j)}, c - c^{(j)}, \gamma)$. Assume that

$$\text{either } c = 0, \text{ either } \lambda < 0. \tag{7.1}$$

We have:

$$\mathbb{E}\left(\sum_j f(\mathcal{X}_j, \mathcal{X})\right) = \mathbb{E}\left(\sum_j \mathbb{E}(f(\mathcal{X}_j, \mathcal{Y}^{(j)} \sqcup \mathcal{X}_j | (\mathcal{X}_j, \Delta_j)))\right). \tag{7.2}$$

Note that Theorem 7.1.1 present two pretty artificial limitations: First we are confident that the result is true without the technical hypothesis (7.1). In fact, Theorem 7.1.1 will be true as soon as the conclusion of Lemma 7.2.1 below holds. However as in this work Theorem 7.1.1 will always be used for convergence when $\lambda \rightarrow -\infty$, it is not a problem here. Second, the conditioning $(\mathcal{X}_j, \Delta_j)$ in the left hand side of (7.2) does not fit well with the fact that we announced that it is a result on the Palm measure of $\mathcal{E}(\alpha, \lambda, c, \gamma)$, view as a point process on \mathbb{M} : there should be only \mathcal{X}_j . However we also believe that Δ_j is a measurable function of \mathcal{X}_j , which could possibly be proven via a result analogous to Proposition 2 in [24].

7.2 Proof

7.2.1 Discrete setting

The idea of the proof is straightforward in the sense that it essentially consists in implementing the heuristic on the discrete multiplicative coalescent. Let $X^n = \text{SCoal}_{\sigma_2}(x, q)$, let \mathbb{G}_n be the set of all

connected subgraphs of the complete graph on $[n]$, we have:

$$\begin{aligned}\mathbb{E}\left(\sum_i f(X_i^n, X^n)\right) &= \mathbb{E}\left(\sum_{G \in \mathbb{G}_n} f(G, X^n) 1_{G \in \text{comp}(X^n)}\right) \\ &= \sum_{G \in \mathbb{G}_n} \mathbb{E}(f(G, X^n) | G \in \text{comp}(X^n)) \mathbb{P}(G \in \text{comp}(X^n)).\end{aligned}$$

Moreover, if G is a deterministic connected graph on a subset $V_G \subset [n]$, we have

$$\mathbb{E}(f(G, X^n) | G \in \text{comp}(X^n)) = \mathbb{E}(f(G, Y^{n,(G)}))$$

where

$$Y^{n,(G)} = \text{SCoal}_{\sigma_2}(x^{(G)}, q)$$

where $x^{(G)}$ is the ordering of the sequence of weight of vertices that are not in V_G . Thus

$$\mathbb{E}\left(\sum_j \mathbb{E}(f(X_j^n, Y^{n,(X_j^n)} \sqcup X_j^n) | X_j^n)\right) = \sum_{G \in \mathbb{G}_n} \mathbb{E}(f(G, Y^{n,(G)} \sqcup G)) \mathbb{P}(G \in \text{comp}(X^n)) = \mathbb{E}\left(\sum_j f(X_j^n, X^n)\right). \quad (7.3)$$

The goal of the rest of the proof will be to take this identity to the limit. However, we first have to use it as such in the following lemma :

Lemma 7.2.1. *Let $(\alpha, \lambda, c) \in \mathcal{I}$. Suppose that Hypothesis (7.1) is verified. Then there exists $x = (x^n) \in \text{sl}_2$, $q = (q_n)_{n \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$ such that: (x, q) satisfies the Aldous-Limic conditions Hypothesis 1 for (α, λ, c) and let $\mathcal{X} = (\mathcal{X}^n)_{n \in \mathbb{N}} = (\text{SCoal}(x^n, q_n))_{n \in \mathbb{N}}$, then, for any $\epsilon > 0$,*

$$(\#\{j : |\mathcal{X}_j^n| \geq \epsilon\})_{n \in \mathbb{N}}$$

is uniformly integrable.

Proof. The case where $c = 0$ is a consequence of Proposition 1.4 in [22]. We prove the lemma for $\lambda < 0$. Let $W_\epsilon^n = \#\{j : |\mathcal{X}_j^n| \geq \epsilon\}$. As

$$\mathbb{E}(W_\epsilon^n 1_{W_\epsilon^n \geq a}) \leq \frac{\mathbb{E}((W_\epsilon^n)^2)}{a}$$

it suffices to prove that $\mathbb{E}((W_\epsilon^n)^2)$ is bounded, uniformly in n . Thanks to (7.3), applied with $f(M, N) = 1_{|M| \geq \epsilon} \sum_j 1_{|N_j| \geq \epsilon}$,

$$\mathbb{E}((W_\epsilon^n)^2) \leq \mathbb{E}(W_\epsilon^n) + \mathbb{E}(W_\epsilon^n)^2.$$

so it suffices to show that $\mathbb{E}(W_\epsilon^n)$ is bounded uniformly in n . Finally,

$$W_\epsilon^n \leq \frac{S(x, q)}{\epsilon^2}$$

so we will focus our efforts on $S(x, q)$ for the rest of the proof. Thanks to Lemma 2.5 of [26], for i and $j \in \mathbb{N}$, $i \neq j$, $q < \sigma_2^{-1}$

$$\mathbb{P}(i \leftrightarrow j) \leq \frac{x_i x_j q}{1 - q \sigma_2}$$

where $\mathbb{P}(i \leftrightarrow j)$ is the probability that i and j are in the same component in $\text{gCoal}(x, q)$. As a consequence,

$$\begin{aligned}\mathbb{E}(S(x, q)) &= \sigma_2 + \sum_{i \neq j} x_i x_j \mathbb{P}(i \leftrightarrow j) \\ &\leq \sigma_2 + \frac{\sigma_2^2 q}{1 - q \sigma_2}.\end{aligned}$$

We take $q = \frac{1}{\sigma_2} + \lambda$, this gives that

$$\mathbb{E}(S(x, q)) \leq \frac{1}{-\lambda}$$

which proves that $\mathbb{E}(S(x, q))$ is uniformly bounded. \square

7.2.2 Taking the discrete identity to the limit

Proof of Theorem 7.1.1. Let $x = (x^n)_{n \in \mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}}$ be as in Lemma 7.2.1. Let $X^n = \text{SCoal}_{\sigma_2}(x, q)$. For $j \in \mathbb{N}$, let $x^{(j)}$ be the ordering of the sequence of weight of vertices not in X_j^n . The crucial part of the proof is to understand, by using Theorem 2.4.9, the limit of $\text{SCoal}_{\sigma_2}(x^{(j)}, q)$, where the simple coalescence is independent from X_j^n . For this we need to find for which parameters $x^{(j)}$ would satisfy the Aldous-Limic conditions 1.

Let us suppose that, thanks to Skorokhod's representation theorem, the convergence of Theorem 2.4.9 occurs almost surely. Let us examine the Aldous-Limic conditions for $x^{(j)}$, in the light of Theorem 3.1.3:

$$\begin{cases} \sigma_2(x^{(j)}) = \sigma_2 - \sigma_2(X_j^n) = \sigma_2(1 - \sigma_2\theta_2(X_j^n)) \xrightarrow[n \rightarrow +\infty]{a.s.} 0 \\ \frac{\sigma_3(x^{(j)})}{\sigma_2(x^{(j)})^3} = \frac{1}{(1 - \sigma_2\theta_2(X_j^n))^3} (\frac{\sigma_3}{\sigma_2^3} - \theta_3(X_j^n)) \xrightarrow[n \rightarrow +\infty]{a.s.} \alpha + \sum_{j \geq 1} c_j^3 - \delta_3(\mathcal{X}_j) \\ \forall i \in [n], \frac{x_i}{\sigma_2(x^{(j)})} \mathbf{1}_{\xi_i \notin (l_j^n, r_j^n]} = \frac{x_i}{\sigma_2(1 - \sigma_2\theta_2(X_j^n))} \mathbf{1}_{\xi_i \notin (l_j^n, r_j^n]} \xrightarrow[n \rightarrow +\infty]{a.s.} c_i \mathbf{1}_{\xi_i \notin (l_j, r_j]} \\ q - \frac{1}{\sigma_2(x^{(j)})} = q - \frac{1}{\sigma_2} - \frac{\theta_2(X_j^n)}{1 - \theta_2(X_j^n)\sigma_2} \xrightarrow[n \rightarrow +\infty]{a.s.} \lambda - \alpha|\mathcal{X}_j| - \delta_2(\mathcal{X}_j) \end{cases} \quad (7.4)$$

The third convergence is due to the fact that, thanks to Lemma 4.4 in [14], \mathcal{Z} is continuous at l_j and r_j . Notice how the modified weight sequence $x^{(j)}$ satisfies the Aldous-Limic conditions precisely with the parameters described in the theorem.

Let

$$E_1 = \left\{ f \in C_b(\mathbb{M} \times \mathbb{M}_2^{\mathbb{N}}, \mathbb{R}^+) : \exists \epsilon > 0 \text{ s.t. } \forall (M, N) \in \mathbb{M} \times \mathbb{M}_2^{\mathbb{N}}, |M| < \epsilon \Rightarrow f(M, N) = 0 \right\}.$$

We show that for every $f \in E_1$ the identity eq. (7.2) is true For $n \in \mathbb{N}$, let $W_\epsilon^n = \#\{j : |X_j^n| \geq \epsilon\}$ with ϵ such that if $|M| < \epsilon$, $f(M, N) = 0$.

First, $\sum_j f(X_j^n, X^n)$ converges toward $\sum_j f(\mathcal{X}_j, \mathcal{X})$ in distribution. In fact:

(i) For all $K \geq 1$, $\sum_{j=0}^K f(X_j^n, X^n) \xrightarrow[n \rightarrow +\infty]{(d)} \sum_{j=0}^K f(\mathcal{X}_j, \mathcal{X})$ because $X^n \xrightarrow[n \rightarrow +\infty]{a.s.} \mathcal{X}$.

(ii) We have that for every $\eta > 0$,

$$\mathbb{P}\left(\sum_{j=K+1}^{+\infty} f(X_j^n, X^n) \geq \eta\right) \leq \mathbb{P}\left(\sum_{j=K+1}^{+\infty} \mathbf{1}_{|X_j^n| > \epsilon} \geq \frac{\eta}{\|f\|_\infty}\right) \leq \mathbb{P}(|X_{K+1}^n| \geq \epsilon)$$

and as a consequence,

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\sum_{j=K+1}^{+\infty} f(X_j^n, X^n) \geq \eta\right) = 0.$$

(iii) Since $\mathbb{E}(\sum_i f(\mathcal{X}_i, \mathcal{X})) < \infty$, $\lim_{K \rightarrow \infty} \mathbb{P}(\sum_{i > K} f(\mathcal{X}_i, \mathcal{X})) = 0$.

This proves that $\sum_j f(X_j^n, X^n)$ converges toward $\sum_j f(\mathcal{X}_j, \mathcal{X})$ in distribution. On the other hand, we have that

$$\sum_j f(X_j^n, X^n) \leq \|f\|_\infty W_\epsilon^n$$

and W_ϵ^n is uniformly integrable thanks to Lemma 7.2.1 so by dominated convergence, $\mathbb{E}(\sum_j f(X_j^n, X^n)) \xrightarrow[n \rightarrow \infty]{} \mathbb{E}(\sum_j f(\mathcal{X}_j, \mathcal{X}))$.

Similarly, we show that $\sum_j \mathbb{E}(f(X_j^n, Y^{n, (X_j^n)} \sqcup X_j^n) | X_j^n)$ converges toward $\mathbb{E}(f(\mathcal{X}_j, \mathcal{Y}^{(j)} \sqcup \mathcal{X}_j) | (\mathcal{X}_j, \Delta_j))$ in distribution.

(i) Recall $X^n \xrightarrow[n \rightarrow \infty]{a.s.} \mathcal{X}$. Thanks to Equation (7.4), for $j \in \mathbb{N}$, we can apply Theorem 2.4.9 to $Y^{n, (j)}$ and

we get that $Y^{n, (j)} \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{Y}^{(j)}$. As a consequence, for all $K \geq 1$, conditionally on $((X_j^n)_{n \in \mathbb{N}}, \mathcal{X}_j)$,

$$\sum_{j=0}^K \mathbb{E}(f(X_j^n, Y^{n, (X_j^n)} \sqcup X_j^n) | X_j^n) \xrightarrow[n \rightarrow +\infty]{(d)} \sum_{j=0}^K \mathbb{E}(f(\mathcal{X}_j, \mathcal{Y}^{(j)} \sqcup \mathcal{X}_j) | (\mathcal{X}_j, \Delta_j))$$

(ii) We have that for every $\eta > 0$,

$$\mathbb{P}\left(\sum_{j=K+1}^{+\infty} \mathbb{E}(f(X_j^n, Y^{n, (X_j^n)} \sqcup X_j^n) | X_j^n) \geq \eta\right) \leq \mathbb{P}\left(\sum_{j=K+1}^{+\infty} 1_{|X_j^n| > \epsilon} \geq \frac{\eta}{\|f\|_\infty}\right) \leq \mathbb{P}(|X_{K+1}^n| \geq \epsilon)$$

and as a consequence,

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\sum_{j=K+1}^{+\infty} \mathbb{E}(f(X_j^n, Y^{n, (X_j^n)} \sqcup X_j^n) | X_j^n) \geq \eta\right) = 0.$$

(iii) Since $\sum_{j \geq K} \mathbb{E}(f(\mathcal{X}_j, \mathcal{Y}^{(j)} \sqcup \mathcal{X}_j) | (\mathcal{X}_j, \Delta_j)) \leq \|f\|_\infty \mathbb{E}(\sum_{j \geq K} 1_{|\mathcal{X}_j| \geq \epsilon})$, for all $\eta > 0$,

$$\lim_{K \rightarrow \infty} \mathbb{P}\left(\sum_{j \geq K} \mathbb{E}(f(\mathcal{X}_j, \mathcal{Y}^{(j)} \sqcup \mathcal{X}_j) | (\mathcal{X}_j, \Delta_j)) \geq \eta\right) = 0.$$

This proves that $\sum_j \mathbb{E}(f(X_j^n, Y^{n, (X_j^n)} \sqcup X_j^n) | X_j^n)$ converges toward $\mathbb{E}(f(\mathcal{X}_j, \mathcal{Y}^{(j)} \sqcup \mathcal{X}_j) | (\mathcal{X}_j, \Delta_j))$. Again, as

$$\sum_j \mathbb{E}(f(X_j^n, Y^{n, (X_j^n)} \sqcup X_j^n) | X_j^n) \leq W_\epsilon^n$$

and $(W_\epsilon^n)_{n \in \mathbb{N}}$ is uniformly integrable, by dominated convergence,

$$\mathbb{E}\left(\sum_j \mathbb{E}(f(X_j^n, Y^{n, (X_j^n)} \sqcup X_j^n) | X_j^n)\right) \xrightarrow{n \rightarrow \infty} \mathbb{E}(f(\mathcal{X}_j, \mathcal{Y}^{(j)} \sqcup \mathcal{X}_j) | (\mathcal{X}_j, \Delta_j))$$

This gives the identity (7.2), for $f \in E_1$.

We show that for every $f \in C_b(\mathbb{M} \times \mathbb{M}_2^{\mathbb{N}}, \mathbb{R}^+)$ the identity (7.2) is true. Let $f \in C_b(\mathbb{M} \times \mathbb{M}_2^{\mathbb{N}}, \mathbb{R}^+)$ and let h be its corresponding function

$$h : \begin{cases} \mathbb{M} \times \{\text{subsets of } \mathbb{N}\} & \rightarrow \mathbb{R}^+ \\ (M, \Delta) & \mapsto \mathbb{E}(f(M, \mathcal{Y}^{(M, \Delta)})) \end{cases}$$

where, if $\delta_2 = \sum_{i \in \Delta} c_i$, $c' = (c_i 1_{i \in \Delta})_{i \in \mathbb{N}}$, $\mathcal{Y}^{(M, \Delta)} \sim \mathcal{E}(\alpha, \lambda - |M| - \delta_2, c - c', \gamma)$. For any $\epsilon > 0$, let $\psi_\epsilon \in C(\mathbb{R}^+, \mathbb{R}^+)$ be a function such that $\psi_\epsilon(x) \leq 1_{x \geq \epsilon}$ and $\psi_\epsilon \nearrow 1_{x > 0}$. Let

$$f_\epsilon : \begin{cases} \mathbb{M} \times \mathbb{M}_2^{\mathbb{N}} & \rightarrow \mathbb{R}^+ \\ (M, N) & \mapsto f(M, N) \psi_\epsilon(|M|) \end{cases}, h_\epsilon : \begin{cases} \mathbb{M} \times \{\text{subsets of } \mathbb{N}\} & \rightarrow \mathbb{R}^+ \\ (M, \Delta) & \mapsto h(M, \Delta) \psi_\epsilon(|M|). \end{cases}$$

Then as almost surely, for any $i \in \mathbb{N}$, $|X_i| > 0$, $\mathbb{E}(\sum_j f_\epsilon(X_j, X)) \nearrow \mathbb{E}(\sum_j f(X_j, X))$ and $\mathbb{E}(\sum_j h_\epsilon(X_j, \Delta_j)) \nearrow \mathbb{E}(\sum_j h(X_j, \Delta_j))$ as $\epsilon \searrow 0$, by monotone convergence. This gives the identity for $f \in C_b(\mathbb{M} \times \mathbb{M}_2^{\mathbb{N}}, \mathbb{R}^+)$. Finally, we can approximate the indicator function of a ball by bounded continuous functions and then extend the identity to all f measurable using a $\pi - \lambda$ argument. \square

Chapter 8

Mixing property for the coalescence of homogeneous multiplicative spaces

8.1 Results

8.2 Prerequisite

We collect in the following lemma some crucial results from Aldous' stochastic calculus analysis of the multiplicative coalescent:

Lemma 8.2.1 ([5], subsection 5.3 and 5.4). *For $t \in \mathbb{R}$, let ψ^t be defined by: for all $f : \mathbb{R}^+ \mapsto \mathbb{R}^+$ measurable,*

$$\mathbb{E}\left(\sum_i f(x_i)\right) = \int_{\mathbb{R}^+} f(s)\psi^t(ds).$$

Let ρ^t be defined by: for all $s > 0$

$$\frac{d\psi^t}{d\rho^t}(s) = s^{-1}.$$

Let ν^t be defined by:

$$\nu^t(ds) = \frac{e^{-t^2 s/2}}{\sqrt{2\pi s^3}} ds.$$

Then, for $t < 0$,

$$\forall s > 0, \rho^t(s, +\infty) \leq \nu^t(s, +\infty)$$

and

$$\rho^t(ds) \geq e^{s^2 t/2 - s^3/6} \nu^t(ds).$$

In particular, when t goes to $-\infty$,

$$\forall p \geq 2, \mathbb{E}(\sigma_p) \sim |t|^{3-2p} \int_0^\infty \frac{e^{-s^2}}{\sqrt{2\pi s^{5/2-p}}} ds.$$

Thanks to Theorem 7.1.1 and to the precise understanding of the distribution $\mathcal{E}(1, 0, 0, 1)$ developed in [1] and [2], we are able to obtain the following rather precise asymptotic result.

Theorem 8.2.2. *Let $(X(t))_{t \in \mathbb{R}} \sim \mathcal{E}(1, \lambda, 0, 1)$. Recall from Definition 2.5.6 the definition of ρ_1, u, x, D . There exists constants K_r such that, as $t \rightarrow -\infty$,*

$$\forall p \geq 2, r \geq 0, \sum_j x_j^p u_i^r \sim_{\mathbb{P}} K_r \sigma_{p+r/2}, \quad (8.1)$$

$$\forall p \geq 2, \sigma_p \sim \mathbb{E}(\sigma_p), \quad (8.2)$$

$$\forall \eta \in (0, 1), \frac{x_0}{\sigma_2} = o_{\mathbb{P}}(\sigma_2^\eta) \quad (8.3)$$

$$\forall \eta \in (0, 1), \sup_i D_i = o_{\mathbb{P}}(\sigma_2^\eta) \quad (8.4)$$

In particular, this theorem allows us to show:

Corollary 8.2.3. *Let $X \sim \mathcal{E}(1, \lambda, 0, 1)$, then for all $t \in \mathbb{R}$, for all f and g bounded and measurable for T_{GHPS} ,*

$$\mathbb{E}(f(\text{Perc}(X(t), s))g(X(t))) \xrightarrow{s \rightarrow +\infty} \mathbb{E}(f(X(t)))\mathbb{E}(g(X(t))).$$

Proof. Thanks to Theorem 8.2.2,

$$\exists \eta > 0, \frac{x_0}{\sigma_2} = o_{\mathbb{P}}(\sigma_2^\eta).$$

Moreover, there exists K_1 such that

$$\frac{\rho_1}{\sigma_2^2} \sim_{\mathbb{P}} K_1 \frac{\mathbb{E}(\sigma_{5/2})}{\mathbb{E}(\sigma_2^3)}.$$

Thus, thanks to Lemma 8.2.1, there exists a constant C such that

$$\frac{\rho_1}{\sigma_2^2} \xrightarrow{t \rightarrow -\infty} C.$$

With the same reasoning we get

$$\frac{\sum_i x_i^3 u_i^2}{\rho_1} \rightarrow 0.$$

Finally, again thanks to Theorem 8.2.2,

$$\exists \eta > 0, \sup_i D_i = o_{\mathbb{P}}(\sigma_2^\eta),$$

Remark that we could determine C precisely from the proof of Theorem 8.2.2 and Lemma 8.2.1. However, by taking $g = 1$, the convergence of the theorem should, once proved, show that $C = 1$. Indeed, thanks to Theorem 6.2.1,

$$\mathbb{E}(f(\text{Coal}(X(t-s), s))g(X(t))) \rightarrow \mathbb{E}(f(\mathcal{X}(t)))\mathbb{E}(g(\mathcal{X}(t))).$$

where $\mathcal{X}(t) \sim \mathcal{E}_0(A, \lambda, 0, C)$ so $C = 1$. Finally, thanks to the duality property of [28], Proposition 14,

$$\mathbb{E}(f(\text{Perc}(X(t), s))g(X(t))) = \mathbb{E}(f(\text{Coal}(X(t-s), s))g(X(t))).$$

□

8.3 First moment of metric statistics

To make things easier, we are going to focus on $\sigma_{p,r}^A = \sum_i x_i^p u_i^r 1_{x_i \leq A}$, ensuring that the component sizes are not too large, given that they will go to 0 when $t \rightarrow -\infty$ anyway.

Lemma 8.3.1. *Let $A > 0$, $p \geq 2$, $r \geq 0$, e be a standard Brownian excursion, $E = \int_0^1 e(u)du$, $\sigma_{p,r}^A = \sum_i x_i^p u_i^r 1_{x_i \leq A}$. Then exists a continuous function ϕ_r on \mathbb{R}_+ such that $\lim_{A \rightarrow 0} \phi(A) = 0$ and*

$$|\mathbb{E}(\sigma_{p,r}^A) - \mathbb{E}(E^r)\mathbb{E}(\sigma_{p+\frac{r}{2}}^A)| \leq \phi_r(A)\mathbb{E}(\sigma_{p+r/2}^A).$$

In particular, there is $B_r > 0$ such that for $A > 0$,

$$\mathbb{E}(\sigma_{p,r}^A) \leq B_r \mathbb{E}(\sigma_{p+r/2}^A)$$

Proof. In subsection 2.1 of [2] the i -th component of $X \sim \mathcal{E}_t(1, 0, 0, 1)$ conditionally on its size x_i is constructed as follows: sample a random real excursion \tilde{e}_i such that, for all $f : ([0, x_i] \rightarrow \mathbb{R}^+) \rightarrow \mathbb{R}$ bounded continuous,

$$\mathbb{E}(f(\tilde{e}_i)|x_i) = \frac{\mathbb{E}(f(e_i) \exp(\int_0^{x_i} e_i(s)ds)|x_i)}{\mathbb{E}(\exp(\int_0^{x_i} e_i(s)ds)|x_i)}.$$

Then, add n_i shortcuts, where n_i has a Poisson distribution with parameter $\int_0^{x_i} \tilde{e}_i(s)ds$. The main point of the proof is to show the asymptotics at $-\infty$ are well understood by considering that all components have 0 shortcuts. For the sake of clarity, from now on, x_i will be deterministic i.e., we consider the average distance u on a connected component of fixed size x , encoded by \tilde{e} and with n shortcuts.

- On one hand, as the addition of shortcuts only decreases the distances on the component, the average distance on the said component is at most the average distance on the tree encoded by \tilde{e} , which has expression: $x^{-1} \int_0^x \tilde{e}(s) ds$. As a consequence:

$$\begin{aligned} \mathbb{E}(u^r) &\leq \mathbb{E}\left(x^{-r} \left(\int_0^x \tilde{e}(s) ds\right)^r\right) \\ &= \frac{\mathbb{E}\left(x^{-r} \left(\int_0^x e_x(s) ds\right)^r \exp\left(\int_0^x e_x(s) ds\right)\right)}{\mathbb{E}\left(\exp\left(\int_0^x e_x(s) ds\right)\right)} \end{aligned}$$

where e_x is an untilted Brownian excursion on $[0, x]$. By Brownian scaling, we then get that:

$$\mathbb{E}(u^r|x) \leq x^{r/2} \frac{\mathbb{E}(E^r \exp(x^{3/2}E))}{\mathbb{E}(\exp(x^{3/2}E))} \leq x^{r/2} \mathbb{E}(E^r \exp(x^{3/2}E))$$

and then

$$\mathbb{E}(u^r) - \mathbb{E}(E^r)x^{r/2} \leq x^{\frac{r}{2}} \mathbb{E}(E^r(\exp(x^{3/2}E) - 1)).$$

- On the other hand, just from the fact that $1 \geq 1_{n=0}$,

$$\begin{aligned} \mathbb{E}(u^r) &\geq \mathbb{E}(u^r 1_{n=0}) \\ &= \mathbb{E}\left(x^{-r} \left(\int_0^x \tilde{e}(s) ds\right)^r 1_{n_i=0}\right) \end{aligned}$$

Moreover, as n is a Poisson random variable with mean $\int_0^x \tilde{e}(s) ds$ conditionally on \tilde{e} , $\mathbb{P}(n = 0|\tilde{e}) = \exp(-\int_0^x \tilde{e}(u) du)$ so

$$\begin{aligned} \mathbb{E}\left(x^{-r} \left(\int_0^x \tilde{e}(s) ds\right)^r 1_{n=0}\right) &= \mathbb{E}\left(\mathbb{E}\left(x^{-r} \left(\int_0^x \tilde{e}(s) ds\right)^r 1_{n=0} \mid \int_0^x \tilde{e}(s) ds\right)\right) \\ &= \mathbb{E}\left(x^{-r} \left(\int_0^x \tilde{e}(s) ds\right)^r \mathbb{P}\left(n = 0 \mid \int_0^x \tilde{e}(u) du\right)\right) \\ &= \mathbb{E}\left(x^{-r} \left(\int_0^x \tilde{e}(s) ds\right)^r \exp\left(-\int_0^x \tilde{e}(s) ds\right)\right) \\ &= x^{-r} \frac{\mathbb{E}\left(\left(\int_0^x e_x(s) ds\right)^r\right)}{\mathbb{E}\left(\exp\left(\int_0^x e_x(u) du\right)\right)} \\ &= x^{r/2} \frac{\mathbb{E}(E^r)}{\mathbb{E}(\exp(x^{3/2}E))} \end{aligned}$$

Finally, we get

$$|\mathbb{E}(u_i^r) - x^{r/2} \mathbb{E}(E^r)| \leq x^{r/2} (\mathbb{E}(E^r(e^{x^{3/2}E} - 1)) \vee (1 - \mathbb{E}(e^{x^{3/2}E})^{-1})).$$

Let us suppose that $x \leq A$ for $A > 0$. Because the right hand side of this inequality is non-increasing in x , if $\phi_r : x \mapsto \mathbb{E}(E^r(e^{x^{3/2}E} - 1)) \vee (1 - \mathbb{E}(e^{x^{3/2}E})^{-1})$,

$$|\mathbb{E}(u^r) - x^{r/2} \mathbb{E}(E^r)| \leq x^{r/2} \phi_r(A).$$

Now let us focus on $\sigma_{p,r}^A$. From now on x_i is no longer deterministic.

$$\begin{aligned} \mathbb{E}(\sigma_{p,r}^A) &= \mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A}\right) \\ &= \mathbb{E}\left(\sum_i \mathbb{E}(x_i^p u_i^r 1_{x_i \leq A} | x_i)\right) \\ &= \mathbb{E}\left(\sum_i x_i^p 1_{x_i \leq A} \mathbb{E}(u_i^r | x_i)\right) \end{aligned}$$

so

$$\begin{aligned}
|\mathbb{E}(\sigma_{p,r}^A) - \mathbb{E}(E^r)\mathbb{E}(\sigma_{p+r/2}^A)| &= \left| \mathbb{E} \left(\sum_i (x_i^p 1_{x_i \leq A} \mathbb{E}(u_i^r | x_i) - \mathbb{E}(E^r) x_i^{p+r/2} 1_{x_i \leq A}) \right) \right| \\
&\leq \mathbb{E} \left(\sum_i x_i^p 1_{x_i \leq A} |\mathbb{E}(u_i^r | x_i) - \mathbb{E}(E^r) x_i^{r/2}| \right) \\
&\leq \phi_r(A) \mathbb{E} \left(\sum_i x_i^{p+r/2} 1_{x_i \leq A} \right) \\
&= \phi_r(A) \mathbb{E}(\sigma_{p+r/2}^A).
\end{aligned}$$

□

In the same spirit, we are able to control the moments of $\sup_i D_i 1_{x_0 \leq A}$ as follows:

Lemma 8.3.2. *Recall from Definition 2.5.6 the definition of D . Let $D_{\max} = \sup_i D_i$. Let $A > 0$, $p \geq 2$ let $\sigma_p^A = \sum_i x_i^p 1_{x_i \leq A}$. Then as $t \rightarrow -\infty$,*

$$\mathbb{E}(D_{\max}^{2p} 1_{x_0 \leq A}) = O(\mathbb{E}(\sigma_p^A))$$

Proof. First, as $x_i \leq x_0$

$$\mathbb{E}(D_{\max}^{2p} 1_{x_0 \leq A}) \leq \mathbb{E} \left(\sum_i D_i^{2p} 1_{x_i \leq A} \right)$$

As adding shortcuts can only decrease the diameter of a component, D_i is bounded by the diameter of the tree encoded by \tilde{e}_i , which is itself bounded by $2 \sup_{[0, x_i]} \tilde{e}_i$. As a consequence,

$$\begin{aligned}
\mathbb{E}(D_{\max}^{2p} 1_{x_0 \leq A}) &\leq 2^{2p} \mathbb{E} \left(\sum_i \mathbb{E} \left(\sup_{s \in [0, x_i]} \tilde{e}(s)^{2p} | x_i \right) 1_{x_i \leq A} \right) \\
&\leq 2^{2p} \mathbb{E} \left(\sum_i x_i^p \mathbb{E} \left(\sup_{s \in [0, 1]} e(s)^{2p} e^{x_i^{3/2} E} | x_i \right) 1_{x_i \leq A} \right)
\end{aligned}$$

where e is a standard Brownian excursion on $[0, 1]$ independent of x_i and $E = \int_0^1 e(s) ds$. Finally,

$$\mathbb{E}(D_{\max}^{2p} 1_{x_0 \leq A}) \leq \mathbb{E} \left(\sup_{s \in [0, 1]} e(s)^{2p} e^{A^{3/2} E} \right) \mathbb{E}(\sigma_p^A).$$

□

8.4 Second moment of general statistics

Now that we have precise asymptotics on the first moments, we want to know more about the second moments to get concentrations results. Our main tool here is Theorem 7.1.1.

Lemma 8.4.1. *Let $A > 0$, let $\sigma_{p,r}^A = \sum_i x_i^p u_i^r 1_{x_i \leq A}$. Then for any $p \geq 2$, $r \geq 0$*

$$\lim_{A \rightarrow 0} \limsup_{t \rightarrow -\infty} \frac{\text{Var}(\sigma_{p,r}^A)}{\mathbb{E}(\sigma_{p,r}^A)^2} = 0.$$

Proof. We use Janson and Spencer's formula (7.2) with $f(M, N) = |M_i|^2 u(M) 1_{|M| \leq A} \sum_{j \in \mathbb{N}} |N_j|^2 u(N_j) 1_{|N_j| \leq A}$. We have $\sum_{i \in \mathbb{N}} f(X_i, X) = (\sigma_{p,r}^A)^2$ and

$$\mathbb{E}((\sigma_{p,r}^A)^2) = \mathbb{E}(\sigma_{2p,2r}^A) + \mathbb{E} \left(\sum_i x_i^p u_i^r 1_{x_i \leq A} \mathbb{E}(\sigma_{p,r}^A(t - x_i) | x_i, u_i) \right)$$

where $\sigma_{p,r}^A(t - x_i)$ is, conditionally on (x_i, u_i) , independent from (x, u) and is sampled according to the law at time $t - x_i$. As a consequence,

$$\text{Var}(\sigma_{p,r}^A) = \mathbb{E}(\sigma_{2p,2r}^A) + \mathbb{E} \left(\sum_i x_i^p u_i^r 1_{x_i \leq A} [\mathbb{E}(\sigma_{p,r}^A(t - x_i) | x_i, u_i) - \mathbb{E}(\sigma_{p,r}^A)] \right)$$

Now we are going to approximate each term from the right-hand side of this equation using Lemma 8.3.1, starting with $\mathbb{E}(\sigma_{2p,2r}^A)$, $\mathbb{E}(\sigma_{p,r}^A(t-x_i)|x_i, u_i)$ and $\mathbb{E}(\sigma_{p,r}^A)$, for $A > 0$ small enough,

$$\begin{aligned}\mathbb{V}ar(\sigma_{p,r}^A) &\leq B_{2r}\mathbb{E}(\sigma_{2p+r}^A) + \\ &\quad + \mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A} [\mathbb{E}(\sigma_{p+r/2}^A(t-x_i)|x_i) - \mathbb{E}(\sigma_{p+r/2}^A)]\right) \\ &\quad + \phi_r(A)\mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A} [\mathbb{E}(\sigma_{p+r/2}^A(t-x_i)|x_i) + \mathbb{E}(\sigma_{p+r/2}^A)]\right)\end{aligned}$$

Moreover as some of these quantities are non-decreasing in A ,

$$\begin{aligned}\mathbb{V}ar(\sigma_{p,r}^A) &\leq B_{2r}\mathbb{E}(\sigma_{2p+r}^A) \\ &\quad + \mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A} [\mathbb{E}(\sigma_{p+r/2}^A(t-x_i)|x_i) - \mathbb{E}(\sigma_{p+r/2}^A)]\right) \\ &\quad + \phi_r(A)\mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A} [\mathbb{E}(\sigma_{p+r/2}^A(t-x_i)|x_i) + \mathbb{E}(\sigma_{p+r/2}^A)]\right).\end{aligned}$$

Moreover,

$$\begin{aligned}\mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A} [\mathbb{E}(\sigma_{p+r/2}^A(t-x_i)|x_i) - \mathbb{E}(\sigma_{p+r/2}^A)]\right) &= \mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A} \mathbb{E}(\sigma_{p+r/2}^A(t-x_i)|x_i)\right) - \mathbb{E}(\sigma_{p,r}^A)\mathbb{E}(\sigma_{p+r/2}^A) \\ &= \mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A} \mathbb{E}(\sigma_{p+r/2}^A(t-x_i)|x_i)\right) - \mathbb{E}(\sigma_{p,r}^A)\mathbb{E}(\sigma_{p+r/2}^A) + \mathbb{E}(\sigma_{p,r}^A)(\mathbb{E}(\sigma_{p+r/2}^A) - \mathbb{E}(\sigma_{p+r/2}^A)) \\ &= \mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A} [\mathbb{E}(\sigma_{p+r/2}^A(t-x_i)|x_i) - \mathbb{E}(\sigma_{p+r/2}^A)]\right) + \mathbb{E}(\sigma_{p,r}^A)(\mathbb{E}(\sigma_{p+r/2}^A) - \mathbb{E}(\sigma_{p+r/2}^A)) \\ &\leq \mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A} [\mathbb{E}(\sigma_{p+r/2}^A(t-x_i)|x_i) - \mathbb{E}(\sigma_{p+r/2}^A)]\right) + B_r\mathbb{E}(\sigma_{p+r/2}^A)(\mathbb{E}(\sigma_{p+r/2}^A) - \mathbb{E}(\sigma_{p+r/2}^A)).\end{aligned}$$

Finally, notice that when two components merge, their contributions to $\sigma_{p+r/2}$ cannot decrease, hence $\mathbb{E}(\sigma_{p+r/2}^A(t-x_i)|x_i) \leq \mathbb{E}(\sigma_{p+r/2}^A)$. We get:

$$\begin{aligned}\mathbb{V}ar(\sigma_{p,r}^A) &\leq B_{2r}\mathbb{E}(\sigma_{2p+r}^A) + B_r\mathbb{E}(\sigma_{p+r/2}^A)(\mathbb{E}(\sigma_{p+r/2}^A) - \mathbb{E}(\sigma_{p+r/2}^A)) + 2\phi_r(A)\mathbb{E}(\sigma_{p+r/2}^A)\mathbb{E}\left(\sum_i x_i^p u_i^r 1_{x_i \leq A}\right) \\ &\leq B_{2r}\mathbb{E}(\sigma_{2p+r}^A) + B_r\mathbb{E}(\sigma_{p+r/2}^A)(\mathbb{E}(\sigma_{p+r/2}^A) - \mathbb{E}(\sigma_{p+r/2}^A)) + 2\phi_r(A)B_r\mathbb{E}(\sigma_{p+r/2}^A)^2.\end{aligned}$$

Then, again by Lemma 8.3.1, for $A > 0$ small enough, there exists C_r ,

$$\begin{aligned}\frac{\mathbb{V}ar(\sigma_{p,r}^A)}{\mathbb{E}(\sigma_{p,r}^A)^2} &\leq C_r \frac{\mathbb{V}ar(\sigma_{p,r}^A)}{\mathbb{E}(\sigma_{p+r/2}^A)^2} \\ &\leq C_r \left(B_{2r} \frac{\mathbb{E}(\sigma_{2p+r}^A)}{\mathbb{E}(\sigma_{p+r/2}^A)^2} + B_r \frac{\mathbb{E}(\sigma_{p+r/2}^A) - \mathbb{E}(\sigma_{p+r/2}^A)}{\mathbb{E}(\sigma_{p+r/2}^A)} + 2\phi_r(A)B_r \right).\end{aligned}$$

By Lemma 8.5.1 and Lemma 8.2.1, $\frac{\mathbb{E}(\sigma_{p+r/2}^A) - \mathbb{E}(\sigma_{p+r/2}^A)}{\mathbb{E}(\sigma_{p+r/2}^A)} \rightarrow 0$ as t goes to $-\infty$. By Lemma 8.2.1, $\frac{\mathbb{E}(\sigma_{2p+r}^A)}{\mathbb{E}(\sigma_{p+r/2}^A)^2} = O(|t|^{-3})$ so all in all,

$$\limsup_{t \rightarrow -\infty} \frac{\mathbb{V}ar(\sigma_{p,r}^A)}{\mathbb{E}(\sigma_{p,r}^A)^2} \leq 2\phi_r(A)B_r.$$

As, from Lemma 8.3.1, $\phi_r(A) \xrightarrow{A \rightarrow 0} 0$, this concludes the proof. \square

8.5 Proof of Theorem 8.2.2

8.5.1 Average distances

Proof of (8.1) in Theorem 8.2.2. Notice that for u, v and $w \in \mathbb{R}$, $\epsilon \in (0, 1]$,

$$|u-1| < \epsilon, |v-1| < \epsilon, |w-1| < \epsilon,$$

then

$$|uvw - 1| < \epsilon(2^3 - 1).$$

As a consequence, for $A > 0$,

$$\begin{aligned} \mathbb{P}\left(\left|\frac{\sigma_{p,r}}{\mathbb{E}(E^r)\sigma_{p+r/2}} - 1\right| \geq \epsilon(2^3 - 1)\right) &= \mathbb{P}\left(\left|\frac{\sigma_{p,r}}{\mathbb{E}(\sigma_{p,r}^A)} \frac{\mathbb{E}(\sigma_{p,r}^A)}{\mathbb{E}(E^r)\mathbb{E}(\sigma_{p+r/2}^A)} \frac{\mathbb{E}(\sigma_{p+r/2}^A)}{\sigma_{p+r/2}} - 1\right| \geq \epsilon(2^3 - 1)\right) \\ &\leq \mathbb{P}\left(\left|\frac{\sigma_{p,r}}{\mathbb{E}(\sigma_{p,r}^A)} - 1\right| \geq \epsilon\right) + \mathbb{P}\left(\left|\frac{\mathbb{E}(\sigma_{p,r}^A)}{\mathbb{E}(E^r)\mathbb{E}(\sigma_{p+r/2}^A)} - 1\right| \geq \epsilon\right) \\ &\quad + \mathbb{P}\left(\left|\frac{\mathbb{E}(\sigma_{p+r/2}^A)}{\sigma_{p+r/2}} - 1\right| \geq \epsilon\right) \\ &= P_1 + P_2 + P_3 \end{aligned}$$

Notice that on the event $\{x_0 \leq A\}$, $\sigma_{p,r}^A = \sigma_{p,r}$. As a consequence,

$$\begin{aligned} P_1 &= \mathbb{P}\left(\left|\frac{\sigma_{p,r}}{\mathbb{E}(\sigma_{p,r}^A)} - 1\right| \geq \epsilon\right) \\ &\leq \mathbb{P}\left(\left|\frac{\sigma_{p,r}}{\mathbb{E}(\sigma_{p,r}^A)} - 1\right| \geq \epsilon \text{ and } x_0 \leq A\right) + \mathbb{P}(x_0 > A) \\ &\leq \mathbb{P}\left(\left|\frac{\sigma_{p,r}^A}{\mathbb{E}(\sigma_{p,r}^A)} - 1\right| \geq \epsilon\right) + \mathbb{P}(x_0 > A) \\ &\leq \frac{\text{Var}(\sigma_{p,r}^A)}{\epsilon^2 \mathbb{E}(\sigma_{p,r}^A)^2} + \mathbb{P}(x_0 > A). \end{aligned}$$

However, from [5], equation (78), $\mathbb{P}(x_0 > A) \xrightarrow[t \rightarrow -\infty]{} 0$, so

$$\limsup_{t \rightarrow -\infty} \mathbb{P}\left(\left|\frac{\sigma_{p,r}}{\mathbb{E}(\sigma_{p,r}^A)} - 1\right| \geq \epsilon\right) \leq \limsup_{t \rightarrow -\infty} \frac{\text{Var}(\sigma_{p,r}^A)}{\epsilon^2 \mathbb{E}(\sigma_{p,r}^A)^2} \quad (8.5)$$

for all $A > 0$, which gives that $\lim_{A \rightarrow 0} \limsup_{t \rightarrow -\infty} \mathbb{P}\left(\left|\frac{\sigma_{p,r}}{\mathbb{E}(\sigma_{p,r}^A)} - 1\right| > \epsilon\right) = 0$. On the other hand,

$$\mathbb{E}(\sigma_{p,r}^A) \sim \mathbb{E}(E^r)\mathbb{E}(\sigma_{p+r/2}^A)$$

by Lemma 8.3.1 so $\lim_{n \rightarrow \infty} P_2 = 0$ Finally,

$$\begin{aligned} P_3 &= \mathbb{P}\left(\left|\frac{\mathbb{E}(\sigma_{p+r/2}^A)}{\sigma_{p+r/2}} - 1\right| \geq \epsilon\right) \\ &\leq \mathbb{P}\left(\left|\frac{\mathbb{E}(\sigma_{p+r/2}^A)}{\sigma_{p+r/2}} - 1\right| \geq \epsilon \text{ and } x_0 \leq A\right) + \mathbb{P}(x_0 > A) \\ &\leq \mathbb{P}\left(\left|\frac{\mathbb{E}(\sigma_{p+r/2}^A)}{\sigma_{p+r/2}^A} - 1\right| \geq \epsilon\right) + \mathbb{P}(x_0 > A) \end{aligned}$$

As

$$\begin{aligned} \mathbb{P}\left(\left|\frac{\mathbb{E}(\sigma_{p+r/2}^A)}{\sigma_{p+r/2}^A} - 1\right| \geq \epsilon\right) &\leq \mathbb{P}\left(\left\{\frac{\sigma_{p+r/2}^A}{\mathbb{E}(\sigma_{p+r/2}^A)} - 1 \leq \frac{-\epsilon}{1+\epsilon}\right\} \cup \left\{\frac{\sigma_{p+r/2}^A}{\mathbb{E}(\sigma_{p+r/2}^A)} - 1 \geq \frac{\epsilon}{1-\epsilon}\right\}\right) \\ &\leq \mathbb{P}\left(\left|\frac{\sigma_{p+r/2}^A}{\mathbb{E}(\sigma_{p+r/2}^A)} - 1\right| \geq \epsilon/2\right) \\ &\leq \frac{4\text{Var}(\sigma_{p,r}^A)}{\epsilon^2 \mathbb{E}(\sigma_{p,r}^A)^2}. \end{aligned}$$

Again, we get

$$\lim_{A \rightarrow 0} \lim_{n \rightarrow \infty} \mathbb{P}\left(\left|\frac{\mathbb{E}(\sigma_{p+r/2}^A)}{\sigma_{p+r/2}^A} - 1\right| \geq \epsilon\right) = 0.$$

All in all,

$$\mathbb{P}\left(\left|\frac{\sigma_{p,r}}{\mathbb{E}(E^r)\sigma_{p+r/2}} - 1\right| \geq \epsilon(2^3 - 1)\right) \xrightarrow[t \rightarrow -\infty]{} 0.$$

As a consequence, the first assertion of Theorem 8.2.2 is true with $K_r = \mathbb{E}(E^r)$. \square

8.5.2 Concentration

We have

Lemma 8.5.1. *Let $A > 0$, $p \geq 2$, $t < 0$, $\sigma_p^A = \sum_i x_i^p(t)$. We have*

$$\mathbb{E}(\sigma_p) - \mathbb{E}(\sigma_p^A) \leq \int_A^\infty s^{p-5/2} \frac{e^{-t^2 s/2}}{\sqrt{2\pi}} ds.$$

In particular,

$$\left| 1 - \frac{\mathbb{E}(\sigma_p^A)}{\mathbb{E}(\sigma_p)} \right| \leq \frac{\int_{t^2 A}^{+\infty} s^{p-5/2} e^{-s/2} ds}{\int_0^{+\infty} s^{p-5/2} e^{-s/2} ds} = o(1).$$

Proof. According to Lemma 8.2.1,

$$\begin{aligned} \mathbb{E}(\sigma_p) - \mathbb{E}(\sigma_p^A) &= \int_A^\infty s^p \psi^t(ds) \\ &= \int_A^\infty s^{p-1} \rho^t(ds) \end{aligned}$$

Remark that, by integration by part, for any measure η on \mathbb{R}_+ ,

$$\begin{aligned} \int_A^\infty s^{p-1} \eta(ds) &= \int_A^\infty \int_0^{s^{p-1}} d\eta(ds) \\ &= \int_{\mathbb{R}_+^2} \mathbf{1}_{s \geq A} \mathbf{1}_{0 \leq u^{1/(p-1)} \leq s} \eta(ds) du \\ &= \int_{\mathbb{R}_+^2} \mathbf{1}_{A^{p-1} \leq u} \mathbf{1}_{s \geq u^{1/(p-1)}} \eta(ds) du + \int_{\mathbb{R}_+^2} \mathbf{1}_{A^{p-1} \geq u} \mathbf{1}_{s \geq A} \eta(ds) du \\ &= \int_{A^{p-1}}^\infty \eta([u^{1/(p-1)}, +\infty)) du + \int_0^{A^{p-1}} \eta([A, +\infty)) du. \end{aligned}$$

As a consequence, as $\rho^t([s, +\infty)) \leq \nu^t([s, +\infty))$,

$$\int_A^\infty s^{p-1} \rho^t(ds) \leq \int_A^\infty s^{p-1} \nu^t(ds).$$

$$\mathbb{E}(\sigma_p) - \mathbb{E}(\sigma_p^A) \leq \int_A^\infty s^{p-5/2} \frac{e^{-t^2 s/2}}{\sqrt{2\pi}} ds.$$

This proves the first assertion of Lemma 8.5.1. The second assertion comes from the change of variable $s \rightarrow s/t^2$ in $\int_A^\infty s^{p-5/2} \frac{e^{-t^2 s/2}}{\sqrt{2\pi}} ds$. \square

proof of (8.2) in Theorem 8.2.2. As a particular case of (8.5) where $r = 0$, we have for $\epsilon > 0$,

$$\lim_{A \rightarrow 0} \limsup_{t \rightarrow -\infty} \mathbb{P}\left(\left| \frac{\sigma_p}{\mathbb{E}(\sigma_p^A)} - 1 \right| \geq \epsilon\right) = 0.$$

On the other hand, for $A > 0$, if $\left| \frac{\sigma_p}{\mathbb{E}(\sigma_p)} - 1 \right| = \left| \frac{\sigma_p}{\mathbb{E}(\sigma_p^A)} \frac{\mathbb{E}(\sigma_p^A)}{\mathbb{E}(\sigma_p)} - 1 \right| \geq 3\epsilon$, we have that $\left| \frac{\sigma_p}{\mathbb{E}(\sigma_p^A)} - 1 \right| \geq \epsilon$ or $\left| \frac{\mathbb{E}(\sigma_p^A)}{\mathbb{E}(\sigma_p)} - 1 \right| \geq \epsilon$. Thanks to Lemma 8.5.1, $\mathbb{E}(\sigma_p^A) \sim \mathbb{E}(\sigma_p)$ so by taking the limit when $n \rightarrow +\infty$, then $A \rightarrow 0$, we get (8.2). \square

8.5.3 Largest mass

Proof of (8.3) in Theorem 8.2.2. We have, for $r \geq 0$,

$$\begin{aligned} \mathbb{P}\left(\frac{x_0}{\mathbb{E}(\sigma_2)^p} > \epsilon\right) &\leq \mathbb{P}(\sigma_r > \epsilon^r \mathbb{E}(\sigma_2)^{pr}) \\ &\leq \frac{\mathbb{E}(\sigma_r)}{\epsilon^r \mathbb{E}(\sigma_2)^{pr}} \end{aligned}$$

which goes to zero as t goes to $-\infty$ as long as $p < \frac{2r-3}{r}$, thanks to Lemma 8.2.1. By taking any $r > 0$, we have that

$$\forall p < 2, x_0 = o_{\mathbb{P}}(\mathbb{E}(\sigma_2)^p)$$

Moreover, as $\sigma_2 \sim_{\mathbb{P}} \mathbb{E}(\sigma_2)$, we have that

$$\forall p < 1, \frac{x_0}{\sigma_2} = o_{\mathbb{P}}(\sigma_2^p)$$

which proves (8.3) in Theorem 8.2.2. □

8.5.4 Maximum diameter

Proof of (8.4) in Theorem 8.2.2. We have, for $p \geq 2$, $q \in (0, 1)$,

$$\begin{aligned} \mathbb{P}(D_{\max} > \mathbb{E}(\sigma_2)^q) &\leq \mathbb{P}(D_{\max} > \mathbb{E}(\sigma_2)^q \text{ and } x_0 \leq A) + \mathbb{P}(x_0 > A) \\ &\leq \mathbb{P}(D_{\max} 1_{x_0 \leq A} > \mathbb{E}(\sigma_2)^q) + \mathbb{P}(x_0 > A) \\ &\leq \frac{\mathbb{E}(D_{\max}^{2p} 1_{x_0 \leq A})}{\mathbb{E}(\sigma_2)^{2qp}} + \mathbb{P}(x_0 > A) \end{aligned}$$

so using Lemma 8.3.2,

$$\begin{aligned} \mathbb{P}(D_{\max} > \mathbb{E}(\sigma_2)^q) &\leq O\left(\frac{\mathbb{E}(\sigma_p^A)}{\mathbb{E}(\sigma_2)^{2qp}}\right) + \mathbb{P}(x_0 > A) \\ &\leq O\left(\frac{\mathbb{E}(\sigma_p)}{\mathbb{E}(\sigma_2)^{2qp}}\right) + \mathbb{P}(x_0 > A) \end{aligned}$$

Finally, using Lemma 8.2.1,

$$\mathbb{P}(D_{\max} > \mathbb{E}(\sigma_2)^q) \leq O(|t|^{3-2p(1-q)}) + \mathbb{P}(x_0 > A)$$

so by taking $\frac{3}{2(1-q)} < p$,

$$\mathbb{P}(D_{\max} > \mathbb{E}(\sigma_2)^q) \xrightarrow{n \rightarrow \infty} 0$$

which proves (8.4) in Theorem 8.2.2. □

Chapter 9

Noise sensitivity for critical Erdős–Rényi random graphs

9.1 Asymptotic Independence

Here we show how to get noise sensitivity results from mixing results like Corollary 8.2.3. Recall from Section 2.5.3 the general context of noise sensitivity. Before giving actual noise sensitivity results we are looking in what situations

$$\mathbb{E}(f(G^{n,\epsilon})g(G^n)) - \mathbb{E}(f(G^n))\mathbb{E}(g(G^n)) \xrightarrow[n \rightarrow +\infty]{} 0.$$

First we show that the couple $(G^{n,\epsilon}, G^n)$ can be approximated by a couple of multiplicative coalescents:

Theorem 9.1.1. *For $n \in \mathbb{N}$, let $G^n \sim G(n, \frac{1}{n} + \frac{t}{n^{4/3}})$ and $G^{n,\epsilon}$ be an ϵ resampling of G^n . Assume that $n^{1/3}\epsilon \xrightarrow[n \rightarrow +\infty]{} +\infty$. We have*

$$(G^{n,\epsilon}, G^n) \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{E}_t(1, 0, 0, 1) \times \mathcal{E}_t(1, 0, 0, 1).$$

with respect to T_{GHPS} .

Proof. Let f be bounded continuous for T_{GHPS} . Thanks to Proposition 3.3 in [29], for any $s \geq 0$, for n large enough

$$\mathbb{E}(f(G^{n,\epsilon})f(G^n)) - \mathbb{E}(f(G^{n,\epsilon}))\mathbb{E}(f(G^n)) \leq \mathbb{E}(f(G^{n,sn^{-1/3}})f(G^n)) - \mathbb{E}(f(G^{n,sn^{-1/3}}))\mathbb{E}(f(G^n)).$$

(reference) As a special case of Theorem 3.7 from [28], $(G^{n,sn^{-1/3}}, G^n)_{n \in \mathbb{N}}$ converges in distribution towards $(\text{Perc}(X(t), s), X(t))$ with respect to T_{GHP} . Thanks to Lemma 6.1.2, this can be extended to a convergence with respect to T_{GHPS} , because $G^{n,\epsilon} \sim G^n$ and $G^n \xrightarrow[n \rightarrow +\infty]{(d)} \mathcal{E}_t(1, 0, 0, 1)$ with respect to T_{GHPS} thanks to Theorem 4.1 of [3]. As a consequence, for all $s \in \mathbb{R}_+$,

$$\limsup_n \mathbb{E}(f(G^{n,\epsilon})f(G^n)) - \mathbb{E}(f(G^{n,\epsilon}))\mathbb{E}(f(G^n)) \leq \mathbb{E}(f(\text{Perc}(X(t), s))f(X(t))) - \mathbb{E}(f(\text{Perc}(X(t), s)))\mathbb{E}(f(X(t))).$$

Thanks to Corollary 8.2.3, this allows us to conclude that

$$\limsup_n \mathbb{E}(f(G^{n,\epsilon})f(G^n)) - \mathbb{E}(f(G^{n,\epsilon}))\mathbb{E}(f(G^n)) = 0.$$

In order to prove the theorem in its full generality, we need to prove the same thing with f and g distinct. Remark that

$$\begin{aligned} & \mathbb{E}((f(G^{n,\epsilon}) + g(G^{n,\epsilon}))(f(G^n) + g(G^n))) - \mathbb{E}((f(G^{n,\epsilon}) - g(G^{n,\epsilon}))(f(G^n) - g(G^n))) \\ &= 2\mathbb{E}(f(G^{n,\epsilon})g(G^n)) + 2\mathbb{E}(g(G^{n,\epsilon})f(G^n)) \\ &= 4\mathbb{E}(f(G^{n,\epsilon})g(G^n)) \end{aligned}$$

because $(G^{n,\epsilon}, G^n) \stackrel{(d)}{=} (G^n, G^{n,\epsilon})$. We can then obtain the convergence of $\mathbb{E}(f(G^{n,\epsilon})g(G^n))$ as a consequence of the convergence of $\mathbb{E}((f(G^{n,\epsilon}) + g(G^{n,\epsilon}))(f(G^n) + g(G^n)))$ and $\mathbb{E}((f(G^{n,\epsilon}) - g(G^{n,\epsilon}))(f(G^n) - g(G^n)))$, for which the first part of the proof apply. \square

Notice that Theorem 9.1.1 is optimal because if there exists $s \in \mathbb{R}$ such that $n^{1/3}\epsilon \xrightarrow{n \rightarrow +\infty} s$, according to Theorem 3.7 from [28], $(G^{n,\epsilon}, G^n)$ converges in distribution toward $(\text{Perc}(X(t), s), X(t))$. Notice also that Theorem 9.1.1 is an improvement of Theorem 2 in [23], where the condition is that $\epsilon \geq n^{-1/3+\delta}$. Moreover, while the core ideas are similar, the method is completely different, because our proof relies heavily on properties of the limiting multiplicative coalescent, which should make it more robust.

9.2 Some random graphs background

In order to retrieve noise sensitivity results for specific properties from Theorem 9.1.1, we first need to show that an event that would be candidate to be noise sensitive, can be read on the largest components alone. Recall that when we talk about the ordering of the components of a graph G , the order is respective to the number of vertices in the components, and if this criteria is not sufficient, we can break ties by looking at which component has the smallest label. Before introducing our result, consider the following definition: we say that a graph G possesses a bicycle if there exists an injective path (v_1, \dots, v_k) in G such that there is an edge between v_0 and one of the (v_3, \dots, v_k) , and there is an edge between v_k and one of the (v_1, \dots, v_{k-2}) . Remark that if there is a component of surplus at least 2 in G , there is a bicycle.

Lemma 9.2.1. *Let A be a graph property such that*

$$\begin{cases} A \text{ is stable under addition of edges between two existing vertices} \\ \neg A \text{ is stable under deletion of vertices.} \end{cases} \quad (9.1)$$

Moreover suppose that either A implies that the graph possesses a component with a bicycle, either A implies that there is a cycle of length greater than $\alpha n^{1/3}$ for some fixed $\alpha > 0$. Let $G^n \sim G(n, p)$ and $(C_j^n)_{j \in \mathbb{N}}$ be its associated ordered sequence of components. Suppose that $pn = 1 + O(n^{-1/3})$. Then

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\bigsqcup_{j \geq K+1} C_j^n \in A\right) = 0.$$

This lemma is the consequence of the three following lemmas:

Lemma 9.2.2. *Let A be a graph property such that (9.1) is true. Let $n \in \mathbb{N}$, $p \in [0, 1]$, $G^n \sim G(n, p)$ and $(C_i^n)_{i \in \mathbb{N}}$ the ordered sequence of component of G^n . Then*

$$\forall K \in \mathbb{N}, m \in \mathbb{N}, \mathbb{P}(\bigsqcup_{i \geq K+1} C_i^n \in A \text{ and } \sum_{1 \leq i \leq K} |C_i^n| \geq m) \leq \mathbb{P}(G^{n-m} \in A).$$

Proof. For V a discrete set, let $\mathbb{G}(V)$ be the set of graphs on a subset of V . For $G \in \mathbb{G}(V)$, let $V(G) \subset V$ be the set of vertices of G , $C(G) = (C_i(G))_{1 \leq i \leq k}$ the ordered sequence of connected components of G and for each $i \in [k]$, $|C_i(G)| = |V(C_i(G))|$. We have that:

$$\mathbb{P}\left(\bigsqcup_{j \geq K+1} C_j^n \in A \text{ and } \sum_{1 \leq i \leq K} |C_i^n| > m\right) = \sum_{i=m+1}^n \sum_{\substack{G \in \mathbb{G}([n]) \\ \# \text{comp}(G)=K \\ |C_1(G)|+\dots+|C_K(G)|=i}} \sum_{\substack{G' \in \mathbb{G}([n] \setminus V(G)) \\ C_1(G') < C_K(G) \\ G' \in A}} \mathbb{P}(G^n = G \cup G'). \quad (9.2)$$

Moreover,

$$\mathbb{P}(G^n = G \cup G') = \mathbb{P}(G^i = G) \mathbb{P}(G^{n-i} = G') (1-p)^{l(n-l)}$$

so

$$\begin{aligned}
\mathbb{P}\left(\bigsqcup_{j \geq K+1} C_j^n \in A \text{ and } \sum_{1 \leq i \leq K} |C_i^n| > m\right) &= \sum_{i=m+1}^n \sum_{\substack{G \in \mathbb{G}([n]) \\ \#\text{comp}(G)=K \\ |C_1(G)|+\dots+|C_K(G)|=i}} \mathbb{P}(G^i = G)(1-p)^{i(n-i)} \\
&\times \sum_{\substack{G' \in \mathbb{G}([n] \setminus V(G)) \\ C_1(G') < C_K(G) \\ G' \in A}} \mathbb{P}(G^{n-i} = G') \\
&= \sum_{i=m+1}^n \sum_{\substack{G \in \mathbb{G}([n]) \\ \#\text{comp}(G)=K \\ |C_1|+\dots+|C_K|=i}} \mathbb{P}(G^i = G)(1-p)^{i(n-i)} \\
&\times \mathbb{P}(\{G^{n-i} \in A\} \cap \{C_1(G^{n-i}) < C_K(G)\}).
\end{aligned}$$

By FKG's inequality, as A is stable when adding edges,

$$\mathbb{P}(\{G^{n-i} \in A\} \cap \{C_1(G^{n-i}) < C_K(G)\}) \leq \mathbb{P}(G^{n-i} \in A)\mathbb{P}(C_1(G^{n-i}) < C_K(G)).$$

Moreover, as A is stable under the deletion of vertices and $i \geq m$,

$$\mathbb{P}(G^{n-i} \in A) \leq \mathbb{P}(G^{n-m} \in A).$$

Finally,

$$\begin{aligned}
\mathbb{P}\left(\bigsqcup_{j \geq K+1} C_j^n \in A \text{ and } \sum_{1 \leq i \leq K} |C_i^n| > m\right) &\leq \mathbb{P}(G^{n-m} \in A) \sum_{i=m+1}^n \sum_{\substack{G \in \mathbb{G}([n]) \\ \#\text{comp}(G)=K \\ |C_1|+\dots+|C_K|=i}} \mathbb{P}(G^i = G)(1-p)^{i(n-i)} \\
&\times \mathbb{P}(C_1(G^{n-i}) < C_K(G)) \\
&= \mathbb{P}(G^{n-m} \in A)\mathbb{P}\left(\sum_{1 \leq i \leq K} |C_i^n| > m\right) \\
&\leq \mathbb{P}(G^{n-m} \in A).
\end{aligned}$$

□

This lemma tells us that, provided that the largest components are large enough, the probability that the property A is true for the smallest components can be controlled by the probability of the same event in a smaller graph, which is by consequence more subcritical. For this reason, we need a tool that gives that for graphs subcritical enough, having a bicycle or having a large cycle are highly improbable.

Lemma 9.2.3. *Let B_n be the number of bicycles of $G^n \sim G(n, p)$. Suppose that $pn < 1$. Then:*

$$\mathbb{E}(B_n) \leq \frac{2p}{(1-pn)^3}.$$

Proof. The number of distinct paths (v_1, \dots, v_k) of length k in the complete graph is $n!/(n-k)! \leq n^k$. On each of these paths, we can join the first vertex v_1 to $k-2$ vertices which are (v_3, \dots, v_k) . The last vertex v_k can be joined to at most any of the $k-2$ vertices (v_1, \dots, v_{k-2}) , with v_1 being possibly unavailable. All in all, the number of bicycles of size k in a graph of size n is at most $n^k(k-2)^2$. As a consequence,

$$\begin{aligned}
\mathbb{E}(B_n) &\leq \sum_{4 \leq k \leq n} \mathbb{E}(X_k) \\
&\leq \sum_{4 \leq k \leq n} n^k(k-2)^2 p^{k+1} \\
&\leq n^2 p^3 \sum_{4 \leq k} k(k-1)(pn)^{k-2} \\
&\leq n^2 p^3 \sum_{2 \leq k} k(k-1)(pn)^{k-2}
\end{aligned}$$

As $pn < 1$,

$$\begin{aligned}\mathbb{E}(B_n) &\leq n^2 p^3 \frac{2}{(1-pn)^3} \\ &\leq \frac{2p}{(1-pn)^3}\end{aligned}$$

□

Lemma 9.2.4. *Let $m, n \in \mathbb{N}$, $p \in (0, 1]$. Let L_n^m be the number of distinct cycles of size greater than m in $G^n \sim G(n, p)$ with $pn < 1$. Then*

$$\mathbb{E}(L_n^m) \leq \frac{1}{\ln(1 - (pn)^m)}.$$

Proof. For $k \leq n$, let l_n^k be the number of distinct cycles of size k in G^n . The number of distinct ordered cycles of length k in the complete graph is $\frac{n!}{2k(n-k)!} \leq \frac{n^k}{k}$. As a consequence, for $m \geq 1$,

$$\begin{aligned}\mathbb{E}(L_n^m) &= \sum_{m \leq k \leq n} \mathbb{E}(l_n^k) \\ &\leq \sum_{m \leq k} \frac{n^k}{k} p^k \\ &= \sum_{0 \leq k} \frac{(pn)^{k+m}}{k+m} \\ &= \sum_{i \geq 1} \sum_{j=0}^{m-1} \frac{(pn)^{mi+j}}{mi+j} \\ &\leq \sum_{i \geq 1} \frac{(pn)^{mi}}{mi} \sum_{j=0}^{m-1} 1 \\ &= \sum_{i \geq 1} \frac{(pn)^{mi}}{mi} m \\ &= \sum_{i \geq 1} \frac{(pn)^{mi}}{i} \\ &= -\ln(1 - (pn)^m).\end{aligned}$$

□

We can now provide a proof of Lemma 9.2.1:

Proof of Lemma 9.2.1. Since there exists λ' such that $pn \leq 1 + \lambda' n^{-1/3}$, for $\lambda > \lambda'$, if $m = \lceil \lambda n^{2/3} \rceil$, we have

$$p(n-m) \leq 1 + (\lambda' - \lambda)n^{-1/3} < 1.$$

Moreover,

$$\begin{aligned}\mathbb{P}\left(\bigsqcup_{j \geq K+1} C_j^n \in A\right) &\leq \mathbb{P}\left(\bigsqcup_{j \geq K+1} C_j^n \in A \text{ and } \sum_{1 \leq i \leq K} |C_i^n| > m\right) + \mathbb{P}\left(\sum_{1 \leq i \leq K} |C_i^n| \leq m\right) \\ &\leq \mathbb{P}(G^{n-m} \in A) + \mathbb{P}\left(\sum_{1 \leq i \leq K} |C_i^n| \leq m\right).\end{aligned}$$

As either A implies that there exists a bicycle, either A implies that there exists a cycle of size greater than $\alpha n^{1/3}$ for $\alpha > 0$,

$$\mathbb{P}(G^{n-m} \in A) \leq \mathbb{E}(B_{n-m}) \vee \mathbb{E}(L_{n-m}^{\alpha n^{1/3}})$$

with B_{n-m} and $L_{n-m}^{\alpha n^{1/3}}$ from Lemma 9.2.3 and lemma 9.2.4. From those lemmas, either

$$\begin{aligned}\mathbb{P}(G^{n-m} \in A) &\leq \frac{2p}{(1-p(n-m))^3} \\ &\leq \frac{2p}{(1-(1+(\lambda'-\lambda)n^{-1/3}))^3} \\ &= \frac{2pn}{(\lambda-\lambda')^3} \\ &\leq \frac{2}{(\lambda-\lambda')^3}\end{aligned}$$

either

$$\begin{aligned}\mathbb{P}(G^{n-m} \in A) &\leq \frac{1}{\ln(1-(p(n-m))^{\alpha n^{1/3}})} \\ &\leq \frac{1}{\ln(1-(1+(\lambda'-\lambda)n^{-1/3})^{\alpha n^{1/3}})}.\end{aligned}$$

Moreover,

$$\begin{aligned}\ln((1+(\lambda'-\lambda)n^{-1/3})^{\alpha n^{1/3}}) &= \alpha n^{1/3} \ln(1+(\lambda'-\lambda)n^{-1/3}) \\ &\sim \alpha(\lambda'-\lambda)\end{aligned}$$

so for some $H > 0$, $(1+(\lambda'-\lambda)n^{-1/3})^{\alpha n^{1/3}} \leq e^{\alpha(\lambda'-\lambda)H}$. As a consequence,

$$\mathbb{P}(G^{n-m} \in A) \leq \frac{1}{\ln(1-e^{-\alpha(\lambda-\lambda')H})}.$$

From this, we can conclude, considering that for any fixed $\lambda > 0$,

$$\limsup_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P}\left(\sum_{1 \leq i \leq K} |C_i^n| \leq \lceil \lambda n^{2/3} \rceil\right) = 0$$

from Theorem 2.2.8, and then taking the successive limits in n , K and $\lambda \rightarrow +\infty$. \square

9.3 Some noise sensitive events

Theorem 9.1.1 says that with a noise of level large with respect to $n^{1/3}$, $G^{n,\epsilon}$ and G^n become asymptotically independent. This is a pretty satisfying result as such, but we will go further and give some examples of such events.

Let \mathbb{K} be the set of finite connected multigraphs such that either $K \in \mathbb{K}$ has minimal degree at least 3, either it is a single vertex with a loop. Recall from [3] Section 6.2 the definition of the core $\text{core}(G)$ and the kernel $\text{ker}(G)$ of a real graph G . Let $E(\text{ker}(G))$ be the multiset of edges of $\text{ker}(G)$. Recall that $\text{core}(G)$ can be reconstructed from $\text{ker}(G)$ given a family of length $(l_G(e))_{e \in E(\text{ker}(G))}$. Lemma 6.7 of [28] implies the following:

Lemma 9.3.1 (Lemma 6.7 of [28]). *Let G be a real graph. For every real graph G' with the same surplus as G , for all $\epsilon > 0$, there exists $\eta > 0$ such that*

$$d_{GHP}(G, G') \leq \eta \Rightarrow \exists \phi : \text{ker}(G) \rightarrow \text{ker}(G') \text{ an isomorphism s.t. } \forall e \in E(\text{ker}(G)), |l_G(e) - l_{G'}(\phi(e))| \leq \epsilon.$$

This lemma is an important tool while proving the following theorem:

Theorem 9.3.2. For $K \in \mathbb{K}$, $a = (a_e)_{e \in E(K)} \in \mathbb{R}_+^{E(K)}$, consider the following set:

$$S_{K,a} = \left\{ X \text{ is a real graph} : \exists \delta \geq a \text{ s.t. } (K, \delta) \prec X_j \right\}$$

where $(K, \delta) \prec G$ denotes the fact that the metric space defined by equipping K with the edge length δ is isometrically embeddable in G . Let

$$S_{K,a}^\infty = \left\{ X \text{ is a collection of real graph} : \exists j \leq k \text{ s.t. } X_j \in S_{K,a} \right\}$$

Moreover, suppose that if K is a single vertex v with a loop e , $a_e > 0$. Then $G^n \in S_{K,a}^\infty$ is noise sensitive for a noise level ϵ such that $\epsilon n^{1/3} \rightarrow +\infty$. In particular, the event that G^n possesses a cycle of size larger than $\alpha > 0$ is noise sensitive for a noise level ϵ such that $\epsilon n^{1/3} \rightarrow +\infty$.

Remark that Theorem 9.3.2 is a generalisation of Theorem 1.3. where the authors obtained, with vastly different methods, the noise sensitivity with the same level ϵ for the property of having a cycle of size in $(\alpha n^{1/3}, \beta n^{1/3})$. In our case, we removed the upper bound $\beta n^{1/3}$ for the sake of simplicity.

Proof. We do the proof for K with minimal degree at least 3, the proof for the single loop is different but follows the same layout. For $k \in \mathbb{N}$, let

$$S_{K,a}^k = \left\{ X \text{ is a collection of real graph} : \exists j \leq k \text{ s.t. } X_j \in S_{K,a} \right\}.$$

First we show that $S_{K,a}$ is closed under T_{GHP} . Let X^n be a sequence of real graphs of $S_{K,a}$ that converges toward X for d_{GHP} plus surplus. As the convergence with respect to the surplus is discrete, we can suppose that for all n , X^n has the same surplus. Then from Lemma 9.3.1, we get that for n large enough,

$$\exists \phi_n : \ker(X) \rightarrow \ker(X^n) \text{ an isomorphism}$$

and

$$\forall e \in E(\ker(X)), |l_X(e) - l_{X^n}(\phi_n(e))| \xrightarrow{n \rightarrow \infty} 0.$$

On the other hand, as for all $n \in \mathbb{N}$, X^n is in $S_{K,a}$ and K has minimal degree at least 3, there exists a multigraph homomorphism $\psi_n : K \rightarrow \ker(X^n)$ such that

$$\forall e \in E(K), |\delta_e^n - l_{X^n}(\psi_n(e))| \xrightarrow{n \rightarrow \infty} 0.$$

As a consequence, for n large enough, there exists a graph homomorphism $\phi_n^{-1} \circ \psi_n : K \rightarrow \ker(X)$. Moreover, because K and $\ker(X)$ are both finite, the number of such homomorphism is finite so, at least along a subsequence, we can suppose that such that $\phi_n^{-1} \circ \psi_n = h$ does not depend on n . Lastly,

$$\forall e \in E(K), |\delta_e^n - l_X(h(e))| \xrightarrow{n \rightarrow \infty} 0$$

so in particular, for any $e \in E(K)$, $l_X(h(e)) \geq a(e)$ so $X \in S_{K,a}$. The same ideas show that if $X^n \xrightarrow{n \rightarrow \infty} X$ and for all $n \in \mathbb{N}$, $X^n \notin S_{K,a}$, there exists $e \in E(K)$, $\delta \geq a$ and a homomorphism $h : K \rightarrow X$ such that $\delta_{h(e)} = a_e$. As a consequence,

$$\partial S_{K,a} \subset \bigcup_{e \in E(K)} \{X \text{ is a real graph and } \exists e' \in \ker(X) \text{ s.t. } l_X(e') = a(e)\}.$$

From [2], Theorem 4, we get that conditionally on the kernel, the length of the edges used to construct the core of $X \sim \mathcal{E}(1, 0, 0, 1)$ are distributed according to a continuous law so $\{X \in \partial S_K^k\}$ has probability 0. As a consequence, thanks to Theorem 9.1.1 and the Portmanteau lemma,

$$\begin{aligned} & \mathbb{P}(G^{n,\epsilon} \text{ and } G^n \in S_K^k) - \mathbb{P}(G^{n,\epsilon} \in S_K^k) \mathbb{P}(G^n \in S_K^k) \\ & \xrightarrow{n \rightarrow +\infty} \mathbb{P}(X \text{ and } Y \in S_K^k) - \mathbb{P}(X \in S_K^k) \mathbb{P}(Y \in S_K^k) = 0 \end{aligned}$$

where $X \stackrel{(d)}{=} Y \sim \mathcal{E}(1, 0, 0, 1)$ and X and Y are independent. Finally, notice that S_K is stable when adding edges between two existing vertices and $\lrcorner S_K$ is stable when deleting vertices. As a consequence, because K has minimal degree at least 3, per Lemma 9.2.1,

$$\lim_{k \rightarrow +\infty} \limsup_{n \rightarrow \infty} \mathbb{P}(G^n \in S_K \setminus S_K^k) = 0.$$

Moreover,

$$\lim_{k \rightarrow +\infty} \mathbb{P}(X \in S_K \setminus S_K^k) = 0,$$

which allows to conclude on the convergence part of the noise sensitivity. Finally $\mathbb{P}(X \in S_K) \in (0, 1)$ again from [2], Theorem 4. \square

Notice that the proof of Theorem 9.3.2 could be used, via the Kuratowski characterization of planar graphs to deduce the noise sensitivity of planarity. However more general results of this type can be easily and directly deduced. In the following theorem, a closed surface will designate a connected compact topological space such that every point admits a neighbourhood homeomorph to \mathbb{R}^2 , see [25].

Theorem 9.3.3. *Let S be a closed surface. The fact that each component of $G^n \sim G(n, p)$ can be embedded in S is noise sensitive for a noise level ϵ such that $\epsilon n^{1/3} \rightarrow +\infty$.*

Proof. Let $P = \left\{ X \in \mathbb{M}_2^{\mathbb{N}} : \exists j \in \mathbb{N} \text{ s.t. } X_j \text{ is NOT embeddable in } S \right\}$ for S a surface. Remark that P is stable when adding edges between existing vertices, $\neg P$ is stable when deleting vertices, and P implies that there is a component with surplus at least 2, because unicyclic components are embeddable in all surfaces. As a consequence of Lemma 9.2.1, $\neg P$ can be read on the largest components of G^n , i.e., if for $k \in \mathbb{N}$, $P^k = \left\{ X \in \mathbb{M}_2^{\mathbb{N}} : \exists j \leq k \text{ s.t. } X_j \text{ is NOT embeddable in } S \right\}$ then $\lim_{k \rightarrow \infty} \limsup_{n \rightarrow +\infty} \mathbb{P}(G^n \in P \setminus P^k) = 0$. Suppose that $(X^n)_{n \in \mathbb{N}}$ is a sequence of P^k such that $X^n \xrightarrow[n \rightarrow \infty]{} X$ for T_{GHPS} . As for n big enough, the kernel of X_j^n is isomorphic to the kernel of X_j , thanks to Lemma 9.3.1, so we get that $X \in P^k$. Moreover, if $(X^n)_{n \in \mathbb{N}}$ is a sequence of $\neg P^k$, for the same reason, $X \in \neg P^k$. As a consequence, P_k is closed and open so $\mathbb{P}(G^n \in P^k) \xrightarrow[n \rightarrow \infty]{} \mathbb{P}(X \in P^k)$. Combined with the fact that $\lim_{k \rightarrow \infty} \limsup_{n \rightarrow +\infty} \mathbb{P}(G^n \in P \setminus P^k) = 0$, this allows to conclude, thanks to Theorem 9.1.1 and the Portmanteau lemma, that

$$\begin{aligned} & \mathbb{P}(G^{n,\epsilon} \text{ and } G^n \in P) - \mathbb{P}(G^{n,\epsilon} \in P)\mathbb{P}(G^n \in P) \\ & \xrightarrow[n \rightarrow +\infty]{} \mathbb{P}(X \text{ and } Y \in P) - \mathbb{P}(X \in P)\mathbb{P}(Y \in P) = 0 \end{aligned}$$

with X and Y independent. Finally, we have that $\mathbb{P}(X \in P) \in (0, 1)$ because from [2], Theorem 4, every 3-regular graph have a positive probability of being a subgraph of the kernel so, $\mathbb{P}(X \in P) > 0$ and there is a positive probability that all components have surplus at most 1, in which case X is embeddable in all surfaces. \square

Part IV

Perspectives

Chapter 10

Toward a characterization of all coalescence processes of compact measured metric spaces

This chapter is devoted to expose ideas and partial results to complete our program. It does not have the same level of rigor than the previous chapters

10.1 What we have, what we want

To understand what we accomplished in this work, we need to restate briefly what would be the ultimate underlying goals, which are not yet fulfilled:

- (i) For all $(\alpha, \lambda, c, \gamma) \in \mathcal{I}$, $X \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$, for $t > 0$ we can define $(\text{Perc}(X(t), s))_{s \in \mathbb{R}^+}$. Moreover,

$$(\text{Perc}(X(t), s), X(t)) \stackrel{(d)}{=} (\text{Coal}(X(t-s), s), X(t))$$

where the coalescence $\text{Coal}(\cdot, s)$ is independent of the coalescence used to obtain $X(t)$ from $X(t-s)$.

- (ii) For all $(\alpha, \lambda, c, \gamma) \in \mathcal{I}$, $X \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$,

$$\forall t \in \mathbb{R}, (\text{Perc}(X(t), s), X(t)) \xrightarrow[s \rightarrow +\infty]{(d)} \mathcal{E}(\alpha, \lambda, c, \gamma) \times \mathcal{E}(\alpha, \lambda, c, \gamma).$$

- (iii) If X is an eternal extreme multiplicative coalescent in $\mathbb{M}_2^{\mathbb{N}}$, that is moreover with full support almost surely, then there exists $(\alpha, \lambda, c, \gamma) \in \mathcal{I}$ such that $X \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$.

For now (i) and (ii) are proven to be true for the specific parameters $(\alpha, \lambda, 0, \gamma) \in \mathcal{I}$. In fact, (i) comes from [28] and (ii) comes from Corollary 8.2.3.

Extending (i) for all parameters could be obtained via Theorem 2.4.9, the Feller property Theorem 2.5.5 and a duality property as in [28]. There should not be any major obstacle.

Then, extending (ii) to the specific pure jump parameters $(0, \lambda, c, \gamma) \in \mathcal{I}^*$ could be done by using Theorem 6.2.1. In that case, we would need an asymptotic result on $(X(t))_{t \in \mathbb{R}} \sim \mathcal{E}(0, \lambda, c, \gamma)$ as t goes to $-\infty$. Luckily, our Theorem 7.1.1 admits all parameters so a similar analysis as the one of chapter 8 could be done. At this point the missing link would be a precise knowledge of the scaling of the metric structure of a component of $X(t)$ conditionally on its size. Note that it is less straightforward than for the homogeneous case as the metric structure of a component of a $X(t)$ depends both on its size and on its inhomogeneity.

Remark that the specificity of the parameters $(0, \lambda, c, \gamma) \in \mathcal{I}^*$ doesn't really play a role apart from the fact that Theorem 6.2.1 only works for parameters $(0, \lambda, c, \gamma) \in \mathcal{I}^*$ and $(\alpha, \lambda, 0, \gamma) \in \mathcal{I}^\circ$. This result itself inherits this limitation from the fact that it is nothing else than the only two existing results in the literature, viewed through the lenses of the Chapter 3, Chapter 4 and Chapter 5. However, if

$$\rho_2 = \sum_j \int_{X_j \times X_j} d_j(u, v)^2 \mu_j(du) \mu_j(dv),$$

we conjecture the following theorem:

Conjecture 1. Let $X = (X^n)_{n \in \mathbb{N}}$ be a sequence of $\overline{\mathbb{M}}^{\mathbb{N}}$, and $q = (q_n)_{n \in \mathbb{N}}$ such that (x, q) satisfies the Aldous-Limic conditions 1 with parameters $(\alpha, \lambda, c) \in \mathcal{I}$. Recall from Definition 2.5.6 the definition of x , ρ_1 and D . Assume that there exists $\gamma \geq 0$ such that:

$$\left\{ \begin{array}{l} \frac{\rho_1}{\sigma_2^2} \xrightarrow[n \rightarrow \infty]{} \gamma \\ \frac{\rho_2}{\sigma_2^2} \xrightarrow[n \rightarrow \infty]{} 0, \\ \frac{\sum_i x_i^3 u_i^2}{\rho_1} \xrightarrow[n \rightarrow \infty]{} 0, \end{array} \right.$$

Then we have that

$$\frac{\sigma_2^2}{\rho_1 + \sigma_2} S\text{Coal}_1(X, q) \xrightarrow[n \rightarrow +\infty]{d} \mathcal{E}_0(\alpha, \lambda, c, \gamma)$$

with respect to the Gromov-weak product topology.

Remark that $\rho_2 \leq D_{\max} \rho_1$ with D_{\max} being the max of the diameters, so $\frac{\rho_2}{\sigma_2^2} \xrightarrow[n \rightarrow \infty]{} 0$ is verified in particular in the case of the maximum of diameters hypothesis of [9] and [10].

Then we could once again pursue, verbatim, the analysis of Chapter 3, Chapter 4 and Chapter 5 and obtain:

Conjecture 2. Let $(X(t))_{t \in \mathbb{R}}$ be an eternal metric multiplicative coalescent. Let $t \in \mathbb{R}$. Suppose that there exists $(s_n)_{n \in \mathbb{N}}$, $s_n \xrightarrow[n \rightarrow \infty]{} +\infty$ such that $(X(t - s_n), s_n)$ satisfies in probability the Aldous and Limic conditions Hypothesis 1 with parameters $(\alpha, \lambda, c, \gamma) \in \mathcal{I}$, and

$$\left\{ \begin{array}{l} \frac{\rho_1}{\sigma_2^2} \xrightarrow[n \rightarrow \infty]{} \gamma \\ \frac{\rho_2}{\sigma_2^2} \xrightarrow[n \rightarrow \infty]{} 0, \\ \frac{\sum_i x_i^3 u_i^2}{\rho_1} \xrightarrow[n \rightarrow \infty]{} 0, \end{array} \right.$$

Then for any f and g measurable for the Gromov-Hausdorff-Prokhorov plus surplus product topology,

$$\mathbb{E}(f(\text{Coal}(X(t-s), s))g(X(t))) \xrightarrow[s \rightarrow \infty]{} \mathbb{E}(f(\mathcal{X}(t)))\mathbb{E}(g(\mathcal{X}(t)))$$

where $\mathcal{X} \sim \mathcal{E}(\alpha, \lambda, c, \gamma)$. In particular the couple $(\text{Coal}(X(t-s), s), X(t))$ converges in distribution towards $\mathcal{E}_t(\alpha, \lambda, c, \gamma) \times \mathcal{E}_t(\alpha, \lambda, c, \gamma)$

One of the consequences of this would of course be new noise sensitivity results, but starting from an inhomogeneous random graph $\text{gCoal}(x, \cdot)$ for $x \in \overline{l_{\downarrow}}$. For this purpose, one would have to examine discrete random graph considerations similar to Section 9.2, but for possibly strongly inhomogeneous random graphs.

In Section 10.2, we provide a possible lead toward a proof of Conjecture 1.

Lastly, equipped with Conjecture 2, one might be tempted to prove directly (iii). However, one would then need to obtain $-\infty$ asymptotic for any generic extreme eternal multiplicative coalescent, preventing one from using the powerful Theorem 7.1.1. Thus we need a completely new way of approaching these asymptotic. Nevertheless, our work provides some hints about how one could carry such a program. First, remark that all asymptotics only concerned by masses are covered in Aldous and Limic's work [6]. Just for the sake of the investigation, let's have a look at Janson and Spencer's formula from Theorem 7.1.1 applied with $f(X_j, X) = x_j^2 u_j \rho_1$. Remark that $\rho_1^2 = \sum_j f(X_j, X)$. Let $\phi(\lambda, c) = \mathbb{E}(\rho_1)$ when $X \sim \mathcal{E}_0(\alpha, \lambda, c, \gamma)$. We have that:

$$\mathbb{E}(\rho_1^2) = \mathbb{E}\left(\sum_j x_j^4 u_j^2\right) + \mathbb{E}\left(\sum_j x_j^2 u_j \phi(\lambda - x_j - \delta_2(X_j), c \setminus c(X_j))\right).$$

It gives that:

$$\text{Var}(\rho_1) = \mathbb{E}\left(\sum_j x_j^4 u_j^2\right) + \mathbb{E}\left(\sum_j x_j^2 u_j (\phi(\lambda - x_j - \delta_2(X_j), c \setminus c(X_j)) - \phi(\lambda, c))\right)$$

This formula is at the heart of the reasoning of Chapter 8. In this chapter we use that, if the masses are small enough, shortcuts in components should be ignored, which makes the difference

$$\phi(\lambda - x_j - \delta_2(X_j), c \setminus c(X_j)) - \phi(\lambda, c)$$

itself small. Let us focus on this specific heuristic for a moment. Suppose that shortcuts can effectively be ignored, to the extent that we can coalesce only the points that are not already on the same components. What happens to average distances when merging two distinct components? Let's suppose that we have two components of size x_1 and x_2 , with average distances u_1 and u_2 , and we identify two points, one from each component at random. Then the expectation of the resulting average distance is $u_{1+2} = \frac{x_1^2 u_1}{(x_1+x_2)^2} + \frac{x_2^2 u_2}{(x_1+x_2)^2} + \frac{2}{(x_1+x_2)^2} (u_1 \frac{x_2}{x_1} + u_2 \frac{x_1}{x_2})$. As a consequence, $(x_1 + x_2)^2 u_{1+2} \geq x_1^2 u_1 + x_2^2 u_2$. What this small computation implies very informally is that $\mathbb{E}_{(\alpha, \lambda, c, \gamma)}(\rho_1)$ is increasing in λ (the details of this are made precise in Lemma 10.4.4). In particular, $\phi(\lambda - x_j - \delta_2(X_j), c \setminus c(X_j)) - \phi(\lambda, c) < 0$ so

$$\text{Var}(\rho_1^2) < \mathbb{E}(\sum_j x_j^4 u_j^2).$$

Note that we are still in an imaginary context where one can "ignore" shortcuts and where Janson and Spencer's Formula from Theorem 7.1.1 applies. But then, as it is expected that $\mathbb{E}(\sum_j x_j^4 u_j^2)$ is of order of magnitude smaller than $\mathbb{E}(\rho_1)$, it seems that the problem of the concentration of ρ_1 become easier. In Section 10.4, we show how to make the "ignore shortcuts" part of this heuristic rigorous by giving a coupling between an eternal multiplicative coalescent and a process that effectively ignores shortcuts and we give leads to exploit this coupling in the spirit of the martingale methods used by Aldous and Limic in [6].

10.2 Toward a generalized universality theorem

10.3 Completion of the discrete process

One of the main features of the encoding from Broutin, Duquesne and Wang is that it allows them to counter one major downside of other descriptions: after one node is visited, it will not be visited again, which greatly infringes on any potentially Markovian property of these processes. To counter that effect, they define a completed version of Z , B and H where a vertex can be visited more than once.

Definition 10.3.1 (Completed process). *Let $x \in l_{\searrow}^2$, let $(\tau_k)_{k \in \mathbb{N}}$ be a collection of exponential random variables such that*

$$\mathbb{E}(\tau_k) = \frac{\sigma_1}{\sigma_2}$$

and let $(J_k)_{k \in \mathbb{N}}$ be a family of independent random variables on $[n]$ such that

$$\forall k \in \mathbb{N}, \mathbb{P}(J_k = i) = \frac{x_i}{\sigma_1}.$$

We define the completed algebraic load of the server at time t :

$$\mathring{Z}(t) = -\frac{t}{\sigma_2} + \sum_k \frac{x_{J_k}}{\sigma_2} 1_{\tau_k \leq t},$$

the load of the server at the same time \mathring{B} and the height process associated

$$\mathring{\alpha}_k(t) = \begin{cases} 1 & \text{if } \tau_k \leq t, \inf_{[\tau_k, t]} Z > Z(\tau_k^-) \\ 0 & \text{else.} \end{cases}, \mathring{H}(t) = \sum_k \sigma_2 \mathring{\alpha}_k(t).$$

In order to retrieve Z from \mathring{Z} , we simply ignore a node that has already appeared, as well as all his descendants. Surprisingly, this operation can be taken to the limit [14].

The work of Broutin, Duquesne and Wang, as it deals transparently with the parameters (α, λ, c) of the coalescence should be a good basis to show a theorem that would improve Theorem 2.5.11 and Theorem 2.5.12.

In order to do this, let us introduce a new form of coalescence:

Definition 10.3.2 (Thin coalescence). Let $X = (X, d, \mu) \in \mathbb{M}_2^{\mathbb{N}}$, let \mathcal{P} be a Poisson point process on $X^2 \times [0, +\infty)$ and for all $t \geq 0$, let

$$\mathcal{SP}_t = \left\{ (u, v) \in X^2 : \begin{cases} \exists i, j \in \mathbb{N}, i \neq j \text{ s.t. } (u, v) \in X_i \times X_j \text{ and } \mathcal{P} \cap X_i \times X_j \times [0, t) = \emptyset \\ \exists s \leq t, (u, v, s) \in \mathcal{P} \end{cases} \right\}.$$

and

$$P_t = \{u \in X : \exists v \in X, (u, v) \text{ or } (v, u) \in P_t\}.$$

We denote

$$\text{Thin}(X, t) = \text{Glue}(P_t, d_t^{\text{SCoal}}).$$

Notice that there is an injective map from $\text{Thin}(X, t)$ to $\text{SCoal}(X, t)$ that preserves the distances.

This will be useful in order to decouple two different issues:

- (i) As each ending point is distributed on a component X_j according to $\frac{\mu_j(\cdot)}{|X_j|}$, the convergence of $\text{Thin}(X, t)$ should depend essentially on the property of this measure, namely its expectation and its variance.
- (ii) The difference between $\text{SCoal}(X, t)$ and $\text{Thin}(X, t)$ depends on how this measure fills the component.

We conjecture that in regard to the metric hypothesis, we could obtain the convergence of $\text{Thin}(\cdot, \cdot)$ for Gromov-Hausdorff-Prokhorov, and then the convergence of $\text{SCoal}(\cdot, \cdot)$ for the Gromov-weak topology, with only assumptions about average quantities such as ρ_1 or ρ_2 . If we want to strengthen the convergence of $\text{SCoal}(\cdot, \cdot)$ into a convergence for the Gromov-Hausdorff-Prokhorov topology, we would then add the Lower-mass-bound property of $\text{SCoal}(\cdot, \cdot)$.

To better understand $\text{Thin}(\cdot, \cdot)$, we introduce:

Definition 10.3.3 (Modified (and completed) height process). Let $X \in \overline{\mathbb{M}}^{\mathbb{N}}$. Recall from Definition 10.3.1 the definition of the completed height process \mathring{H} . For $t > 0$, $k \in \mathbb{N}$, let $I_k(t)$ be the number of child of the vertex visited at time τ_k already visited until t . Let J_k be the label of the component visited at time τ_k . We define $(U_k)_{k \in \mathbb{N}}$ and $(V_k^i)_{i, k \in \mathbb{N}}$ where for $k \in \mathbb{N}$, $U_k \sim \frac{\mu_{J_k}}{|X_{J_k}|}$ is an entry point on the component X_{J_k} and for $i \in \mathbb{N}$, $V_k^i \sim \frac{\mu_{J_k}^i}{|X_{J_k}^i|}$ is an exit point from the component X_{J_k} toward its i -th children. Note that, conditionally on J_k , U_k and for all $i \in \mathbb{N}$, (V_k^i) are independent. Let $D_k^i = d(U_k, V_k^i)$ and

$$D_k(t) = D_k^{I_k(t)}.$$

We finally define

$$\tilde{H}(t) = \sum_k D_k(t) \mathring{\alpha}_k(t),$$

which will be our main object of interest. It represent the total distance through the multiple components traversed from the root of the underlying tree to the vertex visited at time t .

For $X \in \mathbb{M}_2^{\mathbb{N}}$, we define, for $i \in \mathbb{N}$

$$v_i = x_i^{-2} \int_{X_i \times X_i} d_i(u, v)^2 \mu_i(du) \mu_i(dv)$$

and

$$\rho_2 = \sum_i x_i^2 v_i.$$

Then as a first step toward a proof of Conjecture 2, we have:

Lemma 10.3.4 (Tightness of the modified height process). Let $X = (X^n)_{n \in \mathbb{N}}$ be a sequence of $\overline{\mathbb{M}}^{\mathbb{N}}$ and $q = (q_n)_{n \in \mathbb{N}}$ be a real sequence. Suppose that

$$\frac{\rho_2}{\sigma_2^2} = O(1)$$

Recall from Definition 10.3.3 the definition of \mathring{H} and \tilde{H} . Suppose that \mathring{H} is tight for the Skorokhod topology. Then \tilde{H} is tight for the Skorokhod topology. In particular, under the Aldous-Limic conditions 1, if $\frac{\rho_2}{\sigma_2^2} = O(1)$, \tilde{H} is tight for the Skorokhod topology.

Let us briefly state what would be needed to obtain a theorem for $\text{SCoal}(X, q)$ or $\text{Thin}(X, q)$ from this lemma.

- (i) A result of point wise convergence: for $u \in \mathbb{R}$, $\tilde{H}(u) - \dot{H}(u) \xrightarrow{\mathbb{P}} 0$.
- (ii) Adapting the way how Broutin, Duquesne and Wang manage to forget the vertex that reappears in the completed process.
- (iii) Adding to the height process the pinch points that restore their complexity to the components.

We are confident that none of these steps should constitute a major obstacle. They require however some fine understanding of the law of a vertex, conditioned to be on a branch of a given length from the root.

The following quantity will be useful in the rest of this section:

$$E_t = \frac{\sum_{k: \tau_k \leq t} \sum_{i=1}^{N_k} (D_k^i)^2}{\sigma_2}$$

where N_k denotes the total number of children of the vertex visited at time τ_k .

Lemma 10.3.5. *Let $s, t \in \mathbb{R}$, $s \leq t$, then there exists $r \in [s, t]$ such that*

$$|\tilde{H}(t) - \tilde{H}(s)| \leq \sqrt{E_t} (\sqrt{\dot{H}(t) - \dot{H}(r)} + \sqrt{\dot{H}(s) - \dot{H}(r)}).$$

In particular,

$$\max_{[0, t]} \tilde{H} \leq \sqrt{E_t} \sqrt{\max_{[s, t]} \dot{H}}$$

and for all $\delta > 0$,

$$\omega_{\tilde{H}}^t(\delta) \leq 2\sqrt{E_t} \sqrt{\omega_{\dot{H}}^t(\delta)}$$

where $\omega_{\tilde{H}}^t$ is the càdlàg modulus of \tilde{H} on the interval $[0, t]$.

Proof. Let $r \in [s, t]$ be such that $\dot{H}(r) = \inf_{[s, t]} \dot{H}$. In words, it is a time where the closest common ancestor of the vertices visited at time s and t is visited. Of course, we have that for $k \in \mathbb{N}$, if $\hat{\alpha}_k(r) = 1$, then both $\hat{\alpha}_k(s) = 1$ and $\hat{\alpha}_k(t) = 1$. Moreover, for $k \in \mathbb{N}$, $D_k(r) = D_k(s) = D_k(t)$. As a consequence,

$$\begin{aligned} \tilde{H}(s) - \tilde{H}(r) &= \sum_k D_k(s) \hat{\alpha}_k(s) - \sum_k D_k(r) \hat{\alpha}_k(r) \\ &= \sum_k D_k(s) (\hat{\alpha}_k(s) - \hat{\alpha}_k(r)). \end{aligned}$$

Then by using the Cauchy-Schwarz inequality with the vectors $(\sqrt{\sigma_2}(\hat{\alpha}_k(t) - \hat{\alpha}_k(r)))_{k \in \mathbb{N}}$ and $D_k(s)/\sqrt{\sigma_2}$,

$$|\tilde{H}(s) - \tilde{H}(r)| \leq \sqrt{\sum_k D_k(s)^2 / \sigma_2} \sqrt{\sum_k \sigma_2 (\hat{\alpha}_k(s) - \hat{\alpha}_k(r))^2}.$$

Note that $\sum_k (D_k(s))^2 / \sigma_2 \leq E_s \leq E_t$ so

$$|\tilde{H}(s) - \tilde{H}(r)| \leq \sqrt{E_t} \sqrt{\sum_k \sigma_2 (\hat{\alpha}_k(s) - \hat{\alpha}_k(r))^2}.$$

For the same reason

$$|\tilde{H}(t) - \tilde{H}(r)| \leq \sqrt{E_t} \sqrt{\sum_k \sigma_2 (\hat{\alpha}_k(t) - \hat{\alpha}_k(r))^2}.$$

As a consequence,

$$\begin{aligned} |\tilde{H}(t) - \tilde{H}(s)| &\leq |\tilde{H}(t) - \tilde{H}(r)| + |\tilde{H}(s) - \tilde{H}(r)| \\ &\leq \sqrt{E_t} \left(\sqrt{\sum_k \sigma_2 (\hat{\alpha}_k(t) - \hat{\alpha}_k(r))^2} + \sqrt{\sum_k \sigma_2 (\hat{\alpha}_k(s) - \hat{\alpha}_k(r))^2} \right)^2. \end{aligned}$$

Again because of the common ancestry, the quantity $\hat{\alpha}_k(t) - \hat{\alpha}_k(r)$ can only be 0 or 1 so

$$(\hat{\alpha}_k(t) - \hat{\alpha}_k(r))^2 = \hat{\alpha}_k(t) - \hat{\alpha}_k(r).$$

As a consequence,

$$|\tilde{H}(t) - \tilde{H}(s)| \leq \sqrt{E_t}(\sqrt{\dot{H}(t) - \dot{H}(r)} + \sqrt{\dot{H}(s) - \dot{H}(r)})$$

which proves the first inequality. By taking $s = 0$ in this inequality, automatically $r = 0$ and we get that for $t \in \mathbb{R}_+$, $\tilde{H}(t) \leq \sqrt{E_t} \sqrt{\dot{H}(t)}$. As E is non-decreasing, we get that

$$\max_{[0,t]} \tilde{H} \leq \sqrt{E_t} \sqrt{\max_{[0,t]} \dot{H}}.$$

Finally, let $T > 0$ a, b be such that $[a, b] \subset [0, T]$,

$$\sup_{s,t \in [a,b]} |\tilde{H}(t) - \tilde{H}(s)| \leq 2\sqrt{E_T} \sqrt{\sup_{s,t \in [a,b]} |\dot{H}(t) - \dot{H}(s)|}$$

so if

$$\omega_f(\delta) = \begin{cases} \inf & \max \sup |f(t) - f(s)| \\ 0 = t_1 < \dots < t_k = T, & 2 \leq i \leq k, t, s \in [t_{i-1}, t_i] \\ k \in \mathbb{N} & \\ \min_{2 \leq i \leq k} |t_i - t_{i-1}| \geq \delta & \end{cases}$$

is the càdlàg modulus of f on $[0, T]$,

$$\omega_{\tilde{H}}(\delta) \leq 2\sqrt{E_T} \sqrt{\omega_{\dot{H}}(\delta)}$$

□

Of course Lemma 10.3.5 can only be useful if we have the mean to control E :

Lemma 10.3.6. *Let $t \in \mathbb{R}_+$,*

$$\mathbb{E}(E_t) = t \sum_{i=1}^n \frac{x_i^2 v_i}{\sigma_2^2}$$

Proof. We have that for any $k \in \mathbb{N}$,

$$\begin{aligned} \mathbb{E}\left(\sum_{i=1}^{N_k} (D_k^i)^2\right) &= \mathbb{E}\left(\mathbb{E}\left(\sum_{i=1}^{N_k} (D_k^i)^2 \mid N_k, J_k\right)\right) \\ &= \mathbb{E}\left(\sum_{i=1}^{N_k} v_{J_k}\right) \\ &= \mathbb{E}(N_k v_{J_k}) \\ &= \sum_{j=1}^n \mathbb{P}(J_k = j) v_j \mathbb{E}(N_k \mid J_k = j) \\ &= \sum_{j=1}^n x_j v_j \mathbb{P}(J_k = j) \\ &= \sum_{j=1}^n \frac{x_j^2}{\sigma_1} v_j. \end{aligned}$$

As a consequence, as $\sum_{i=1}^{N_k} (D_k^i)^2$ is independent from τ_k ,

$$\begin{aligned}
\mathbb{E}(\sigma_2 E_t) &= \sum_{k \in \mathbb{N}} \mathbb{E} \left(\sum_{i=1}^{N_k} (D_k^i)^2 1_{\tau_k \leq t} \right) \\
&= \sum_{k \in \mathbb{N}} \mathbb{E} \left(\mathbb{E} \left(\sum_{i=1}^{N_k} (D_k^i)^2 \right) 1_{\tau_k \leq t} \right) \\
&= \sum_{j=1}^n \frac{x_j^2}{\sigma_1} v_j \mathbb{E}(\#\{k : \tau_k \leq t\}) \\
&= t \sum_{j=1}^n \frac{x_j^2 v_j}{\sigma_2}.
\end{aligned}$$

□

10.4 Coupling with an eternal spanning forest

One strong hindsight from the proof of Theorem 8.2.2, is that the components that have positive complexity, which means they are not trees, despite being infinite in number at every time, do not contribute to the relevant asymptotic at the entrance boundary. In fact, in the proof of Theorem 8.2.2, those components are unfolded and become trees, and we show that this unfolding does not have any impact on the asymptotic. This fact is related to the fact that this infinite number of components with positive complexity, have a finite total mass, which is way smaller than the total infinite mass of all the space. This suggests that any unfolding of complex components into trees might do the trick. Here we present an unfolding that is easier to handle in regard to the evolution of the coalescence. Our goal is to be able to apply martingale methods to the unfolded process to obtain the entrance boundary asymptotics. Note that this unfolding is very similar to the one in [3], but used in a very different context, whereas here we want an unfolded process very similar to the original process in $-\infty$, the authors of [3] show that the unfolded process poses a limit in $+\infty$ that is very different, and it is the basis of their investigation of the scaling limit of the minimal spanning tree of the complete graph.

Definition 10.4.1 (Tree coalescence). *Let $X = (X, d, \mu) \in \mathbb{M}_2^{\mathbb{N}}$, let \mathcal{P} be a Poisson point process of intensity $\frac{1}{2}\mu^2 \times \text{Leb}_{\mathbb{R}_+}$ on $X^2 \times [0, +\infty)$ and for all $t \geq 0$, let*

$$\mathcal{TP}_t = \{(u, v) \in X^2 : \exists r \leq t, (u, v, r) \in \mathcal{P} \text{ and } \forall r' < r, d_{r'}^{\text{Coal}}(u, v) = +\infty\}.$$

For all $t > 0$, we define the semi metric d_t^{TCoal} on X as:

$$\forall u, v \in X, d_t^{\text{TCoal}}(u, v) = \inf \left\{ \sum_{i=1}^k d(u_i, v_i) : u_1 = u, v_k = v, \forall i \leq k-1, (v_i, u_{i+1}) \in \mathcal{TP}_t \right\}.$$

We define

$$\text{TCoal}(X, t) = \text{Glue}(X, d_t^{\text{TCoal}}).$$

This tree coalescence is such that, when fed with a sequence of real trees $(X_j)_{j \in \mathbb{N}}$, $\text{TCoal}(X, q)$ stays a sequence of real trees. If $T = (T(t))_{t \in \mathbb{R}}$ is a Markov process in $\mathbb{M}_2^{\mathbb{N}}$ such that for all $s, t \in \mathbb{R}$,

$$(\text{TCoal}(T(t), s), T(t)) \stackrel{(d)}{=} (T(t+s), T(t)),$$

we say that it is an eternal spanning forest. Remark that $(T(t))_{t \in \mathbb{R}}$ is not necessarily a forest.

We can now present our unfolding, which is nothing more than the obvious coupling with the tree coalescence, but started from $-\infty$.

Recall that there is a natural coupling between $\text{Coal}(\cdot, \cdot)$ and $\text{TCoal}(\cdot, \cdot)$ in which we simply forget gluing points of the coalescence that increase the surplus when performing $\text{TCoal}(\cdot, \cdot)$.

Theorem 10.4.2. *Let X be an extreme eternal metric multiplicative coalescent in $\mathbb{M}_2^{\mathbb{N}}$. There exists T , an eternal spanning forest such that for all $t \in \mathbb{R}$,*

$$\text{Coal}(T(t-s), s) \xrightarrow[s \rightarrow +\infty]{(d)} X(t)$$

for T_{GHPS} .

Proof. For all $t \in \mathbb{R}$ and $s \geq 0$, let $T^s(t) = \text{TCoal}(X(t-s), s)$. Remark that, thanks to Lemma 4.2.7 the components of $T^s(t)$ and $X(t)$ are the same pointwise. Let $(\mathcal{V}_j(t))_{j \in \mathbb{N}}$ be the sets of points of these components. For $j \in \mathbb{N}$, $t \in \mathbb{R}$, and any time interval $I \subset (-\infty, t]$, let $\mathcal{S}_j^t(I)$ be the number of extra points added on $\mathcal{V}_j(t)$ by the coalescence during I that are not added by the tree-coalescence. Remark that if I and J are such that $I \cap J = \emptyset$, $\mathcal{S}_j^t(I \cup J) = \mathcal{S}_j^t(I) + \mathcal{S}_j^t(J)$. Now consider the following fact:

- (i) For all $s, t \in \mathbb{R}$, $s > 0$, if $d_{T^s(t)} \neq d_{X(t)}$ on $\mathcal{V}_j(t)$, by definition of $\text{TCoal}(\cdot, \cdot)$, this means that $\mathcal{S}_j^t((t-s, t]) > 0$.

From this we can generalise a bit:

- (ii) For all $r, s, t \in \mathbb{R}$, $r > s > 0$, if $d_{T^{r-s}(t-s)} = d_{T^0(t-s)} = d_{X(t-s)}$ on $\mathcal{V}_j(t)$, because they start from the same space, this shows that $d_{T^s(t)} = d_{T^r(t)}$ on $\mathcal{V}_j(t)$. As a consequence, if $d_{T^s(t)} \neq d_{T^r(t)}$, this implies that $d_{T^{r-s}(t-s)} \neq d_{X(t-s)}$ on \mathcal{V}_j , and thus by (i), this means that $\mathcal{S}_j^t((t-r, t-s]) > 0$.

From (ii) we get that for $t \in \mathbb{R}$, $j \in \mathbb{N}$,

$$\forall s > 0, \exists r > s \text{ such that } d_{T^s(t)} \neq d_{T^r(t)} \text{ on } \mathcal{V}_j(t),$$

implies that $\mathcal{S}_j^t((-\infty, t]) = +\infty$. However, $\mathcal{S}_j^t((s, t])$ corresponds to the surplus of the j -th component of $\text{gCoal}(x(t-s), s)$. As X is extreme, thanks to Theorem 2.2.10, for all $(s_n)_{n \in \mathbb{N}}$ such that $s_n \xrightarrow[n \rightarrow \infty]{} \infty$ $((x(s_n))_{n \in \mathbb{N}}, (s_n)_{n \in \mathbb{N}})$ satisfies the Aldous-Limic conditions Hypothesis 1 almost surely. As a consequence, from Theorem 2.4.9, the surplus of the j -th component of $\text{gCoal}(x(t-s_n), s_n)$ is tight. In particular, for all $\epsilon > 0$, there exists $K > 0$ such that

$$\mathbb{P}(\{\mathcal{S}_j^t((-\infty, t]) = +\infty\}) \leq \lim_{n \rightarrow +\infty} \mathbb{P}(\{\mathcal{S}_j^t((t-s_n, t]) \geq K\}) \leq \epsilon$$

so $\mathbb{P}(\{\mathcal{S}_j^t((-\infty, t]) = +\infty\}) = 0$. As a consequence, with probability one there exists s_j such that

$$\forall r > s_j, d_{T^{s_j}(t)} = d_{T^r(t)} \text{ on } \mathcal{V}_j.$$

Let $T(t) = (\mathcal{V}_j(t), d_{T^{s_j}(t)})_{j \in \mathbb{N}}$.

By construction, we have that $T(t) = \text{TCoal}(T(t-s), s)$. Moreover, we have that for any $j \in \mathbb{N}$,

$$d_{T(t-s_j)} = d_{T^0(t-s_j)} = d_{X(t-s_j)} \text{ on } \mathcal{V}_j.$$

□

Finally, let us show how one might use this coupling. One argument at the heart of Theorem 2.2.10 is an \mathcal{L}_2 convergence theorem for (reversed) martingales:

Lemma 10.4.3 (Lemma 25 of [6]). *Let $(Y(t))_{t < 0}$ be a process adapted to $(\mathcal{F}(t))_{t < 0}$. For $t < 0$, let $A(t)$ be a càdlàg adapted process such that for all $s < t < 0$ $M(t) = Y(t) - Y(s) - \int_s^t A(u)du$ is a local martingale. Let $B(t)$ be a càdlàg adapted process such that for all $s < t < 0$ $M^2(t) - M^2(s) - \int_s^t B(u)du$ is a local martingale. Suppose that:*

$$\int_{-\infty}^0 |A(t)| dt < \infty$$

and

$$\int_{-\infty}^0 |B(t)| dt < \infty.$$

Then $\lim_{t \rightarrow -\infty} Y(t)$ exists and is finite a.s.

However, contrary to the case of masses, dealing with metric statistics makes martingales methods harder. In fact while expectation of the average distance of the merging of two distinct components is easy to understand, the consequence of coalescing two points inside the same component are far less predictable. This is where $\text{TCoal}(\cdot, \cdot)$ simplifies things so much: there is no such event. In order to illustrate this, we give the following lemma where we compute the compensator of ρ_1 , in a particular case of tree-coalescence.

Lemma 10.4.4. Let $T \in \overline{\mathbb{M}}^{\mathbb{N}}$ such that T has a finite number of components and for all $t \geq 0$, $\mathcal{T}(t) = T\text{Coal}(T, t)$. For all $j \in \mathbb{N}^*$, let $x_j = (x_j(t))_{t \in \mathbb{R}} = (|\mathcal{T}_j(t)|)_{t \in \mathbb{R}}$, $u_j = (u_j(t))_{t \in \mathbb{R}}$ such that

$$u_j(t) = x_j(t)^{-2} \int_{\mathcal{T}_j(t)^2} d_j^t(u, v) \mu_j(du) \mu_j(dv).$$

For $t \in \mathbb{R}$, let \mathcal{F}_t be the sigma algebra generated by $(T(s))_{s \in [0, t]}$. Let

$$\rho_1(t) = \sum_j x_j^2(t) u_j(t),$$

$$\sigma_2(t) = \sum_j x_j^2(t)$$

and

$$\nu(t) = \sum_j x_j^4(t) u_j(t)$$

then $\rho_1(t) - \int_0^t 2(\sigma_2(s)\rho_1(s) - \nu(s))ds$ is a martingale, in other words, $(2(\sigma_2(t)\rho_1(t) - \nu(t)))_{t \geq 0}$ is the compensator of ρ_1 .

Proof. Let f be a function from $\mathbb{M}_2^{\mathbb{N}}$ to \mathbb{R} . For $X \in \overline{\mathbb{M}}^{\mathbb{N}}$, let

$$L_f(X) = \frac{1}{2} \sum_{i \neq j} |X_i| |X_j| \mathbb{E}(f(X^{i+j}) - f(X))$$

where X^{i+j} is the space where we have coalesced two independent points U and V , with $U \sim \frac{\mu_i(\cdot)}{|X_i|}$ and $V \sim \frac{\mu_j(\cdot)}{|X_j|}$, and the expectation is taken relatively to U and V . In the same spirit than of Section 4 of [5], it is not hard to prove that

$$f(T(t)) - \int_0^t L_f(T(s)) ds$$

is a martingale. Let $f(X) = \sum_j \int_{X_j \times X_j} d(u, v) \mu_j(du) \mu_j(dv)$. Then

$$f(X^{i+j}) - f(X) = 2 \int_{X_i \times X_j} (d_i(u, U) + d_j(V, v)) \mu_i(du) \mu_j(dv)$$

so

$$\begin{aligned} \mathbb{E}(f(X^{i+j}) - f(X)) &= 2 \int_{X_i \times X_i \times X_j \times X_j} (d_i(u, u') + d_j(v', v)) \mu_i(du) \frac{\mu_i(du')}{|X_i|} \frac{\mu_j(dv')}{|X_j|} \mu_j(dv) \\ &= \frac{4}{|X_i| |X_j|} (|X_j|^2 \int_{X_i \times X_i} d_i(u, u') \mu_i(du) \mu_i(du') + |X_i|^2 \int_{X_j \times X_j} d_j(v', v) \mu_j(dv') \mu_j(dv)). \end{aligned}$$

As a consequence,

$$\begin{aligned} L_f(X) &= 2 \sum_{i \neq j} |X_j|^2 \int_{X_i \times X_i} d_i(u, u') \mu_i(du) \mu_i(du') \\ &= 2 \left(\sum_i |X_i|^2 \sum_j x_j^2 u_j - \sum_i x_i^4 u_i \right) \end{aligned}$$

with of course $u_i = |X_i|^{-2} \int_{X_i \times X_i} d_i(u, u') \mu_i(du) \mu_i(du')$. This proves the lemma because

$$L_f(T(t)) = 2(\sigma_2(t)\rho_1(t) - \nu(t)).$$

□

Appendix A

Lower mass bound and randomness

Here is the proof of a result that is very natural and regularly used in the literature but that we are not sure it is rigorously proven in this particular setting.

Proof of Lemma 2.3.7. We only prove (i) \Rightarrow (ii), since this is the non trivial part, and the one we shall use. Recall from Theorem 2.3.5 the definition of the Gromov-Prokhorov distance d_{GP} . Remark that d_{GHP} is equivalent to $\widetilde{d_{GHP}}$ given by

$$\widetilde{d_{GHP}}(X, Y) = \inf_{(\phi_X, \phi_Y, Z)} \inf \left\{ \epsilon' > 0 : \begin{cases} \mu_X(\phi_X^{-1}(F)) \leq \mu_Y(\phi_Y^{-1}(F^{\epsilon'})) + \epsilon', \forall F \text{ closed} \\ \phi_X(X) \subset \phi_Y(Y)^{\epsilon'} \text{ and } \phi_Y(Y) \subset \phi_X(X)^{\epsilon'} \end{cases} \right\}.$$

Suppose that

$$X^n \xrightarrow[n \rightarrow \infty]{(d)} X$$

with respect to d_{GP} . For $\delta > 0$, $\epsilon > 0$, let

$$E_{\delta, \epsilon} = \{Y \in \mathbb{M} : m_\delta(Y) \geq \epsilon\}.$$

Let F be closed under d_{GHP} . For all $\delta > 0$, $\epsilon > 0$,

$$\mathbb{P}(X^n \in F) = \mathbb{P}(X^n \in F \cap E_{\delta, \epsilon}) + \mathbb{P}(X^n \in F \cap \neg E_{\delta, \epsilon})$$

so for n large enough,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{P}(X^n \in F) &\leq \limsup_{n \rightarrow \infty} \mathbb{P}(X^n \in F \cap E_{\delta, \epsilon}) + \limsup_{n \rightarrow \infty} \mathbb{P}(X^n \notin E_{\delta, \epsilon}) \\ &\leq \limsup_{n \rightarrow \infty} \mathbb{P}(X \in \overline{F \cap E_{\delta, \epsilon}}^{GP}) + \limsup_{n \rightarrow \infty} \mathbb{P}(X^n \notin E_{\delta, \epsilon}) \end{aligned}$$

where $\overline{F \cap E_{\delta, \epsilon}}^{GP}$ is the closure of $F \cap E_{\delta, \epsilon}$ under Gromov-weak. Let us show that if $Y \in \overline{F \cap E_{\delta, \epsilon}}^{GP}$, it is at distance at most 2δ for $\widetilde{d_{GHP}}$, of an element of F or $m_\delta(Y) \geq \epsilon$. First, there exists $(Y^n)_{n \in \mathbb{N}}$ in $F \cap E_{\delta, \epsilon}$ such that $Y^n \xrightarrow[n \rightarrow \infty]{d_{GP}} Y$. For n large enough,

$$d_{GP}(Y^n, Y) < \delta \wedge \epsilon$$

and as $Y^n \in E_{\delta, \epsilon}$, $m_\delta(Y^n) \geq \epsilon$. Let (ϕ_{Y^n}, ϕ_Y, Z) be such that $\phi_{Y^n} : Y^n \rightarrow Z$ and $\phi_Y : Y \rightarrow Z$ are isometric embeddings that realize the infimum in $\widetilde{d_{GHP}}(Y^n, Y)$. Suppose that $\widetilde{d_{GHP}}(Y^n, Y) > \delta$. This means that either $\phi_Y(Y) \not\subset \phi_{Y^n}(Y^n)^{2\delta}$, either $\phi_{Y^n}(Y^n) \not\subset \phi_Y(Y)^{2\delta}$. Suppose that $\phi_{Y^n}(Y^n) \not\subset \phi_Y(Y)^{2\delta}$. There exists $u \in \phi_{Y^n}(Y^n)$ such that $\overline{B(u, \delta)} \cap \phi_Y(Y) = \emptyset$, but as $d_{GP}(Y^n, Y) < \delta \wedge \epsilon$, this implies that $\mu_{Y^n}(\phi_{Y^n}^{-1}(\overline{B(u, \delta)})) \leq \epsilon$ which contradicts the fact that $m_\delta(Y^n) \geq \epsilon$. With the same reasoning if $\phi_Y(Y) \not\subset \phi_{Y^n}(Y^n)^{2\delta}$, $m_\delta(Y) \geq \epsilon$. Thus, either

$$\widetilde{d_{GHP}}(Y^n, Y) \leq 2\delta \text{ or } m_\delta(Y) \geq \epsilon.$$

All in all,

$$\limsup_{n \rightarrow \infty} \mathbb{P}(X^n \in F) \leq \mathbb{P}(\exists Y \in F : \widetilde{d_{GHP}}(X, Y) \leq 2\delta) + \limsup_{n \rightarrow \infty} \mathbb{P}(X^n \notin E_{\delta, \epsilon}) + \mathbb{P}(m_\delta(X) \geq \epsilon).$$

We suppose that for $\delta > 0$ fixed, $\limsup_{n \rightarrow \infty} \mathbb{P}(X^n \in E_{\delta, \epsilon}) \xrightarrow{\epsilon \rightarrow 0} 0$, which is the lower mass bound property in probability. Also, $\mathbb{P}(m_\delta(X) \geq \epsilon) \xrightarrow{\epsilon \rightarrow 0} 1$ since X is a.s. compact with full support. Then

$$\limsup_{n \rightarrow \infty} \mathbb{P}(X^n \in F) \leq \mathbb{P}(\exists Y \in F : \widetilde{d_{GHP}}(X, Y) \leq \delta).$$

Finally, $\mathbb{P}(\exists Y \in F : \widetilde{d_{GHP}}(X, Y) \leq \delta) \xrightarrow{\delta \rightarrow 0} \mathbb{P}(X \in F)$ since F is closed under $\widetilde{d_{GHP}}$. □

Appendix B

Feller property for metric coalescence

Proof of Theorem 2.5.5. The main difference between Theorem 2.5.5 and Theorem 4.1 in [28] is the topology. To fulfil the hypothesis of Theorem 4.1, the convergence

$$X^n \xrightarrow[n \rightarrow +\infty]{(d)} X^\infty,$$

should be in regard to L_{GHP}^2 , which is defined in [28]. Thanks to Proposition 4.6 in [28], the convergence for L_{GHP}^2 is a consequence of the convergence for T_{GHP}^2 . Then, the proof of Theorem 3.1 shows the convergence

$$(\text{Coal}(X^n, t_i))_{i \in [k]} \xrightarrow[n \rightarrow \infty]{(d)} (\text{Coal}(X^\infty, t_i))_{i \in [k]}$$

with respect to the L_{GP}^2 topology, which is L_{GHP}^2 with d_{GHP} replaced by d_{GP} defined in Theorem 2.3.5. This convergence implies the convergence for the Gromov-weak product topology under the condition

$$\forall t \in \mathbb{R}_+, \mathbb{P}(\exists i, j \in \mathbb{N}, i \neq j \text{ s.t. } |\text{Coal}(X^\infty, t)_i| = |\text{Coal}(X^\infty, t)_j|) = 0.$$

We can now use the lower-mass bound property

$$\forall t \in \mathbb{R}_+, \forall \delta > 0, \limsup_n \mathbb{P}(m_\delta^k(\mathcal{X}^n(t)) < \epsilon) \xrightarrow{\epsilon \rightarrow 0} 0$$

to strengthen this convergence in convergence in a convergence for T_{GHP} . This gives that for $t \in \mathbb{R}_+$, $\text{Coal}(X^\infty, t)$ is a.s. in $\mathbb{M}_2^{\mathbb{N}}$ and for any family $(t_0, \dots, t_k) \in \mathbb{R}_+^{k+1}$,

$$(\text{Coal}(X^n, t_i))_{i \in [k]} \xrightarrow[n \rightarrow \infty]{(d)} (\text{Coal}(X^\infty, t_i))_{i \in [k]}$$

for T_{GHP}^2 . □

We show why Theorem 2.5.5 implies the existence of $\mathcal{E}(\alpha, \lambda, c, \gamma)$, an eternal multiplicative coalescent with marginals $(\mathcal{E}_t(\alpha, \lambda, c, \gamma))$, for $(\alpha, \lambda, c, \gamma) \in \mathcal{I} \times \mathbb{R}_+$.

Let $x = (x^n)_{n \in \mathbb{N}}$, $q = (q_n)_{n \in \mathbb{N}}$ such that (x, q) satisfies the Aldous-Limic conditions Hypothesis 1 with parameters $(\alpha, \lambda, c, \gamma) \in \mathcal{I} \times \mathbb{R}_+$. We want to apply Theorem 2.5.5 with, for $n \in \mathbb{N}$, $X^n = \text{SCoal}(x^n, q_n)$, which is such that

$$X^n \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{E}_0(\alpha, \lambda, c, \gamma)$$

according to Theorem 2.4.9. The main point to verify in order to apply Theorem 2.5.5 and redo the discussion following Theorem 2.2.5, since

$$\forall t \in \mathbb{R}_+, \mathbb{P}(\exists i, j \in \mathbb{N}, i \neq j \text{ s.t. } |\text{Coal}(X^\infty, t)_i| = |\text{Coal}(X^\infty, t)_j|) = 0$$

is clearly satisfied, is to show that for $k \in \mathbb{N}$,

$$\forall t \in \mathbb{R}, \forall \delta > 0, \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(m_\delta^k(\mathcal{X}^n(t)) < \epsilon) = 0.$$

Here, $\mathcal{X}^n(t) = \text{Coal}(\text{SCoal}(x^n, q_n), t)$. Remark that, even if $\text{Coal}(\text{SCoal}(x^n, q_n), t) \neq \text{SCoal}(x^n, q_n + t)$,

$$m_\delta^k(\mathcal{X}^n(t)) \geq m_\delta^k(\text{SCoal}(x^n, q_n + t))$$

because with the obvious coupling, the fact that we use the complete coalescence between q_n and $q_n + t$ can only decrease the distances, and thus augment the mass of small balls. As a consequence, for $\epsilon > 0$,

$$\limsup_{n \rightarrow \infty} \mathbb{P}(m_\delta^k(\mathcal{X}^n(t)) < \epsilon) \leq \limsup_{n \rightarrow \infty} \mathbb{P}(m_\delta^k(\text{SCoal}(x^n, q_n + t)) < \epsilon).$$

It is then easy to see that $(x, q + t)$ verifies the Aldous-Limic conditions Hypothesis 1 so thanks to Theorem 2.4.9,

$$m_\delta^k(\text{SCoal}(x^n, q_n + t)) \xrightarrow[n \rightarrow \infty]{(d)} m_\delta^k(X^\infty(t))$$

where $X^\infty(t) \sim \mathcal{E}_t(\alpha, \lambda, c, \gamma)$, so

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(m_\delta^k(\mathcal{X}^n(t)) < \epsilon) \leq \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(m_\delta^k(\text{SCoal}(x^n, q_n + t)) < \epsilon) = 0.$$

List of notation and terminology

$(\Delta_j)_{j \in \mathbb{N}}$	The sequence of $\{i : \zeta_i \in (l_j, r_j]\}$, for $j \in \mathbb{N}$	44
$(\delta_r^{(j)})_{j \in \mathbb{N}}$	The sequence of $\sum_{i \in \Delta_j} c_i^r$, for $j \in \mathbb{N}$	44
$(c^{(j)})_{j \in \mathbb{N}}$	The sequence such that $c_i^{(j)} = c_i$ if $i \in \Delta_j$, 0 otherwise, for $j \in \mathbb{N}$	44
D	For $X \in \mathbb{M}_2^{\mathbb{N}}$, $D = (D_i)_{i \in \mathbb{N}}$ is the sequence of diameters of the components $(X_i)_{i \in \mathbb{N}}$	48
T_{GHP}^2	The topology associated with d_{GHP}^2	40
$[n]$	The discrete set $\{0, 1, \dots, n\}$	33
$\text{Coal}(X, q)$	Metric (complete) coalescence of X , during a time q	45
$\text{Glue}(X, d)$	Gluing of X along a semi-metric d	45
H	Height process of the client-server construction.....	43
$\text{Perc}(X, q)$	Dynamical percolation on X , during a time q	51
$\text{SCoal}(X, q)$	Simple coalescence of X , during a time q	45
$\text{SCoal}_a(X, q)$	Simple coalescence of X , during a time q , with edge length a	45
$\text{SCoal}_a^V(X, q)$	Coalescence of X , restricted to the vertex set V , conditioned to be connected.....	49
$\text{TCoal}(X, q)$	Tree coalescence of X , during a time q	115
$\text{Thin}(X, q)$	Thin coalescence of X , during a time q	112
$\tilde{\mathbb{M}}$	Space of complete, separable, with finite measure, full support, metric spaces.....	40
Z	Algebraic load of the server.....	43
\mathcal{Z}	Algebraic load of the continuous process.....	36
\mathcal{H}	Height process of the continuous process.....	43
\mathcal{B}	Load of the continuous process.....	36
$\text{comp}(X)$	Set of components of the (semi-metric) space X	40
\mathring{H}	Height process of the completed process.....	111
\tilde{H}	Modified height process of the completed process, where the increments are affected by the distances on the components.....	111
\mathring{Z}	Completed algebraic load of the server.....	111
\mathring{B}	Completed load of the server.....	111
$\text{core}(X)$	Core of the real graph X	42
$E(x, q)$	Set of edges of $\text{gCoal}(x, q)$	35
$\text{gCoal}(x, q)$	Inhomogeneous random graph, or graph of Aldous' multiplicative coalescent started from x , performed during a time q	35
$\mathring{\alpha}_k(t)$	Is equal to one if the client served at time t is a descendant of the client served at time τ_k for the completed process.....	43
$\alpha_i(t)$	Is equal to one if the client served at time t is a descendant of i	43
$m_\delta^l(X)$	The δ -lower mass of the l largest components of $X \in \mathbb{M}_2^{\mathbb{N}}$, $\inf_{i \in [l]} m_\delta(X_i)$	41
$m_\delta(X)$	The δ -lower mass of $X \in \mathbb{M}$, $\inf\{\mu(\{v : d(u, v) \leq \delta\}) : u \in X\}$	41
B	Load of the server.....	43
$\text{mCoal}(x, q)$	Masses of Aldous' coalescent started from x , performed during a time q	35
$\mathcal{M}_t(\alpha, \lambda, c)$	Distribution of the mass of extreme eternal Aldous's multiplicative coalescent at time t	38
\mathcal{I}	Set of admissible parameters for the Aldous-Limic condition Hypothesis 1.....	36
\mathcal{I}°	Subset of \mathcal{I} that generate purely homogeneous spaces.....	36
\mathcal{I}^*	Subset of \mathcal{I} that generate purely inhomogeneous spaces.....	36
ρ_r	For $X \in \mathbb{M}_2^{\mathbb{N}}$, $\rho_r = \sum_i \int_{X_i \times X_i} d_i(u, v)^r \mu_i(du) \mu_i(dv)$	48
\mathbb{F}	Set of the polynomials, used in the definition of the Gromov-weak topology.....	41
\overline{l}_\searrow	Space of non-increasing and non-negative sequence x such that $x_i = 0$ for i large enough.....	36
\overline{sl}	Space of sequence of \overline{l}_\searrow	39
sl_2	Space of sequence of l_\searrow^2	36

$\sigma_{p,r}$	For x, y two sequences of \mathbb{R}_+ , $\sigma_{p,r}(x, y) = \sum_{i \in \mathbb{N}} x_i^p y_i^r$	57
σ_r	For $x \in l_{\searrow}^r$, $\sigma_r(x) = \sum_{i \in \mathbb{N}} x_i^r$	35
$\mathcal{E}_t(\alpha, \lambda, c, \gamma)$	Distribution of limiting continuous multiplicative coalescent for the parameters $(\alpha, \lambda, c, \gamma)$, at time t	44
$S(x, q)$	l^2 -norm of $\text{mCoal}(x, q)$	35
s_j	Surplus of the j -th component	42
$\text{scl}(\alpha, \beta)$	Scaling operator, the distance is multiplied by α and the measure by β	40
d_{GHP}	Gromov-Hausdorff-Prokhorov distance	40
u	For $X \in \mathbb{M}_2^{\mathbb{N}}$, $u = (u_i)_{i \in \mathbb{N}}$ is the sequence of average distances of the components $(X_i)_{i \in \mathbb{N}}$	48
\mathbb{M}	Space of compact measured metric spaces	40
l_{\searrow}^r	Space of non-increasing and non-negative sequences x such that $\sum_i x_i^r < \infty$	35

Bibliography

- [1] L. Addario-Berry, N. Broutin, and C. Goldschmidt. “The continuum limit of critical random graphs”. In: *Probability Theory and Related Fields* 152.3 (Apr. 2012), pp. 367–406. ISSN: 1432-2064. DOI: 10.1007/s00440-010-0325-4. URL: <https://doi.org/10.1007/s00440-010-0325-4>.
- [2] L. Addario-Berry, N. Broutin, and C. Goldschmidt. “Critical Random Graphs: Limiting Constructions and Distributional Properties”. In: *Electronic Journal of Probability* 15.none (2010), pp. 741–775. DOI: 10.1214/EJP.v15-772. URL: <https://doi.org/10.1214/EJP.v15-772>.
- [3] L. Addario-Berry, N. Broutin, C. Goldschmidt, and G. Miermont. “The scaling limit of the minimum spanning tree of the complete graph”. In: *The Annals of Probability* 45.5 (2017), pp. 3075–3144. DOI: 10.1214/16-AOP1132. URL: <https://doi.org/10.1214/16-AOP1132>.
- [4] D. Aldous. “The Continuum Random Tree III”. In: *The Annals of Probability* 21.1 (1993), pp. 248–289. DOI: 10.1214/aop/1176989404. URL: <https://doi.org/10.1214/aop/1176989404>.
- [5] D. Aldous. “Brownian excursions, critical random graphs and the multiplicative coalescent”. In: *Ann. Probab.* 25.2 (Apr. 1997), pp. 812–854. DOI: 10.1214/aop/1024404421. URL: <https://doi.org/10.1214/aop/1024404421>.
- [6] D. Aldous and V. Limic. “The Entrance Boundary of the Multiplicative Coalescent”. In: *Electron. J. Probab.* 3 (1998), 59 pp. DOI: 10.1214/EJP.v3-25. URL: <https://doi.org/10.1214/EJP.v3-25>.
- [7] S. Athreya, W. Loehr, and A. Winter. “The gap between Gromov-vague and Gromov-Hausdorff-vague topology”. In: *arXiv:1407.6309* (July 2014). DOI: 10.1016/j.spa.2016.02.009.
- [8] I. Benjamini, G. Kalai, and O. Schramm. “Noise sensitivity of Boolean functions and applications to percolation”. In: *Publications Mathématiques de l’Institut des Hautes Études Scientifiques* 90.1 (Dec. 1999), pp. 5–43. ISSN: 1618-1913. DOI: 10.1007/BF02698830. URL: <https://doi.org/10.1007/BF02698830>.
- [9] S. Bhamidi, N. Broutin, S. Sen, and X. Wang. “Scaling limits of random graph models at criticality: Universality and the basin of attraction of the Erdős-Rényi random graph”. In: (2014). arXiv: 1411.3417 [math.PR].
- [10] S. Bhamidi, S. Dhara, R. van der Hofstad, and S. Sen. “Universality for critical heavy-tailed network models: Metric structure of maximal components”. In: *Electron. J. Probab.* 25 (2020), 57 pp. DOI: 10.1214/19-EJP408. URL: <https://doi.org/10.1214/19-EJP408>.
- [11] S. Bhamidi, R. van der Hofstad, and S. Sen. “The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs”. In: *Probability Theory and Related Fields* 170.1 (Feb. 2018), pp. 387–474. ISSN: 1432-2064. DOI: 10.1007/s00440-017-0760-6. URL: <https://doi.org/10.1007/s00440-017-0760-6>.
- [12] S. Bhamidi, S. Sen, and X. Wang. “Continuum limit of critical inhomogeneous random graphs”. English. In: *Probab. Theory Relat. Fields* 169.1-2 (2017), pp. 565–641. ISSN: 0178-8051. DOI: 10.1007/s00440-016-0737-x.
- [13] B. Bollobás. *Random Graphs*. 2nd ed. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2001. DOI: 10.1017/CB09780511814068.
- [14] N. Broutin, T. Duquesne, and M. Wang. “Limits of multiplicative inhomogeneous random graphs and Lévy trees: Limit theorems”. In: (2020). arXiv: 2002.02769 [math.PR].
- [15] N. Broutin, T. Duquesne, and M. Wang. “Limits of multiplicative inhomogeneous random graphs and Lévy trees: The continuum graphs”. In: (2020). arXiv: 1804.05871 [math.PR].

- [16] N. Broutin and J.-F. Marckert. “A new encoding of coalescent processes. Applications to the additive and multiplicative cases”. In: *Probability Theory and Related Fields* 166.1 (2016), pp. 515–552. DOI: 10.1007/s00440-015-0665-1. URL: <https://hal.inria.fr/hal-01092562>.
- [17] S. Dhara, R. van der Hofstad, J. S. H. van Leeuwen, and S. Sen. “Critical window for the configuration model: finite third moment degrees”. English. In: *Electron. J. Probab.* 22 (2017). Id/No 16, p. 33. ISSN: 1083-6489. DOI: 10.1214/17-EJP29.
- [18] E. B. Dynkin. “Sufficient Statistics and Extreme Points”. In: *The Annals of Probability* 6.5 (1978), pp. 705–730. DOI: 10.1214/aop/1176995424. URL: <https://doi.org/10.1214/aop/1176995424>.
- [19] P. Erdos and A. Renyi. “On the evolution of random graphs”. In: *Publ. Math. Inst. Hungary. Acad. Sci.* 5 (1960), pp. 17–61.
- [20] J.-F. L. Gall. “Random trees and applications”. In: *Probability Surveys* 2.none (2005), pp. 245–311. DOI: 10.1214/154957805100000140. URL: <https://doi.org/10.1214/154957805100000140>.
- [21] A. Greven, P. Pfaffelhuber, and A. Winter. “Convergence in distribution of random metric measure spaces (Λ -coalescent measure trees)”. English. In: *Probab. Theory Relat. Fields* 145.1-2 (2009), pp. 285–322. ISSN: 0178-8051. DOI: 10.1007/s00440-008-0169-3.
- [22] S. Janson and J. Spencer. “A Point Process Describing the Component Sizes in the Critical Window of the Random Graph Evolution”. In: *Combinatorics, Probability and Computing* 16.4 (2007), pp. 631–658. DOI: 10.1017/S0963548306008327.
- [23] E. Lubetzky and Y. Peled. “Noise sensitivity of critical random graphs”. In: (2020). arXiv: 2009.01707 [math.PR].
- [24] G. Miermont. “Self-similar fragmentations derived from the stable tree II: splitting at nodes”. In: *Probability Theory and Related Fields* 131.3 (Mar. 2005), pp. 341–375. ISSN: 1432-2064. DOI: 10.1007/s00440-004-0373-8. URL: <https://doi.org/10.1007/s00440-004-0373-8>.
- [25] B. Mohar and C. Thomassen. *Graphs on surfaces*. English. Baltimore, MD: Johns Hopkins University Press, 2001, pp. xii + 291. ISBN: 0-8018-6689-8.
- [26] B. Ráth. “Feller property of the multiplicative coalescent with linear deletion”. English. In: *Bernoulli* 25.1 (2019), pp. 221–240. ISSN: 1350-7265. DOI: 10.3150/17-BEJ984.
- [27] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 1*. Cambridge Mathematical Library. Foundations, Reprint of the second (1994) edition. Cambridge: Cambridge University Press, 2000, pp. xx+386. ISBN: 0-521-77594-9.
- [28] R. Rossignol. “Scaling limit of dynamical percolation on critical Erdős–Rényi random graphs”. In: *Ann. Probab.* 49.1 (Jan. 2021), pp. 322–399. DOI: 10.1214/20-AOP1472. URL: <https://doi.org/10.1214/20-AOP1472>.
- [29] R. Rossignol. “Noise-stability and central limit theorems for effective resistance of random electric networks”. English. In: *Ann. Probab.* 44.2 (2016), pp. 1053–1106. ISSN: 0091-1798. DOI: 10.1214/14-AOP996.

On étudie les limites d'échelles de graphes aléatoires construits par coalescence, dont les représentants les plus classiques sont les graphes aléatoires d'Erdős–Rényi. Il a été montré que ces graphes, sous certaines hypothèses de criticités, convergent en limite d'échelle vers les membres d'une famille de lois limites. De plus, cette convergence se renforce en la convergence de tout le processus de coalescence. Les limites consistent en des processus de coalescence éternels sur des collections infinies d'espaces métriques mesurés compacts continus, tous ces processus ayant le même noyau de transition, qu'on appelle le coalescent multiplicatif métrique. La loi d'un processus éternel ayant ce noyau est appelée une loi d'entrée. Dans cette thèse, on cherche à caractériser le coalescent multiplicatif métrique en termes de lois d'entrées extrémales. Pour ce faire on exprime une loi d'entrée comme la coalescence d'une condition initiale loin dans le passé. On cherche donc à comprendre d'une part le comportement d'une loi d'entrée loin dans le passé, et les conditions sous lesquelles la coalescence d'une suite de conditions initiales de plus en plus lointaines converge vers une loi limite connue. Dans ce cadre, on a réussi à montrer que lorsque le graphe considéré est suffisamment homogène, cette convergence a lieu, ce qui montre que les coalescents homogènes sont extrémaux. Pour obtenir ceci, on calcule la loi du graphe limite conditionnellement à la présence d'une composante donnée, ce qui est un résultat intéressant en soi. En outre, la convergence dans le cas homogène possède des applications à la théorie de la sensibilité au bruit.

Pour établir ces démonstrations, on développe divers outils d'apparence relativement technique, mais qui permettent de manipuler plus aisément la loi limite : on montre que la convergence en limite d'échelle s'accompagne de la convergence de familles de réels indexées par les sommets, on montre qu'on peut approximer le graphe limite par la coalescence d'un nombre fini d'espaces métriques, et on montre que pour un coalescent éternel, la borne de masse inférieure est nécessairement vérifiée en probabilité, ce qui permet de transformer des convergences pour la topologie Gromov-faible en convergence pour la topologie Gromov-Hausdorff-Prokhorov.

Enfin, on présente une série de résultats partiels dont le but serait à terme la caractérisation complète de la frontière d'entrée du coalescent multiplicatif métrique.

We study the scaling limits of random graphs constructed by coalescence, the most classical of which are the Erdős–Rényi random graphs. It has been shown that these graphs, under certain criticality assumptions, converge in the scaling limit towards the members of a family of limit distributions. Moreover, this convergence can be strengthened to the convergence of the whole coalescence process. The limit process consists in an eternal coalescence of an infinite collection of continuous compact measured metric spaces, all these processes admitting the same transition kernel which we call the metric multiplicative coalescent. The distribution of an eternal stochastic process with this kernel is called an entrance law. In this thesis, we try to characterize this family of distributions in terms of extreme entrance laws. To do this we see a limit distribution as the coalescence from an initial condition far in the past. We therefore seek to understand on one hand the behavior of an entrance law far in the past and on the other hand the conditions under which the coalescence of an increasingly far in the past sequence of initial conditions converges to a known limit distribution. In this context, we showed that when the random graph considered is sufficiently homogeneous, this convergence takes place, which shows that homogeneous coalescents are extreme. To obtain this, we compute the distribution of the limiting random graph conditionally on the presence of a given component, which is an interesting result in itself. In addition, convergence in the homogeneous case has applications to noise sensitivity theory.

To establish these demonstrations, we develop various technical tools which allow to handle the limit distribution more easily: we show that the scaling limit convergence happens jointly with the convergence of families of real numbers indexed by the vertices. We show that we can approximate the limiting random graph by the coalescence of a finite number of measured metric spaces, and we show that for an eternal coalescent, the lower mass bound is necessarily satisfied in probability, which allows to transform convergences for the Gromov-weak topology into convergences for the Gromov-Hausdorff-Prokhorov topology.

Finally, we present some partial results whose ultimate goal would be the complete characterization of the family of entrance laws of the metric multiplicative coalescent.