UNIVERSITÉ PARIS NORD

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PARIS NORD

Discipline : Mathématiques

Présentée par

Linglong Yuan

Comportements en temps petits des Λ n-coalescents avec l'accent sur les longueurs des branches externes

Soutenue le 3 Décembre 2013 devant le jury composé de :

Président du jury	Jean-François Delmas
Directeur de thèse	Jean-Stéphane Dhersin
Rapporteur	Julien Berestycki
	Götz Kersting
Examinateur	Sylvie Méléard
	Yueyun Hu

致　妈妈陈莲芳

Résumé

Le Λ-coalescent est un processus stochastique utilisé pour modéliser l'arbre généalogique d'une population composée d'une infinité d'individus, admettant des collisions multiples de lignages. Motivée par les applications en biologie, cette thèse étudie principalement les longueurs des branches d'une sous-famille de coalescents associée à n individus (les Λ n-coalescents) qui est largement utilisée par les biologistes et les probabilistes. Nous notons que les Λ-coalescents sont caractérisés par des mesures finies Λ définies sur $[0, 1]$.

Dans un premier temps, nous nous intéressons au cas particulier des Λ n-coalescents avec Λ =Beta$(2 - \alpha, \alpha)$, la mesure Beta lorsque $1 < \alpha < 2$. Pour simplifier les notations, on les appelle Beta$(2-\alpha, \alpha)$ n-coalescents. Möhle a conjecturé dans une communication à Marseille en 2008 que dans ce cas-là, la longueur d'une branche externe choisie au hasard est de l'ordre de $n^{1-\alpha}$ quand n tend vers ∞. Nous montrons que cette longueur normalisée par $n^{\alpha-1}$ converge vers une loi limite que nous expliciterons. Nous présentons deux méthodes pour montrer ce résultat. Ma contribution est d'introduire la construction récursive du Λ n-coalescent qui nous permet d'établir l'une des méthodes. Cette construction consiste à ajouter l'individu n au Λ $(n-1)$-coalescent d'une certaine manière pour constituer un Λ n-coalescent. L'avantage de cette construction est que, d'une part, on peut voir directement comment la branche externe de l'individu n s'attache au coalescent, ce qui facilite le calcul de cette longueur. D'autre part, cette construction donne un processus d'auto-apprentissage des individus en fonction du coalescent construit par les précédents, ce qui offre un autre point de vue qui peut être utile pour d'autres problèmes. L'article (14) basé sur ce travail, en collaboration avec Jean-Stéphane Dhersin, Fabian Freund et Arno Siri-Jégousse, est publié en 2013 dans le revue *Stochastic Processes and their Applications*. Dans cet article, les résultats sont prouvés pour une classe relativement générale de coalescents qui incluent les Beta$(2 - \alpha, \alpha)$-coalescents avec $1 < \alpha < 2$.

Dans un deuxième travail, toujours dans le cadre des Beta$(2-\alpha, \alpha)$ n-coalescents avec $1 < \alpha < 2$, nous étudions le comportement asymptotique du moment d'ordre 2 de la longueur totale des branches externes en utilisant une méthode récursive. Le comportement est décrit d'une façon explicite. On étudie les moments de la longueur d'une seule branche externe ainsi que la covariance entre deux longueurs de branches externes. L'article correspondant est soumis à *Journal of Applied Probability.* Il existe une version plus générale (16) de cet article qui s'intéresse à une classe plus large de Λ-coalescents que les

Beta$(2-\alpha,\alpha)$ n-coalescents avec $1<\alpha<2$.

Un troisième travail cherche à clôturer le problème de la longueur d'une branche externe en proposant une caractérisation pour beaucoup de Λ n-coalescents. Sous certaines conditions sur la mesure Λ (satisfaites en particulier par le coalescent de Bolthausen-Sznitman et les coalescents avec poussière comme les Beta$(2-\alpha,\alpha)$ n-coalescents avec $0<\alpha<1$), nous obtenons un facteur de normalisation "universel" et nous proposons une conjecture pour un résultat d'une classe plus générale de mesures Λ incluant les Beta$(2-\alpha,\alpha)$ n-coalescents avec $1<\alpha<2$. L'outil principal est une nouvelle construction (*measure division construction*) du Λ n-coalescent. Pour cette construction, on divise d'abord Λ en deux parties Λ_1, Λ_2 telles que $\Lambda=\Lambda_1+\Lambda_2$. On part alors d'un Λ_1 n-coalescent que l'on modifie selon la mesure Λ_2 pour avoir finalement un Λ n-coalescent. Si la masse totale de Λ_1 est petite, alors le Λ n-coalescent est proche du Λ_2 coalescent. En ce sens, on peut espérer que même si les coalescents sont perturbés par une autre mesure, certaines propriétés restent encore vraies. Cette méthodologie aboutit à montrer le résultat principal de ce travail qui est dans l'article Y (2013) soumis à *Markov Processes and Related Fields*.

Le dernier travail est une continuation de l'article de (6). Ce dernier étudie le comportement asymptotique en temps petits du processus de fréquence asymptotique ordonnée (Beta$(2-\alpha,\alpha)$ ranked coalescent process) dans le cas où $1<\alpha<2$. Nous utilisons ce processus de fréquence asymptotique ordonnée et la construction de la boîte de peinture de Kingman pour retrouver les Beta$(2-\alpha,\alpha)$ n-coalescents. Ceci nous permet de déduire la loi limite jointe de plusieurs longueurs de branches externes dans le cas des Beta$(2-\alpha,\alpha)$ n-coalescents avec $1<\alpha<2$. Certaines quantités comme la taille du clade minimal et la taille du bloc maximal en temps petits ont aussi été étudiées. Il s'agit d'un travail en collaboration avec Arno Siri-Jégousse. L'article est soumis à *ALEA (Latin American Journal of Probability and Mathematical Statistics)*.

La thèse se compose d'une introduction des résultats, suivie des 4 articles parus ou soumis dans leur version originale.

Notations

1. $Exp(r), r > 0$: la variable exponentielle de paramètre r.
2. $Geo(r), r > 0$: la variable géométrique de paramètre r.
3. $\mathbb{N} := \{1, 2, \cdots\}$.
4. $\mathbb{N}_n := \{1, 2, \cdots, n\}$.
5. $|\pi|$: le nombre de blocs de π et π est une partition d'un ensemble.
6. $\mathcal{P}$: l'ensemble de toutes les partitions de $\mathbb{N}$.
7. $\mathcal{P}_n$: l'ensemble de toutes les partitions de $\mathbb{N}_n$.
8. Λ : une measure finie sur $[0, 1]$.
9. $\nu(dx) = x^{-2}\Lambda(dx)$.
10. $\rho(t) = \int_t^1 \nu(dx)$.
11. $\lambda_{b,k} := \int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx),\ 2 \leq k \leq b$.
12. $g_n := \sum_{k=2}^n \binom{n}{k}\lambda_{n,k},\ n \geq 2$.
13. $\mu_{-1} := \int_0^1 x^{-1}\Lambda(dx)$.
14. $\mu_{-2} := \int_0^1 x^{-2}\Lambda(dx)$.
15. $\mu^{(n)} = \int_{1/n}^1 x^{-1}\Lambda(dx)$.
16. $\bar{\mu}^{(n)} = \int_{1/n}^1 x^{-2}\Lambda(dx)$.
17. $\Pi = (\Pi(t), t \geq 0)$: le Λ-coalescent.
18. $\Pi^{(n)} = (\Pi^{(n)}(t), t \geq 0)$: le Λ n-coalescent.
19. $(K_t, t \geq 0)$: le processus de comptage de blocs de Π.
20. $(K_t^{(n)}, t \geq 0)$: le processus de comptage de blocs de $\Pi^{(n)}$.
21. $T_i^{(n)}$: la longueur de la branche externe de i pour $i = 1, 2, \cdots, n$.
22. $T^{(n)}$: la longueur d'une branche externe choisie uniformément.
23. $\tau^{(n)}$: le nombre total de sauts de $\Pi^{(n)}$.
24. $\sigma^{(n)}$: le nombre de sauts jusqu'à la coalescence de $\{1\}$.
25. $fr(B)$: la fréquence asymptotique de B et B est un bloc de $\mathbb{N}$.
26. $||\pi||_{\downarrow}$: la fréquence asymptotique ordonnée d'une partition π de $\mathbb{N}$.

Remerciements

Sept ans se sont écoulés depuis mon arrivée en France. Impressionné au premier regard par le beau paysage de ce pays, je me suis ensuite plongé dans les études du cursus polytechnicien. C'est grâce à Sylvie Méléard, qui nous a présenté la théorie élégante de probabilité lors du cours de tronc commun, que j'ai commencé à aimer le monde des probabilités. Elle me fait l'honneur d'être dans le jury de cette thèse et je la remercie. Jean-François Delmas et Amaury Lambert m'ont aidé en petites classes à résoudre les exercices et à découvrir les techniques de cette matière. Grâce au travail de ces excellents chercheurs, il était un vrai plaisir de penser aux problèmes probabilistes en se baladant sur le plateau de l'X. Fasciné par la beauté de la théorie des probabilités, je me demandais à quel niveau elle pourrait influencer le monde humain et le monde des mathématiques. Cette naïve intérogation de soi fait toujours partie de mes plus grandes motivations pour continuer la recherche en probabilité. Ce merveilleux souvenir de l'X restera toute ma vie au sein de mon coeur. Je remercie Jean-François Delmas d'être aujourd'hui membre du jury.

Ma vie s'est déplacée ensuite à Villetaneuse, accueillie par l'équipe de Proba-Stat dirigée par Yueyun Hu. Son encouragement pour ma recherche m'a accompagné durant toute la thèse. Je le remercie également d'être dans mon jury. Le plus grand remerciement revient à mon directeur, Jean-Stéphane Dhersin. Il m'a guidé dans le domaine des coalescents avec ses connaissances et sa chaleureuse attitude. Il m'a appris les exigences d'une bonne recherche et comment perfectionner un article “cru”. En un mot, il m'a ouvert la porte du monde de la recherche. Lors d'un stage à Francfort qui s'est révélé très fructueux, j'ai eu l'occasion de travailler avec Götz Kersting et Anton Wakolbinger, auxquels je dois un grand remerciement. En particulier, je remercie Götz Kersting d'avoir accepté de rapporter cette thèse. Je remercie également Julien Berestycki, l'autre rapporteur. Les travaux de ces deux rapporteurs ont été pour moi une précieuse source d'inspiration.

Faire une thèse sans être avec des amis est impossible et je suis très heureux et fier d'en avoir de très bons. D'abord, j'aimerais remercier Arno et Fabian, qui sont non seulement mes deux chers collaborateurs mais aussi sont de très bons amis. Arno m'a donné de nombreuses recommendations qui ont été fortes utiles au moment de la rédaction de la thèse. J'ai bien aimé la poudre magique que Bakari, mon cher frère de thèse, m'a fait partager pour faire la cuisine malienne. Il est toujours prêt à m'aider quand j'ai des soucis dans la vie. Il en est de même de Phong, un ami vietnamien et avec qui nous avons partagé de nombreux moments festifs et de fructueuses discussions mathématiques. Ma vie à Paris a beaucoup été enrichie par Thomas qui m'a invité à des soirées avec ses amis, parmi lesquels Xan, très chaleureux et sympathique. Avec Olivier nous avons eu de belles expériences ensemble, que ce soit tous invités par Jean-François Delmas aux Journées MAS, ou lors d'un très agréable séjour à Francfort. Mon remerciement revient aussi à Clément qui est maintenant Maître de Conférences à Paris 13 et qui m'a aussi beaucoup aidé dans la rédaction du manuscrit. Les discussions et les échanges que l'on a eu à Paris 13 sont toujours inspirants.

Il y a aussi Patrick, Cécile, Miraine, Mathieu à SMILE ; Boubacar, Mamadou, Julien à Marseille ; Camille, Manon, Clément, Charline à l'X. J'ai aussi mes amis chinois de l'X, Zhiping Rao, Zixian Jiang, Lingxuan Shao, pour sortir de temps en temps.

Je dédie cette thèse à ma très chère mère qui m'a toujours soutenu et de qui je tire mes forces pour avancer dans la vie, depuis le décès de mon père en 1996.

Table des matières

Chapitre 1

Introduction et résultats principaux

1.1 Quelques éléments d'Histoire de l'évolution et de la coalescence

Notre monde est rempli d'une diversité extraordinaire d'organismes vivants. Des tentatives d'explication de ce phénomène se trouvent dans les écrits de nombreux philosophes. Anaximandre (vers 610-514 av. J.-C.), philosophe de la Grèce Antique, est considéré comme étant le premier à expliquer l'origine des êtres d'un point de vue « scientifique ». Il affirme que tout animal vivant sur les terres a un ancêtre lointain dans la mer. Par l'action du soleil sur l'humidité, les terres sont apparues et certains animaux s'y sont adaptés pour vivre en dehors de la mer. Presque à la même époque, le philosophe Empedocle (vers 610-514 av. J.-C.) dit que les premiers animaux et plantes sont des morceaux des êtres actuels. Lorsque deux animaux (ou plantes) se croisent par hasard, ils se connectent pour créer une nouvelle espèce. L'aléa joue donc un rôle important. Platon (vers 428-348 av.J.-C.) considère que la forme d'un objet quelconque est une version imparfaite de celle idéale créée par le Démiurge. Aristote (384-322 av. J.-C.) développe cette théorie en disant en particulier que tous les caractères d'une espèce sont conçus pour une cause finale, ce qui entraîne beaucoup d'interprétations religieuses. Le point de vue de Platon et Aristote est devenu dominant au Moyen-Âge. Plusieurs autres théories de la Grèce Antique disparaissent à cette époque dans le monde occidental mais sont reprises dans le monde islamique.

Au 17ème siècle, le philosophe français René Descartes (1596-1650) a l'idée que l'univers se développe d'une façon mécanique en incluant les êtres vivants, sans aucune intervention divine. Au contraire, Gottfried Wilhelm Leibniz (1646-1716) considère que l'univers évolue de manière perpétuelle et sans contrainte en accumulant la beauté et la perfection des œuvres de Dieu. Assez différemment, le taxinomiste britannique John Ray (1627-1706) soutient que toute espèce est caractérisée par une partie essentielle qui ne change pas au cours du temps.

L'étape suivante est sans doute due au philosophe français, Georges-Louis Buffon (1707-1788), qui présente sa théorie sur l'origine de la terre et les doutes sur la fixité des espèces (l'idée de John Ray) dans son livre *l'Histoire naturelle* . Selon lui, les ancêtres des êtres présents sont nés spontanément et les descendants ont hérité et aussi dégradé les caractères de leurs ancêtres en raison du changement environnemental. La dégénération mène les espèces à converger vers leurs formes primitives. Mais un être dégénéré peut reprendre sa forme originale s'il est placé dans un environnement favorable. Le philosophe français Denis Diderot (1713-1784) a étendu l'idée de Buffon en ajoutant que les nouvelles formes des espèces apparaissent au travers d'un processus expérimental en fonction des essais et des erreurs, ce qui est considéré comme une anticipation partielle de la sélection naturelle.

Au 18ème siècle, certains naturalistes croient que les hommes et les primates ont les mêmes ancêtres et que tout être vivant a développé ses propres méthodes pour évoluer sur une longue terme. Erasmus Darwin (1731-1802), le grand père de Charles Darwin, suggère que tout animal au sang chaud est descendu d'une espèce minuscule.

Au 19ème siècle, Jean-Baptiste Lamarck (1744-1829), naturaliste français, propose une théorie qui consiste à dire que le développement des organes et leur force d'action sont constamment liés à l'emploi de ces organes. Les individus transmettent ensuite leurs caractères acquis aux descendants. Mais la théorie de l'hérédité n'est pas fournie dans le *Lamarckisme.* En 1859, Charles Darwin (1809-1882) publie sa théorie de l'évolution dans son fameux ouvrage *On the Origin of Species.* La principale différence entre le *Lamarckisme* et le *Darwinisme* est que pour ce dernier, les changements des traits des organes sont dus à la sélection naturelle, et non à l'emploi des organes. Plus précisément, les individus dans une population ont des traits différents qui sont soumis au changement environnemental. Seuls les individus les plus adaptés peuvent survivre. Ici un individu adapté est un individu ayant des traits héréditaires qui aident à la reproductivité. Lorsque l'environnement change, certains traits sont choisis et seront plus tard partagés par toute la population. Comme les changements géologiques et géographiques engendrent des environnements diversifiés et que les êtres vivants se déplacent au cours des générations, les espèces évoluent suivant des directions différentes.

Le défaut de cette théorie est que la sélection naturelle s'effectue sur des variations des traits dont l'origine n'est pas précisée. Darwin écrit dans son livre que

Our ignorance of the laws of variation is profound. Not in one case out of a hundred can we pretend to assign any reason why this or that part has varied.

Le point de vue accepté à l'époque à propos du mécanisme de l'hérédité est *l'hérédité par mélange.* Cela consiste à dire que le phénotype d'un individu provient du mélange dans certaines proportions des influences parentales. Mais on peut facilement en déduire que, dans un environnement donné, la variance des traits d'une espèce diminue au cours du temps. Finalement, les individus de la même espèce sont identiques. A cause de ce défaut, le *Darwinisme* a beaucoup de mal à se faire accepter à l'époque.

C'est grâce au moine silésien Gregor Mendel (1822-1884) que le mécanisme d'hérédité est découvert. En 1865, Mendel publie l'article *Experiments on Plant Hybridization* basé sur l'expérimentation sur des pois. Il montre que la variation génétique est à l'origine de la variation des traits. Il y décrit ce qui deviendront les lois de Mendel. Ces dernières engendrent le principe de Hardy-Weinberg qui dit que la variation génétique est préservée dans certaines populations isolées sans mutations et sans sélection naturelle. Par conséquent, les lois de Mendel remédient au défaut du *Darwinisme* sur le problème de variation des traits. Mais elles n'étaient malheureusement pas connues par Charles Darwin. En 1900, elles sont retrouvées et rapidement acceptées par les scientifiques. Dans les années 1920-1940, Ronald Fisher (1890-1962), John Burdon Sanderson Haldane (1892-1964) et Sewall Wright (1889-1988) initient la recherche sur la génétique des populations, laquelle consiste à étudier les phénomènes d'adaptation et de spéciation en utilisant la sélection naturelle et les lois de Mendel qui s'appliquent à travers différents ingrédients : les mutations, les recombinaisons, les immigrations et aussi les dérives génétiques.

Grâce au développement de nouvelles technologies en Physique et en Informatique au début du 20ème siècle, les observations peuvent s'effectuer au niveau moléculaire. En 1953, James Watson (1928-) et Francis Crick (1916-2004) relèvent que les gènes sont de « petits morceaux d'ADN ». Une ADN est formée d'une double hélice où une suite de bases génétiques couple une autre suite complémentaire. Très rapidement, les scientifiques comprennent comment l'ADN code les protéines via les codes génétiques. On s'intéresse ensuite aux variations génétiques au niveau moléculaire. Il est montré que la plupart des différences entre les génomes des espèces vivantes sont sélectivement neutres, i.e., ces différences génomiques n'influencent pas les niveaux de fitness. Motoo Kimura (1924-1994) soutient l'hypothèse qu'il y a beaucoup de variations génétiques mais que seule une très petite fraction est soumise à la sélection naturelle, et que la plupart des changements sont neutres en fitness. La théorie de Kimura, autrement appelée la théorie neutre de l'évolution moléculaire, fournit un modèle sur lequel les données peuvent être testées. Plus précisément, on étudie théoriquement la configuration génétique d'une espèce lorsque les changements génétiques sont supposés neutres. Des tests statistiques permettent alors de tester la neutralité. Si le modèle est neutre, l'évolution est très peu influencée par la sélection naturelle et beaucoup par les dérives génétiques sur les allèles neutres. Le débat sur l'importance relative de la dérive génétique par rapport à la sélection naturelle est toujours d'actualité.

Les premières idées à propos de la coalescence apparaissent dans Robert C. (32), mais c'est à Sir John Frank Charles Kingman (1939-) que l'on doit une théorie mathémathiquement élégante pour étudier l'arbre généalogique d'un échantillon d'une population qui permet de mesurer les variations génétiques et tester la neutralité de la population. Dans une série d'articles publiés en 1982, Kingman introduit le processus de coalescence, appelé plus tard le coalescent de Kingman, et qui peut être associé à l'arbre généalogique d'une population de grande taille. Nous invitons le lecteur à se référer à (34), (18), (41), (42), (40), ou encore (47) pour plus d'informations. Un fait étonnant est que la classe

FIGURE 1.1 – Gregor Mendel (1822-1884)

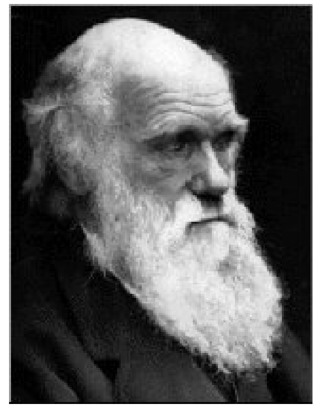


FIGURE 1.2 – Charles Darwin (1809-1882)

des populations où ce résultat s'applique est assez générale, et aussi que le mécanisme précis de la reproduction n'est pas important. L'hypothèse cruciale de la théorie neutre est qu'indépendamment de l'arbre généalogique donné, les mutations arrivent à un taux constant sur cet arbre. C'est sur cette hypothèse que repose le test de Fu et Li introduit par Yun-Xin Fu et Wen-Hsiung Li dans leur article de 1993. En pratique, on considère que les mutations arrivent sur le coalescent de Kingman à un taux constant que l'on veut estimer. Le test de Fu et Li s'appuie alors sur les longueurs de certaines branches de l'arbre généalogique. Ce sont ces quantités qui sont étudiées dans cette thèse.

FIGURE 1.3 – Motoo Kimura (1924-1994)

Le coalescent de Kingman est un processus qui n'admet que des collisions binaires. Des études sur les espèces marines montrent que certains individus peuvent avoir beaucoup d'enfants, et que la probabilité pour que la plupart des individus de la génération suivante proviennent d'un seul individu est assez grande. Cela est découvert par (33), (20), (12), (1). Dans ce cas-là, le coalescent de Kingman n'est plus adapté pour modéliser l'arbre généalogique. En 1999, Jim Pitman et Serik Sagitov introduisent indépendamment et en même temps le Λ-coalescent. Le « Λ » correspond à une mesure finie sur $[0, 1]$ qui caractérise ce processus. Ces processus à coalescences multiples incluent le coalescent de Kingman pour lequel la mesure Λ est la mesure de Dirac au point $\{0\}$. Les Λ-coalescents ne sont qu'une sous-classe des coalescents échangeables généraux appelés Ξ-coalescents introduits par Martin Möhle, Serik Sagitov et Jason Schweinsberg en 2000-2001 et qui admet des coalescences multiples et simultanées. Il est important de remarquer que le coalescent de Kingman est un cas particulier de Λ-coalescent qui est encore un cas particulier de Ξ-coalescent. Il y a aussi d'autres extensions du processus de coalescence dont un excellent exemple est de supposer que les individus sont localisés sur plusieurs endroits séparés avec immigrations possibles et ensuite de sélectionner aléatoirement un échantillon parmi toute la population pour étudier l'arbre généalogique (on peut se référer à (21) pour plus de détails).

1.2 Le Λ-coalescent

1.2.1 Consistance et échangeabilité

Le coalescent de Kingman est apparu avant le Λ-coalescent. Mais ce dernier étant plus général, on va l'introduire et voir le coalescent de Kingman comme un cas particulier.

Définition 1.1. *Soit* Λ *une mesure finie sur* $[0, 1]$. *Le* Λ*-coalescent est un processus de Markov* $\Pi = (\Pi(t), t \geq 0)$ *à temps continu à valeurs dans* $\mathcal{P}$, *l'ensemble des partitions de* $\mathbb{N} := \{1, 2, \cdots\}$, *tel que son état initial* $\Pi(0) = \{\{1\}, \{2\}, \cdots\}$, *la partition en singletons. La loi de* Π *est caractérisée par* $\Pi^{(n)} = \{\Pi^{(n)}(t), t \geq 0\}$ *qui est la restriction de* Π *sur* $\mathbb{N}_n := \{1, 2, \cdots, n\}$, *et à valeurs dans* $\mathcal{P}_n$, *l'ensemble des partitions de* $\mathbb{N}_n$. $\Pi^{(n)}(0) = \{\{1\}, \{2\}, \cdots, \{n\}\}$ *et si à l'instant* $t \geq 0$, $\Pi^{(n)}(t)$ *a* b *blocs* ($b \geq 2$), *alors chaque* k*-uplet de blocs* ($2 \leq k \leq b$) *coalesce en un seul bloc indépendamment des autres au taux*

$$\lambda_{b,k} := \int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx) = \int_0^1 x^k(1-x)^{b-k}\nu(dx), \tag{1.1}$$

où $\nu(dx) = x^{-2}\Lambda(dx)$. *Ainsi* $\Pi^{(n)}$ *effectue sa prochaine coalescence au taux*

$$g_b := \sum_{k=2}^{b} \binom{b}{k} \lambda_{b,k}. \tag{1.2}$$

Le processus $\Pi^{(n)}$ *est appelé le* Λ n*-coalescent. On note* $|\Pi(t)|$ *le nombre de blocs de* $\Pi(t)$, $|\Pi^{(n)}(t)|$ *le nombre de blocs de* $\Pi^{(n)}(t)$. *On définit les processus de comptage de blocs* $K = (K(t), t \geq 0)$ *et* $K^{(n)} = (K^{(n)}(t), t \geq 0)$ *où* $K(t) = |\Pi(t)|$, $K^{(n)}(t) = |\Pi^{(n)}(t)|$.

Remarque 1.1. *Si on munit $\mathcal{P}$ d'une topologie telle qu'une suite d'éléments $\{p_1, p_2, \cdots, \}$ dans $\mathcal{P}$ converge vers p si et seulement si pour tout $n \geq 1$, il existe N tel que les restrictions de p_i et p sur $\mathbb{N}_n$ sont égaux pour $i \geq N$, alors Π est un processus càdlàg par rapport à cette topologie. Une telle topologie peut se trouver dans le Lemme 9 de (22).*

Remarque 1.2. *Le fait que la loi de Π peut être établie à partir de celles de $\{\Pi^{(n)}, n \geq 2\}$ est dû au Théorème d'extension de Kolmogrov. On invite le lecteur à consulter la page 33 du (7). On peut aussi le voir directement à partir d'une construction récursive qui sera décrite plus tard.*

Il y a deux propriétés essentielles pour le Λ-coalescent.

Propriété 1 : Consistance.

Soit $2 \leq n < m \leq \infty$ et ς la restriction de $\mathbb{N}_m$ sur $\mathbb{N}_n$ (on note $\mathbb{N}_\infty = \mathbb{N}$, $\Pi^{(\infty)} = \Pi$). Alors

$$\varsigma \circ \Pi^{(m)} \overset{(d)}{=} \Pi^{(n)}.$$

Remarque 1.3. *La consistance découle en fait de la relation suivante :*

$$\lambda_{b,k} = \lambda_{b+1,k} + \lambda_{b+1,k+1}. \tag{1.3}$$

Celle-ci caractérise une mesure Λ, et donne la formule (1.1) (voir (48)). On peut expliquer (1.3) de la façon suivante : supposons qu'un Λ n-coalescent a $b+1$ blocs numérotés de 1 à $b+1$ à l'instant t. L'événement que k ($2 \leq k \leq b$) blocs soient impliqués à la prochaine coalescence parmi les blocs de 1 à b peut être décomposé en deux sous-événements : le premier est d'avoir une coalescence de $k+1$ concernant les k blocs précédents et aussi le bloc $b+1$ et le deuxième est d'avoir une coalescence ne concernant que les k blocs. Alors la consistance implique que le taux de l'événement tout entier est la somme des taux des deux sous-événements, d'où vient (1.3).

Propriété 2 : Echangeabilité.

Soit ω une permutation de $\mathbb{N}_n$, alors

$$\omega \circ \Pi^{(n)} \overset{(d)}{=} \Pi^{(n)}. \tag{1.4}$$

Remarque 1.4. *La propriété d'échangeabilité traduit la neutralité entre les individus.*

Exemples importants de Λ-coalescents

— Le coalescent de Kingman.

Le coalescent de Kingman est le cas particulier de Λ-coalescent le plus utilisé dans les applications ((38)). La mesure Λ du coalescent de Kingman est la mesure de Dirac au point $\{0\}$. Dans ce cas-là, toutes les coalescences sont binaires.

Dans la simulation de la Figure 1.4, on voit que le nombre de blocs décroît très rapidement. Le temps pour aller de l'infini aux deux plus anciens ancêtres communs

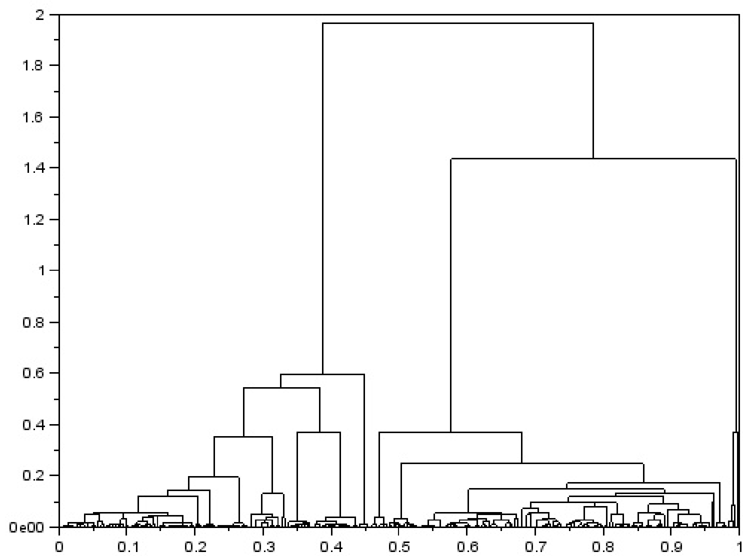


FIGURE 1.4 – Simulation du 300-coalescent de Kingman.

est comparable au temps pour aller de ces deux ancêtres au MRCA (*most recent common ancestor*).

— Les Beta($2-\alpha,\alpha$)-coalescents avec $0<\alpha<2$. Ces processus avec différentes valeurs de α sont aussi beaucoup étudiés dans la littérature. La mesure Λ du Beta($2-\alpha,\alpha$)-coalescent est comme suit :

$$\Lambda(dx) = \frac{\int x^{1-\alpha}(1-x)^{\alpha-1}\mathbf{1}_{0\leq x\leq 1}dx}{Beta(2-\alpha,\alpha)},$$

où $Beta(2-\alpha,\alpha)$ est la fonction Beta avec paramètres $2-\alpha$ et α.
Pour simplifier les notations, on appelle le Beta($2-\alpha,\alpha$)-coalescent aussi le Beta-coalescent. Le Beta-coalescent modélise l'arbre généalogique d'une population pour laquelle un individu peut avoir beaucoup d'enfants. Plus α est petit, plus la probabilité pour qu'un individu dans la population puisse avoir des enfants qui forment une partie considérable de la génération suivante est élevée. On peut voir dans la simulation des Figures 1.5 et 1.6 que lorsque α est petit, les coalescences impliquent plus de blocs.
Si $\alpha=1$, le Beta-coalescent est appelé en particulier coalescent de Bolthausen-Sznitman. Il fut pour la première fois appliqué en Physique pour modéliser un phénomène de verres de spin, alors même que la notion de Λ-coalescent n'avait pas

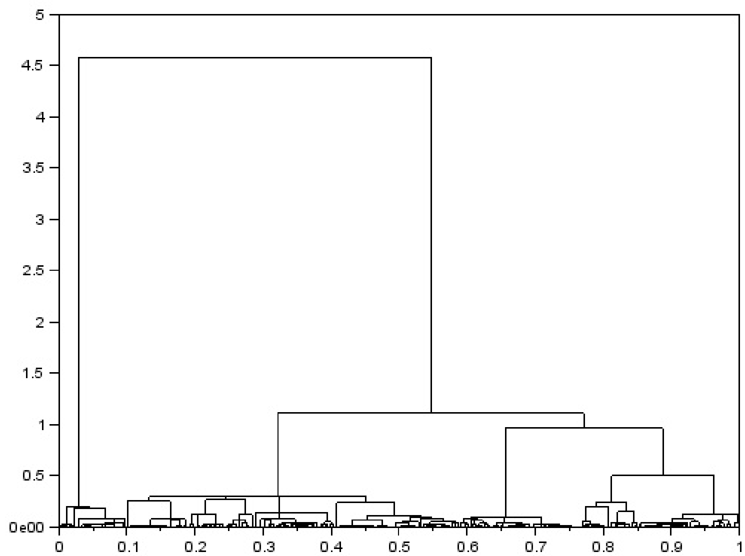


FIGURE 1.5 – Simulation du Beta(0,2, 1,8) 300-coalescent

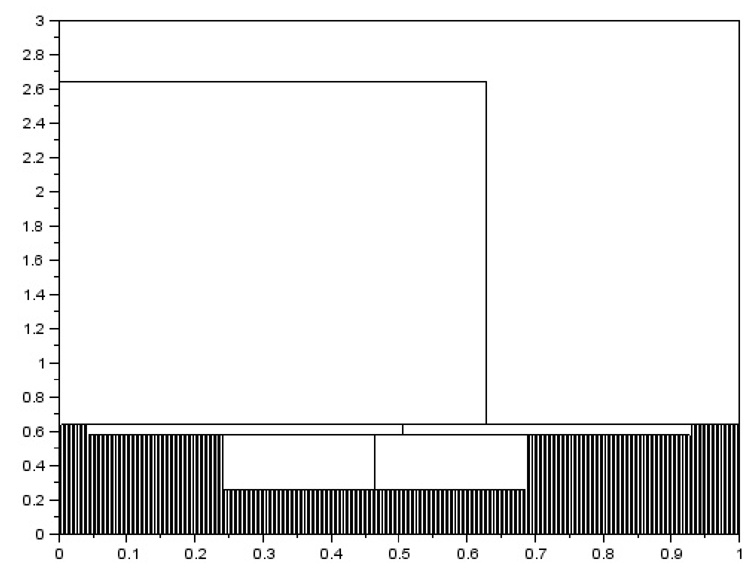


FIGURE 1.6 – Simulation du Beta(1,8, 0,2) 300-coalescent

été introduite ((11)).

Le Beta-coalescent apparaît aussi dans le contexte des processus de branchements ((52)), des processus à valeurs mesures ((24), (17), (8), (25), (26)), des arbres aléatoires ((3), (5)), etc.

1.2.2 Descente de l'infini

Pour tout Λ-coalescent Π, on part de la partition en singletons et les blocs coalescent au fur et à mesure. Donc on voit que le processus de comptage de blocs est décroissant. On remarque que si pour un temps $t_0 > 0$, $K(t_0)$ est infini, par la propriété de Markov, $K(t)$ est infini pour tout $t \geq 0$. Par conséquent, soit $K(t)$ est infini pour tout $t \geq 0$, soit Π descend de l'infini, i.e, $K(t) < \infty$ pour tout $t > 0$ presque sûrement. Donc on souhaite avoir un critère pour distinguer les deux cas. Ce critère est donné par (51).

Théorème 1.1. *Le Λ-coalescent Π descend de l'infini si et seulement si $\sum_{b=2}^{\infty} \gamma_b^{-1} < \infty$, où $\gamma_b = \sum_{i=2}^{b}(k-1)\binom{b}{k}\lambda_{b,k}$.*

Pour comprendre ce théorème intuitivement, on peut penser à la chose suivante. Remarquons que $g_b = \sum_{k=2}^{b} \binom{b}{k}\lambda_{b,k}$ est le taux de coalescence lorsque l'on a b blocs. Si une coalescence de k blocs arrive, on en perd $k-1$. L'idée est de décomposer une perte de k en $k-1$ pertes de 1. Imaginons un processus décroissant $\bar{K}^{(n)}$ à valeur dans $\mathbb{N}_n$ tel que $\bar{K}^{(n)}(0) = n$ et lorsqu'il a b blocs, la prochaine perte est 1 et aura lieu au taux γ_b, ainsi de suite. Soit $\bar{K}$ la limite de convergence fini-dimensionelle de $(\bar{K}^{(n)}, n \geq 2)$ et donc $\bar{K}(0) = \infty$. Alors $\bar{K}$ descend de l'infini si et seulement si la durée de vie de $\bar{K}^{(n)}$ tend vers une limite finie lorsque n tend vers ∞. Soit $\{e_i\}_{i\geq 1}$ les variables i.i.d exponentielles de paramètre 1. Alors la durée de vie de $\bar{K}^{(n)}$ est égale en loi à

$$\sum_{i=2}^{n} \frac{e_i}{\gamma_i},$$

qui tend vers une limite finie si et seulement si $\sum_{b=2}^{\infty} \gamma_b^{-1} < \infty$.

Ensuite on voit que $|\Pi|$ et $\bar{K}$ sont comparables en terme de la vitesse de la descente. Supposons que les deux processus ont b blocs au temps t. Alors pour Δt très petit,

$$\begin{aligned} b - \mathbb{E}[|\Pi(t+\Delta t)|] &= g_b \left(\sum_{k=2}^{b} (k-1) \frac{\binom{b}{k}\lambda_{b,k}}{g_b} \right) \Delta t + o(\Delta t) \\ &= \sum_{k=2}^{b} \gamma_{b,k} \Delta t + o(\Delta t), \end{aligned}$$

où $g_b\Delta t + o(\Delta t)$ est la probabilité d'avoir une seule coalescence.

Pour $\bar{K}$, on perd un bloc à chaque coalescence, donc

$$b - \mathbb{E}[|\bar{K}(t + \Delta t)|] = \sum_{k=2}^{b} \gamma_{b,k} \Delta t + o(\Delta t),$$

où le terme à droite est la probabilité d'avoir une seule coalescence. On voit clairement que, au moins au sens de l'espérance, les vitesses de la descente sont comparables. Donc il n'est pas étonnant d'apprendre que Π descend de l'infini si et seulement si $\bar{K}$ descend de l'infini.

Les Λ-coalescents qui descendent de l'infini incluent les Beta$(2-\alpha, \alpha)$-coalescents avec $1 < \alpha < 2$ ainsi que le coalescent de Kingman. Les Beta$(2-\alpha, \alpha)$-coalescents avec $0 < \alpha \leq 1$ restent toujours infinis.

1.2.3 La boîte de peinture et la fréquence asymptotique ordonnée

On introduit d'abord la notion de partition échangeable de $\mathbb{N}_n$ et de $\mathbb{N}$.

Définition 1.2. *Soit ω_n une permutation quelconque de $\mathbb{N}_n$ et π_n une partition aléatoire de $\mathbb{N}_n$. Alors π_n est une partition échangeable de $\mathbb{N}_n$ si et seulement si*

$$\omega_n \circ \pi_n \overset{(d)}{=} \pi_n.$$

Soit ς_n la restriction de $\mathbb{N}$ vers $\mathbb{N}_n$ et π une partition aléatoire de $\mathbb{N}$, alors π est une partition échangeable de $\mathbb{N}$ si et seulement si pour tout $n \geq 1$

$$\omega_n \circ \varsigma_n \circ \pi \overset{(d)}{=} \varsigma_n \circ \pi.$$

Il est évident que $\Pi(t)$ est une partition échangeable pour tout $t \geq 0$. (38) prouve qu'une partition aléatoire échangeable de $\mathbb{N}$ peut être obtenue via la boîte de peinture (*paintbox*) à partir d'une partition de la masse 1.

Procédure de la boîte de peinture :

Soit l'ensemble de partitions des masse 1 défini comme

$$P_m := \{s = (s_1, s_2, \cdots) : s_1 \geq s_2 \geq s_3 \cdots ; \sum_{i=1}^{\infty} s_i \leq 1.\}$$

On définit $s_0 := 1 - \sum_{i=1}^{\infty} s_i$ appelée la « poussière ». Soit $s \in P_m$, on divise l'intervalle $[0, 1]$ en sous-intervalles dont les longueurs sont respectivement $s_0, s_1, s_2, \cdots$. L'intervalle correspondant à s_0 est appelé l'intervalle de poussière. Chaque intervalle est dotée d'une couleur différente des autres, sauf l'intervalle de poussière. Soit $\{U_1, U_2, U_3, \cdots\}$ une suite de variables uniformes indépendantes sur $[0, 1]$ et aussi indépendantes de s. Si U_i tombe dans un intervalle autre que l'intervalle de poussière, alors on donne la couleur de cet intervalle à i. Sinon, on ne donne aucune couleur à i. Les is sans couleurs deviennent des singletons et chaque couleur forme un bloc en regroupant certains is (la probabilité pour que la taille du groupe d'une couleur soit finie est 0). On note cette partition π_s et sa loi ς_s. Il est facile de voir que π_s est une partition échangeable.

Avant d'énoncer le Théorème de la boîte de peinture de Kingman, on introduit la notion de la fréquence asymptotique ordonnée d'une partition de $\mathbb{N}$.

Définition 1.3 (Fréquence asymptotique ordonnée)**.** *Soit B un sous-ensemble de $\mathbb{N}$. S'il existe une valeur b non-negative telle que quand n tend vers ∞*

$$\frac{1}{n}\#\{B \cap \mathbb{N}_n\} \longrightarrow b,$$

alors on dit que B admet une fréquence asymptotique notée $fr(B) := b$.

Soit π une partition de $\mathbb{N}$. On dit que π admet une fréquence asymptotique si tout bloc de π admet une fréquence asymptotique. On ordonne les fréquences asymptotiques des blocs de π en ordre décroissant et note cette suite $||\pi||_\downarrow$. Alors $||\pi||_\downarrow$ est appelée la fréquence asymptotique ordonnée de π.

Théorème 1.2 (Le Théorème de la boîte de peinture de Kingman)**.** *Soit π une partition aléatoire échangeable de $\mathbb{N}$. Alors π admet une fréquence asymptotique ordonnée presque sûrement et la loi de π peut être exprimée par*

$$\mathbb{P}(\pi \in \cdot) = \int_{P_m} \mathbb{P}(||\pi||_\downarrow \in ds)\varsigma_s(\cdot),$$

où ς_s est la loi de la partition induite par s via la procédure de la boîte de peinture.

Remarque 1.5. *L'avantage de ce théorème est que la loi de π est caractérisée par $||\pi||_\downarrow$, un élément de P_m. Sachant $||\pi||_\downarrow$, on peut retrouver π en utilisant la procédure de la boîte de peinture. Lorsque l'on connait bien la loi de $||\pi||_\downarrow$, les études de π deviennent abordable, sachant que la procédure de la boîte de peinture est un problème d'occupation de cases sur* $[0, 1]$ *(finies ou dénombrablement infinies).*

Définition 1.4. *Soit* Π *un* Λ*-coalescent.* Π *est sans poussière si et seulement si pour $t \geq 0$ quelconque, $||\Pi(t)||_\downarrow$ est sans poussière presque sûrement. De même,* Π *est avec poussière si et seulement si pour $t \geq 0$ quelconque, $||\Pi(t)||_\downarrow$ est avec poussière presque sûrement.*

En effet, Π est soit sans poussière, soit avec poussière. Il n'y a pas de troisième possibilité. Ce fait est montré par (48).

Théorème 1.3. *Un* Λ*-coalescent* Π *est sans poussière si et seulement si*

$$\int_0^1 x^{-1}\Lambda(dx) = \infty.$$

Sinon, Π *est avec poussière.*

D'après ce résultat, le coalescent de Kingman et les Beta$(2-\alpha,\alpha)$-coalescents avec $1 \leq \alpha < 2$ sont sans poussière et les Beta$(2-\alpha,\alpha)$-coalescents avec $0 < \alpha < 1$ sont avec poussière.

Remarque 1.6. *Un* Λ*-coalescent qui descend de l'infini est un coalescent sans poussière. La réciproque est fausse. On peut prendre le coalescent de Bolthausen-Sznitman qui ne descend pas de l'infini mais est sans poussière.*

Corollaire 1.1. *Soit* Π *un* Λ*-coalescent. Il existe un processus* $\Theta = (\Theta(t), t \geq 0)$ *à valeurs dans* P_m *tel que presque sûrement* $\Pi(t)$ *admet une fréquence asymptotique ordonnée* $\Theta(t)$ *pour tout* $t \geq 0$*. On appelle* Θ *le processus de fréquence asymptotique ordonnée de* Π *(ranked* Λ*-coalescent).*

Démonstration. On divise la discussion en deux parties.

1. Π **est un processus sans poussière.** On peut se référer à la Proposition 30 de (22) dans ce cadre-là. Ici, on donne une explication essentiellement identique mais plus détaillée. Soit E_r l'événement où $\Pi(t)$ admet une fréquence asymptotique ordonnée sans poussière pour toute valeur rationnelle non-negative t. Alors d'après le Théorème 1.2, on a

$$\mathbb{P}(E_r) = 1.$$

Soit s une valeur positive irrationnelle, on peut toujours trouver une valeur rationnelle s' telle que $0 < s' < s$. Supposons que $\Pi(s) = \{B_k\}_{k\geq 1}$ et $B_k = \cup_{i=1}^{\infty} A_{k,i}$ où $A_{k,i}$ est un bloc vide ou bien un bloc ayant une fréquence asymptotique positive au temps s'. D'après le Lemme de Fatou, on a pour tout $k \geq 1$

$$\liminf_{n\to\infty} \frac{1}{n}\#\{B_k \cap \mathbb{N}_n\} \geq \sum_{i=1}^{\infty} fr(A_{k,i}).$$

Sachant E_r, $\sum_{k=1}^{\infty}\sum_{i=1}^{\infty} fr(A_{k,i}) = 1$. De plus, $\sum_{k=1}^{\infty} \frac{1}{n}\#\{B_k \cap \mathbb{N}_n\} = 1$. Alors on a pour tout $k \geq 1$

$$\lim_{n\to\infty} \frac{1}{n}\#\{B_k \cap \mathbb{N}_n\} = \sum_{i=1}^{\infty} fr(A_{k,i}). \tag{1.5}$$

Comme s est arbitraire, on peut conclure en notant $\Theta = \{||\Pi(t)||_{\downarrow}, t \geq 0\}$.

2. Π est un processus avec poussière. Lorsque Π est avec poussière, on a une autre construction via un processus de subordinateur. Cette construction se trouve à la page 7 de (31) et son origine est dans (48).

Soit $\tilde{\nu}$ la mesure image de ν (rappelons que $\nu(dx) = x^{-2}\Lambda(dx)$) sous la transformation $x \to -\ln(1-x)$. Soit $(\tilde{S}_t, t \geq 0)$ le subordinateur de la mesure de Lévy $\tilde{\nu}$ et $S_t = e^{-\tilde{S}_t}$. Alors $(S_t, t \geq 0)$ est un processus décroissant de sauts purs tendant vers 0 et $S_0 = 1$. Soit $\{U_i\}_{i\geq 1}$ des variables uniformes indépendantes sur $(0,1]$ et aussi indépendantes de $(S_t, t \geq 0)$. Pour tout $i \geq 1$, on associe U_i au singleton $\{i\}$. Fixons un $n \in \mathbb{N}$. Soit t_1 le premier instant où on a au moins un élément de $\{U_i\}_{1\leq i\leq n}$ dans $(S_{t_1}, S_{t_1-}]$. Alors $\Pi^{(n)}(t)$ peut être défini de la façon suivante : $\Pi^{(n)}(s) = \{\{1\},\{2\},\cdots,\{n\}\}$ pour $0 \leq s < t_1$; A l'instant t_1, supposons que les $\{U_i\}_{i\in I_1}$ sont comprises entre $(S_{t_1}, S_{t_1-}]$ où I_1 est un sous-ensemble de $\mathbb{N}_n$. On fusionne les singletons de I_1 pour avoir un grand bloc qui sera envoyé uniformément sur $(0, S_{t_1}]$. On pose ensuite $\Pi^{(n)}(t_1) = \{I_1, \{i\}_{i\notin I_1}\}$. Soit $t_2 > t_1$ le prochain instant où $(S_{t_2}, S_{t_2-}]$ contient au moins un bloc (les singletons de I_1^c ou le nouveau bloc représentant I_1), alors on pose $\Pi^{(n)}(s) = \Pi^{(n)}(t_1)$ pour $t_1 \leq s < t_2$. On fusionne ensuite les blocs sur $(S_{t_2}, S_{t_2-}]$ en un grand bloc qui sera envoyé uniformément sur $(0, S_{t_2}]$, et ainsi de suite. En résumé, $\Pi^{(n)}(t)$ est l'ensemble des blocs sur $(0, S_t]$ et de l'union des blocs localisés sur $(S_t, S_{t-}]$ pour tout $t \geq 0$.

Cette construction est consistante. Etant donnés $(S_t, t \geq 0)$ et $\Pi^{(n)}$, on peut envoyer le singleton $\{n+1\}$ uniformément sur $(0,1]$ et ensuite suivre les règles de coalescence pour obtenir $\Pi^{(n+1)}$ sans modifier la structure de $\Pi^{(n)}$. Après une infinité d'ajouts de nouveaux individus, on obtient le Λ-coalescent Π.

Soit $0 < s_1 < s_2 < \cdots$ les instants de sauts de $(S_t, t \geq 0)$. Alors

$$(0,1] = \cup_{i\geq 1}(S_{s_i}, S_{s_i-}].$$

Notons que pour obtenir Π, chaque singleton est envoyé uniformément sur $(0,1]$. On fusionne les singletons qui tombent dans le même intervalle. Ceci nous donne une partition aléatoire de $\mathbb{N}$ notée π. Il est clair que π est une partition échangeable. Si on note B_i l'union des singletons qui sont tombés sur $(S_{s_i}, S_{s_i-}]$, alors

$$\pi = \{B_i\}_{i\geq 1}.$$

Soit E_π l'événement où π admet une fréquence asymptotique. Alors

$$\mathbb{P}(E_\pi) = 1.$$

On fixe $t \geq 0$ et soit D_0 le sous-ensemble de $\mathbb{N}$ tel que si $i \in D$, alors $\{i\} \in \Pi(t)$. Soit $\Pi(t) = \{D_0, D_1, D_2, \cdots\}$ où $\{D_i\}_{i\geq 1}$ sont des blocs contenant plusieurs entiers. D'après la construction de Π, sachant E_π, il existe des sous-ensembles disjoints $\{I_i\}_{i\geq 0}$ tels que $\mathbb{N} = \cup_{i\geq 0} I_i$ et

$$D_i = \cup_{j\in I_i} B_j,\ D_0 = \cup_{j\in I_0} B_j.$$

D'après le lemme de Fatou, pour tout $i \geq 1$

$$\liminf_{n\to\infty} \frac{1}{n}\#\{D_i \cap \mathbb{N}_n\} \geq \sum_{j\in I_i} fr(B_j),$$
$$\text{et } \liminf_{n\to\infty} \frac{1}{n}\#\{D_0 \cap \mathbb{N}_n\} \geq \sum_{j\in I_0} fr(B_j).$$

Sachant E_π, on a

$$\sum_{i\geq 1}\sum_{j\in I_i} fr(B_j) + \sum_{j\in I_0} fr(B_j) = 1.$$

On en déduit que

$$fr(D_i) = \sum_{j\in I_i} fr(B_j),\ \forall i \geq 1; fr(D_0) = \sum_{j\in I_0} fr(B_j); fr(D_0) + \sum_{i\geq 1} fr(D_i) = 1.$$

On peut conclure que sachant E_π, $\Pi(t)$ admet une fréquence asymptotique ordonnée pour tout t. □

Remarque 1.7. *Soit Π un Λ-processus avec poussière. Presque sûrement $\Pi(t)$ admet une fréquence asymptotique ordonnée $\Theta(t)$ pour tout $t \geq 0$. Supposons que $\Theta(t) = \{s_1, s_2, \cdots\}$. Alors $1 - \sum_{i\geq 1} s_i$ est la fréquence asymptotique de l'union des singletons au temps t.*

L'inconvénient du processus Θ est que l'on ne voit pas d'une façon explicite comment les blocs coalescent. Dans le cas Kingman, Aldous (1999) trouve une représentation graphique telle que non seulement on voit l'évolution des coalescences mais aussi la masse (ou fréquence asymptotique) de chaque bloc. Une autre représentation graphique de ce genre pour les Λ-coalescents avec $\Lambda(\{0\}) = 0$ est fournie par Bertoin et Le Gall (2003) en utilisant les flots stochastiques de ponts.

1.3 Première application : l'arbre généalogique du modèle de Cannings

Introduisons tout d'abord un modèle d'évolution général (Cannings (1974, 1975)) d'une population de taille fixe N, haploïde, évoluant en générations discrètes (ne se chevauchant pas). Étiquetons les individus au hasard de 1 à N. Soit $r \in \mathbb{Z}$ et $1 \leq i \leq N$. Soit $\Upsilon_i^{(r+1)}$ la taille de la descendance à la génération $r+1$ de l'individu i de la génération r. La taille fixe de la population au cours du temps impose que

$$\Upsilon_1^{(r+1)} + \Upsilon_2^{(r+1)} + \cdots + \Upsilon_N^{(r+1)} = N, r \in \mathbb{Z}.$$

Notons $\Upsilon^r = (\Upsilon_1^{(r)}, \cdots, \Upsilon_N^{(r)})$. Il y a deux propriétés essentielles de $\{\Upsilon^{(r)}\}_{r\in\mathbb{Z}}$ qui caractérisent le modèle de Cannings.

— Homogénéité : les $\{\Upsilon^{(r)}\}_{r\in\mathbb{Z}}$ sont i.i.d de même loi qu'une variable notée $\Upsilon = (\Upsilon_1, \cdots, \Upsilon_N)$.
— Echangeabilité : Soit ω une permutation de $\mathbb{N}_N$, alors $(\Upsilon^{(r)}_{\omega(1)}, \cdots, \Upsilon^{(r)}_{\omega(N)}) \overset{(d)}{=} \Upsilon^{(r)}$, pour tout $r \in \mathbb{Z}$.

Une version simplificatrice du modèle de Cannings est le modèle de Wright-Fisher qui joue un grand rôle dans les travaux de Wright et Fisher dans les années 30. Dans ce modèle, Υ suit la loi binomiale multinomiale : soit $0 \le i_1, i_2, \cdots, i_N \le N, i_1 + i_2 + \cdots + i_N = N$,

$$\mathbb{P}(\Upsilon_1 = i_1, \Upsilon_2 = i_2, \cdots, \Upsilon_N = i_N) = \frac{N!}{i_1! i_2! \cdots i_N!}.$$

On considère la généalogie d'un échantillon de taille n $(2 \le n \le N)$ choisie d'une manière uniforme à une génération donnée dans la population. On souhaite avoir une limite de l'arbre généalogique de cet échantillon à n fixé lorsque N tend vers ∞ sous une échelle de temps appropriée. Numérotons les individus dans l'échantillon de 1 à n. Introduisons la relation d'équivalence suivante : soient $1 \le i, j \le N$. S'ils ont un ancêtre commun r générations auparavant, alors on note

$$i \overset{r}{\sim} j.$$

Nous pouvons alors définir la chaîne de Markov décrite par les classes d'équivalence correspondantes, à valeurs dans $\mathcal{P}_n$, l'ensemble des partitions de $\mathbb{N}_n$. Nous la noterons $(\mathfrak{R}^{(N,n)}_r, r \in \mathbb{N})$. $\mathfrak{R}^{(N,n)}_0$ est la partition triviale ne contenant que des singletons.

Définissons la probabilité c_N de l'événement où i et j ont un ancêtre commun dans la génération précédente.

$$c_N := \mathbb{P}(i \overset{1}{\sim} j) = \mathbb{E}\left[\frac{\sum_{k=1}^N \binom{\Upsilon_k}{2}}{\binom{N}{2})}\right] = \mathbb{E}\left[\frac{\Upsilon_1(\Upsilon_1 - 1)}{N-1}\right].$$

Cette formule montre que c_N ne dépend pas de i, j, en raison de l'échangeabilité des individus. On va donc considérer $1, 2$ au lieu de i, j quelconques. Imposons une hypothèse biologiquement naturelle

$$c_N \to 0 \text{ lorsque } N \to \infty. \tag{1.6}$$

Nous avons alors

$$\mathbb{P}(1 \overset{1}{\sim} 2) = c_N$$

et

$$\mathbb{P}(1 \overset{r}{\sim} 2) = 1 - (1 - c_N)^r \sim 1 - \mathrm{e}^{-rc_N}. \tag{1.7}$$

Pour tout $t \ge 0$, ce calcul mène à une limite non-triviale lorsque $r = \lfloor t/c_N \rfloor$. Ceci nous donne une idée de la renormalisation en temps que nous allons devoir effectuer pour obtenir un processus limite, lequel sera une chaîne de Markov en temps continu à valeurs dans $\mathcal{P}_n$. D'après (1.7), il apparaît que dans ce processus deux lignées fusionnent après un temps exponentiel (de paramètre 1).

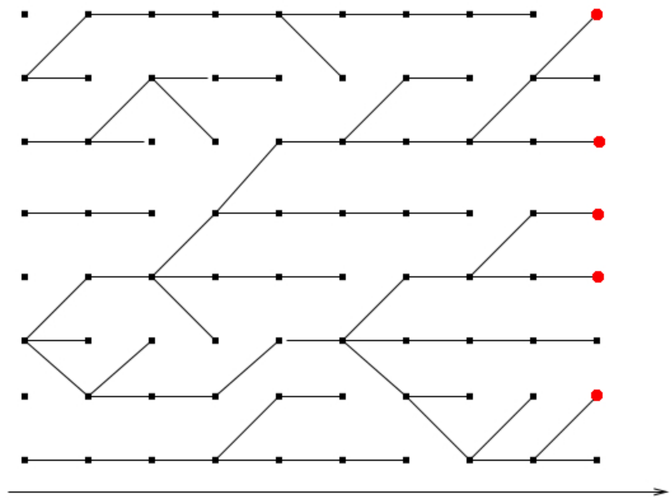


FIGURE 1.7 – Le modèle de Cannings avec $N = 8$. La taille de l'échantillon est 5.

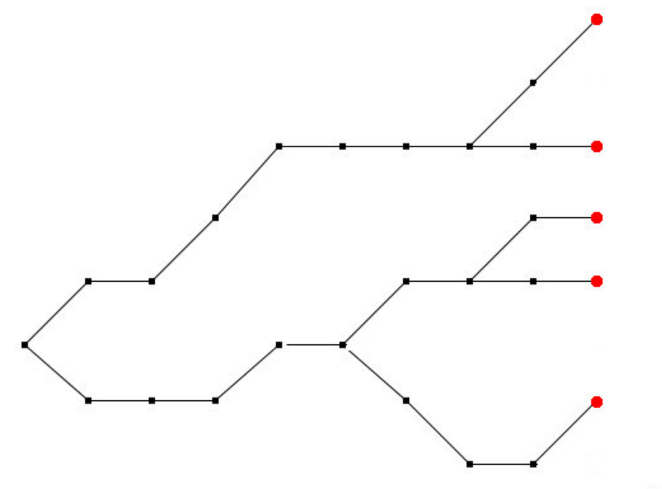


FIGURE 1.8 – Genéalogies des individus de l'échantillon dans la Figure 1.7.

Est-il possible que plus de deux lignées fusionnent au même instant ? Pour répondre à cette question, introduisons d_N, la probabilité que trois individus choisis au hasard aient le même ancêtre à la génération précédente :

$$d_N = \mathbb{P}(1 \overset{1}{\sim} 2 \overset{1}{\sim} 3) = \frac{\sum_{i=1}^N \mathbb{E}\left[\binom{\Upsilon_i}{3}\right]}{\binom{N}{3}} = \frac{\mathbb{E}[\Upsilon_1(\Upsilon_1-1)(\Upsilon_1-2)]}{(N-1)(N-2)}.$$

Nous allons nous placer dans le cas où

$$\frac{d_N}{c_N} \to 0 \text{ lorsque } N \to \infty. \tag{1.8}$$

Remarque 1.8. *Un simple calcul montre que (1.8) implique (1.6) (on peut voir l'argument de Lemme 5.5 de (46) pour plus de détails).*

En utilisant l'échangeabilité, il apparaît que la probabilité que, à la génération où deux lignées fusionnent, une troisième se joigne à l'événement de coalescence est

$$\mathbb{P}(\{1 \overset{1}{\sim} 2\} \cap \{\exists\, 3 \le k \le n,\ k \overset{1}{\sim} 1\}) \le (n-2)\mathbb{P}(1 \overset{1}{\sim} 2 \overset{1}{\sim} 3) = o(\mathbb{P}(1 \overset{1}{\sim} 2).$$

Sous l'hypothèse (1.8), il n'y aura donc pas de collisions multiples dans le processus limite.

Le lecteur ne sera donc pas surpris par le résultat suivant, qui se trouve dans (43).

Théorème 1.4. $(\mathfrak{R}^{(N,n)}_{\lfloor c_N^{-1} t\rfloor}, t \ge 0)$ *converge au sens des lois fini-dimensionnelles vers le* n*-coalescent de Kingman* $(\Pi^{(n)}_t, t \ge 0)$ *si et seulement si* (1.8) *est vérifiée.*

Ce résultat généralise celui de (39) où l'on a besoin que $\sup_{N\ge 2} \mathbb{E}[(\Upsilon_1)^k] < \infty$ pour tout $k \ge 1$. (41) obtient une condition suffisante pour avoir une convergence étroite de $(\mathfrak{R}^{(N,n)}_{\lfloor c_N^{-1} t\rfloor}, t \ge 0)$ vers le n-coalescent de Kingman dans l'espace $\mathcal{D}(\mathbb{R}_+, \mathcal{P}^n)$.

Que se passe-t-il lorsque (1.8) n'est pas satisfaite ? Dans ce cas, la probabilité d'avoir une coalescence binaire n'est pas significativement plus grande que les probabilités de coalescences multiples. On peut espérer avoir d'autres processus limites qui admettent les coalescences multiples. (50) définit le processus de Λ-coalescent et propose trois conditions suffisantes pour que $(\mathfrak{R}^{(N,n)}_{\lfloor c_N^{-1} t\rfloor}, t \ge 0)$ converge au sens des lois fini-dimensionnelles vers un Λ n-coalescent. En particulier, le facteur de normalisation du temps est toujours c_N. (45) élargissent la classe des processus limites en incluant le Ξ-coalescent, dont l'on ne parlera pas ici.

1.4 Deuxième application : test de neutralité de Fu et Li

L'information génétique est contenue dans le génome qui est l'ensemble du matériel génétique d'un individu ou d'une espèce codé dans son ADN (à l'exception de certains virus dont le génome est porté par des molécules d'ARN). Les éléments de base codants sont des acides nucléiques ne contenant que quatre types représentés par quatre lettres A, T, G, C. Un acide nucléique est un polymère dont l'unité de base est le nucléotide qui s'occupe de l'information génétique.

Une mutation est une modification de l'information génétique dans le génome d'une cellule ou d'un virus. Sans mutations, il ne peut y avoir de nouveaux traits dans la population. Il y a plusieurs types de mutations mais, pour simplifier, nous ne considérerons que des mutations ponctuelles, affectant une base de nucléotide.

La variation génétique est essentiellement agrandie par les *mutations* qui interviennent dans la population (il faut aussi compter les recombinaisons, les immigrations, les dérives génétiques). Les taux de mutations varient suivant les espèces et l'emplacement des bases touchées. Ils sont assez faibles. Les plus élevés sont de l'ordre de 10^{-4} mutations par paire de bases par génération dans l'ARN de certains virus et l'on observe des taux de l'ordre de 10^{-10} chez les humains (nous invitons le lecteur à se reporter au Chapitre 12 de (2) pour plus d'informations).

Pour ajouter des mutations dans un modèle de Cannings à N individus, nous supposons que la probabilité de muter pour chaque individu (ou chromosome, gène, nucléotide, etc) à chaque génération est constante égale à μ. En remontant une lignée ancestrale, le nombre de générations à attendre avant de voir une mutation est donc géométrique de paramètre μ. Rappelons que c_N désigne la probabilité que deux individus donnés aient un ancêtre commun à la génération précédente. En supposant que

$$\frac{\mu}{c_N} \to \theta, \theta > 0, \text{ lorsque } N \to \infty$$

et en renormalisant le temps de manière à obtenir un coalescent de Kingman, le temps d'attente pour observer une mutation sur chaque lignée est exponentiel de paramètre θ. L'objet limite peut être obtenu en considérant un processus ponctuel de Poisson d'intensité θdl sur les branches de l'arbre de Kingman.

L'indépendance entre l'arbre de généalogie et les arrivées des mutations implique que les mutations n'ont pas d'influences sur l'évolution de la population et l'évolution des mutés est soumise à la dérive génétique. Ce mécanisme de production de mutations est remis en cause par ceux qui croient que l'effet de la sélection naturelle est plus important que celui de la dérive génétique. Il existe beaucoup de tests de neutralité pour le coalescent de Kingman. On peut citer le test de (23), de (55), de (54) et aussi de (29). En particulier, le test de (29), qui s'appuie sur les longueurs des branches du n-coalescent de Kingman, a beaucoup motivé cette thèse pour étudier ces longueurs des autres Λ n-coalescents.

Dans un Λ n-coalescent $\Pi^{(n)}$, l'individu i est attaché à $\Pi^{(n)}$ par une branche *externe* dont la longueur est notée par $T_i^{(n)}$ pour $1 \leq i \leq n$. Il existe donc n branches externes. Les autres branches sont *internes*. Les mutations arrivées sur les branches externes sont dites *mutations externes* et les mutations sur les branches internes sont dites *mutations internes*. Supposons que les mutations arrivent sur les branches au taux θ. Soit

— $L_{ext}^{(n)}$: la longueur totale des branches externes.
— $\eta_e^{(n)}$: le nombre total de mutations externes.
— $L_i^{(n)}$: la longueur totale des branches internes.
— $\eta_i^{(n)}$: le nombre total de mutations internes.
— $L_{total}^{(n)}$: la longueur totale des branches .
— $\eta^{(n)}$: le nombre total de mutations.

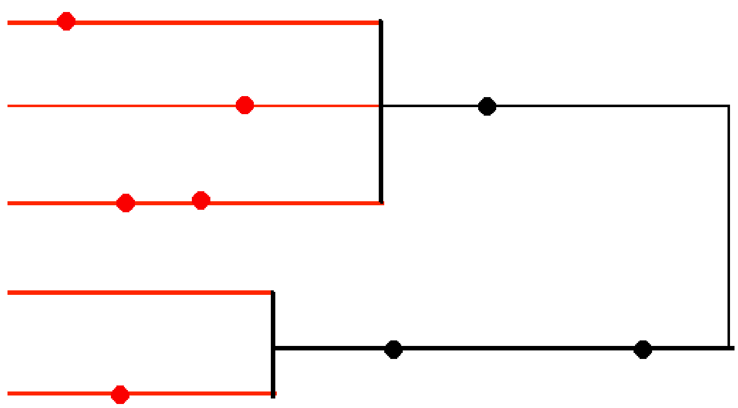


FIGURE 1.9 – les branches rouges sont des branches externes, sur lesquelles les points ronds sont des mutations externes. Les branches noires horizontales sont des branches internes, sur lesquelles les points ronds sont des mutations externes.

Fu et Li argumentent que la plupart des mutations externes sont récentes et la plupart des mutations internes sont vieilles. Si les vieilles mutations sont principalement nuisibles, alors on voit plus de mutations externes dans la généalogie, car les individus avec beaucoup de vieilles mutations nuisibles ont très peu de descendants. S'il y a une mutation qui est fixée récemment, alors on voit encore plus de mutations externes, car la plupart des mutations sont récentes (les vieilles mutations sont expulsées par la mutation fixée). D'autre part, si la sélection qui favorise les hétérozygotes d'allèles que les homozygotes, alors on pourrait avoir moins de mutations externes. Donc il est raisonnable de comparer les nombres de mutations externes et internes pour détecter la neutralité du modèle.

Un simple calcul sur le n-coalescent de Kingman donne (voir (29)) :

$$\mathbb{E}[L^{(n)}_{ext}] = 2,\ \mathbb{E}[L^{(n)}] = 2\sum_{k=1}^{n-1}\frac{1}{k}.$$

On note $a_n = \sum_{k=1}^{n-1}\frac{1}{k}$. Le test statistique proposé par Fu et Li est :

$$D_n = \frac{\eta^n - a_n\eta^n_e}{\sqrt{\mathrm{Var}\{\eta^n - a_n\eta^n_e\}}}.$$

En pratique, il faut tout d'abord estimer le taux θ. La loi limite de D_n est simulée par Fu et Li . Par exemple, on a

$$\mathbb{P}(D_n \leq -1.87) \to 0.01, n \to \infty.$$

Donc pour n assez grand, si D_n est inférieur à -1.87, alors on a qu'avec probabilité 0.99 le modèle n'est pas neutre.

On constate que a_n est un paramètre important dans D et aussi que

$$a_n = \frac{\mathbb{E}[L^{(n)}]}{\mathbb{E}[L^{(n)}_{ext}]}.$$

Les généalogies de certaines populations n'entrent plus dans le cadre du coalescent de Kingman. Elles sont mieux modélisées par les Λ-coalescents, voire les Ξ-coalescents. On peut appliquer directement le test de Fu et Li sans aucune difficulté, à condition de connaître le comportement limite de $L^{(n)}$ et de $L^{(n)}_{ext}$. On remarque aussi que la loi de $L^{(n)}_{ext}$ est étroitement liée à celle de la longueur d'une branche externe $T^{(n)}_i$. Ma thèse a pour but de caractériser les longueurs des branches de certains Λ n-coalescents.

1.5 Présentation des résultats

1.5.1 Premier article : sur la longueur d'une branche externe du Beta-coalescent

On the length of an external branch in the beta-coalescent.

Cet article a été écrit en collaboration avec Jean-Stéphane Dhersin, Fabian Freund, Arno Siri-Jégousse, et publié en 2013 dans *Stochastic Processes and their Applications.*

On s'intéresse à la loi limite de $T^{(n)}$, qui est la longueur d'une branche externe choisie au hasard, pour certains Λ n-coalescents. Par l'échangeabilité des singletons, on a

$$T^{(n)}_i \overset{(d)}{=} T^{(n)},\ 1 \leq i \leq n.$$

Donc il suffit de regarder $T^{(n)}_1$, appelée *l'unicité de l'individu 1* ((49)), qui n'appartient qu'à l'individu 1 dans l'arbre et qui sert à mesurer la variation génétique. La loi limite de $T_1^{(n)}$ est pour la première fois abordée par (10) dans le cas du coalescent de Kingman. (13) montrent ensuite que

$$nT_1^{(n)} \stackrel{(d)}{\to} T_{kin}, \text{ lorsque } n \to \infty,$$

où T_{kin} est une variable de densité $\frac{8}{(2+t)^3}\mathbf{1}_{t\geq 0}$. (31) trouvent la loi limite de $T_1^{(n)}$ lorsque $\int_0^1 x^{-2}\Lambda(dx) < \infty$. (48) et (44) étudient le cas plus général où $\int_0^1 x^{-1}\Lambda(dx) < \infty$, qui correspondent aux Λ-coalescents avec poussière. Plus précisément,

$$\left(\int_0^1 x^{-1}\Lambda(dx)\right) T_1^{(n)} \stackrel{(d)}{\to} Exp(1), \text{ lorsque } n \to \infty.$$

Remarquons que les Beta$(2-\alpha,\alpha)$-coalescents avec $0 < \alpha < 1$ sont bien des coalescents avec poussière. (27) montrent que dans le cas où $\alpha = 1$, i.e., le coalescent de Bolthausen-Sznitman, on a

$$(\ln n)T_1^{(n)} \stackrel{(d)}{\to} Exp(1), \text{ lorsque } n \to \infty.$$

La contribution majeure de cet article est de résoudre ce problème dans le cas des Beta$(2-\alpha,\alpha)$-coalescents avec $1 < \alpha < 2$. On se place en fait dans un cadre plus général. Définissons

$$\rho(t) = \int_t^1 x^{-2}\Lambda(dx), 0 < t \leq 1,$$

et supposons qu'il existe $C_0 > 0, 1 < \alpha < 2, \zeta > 1 - 1/\alpha$ tels que

$$\rho(t) = C_0 t^{-\alpha} + o(t^{-\alpha+\zeta}), \text{ lorsque } t \to 0+. \tag{1.9}$$

Pour les Beta$(2-\alpha,\alpha)$-coalescents avec $1 < \alpha < 2$, on a $C_0 = \frac{1}{\Gamma(\alpha+1)\Gamma(2-\alpha)}$ et $\zeta = 1$.

Le théorème principal est le suivant :

Théorème 1.5. *On a la convergence suivante :*

$$n^{\alpha-1}T_1^{(n)} \stackrel{(d)}{\to} T, \text{ lorsque } n \to \infty,$$

où T est une variable à valeurs positives. La densité de T est donnée par

$$f_T(t) = \frac{\alpha C_0 \Gamma(2-\alpha)}{\alpha - 1}(1 + C_0\Gamma(2-\alpha)t)^{-\frac{\alpha}{\alpha-1}-1}, \quad t \geq 0.$$

En particulier, dans le cas du Beta$(2-\alpha,\alpha)$-coalescent avec $1 < \alpha < 2$, la densité est

$$f_T(t) = \frac{1}{(\alpha-1)\Gamma(\alpha)}(1 + \frac{t}{\alpha\Gamma(\alpha)})^{-\frac{\alpha}{\alpha-1}-1}, \quad t \geq 0. \tag{1.10}$$

On présente deux méthodes dans cet article. La première est de considérer d'abord $\sigma_1^{(n)}$, le nombre de sauts jusqu'à la coalescence de $\{1\}$. Une fois $\sigma_1^{(n)}$ connue, il suffit de sommer les temps écoulés entre les sauts. En fait, $\sigma_1^{(n)}$ a sa propre importance en biologie. Soit $\tau^{(n)}$ le nombre total de sauts du processus. Alors la valeur $\frac{\sigma_1^{(n)}}{\tau^{(n)}}$ donne la position de la coalescence de $\{1\}$ parmi toutes les coalescences. Dans le cas de Kingman (voir (13)), on a

$$\frac{\sigma_1^{(n)}}{n} \stackrel{(d)}{\to} Beta(1,2), \text{ lorsque } n \to \infty.$$

Dans le cas du coalescent de Bolthausen-Sznitman, (27) montrent que

$$\frac{(\ln n)\sigma_1^{(n)}}{n} \stackrel{(d)}{\to} Beta(1,1), \text{ lorsque } n \to \infty.$$

Définissons $\mu_{-1} := \int_0^1 x^{-1}\Lambda(dx), \mu_{-2} := \int_0^1 x^{-2}\Lambda(dx)$ et supposons que $\mu_{-2} < \infty$, alors (31) montrent que

$$\sigma_1^{(n)} \stackrel{(d)}{\to} Geo(\frac{\mu_{-1}}{\mu_{-2}}), \text{ lorsque } n \to \infty,$$

où $Geo(\cdot)$ désigne une variable géométrique.

Dans notre cadre, on obtient

Proposition 1.1. *Supposons que Λ satisfait (1.9), alors*

$$\frac{\sigma_1^{(n)}}{n(\alpha-1)} \stackrel{(d)}{\to} Beta(1,\alpha), \text{ lorsque } n \to \infty.$$

Ma contribution est d'introduire une construction du Λ n-coalescent et une méthode reposant sur cette construction pour montrer le Théorème 1.5. Cette construction est appelée *construction récursive du Λ n-coalescent*. Le processus $\Pi^{(n)}$ est construit pour n individus de 1 à n. La consistance du processus donne que la restriction de $\Pi^{(n)}$ à $\mathbb{N}_{n-1}$ est $\Pi^{(n-1)}$. Donc une question naturelle est de se demander comment construire $\Pi^{(n)}$ sachant $\Pi^{(n-1)}$. Pour cela, il faut savoir où la branche externe de $\{n\}$ s'attache à $\Pi^{(n-1)}$.

On va un peu modifier cette question. Soit $\Pi^{(2,n)}$ le processus construit à partir des individus de 2 à n. On va attacher $\{1\}$ pour construire $\Pi^{(n)}$. La modification est due au fait que l'on est intéressé par le comportement de $\{1\}$. L'avantage de cette construction est que lorsque l'on attache $\{1\}$ à $\Pi^{(2,n)}$, on obtient directement la longueur de la branche externe de $\{1\}$. En adaptant cette construction, il s'avère que la longueur de la branche externe de $\{1\}$ suit la même loi que l'instant du premier saut d'un processus de Cox qui est caractérisé par $\Pi^{(2,n)}$. Le résultat essentiel est le suivant :

Proposition 1.2. *Supposons que Λ satisfait (1.9), alors*

$$\mathbb{P}(n^{\alpha-1}T_1^{(n)} \geq t|\Pi^{(2,n)}) \xrightarrow{P} (1 + C_0\Gamma(2-\alpha)t)^{-\alpha/(\alpha-1)}, \ \textit{lorsque} \ n \to \infty.$$

Ce résultat est plus fort que le Théorème 1.5, car ici on voit l'indépendance asymptotique entre $n^{\alpha-1}T_1^{(n)}$ et $\Pi^{(2,n)}$. Un autre résultat intéressant concerne le comportement asymptotique du processus de comptage de blocs $K^{(n)}$:

Proposition 1.3. *Si Λ satisfait (1.9), alors pour tout $\varepsilon > 0$, on a*

$$\mathbb{P}(\sup_{t\geq 0} |n^{-1}K^{(n)}(tn^{1-\alpha}) - (1 + C_0\Gamma(2-\alpha)t)^{-1/(\alpha-1)}| > \varepsilon) \to 0, \ \textit{lorsque} \ n \to \infty. \quad (1.11)$$

Remarquons que ce résultat est plus fort que celui contenu dans l'article. Il suffit d'utiliser le fait que $K^{(n)}$ est décroissante et la courbe limite est aussi décroissante et tendant vers 0.

1.5.2 Deuxième article : sur la longueur totale des branches externes du Beta-coalescent

On the total length of external branches for Beta-coalescent.

Cet article a été écrit avec Jean-Stéphane Dhersin et soumis dans *Journal of Applied Probability*. On étudie les moments de $L_{ext}^{(n)}$, la longueur totale de branches externes dans le cas des Beta$(2-\alpha, \alpha)$-coalescents avec $1 < \alpha < 2$. Une version pour une classe plus générale de Λ-coalescents mais qui demande de beaucoup plus de calculs se trouve sur Arxiv ((16)). Dans le cas Kingman, (35) montrent que

$$\frac{1}{2}\sqrt{\frac{n}{\ln n}}(L_{ext}^{(n)} - 2) \xrightarrow{(d)} N(0,1).$$

Pour le coalescent de Bolthausen-Sznitman, (36) prouvent que

$$\frac{(\ln n)^2}{n}L_{ext}^{(n)} - \ln n - \ln\ln n \xrightarrow{(d)} Z - 1,$$

où Z est une variable stable d'indice 1. Plus précisément,

$$\mathbb{E}[e^{i\theta Z}] = exp(-\frac{\pi}{2}|\theta| + i\theta\ln|\theta|), \theta \in \mathbb{R}.$$

Dans le cas des Beta$(2-\alpha, \alpha)$-coalescents avec $1 < \alpha < 2$, la convergence suivante est prouvée par (4), (7) et (5) :

$$n^{\alpha-2}L_{ext}^{(n)} \xrightarrow{p.s,P} \alpha(\alpha-1)\Gamma(\alpha).$$

La convergence presque sûre est due au fait que les $\{\Pi^{(n)}, n \geq 2\}$ sont des restrictions de Π sur $\mathbb{N}_n$. (37) ont beaucoup avancé sur ce problème. Ils prouvent que

$$\frac{L_{ext}^{(n)} - \alpha(\alpha-1)\Gamma(\alpha)n^{2-\alpha}}{n^{1/\alpha+1-\alpha}} \xrightarrow{(d)} \frac{\alpha(2-\alpha)(\alpha-1)^{1/\alpha+1}\Gamma(\alpha)}{\Gamma(2-\alpha)^{1/\alpha}}\zeta, \tag{1.12}$$

où ζ est la variable stable d'indice α telle que $\mathbb{E}[\zeta] = 0$, et pour x tendant vers ∞, $\mathbb{P}(\zeta > x) = o(x^{-\alpha}), \mathbb{P}(\zeta < -x) \sim x^{-\alpha}$.

Cet article a aussi pour but d'établir un théorème limite pour $L_{ext}^{(n)}$. L'idée est d'étudier les comportements asymptotiques du moment d'ordre 2 de $L_{ext}^{(n)}$ et de trouver le facteur de normalisation qui pourrait être utile dans le théorème limite. Le résultat principal est le suivant :

Théorème 1.6. — *Dans le cas des Beta*$(2-\alpha, \alpha)$*-coalescents avec* $1 < \alpha < 2$*, on a*

$$\lim_{n\to\infty} n^{3\alpha-5}\mathbb{E}[(L_{ext}^{(n)} - \alpha(\alpha-1)\Gamma(\alpha)n^{2-\alpha})^2] = \Delta(\alpha)/2,$$

où

$$\begin{aligned}\Delta(\alpha) &= \frac{((\alpha-1)\Gamma(\alpha+1))^3}{3-\alpha}\int_0^1 ((1-x)^{2-\alpha}-1)^2\nu(dx) \\ &= \frac{((\alpha-1)\Gamma(\alpha+1))^2\Gamma(4-\alpha)}{(3-\alpha)\Gamma(4-2\alpha)}.\end{aligned} \tag{1.13}$$

— *Par conséquent,* $n^{\alpha-2}L_{ext}^{(n)}$ *converge dans* L^2 *vers* $\alpha(\alpha-1)\Gamma(\alpha)$.

Le terme $\int_0^1((1-x)^{2-\alpha}-1)^2\nu(dx)$ se retrouve aussi chez les processus plus généraux (voir (16)). A partir de ce résultat, on pourrait espérer que le terme suivant

$$n^{(3\alpha-5)/2}\left(L_{ext}^{(n)} - \alpha(\alpha-1)\Gamma(\alpha)n^{2-\alpha}\right)$$

converge en distribution vers une variable limite. Mais (1.12) montre que cela est faux. Cela souligne la limite de la méthode des moments.

Le Théorème 1.6 est fondé sur les deux résultats suivants.

Proposition 1.4. *Dans le cas des Beta*$(2-\alpha, \alpha)$*-coalescents avec* $1 < \alpha < 2$*, on a la covariance asymptotique entre deux longueurs de branches externes :*

$$\lim_{n\to\infty} n^{3(\alpha-1)}Cov(T_1^{(n)}, T_2^{(n)}) = \Delta(\alpha).$$

Proposition 1.5. *Dans le cas des Beta*$(2-\alpha, \alpha)$*-coalescents avec* $1 < \alpha < 2$*, on a*

1. *Si* $0 \leq \beta < \frac{\alpha}{\alpha-1}$*, alors* $\lim_{n\to\infty}\mathbb{E}[(n^{\alpha-1}T_1^{(n)})^\beta] = \mathbb{E}[T^\beta] < \infty$.
2. *Si* $\beta \geq \frac{\alpha}{\alpha-1}$*, alors* $\lim_{n\to\infty}\mathbb{E}[(n^{\alpha-1}T_1^{(n)})^\beta] = \infty$.

1.5.3 Troisième article : sur la construction par division de mesure de Λ-coalescents

On the measure division construction of Λ-coalescents.

Cet article est soumis à *Markov Processes and Related Fields.*

Le but de cet article est d'introduire une nouvelle construction de Λ n-coalescents appelée construction par division de mesure (*measure division construction*) pour étudier le cas où une « bonne » mesure Λ est perturbée par un bruit. De nombreux résultats sont connus pour certains Λ-coalescents tels que les Beta-coalescents, le coalescent de Kingman, etc. Mais que se passe-t-il si l'on considère le cas d'une mesure Λ' qui est proche d'une bonne mesure Λ ?

Supposons que $\Lambda = \Lambda_1 + \Lambda_2$ où Λ_1 et Λ_2 sont deux mesures finies sur $[0,1]$. Soit $\Pi^{(1,n)}$ le Λ_1 n-coalescent et $\Pi^{(2,n)}$ le Λ_2 n-coalescent. La construction par division de mesure consiste à prendre une réalisation de $\Pi^{(1,n)}$, et la modifier en utilisant Λ_2 pour obtenir finalement $\Pi^{(n)}$. Si Λ_1 est « petite », on peut montrer que $\Pi^{(n)}$ et $\Pi^{(2,n)}$ sont proches, si l'on compare certaines quantités. Pour tout $r > 0$, on note $Exp(r)$ la variable exponentielle de paramètre r. On a le théorème suivant :

Théorème 1.7. *Soit* $\mu^{(n)} = \int_{1/n}^{1} x^{-1}\Lambda(dx)$. *Si*

$$\lim_{n\to\infty} \frac{g_n}{n\mu^{(n)}} = 0, \tag{1.14}$$

alors $\mu^{(n)}T^{(n)}$ *converge en loi vers* $Exp(1)$.

La classe des Λ n-coalescents satisfaisant (1.14) est assez large. Les exemples classiques sont les suivants :

1. $\int_0^1 x^{-1}\Lambda(dx) < \infty$: Les coalescents avec poussière.
2. Le n-coalescent de Bolthausen-Sznitman.
3. Λ a une densité f_Λ sur $[0,r)$ où $0 < r < 1$ et il existe un nombre positif M tel que $f_\Lambda < M$ sur $[0,r)$.
4. Λ a une densité $f_\Lambda(x) = p(\ln \frac{1}{x})^q$ sur $[0,r)$ où $0 < r < 1$ et p,q sont des constantes positives.

Rappelons que le résultat dans le premier exemple se trouve dans (44) et le deuxième exemple dans (27). Notre article recouvre ces résultats à l'aide d'une méthode unique.

Bien évidemment, (1.14) ne couvre pas tous les Λ n-coalescents classiques. Par exemple, le n-coalescent de Kingman et les Beta$(2-\alpha,\alpha)$ n-coalescents avec $1 < \alpha < 2$ n'y sont pas compris. Dans le cas Kingman, le terme $\frac{g_n}{n\mu^{(n)}}$ est non défini. Dans le cas Beta avec $1 < \alpha < 2$,

$$\frac{g_n}{n\mu^{(n)}} \longrightarrow (\alpha-1)\Gamma(2-\alpha)/\alpha.$$

Vue que l'on connait bien la loi limite de $n^{\alpha-1}T^{(n)}$ dans ce cas Beta, une conjecture peut être établie :

Conjecture $(*)$: *Soit $c > 0$. Si*

$$\lim_{n\to\infty} \frac{g_n}{n\mu^{(n)}} = c,$$

alors $\mu^{(n)}T^{(n)} \stackrel{(d)}{\to} T_c$, où T_c est une variable aléatoire avec la densité

$$\Gamma(2-\alpha^*)(1+cx)^{-\frac{\alpha^*}{\alpha^*-1}-1}\mathbf{1}_{x\geq 0}.$$

Ici α^* est la solution unique de l'équation suivante

$$\frac{(\alpha-1)\Gamma(2-\alpha)}{\alpha} = c.$$

Remarque 1.9. *Cette conjecture couvre plus généralement les Λ-coalescents satisfaisant (1.9).*

Remarque 1.10. *Le Théorème 1.7 et la conjecture $(*)$ donnent l'idée que la longueur d'une branche externe est de l'ordre $1/\mu^{(n)}$ pour tous les Λ-coalescents classiques sauf le coalescent de Kingman. Mais on peut argumenter que le coalescent de Kingman suit aussi cette règle. Le n-coalescent de Kingman est la limite au sens de convergence fini-dimensionnelle des Beta$(2-\alpha,\alpha)$ n-coalescents lorsque α tend vers 2. La longueur d'une branche externe dans le cas Beta quand $1 < \alpha < 2$ est de l'ordre $n^{1-\alpha}$. Donc il n'est pas étonnant que la longueur d'une branche externe dans le Kingman soit de l'ordre n^{-1}. Par conséquent, $\mu^{(n)}$ dans le cas Kingman peut être considéré comme n, et ainsi que le coalescent de Kingman satisfait cette règle universelle.*

On est ensuite intéressé par la loi jointe des longueurs de plusieurs branches externes.

Théorème 1.8. *Si Λ satisfait la condition (1.14) et $\int_0^1 x^{-1}\Lambda(dx) = \infty$, alors :*

$$\mu^{(n)}(T_1^{(n)}, T_2^{(n)}, \cdots, T_n^{(n)}, 0, 0, \cdots) \stackrel{(d)}{\to} (e_1, e_2, \cdots), \tag{1.15}$$

où $(e_i)_{i\in\mathbb{N}}$ sont indépendants et de loi exponentielle de paramètre 1.

Remarque 1.11. *Le même résultat a été trouvé par (15) dans le cas du coalescent de Bolthausen-Sznitman. Le condition $\int_0^1 x^{-1}\Lambda(dx) = \infty$ est indispensable pour que (1.15) soit vraie (voir (44) pour le cas contraire).*

Les trois corollaires suivants sont aussi vrais pour le coalescent de Bolthausen-Sznitman (voir (15), (19), (30))

Corollaire 1.2. *Si Λ satisfait la condition* (1.14)*, alors pour tout $r \in \mathbb{R}^+$,*

$$\lim_{n\to\infty} \mathbb{E}[(\mu^{(n)}T^{(n)})^r] = \mathbb{E}[e_1^r],$$

où e_1 suit une loi exponentielle de paramètre 1*. De plus, si $\int_0^1 x^{-1}\Lambda(dx) = \infty$, alors pour tout $k \in \mathbb{N}$ et $(r_1, r_2, \cdots, r_k) \in \{\mathbb{R}^+\}^k$, on a :*

$$\lim_{n\to\infty} \mathbb{E}[\prod_{i=1}^{k}(\mu^{(n)}T_i^{(n)})^{r_i}] = \mathbb{E}[\prod_{i=1}^{k} e_i^{r_i}], \tag{1.16}$$

où $(e_i)_{i\in\mathbb{N}}$ sont indépendants et de loi exponentielle de paramètre 1*.*

Corollaire 1.3. *Si Λ satisfait la condition* (1.14) *et $\int_0^1 x^{-1}\Lambda(dx) = \infty$, alors la longueur totale des branches externes $L_{ext}^{(n)}$ satisfait : $\mu^{(n)}L_{ext}^{(n)}/n$ converge dans L^2 vers* 1*.*

Corollaire 1.4. *Si Λ satisfait la condition* (1.14) *et $\int_0^1 x^{-1}\Lambda(dx) = \infty$, alors la longueur totale des branches $L_{total}^{(n)}$ satisfait : $\mu^{(n)}L_{total}^{(n)}/n$ converge en probabilité vers* 1*.*

Remarque 1.12. *Si $\int_0^1 x^{-1}\Lambda(dx) < \infty$, alors (1.16) et les Corollaires 1.3, 1.4 ne sont plus vrais (voir encore (44)).*

En conclusion, cet article étudie les différentes longueurs du Λ n-coalescent satisfaisant (1.14) et $\int_0^1 x^{-1}\Lambda(dx) = \infty$. La forme du processus est à peu près claire sachant les lois limites de $T^{(n)}, L_{ext}^{(n)}$ et $L_{total}^{(n)}$. Ce qui est faisable par la suite est de vérifier la conjecture. Si la conjecture est vraie, alors on a une classe bien plus générale que les Beta$(2-\alpha, \alpha)$-coalescents : il suffit de regarder les différentes limites de $g_n/(n\mu^{(n)})$.

1.5.4 Quatrième article : résultats asymptotiques sur la taille du clade minimal et sur des fonctionelles liées au Beta-coalescent

Asymptotics of the minimal clade size and related functionals of certain Beta-coalescents.

Cet article est en collaboration avec Arno Siri-Jégousse. On peut considérer l'article comme une extension au cas fini de (6).

Considérons les Beta$(2-\alpha, \alpha)$-coalescents avec $1 < \alpha < 2$. (6) ont étudié le processus de fréquence asymptotique ordonnée Θ qui pourrait générer la loi de $\Pi(t)$ et $\Pi^{(n)}(t)$ pour tout $t \geq 0$ en utilisant la procédure de la boîte de penture de Kingman. On se réfère à la section 0.2.3 pour plus de détails.

Citons une version un peu modifiée du Théorème 1.4 de (6) qui est essentiel dans notre article.

Théorème 1.9. *(Théorème 1.4 de (6)) Soit μ la distribution de Slack (voir (53)) de transformée de Laplace*

$$\mathcal{L}_\mu(\lambda) := \int_0^\infty e^{-\lambda x}\mu(dx) = 1 - (1+\lambda^{1-\alpha})^{-\alpha/(\alpha-1)},\ \lambda > 0.$$

Soit $t \geq 0, x \geq 0$. On note $F(x) = \mu((0,x])$. Soit $K(t,x)$ le nombre d'éléments de $\Theta(t)$ qui ont des fréquences asymptotiques au plus x. Alors dans le cas des Beta($2-\alpha,\alpha$)-coalescents avec $1 < \alpha < 2$,

$$\lim_{t\downarrow 0}\sup_{x\geq 0}|t^{1/(\alpha-1)}K(t,t^{1/(\alpha-1)}x) - (\alpha\Gamma(\alpha))^{1/(\alpha-1)}F((\alpha\Gamma(\alpha))^{1/(\alpha-1)}x)| = 0, p.s.$$

Si on connait la valeur de $K(t,x)$ pour tout $x \geq 0$, alors $\Theta(t)$ est déterminée. Ce théorème caractérise donc bien le comportement asymptotique de Θ. Pour donner une application, si on veut savoir la loi limite de $n^{\alpha-1}T_1^{(n)}$, il suffit de calculer la limite de la probabilité

$$\mathbb{P}(T_1^{(n)} > n^{1-\alpha}t),\ t \geq 0.$$

En utilisant la procédure de la boîte de peinture, on jette les particules de 1 à n uniformément sur $[0,1]$ qui est divisée en $|\Theta(n^{1-\alpha}t)|$ cases dont les longueurs correspondent bijectivement aux valeurs de tous les éléments de $\Theta(n^{1-\alpha}t)$. Alors l'événement où la particule 1 se trouve toute seule dans une case est équivalent à $\{T_1^{(n)} > n^{1-\alpha}t\}$. Mais la limite de la probabilité du premier événement est plus facile à trouver en appliquant le Théorème 1.9. Plus généralement, on a

Théorème 1.10. *Pour tout $k \geq 1$*

$$n^{\alpha-1}(T_1^{(n)},\cdots,T_k^{(n)}) \xrightarrow{(d)} (T_1,\cdots,T_k), \tag{1.17}$$

où $(T_1,\cdots,T_k)$ sont des copies i.i.d de T qui est une variable de densité donnée par (1.10).

On s'intéresse ensuite à la quantité $Q^{(n)}$, le nombre de blocs impliqués dans la coalescence de $\{1\}$. Cette quantité mesure la fécondité de la population.

Théorème 1.11. *$Q^{(n)}$ converge en loi vers une variable aléatoire Q telle que pour tout $k \geq 2$*

$$q_k := \mathbb{P}(Q = k) = \frac{(\alpha-1)\Gamma(k-\alpha)}{\Gamma(k)\Gamma(2-\alpha)}. \tag{1.18}$$

La preuve de cette proposition est simple. Il suffit de savoir le nombre total de blocs juste avant la coalescence de $\{1\}$. Ensuite on utilise la propriété de Markov fort de $\Pi^{(n)}$ pour calculer le nombre de blocs qui serons impliqués immédiatement.

Remarquons que dans le coalescent de Kingman, $Q^{(n)} = Q = 2$ presque surement. La finitude de la loi limite n'est pas toujours vraie pour les Λ-coalescents.

Proposition 1.6. *Considérons un Λ n-coalescent dont la measure caractéristique Λ satisfait $\int_0^1 x^{-1}\Lambda(dx) < \infty$. Alors $Q^{(n)} \xrightarrow{(d)} \infty$.*

Ce résultat est même vrai pour le n-coalescent Bolthausen-Sznitman (voir la Remarque 3.2 de l'article).

Le résultat suivant concerne la taille du clade minimal notée $Y^{(n)}$. La définition de $Y^{(n)}$ est très liée à $Q^{(n)}$. $Y^{(n)}$ est la taille du bloc contenant 1 au moment de la coalescence de $\{1\}$. Autrement dit, $Y^{(n)}$ est la somme des tailles de $Q^{(n)}$ blocs à la coalescence de $\{1\}$. La notion du clade minimal est introduite par (10) dans le cas Kingman. Ils montrent que

$$\mathbb{P}(Y^{(n)} = k) = \frac{4}{(k+1)k(k-1)}, k = 2, \cdots, n-1;\ \mathbb{P}(Y^{(n)} = n) = \frac{2}{n(n-1)}.$$

(28) étudient le cas du coalescent de Bolthausen-Sznitman. Ils trouvent le résultat suivant :

$$\frac{\ln Y^{(n)}}{\ln n} \xrightarrow{(d)} U_{[0,1]},$$

où $U_{[0,1]}$ est une variable uniforme sur $[0,1]$. Dans notre cas, on a

Théorème 1.12. *Soient $t \geq 0, k \geq 1$. Définissons un processus stochastique $(\beta(t), t \geq 0)$ tel que*

$$\mathbb{P}(\beta(t) = k) := \frac{1}{\Gamma(k)} \left(\frac{t}{\Gamma(\alpha+1)}\right)^{\frac{k-1}{\alpha-1}} \int_0^\infty e^{-x\left(\frac{t}{\Gamma(\alpha+1)}\right)^{\frac{1}{\alpha-1}}} x^k \mu(dx). \tag{1.19}$$

Soient $(\beta_i(t), t \geq 0)_{i\geq 1}$ les copies i.i.d de $(\beta(t), t \geq 0)$. Supposons que $Q, T, (\beta_i(t), t \geq 0)_{i\geq 1}$ sont indépendantes. Alors

$$Y^{(n)} \xrightarrow{(d)} Y = 1 + \sum_{i=1}^{Q-1} \beta_i(T). \tag{1.20}$$

Donc, Y est une variable aléatoire à valeurs dans $\{2, 3, \cdots\}$ et pour tout $l \geq 2$,

$$\mathbb{P}(Y = l) = \int_0^\infty \sum_{k=2}^{l} q_k \sum_{i_1+\cdots+i_{k-1}=l-1} \Pi_{j=1}^{k-1} \mathbb{P}(\beta(t) = i_j)(1 + \frac{t}{\Gamma(\alpha+1)})^{-\frac{\alpha}{\alpha-1}-1} dt,$$

Au premier temps, remarquons que Q et $(\beta(t), t \geq 0)$ sont impliqués dans le théorème. En effet, les deux objets sont étroitement liés, ce qui est révélé dans le corollaire suivant. Notons aussi que $\lim_{n\to\infty} \mathbb{P}(\beta(t) = 1) = 1$ et on va montrer que $\mathbb{P}(\beta(t) = k)$ pour $k \geq 2$ est proportionnelle à t quand t tends vers $0+$.

Corollary 1.1. 1) *Pour tout $k \geq 2$,*

$$q_k = (\alpha - 1)\Gamma(\alpha) \lim_{t \to 0+} \frac{\mathbb{P}(\beta(t) = k)}{t}. \tag{1.21}$$

2) *La transformée de Laplace de Q est*

$$\mathbb{E}[e^{-\lambda Q}] = \lim_{t \to 0+} \mathbb{E}[(\alpha - 1)\Gamma(\alpha) \frac{e^{-\lambda \beta(t)} \mathbf{1}_{\beta(t) \geq 2}}{t}] = e^{-\lambda}(1 - (1 - e^{-\lambda})^{\alpha - 1}) \tag{1.22}$$

pour tout $\lambda \geq 0$.

Deuxièmement, la loi de Y semble très compliquée, ce qui endommage l'applicabilité du resultat. Mais la clarification donnée ci-dessous peut améliorer cette situation.

Corollary 1.2. *Lorsque k tend vers ∞, on a $\mathbb{P}(Y > k) \sim \frac{\int_0^\infty t^{\alpha-1} f_T(t) dt}{((\alpha-1)\Gamma(\alpha))^{\alpha-1}\Gamma(1-(\alpha-1)^2)} k^{-(\alpha-1)^2}$.*

Si α tend vers 1, alors $k^{-(\alpha-1)^2}$ converge vers 1. Ceci est consistant aver le cas de Bolthausen-Sznitman où $Y = \infty$. Mais si α tend vers 2, $k^{-(\alpha-1)^2}$ converge vers k^{-1}. Ceci n'est plus consistant avec le cas de Kingman. Ce corollaire montre une sorte de « discontinuité » du Beta-coalescent vers le coalescent de Kingman.

Remarquons que $Y^{(n)}$ est la taille du bloc qui contient 1 au moment où $\{1\}$ coalesce avec d'autres blocs. Une question naturelle et ayant un sens en biologie est de se demander la taille du plus grand bloc à cet instant-là. Tout d'abord, on considère la taille du plus grand bloc au temps $n^{1-\alpha}t$ pour $t > 0$.

Théorème 1.13. *Soit $W^{(n)}(t)$ la taille du plus grand bloc de $\Pi^{(n)}$ au temps t. Alors, lorsque n tend vers ∞,*

$$\frac{W^{(n)}((\alpha - 1)\Gamma(\alpha + 1)n^{1-\alpha}t)}{n^{1/\alpha}} \xrightarrow{(d)} W(t), \tag{1.23}$$

où $W(t)$ est une variable aléatoire suivant la loi de Gumbel de type 2, de fonction de répartition

$$\mathbb{P}(W(t) \leq x) = e^{-x^{-\alpha}(\alpha-1)t/\Gamma(2-\alpha)}. \tag{1.24}$$

La preuve de ce théorème utilise un résultat remarquable (voir (9)) qui met en relation Θ et les processus de branchements stables. Cette preuve est très proche de celle de la proposition 1.6 de (6).

Corollaire 1.5. *Soit $\tilde{W}^{(n)} = W^{(n)}(T_1^{(n)})$ la taille du plus grand bloc quand $\{1\}$ est coalescé. Alors , lorsque n tend vers ∞,*

$$\frac{\tilde{W}^{(n)}}{n^{1/\alpha}} \xrightarrow{(d)} \tilde{W}, \tag{1.25}$$

où $\tilde{W}$ est une variable aléatoire telle que pour tout $x \geq 0$,

$$\mathbb{P}(\tilde{W} \leq x) = \int_0^\infty e^{-x^{-\alpha} \frac{t}{\Gamma(\alpha+1)\Gamma(2-\alpha)}} f_T(t) dt.$$

Remarque 1.13. *Introduisons la quantité* $W^{(2,n)}(t)$ *qui est la taille du plus grand bloc au temps* t *de* $\Pi^{(2,n)}$. *Alors à la limite, on aura le même résultat que le Théorème 1.13 :*

$$\frac{W^{(2,n)}((\alpha-1)\Gamma(\alpha+1)n^{1-\alpha}t)}{n^{1/\alpha}} \xrightarrow{(d)} W(t), \tag{1.26}$$

La raison est que la taille du bloc ne dépend pas des notations des individus. En comparant (1.26) et (1.25), à la limite, la branche externe de $\{1\}$*est comme connectée à* $\Pi^{(2,n)}$ *d'une façon indépendante. Ce fait est en fait prouvé par la Proposition 1.2. Donc ce corollaire est une conséquence directe du Théorème 1.13 et de la Proposition 1.2.*

Chapitre 2

Sur la longueur d'une branche externe du Beta-coalescent

Version non modifiée de l'article *On the length of an external branch length of Beta-coalescent*
publié en 2013 dans
Stochastic Processes and their Applications, 123 (2013) 1691-1715.

2.1 Introduction

2.1.1 Motivation and main results

In modern genetics, it is possible to sequence whole genomes of individuals. In order to put these informations to maximal use, it is important to have well-fitting models for the gene genealogies of a sample of individuals. The standard model for gene genealogies of a sample of n individuals is Kingman's n-coalescent (see (25), (26)). Kingman's n-coalescent is a continuous-time Markov process with state space $\mathcal{P}^{(n)}$, the set of partitions of $\{1, \ldots, n\}$. The process starts in the trivial partition $(\{1\}, \ldots, \{n\})$ and transitions are only possible as mergers of exactly two blocks of the current state. Each such binary merger occurs with rate 1. These mergers are also called collisions.
For many populations, Kingman's n-coalescent describes the genealogy quite well. Kingman showed in (26) that the ancestral trees of a sample of size n in populations with size N evolving by a Wright-Fisher model will converge weakly to Kingman's n-coalescent for $N \to \infty$ (after a suitable time-change). This result is relatively robust if population evolution deviates from the Wright-Fisher model (see (26) or (27)). However, there is evidence that there are populations where the gene genealogies of a sample are not described well by Kingman's n-coalescent. Examples of such populations can be found in maritime species, where one individual can have a huge number of offspring with non-negligible probability (see (1), (10) (21), (19) and (15)).
A whole class of potential models for gene genealogies of a sample was introduced independently by Pitman and Sagitov (see (30) and (31)): The class of n-coalescents with multiple collisions. A n-coalescent with multiple collisions is a continuous-time Markov process with state space $\mathcal{P}^{(n)}$, where all possible transitions are done by merging two or more blocks of the current state into one new block. Every n-coalescent $\Pi^{(n)}$ is exchangeable, meaning $\tau \circ \Pi^{(n)} \stackrel{d}{=} \Pi^{(n)}$ for every permutation τ of $\{1, \ldots, n\}$. The transition rate of a merger/collision of k of b present blocks is given by

$$\lambda_{b,k} = \int_0^1 x^k (1-x)^{b-k} x^{-2} \Lambda(dx) \tag{2.1}$$

for a finite measure Λ on $[0,1]$ (this definition is due to (30)). Since the process is characterized by the measure Λ, it is also called a Λ-n-coalescent. Note that Kingman's n-coalescent is a Λ-n-coalescent with Λ being the Dirac measure δ_0 in 0.
An important subclass of Λ-n-coalescents are Beta n-coalescents characterized by Λ being a Beta distribution, especially for the choice of parameters $2-\alpha$ and α for $\alpha \in (0,2)$. The class of $Beta(2-\alpha, \alpha)$-n-coalescents appears as ancestral trees in various settings. They appear in the context of supercritical Galton-Watson processes (see (32)), of continuous-state branching processes (see (7)) and of continuous random trees (see (2)). They also seem to be a class where suitable models for the ancestral tree can be found for samples from species who do not fit well with the Kingman-setting (see (6; 15)). Note that for

$\alpha \to 2$, the rates of the $Beta(2-\alpha,\alpha)$-n-coalescent converge to the rates of Kingman's n-coalescent. In this sense, Kingman's n-coalescent can be seen as a border case of this class of Beta n-coalescents.
For $\alpha = 1$, $Beta(2-\alpha,\alpha)$ is the uniform distribution on $[0,1]$. The corresponding n-coalescent is the Bolthausen-Sznitman n-coalescent. It appears in the field of spin glasses (see (9), (11)) and is also connected to random recursive trees (see (20)). It is also seems to be a suitable model for a gene genealogy if selection is acting on the genome (e.g. see (3; 12; 29)).
Let us denote by $\Pi^{(n)} = (\Pi_t^{(n)})_{t\geq 0}$ a n-coalescent. In this paper, we are interested in three functionals of n-coalescents

- — the length $T^{(n)}$ of a randomly chosen external branch ;
- — the number $\sigma^{(n)}$ of collisions which occur in the n-coalescent until the end of a randomly chosen external branch ;
- — the block counting process $R^{(n)} = (R_t^{(n)})_{t\geq 0}$: $R_t^{(n)} = |\Pi_t^{(n)}|$ is the number of blocks of $\Pi_t^{(n)}$.

Note that $T^{(n)}$ can also be characterized as the waiting time for the first collision of a randomly chosen individual and $\sigma^{(n)}$ as the number of collisions we have to wait to see the randomly chosen individual merge. For $i \in \{1,\ldots,n\}$ define

$$T_i^{(n)} := \inf\left\{t|\,\{i\} \notin \Pi_t^{(n)}\right\}$$

as the length of the ith external branch and

$$\sigma_i^{(n)} := \inf\{k|\{i\} \notin \pi_k\}$$

as the number of collisions until the end of the ith external branch, where π_k is the state of the n-coalescent after k jumps. Due to the exchangeability of the n-coalescent, we have $T^{(n)} \stackrel{d}{=} T_1^{(n)}$ and $\sigma^{(n)} \stackrel{d}{=} \sigma_1^{(n)}$. Since we are only interested in distributional results, for the remainder of the article we will identify $T^{(n)}$ with $T_1^{(n)}$ and $\sigma^{(n)}$ with $\sigma_1^{(n)}$.
If the n-coalescent is used as a model for an ancestral tree of a sample of individuals/genes, the functionals $T^{(n)}$ and $\sigma^{(n)}$ can be interpreted biologically. The length of an external branch measures the uniqueness of the individual linked to that branch compared to the sample, since it gives the time this individual has to evolve by mutations that do not affect the rest of the sample (see the introduction of (13) for more information). It was first introduced by Fu and Li in (17), where they compare mutations on external and internal branches of Kingman's n-coalescent in order to test for the neutrality of mutations.
The functional $\sigma^{(n)}$ was first introduced in (13), though $n-\sigma^{(n)}$ was also analyzed in (8) as the level of coalescence of the chosen individual with the rest of the sample. In both articles, the functionals were defined for Kingman's n-coalescent.
For the biological interpretation of $\sigma^{(n)}$, we see the n-coalescent as an ancestral tree of a sample of size n. Each collision in the n-coalescent then resembles the emergence of an ancestor of the sample. $\sigma^{(n)}-1$ is the number of ancestors of the sample which emerge before

the most recent ancestor of the randomly chosen individual/gene emerges. In this line of thought, $\sigma^{(n)}$ gives the temporal position of the first ancestor of the chosen individual/gene among all ancestors of the sample, which are the $\tau^{(n)}$ collisions in the n-coalescent. Thus, $\frac{\sigma^{(n)}}{\tau^{(n)}}$ gives the relative temporal position of the first ancestor of the chosen individual/gene among all ancestors of the sample (until the most recent common ancestor). In this sense, we interprete $\frac{\sigma^{(n)}}{\tau^{(n)}}$ as a measure of how ancient the chosen individual/gene is compared to the rest of the sample.

In this article, we focus on the asymptotics of $T^{(n)}$, $\sigma^{(n)}$ and $R^{(n)}$ for $n \to \infty$. The asymptotics of these functionals are already known for some Λ-n-coalescents. For $T^{(n)}$, we have

— $\Lambda = \delta_0$ (Kingman coalescent): $nT^{(n)} \xrightarrow{d} T$, where T has density $t \mapsto \frac{8}{(2+t)^3}$ (see (8), (13), (23)),

— $\Lambda = Beta(1,1)$ (Bolthausen-Sznitman coalescent): $\log(n)T^{(n)} \xrightarrow{d} Exp(1)$ (see (16)),

— Λ with $\mu_{-1} = \int_0^1 x^{-1}\Lambda(dx) < \infty$: $T^{(n)} \xrightarrow{d} Exp(\mu_{-1})$ (see (28), see also (18))

for $n \to \infty$. For $\sigma^{(n)}$, we have

— $\Lambda = \delta_0$: $\sigma^{(n)}/n \xrightarrow{d} Beta(1,2)$ (see (13)),

— $\Lambda = Beta(1,1)$: $\frac{\log(n)}{n}\sigma^{(n)} \xrightarrow{d} Beta(1,1)$ (see (16)),

— Λ with $\mu_{-2} = \int_0^1 x^{-2}\Lambda(dx) < \infty$: $\sigma^{(n)} \xrightarrow{d} Geo(\frac{\mu_{-1}}{\mu_{-2}})$ (see (18))

for $n \to \infty$, where $Geo(p)$ is the geometric distribution on $\mathbb{N}$ with parameter p.

We will analyze the asymptotics for $T^{(n)}$, $\sigma^{(n)}$ and $R^{(n)}$ for Λ-n-coalescents with Λ fulfilling

$$\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta}),\ t \to 0$$

for some $C_0 > 0$, $\alpha \in (1,2)$ and $\zeta > 1 - 1/\alpha$, where $\rho(t) = \int_t^1 x^{-2}\Lambda(dx)$. Note that this class of n-coalescents includes all $Beta(a,b)$-n-coalescents with parameters $a \in (0,1)$ and $b > 0$. In this class of n-coalescents, we have the following asymptotics for $T^{(n)}$, $\sigma^{(n)}$ and $R^{(n)}$:

— $\frac{\sigma^{(n)}}{n(\alpha-1)} \xrightarrow{d} \sigma$,

— $n^{\alpha-1}T^{(n)} \xrightarrow{d} \frac{1}{C_0\Gamma(2-\alpha)}((1-\sigma)^{1-\alpha} - 1)$,

— for any $t_0 > 0, \varepsilon > 0$, $\mathbb{P}(\sup_{0\le t\le t_0} |n^{-1}R^{(n)}_{tn^{1-\alpha}} - (1 + C_0\Gamma(2-\alpha)t)^{-1/(\alpha-1)}| > \varepsilon) \to 0$,

for $n \to \infty$, where $\sigma \stackrel{d}{=} Beta(1,\alpha)$. We also prove the following asymptotic result for the block counting process of the Kingman coalescent:For any $t_0 > 0, \varepsilon > 0$, we have $\mathbb{P}(\sup_{0\le t\le t_0} |n^{-1}R^{(n)}_{tn^{-1}} - (1+t/2)^{-1}| > \varepsilon) \to 0$ when $n \to \infty$.

Note that if we see the Bolthausen-Sznitman n-coalescent as a $Beta(1,1)$-coalescent and Kingman's n-coalescent as the borderline case of a Beta distribution with parameter $\alpha = 1$, the convergence results for $\sigma^{(n)}$ shows a nice continuity in the parameters of the limit distributions in the range of $Beta(2-\alpha,\alpha)$-n-coalescents with $\alpha \in [1,2]$. Our convergence result itself is even somewhat true in the border cases 1 and 2 (if one wages

$(\alpha-1)^{-1} \to \infty$ for $\alpha \to 1$ against $\log(n) \to \infty$ for $n \to \infty$ for the Bolthausen-Sznitman n-coalescents). Also note that T obtained as the limit variable of $nT^{(n)}$ in Kingman's case has the same law as $2((1-\sigma)^{-1}-1)$, which gives again a nice continuity in results. The continuity also appears for the block counting process: replacing α by 2 in the formula for the $Beta(2-\alpha,\alpha)$-n-coalescent gives the formula for Kingman.

Finally we can remark that together with the known asymptotics $\frac{\tau^{(n)}}{n} \xrightarrow{d} \alpha - 1$ for this class of n-coalescents (see (14), (19) and (22)), we have $\sigma^{(n)}/\tau^{(n)} \xrightarrow{d} \sigma$.
To prove these results, we will exploit some techniques from (14). In (14), they were used to analyze the asymptotics of a part of the height of a n-coalescent and the number of collisions in a n-coalescent for the same class of n-coalescents as analyzed in the present paper. For the convergence result for $T^{(n)}$, we present two proofs. One mimic the approach in (13), using the convergence result for $\sigma^{(n)}$ and $T^{(n)} = \sum_{i=1}^{\sigma^{(n)}} T_i$, where T_i is the waiting time between the $i-1$th and ith collision/jump of the n-coalescent. The other proof is based on the representation of $T^{(n)}$ as the first jump time of a Cox process driven by a random rate process which depends only on the block counting process associated with the remaining individuals labelled $\{2, 3, \ldots, n\}$. We use a recursive construction suitable for any λ n-coalescent: This construction consists in adding individual i to a coalescent process constructed by individuals from 1 to n except i such that consistence relationship is fulfilled.

2.1.2 Organization of the paper

In section 2, we recall some known technical results which can all be found in (14). In section 3, we obtain the asymptotic result about $\sigma^{(n)}$ and also about the ratio between $\sigma^{(n)}$ and $\tau^{(n)}$. Section 4 studies the small time behavior of the block counting process $R^{(n)}$. Depending on the property of $R^{(n)}$, our first method taking $T^{(n)}$ as the first jump time of a Cox process gives the asymptotic behavior of $T^{(n)}$, hence of $T^{(n)}$ in section 5. In section 6, another method is provided by taking into account the fact that $T^{(n)}$ is the sum of $\sigma^{(n)}$ initial waiting times for the coalescent process $\Pi^{(n)}$ to jump from one state to the following.

2.2 Preliminaries

In this Section, we recall some results from (14).

Consider a n-coalescent with multiple collisions characterized by a finite measure Λ on $[0,1]$. Let $\nu(dx) = x^{-2}\Lambda(dx)$ and $\rho(t) = \nu[t,1]$. When the process has k blocks, the next coalescence event comes at rate g_k given by

$$g_k = \sum_{\ell=1}^{k-1} \binom{k}{\ell+1} \lambda_{k,\ell+1} = \int_{(0,1)} \Big(1 - (1-x)^k - kx(1-x)^{k-1}\Big) \frac{\Lambda(dx)}{x^2}. \qquad (2.2)$$

For $n \geq 1$, $x \in (0,1)$, let $B_{n,x}$ be a binomial r.v. with parameter (n,x). Recall that for $1 \leq k \leq n$, we have

$$\mathbb{P}(B_{n,x} \geq k) = \frac{n!}{(k-1)!(n-k)!} \int_0^x t^{k-1}(1-t)^{n-k}\, dt. \tag{2.3}$$

Use the first equality in 2.2 and 2.3 to get

$$\begin{aligned} g_n &= \int_0^1 \sum_{k=2}^{n} \binom{n}{k} x^k (1-x)^{n-k} \nu(dx) \\ &= \int_0^1 \mathbb{P}(B_{n,x} \geq 2)\nu(dx) \\ &= n(n-1)\int_0^1 (1-t)^{n-2} t\rho(t)\, dt. \end{aligned}$$

All along this paper, the following hypothesis will be assumed

$$\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta}). \tag{2.4}$$

For some $C_0 > 0$,$\alpha \in (1,2)$ and $\zeta > 1 - 1/\alpha$, Lemma 2.2 of (14) gives us that, for $n \geq 2$,

$$g_n = C_0\Gamma(2-\alpha)n^{\alpha} + O(n^{\alpha-\min(\zeta,1)}). \tag{2.5}$$

Recall that we call $\tau^{(n)}$ the number of coalescence events until reaching the common ancestor of the initial population (of size n). For $k \geq 0$, denote by $Y_k^{(n)}$ the number of blocks remaining after k jumps. Notice that $Y^{(n)}$ is a decreasing Markov chain with $Y_0^{(n)} = n$ and $Y_k^{(n)} = 1$ for $k \geq \tau^{(n)}$. Let $X_k^{(n)} = Y_{k-1}^{(n)} - Y_k^{(n)}$ be the number of blocks we lose during the kth coalescence event. We write $X_0^{(n)} = 0$.

The Markov property makes that the law of the first jump $X_1^{(n)}$ will be of much interest. We will look at some properties of $X_1^{(n)}$. Notice that

$$\mathbb{P}(X_1^{(n)} = k) = \frac{1}{g_n}\int_0^1 \mathbb{P}(B_{n,x} = k+1)\nu(dx) \tag{2.6}$$

and thus

$$\mathbb{P}(X_1^{(n)} \geq k) = \frac{\int_0^1 \mathbb{P}(B_{n,x} \geq k+1)\nu(dx)}{g_n} = \frac{(n-2)!}{k!(n-k-1)!}\frac{\int_0^1 (1-t)^{n-k-1}t^k\rho(t)\, dt}{\int_0^1 (1-t)^{n-2}t\rho(t)\, dt}. \tag{2.7}$$

Under the same assumptions on $\rho(t)$, setting $\varepsilon_0 > 0$ and

$$\varphi_n = \begin{cases} n^{-\zeta} & \text{if} \quad \zeta < \alpha - 1, \\ n^{1-\alpha+\varepsilon_0} & \text{if} \quad \zeta = \alpha - 1, \\ n^{1-\alpha} & \text{if} \quad \zeta > \alpha - 1, \end{cases} \tag{2.8}$$

Lemma 2.3 of (14) tells us there exists a constant $C_{2.9}$ s.t. for all $n \geq 2$, we have

$$\left|\mathbb{E}[X_1^{(n)}] - \frac{1}{\alpha - 1}\right| \leq C_{2.9}\varphi_n. \tag{2.9}$$

Moreover, from Lemma 2.4 of (14), there exists a constant $C_{2.10}$ s.t. for all $n \geq 2$, we have

$$\mathbb{E}\left[\left(X_1^{(n)}\right)^2\right] \leq C_{2.10}\frac{n^2}{g_n}. \tag{2.10}$$

We consider ϕ_n the Laplace transform of $X_1^{(n)}$: for $u \geq 0$, $\phi_n(u) = \mathbb{E}[\mathrm{e}^{-uX_1^{(n)}}]$. Assume that $\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta})$ for some $C_0 > 0$,$\alpha \in (1,2)$ and $\zeta > 0$. Let $\varepsilon_0 > 0$. Recall φ_n given by 2.8. Then we have (see (14), Lemma 2.5) , for $n \geq 2$,

$$\phi_n(u) = 1 - \frac{u}{\alpha - 1} + \frac{u^\alpha}{\alpha - 1} + R(n, u), \tag{2.11}$$

where $R(n, u) = \left(u\varphi_n + u^2\right) h(n, u)$ with $\sup_{u\in[0,K],n\geq 2} |h(n, u)| < \infty$ for all $K > 0$.

Moreover, if we assume that $\zeta > 1 - 1/\alpha$ and set $\eta \geq \frac{1}{\alpha}$, then (from (14), Lemma 3.2) there exist $\varepsilon_1 > 0$ and $C_{2.12}(K)$ a finite constant such that for all $n \geq 1$ and $u \in [0, K]$, a.s. with $a_n = n^{-\eta}$,

$$\sum_{i=1}^{\tau_n} \left|R(Y_{i-1}^{(n)}, ua_n)\right| \leq C_{2.12}(K)n^{-\varepsilon_1}. \tag{2.12}$$

We will also use the following result : Let $V = (V_t, t \geq 0)$ be a α-stable Lévy process with no positive jumps (see chap. VII in (4)) with Laplace exponent $\psi(u) = u^\alpha/(\alpha - 1)$: for all $u \geq 0$, $\mathbb{E}[\mathrm{e}^{-uV_t}] = \mathrm{e}^{tu^\alpha/(\alpha-1)}$. We assume that $\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta})$ for some $C_0 > 0$ and $\zeta > 1 - 1/\alpha$. Recall that $\tau^{(n)}$ is the number of coalescing events in the n-coalescent until reaching its absorbing state. Let

$$V_t^{(n)} = n^{-1/\alpha} \sum_{k=1}^{\lfloor nt\rfloor \wedge \tau^{(n)}} (X_k^{(n)} - \frac{1}{\alpha - 1})$$

for $t \in [0, \alpha - 1)$, and

$$V_{\alpha-1}^{(n)} = n^{-1/\alpha} \sum_{k=1}^{\tau^{(n)}} (X_k^{(n)} - \frac{1}{\alpha - 1}) = n^{-1/\alpha}\left(n - 1 - \frac{\tau^{(n)}}{\alpha - 1}\right).$$

Then,

$$(V_t^{(n)}, t \in [0, \alpha - 1]) \to (V_t, t \in [0, \alpha - 1]) \tag{2.13}$$

in the sense of convergence in law of the finite-dimensional marginals (see (14), Corollary 3.5, see also (19; 22)).

2.3 The number of collisions before the time $T^{(n)}$

Consider a n-coalescent with multiple collisions characterized by a finite measure Λ on $[0,1]$. Recall that $\nu(dx) = x^{-2}\Lambda(dx)$ and $\rho(t) = \nu[t,1]$ which is assumed to satisfy (2.4). A n-coalescent takes its values in $\mathcal{P}^{(n)}$, the set of partitions of $\{1,\ldots,n\}$. For $i \geq 0$, let $\pi_i = \pi_i^{(n)}$ be the state of the process after the ith coalescence event.

Pick at random an individual from the initial population and denote by $T^{(n)}$ the length of the external branch starting from it. Because of exchangeability, $T^{(n)}$ has the same law as the length $T_1^{(n)}$ of the external branch starting from the initial individual labelled by $\{1\}$. A quantity of interest will be $\sigma^{(n)}$, the number of coalescence events happening before time $T^{(n)}$. Again because of exchangeability, $\sigma^{(n)}$ has the same law as

$$\sigma_1^{(n)} = \inf\{i > 0, \{1\} \notin \pi_i\},$$

the number of collisions happening before individual 1 collides for the first time. We can write

$$T_1^{(n)} = \sum_{i=1}^{\sigma_1^{(n)}} \frac{e_i}{g_{Y_{i-1}^{(n)}}} \tag{2.14}$$

where the e_i's are i.i.d. exponential random variables with mean 1. Note that the formula also holds true for $T^{(n)}$ and $\sigma^{(n)}$ (just omit the subscripts). For the remainder of the chapter, we will identify $\sigma^{(n)}$ with $\sigma_1^{(n)}$.

In this section, we will determinate the asymptotic law of $\sigma^{(n)}$ for a class of coalescents containing the *Beta*-coalescent with $\alpha \in (1,2)$.

Theorem 2.1. *We assume that $\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta})$ for some $C_0 > 0$, $\alpha \in (1,2)$ and $\zeta > 1 - 1/\alpha$. Then*

$$\frac{\sigma^{(n)}}{n(\alpha-1)} \xrightarrow[n\to\infty]{d} \sigma, \tag{2.15}$$

for $n \to \infty$, where $\sigma \stackrel{d}{=} Beta(1,\alpha)$.

Recall that in this class of n-coalescents, we also have $\tau^{(n)}/n \stackrel{d}{\to} \alpha - 1$ (from (14), see also (19) and (22)) for $n \to \infty$. Slutsky's theorem gives a convergence result for $\sigma^{(n)}/\tau^{(n)}$, which measures how ancient the chosen individual is compared to the rest of the sample

Corollary 2.1. *We assume that $\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta})$ for some $C_0 > 0$, $\alpha \in (1,2)$ and $\zeta > 1 - 1/\alpha$. Then*

$$\frac{\sigma^{(n)}}{\tau^{(n)}} \stackrel{d}{\to} \sigma,$$

for $n \to \infty$, where $\sigma \stackrel{d}{=} Beta(1,\alpha)$.

proof of Theorem 2.1. For convenience, we set

1. $\binom{a}{b} = 0$, if $0 \leq a < b, a \in \mathbb{Z}_+, b \in \mathbb{Z}_+, \mathbb{Z}_+ = \{0, 1, 2, ...\}$.
2. $\log(0) = -\infty$.

Notice that $\sigma^{(n)} \leq \tau^{(n)}$. Let $\mathcal{Y} = (\mathcal{Y}_k, k \geq 0)$ denotes the filtration generated by $Y^{(n)}$. For any $t \geq 0$, we have

$$\begin{aligned}\mathbb{P}(\sigma^{(n)} > nt|\mathcal{Y}) &= \mathbb{P}(\sigma^{(n)} > nt, \tau^{(n)} > nt|\mathcal{Y})\\ &= \prod_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \mathbb{P}(\{1\} \in \pi_i | \{1\} \in \pi_{i-1}, \mathcal{Y})\\ &= \prod_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \frac{\binom{Y_{i-1}^{(n)}-1}{X_i^{(n)}+1}}{\binom{Y_{i-1}^{(n)}}{X_i^{(n)}+1}}\\ &= \prod_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \frac{Y_{i-1}^{(n)} - (X_i^{(n)} + 1)}{Y_{i-1}^{(n)}}.\end{aligned}$$

Notice that $\frac{X_i^{(n)}+1}{Y_{i-1}^{(n)}} < 1$ if $i \neq \tau^{(n)}$, and $\frac{X_i^{(n)}+1}{Y_{i-1}^{(n)}} = 1$ if $i = \tau^{(n)}$.
We can hence write

$$\log\left(\mathbb{P}(\sigma^{(n)} > nt|\mathcal{Y})\right) = \sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \log\left(1 - \frac{X_i^{(n)} + 1}{Y_{i-1}^{(n)}}\right)$$

and proceed to a power series expansion :

$$\log\left(\mathbb{P}(\sigma^{(n)} > nt|\mathcal{Y})\right) = I_{nt}^{(1)} + I_{nt}^{(2)},$$

with

$$I_{nt}^{(1)} = -\sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(X_i^{(n)} + 1\right)\left(Y_{i-1}^{(n)}\right)^{-1},$$

and

$$I_{nt}^{(2)} = \sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} (\ln(1 - \frac{X_i^{(n)} + 1}{Y_{i-1}^{(n)}}) + \frac{X_i^{(n)} + 1}{Y_{i-1}^{(n)}}),$$

where $I_{nt}^{(2)}$ can be $-\infty$ if $\frac{X_i^{(n)}+1}{Y_{i-1}^{(n)}} = 1$. Let us look further at $I_{nt}^{(1)}$. The idea is to replace $X_i^{(n)}$ by the limit of its expectation.

$$I_{nt}^{(1)} = J_{nt}^{(1)} + J_{nt}^{(2)},$$

with

$$J_{nt}^{(1)} = -\sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(\frac{1}{\alpha-1} + 1\right)\left(Y_{i-1}^{(n)}\right)^{-1},$$

and

$$J_{nt}^{(2)} = -\sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(X_i^{(n)} - \frac{1}{\alpha-1}\right)\left(Y_{i-1}^{(n)}\right)^{-1}.$$

We will use three lemmas whose proofs are given in the rest of the Section.

Lemma 2.1, with $\eta = 1$, tells us that, when $0 < t < \alpha - 1$

$$J_{nt}^{(1)} \xrightarrow{\mathbb{P}} -\frac{\alpha}{\alpha-1}\int_0^t \left(1 - \frac{x}{\alpha-1}\right)^{-1} dx = \alpha \log\left(1 - \frac{t}{\alpha-1}\right). \tag{2.16}$$

Lemma 2.2 gives, for $0 < t < \alpha - 1$,

$$J_{nt}^{(2)} = -n^{1/\alpha-1} n^{1-1/\alpha} \sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(X_i^{(n)} - \frac{1}{\alpha-1}\right)\left(Y_{i-1}^{(n)}\right)^{-1} \xrightarrow{\mathbb{P}} 0. \tag{2.17}$$

Finally, Lemma 2.3 gives, for $t < \alpha - 1$,

$$I_{nt}^{(2)} \xrightarrow{\mathbb{P}} 0. \tag{2.18}$$

Adding (2.16), (2.17) and (2.18), we get that for $t < \alpha - 1$

$$\log\left(\mathbb{P}(\sigma^{(n)} > nt|\mathcal{Y})\right) \xrightarrow{\mathbb{P}} \alpha \log\left(1 - \frac{t}{\alpha-1}\right),$$

and thus

$$\mathbb{P}(\sigma^{(n)} > nt|\mathcal{Y}) \xrightarrow{\mathbb{P}} \left(1 - \frac{t}{\alpha-1}\right)^{\alpha}.$$

While we know that $\mathbb{P}(\sigma^{(n)} > nt|\mathcal{Y}) \leq 1$,then

$$\mathbb{P}(\sigma^{(n)} > nt) = \mathbb{E}[\mathbb{P}(\sigma^{(n)} > nt|\mathcal{Y})] \to \left(1 - \frac{t}{\alpha-1}\right)^{\alpha}.$$

We thus obtain that, for $x \in (0,1)$,

$$\mathbb{P}(\sigma^{(n)} > n(\alpha-1)x) \to (1-x)^{\alpha}.$$

and then that $\dfrac{\sigma^{(n)}}{n(\alpha-1)}$ converges in distribution to a $Beta(1,\alpha)$ law. □

Lemma 2.1. *We set $\nu_\eta(t) = \int_0^t \left(1 - \frac{x}{\alpha-1}\right)^{-\eta} dx, \eta \in \mathbb{R}$. We assume that $\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta})$ for some $C_0 > 0$, $\alpha \in (1,2)$ and $\zeta > 1 - 1/\alpha$. For any $0 < t < \alpha - 1$ and $\eta \in \mathbb{R}$, we have*

1. *Let $t_0 \in [0, \alpha-1)$ and $\delta > 0$. The following convergence in probability holds when $n \to \infty$:*
$$n^{(\alpha-1)/2-\delta} \sup_{0 \le t \le t_0} |n^{\eta-1} \sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(Y_{i-1}^{(n)}\right)^{-\eta} - \nu_\eta(t)| \to 0.$$
2. *Let $t \in [0, \alpha-1)$. The following convergence in distribution holds when $n \to \infty$:*
$$n^{\eta-1/\alpha} (\sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(Y_{i-1}^{(n)}\right)^{-\eta} - n^{1-\eta}\nu_\eta(t)) \to \eta \int_0^t dr (1 - \frac{r}{\gamma})^{-\eta-1} V_r.$$

Proof. The case $\eta = \alpha - 1$ is given by Theorem 5.1 in (14). Following the same arguments, it is easy to get the general result. □

Lemma 2.2. *For any $t < \alpha - 1$, the following convergence in distribution holds :*

$$n^{1-1/\alpha} \sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(X_i^{(n)} - \frac{1}{\alpha-1}\right) \left(Y_{i-1}^{(n)}\right)^{-1} \xrightarrow{d} (v_\alpha(t))^{1/\alpha} V_1,$$

where $(V_t)_{t \ge 0}$ is an α-stable Lévy process with no positive jumps and $v_\alpha(t) = \int_0^t \left(1 - \frac{x}{\alpha-1}\right)^{-\alpha} dx$.

Proof. Let $\delta \in (0, \alpha-1)$, $t_0 = \alpha - 1 - \delta$ and $t \in [0, t_0]$.

Let $\varepsilon \in (0, 1 - \frac{t}{\alpha-1})$ and $\beta = 1 - \frac{t}{\alpha-1} - \varepsilon > 0$. We have

$$n^{1-1/\alpha} \sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(X_i^{(n)} - \frac{1}{\alpha-1}\right) \left(Y_{i-1}^{(n)}\right)^{-1} = A_{nt} + B_{nt},$$

with

$$A_{nt} = n^{1-1/\alpha} \sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(X_i^{(n)} - \frac{1}{\alpha-1}\right) \left(Y_{i-1}^{(n)}\right)^{-1} \mathbf{1}_{\{Y_{i-1}^{(n)} \ge n\beta\}},$$

and

$$B_{nt} = n^{1-1/\alpha} \sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(X_i^{(n)} - \frac{1}{\alpha-1}\right) \left(Y_{i-1}^{(n)}\right)^{-1} \mathbf{1}_{\{Y_{i-1}^{(n)} < n\beta\}}.$$

We will show that B_{nt} converges to 0 in probability and that A_{nt} weakly converges to $(v_\alpha(t))^{1/\alpha} V_1$ as $n \to \infty$.

Convergence of A_{nt}. Let $Z_i^{(n)} = n \left(Y_{i-1}^{(n)}\right)^{-1} \mathbf{1}_{\{Y_{i-1}^{(n)} \ge n\beta\}}$. We have that $\sup_{n,i \ge 1} Z_i^{(n)} \le \beta^{-1}$ a.s..

By using (2.11), it is enough to prove that

$$\mathbb{E}[\exp(-uA_{nt})] \xrightarrow[n\to\infty]{} \mathrm{e}^{v_\alpha(t)u^\alpha/(\alpha-1)},$$

for any u positive. Where $\mathrm{e}^{v_\alpha(t)u^\alpha/(\alpha-1)}$ is the Laplace transform of $(v_\alpha(t))^{1/\alpha}V_1$.

Taking $uZ_i^{(n)}$ as $Z_i^{(n)}$, we shall only consider the case $u=1$.

Let us consider $n^{-1}\sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}}\left(Z_i^{(n)}\right)^\alpha = n^{\alpha-1}\sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}}\left(Y_{i-1}^{(n)}\right)^{-\alpha}\mathbf{1}_{\{Y_{i-1}^{(n)}\geq n\beta\}}$. We have, because the process $(Y_i^{(n)}, i\geq 0)$ is decreasing, that

$$\begin{aligned}\mathbb{P}(n^{-1}\sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}}\left(Z_i^{(n)}\right)^\alpha \neq n^{\alpha-1}\sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}}\left(Y_{i-1}^{(n)}\right)^{-\alpha}) &= \mathbb{P}(\exists i; Y_{i-1}^{(n)} < n\beta) \\ &\leq \mathbb{P}(\{Y^{(n)}_{(\lfloor nt\rfloor\wedge\tau^{(n)})-1} < n\beta) \\ &\leq \mathbb{P}(n^{-1}\sum_{j=1}^{(\lfloor nt\rfloor\wedge\tau^{(n)})-1}(X_j^{(n)} - \frac{1}{\alpha-1}) \geq \varepsilon).\end{aligned}$$

Use (2.13) to get that the right-hand side of the last inequality converges to 0 as n goes to infinity. Using also Lemma 2.1 with $\eta=\alpha$, we have that

$$n^{\alpha-1}\sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}}\left(Y_{i-1}^{(n)}\right)^{-\alpha} \xrightarrow{\mathbb{P}} v_\alpha(t),$$

as $n\to\infty$. We can thus deduce that

$$n^{-1}\sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}}\left(Z_i^{(n)}\right)^\alpha \xrightarrow{\mathbb{P}} v_\alpha(t), \tag{2.19}$$

as $n\to\infty$.

For $a>0$, we set

$$M_{n,k}^{(a)} = \exp\left(\sum_{i=1}^{k}\left(-n^{-1/\alpha}aZ_i^{(n)}X_i^{(n)} - \log\phi_{Y_{i-1}^{(n)}}(n^{-1/\alpha}aZ_i^{(n)})\right)\right).$$

The process $(M_{n,k}^{(a)}, k\geq 1)$ is a bounded martingale w.r.t. the filtration $\mathcal{Y}$. Notice that $\mathbb{E}[M_{n,k}^{(a)}]=1$. As $X_i^n=0$ and $Z_i^{(n)}=0$ for $i>\tau^{(n)}$, we also have

$$M_{n,k}^{(a)} = \exp\left(\sum_{i=1}^{k\wedge\tau^{(n)}}\left(-n^{-1/\alpha}aZ_i^{(n)}X_i^{(n)} - \log\phi_{Y_{i-1}^{(n)}}(n^{-1/\alpha}aZ_i^{(n)})\right)\right).$$

Using $R(n,u)$ defined in 2.11, we get that :

$$\begin{aligned}
&M^{(a)}_{n,\lfloor nt\rfloor}\\
&= \exp\left(-\sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}} n^{-1/\alpha} a Z_i^{(n)}(X_i^{(n)}-\frac{1}{\alpha-1}) - \sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}} \frac{n^{-1}\left(aZ_i^{(n)}\right)^{\alpha}}{\alpha-1} - \sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}} R(Y_{k-1}^{(n)}, n^{-1/\alpha} a Z_i^{(n)})\right)\\
&= \exp\left(-aA_{nt}\right)\exp\left(-n^{-1}\sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}} \frac{\left(aZ_i^{(n)}\right)^{\alpha}}{\alpha-1} - \sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}} R(Y_{k-1}^{(n)}, n^{-1/\alpha} a Z_i^{(n)})\right).
\end{aligned}$$

Let

$$\Lambda_n = -n^{-1}\sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}} \frac{\left(Z_i^{(n)}\right)^{\alpha}}{\alpha-1} + \frac{v_\alpha(t)}{\alpha-1},$$

and write

$$\mathbb{E}\left[\exp\left(-A_{nt}\right)\right] = A_1 + A_2,$$

with $A_1 = \mathbb{E}\left[\mathrm{e}^{-A_{nt}}\left(1-\mathrm{e}^{\Lambda_n}\right)\right]$ and $A_2 = \mathbb{E}\left[\mathrm{e}^{-A_{nt}}\,\mathrm{e}^{\Lambda_n}\right]$.

First of all, let us prove that A_1 converges to 0 when n tends to ∞. Recall that the r.v $Z_i^{(n)}$ are uniformly bounded by β^{-1} a.s.. Thanks to (2.12), we have

$$\mathbb{E}[\mathrm{e}^{-2A_{nt}}] = \mathbb{E}\left[M^{(2)}_{n,\lfloor nt\rfloor}\exp\left(n^{-1}\sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}} \frac{\left(2Z_i^{(n)}\right)^{\alpha}}{\alpha-1} + \sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}} R(Y_{k-1}^{(n)}, 2n^{-1/\alpha} Z_i^{(n)})\right)\right] \leq M,$$

where M is a finite constant which does not depend on n. By Cauchy-Schwarz' inequality, we get that

$$(A_1)^2 \leq \left(\mathbb{E}\left[\mathrm{e}^{-A_{nt}}\left|1-\mathrm{e}^{\Lambda_n}\right|\right]\right)^2 \leq \mathbb{E}\left[\mathrm{e}^{-2A_{nt}}\right]\mathbb{E}\left[\left(1-\mathrm{e}^{\Lambda_n}\right)^2\right] \leq M\mathbb{E}\left[\left(1-\mathrm{e}^{\Lambda_n}\right)^2\right]. \quad (2.20)$$

The quantity Λ_n is bounded and goes to 0 in probability when n goes to infinity (see (2.19)). Therefore, the right-hand side of 2.20 converges to 0. This implies that $\lim_{n\to\infty} A_1 = 0$.

Let us now consider the convergence of A_2. Remark that

$$A_2 = \mathbb{E}\left[M^{(1)}_{n,\lfloor nt\rfloor}\exp\left(\frac{v_\alpha(t)}{\alpha-1} + \sum_{k=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}} R(Y_{k-1}^{(n)}, n^{-1/\alpha} Z_i^{(n)})\right)\right].$$

Recall that $\mathbb{E}[M^{(1)}_{n,\lfloor nt\rfloor}] = 1$. Using (2.12), we get

$$\exp\left(-C_{2.12}(\beta^{-1})n^{-\varepsilon_1} + \frac{v_\alpha(t)}{\alpha-1}\right) \leq A_2 \leq \exp\left(C_{2.12}(\beta^{-1})n^{-\varepsilon_1} + \frac{v_\alpha(t)}{\alpha-1}\right).$$

We get that $\lim_{n\to\infty} A_2 = \mathrm{e}^{v_\alpha(t)/(\alpha-1)}$, which achieves the proof.
Convergence of B_{nt}. Here, we will use a similar approach as the one we used on the first half of p.48. The process $(Y_i^{(n)}, i \geq 0)$ is decreasing. So if for some $i \leq \lfloor nt \rfloor$, $Y_{i-1}^{(n)} < n\beta$, then we have $Y_{\lfloor nt \rfloor -1}^{(n)} < n\beta$. Thus we get $B_{nt} = B_{nt}\mathbf{1}_{\{Y^{(n)}_{(\lfloor nt\rfloor\wedge\tau^{(n)})-1} < n\beta\}}$. Moreover,

$$\{Y^{(n)}_{(\lfloor nt\rfloor\wedge\tau^{(n)})-1} < n\beta\} \subset \{n^{-1}\sum_{j=1}^{(\lfloor nt\rfloor\wedge\tau^{(n)})-1}(X_j^{(n)} - \frac{1}{\alpha-1}) \geq \varepsilon\},$$

and then for any $\varepsilon' > 0$

$$\begin{aligned}\mathbb{P}(|B_{nt}| \geq \varepsilon') &= \mathbb{P}(\mathbf{1}_{\{Y^{(n)}_{(\lfloor nt\rfloor\wedge\tau^{(n)})-1} < n\beta\}}|B_{nt}| \geq \varepsilon') \\ &\leq \mathbb{P}(\{Y^{(n)}_{(\lfloor nt\rfloor\wedge\tau^{(n)})-1} < n\beta) \\ &\leq \mathbb{P}(n^{-1}\sum_{j=1}^{(\lfloor nt\rfloor\wedge\tau^{(n)})-1}(X_j^{(n)} - \frac{1}{\alpha-1}) \geq \varepsilon).\end{aligned}$$

Use (2.13) to get that the right-hand side of the last inequality converges to 0 as n goes to infinity.

□

Now we deal with $I_{nt}^{(2)}$.

Lemma 2.3. *We assume that $\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta})$ for some $C_0 > 0$ and $\zeta > 1 - 1/\alpha$. Then, for any $t < \alpha - 1$, we have*

$$|I_{nt}^{(2)}| = \sum_{i=1}^{\lfloor nt\rfloor\wedge\tau^{(n)}} -(\ln(1 - \frac{X_i^{(n)}+1}{Y_{i-1}^{(n)}}) + \frac{X_i^{(n)}+1}{Y_{i-1}^{(n)}}) \xrightarrow{\mathbb{P}} 0,$$

when $n \to \infty$.

Proof. Let $0 \leq t < \alpha - 1$. First of all, remark that:

$$\frac{Y_{\lfloor nt\rfloor}^{(n)}}{n} \xrightarrow{\mathbb{P}} \frac{\alpha-1-t}{\alpha-1}. \tag{2.21}$$

Indeed,

$$\frac{Y_{\lfloor nt\rfloor}^{(n)}}{n} = \frac{n - (\lfloor nt\rfloor)/(\alpha-1)}{n} - \frac{\sum_{k=1}^{\lfloor nt\rfloor}(X_k^{(n)} - 1/(\alpha-1))}{n},$$

and we conclude using (2.13) and the convergence of $\mathbb{P}(\tau^{(n)} > \lfloor nt\rfloor)$ to 1. Let us write

$$|I_{nt}^{(2)}| = A_n + B_n,$$

with

$$A_n = |I_{nt}^{(2)}|\mathbf{1}_{\left\{Y_{\lfloor nt\rfloor}^{(n)}<(1-t/(\alpha-1))n/2\right\}}$$

and

$$B_n = |I_{nt}^{(2)}|\mathbf{1}_{\left\{Y_{\lfloor nt\rfloor}^{(n)}\geq(1-t/(\alpha-1))n/2\right\}}.$$

The convergence (2.21) implies that A_n tends to 0 in probability. To prove the convergence of B_n, let us first notice that for $a \in (0,1)$, there exists a constant $C(a)$ such that, if $B_{n,x}$ is a binomial r.v. with parameter (n,x), then

$$0 < -\int_0^1 \mathbb{E}\left[\mathbf{1}_{\{2\leq B_{n,x}\leq(1-a)n\}}(\ln(1-\frac{B_{n,x}}{n})+\frac{B_{n,x}}{n})\right]\nu(dx) \leq C(a). \tag{2.22}$$

Indeed, there exists a constant $C'(a)$ such that for $u \in (0, 1-a)$, $0 < -\ln(1-u) - u \leq C'(a)u^2$. Hence,

$$\begin{aligned}
0 &< -\int_0^1 \mathbb{E}\left[\mathbf{1}_{\{2\leq B_{n,x}\leq(1-a)n\}}(\ln(1-\frac{B_{n,x}}{n})+\frac{B_{n,x}}{n})\right]\nu(dx)\\
&\leq C'(a)\int_0^1 \mathbb{E}\left[\mathbf{1}_{\{2\leq B_{n,x}\leq(1-a)n\}}(\frac{B_{n,x}}{n})^2\right]\nu(dx)\\
&\leq C'(a)\int_0^1 \mathbb{E}\left[(\frac{B_{n,x}}{n})^2\right]\nu(dx)\\
&\leq 2C'(a)\int_0^1 \mathbb{E}\left[\frac{B_{n,x}(B_{n,x}-1)}{n^2}\right]\nu(dx)\\
&= 2C'(a)\frac{\int_0^1 n(n-1)x^2\nu(dx)}{n^2} =: C(a).
\end{aligned}$$

Let us set $a = (1-t/(\alpha-1))/2$. Hence $B_n = |I_{nt}^{(2)}|\mathbf{1}_{\left\{Y_{\lfloor nt\rfloor}^{(n)}\geq an\right\}}$. Notice that if n is large enough such that $an \geq 2$, then if $Y_{\lfloor nt\rfloor}^{(n)} \geq an$ we have $\tau^{(n)} > nt$. Moreover, if $Y_{\lfloor nt\rfloor}^{(n)} \geq an$, for $i \leq nt$, we have $Y_i^{(n)} \geq an \geq aY_{i-1}^{(n)}$ and $X_i^{(n)} = Y_{i-1}^{(n)} - Y_i^{(n)} \leq (1-a)Y_{i-1}^{(n)} < (1-a/2)Y_{i-1}^{(n)}$. Using (2.6), (2.22) and (2.5), we get that

$$\begin{aligned}
&\mathbb{E}[B_n]\\
&\leq \sum_{i=1}^{\lfloor nt\rfloor}\mathbb{E}[\mathbb{E}[-(\ln(1-\frac{X_i^{(n)}+1}{Y_{i-1}^{(n)}})+\frac{X_i^{(n)}+1}{Y_{i-1}^{(n)}})\mathbf{1}_{\{1\leq X_i^{(n)}\leq(1-a)Y_{i-1}^{(n)}\}}\mathbf{1}_{\{Y_{i-1}^{(n)}\geq an\}}|Y_{i-1}^{(n)}]]\\
&\leq \sum_{i=1}^{\lfloor nt\rfloor}\mathbb{E}[-\mathbb{E}[\int_0^1 \mathbf{1}_{\{2\leq B_{Y_{i-1}^{(n)},x}\leq(1-a/2)Y_{i-1}^{(n)}\}}\mathbf{1}_{\{Y_{i-1}^{(n)}\geq an\}}\frac{1}{g_{Y_{i-1}^{(n)}}}(\ln(1-\frac{B_{Y_{i-1}^{(n)},x}}{Y_{i-1}^{(n)}})+\frac{B_{Y_{i-1}^{(n)},x}}{Y_{i-1}^{(n)}})\nu(dx)|Y_{i-1}^{(n)}]]\\
&\leq \frac{C(a/2)nt}{g_{an}} \to 0,
\end{aligned}$$

when n tends to ∞. This achieves the proof of the Lemma. □

2.4 A result on small-time behavior of the block process

We now turn to the study of the length of an external branch picked at random, denoted by $T^{(n)}$. For any integer k between 1 and $\tau^{(n)}$, define $A_k^{(n)}$ as the time when the kth jump is achieved. This variable can be expressed as a sum of k independent exponential random variables. More precisely,

$$A_k^{(n)} = \sum_{i=1}^{k\wedge\tau^{(n)}} \frac{e_i}{g_{Y_{i-1}^{(n)}}},$$

where the e_i's are independent standard exponential variables. Notice that $T^{(n)} = A_{\sigma^{(n)}}^{(n)}$. We will first study asymptotics of $A_k^{(n)}$. For this, we use a two-step approximation method close to Section 4 of (14). Define first

$$\tilde{A}_k^{(n)} = \sum_{i=1}^{k\wedge\tau^{(n)}} \frac{1}{g_{Y_{i-1}^{(n)}}},$$

obtained replacing the e_i's by their mean,and

$$\hat{A}_k^{(n)} = \frac{1}{C_0\Gamma(2-\alpha)} \sum_{i=1}^{k\wedge\tau^{(n)}} (Y_{i-1}^{(n)})^{-\alpha},$$

obtained replacing g_b by its equivalent in (2.5).

Proposition 2.1. *We assume that $\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta})$ for some $C_0 > 0$ and $\zeta > 1 - 1/\alpha$. Then, for any $t < \alpha - 1$, we have*

$$n^{\alpha-1} A_{\lfloor nt\rfloor}^{(n)} \xrightarrow{\mathbb{P}} \frac{1}{C_0\Gamma(2-\alpha)}((1-\frac{t}{\alpha-1})^{1-\alpha} - 1),$$

when $n \to \infty$.

The proof is a straight consequence of Lemma 2.1 with $\eta = \alpha$ and the following Lemmas 2.4 and 2.5.

Lemma 2.4. *Under the assumptions of Proposition 2.1, we have*

$$n^{\alpha-1}(\tilde{A}_{\lfloor nt\rfloor}^{(n)} - \hat{A}_{\lfloor nt\rfloor}^{(n)}) \xrightarrow{\mathbb{P}} 0,$$

when $n \to \infty$.

Proof. Use (2.5) to get

$$\tilde{A}^{(n)}_{\lfloor nt \rfloor} - \hat{A}^{(n)}_{\lfloor nt \rfloor} = \sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(Y^{(n)}_{i-1}\right)^{-\alpha} O\left(\left(Y^{(n)}_{i-1}\right)^{-\min(\zeta,1)}\right).$$

The result then follows from Lemma 2.1 with $\eta = \alpha + \min(\zeta, 1)$. □

Lemma 2.5. *Under the assumptions of Proposition 2.1, we have*

$$n^{\alpha-1}(A^{(n)}_{\lfloor nt \rfloor} - \tilde{A}^{(n)}_{\lfloor nt \rfloor)}) \xrightarrow{\mathbb{P}} 0,$$

when $n \to \infty$.

Proof. Recall that $\mathcal{Y} = (\mathcal{Y}_k, k \geq 0)$ denotes the filtration generated by Y. Conditionally on $\mathcal{Y}$, the random variables $\dfrac{e_i - 1}{g_{Y^{(n)}_{i-1}}}$ are independent with zero mean. We deduce that

$$\mathbb{E}\left[\sup_{t\geq 0}(n^{\alpha-1}(A^{(n)}_{\lfloor nt \rfloor} - \tilde{A}^{(n)}_{\lfloor nt \rfloor}))^2 | \mathcal{Y}\right] = n^{2\alpha-2}\mathbb{E}\left[\sup_{t\geq 0}\left(\sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \frac{e_i - 1}{g_{Y^{(n)}_{i-1}}}\right)^2 | \mathcal{Y}\right]$$

$$\leq 4n^{2\alpha-2} \sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(\frac{1}{g_{Y^{(n)}_{i-1}}}\right)^2,$$

where we used Doob's inequality for the inequality. Thanks to 2.5 and Lemma 2.1 with $\eta = 2\alpha$, we get the $4n^{2\alpha-2}\sum_{i=1}^{\lfloor nt \rfloor \wedge \tau^{(n)}} \left(\frac{1}{g_{Y^{(n)}_{i-1}}}\right)^2$ converges to 0 in probability.

□

Heuristically, combining Theorem 2.1 and Proposition 2.1, we should get that $n^{\alpha-1}T^{(n)} = n^{\alpha-1}A^{(n)}_{\sigma^{(n)}}$ converges in law to $\frac{1}{C_0\Gamma(2-\alpha)}((1-\sigma)^{1-\alpha} - 1)$. This line of proof will be followed in the last section. However, in the next we will first present another way to prove this result with a method based on exchangeable coalescents consistency property. As a first step to this approach, we end this session with a result about small-time behavior of the block-counting process.

Let $R^{(n)}_t$ denote the number of blocks of the n-coalescent $\Pi^{(n)}$ at time t. The initial value $R^{(n)}_0$ is n. We show that the limit law of the process $R^{(n)}$ is deterministic under a certain time rescaling

Theorem 2.2. *We assume that* $\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta})$ *for some* $C_0 > 0$ *and* $\zeta > 1 - 1/\alpha$. *For any* $t_0 > 0, \varepsilon > 0$ *, we have*

$$\mathbb{P}(\sup_{0\leq t\leq t_0} |n^{-1}R^{(n)}_{tn^{1-\alpha}} - (1 + C_0\Gamma(2-\alpha)t)^{-1/(\alpha-1)}| > \varepsilon) \to 0, \quad (2.23)$$

when $n \to \infty$.

Proof. Let $0 < r < \alpha - 1$, we have the following relation :

$$R^{(n)}_{A^{(n)}_{\lfloor nr \rfloor}} = Y^{(n)}_{\lfloor nr \rfloor} = n - \sum_{j=1}^{\lfloor nr \rfloor \wedge \tau^{(n)}} X^{(n)}_j$$

Let $t \in [0, t_0]$, and define

$$r(t) = (\alpha - 1)(1 - (1 + C_0\Gamma(2-\alpha)t)^{-1/(\alpha-1)}), \tag{2.24}$$

on $[0, t_0]$. Notice that

$$\frac{1}{C_0\Gamma(2-\alpha)}((1 - \frac{r(t)}{\alpha - 1})^{1-\alpha} - 1) = t.$$

Then thanks to Proposition 2.1, $n^{\alpha-1}A^{(n)}_{\lfloor nr(t) \rfloor}$ converges in probability to t.

Using the remark at the beginning of the proof in Lemma 2.3, we get the convergence

$$n^{-1}R^{(n)}_{A^{(n)}_{\lfloor nr(t) \rfloor}} = \frac{Y^{(n)}_{\lfloor nr(t) \rfloor}}{n} \xrightarrow{\mathbb{P}} (1 - \frac{r(t)}{\alpha - 1}) = (1 + C_0\Gamma(2-\alpha)t)^{-1/(\alpha-1)},$$

when $n \to \infty$. Moreover, since $R^{(n)}_t$ is decreasing, then for any $0 < \delta < 1$,

$$\lim_{n\to\infty} \mathbb{P}(R^{(n)}_{A^{(n)}_{\lfloor nr(t-\delta t) \rfloor}} \leq R^{(n)}_{tn^{1-\alpha+1}} \leq R^{(n)}_{A^{(n)}_{\lfloor nr(t+\delta t) \rfloor}}) = 1.$$

The constant δ being arbitrary, we thus obtain the convergence in probability of $n^{-1}R^{(n)}_{tn^{1-\alpha}}$ to $(1 + C_0\Gamma(2-\alpha)t)^{-1/(\alpha-1)}$.

We obtain (2.25) using again the fact that $R^{(n)}$ is a decreasing process.

□

In fact, the symptotic result concerning block counting process of Kingman coalescent is also valid. The method is almost identical to that emploied in the above Theorem. In the context of Kingman coalescent, we use the same notations $\Pi^{(n)}$, $A^{(n)}_i$, $R^{(n)}$.

Theorem 2.3. *In the setting of the Kingman coalescent, for any $t_0 > 0, \varepsilon > 0$, we have*

$$\mathbb{P}(\sup_{0\leq t\leq t_0} |n^{-1}R^{(n)}_{tn^{-1}} - (1 + t/2)^{-1}| > \varepsilon) \to 0 \tag{2.25}$$

when $n \to \infty$.

Remark that this Theorem shows a nice continuity from Beta coalescent(the process that we consider is more general which contains Beta coalescent) to Kingman coalescent. In Theorem 2.2, if we let $\alpha = 2$, the result is just we want to prove in this Corollary.

Proof. Recall that $A_i^{(n)}$ is the time when i-th jump is achieved. When $\Pi^{(n)}$ has b individuals at some time t, then the process encounters the following coalescence at rate $\binom{b}{2}$ where two randomly chosen individuals will be coalesced. $\Pi^{(n)}$ remains 1 when all individuals are coalesced.

For $0 < t < \alpha - 1$, we have

$$A_{\lfloor nt \rfloor}^{(n)} = \sum_{k=n}^{n-\lfloor nt \rfloor+1} \frac{e_k}{\binom{b}{2}}$$

where e_is are i.i.d unit exponential variables. Notice that

$$\mathbb{E}[nA_{\lfloor nt \rfloor}^{(n)}] = \sum_{k=n}^{n-\lfloor nt \rfloor+1} \frac{n}{\binom{k}{2}} = 2(\frac{1}{n-\lfloor nt \rfloor} - \frac{1}{n})n \to 2(\frac{1}{1-t} - 1),$$

as n tends to ∞. There exist a constant $K > 0$, such that,

$$\mathrm{Var}(nA_{\lfloor nt \rfloor}^{(n)}) = \sum_{k=n}^{n-\lfloor nt \rfloor+1} n^2(\frac{1}{\binom{k}{2}})^2 \le \frac{K}{n}.$$

So we deduce that

$$nA_{\lfloor nt \rfloor}^{(n)} \overset{L^2}{\to} 2(\frac{1}{1-t} - 1) = \frac{2t}{1-t} := f(t)$$

as n converges to ∞.

We denote by $f^{-1}(t) := t/(t+2)$ the inverse function of $f(t)$.

Similarly, $R^{(n)}$ is decreasing, so

$$\mathbb{P}(R^{(n)}_{A_{\lfloor nf^{-1}(t-\delta)^{(n)} \rfloor}} \le R^{(n)}_{tn^{-1}} \le R^{(n)}_{A_{\lfloor nf^{-1}(t+\delta)^{(n)} \rfloor}}) \to 1,$$

as n tends to ∞ for any $0 < \delta < t$.

So $\frac{R^{(n)}_{tn^{-1}}}{n} - \frac{R^{(n)}_{A^{(n)}_{\lfloor nf^{-1}(t) \rfloor}}}{n} \overset{d}{\to} 0$.

Furthermore,

$$\frac{R^{(n)}_{A^{(n)}_{nf^{-1}(t)}}}{n} = \frac{Y^{(n)}_{\lfloor nf^{-1}(t) \rfloor}}{n} = \frac{n - \lfloor nf^{-1}(t) \rfloor}{n} \to 1 - f^{-1}(t) = \frac{1}{1+t/2},$$

as n tends to ∞. So $\frac{R^{(n)}_{tn^{-1}}}{n} \overset{d}{\to} \frac{1}{1+t/2}$.

Using again the decreasing property of $R^{(n)}_t$, we finish the proof. □

2.5 The length of an external branch picked at random

Dynamics of any exchangeable coalescent with multiple mergers are characterized by rates $\lambda_{b,k}$ which suit a consistent relationship (this is Pitman's structure theorem, see (30), Lemma 18):

$$\lambda_{b,k} = \lambda_{b+1,k+1} + \lambda_{b+1,k}. \tag{2.26}$$

This relationship comes from the fact that k given merging blocks among b can coalesce in two ways while revealing an extra block : either the coalescence event implies the extra block (and then $k+1$ blocks will merge) either not. Thus we get a recursive construction of the n-coalescent process $\Pi^{(n)}$.

Let us define $\Pi^{(n,2)}$ as the coalescent process of individuals labelled from 2 to n. Now we consider the individual labelled by 1. The lineage of this individual can be 'connected' to $\Pi^{(n,2)}$

— either at any of its jump times, in which case block $\{1\}$ participates to a multiple merger implying at least 3 blocks, and we call this collision "Type 1" (see Figure 2.1),

— or at any other time to one of the present blocks and then participates to a binary collision, and we call it "Type 2" (see Figure 2.2).

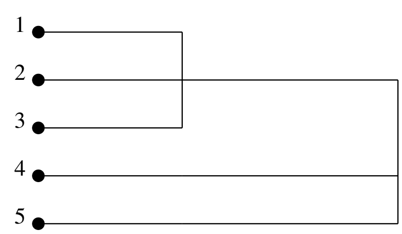


Figure 2.1: $n = 5$. Individual 1 is chosen. Type 1: individual 1 encounters a multiple collision.

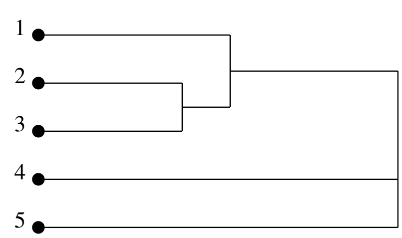


Figure 2.2: $n = 5$. Individual 1 is chosen. Type 2: individual 1 encounters a binary collision

From now on, our analysis is conditional on $\Pi^{(n,2)}$. Between two jump times of $\Pi^{(n,2)}$, assuming that there are b blocks in $\Pi^{(n,2)}$, the extra block coalesces at rate $b\lambda_{b+1,2}$. If the extra block remains unconnected just before a coalescence event implying k blocks among b, then it will participate to this event with probability

$$\frac{\binom{b}{k}\int_0^1 x^{k+1}(1-x)^{b-k}\nu(dx)/g_b}{\binom{b}{k}\int_0^1 x^{k}(1-x)^{b-k}\nu(dx)/g_b} = 1 - \frac{\lambda_{b+1,k}}{\lambda_{b,k}}. \tag{2.27}$$

This equality comes from (2.26). Let us see how to get the law of $T^{(n)}$, the coalescence time of individual 1. We define by $R^{(n,2)}$ the block counting process of $\Pi^{(n,2)}$. Notice that it has the same law as $R^{(n-1)}$. We introduce

— $T_c^{(n)}$ the first jump time of a Poisson process $\eta_c^{(n)}$ directed by the measure $\nu_c^{(n)} = R_t^{(n,2)}\lambda_{R_t^{(n,2)}+1,2}dt$;

— $T_d^{(n)}$ the time of the first appearance of 'Head' in the following coin flip, independent of $\eta_c^{(n)}$: at each jump time t of $R_t^{(n,2)}$, we toss a coin, and get 'Head' with probability $1-\frac{\lambda_{R_{t-}^{(n,2)}+1,R_{t-}^{(n,2)}-R_t^{(n,2)}+1}}{\lambda_{R_{t-}^{(n,2)},R_{t-}^{(n,2)}-R_t^{(n,2)}+1}}$ and 'Tail' with probability $\frac{\lambda_{R_{t-}^{(n,2)}+1,R_{t-}^{(n,2)}-R_t^{(n,2)}+1}}{\lambda_{R_{t-}^{(n,2)},R_{t-}^{(n,2)}-R_t^{(n,2)}+1}}$ (see (2.27)).

Then, conditionally on $\Pi^{(n,2)}$, $T^{(n)}$ and $T_c^{(n)}\wedge T_d^{(n)}$ have the same law.

Remark 2.1. *A more formal way to interpret $T^{(n)}$ is as follow. Let $\xi^{(n)}$ be Cox process directed by random measure $\nu_c^{(n)}+\nu_d^{(n)}$, where $\nu_d^{(n)} = \sum_{\{t \text{ is a jump time}\}} \frac{\lambda_{R_{t-}^{(n,2)}+1,R_{t-}^{(n,2)}-R_t^{(n,2)}+1}}{\lambda_{R_{t-}^{(n,2)},R_{t-}^{(n,2)}-R_t^{(n,2)}+1}}\delta_t$, and δ_t is the Dirac measure in t (see (24, p.226)). Then $T^{(n)}$ has the same law as the first jump time of $\xi^{(n)}$*

Let us now give our main result

Theorem 2.4. *The following convergence holds :*

$$n^{\alpha-1}T^{(n)} \xrightarrow{d} T = \frac{1}{C_0\Gamma(2-\alpha)}((1-\sigma)^{1-\alpha}-1),$$

for $n\to\infty$. The density function of T is

$$f_T(t) = \frac{\alpha C_0\Gamma(2-\alpha)}{\alpha-1}(1+C_0\Gamma(2-\alpha)t)^{-\frac{\alpha}{\alpha-1}-1}, \quad t\geq 0.$$

In particular, in the $Beta(2-\alpha,\alpha)$ case, the density is

$$f_T(t) = \frac{1}{(\alpha-1)\Gamma(\alpha)}(1+\frac{t}{\alpha\Gamma(\alpha)})^{-\frac{\alpha}{\alpha-1}-1}, \quad t\geq 0.$$

Proof. For the sake of simplicity, we will make the proof only in the $Beta(2-\alpha,\alpha)$ case. The proof can be extended to the more general case where (2.4) is satisfied with the details omited here. In this special case, $C_0 = (\alpha\Gamma(\alpha)\Gamma(2-\alpha))^{-1}$ and dynamics are given by

$$\lambda_{b,k} = \frac{B(k-\alpha,b-k+\alpha)}{B(\alpha,2-\alpha)},$$

where $B(a,b)$ is a Beta function of parameters a and b.

Define $r_t^{(n,2)}$ as the number of jumps of the process $\Pi^{(n,2)}$ up to time $n^{1-\alpha}t$. It is a straightforward consequence of Proposition 2.1 that

$$\frac{r_t^{(2,n)}}{n} \xrightarrow{\mathbb{P}} r(t), \quad n \to \infty \tag{2.28}$$

for $t \to \infty$, where $r(t)$ is defined in (2.24).

For $i \geq 0$, in the process $\Pi^{(n,2)}$, we denote by $Y_i^{(n,2)}$ the number of blocks remaining after i jumps which equals 1 from the time all individuals are coalesced to 1, and $Y_0^{(n,2)} = n-1$. Let $X_i^{(n,2)} = Y_{i-1}^{(n,2)} - Y_i^{(n,2)}$ be the number of blocks we lose during the ith coalescent event. We write $X_0^{(n)} = 0$. Notice that $(Y^{(n,2)}, X^{(n,2)})$ has the same law as $((Y^{(n-1)}, X^{(n-1)}))$.

Using the description given above, we have

$$\begin{aligned}
& \mathbb{P}(n^{\alpha-1}T^{(n)} > t) \\
= \ & \mathbb{E}[\mathbb{P}(n^{\alpha-1}T_c^{(n)} \wedge T_d^{(n)} > t | \Pi^{(n,2)})] \\
= \ & \mathbb{E}[\mathbb{P}(n^{\alpha-1}T_c^{(n)} > t | \Pi^{(n,2)})\mathbb{P}(n^{\alpha-1}T_d^{(n)} > t | \Pi^{(n,2)})] \\
= \ & \mathbb{E}[\exp(-\int_0^t \int_0^1 n^{1-\alpha} R^{(n,2)}_{sn^{1-\alpha}} x^2 (1-x)^{R^{(n,2)}_{sn^{1-\alpha}}-1} \nu(dx) ds) \prod_{i=1}^{r_t^{(n)}} \frac{\lambda_{1+Y_{i-1}^{(n,2)}, 1+X_i^{(n,2)}}}{\lambda_{Y_{i-1}^{(n,2)}, 1+X_i^{(n,2)}}}] \\
= \ & \mathbb{E}[\exp(-\int_0^t n^{1-\alpha} R^{(n,2)}_{sn^{1-\alpha}} \frac{B(2-\alpha, R^{(n,2)}_{sn^{1-\alpha}} + \alpha - 1)}{B(2-\alpha,\alpha)} ds) \prod_{i=1}^{r_t^{(n)}} \frac{Y_{i-1}^{(n,2)} - X_i^{(n,2)} + \alpha - 1}{Y_{i-1}^{(n,2)}}].
\end{aligned}$$

We decompose the term in the expectation into two parts: the exponential on one side and the product on the other.

Let us first look at the exponential term. Using Stirling's formula we get that, for $0 \leq s \leq t$,

$$n^{1-\alpha} R^{(n,2)}_{sn^{1-\alpha}} \frac{B(2-\alpha, R^{(n,2)}_{sn^{1-\alpha}} + \alpha - 1)}{B(2-\alpha,\alpha)} = n^{1-\alpha} \frac{(R^{(n,2)}_{sn^{1-\alpha}})^{\alpha-1}}{\Gamma(\alpha)} + (\frac{R^{(n,2)}_{sn^{1-\alpha}}}{n})^{\alpha-1} f(R^{(n,2)}_{sn^{1-\alpha}}),$$

where $f = f(t)_{\{t\geq 0\}}$ is a deterministic function which converges to 0 as t converges to ∞. The sequence $(R^{(n,2)}_{sn^{1-\alpha}}, n \geq 2)$ is decreasing so, thanks to Theorem 2.2, we deduce that $\sup_{0\leq s\leq t}(\frac{R^{(n,2)}_{sn^{1-\alpha}}}{n})^{\alpha-1} f(R^{(n,2)}_{sn^{1-\alpha}})$ converges in probability to 0 as n tends to ∞. Consequently, using again Theorem 2.2, we get that

$$\exp(-\int_0^t n^{1-\alpha} R^{(n,2)}_{sn^{1-\alpha}} \frac{B(2-\alpha, R^{(n,2)}_{sn^{1-\alpha}} + \alpha - 1)}{B(2-\alpha,\alpha)} ds) \xrightarrow{\mathbb{P}} (1 + \frac{t}{\alpha\Gamma(\alpha)})^{-\alpha}, \quad n \to \infty. \tag{2.29}$$

Convergence of the product term is obtained by the same method as in proof of Theorem 2.1, combined with the convergence in (2.28). To avoid showing almost the same reasoning, we leave the details to readers. This way, we have

$$\prod_{i=1}^{r_t^{(n,2)}} \frac{Y_{i-1}^{(n,2)} - X_i^{(n,2)} + \alpha - 1}{Y_{i-1}^{(n,2)}} \xrightarrow{\mathbb{P}} (1 + \frac{t}{\alpha\Gamma(\alpha)})^{-\frac{\alpha(2-\alpha)}{\alpha-1}}, \quad n \to \infty. \tag{2.30}$$

The product of (2.29) and (2.30) then converges in probability to $(1 + \frac{t}{\alpha\Gamma(\alpha)})^{-\alpha/(\alpha-1)}$. Since this product is bounded, we get that

$$\mathbb{P}(n^{\alpha-1}T^{(n)} > t) \to (1 + \frac{t}{\alpha\Gamma(\alpha)})^{-\frac{\alpha}{\alpha-1}}, \quad n \to \infty.$$

We achieve the proof. □

As a consequence of Theorem 2.4, we can get an asymptotic result on the size of the population at the moment of collision of individual 1.

Corollary 2.2. *The following convergence holds :*

$$n^{-1}Y_{\sigma^{(n)}}^{(n)} \xrightarrow{d} (1 + C_0\Gamma(2-\alpha)T)^{-1/(\alpha-1)} = 1 - \sigma$$

for $n \to \infty$. Moreover, the density function of this limit is $\alpha x^{\alpha-1}\mathbf{1}_{\{0 \le x \le 1\}}$.

Proof. In terms of block counting process, we have $Y_{\sigma^{(n)}}^{(n)} = R_{T^{(n)}}^{(n)}$. Notice that $R_{T^{(n)}}^{(n)} = R_{n^{1-\alpha}(n^{\alpha-1}T^{(n)})}^{(n)}$. Using Theorem 2.4, we known that $n^{\alpha-1}T^{(n)}$ converges in distribution to T. Hence, if $t_0 > 0$, we deduce from Theorem 2.2 that

$$\mathbf{1}_{\{n^{\alpha-1}T^{(n)} < t_0\}} \frac{R_{n^{1-\alpha}(n^{\alpha-1}T^{(n)})}^{(n)}}{n} \xrightarrow{d} \mathbf{1}_{\{T < t_0\}}(1 + C_0\Gamma(2-\alpha)T)^{-1/(\alpha-1)}.$$

This achieves the proof. □

2.6 An alternative proof for Theorem 2.4

In this section, we present an alternative proof for Theorem 2.4 using the convergence results for $\sigma^{(n)}$ from Theorem 2.1. First, we need a stronger version of Proposition 2.1 which gives weak convergence in the path space. Recall that

$$A_k^{(n)} = \sum_{i=1}^{k\wedge\tau^{(n)}} \frac{e_i}{g_{Y_{i-1}^{(n)}}},$$

where the e_i's are independent standard exponential variables.

Proposition 2.2. *We assume that $\rho(t) = C_0 t^{-\alpha} + O(t^{-\alpha+\zeta})$ for some $C_0 > 0$ and $\zeta > 1 - 1/\alpha$. Then, for any $t < \alpha - 1$, we have*

$$(n^{\alpha-1} A^{(n)}_{\lfloor ns \rfloor})_{s \le t} \xrightarrow{d} \left(\frac{1}{C_0 \Gamma(2-\alpha)}((1 - \frac{s}{\alpha-1})^{1-\alpha} - 1)\right)_{s \le t}, \tag{2.31}$$

in the sense of convergence in the path space $D[0,t]$ for $n \to \infty$.

Proof. Note that Theorem 2.1 states

$$n^{\alpha-1} A^{(n)}_{\lfloor ns \rfloor} \xrightarrow{\mathbb{P}} \left(\frac{1}{C_0 \Gamma(2-\alpha)}((1 - \frac{s}{\alpha-1})^{1-\alpha} - 1)\right),$$

for $0 < s < \alpha - 1$ and $n \to \infty$. So for every fixed $s \in [0,t]$, we have pointwise convergence in probability in (2.31). This implies weak convergence of all finite dimensional distributions due to the subsequence criterion for weak convergence. In order to show weak convergence in the path space, we will show tightness for the distributions from (2.31). Since the limit process is continuous, it suffices to show that the condition (i) of (5, Theorem 7.3) and condition (7.12) from (5, Corollary 7.4) are fulfilled (see (5, Corollary 13.4)). For the present processes, these conditions translate to showing that for every $\epsilon > 0$ and $\eta > 0$,

(i) there exists $a > 0$ s.t. $P(n^{\alpha-1} A^{(n)}_{\lfloor 0 \rfloor} \ge a) \le \eta$ for n big enough and

(ii) there exists a $0 < \delta < 1$ so that

$$\delta^{-1} P\left(n^{\alpha-1}(A^{(n)}_{\lfloor n \cdot \min(t_1+\delta, t) \rfloor} - A^{(n)}_{\lfloor n \cdot t_1 \rfloor}) \ge \epsilon\right) \le \eta,$$

for n big enough and any $t_1 \in [0,t]$.

Condition (i) is trivially fulfilled, for condition (ii) we can use Theorem 2.1 to show that for $n \to \infty$,

$$P\left(n^{\alpha-1}(A^{(n)}_{\lfloor n \cdot \min(t_1+\delta, t) \rfloor} - A^{(n)}_{\lfloor n \cdot t_1 \rfloor}) \ge \epsilon\right) \to P(f(\min(t_1+\delta, t)) - f(t_1) \ge \epsilon),$$

where $f(s) := \frac{1}{C_0\Gamma(2-\alpha)}((1 - \frac{s}{\alpha-1})^{1-\alpha} - 1)$. Note that $P(f(\min(t_1+\delta,t)) - f(t_1) \ge \epsilon) \le P(f(t) - f(t-\delta) \ge \epsilon) \in \{0, 1\}$. Since f is continuous, you can now choose δ small enough that $f(t) - f(t-\delta) < \epsilon$ and then n big enough to fulfill (ii). Thus, we have shown tightness of the distributions in (2.31) which establishes the desired weak convergence □

Now we come to the alternative proof of Theorem 2.4.

Alternative proof of Theorem 2.4. Fix $t \in [0, \alpha-1)$. We have

$$(\sigma^{(n)}/(n(\alpha-1)), (n^{\alpha-1} A^{(n)}_{\lfloor ns \rfloor})_{s \le t}) \xrightarrow{d} \left(\sigma, \left(\frac{1}{C_0\Gamma(2-\alpha)}((1 - \frac{s}{\alpha-1})^{1-\alpha} - 1)\right)_{s \le t}\right),$$

for $n \to \infty$. Due to Skorohod-coupling, we can assume that this convergence also holds almost surely. Since $s \mapsto (\frac{1}{C_0\Gamma(2-\alpha)}((1 - \frac{s}{\alpha-1})^{1-\alpha} - 1))$ is continuous on $[0,t]$, the almost sure convergence of $(n^{\alpha-1}A^{(n)}_{\lfloor nt\rfloor})_{s\leq t}$ in $D[0,t]$ is even almost sure uniform convergence on $[0,t]$ (see (5, p. 124)). For any series $(x_n)_{n\in\mathbb{N}}$ on [0,t] with $x_n \to x$, we thus have $n^{\alpha-1}A^{(n)}_{\lfloor nx_n\rfloor} \to \frac{1}{C_0\Gamma(2-\alpha)}((1-\frac{x}{\alpha-1})^{1-\alpha}-1)$ almost surely for $n\to\infty$. The only problem left is that $\sigma^{(n)}$, σ may take values in $[0,\alpha-1)$ and not only in some subset $[0,t]$. To remedy this, note that if we restrict all random variables on $\{\sigma\leq\alpha-1-\frac{2}{k}\}$ for $k\in\mathbb{N}$, we have $\sigma^{(n)}(\omega)/n\leq\alpha-1-\frac{1}{k}$ for $n=n(\omega)$ big enough for almost all $\omega\in\{\sigma\leq\alpha-1-\frac{1}{k}\}$. Thus, by using the Skorohod-coupling for the series $(\sigma^{(n)}/(n(\alpha-1)),(n^{\alpha-1}A^{(n)}_{\lfloor ns\rfloor})_{s\leq\alpha-1-\frac{1}{k}})$, we have

$$n^{\alpha-1}A^{(n)}_{\lfloor n\cdot\sigma^{(n)}/n\rfloor}\xrightarrow{d}\frac{1}{C_0\Gamma(2-\alpha)}((1-\frac{\sigma}{\alpha-1})^{1-\alpha}-1),$$

almost surely on $\{\sigma\leq\alpha-1-\frac{2}{k}\}$ for the coupled versions of these random variables (note that $\sigma^{(n)}\leq\tau^{(n)}$). Since σ is Beta-distributed, we have, for $k\to\infty$,

$$P(\{n^{\alpha-1}A^{(n)}_{\lfloor n\cdot\sigma^{(n)}/n\rfloor}\in\cdot\}\cap\{\sigma\leq\alpha-1-\frac{2}{k}\})\sim P(n^{\alpha-1}A^{(n)}_{\lfloor n\cdot\sigma^{(n)}/n\rfloor}\in\cdot).$$

This shows

$$n^{\alpha-1}T^{(n)}=n^{\alpha-1}A^{(n)}_{\sigma^{(n)}}\xrightarrow{d}\frac{1}{C_0\Gamma(2-\alpha)}((1-\frac{\sigma}{\alpha-1})^{1-\alpha}-1).$$

The only thing left to prove is that T has density f_T. This is done by computing the distribution function

$$\begin{aligned}
&\mathbb{P}(\frac{1}{C_0\Gamma(2-\alpha)}((1-\frac{\sigma}{\alpha-1})^{1-\alpha}-1)\leq t)\\
=\;&\mathbb{P}(\sigma\leq 1-(1+C_0\Gamma(2-\alpha)t)^{\frac{1}{\alpha-1}})\\
=\;&\int_0^{1-(1+C_0\Gamma(2-\alpha)t)^{\frac{1}{\alpha-1}}}\alpha(1-x)^{\alpha}dx\\
=\;&1-(1+C_0\Gamma(2-\alpha)t)^{-\frac{\alpha}{\alpha-1}}.
\end{aligned}$$

and finally by differentiating. □

Acknowledgment

Jean-Stéthane Dhersin and Linglong Yuan benefited from the support of the "Agence Nationale de la Recherche": ANR MANEGE (ANR-09-BLAN-0215).

Arno Siri-Jégousse and Fabian Freund would like to thank the Hausdorff Research Institute for Mathematices (Bonn, Germany) for its support. Fabian Freund would like to thank Luca Ferretti for a discussion concerning the interpretation of $\sigma^{(n)}$.

Bibliographie

[1] Árnason, E. (2004). Mitochondrial cytochrome b variation in the high-fecundity atlantic cod : trans-atlantic clines and shallow gene genealogy. *Genetics.*, 15:166 : 1871–1885.

[2] Berestycki, J., Berestycki, N. et Schweinsberg, J. (2007). Beta-coalescents and continuous stable random trees. *Ann. Probab.*, 35(5):1835–1887.

[3] Berestycki, J., Berestycki, N. et Schweinsberg, J. (2013). The genealogy of branching brownian motion with absorption. *The Annals of Probability*, 41(2):527–618.

[4] Bertoin, J. (1996). *Lévy processes*, volume 121 de *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge.

[5] Billingsley, P. (1999). *Convergence of probability measures*. Wiley Series in Probability and Statistics : Probability and Statistics. John Wiley &amp ; Sons Inc., New York, second édition. A Wiley-Interscience Publication.

[6] Birkner, M. et Blath, J. (2008). Computing likelihoods for coalescents with multiple collisions in the infinitely-many-sites model. *Journal of Mathematical Biology*.

[7] Birkner, M., Blath, J., Capaldo, M., Etheridge, A., Möhle, M., Schweinsberg, J. et Wakolbinger, A. (2005). Alpha-stable branching and beta-coalescents. *Electron. J. Probab.*, 10:no. 9, 303–325 (electronic).

[8] Blum, M. G. B. et François, O. (2005). Minimal clade size and external branch length under the neutral coalescent. *Adv. in Appl. Probab.*, 37(3):647–662.

[9] Bolthausen, E. et Sznitman, A. (1998). On ruelle's probability cascades and an abstract cavity method. *Communications in mathematical physics*, 197(2):247–276.

[10] Boom, J., Boulding, E. et Beckenbach, A. (1994). Mitochondrial DNA variation in introduced populations of Pacific oyster, Crassostrea gigas, in British Columbia. *Canadian journal of fisheries and aquatic sciences(Print)*, 51(7):1608–1614.

[11] Bovier, A. et Kurkova, I. (2007). Much ado about derrida's grem. *Spin glasses*, pages 81–115.

[12] Brunet, É. et Derrida, B. (2013). Genealogies in simple models of evolution. *Journal of Statistical Mechanics : Theory and Experiment*, 2013(01):P01006.

[13] Caliebe, A., Neininger, R., Krawczak, M. et Roesler, U. (2007). On the length distribution of external branches in coalescence trees : genetic diversity within species. *Theoretical Population Biology*, 72(2):245–252.

[14] Delmas, J.-F., Dhersin, J.-S. et Siri-Jégousse, A. (2008). Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.*, 18(2):997–1025.

[15] Eldon, B. et Wakeley, J. (2006). Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics*, 172:2621–2633.

[16] Freund, F. et Möhle, M. (2009). On the time back to the most recent common ancestor and the external branch length of the Bolthausen-Sznitman coalescent. *Markov Process. Related Fields*, 15(3):387–416.

[17] Fu, Y. et Li, W. (1993). Statistical tests of neutrality of mutations. *Genetics*, 133:693–709.

[18] Gnedin, A., Iksanov, A. et Möhle, M. (2008). On asymptotics of exchangeable coalescents with multiple collisions. *J. Appl. Probab.*, 45(4):1186–1195.

[19] Gnedin, A. et Yakubovich, Y. (2007). On the number of collisions in Λ-coalescents. *Electron. J. Probab.*, 12:no. 56, 1547–1567 (electronic).

[20] Goldschmidt, C. et Martin, J. B. (2005). Random recursive trees and the Bolthausen-Sznitman coalescent. *Electron. J. Probab.*, 10:no. 21, 718–745 (electronic).

[21] Hedgecock, D. (1994). Does variance in reproductive success limit effective population sizes of marine organisms ? *In Genetics and Evolution of Aquatic Organisms*, pages 1222–1344. Chapman and Hall, London.

[22] Iksanov, A. et Möhle, M. (2008). On the number of jumps of random walks with a barrier. *Advances in Applied Probability*, 40(1):206–228.

[23] Janson, S. et Kersting, G. (2011). The external lengths in kingman's coalescent.

[24] Kallenberg, O. (2002). *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second édition.

[25] Kingman, J. F. C. (1982a). The coalescent. *Stochastic Process. Appl.*, 13(3):235–248.

[26] Kingman, J. F. C. (1982b). On the genealogy of large populations. *J. Appl. Probab.*, (Special Vol. 19A):27–43. Essays in statistical science.

[27] Möhle, M. (2000). Total variation distances and rates of convergence for ancestral coalescent processes in exchangeable population models. *Adv. in Appl. Probab.*, 32(4): 983–993.

[28] Möhle, M. (2010). Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson-Dirichlet coalescent. *Stochastic Process. Appl.*, 120(11):2159–2173.

[29] Neher, R. A. et Hallatschek, O. (2013). Genealogies of rapidly adapting populations. *Proceedings of the National Academy of Sciences*, 110(2):437–442.

[30] Pitman, J. (1999). Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902.

[31] Sagitov, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 36(4):1116–1125.

[32] Schweinsberg, J. (2003). Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Process. Appl.*, 106(1):107–139.

Chapitre 3

Sur la longueur totale des branches externes de Beta-coalescente

Version non modifiée de l'article *On the total length of the external branches for Beta-coalescent.*

soumis à

Journal of Applied Probability.

3.1 Introduction

3.1.1 Motivation

In a Wright-Fisher haploid population model with size N, we sample n individuals at present from the total population, and look backward to see the ancestral tree until we get the most recent common ancestor (MRCA). If time is well rescaled and the size N of population becomes large, then the genealogy of the sample of size n converges weakly to the Kingman n-coalescent (see (33),(34)). During the evolution of the population, mutations may occur. We consider the infinite sites model introduced by (32). In this model, each mutation is produced at a new site which is never seen before and will never be seen in the future. The neutrality of mutations means that all mutants are equally privileged by the environment. Under the infinite sites model, to detect or reject the neutrality when the genealogy is given by the Kingman coalescent, (22) have proposed a statistical test based on the total mutation numbers on the external branches and internal branches. Mutations on external branches affect only single individuals, so in practice they can be picked out according to the model setting. In this test, the ratio $L_{ext}^{(n)}/L^{(n)}$ between the total external branch length $L_{ext}^{(n)}$ and the total length $L^{(n)}$ measures in some sense the weight of mutations occurred on external branches among all. It then makes the study of these quantities relevant.

For many populations, Kingman coalescent describes the genealogy quite well. But for some others, when descendants of one individual can occupy a big ratio of the next generation with non-negligible probability, it is no more relevant. It is for example the case of some marine species (see (1), (9), (19), (24), (26)). In this case, if time is well rescaled and the size of population becomes large, the ancestral tree converges weakly to the Λ-coalescent which is associated with a finite measure Λ on $[0, 1]$. This coalescent allows multiple collisions. It has first been introduced by (38) and(39). Among Λ-coalescents, a special and important subclass is called Beta(a, b)-coalescents characterized by Λ being a Beta distribution $Beta(a, b)$. The most popular ones are those with parameters $2-\alpha$ and α where $\alpha \in (0, 2)$.

Beta-coalescents arise not only in the context of biology. They also have connections with supercritical Galton-Watson process (see (40)), with continuous-state branching processes (see (6), (2), (20)), with continuous random trees (see (4)). If $\alpha = 1$, we recover the Bolthausen-Sznitman coalescent which appears in the field of spin glasses (see (8), (10)) and is also connected to random recursive trees (see (25)). The Kingman coalescent is also obtained from the $Beta(2-\alpha, \alpha)$-coalescent by letting α tend to 2.

For $Beta(2-\alpha, \alpha)$-coalescents with $1 < \alpha < 2$, a central limit theorem of the total external branch length $L_{ext}^{(n)}$ is known (see (31)). The aim of this paper is to study its moments. The results obtained can be extended to more general coalescent processes (see (16)). We should say that in this case, the moment method is not able to obtain the right convergence speed in the central limit theorem, which illustrates some limitations of moment calculations.

3.1.2 Introduction and main results

Let $\mathcal{E}$ be the set of partitions of $\mathbb{N} := \{1, 2, 3, ...\}$ and, for $n \in \mathbb{N}$, $\mathcal{E}_n$ be the set of partitions of $\mathbb{N}_n := \{1, 2, \cdots, n\}$. We denote by $\rho^{(n)}$ the natural restriction on $\mathcal{E}_n$: if $1 \leq n \leq m \leq +\infty$ and $\pi = \{A_i\}_{i\in I}$ is a partition of $\mathbb{N}_m$, then $\rho^{(n)}\pi$ is the partition of $\mathbb{N}_n$ defined by $\rho^{(n)}\pi = \{A_i \bigcap \mathbb{N}_n\}_{i\in I}$. For a finite measure Λ on $[0, 1]$, we denote by $\Pi = (\Pi_t)_{t\geq 0}$ the Λ-coalescent process introduced independently by (38) and (39). The process $(\Pi_t)_{t\geq 0}$ is a càd-làg continuous time Markovian process taking values in $\mathcal{E}$ with $\Pi_0 = \{\{1\}, \{2\}, \{3\}, ...\}$. It is characterized by the càd-làg Λ n-coalescent processes $(\Pi_t^{(n)})_{t\geq 0} := (\rho^{(n)}\Pi_t)_{t\geq 0}, n \in \mathbb{N}$. For $n \leq m \leq +\infty$, we have $(\Pi_t^{(n)})_{t\geq 0} = (\rho^{(n)}\Pi_t^{(m)})_{t\geq 0}$ (where $\Pi^{(+\infty)} = \Pi$).

Let $\nu(dx) = x^{-2}\Lambda(dx)$. For $2 \leq a \leq b$, we set

$$\lambda_{b,a} = \int_0^1 x^{a-2}(1-x)^{b-a}\Lambda(dx) = \int_0^1 x^a(1-x)^{b-a}\nu(dx).$$

$\Pi^{(n)}$ is a Markovian process with values in $\mathcal{E}_n$, and its transition rates are given by: for $\xi, \eta \in \mathcal{E}_n$, $q_{\xi,\eta} = \lambda_{b,a}$ if η is obtained by merging a of the $b = |\xi|$ blocks of ξ and letting the $b - a$ others unchanged, and $q_{\xi,\eta} = 0$ otherwise. We say that a individuals (or blocks) of ξ have been coalesced in one single individual of η. Remark that the process $\Pi^{(n)}$ is an exchangeable process, which means that, for any permutation τ of $\mathbb{N}_n$, $\tau \circ \Pi^{(n)} \overset{(d)}{=} \Pi^{(n)}$.

The process $\Pi^{(n)}$ finally reaches one block. This final individual is called the most recent common ancestor (MRCA). We denote by $\tau^{(n)}$ the number of collisions it takes for the n individuals to be coalesced to the MRCA.

We define by $R^{(n)} = (R_t^{(n)})_{t\geq 0}$ the block counting process of $(\Pi_t^{(n)})_{t\geq 0}$: $R_t^{(n)} = |\Pi_t^{(n)}|$, which equals the number of blocks/individuals at time t. Then $R^{(n)}$ is a continuous time Markovian process taking values in $\mathbb{N}_n$, decreasing from n to 1. At state b, for $a = 2, ..., b$, each of the $\binom{b}{a}$ groups with a individuals coalesces independently at rate $\lambda_{b,a}$. Hence, the time the process $(R_t^{(n)})_{t\geq 0}$ stays at state b is exponential with parameter:

$$g_b = \sum_{a=2}^{b} \binom{b}{a} \lambda_{b,a} = \int_0^1 (1 - (1-x)^b - bx(1-x)^{b-1})\nu(dx) = b(b-1)\int_0^1 t(1-t)^{b-2}\rho(t)dt, \tag{3.1}$$

where $\rho(t) = \int_t^1 \nu(dx)$. We denote by $Y^{(n)} = (Y_k^{(n)})_{k\geq 0}$ the discrete time Markov chain associated with $R^{(n)}$. This is a decreasing process from $Y_0^{(n)} = n$ which reaches 1 at the $\tau^{(n)}$-th jump. The probability transitions of the Markov chain $Y^{(n)}$ are given by: for $b \geq 2$, $k \geq 1$ and $1 \leq l \leq b-1$,

$$p_{b,b-l} := \mathbb{P}(Y_k^{(n)} = b - l | Y_{k-1}^{(n)} = b) = \frac{\binom{b}{l+1}\lambda_{b,l+1}}{g_b}, \tag{3.2}$$

and 1 is an absorbing state.

We introduce the discrete time process $X_k^{(n)} := Y_{k-1}^{(n)} - Y_k^{(n)}$, $k \geq 1$ with $X_0^{(n)} = 0$. This process counts the number of blocks we lose at the k-th jump. For $i \in \{1, \ldots, n\}$, we define

$$T_i^{(n)} := \inf \left\{ t | \, \{i\} \notin \Pi_t^{(n)} \right\}$$

as the length of the i-th external branch and $T^{(n)}$ the length of a randomly chosen external branch. By exchangeability, $T_i^{(n)} \stackrel{(d)}{=} T^{(n)}$. We denote by $L_{ext}^{(n)} := \sum_{i=1}^n T_i^{(n)}$ the total external branch length of $\Pi^{(n)}$, and by $L^{(n)}$ the total branch length.

For several measures Λ, many asymptotic results on the external branches and their total external lengths of the Λ n-coalescent are already known.

1. If $\Lambda = \delta_0$, Dirac measure on 0, $\Pi^{(n)}$ is the Kingman n-coalescent. Then,
   (a) $nT^{(n)}$ converges in distribution to T which is a random variable with density $f_T(x) = \frac{8}{(2+x)^3}\mathbf{1}_{x\geq 0}$ (See (7), (12), (27)).
   (b) $L_{ext}^{(n)}$ converges in L^2 to 2 (see (22), (18)). A central limit theorem is also proved in (27).
2. If Λ is the uniform probability measure on $[0, 1]$, $\Pi^{(n)}$ is the Bolthausen-Sznitman n-coalescent. Then $(\log n)T^{(n)}$ converges in distribution to an exponential variable with parameter 1 (see (21), (41)). For moment results of $L_{ext}^{(n)}$, we refer to (14) and for central limit theorem, we refer to (30).
3. If $\nu_{-1} = \int_0^1 x^{-1}\Lambda(dx) < +\infty$, which includes the case of the $Beta(2-\alpha, \alpha)$-coalescent with $0 < \alpha < 1$, then
   (a) $T^{(n)}$ converges in distribution to an exponential variable with parameter ν_{-1} (see (23; 37)).
   (b) $L^{(n)}/n$ converges in distribution to a random variable L whose distribution coincides with that of $\int_0^{+\infty} e^{-X_t}dt$, where X_t is a certain subordinator (see page 1405 in (17) and (36)), and $L_{ext}^{(n)}/L^{(n)}$ converges in probability to 1 (see (37)).
4. If Λ is the Beta$(2-\alpha, \alpha)$ measure with $1 < \alpha < 2$, then we get the $Beta(2-\alpha, \alpha)$-coalescents. Note that $n^{\alpha-1}T^{(n)}$ converges in distribution to T which is a random variable with density function (see(15))

$$f_T(x) = \frac{1}{(\alpha-1)\Gamma(\alpha)}(1 + \frac{x}{\alpha\Gamma(\alpha)})^{-\frac{\alpha}{\alpha-1}-1}\mathbf{1}_{x\geq 0}. \tag{3.3}$$

   For central limit theorems of $L_{ext}^{(n)}$ and $L^{(n)}$, we refer to (31; 29).

In the rest of the paper, we only consider the Beta$(2-\alpha,\alpha)$ coalescents, $1<\alpha<2$. In that case, we have

$$\nu(dx)=\frac{1}{\Gamma(\alpha)\Gamma(2-\alpha)}x^{-1-\alpha}(1-x)^{\alpha-1}\,dx.$$

T denotes a random variable with density (3.3). If $(a_n)_{n\geq1}$ and $(b_n)_{n\geq1}$ are two real sequences, we define $a_n\sim b_n$ when $\lim_{n\to+\infty}a_n/b_n=1$ is true.

Theorem 3.1. *1. The total external branch length $L_{ext}^{(n)}$ satisfies*

$$\lim_{n\to+\infty}n^{3\alpha-5}\mathbb{E}[(L_{ext}^{(n)}-n^{2-\alpha}\mathbb{E}[T])^2]=\Delta(\alpha),$$

where $\mathbb{E}[T]=\alpha(\alpha-1)\Gamma(\alpha)$ and $\Delta(\alpha)=\frac{((\alpha-1)\Gamma(\alpha+1))^2\Gamma(4-\alpha)}{(3-\alpha)\Gamma(4-2\alpha)}$.

2. As a consequence, $n^{\alpha-2}L_{ext}^{(n)}\overset{(L^2)}{\to}\mathbb{E}[T]$.

Remark 3.1. — *For the second part of the theorem, the convergence in probability and almost surely can be found from (4), (5), (3).*

— *The first part of the theorem gives $n^{(5-3\alpha)/2}$ as the convergence speed for $L_{ext}^{(n)}$ tending to $n^{2-\alpha}\mathbb{E}[T]$ in the sense of second moment. But as shown in (31),*

$$\frac{L_{ext}^{(n)}-n^{2-\alpha}\mathbb{E}[T]}{n^{1/\alpha+1-\alpha}}\overset{(d)}{\to}\frac{\alpha(2-\alpha)(\alpha-1)^{1/\alpha+1}\Gamma(\alpha)}{\Gamma(2-\alpha)^{1/\alpha}}\zeta,$$

where ζ is a stable random variable with parameter α. Our moment method fails to get the right speed of convergence in distribution.
To prove this result, the first idea is to write

$$\mathbb{E}[(L_{ext}^{(n)}-n^{2-\alpha}\mathbb{E}[T])^2]=n\mathrm{Var}(T_1^{(n)})+n(n-1)\mathrm{cov}(T_1^{(n)},T_2^{(n)})+(n\mathbb{E}[T_1^{(n)}]-n^{2-\alpha}\mathbb{E}[T])^2. \tag{3.4}$$

Hence we have to get results on the moments of the external branches. This is given by the next theorems. The first one gives the asymptotic behaviour for the covariance of two external branch lengths.

Theorem 3.2. *The asymptotic covariance of two external branch lengths is given by:*

$$\lim_{n\to+\infty}n^{3(\alpha-1)}\mathrm{cov}(T_1^{(n)},T_2^{(n)})=\frac{\int_0^1((1-x)^{2-\alpha}-1)^2\nu(dx)}{3-\alpha}((\alpha-1)\Gamma(\alpha+1))^3=\Delta(\alpha).$$

Remark 3.2. *$\Delta(\alpha)$ is the limit only in the case of Beta$(2-\alpha,\alpha)$-coalescents, but the result can be extended to more general Λ-coalescent (see (16)).*

Notice that $\Delta(\alpha)$ *is strictly positive implies that* $\mathrm{cov}(T_1^{(n)}, T_2^{(n)})$ *is of order* $n^{3-3\alpha}$ *and* $T_1^{(n)}$, $T_2^{(n)}$ *are positively correlated in the limit which is similar to Boltausen-Sznitman coalescent and opposite of Kingman coalescent (negatively correlated) (see (14)). To prove this theorem, we have to give the asymptotic behaviours of* $\mathbb{E}[T_1^{(n)} T_2^{(n)}]$ *and* $\mathbb{E}[T_1^{(n)}]$ *(Theorem 3.4). We also get from Theorem 3.4 that the third term in (3.4) satisfies*

$$(n\mathbb{E}[T_1^{(n)}] - n^{2-\alpha}\mathbb{E}[T])^2 = O(n^{6-4\alpha}). \tag{3.5}$$

The second one gives the asymptotic behaviour of moments of one external branch length, hence we can estimate $nVar(T_1^{(n)})$. We then see that $n(n-1)\mathrm{cov}(T_1^{(n)}, T_2^{(n)})$ is dominant in $\mathbb{E}[(L_{ext}^{(n)} - n^{2-\alpha}\mathbb{E}[T])^2]$ (see (3.4)). Then we can conclude for Theorem 3.1.

Theorem 3.3. *For Beta*$(2-\alpha, \alpha)$*-coalescent, we have*

1. *If* $0 \le \beta < \frac{\alpha}{\alpha-1}$, *then* $\lim_{n\to+\infty} \mathbb{E}[(n^{\alpha-1}T_1^{(n)})^\beta] = \mathbb{E}[T^\beta]$.
2. *If* $\beta \ge \frac{\alpha}{\alpha-1}$, *then* $\lim_{n\to+\infty} \mathbb{E}[(n^{\alpha-1}T_1^{(n)})^\beta] = +\infty$.

3.1.3 Organization of this paper

In sections 2 and 3, we give estimates of $\mathbb{E}[T_1^{(n)}]$ and $\mathbb{E}[T_1^{(n)} T_2^{(n)}]$ respectively. Both $\mathbb{E}[T_1^{(n)}]$ and $\mathbb{E}[T_1^{(n)} T_2^{(n)}]$ satisfy the same kind of recurrence which allows to get their estimates and they lead to an estimate of $\mathrm{cov}(T_1^{(n)}, T_2^{(n)})$ in section 3. The main tool is Lemma 3.6 given in appendix A. In section 4, we deal with Theorem 3.3. Section 5 is the appendix where are given some proofs omitted before.

3.2 First moment of $T_1^{(n)}$ by recursive method

3.2.1 Preliminaries

For $s > -\alpha$, we define the measure

$$\nu^{(s)}(dx) := (1-x)^s \nu(dx) = \frac{1}{\Gamma(\alpha)\Gamma(2-\alpha)} x^{-1-\alpha}(1-x)^{\alpha-1+s}\, dx; \tag{3.6}$$

The collision rates of the Λ-coalescent associated with the measure $\nu^{(s)}$ is given by

$$g_n^{(s)} := \int_0^1 (1-(1-x)^n - nx(1-x)^{n-1})\nu^{(s)}(dx) \sim \frac{n^\alpha}{\Gamma(\alpha+1)}$$

when n tends to ∞.

We introduce the quantity $\rho^{(s)}(t) := \int_t^1 \nu^{(s)}(dx)$.

Lemma 3.1. *For $s > -\alpha$, we have when t tends to 0:*

1. $\rho^{(s)}(t) = \frac{t^{-\alpha}}{\Gamma(\alpha+1)\Gamma(2-\alpha)} - \frac{(\alpha-1+s)t^{1-\alpha}}{(\alpha-1)\Gamma(\alpha)\Gamma(2-\alpha)} + o(t^{1-\alpha})$;
2. $\int_t^1 \rho^{(s)}(x)dx = \frac{t^{1-\alpha}}{(\alpha-1)\Gamma(\alpha+1)\Gamma(2-\alpha)} + \frac{\int_0^1 x^{-\alpha}((1-x)^{\alpha-1+s}-1)dx}{\Gamma(\alpha)\Gamma(2-\alpha)} - \frac{1}{(\alpha-1)\Gamma(\alpha)\Gamma(2-\alpha)} + O(t^{2-\alpha})$;
3. $\lim_{t\to 0+}\left(\int_t^1 \rho^{(s)}(x)dx - \frac{t^{1-\alpha}}{(\alpha-1)\Gamma(\alpha+1)\Gamma(2-\alpha)}\right)$ *exists, and its value is*

$$C^{(s)} = \frac{\int_0^1 x^{-\alpha}((1-x)^{\alpha-1+s}-1)dx}{\Gamma(\alpha)\Gamma(2-\alpha)} - \frac{1}{(\alpha-1)\Gamma(\alpha)\Gamma(2-\alpha)}.$$

In particular, if $s \geq 1-\alpha$, $C^{(s)} = \frac{\Gamma(\alpha+s)}{\Gamma(s+1)\Gamma(\alpha)(1-\alpha)}$.

Proof. The result for $\rho^{(s)}(t)$ is straightforward since

$$\rho^{(s)}(t) = \int_t^1 \frac{1}{\Gamma(\alpha)\Gamma(2-\alpha)} x^{-1-\alpha}(1-x)^{\alpha-1}\,dx.$$

For $\int_t^1 \rho^{(s)}(x)dx$, using integration by parts, we have

$$\begin{aligned}\int_t^1 \rho^{(s)}(x)dx &= -t\rho^{(s)}(t) + \frac{\int_t^1 x^{-\alpha}(1-x)^{\alpha-1+s}dx}{\Gamma(\alpha)\Gamma(2-\alpha)}\\ &= -\frac{t^{1-\alpha}}{\alpha\Gamma(\alpha)\Gamma(2-\alpha)} + \frac{\int_t^1(x^{-\alpha}(1-x)^{\alpha-1+s}-1)dx}{\Gamma(\alpha)\Gamma(2-\alpha)} + \frac{\int_t^1 x^{-\alpha}dx}{\Gamma(\alpha)\Gamma(2-\alpha)} + O(t^{2-\alpha})\\ &= \frac{t^{1-\alpha}}{(\alpha-1)\Gamma(\alpha+1)\Gamma(2-\alpha)} + \frac{\int_0^1 x^{-\alpha}((1-x)^{\alpha-1+s}-1)dx}{\Gamma(\alpha)\Gamma(2-\alpha)}\\ &\quad - \frac{1}{(\alpha-1)\Gamma(\alpha)\Gamma(2-\alpha)} + O(t^{2-\alpha}),\end{aligned}$$

which gives also the existence and the first definition of $C^{(s)}$.

If $s = 1-\alpha$, $C^{(s)} = \frac{1}{(1-\alpha)\Gamma(\alpha)\Gamma(2-\alpha)}$. If $s > 1-\alpha$, using again integration by parts obtains $C^{(s)} = \frac{\int_0^1 x^{-\alpha}((1-x)^{\alpha-1+s}-1)dx}{\Gamma(\alpha)\Gamma(2-\alpha)} - \frac{1}{(\alpha-1)\Gamma(\alpha)\Gamma(2-\alpha)} = \frac{\Gamma(\alpha+s)}{\Gamma(s+1)\Gamma(\alpha)(1-\alpha)}$. □

We then define two values $A := \displaystyle\int_0^1 ((1-x)^{1-\alpha} - 1 - (\alpha-1)x)\nu^{(1)}(dx)$, $B := \displaystyle\int_0^1 ((1-x)^{2(1-\alpha)} - 1 - 2(\alpha-1)x)\nu^{(2)}(dx)$, which will be used many times later.

Lemma 3.2. *If A, B are defined as above, then $A = \alpha(\alpha^2-\alpha-1)\Gamma(\alpha-1)$ and $B = \frac{1}{(\alpha-1)}(\frac{\Gamma(4-\alpha)}{\Gamma(4-2\alpha)} + (\alpha^2-\alpha-1)\Gamma(\alpha+2))$.*

Proof. Using integration by parts two times,

$$
\begin{aligned}
&A\\
&= \frac{\alpha}{\Gamma(2-\alpha)}\frac{1}{\alpha(\alpha-1)}\int_0^1 x^{1-\alpha}\left(-\alpha(\alpha-1)(1-x)^{\alpha-2}+2\alpha(\alpha-1)(1-x)^{\alpha-1}-\alpha(\alpha-1)^2x(1-x)^{\alpha-2}\right)dx\\
&= \frac{1}{\Gamma(2-\alpha)(\alpha-1)}\left(-\Gamma(\alpha+1)\Gamma(2-\alpha)+2(\alpha-1)\Gamma(\alpha+1)\Gamma(2-\alpha)-(\alpha-1)\Gamma(3-\alpha)\Gamma(\alpha+1)\right)\\
&= \alpha(\alpha^2-\alpha-1)\Gamma(\alpha-1).
\end{aligned}
$$

In the same way, one gets $B = \frac{1}{(\alpha-1)}(\frac{\Gamma(4-\alpha)}{\Gamma(4-2\alpha)} + (\alpha^2-\alpha-1)\Gamma(\alpha+2))$.

□

3.2.2 The main result

Theorem 3.4.

$$\mathbb{E}[T_1^{(n)}] = (\alpha-1)\Gamma(\alpha+1)n^{1-\alpha} + \frac{(\alpha-1)^2(\Gamma(\alpha+1))^2}{2-\alpha}\left(A+(\alpha-1)C^{(1)}-C^{(0)}\right)n^{2(1-\alpha)} + o(n^{2(1-\alpha)}).$$

The idea is to use the recurrence satisfied by $\mathbb{E}[T_1^{(n)}]$(see (14)):

$$\mathbb{E}[T_1^{(n)}] = \frac{1}{g_n} + \sum_{k=2}^{n-1} p_{n,k}\frac{k-1}{n}\mathbb{E}[T_1^{(k)}]. \tag{3.7}$$

Let $L = (\alpha-1)\Gamma(\alpha+1)$ and $Q = \frac{(\alpha-1)^2(\Gamma(\alpha+1))^2}{2-\alpha}(A+(\alpha-1)C^{(1)}-C^{(0)})$. We transform the recurrence (3.7) to

$$
\begin{aligned}
\left(\mathbb{E}[n^{\alpha-1}T_1^{(n)}]-L\right)n^{\alpha-1} - Q = &\left(\frac{n^{\alpha-1}}{g_n} - (1-\sum_{k=2}^{n-1}p_{n,k}\frac{k-1}{n}(\frac{n}{k})^{\alpha-1})L\right)n^{\alpha-1}\\
&- Q(1-\sum_{k=2}^{n-1}p_{n,k}\frac{k-1}{n}(\frac{n}{k})^{2(\alpha-1)})\\
&+\sum_{k=2}^{n-1}(\frac{n}{k})^{2(\alpha-1)}p_{n,k}\frac{k-1}{n}\left(k^{\alpha-1}(\mathbb{E}[k^{\alpha-1}T_1^{(k)}]-L)-Q\right).
\end{aligned} \tag{3.8}
$$

Hence we get a recurrence

$$a_n = b_n + \sum_{k=2}^{n-1} q_{n,k}a_k, \tag{3.9}$$

with

$$a_n = \left(\mathbb{E}[n^{\alpha-1}T_1^{(n)}] - L\right) n^{\alpha-1} - Q,$$

$$b_n = \left(\frac{n^{\alpha-1}}{g_n} - (1 - \sum_{k=2}^{n-1} p_{n,k}\frac{k-1}{n}(\frac{n}{k})^{\alpha-1})L\right) n^{\alpha-1} - Q(1 - \sum_{k=2}^{n-1} p_{n,k}\frac{k-1}{n}(\frac{n}{k})^{2(\alpha-1)}),$$

$$q_{n,k} = (\frac{n}{k})^{2(\alpha-1)} p_{n,k}\frac{k-1}{n}.$$

With this notations, the theorem can be written $\lim_{n\to+\infty} a_n = 0$. It is then natural to study the behaviour of b_n when n tends to ∞. To this aim, we should get asymptotics of $1/g_n$, and $\sum_{k=2}^{n-1} p_{n,k}\frac{(k-1)_l}{(n)_l}(\frac{n}{k})^r$, $r \geq 0$ and $l \in \mathbb{N}$, where $(n)_l$ is (the same for $(k-1)_l$):

$$(n)_l = \begin{cases} n(n-1)(n-2)\cdots(n-l+1) & \text{if } n \geq l \geq 1, \\ 0 & \text{if } l > n \geq 1. \end{cases}$$

Asymptotics of $1/g_n$

For any $c, d \in \mathbb{R}$, we have

$$\frac{\Gamma(n+c)}{\Gamma(n+d)} = n^{c-d}(1 + (c-d)\frac{c+d-1}{2}n^{-1} + O(n^{-2})). \tag{3.10}$$

This is a straightforward consequence of Stirling's formula:

$$\Gamma(z) = \sqrt{2\pi}z^{z-1/2}e^{-z}(1 + \frac{1}{12z} + O(\frac{1}{z^2})), z > 0. \tag{3.11}$$

Then we can proceed to: For any real numbers a and $b > -1$,

$$\int_0^1 (1-t)^{n+a}t^b dx = \frac{\Gamma(n+a+1)\Gamma(b+1)}{\Gamma(n+a+b+2)} = \Gamma(b+1)n^{-1-b}\left(1 + (-1-b)\frac{b+2a+2}{2}n^{-1} + O(n^{-2})\right). \tag{3.12}$$

Using (3.12), we get the following lemma.

Lemma 3.3. *For Beta($2-\alpha, \alpha$)-coalescents, we have*

$$g_n = \frac{n^\alpha}{\Gamma(\alpha+1)} - (\frac{\alpha(\alpha-1)}{2\Gamma(\alpha+1)} + \frac{2-\alpha}{\Gamma(\alpha)})n^{\alpha-1} + o(n^{\alpha-1}),$$

and

$$\frac{1}{g_n} = \Gamma(\alpha+1)\left(1 + (-\alpha^2/2 + 3\alpha/2)n^{-1} + o(n^{-1})\right) n^{-\alpha}. \tag{3.13}$$

Proof. It is straightforward using Lemma 3.1 and $g_n^{(s)} = n(n-1)\int_0^1 t(1-t)^{n-2}\rho^{(s)}(t)dt$ for any $s > -\alpha$.

□

Calculus of $\sum_{k=2}^{n-1} p_{n,k}\frac{(k-1)_l}{(n)_l}(\frac{n}{k})^r$

Lemma 3.4. *Consider any Λ-coalescent process, associated with measure ν. Let $l \in \{1, 2, \cdots, n-2\}$ fixed. Then for any real function f:*

$$\sum_{k=2}^{n-1} p_{n,k}\frac{(k-1)_l}{(n)_l} f(k) = \mathbb{E}[\frac{(n-1-X_1^{(n)})_l}{(n)_l}]\mathbb{E}^{\nu^{(l)}}[f(n-X_1^{(n-l)})],$$

where $\mathbb{E}^{\nu^{(l)}}[*]$ means that the Λ-coalescent is associated with the measure $\nu^{(l)}$.

Proof. Recall the definitions of g_n and $p_{n,k}$(see (3.1), (3.2)). We have

$$\begin{aligned}\sum_{k=2}^{n-1} p_{n,k}\frac{(k-1)_l}{(n)_l} &= \sum_{k=l+1}^{n-1}\frac{\int_0^1 \binom{n-l}{n-k+1} x^{n-k+1}(1-x)^{k-1}\nu(dx)}{g_n}\\ &= \sum_{k=l+1}^{n-1}\frac{\int_0^1 \binom{n-l}{n-k+1} x^{n-k+1}(1-x)^{k-1-l}\nu^{(l)}(dx)}{g_n}\\ &= \sum_{k=1}^{n-1-l}\frac{\int_0^1 \binom{n-l}{n-k-l+1} x^{n-k-l+1}(1-x)^{k-1}\nu^{(l)}(dx)}{g_n} = \frac{g_{n-l}^{(l)}}{g_n}. \end{aligned} \tag{3.14}$$

Then,

$$\begin{aligned}\sum_{k=2}^{n-1} p_{n,k}\frac{(k-1)_l}{(n)_l} f(k) &= \left(\sum_{k=2}^{n-1} p_{n,k}\frac{(k-1)_l}{(n)_l}\right)\frac{\sum_{k=2}^{n-1} p_{n,k}\frac{(k-1)_l}{(n)_l} f(k)}{\sum_{k=2}^{n-1} p_{n,k}\frac{(k-1)_l}{(n)_l}}\\ &= \mathbb{E}[\frac{(n-1-X_1^{(n)})_l}{(n)_l}]\frac{\sum_{k=l+1}^{n-1}\int_0^1 \binom{n-l}{n-k+1} x^{n-k+1}(1-x)^{k-1-l} f(k)\nu^{(l)}(dx)}{g_{n-l}^{(l)}}\\ &= \mathbb{E}[\frac{(n-1-X_1^{(n)})_l}{(n)_l}]\frac{\sum_{k=1}^{n-1-l}\int_0^1 \binom{n-l}{n-k-l+1} x^{n-k-l+1}(1-x)^{k-1} f(k+l)\nu^{(l)}(dx)}{g_{n-l}^{(l)}}\\ &= \mathbb{E}[\frac{(n-1-X_1^{(n)})_l}{(n)_l}]\mathbb{E}^{\nu^{(l)}}[f(Y_1^{(n-l)}+l)] = \mathbb{E}[\frac{(n-1-X_1^{(n)})_l}{(n)_l}]\mathbb{E}^{\nu^{(l)}}[f(n-X_1^{(n-l)})].\end{aligned}$$

This achieves the proof of the lemma. □

In consequence,

$$\sum_{k=2}^{n-1} p_{n,k}\frac{(k-1)_l}{(n)_l}(\frac{n}{k})^r = \mathbb{E}[\frac{(n-1-X_1^{(n)})_l}{(n)_l}]\mathbb{E}^{\nu^{(l)}}[(\frac{n}{n-X_1^{(n-l)}})^r]. \tag{3.15}$$

We have to study $\mathbb{E}[\frac{(n-1-X_1^{(n)})_l}{(n)_l}]$ and $\mathbb{E}^{\nu^{(l)}}[(\frac{n}{n-X_1^{(n-l)}})^r]$. The latter is given by Proposition 3.2 in appendix A. The following lemma studies the former.

Lemma 3.5. *Consider a Beta$(2-\alpha,\alpha)$ n-coalescent. Let $l \in \{1,2,\cdots,n-2\}$ fixed. We have*

$$\mathbb{E}[\frac{(n-1-X_1^{(n)})_l}{(n)_l}] = 1 - \frac{l\alpha}{n(\alpha-1)} + \Gamma(\alpha+1)\left(\sum_{j=2}^{l}\binom{l}{j}(-1)^j\int_0^1 x^j\nu(dx) - C^{(0)}l\right)n^{-\alpha} + o(n^{-\alpha}),$$

Proof. We have

$$\mathbb{E}[\frac{(n-1-X_1^{(n)})_l}{(n)_l}] = \mathbb{E}[1-\sum_{i=0}^{l-1}\frac{X_1^{(n)}+1}{n-i} + \sum_{j=2}^{l}\sum_{i_1,\cdots,i_j \text{ all different}}(-1)^j\frac{(X_1^{(n)}+1)^j}{(n-i_1)(n-i_2)\cdots(n-i_j)}].$$

For $\mathbb{E}[\sum_{i=0}^{l-1}\frac{X_1^{(n)}+1}{n-i}]$, we use Lemma 3.7 in appendix B. While using Lemma 3.8, we get

$$\begin{aligned}&\mathbb{E}[\sum_{j=2}^{l}\sum_{i_1,\cdots,i_j \text{ all different}}(-1)^j\frac{(X_1^{(n)}+1)^j}{(n-i_1)(n-i_2)\cdots(n-i_j)}]\\ =&n^{-\alpha}\Gamma(\alpha+1)\sum_{j=2}^{l}\binom{l}{j}(-1)^j\int_0^1 x^j\nu(dx) + O(n^{-\min\{1+\alpha,j\}}).\end{aligned}$$

Then we conclude.

□

Now we can give the estimate of $\sum_{k=2}^{n-1}p_{n,k}\frac{(k-1)_l}{(n)_l}(\frac{n}{k})^r$ using (3.15), Lemma 3.5 and Proposition 3.2.

Proposition 3.1. *Consider a Beta$(2-\alpha,\alpha)$ n-coalescent. Let $l \in \{1,2,\cdots,n-2\}$ and $r \in [0,\alpha+l)$ fixed. We have*

$$\begin{aligned}&\sum_{k=2}^{n-1}p_{n,k}\frac{(k-1)_l}{(n)_l}(\frac{n}{k})^r\\ =&1+\frac{(r-l\alpha)}{n(\alpha-1)} + o(n^{-\alpha})\\ +&\quad\Gamma(\alpha+1)\left(\int_0^1((1-x)^{-r}-1-rx)\nu^{(l)}(dx) + \sum_{j=2}^{l}\binom{l}{j}(-1)^j\int_0^1 x^j\nu(dx) + rC^{(l)} - lC^{(0)}\right)n^{-\alpha}.\end{aligned}$$

3.2.3 Proof of Theorem 3.4.

Recall the transformation (3.8) and the associated recurrence (3.9). The aim is to prove that $\lim_{n\to+\infty} a_n = 0$ for a_n in (3.9). Using Proposition 3.1, we get

$$1-\sum_{k=2}^{n-1} p_{n,k}\frac{k-1}{n}(\frac{n}{k})^{\alpha-1} = \frac{1}{n(\alpha-1)} - \Gamma(\alpha+1)\left(A+(\alpha-1)C^{(1)}-C^{(0)}\right)n^{-\alpha}+o(n^{-\alpha}),$$

and

$$1-\sum_{k=2}^{n-1} p_{n,k}\frac{k-1}{n}(\frac{n}{k})^{2(\alpha-1)} = \frac{2-\alpha}{n(\alpha-1)}+O(n^{-\alpha}).$$

Hence we deduce that $b_n = o(n^{-1})$.

Let $\varepsilon > 0$ such that $2(\alpha-1)+\epsilon < \alpha$. We have $1-\sum_{k=2}^{n-1} p_{n,k}\frac{k-1}{n}(\frac{n}{k})^{2(\alpha-1)+\varepsilon} = O(n^{-1}) > 0$. The recurrence (3.9) satisfies the assumptions of Lemma 3.6 which leads to $\lim_{n\to+\infty} a_n = 0$. Then we can conclude.

3.3 Estimate of $\mathbb{E}[T_1^{(n)}T_2^{(n)}]$ and proof of Theorem 3.2

Using Theorem 1.1 in (14), we have

$$\mathbb{E}[T_1^{(n)}T_2^{(n)}] = \frac{2\mathbb{E}[T_1^{(n)}]}{g_n} + \sum_{k=2}^{n-1} p_{n,k}\frac{(k-1)_2}{(n)_2}\mathbb{E}[T_1^{(k)}T_2^{(k)}]. \tag{3.16}$$

As a consequence of (3.13) and Theorem 3.4, we have

$$\frac{2\mathbb{E}[T_1^{(n)}]}{g_n} = 2(\Gamma(\alpha+1))^2 n^{1-2\alpha}\left(\alpha-1+\frac{(\alpha-1)^2\Gamma(\alpha+1)}{2-\alpha}(A+(\alpha-1)C^{(1)}-C^{(0)})n^{1-\alpha}\right)+o(n^{2-3\alpha}).$$

Using the recurrence method described in the previous section, a direct calculation gives that

$$\begin{aligned}
&\mathbb{E}[T_1^{(n)}T_2^{(n)}]\\
&= ((\alpha-1)\Gamma(\alpha+1))^2 n^{2(1-\alpha)}\\
&\quad + \frac{\alpha-1}{3-\alpha}((\alpha-1)\Gamma(\alpha+1))^3\left(B+2(\alpha-1)C^{(2)}+1-2C^{(0)}+\frac{2}{2-\alpha}(A+(\alpha-1)C^{(1)}-C^{(0)})\right)n^{3(1-\alpha)}\\
&\quad + o(n^{3(1-\alpha)}).
\end{aligned}$$

Now together with Theorem 3.4, we can get the estimate of $\mathrm{cov}(T_1^{(n)}, T_2^{(n)})$.

$$\mathrm{cov}(T_1^{(n)}, T_2^{(n)}) = \frac{((\alpha-1)\Gamma(\alpha+1))^3}{3-\alpha}\left(B - 2A + 2(\alpha-1)(C^{(2)} - C^{(1)}) + 1\right) n^{3(1-\alpha)} + o(n^{3(1-\alpha)}).$$

Then

$$\Delta(\alpha) = \frac{((\alpha-1)\Gamma(\alpha+1))^3}{3-\alpha}\left(B - 2A + 2(\alpha-1)(C^{(2)} - C^{(1)}) + 1\right). \tag{3.17}$$

It is straightforward to see that $\Delta(\alpha) = \frac{((\alpha-1)\Gamma(\alpha+1))^2\Gamma(4-\alpha)}{(3-\alpha)\Gamma(4-2\alpha)}$ by recalling the values of $A, B, C^{(1)}$ and $C^{(2)}$. We prove then that $\Delta(\alpha) = \frac{\int_0^1((1-x)^{2-\alpha}-1)^2\nu(dx)}{3-\alpha}((\alpha-1)\Gamma(\alpha+1))^3$. Notice that

$$B - 2A = \int_0^1 \left((1-x)^{2(2-\alpha)} - 2(1-x)^{2-\alpha} + 1 - x^2 + 2(\alpha-1)x^2(1-x)\right)\nu(dx).$$

By definition,

$$C^{(2)} - C^{(1)} = \lim_{t\to+\infty}\int_t^1 (\rho^{(2)}(x) - \rho^{(1)}(x))dx = \lim_{t\to 0}\int_t^1 x(\nu^{(2)}(dx) - \nu^{(1)}(dx)) = \int_0^1 -x^2(1-x)\nu(dx),$$

and $\int_0^1 x^2\nu(dx) = 1$. Then it allows to conclude.

3.4 Proof of Theorem 3.3

Notice that $n^{\alpha-1}T_1^{(n)} \overset{(d)}{\to} T$ and if $\beta \geq \frac{\alpha}{\alpha-1}$, one gets $\mathbb{E}[T^\beta] = +\infty$, hence $\mathbb{E}[(n^{\alpha-1}T_1^{(n)})^\beta]$ converges to $+\infty$ (see Lemma 4.11 of (28)). If $0 \leq \beta_1 < \beta_2 < \frac{\alpha}{\alpha-1}$ and $(\mathbb{E}[(n^{\alpha-1}T_1^{(n)})^{\beta_2}], n \geq 2)$ is bounded, then $((n^{\alpha-1}T_1^{(n)})^{\beta_1}, n \geq 2)$ is uniformly integrable (see Lemma 4.11 of (28) and Problem 14 in section 8.3 (11)). Then we need only to prove that for any $\beta \in [2, \frac{\alpha}{\alpha-1})$, $(\mathbb{E}[(n^{\alpha-1}T_1^{(n)})^\beta], n \geq 2)$ is bounded.

We will prove by induction on n that there exists a constant $C > 0$ such that for all $n \geq 2$, $(\mathbb{E}[n^{\alpha-1}T_1^{(n)}])^\beta \leq C$. We first assume that, for all $2 \leq k \leq n-1$, $(\mathbb{E}[k^{\alpha-1}T_1^{(k)}])^\beta \leq C$ and then will prove that (if C is large enough) $(\mathbb{E}[n^{\alpha-1}T_1^{(n)}])^\beta \leq C$.

Writing the decomposition of $T_1^{(n)}$ at the first coalescence, we have

$$T_1^{(n)} = \frac{e_0}{g_n} + \sum_{k=2}^{n-1}\mathbf{1}_{\{H_{n,k}\}}\bar{T}_1^{(k)},$$

where:

— $H_{n,k}$ is the event: {From n individuals, we have k individuals after the first coalescence, and individual 1 is not involved in this collision}, $2 \leq k \leq n-1$;

— e_0 is a unit exponential random variable, $\bar{T}_1^{(k)} \overset{(d)}{=} T_1^{(k)}$, and all these random variables e_0, $\bar{T}_1^{(k)}$, $\mathbf{1}_{\{H_{n,k}\}}$ are independent. One notices that $\mathbb{P}(H_{n,k}) = p_{n,k}\frac{k-1}{n}$ (see (3.7)).

Using Lemma 3.11 in Appendix D, we have the following inequality.

$$\mathbb{E}[(T_1^{(n)})^\beta] = \mathbb{E}[((\frac{e_0}{g_n} + \sum_{k=2}^{n-1} \mathbf{1}_{\{H_{n,k}\}}\bar{T}_1^{(k)}))^\beta] \leq I_{n,1} + I_{n,2} + I_{n,3} + I_{n,4} \tag{3.18}$$

where

$$I_{n,1} = \mathbb{E}[(\frac{e_0}{g_n})^\beta], \quad I_{n,2} = \mathbb{E}[(\sum_{k=2}^{n-1} \mathbf{1}_{\{H_{n,k}\}}\bar{T}_1^{(k)})^\beta],$$

$$I_{n,3} = \mathbb{E}[\beta 2^{\beta-1}\frac{e_0}{g_n}(\sum_{k=2}^{n-1} \mathbf{1}_{\{H_{n,k}\}}\bar{T}_1^{(k)})^{\beta-1}] \text{ and } I_{n,4} = \mathbb{E}[\beta 2^{\beta-1}(\frac{e_0}{g_n})^{\beta-1}\sum_{k=2}^{n-1} \mathbf{1}_{\{H_{n,k}\}}\bar{T}_1^{(k)}].$$

We first bound $I_{n,1}$. Recall that $g_n \sim \dfrac{n^\alpha}{\Gamma(\alpha+1)}$. Hence there exists a constant $K_1 > 0$ (which depends on β) such that for any $n \geq 2$,

$$n^{(\alpha-1)\beta} I_{n,1} \leq \frac{K_1}{n}. \tag{3.19}$$

We now consider $I_{n,2}$. Notice that $(\alpha-1)\beta < \alpha+1$. Hence, using Proposition 3.1, we have

$$n^{(\alpha-1)\beta} I_{n,2} = n^{-(\alpha-1)\beta}\sum_{k=2}^{n-1} p_{n,k}\frac{k-1}{n}(\frac{n}{k})^{(\alpha-1)\beta}\mathbb{E}[(k^{\alpha-1}T_1^{(k)})^\beta] \tag{3.20}$$

$$\leq C\sum_{k=2}^{n-1} p_{n,k}\frac{k-1}{n}(\frac{n}{k})^{(\alpha-1)\beta} \tag{3.21}$$

$$= C(1 - \frac{\alpha-(\alpha-1)\beta}{n(\alpha-1)} + o(n^{-1})) \qquad \leq C(1 - \frac{\alpha-(\alpha-1)\beta}{2n(\alpha-1)}), \tag{3.22}$$

for $n \geq N$, where N is a fixed positive integer.

We now proceed to $I_{n,3}$. Notice that for $2 \leq k \leq n-1$,

$$\mathbb{E}[(k^{\alpha-1}T_1^{(k)})^{\beta-1}] \leq (\mathbb{E}[(k^{\alpha-1}T_1^{(k)})^\beta])^{\frac{\beta-1}{\beta}} \leq C^{\frac{\beta-1}{\beta}}.$$

Hence we have

$$\begin{aligned}
n^{(\alpha-1)\beta} I_{n,3} &= n^{(\alpha-1)\beta}\mathbb{E}[\beta 2^{\beta-1}\frac{e_0}{g_n}\sum_{k=2}^{n-1}\mathbf{1}_{\{H_{n,k}\}}(\bar{T}_1^{(k)})^{\beta-1}] \\
&\leq C^{\frac{\beta-1}{\beta}}\beta 2^{\beta-1} n^{\alpha-1} g_n^{-1}\sum_{k=2}^{n-1} p_{n,k}\frac{k-1}{n}\left(\frac{n}{k}\right)^{(\alpha-1)(\beta-1)} \\
&= C^{\frac{\beta-1}{\beta}} n^{\alpha-1}\beta 2^{\beta-1} g_n^{-1}(1-\frac{\alpha-(\alpha-1)(\beta-1)}{n(\alpha-1)}+o(n^{-1})) \\
&\leq \frac{C^{\frac{\beta-1}{\beta}}K_2}{n}, \qquad (3.23)
\end{aligned}$$

where K_2 is a positive constant. In the second equality, we have used Proposition 3.1.

While for any $n \geq 2$,

$$\begin{aligned}
n^{(\alpha-1)\beta} I_{n,4} &= n^{(\alpha-1)\beta}\mathbb{E}[\beta 2^{\beta-1}(\frac{e_0}{g_n})^{\beta-1}\sum_{k=2}^{n-1}\mathbf{1}_{\{H_{n,k}\}}\bar{T}_1^{(k)}] \\
&\leq \beta 2^{\beta-1}\mathbb{E}[e_0^{\beta-1}](g_n)^{1-\beta} n^{(\alpha-1)(\beta-1)}\mathbb{E}[n^{\alpha-1}T_1^{(n)}] \\
&\leq \frac{K_3}{n^{\beta-1}} \leq \frac{K_3}{n}, \qquad (3.24)
\end{aligned}$$

where K_3 is a positive constant. We have used Lemma 3.4 to bound $\mathbb{E}[n^{\alpha-1}T_1^{(n)}]$.

Using (3.18),(3.19),(3.20),(3.23),(3.24), we have proved that for any $n, n \geq N$, if there exists $C > 0$ such that for all $2 \leq k \leq n-1$, $\mathbb{E}[\left(k^{\alpha-1}T_1^{(k)}\right)^{\beta}] \leq C$, then

$$\mathbb{E}[(n^{\alpha-1}T_1^{(n)})^{\beta}] \leq \frac{C+\left(K_1 - C\frac{\alpha-(\alpha-1)\beta}{2(\alpha-1)} + C^{\frac{\beta-1}{\beta}}K_2 + K_3\right)}{n}. \qquad (3.25)$$

Let C large enough such that

$$K_1 - C\frac{\alpha-(\alpha-1)\beta}{2(\alpha-1)} + C^{\frac{\beta-1}{\beta}}K_2 + K_3 < 0, \qquad (3.26)$$

Then $\mathbb{E}[(n^{\alpha-1}T_1^{(n)})^{\beta}] \leq C$, which allows to conclude.

3.5 Appendix

A) The main recurrence tool

Lemma 3.6. *We consider the recurrence $a_n = b_n + \sum_{k=1}^{n-1} q_{n,k} a_k$. We assume that $b_n = o(n^{-1})$ and that there exist $\varepsilon > 0$ and $C > 0$ such that $1 - \sum_{k=1}^{n-1} q_{n,k} (\frac{n}{k})^{\varepsilon} \geq Cn^{-1}$ for n large enough. Then $\lim_{n \to +\infty} a_n = 0$.*

Proof. Let $(\bar{c}_n)_{n\geq 1}$ be an increasing sequence such that

$$\lim_{n\to+\infty} \bar{c}_n = +\infty; \lim_{n\to+\infty} nb_n\bar{c}_n = 0.$$

Define another sequence $(c_n)_{n\geq 1}$ by: $c_1 = \bar{c}_1$. For $n \geq 1$,

$$c_{n+1} = \min\{c_n(\frac{n+1}{n})^{\varepsilon}, \bar{c}_{n+1}\},$$

Then we have $\lim_{n\to+\infty} c_n = +\infty, c_n b_n = o(n^{-1})$ and for any $1 \leq k \leq n-1$, $\frac{c_n}{c_k} \leq (\frac{n}{k})^{\varepsilon}$. In consequence, $1 - \sum_{k=1}^{n-1} q_{n,k} \frac{c_n}{c_k} \geq Cn^{-1}$ for n large enough. Let $n_1 > 0$ such that for $n > n_1$, we have $1 - \sum_{k=1}^{n-1} q_{n,k} \frac{c_n}{c_k} > \frac{C}{n}$ and $c_n b_n < \frac{C}{2n}$ and pick a number C' such that $C' > \max\{1, c_k a_k; 1 \leq k \leq n_1\}$. We transform the original recurrence to

$$c_n a_n = c_n b_n + \sum_{k=1}^{n-1} \left(q_{n,k} \frac{c_n}{c_k} \right) c_k a_k.$$

Then $c_{n_1+1} a_{n_1+1} \leq \frac{C}{2(n_1+1)} + (1 - \frac{C}{n_1+1})C' \leq C'$. By induction, we prove that the sequence $(c_n a_n)_{n\geq 1}$ is bounded by C'. Since c_n tends to the infinity, we get $\lim_{n\to+\infty} a_n = 0$. □

Remark 3.3. *We refer to (35) for a rather detailed survey on this kind of recurrence relationships.*

B) Asymptotic behaviours of $X_1^{(n)}$

Lemma 3.7. *Consider the coalescent process with related measure $\nu^{(s)}$ where $s > -\alpha$. Then*

$$\mathbb{E}^{\nu^{(s)}}[X_1^{(n)}] = \frac{1}{\alpha - 1} + \Gamma(\alpha+1)C^{(s)}n^{1-\alpha} + o(n^{1-\alpha}),$$

Proof. We have (see (13)):

$$\mathbb{E}^{\nu^{(s)}}[X_1^{(n)}] = \frac{\int_0^1 (1-t)^{n-2} (\int_t^1 \rho^{(s)}(r)dr)dt}{\int_0^1 (1-t)^{n-2} t\rho^{(s)}(t)dt}$$

Lemma 3.1 gives the developments of $\rho^{(s)}(t)$ and $\int_t^1 \rho^{(s)}(r)dr$. Using (3.10), we get

$$\int_0^1 (1-t)^{n-2} (\int_t^1 \rho^{(s)}(r)dr)dt = \frac{n^{\alpha-2}}{(\alpha-1)\Gamma(\alpha+1)} + C^{(s)}n^{-1} + o(n^{-1}),$$

and $\int_0^1 (1-t)^{n-2} t\rho^{(s)}(t)dt = \frac{n^{\alpha-2}}{\Gamma(\alpha+1)} + O(n^{\alpha-3})$. Then we can conclude. □

Lemma 3.8. *If $s > -\alpha$ and $k \geq 2$,*

$$\mathbb{E}^{\nu^{(s)}}[(\frac{X_1^{(n)}}{n})^k] = \Gamma(\alpha+1)\int_0^1 x^k \nu^{(s)}(dx) n^{-\alpha} + O(n^{-\min\{1+\alpha,k\}}).$$

Proof. Let $B_{n,x}$ denote a binomial random variable with parameter $(n,x), n \geq 2, 0 \leq x \leq 1$. Recall that for $2 \leq i \leq n$, $\mathbb{P}^{\nu^{(s)}}(X_1^{(n)} = i-1) = \int_0^1 \binom{n}{i} x^i (1-x)^{n-i} \nu^{(s)}(dx)/g_n^{(s)} = \int_0^1 \mathbb{P}(B_{n,x} = i)\nu^{(s)}(dx)/g_n^{(s)}$. Here $\mathbb{P}^{\nu^{(s)}}$ means that $X_1^{(n)}$ is related to the coalescent process with measure $\nu^{(s)}$.

$$\begin{aligned}\mathbb{E}^{\nu^{(s)}}[(\frac{X_1^{(n)}}{n})^k] &= \int_0^1 \mathbb{E}[(\frac{B_{n,x}-1}{n})^k \mathbf{1}_{B_{n,x}\geq 1}])\nu^{(s)}(dx)/g_n^{(s)} \\ &= \int_0^1 n^{-k}\mathbb{E}[(B_{n,x}^k - B_{n,x}) \\ &\quad + \sum_{i=1}^{k-1}\binom{k}{i}(-1)^i(B_{n,x}^{k-i} - B_{n,x}) + (-1)^k(1-B_{n,x})\mathbf{1}_{B_{n,x}\geq 1})]\nu^{(s)}(dx)/g_n^{(s)}.\end{aligned}$$

Using Lemma 3.9 in Appendix C, we get $\mathbb{E}[(B_{n,x}^k - B_{n,x})] = (nx)^k + O(n^{k-1})x^2$. Then

$$\begin{aligned}\mathbb{E}^{\nu^{(s)}}[(\frac{X_1^{(n)}}{n})^k] &= \frac{\int_0^1 n^{-k}\left((nx)^k + O(n^{k-1})x^2\right)\nu^{(s)}(dx)}{g_n^{(s)}} + \frac{n^{-k}\int_0^1(-1)^k(1-nx-(1-x)^n)\nu^{(s)}(dx)}{g_n^{(s)}} \\ &= \Gamma(\alpha+1)\int_0^1 x^k\nu^{(s)}(dx)n^{-\alpha} + O(n^{-\min\{1+\alpha,k\}}).\end{aligned}$$

In the second equality, we have used $g_n^{(s)} \sim \frac{n^\alpha}{\Gamma(\alpha+1)}$ and also the fact that $\int_0^1(1-nx-(1-x)^n)\nu^{(s)}(dx) \leq g_n^{(s)} = \int_0^1(1-nx(1-x)^{n-1}-(1-x)^n)\nu^{(s)}(dx)$. This achieves the proof. □

Proposition 3.2. *For $s \in \mathbb{N}\bigcup\{0\}$ and $0 \leq r < \alpha + s$, we have*

$$\mathbb{E}^{\nu^{(s)}}[\left(\frac{n}{n-X_1^{(n-s)}}\right)^r] = 1 + \frac{r}{n(\alpha-1)} + \Gamma(\alpha+1)\left(\int_0^1((1-x)^{-r}-1-rx)\nu^{(s)}(dx) + rC^{(s)}\right)n^{-\alpha} + o(n^{-\alpha}).$$

Proof. By Taylor expansion formula, for $m \geq 2$ and $n \geq s+2$, we have,

$$\mathbb{E}^{\nu^{(s)}}[\left(\frac{n}{n-X_1^{(n-s)}}\right)^r] = \mathbb{E}^{\nu^{(s)}}[\left(\frac{1}{1-\frac{X_1^{(n-s)}}{n}}\right)^r]$$

$$= \mathbb{E}^{\nu^{(s)}}[1 + r\frac{X_1^{(n-s)}}{n} + \sum_{k=2}^m \frac{\Gamma(k+r)}{\Gamma(r)k!}(\frac{X_1^{(n-s)}}{n})^k + \frac{\Gamma(m+1+r)}{\Gamma(r)m!}\int_0^{\frac{X_1^{(n-s)}}{n}}(1-t)^{-r-m-1}(\frac{X_1^{(n-s)}}{n}-t)^m dt].$$

Using Lemma 3.7 and Lemma 3.8, we have for $m \geq 2$,

$$\lim_{n\to+\infty} n^\alpha \mathbb{E}^{\nu^{(s)}}[\sum_{k=2}^m \frac{\Gamma(k+r)}{\Gamma(r)k!}(\frac{X_1^{(n-s)}}{n})^k] = \Gamma(\alpha+1)\sum_{k=2}^m \frac{\Gamma(k+r)}{\Gamma(r)k!}\int_0^1 x^k \nu^{(s)}(x).$$

In consequence,

$$\lim_{m\to+\infty}\lim_{n\to+\infty} n^\alpha \mathbb{E}^{\nu^{(s)}}[\sum_{k=2}^m \frac{\Gamma(k+r)}{\Gamma(r)k!}(\frac{X_1^{(n-s)}}{n})^k] = \Gamma(\alpha+1)\int_0^1 ((1-x)^{-r}-1-rx)\nu^{(s)}(dx).$$

It remains to estimate $\frac{\Gamma(m+1+r)}{\Gamma(r)m!}\mathbb{E}^{\nu^{(s)}}[\int_0^{\frac{X_1^{(n-s)}}{n}}(1-t)^{-r-m-1}(\frac{X_1^{(n-s)}}{n}-t)^m dt]$, which is the sum of two terms $P_1(m,n,s,y)$ and $P_2(m,n,s,y)$, with $0<y<1$, defined by

$$P_1(m,n,s,y) = \frac{\Gamma(m+1+r)}{\Gamma(r)m!}\mathbb{E}^{\nu^{(s)}}[\int_0^{\frac{X_1^{(n-s)}}{n}}(1-t)^{-r-m-1}(\frac{X_1^{(n-s)}}{n}-t)^m dt \mathbf{1}_{X_1^{(n-s)}\geq ny}],$$

$$P_2(m,n,s,y) = \frac{\Gamma(m+1+r)}{\Gamma(r)m!}\mathbb{E}^{\nu^{(s)}}[\int_0^{\frac{X_1^{(n-s)}}{n}}(1-t)^{-r-m-1}(\frac{X_1^{(n-s)}}{n}-t)^m dt \mathbf{1}_{X_1^{(n-s)}<ny}].$$

We first focus on $P_1(m,n,s,y)$. By Proposition 3.3 in Appendix C, we have

$$\begin{aligned} P_1(m,n,s,y) \leq & \mathbb{E}^{\nu^{(s)}}[\left(\frac{n}{n-X_1^{(n-s)}}\right)^r \mathbf{1}_{X_1^{(n-s)}\geq ny}] \\ \leq & \mathbb{E}^{\nu^{(s)}}[\left(\frac{n-s}{n-s-X_1^{(n-s)}}\right)^r \mathbf{1}_{X_1^{(n-s)}\geq (n-s)y}] \\ \leq & n^{-\alpha}K_4 y^{-\alpha}(1-y)^{\bar{r}-r}, \end{aligned} \tag{3.27}$$

where $\bar{r}\in(r,\alpha+s)$ and K_4 is a number depending only on $\bar{r}$ and $\nu^{(s)}$ (it is important that it does not depend on y).

We now give an upper bound for $P_2(m,n,s,y)$. We have

$$n^\alpha P_2(m,n,s,y) = n^\alpha \frac{\Gamma(m+1+r)}{\Gamma(r)m!}\mathbb{E}^{\nu^{(s)}}[\int_0^{\frac{X_1^{(n-s)}}{n}}(1-t)^{-r-1}(\frac{X_1^{(n-s)}/n-t}{1-t})^m dt \mathbf{1}_{X_1^{(n-s)}<ny}].$$

For $t \in [0,x)$ with $0 < x \leq 1$, we have $\frac{x-t}{1-t} \leq x$. Then $\int_0^{\frac{X_1^{(n-s)}}{n}} (\frac{X_1^{(n-s)}/n-t}{1-t})^m dt \leq (\frac{X_1^{(n-s)}}{n})^{m+1}$.

Hence, using Lemma 3.8, for $m > 2$,

$$n^\alpha P_2(m,n,s,y) \leq n^\alpha \frac{\Gamma(m+1+r)}{\Gamma(r)m!}(1-y)^{-r-1}\mathbb{E}[(X_1^{(n-s)}/n)^{m+1}]$$
$$= (1-y)^{-r-1}\frac{\Gamma(m+1+r)}{\Gamma(r)m!}\left(\Gamma(\alpha+1)\int_0^1 x^{m+1}\nu^{(s)}(dx) + O(n^{-1})\right).$$

Using Lemma 3.10 in Appendix C, we have

$$\int_0^1 x^{m+1}\nu^{(s)}(dx) = \int_0^1 x^{m+1}(1-x)^{\bar{r}}\nu^{(-\bar{r}+s)}(dx) \leq K_5 m^{-\bar{r}},$$

where K_5 is a positive real number depending only on $\bar{r}$ and $\nu^{(s)}$.

Notice that $\frac{\Gamma(m+r+1)}{\Gamma(r)m!} \sim \frac{m^r}{\Gamma(r)}$. Hence

$$P_2(m,n,s,y) \leq n^{-\alpha}(1-s)^{-r-1}m^r(O(m^{-\bar{r}}) + o(n^{-1})). \tag{3.28}$$

Combining (3.27) and (3.28), we deduce that

$$\lim_{m\to+\infty}\limsup_{n\to+\infty} n^\alpha(P_1(m,n,s,y) + P_2(m,n,s,y)) = 0.$$

This convergence together with Lemma 3.7 and 3.8 yield this proposition. □

C) Some necessary results for Appendix B

Lemma 3.9. *Let $B_{n,x}$ be a binomial random variable with parameter $(n,x), n \geq 2, 0 \leq x \leq 1$. Let k be an integer such that $2 \leq k \leq n$. Then*

$$nx + n(n-1)\cdots(n-k+1)x^k \leq \mathbb{E}[B_{n,x}^k] \leq (nx)^k + \binom{k}{2}n^{k-1}x^2,$$

Proof. Write $B_{n,x} = Y_1 + \cdots + Y_n$, where $Y_1, \cdots, Y_n$ are independent Bernoulli random variables. Let $S := \{\{i_1, \cdots, i_k\}; 1 \leq i_1, \cdots, i_k \leq n\}$. Then

$$\mathbb{E}[\sum_{\{i_1,\cdots,i_k\}\in S_1} Y_{i_1}\cdots Y_{i_k}] + \mathbb{E}[\sum_{\{i_1,\cdots,i_k\}\in S_3} Y_{i_1}\cdots Y_{i_k}] \leq \mathbb{E}[(B_{n,x})^k]$$
$$\leq \mathbb{E}[\sum_{\{i_1,\cdots,i_k\}\in S_2} Y_{i_1}\cdots Y_{i_k}] + \mathbb{E}[\sum_{\{i_1,\cdots,i_k\}\in S_3} Y_{i_1}\cdots Y_{i_k}],$$

where

1. $S_1 := \{\{i_1, \cdots, i_n\} \in A; i_1 = \cdots = i_k\}$. Then $\mathbb{E}[\sum_{\{i_1,\cdots,i_k\}\in S_1} Y_{i_1}\cdots Y_{i_k}] = nx$.
2. $S_2 := \{\{i_1, \cdots, i_n\} \in A; \exists 1 \le p < q \le k, i_p = i_q\}$. Then $\mathbb{E}[\sum_{\{i_1,\cdots,i_k\}\in S_2} Y_{i_1}\cdots Y_{i_k}] \le \binom{k}{2} n^{k-1}x^2$.
3. $S_3 := \{\{i_1, \cdots, i_n\} \in A; \forall 1 \le p < q \le k, i_p \neq i_q\}$. Then $\mathbb{E}[\sum_{\{i_1,\cdots,i_k\}\in S_3} Y_{i_1}\cdots Y_{i_k}] = n(n-1)\cdots(n-k+1)x^k$.

Then we can conclude. □

Lemma 3.10. *Consider any Λ-coalescent such that $\rho(t) = Ct^{-\alpha} + o(t^{-\alpha})$. Then for every $s \ge 0$, $n \ge 2$, $\int_0^1 x^n(1-x)^s\nu(dx) \le K_6 n^{-s}$, where K_6 is a positive constant which depends only on s and ν.*

Proof. It is clear that there exists $K_7 > 0$ such that $\rho(t) \le K_7 t^{-\alpha}$, for all $0 < t \le 1$. Then

$$\begin{aligned}\int_0^1 x^n(1-x)^s\nu(dx) &= \int_0^1 \rho(t)(n-(n+s)t)t^{n-1}(1-t)^{s-1}dt\\ &\le \int_0^1 \rho(t)(n-nt)t^{n-1}(1-t)^{s-1}dt\\ &\le nK_7\int_0^1 t^{n-1-\alpha}(1-t)^s dt = nK_7\frac{\Gamma(n-\alpha)\Gamma(s+1)}{\Gamma(n-\alpha+s+1)} \le K_6 n^{-s},\end{aligned}$$

for some K_6 which only depends on K_7 and s. This achieves the proof of the lemma. □

Proposition 3.3. *Let $s > -\alpha$ and $0 \le r < \alpha + s$, $\bar{r} \in (r, \alpha + s)$. Then there exists a constant K_{11} depending only on $\bar{r}$ and s such that for all $y \in (0,1)$, $n \ge 2$,*

$$\mathbb{E}^{\nu^{(s)}}[\left(\frac{n}{n-X_1^{(n)}}\right)^r \mathbf{1}_{X_1^{(n)}\ge ny}] \le n^{-\alpha}K_{11}y^{-\alpha}(1-y)^{\bar{r}-r}.$$

Proof. Define $\lceil x\rceil = \min\{m \in \mathbb{Z}; m \ge x\}$. We have

$$\mathbb{E}^{\nu^{(s)}}[\left(\frac{n}{n-X_1^{(n)}}\right)^r \mathbf{1}_{X_1^{(n)}\ge ny}] = \sum_{k=\lceil ny\rceil}^{n-1}\int_0^1 \binom{n}{k+1}x^{k+1}(1-x)^{n-k-1}(\frac{n}{n-k})^r\nu^{(s)}(dx)/g_n^{(s)}.$$

Using (3.10), there exist two positive constants K_8, K_9 such that for all $k \in \{1, 2, \ldots, n-1\}$,

$$K_8\frac{\Gamma(n+1+r)}{\Gamma(k+2)\Gamma(n-k+r)} \le \binom{n}{k+1}(\frac{n}{n-k})^r \le K_9\frac{\Gamma(n+1+r)}{\Gamma(k+2)\Gamma(n-k+r)}.$$

Moreover using integration by parts, for $1 \leq l \leq n-1$ and $0 \leq x \leq 1$, we have:

$$\begin{aligned}
&\sum_{k=l}^{n-1} \frac{\Gamma(n+1+r)}{\Gamma(k+2)\Gamma(n-k+r)} x^{k+1}(1-x)^{n-k-1+r} \\
=&\frac{\Gamma(n+1+r)}{\Gamma(l+1)\Gamma(n-l+r)} \int_0^x t^l (1-t)^{n-l+r-1} dt + \frac{\Gamma(n+1+r)}{\Gamma(n+1)\Gamma(1+r)} x^n (1-x)^r \\
&- \frac{\Gamma(n+1+r)}{\Gamma(n)\Gamma(1+r)} \int_0^x t^{n-1}(1-t)^r dt. \qquad (3.29)
\end{aligned}$$

Lemma 3.1 says $\rho^{(-\bar{r}+s)}(t) = \frac{t^{-\alpha}}{\Gamma(2-\alpha)\Gamma(\alpha+1)} + o(t^{-\alpha})$. Then there exists $K_{10} > 0$, such that $\rho^{(-\bar{r}+s)}(t) \leq K_{10} t^{-\alpha}$ for all $t \in (0,1]$.

$$\begin{aligned}
&\mathbb{E}^{\nu^{(s)}}[\left(\frac{n}{n-X_1^{(n)}}\right)^{\bar{r}} \mathbf{1}_{X_1^{(n)} \geq ny}] \\
&= \sum_{k=\lceil ny \rceil}^{n-1} \frac{\int_0^1 \binom{n}{k+1} (\frac{n}{n-k})^{\bar{r}} x^{k+1}(1-x)^{n-k-1} \nu^{(s)}(dx)}{g_n^{(s)}} \\
&= \sum_{k=\lceil ny \rceil}^{n-1} \frac{\int_0^1 \binom{n}{k+1} (\frac{n}{n-k})^{\bar{r}} x^{k+1}(1-x)^{n-k-1+\bar{r}} \nu^{(-\bar{r}+s)}(dx)}{g_n^{(s)}} \\
&\leq K_9 \frac{\int_0^1 \frac{\Gamma(n+1+\bar{r})}{\Gamma(\lceil ny \rceil+1)\Gamma(n-\lceil ny \rceil+\bar{r})} \int_0^x t^{\lceil ny \rceil}(1-t)^{n-\lceil ny \rceil+\bar{r}-1} dt \nu^{(-\bar{r}+s)}(dx)}{g_n^{(s)}} \\
&\quad + K_9 \frac{\int_0^1 \frac{\Gamma(n+1+\bar{r})}{\Gamma(n+1)\Gamma(1+\bar{r})} x^n (1-x)^{\bar{r}} \nu^{(-\bar{r}+s)}(dx)}{g_n^{(s)}} \\
&\leq K_9 \frac{\int_0^1 \frac{\Gamma(n+1+\bar{r})}{\Gamma(\lceil ny \rceil+1)\Gamma(n-\lceil ny \rceil+\bar{r})} \rho^{(-\bar{r}+s)}(t) t^{\lceil ny \rceil}(1-t)^{n-\lceil ny \rceil+\bar{r}-1} dt}{g_n^{(s)}} \\
&\quad + K_9 \frac{\int_0^1 \frac{\Gamma(n+1+\bar{r})}{\Gamma(n+1)\Gamma(1+\bar{r})} x^n (1-x)^{\bar{r}} \nu^{(-\bar{r}+s)}(dx)}{g_n^{(s)}} \\
&\leq K_9 K_{10} \frac{\frac{\Gamma(n+1+\bar{r})\Gamma(\lceil ny \rceil+1-\alpha)}{\Gamma(\lceil ny \rceil+1)\Gamma(n+1+\bar{r}-\alpha)}}{g_n^{(s)}} + K_6 K_9 \frac{\frac{\Gamma(n+1+\bar{r})}{\Gamma(n+1)\Gamma(1+\bar{r})} n^{-\bar{r}}}{g_n^{(s)}} \\
&\leq K_{11} s^{-\alpha} n^{-\alpha},
\end{aligned}$$

where for the first inequality, we use (3.29) with $l = \lceil ny \rceil$, in the second inequality, we have used an argument of integration by parts and for the third inequality, we bound $\rho^{(-\bar{r}+s)}(x)$ by $K_{10} x^{-\alpha}$ and we also use Lemma 3.10. For the last inequality, we use (3.10).

Here K_{11} is a constant which depends only on $\bar{r}$ and $\nu^{(s)}$. Then we get

$$\mathbb{E}^{\nu^{(s)}}[\left(\frac{n}{n-X_1^{(n)}}\right)^r \mathbf{1}_{X_1^{(n)}\geq ny}] \leq (1-y)^{\bar{r}-r}\mathbb{E}^{\nu^{(s)}}[\left(\frac{n}{n-X_1^{(n)}}\right)^{\bar{r}} \mathbf{1}_{X_1^{(n)}\geq ny}] \leq K_{11}y^{-\alpha}(1-y)^{\bar{r}-r}n^{-\alpha},$$

which achieves the proof of the lemma. □

Remark 3.4. *If $r \geq \alpha + s$, this lemma is false. Assume that $s = 0, r \geq \alpha$ and for any fixed $0 < y < 1$, $n \geq \frac{1}{1-y}$, we have $ny \leq n - 1$ and it follows that*

$$\begin{aligned}
&P_1(m, n, s, y)\\
&\geq \mathbb{E}[\left(\left(\frac{n}{n-X_1^{(n)}}\right)^r - 1 - r\frac{X_1^{(n)}}{n} - \sum_{k=2}^{m}\frac{\prod_{i=0}^{k-1}(r+i)}{k!}(\frac{X_1^{(n)}}{n})^k\right)\mathbf{1}_{X_1^{(n)}=n-1}]\\
&= \mathbb{P}(X_1^{(n)} = n-1)\left(n^r - 1 - r\frac{n-1}{n} - \sum_{k=2}^{m}\frac{\prod_{i=0}^{k-1}(r+i)}{k!}(\frac{n-1}{n})^k\right)\\
&= \frac{\int_0^1 x^n\nu(dx)}{g_n}\left(n^r - 1 - r\frac{n-1}{n} - \sum_{k=2}^{m}\frac{\prod_{i=0}^{k-1}(r+i)}{k!}(\frac{n-1}{n})^k\right)\\
&\sim Cn^{-2\alpha}\left(n^r - 1 - r\frac{n-1}{n} - \sum_{k=2}^{m}\frac{\prod_{i=0}^{k-1}(r+i)}{k!}(\frac{n-1}{n})^k\right)
\end{aligned}$$

where C is a positive number. Then

$$\liminf_{n\to+\infty} n^\alpha P_1(m, n, s, y) \geq C, \forall 0 < y < 1.$$

Hence this remark justifies the constraint $0 \leq r < \alpha + s$.

D) Results that are used to prove Theorem 3.3.

Lemma 3.11. *Let $a > 0, b > 0$, $\beta \geq 1$. Then $0 < (a + b)^\beta \leq a^\beta + b^\beta + \beta 2^{\beta-1}ab^{\beta-1} + \beta 2^{\beta-1}ba^{\beta-1}$.*

Proof. If $0 \leq m \leq 1$, then

$$(1 + m)^\beta \leq 1 + \beta 2^{\beta-1}m \leq 1 + m^\beta + \beta 2^{\beta-1}m + \beta 2^{\beta-1}m^{\beta-1}.$$

We use that the function $m \mapsto (1 + m)^\beta$ is convex and that $\beta 2^{\beta-1}$ is the derivative of $(1 + m)^\beta$ at $m = 1$.

If $1 < m$, then

$$(1 + m)^\beta = m^\beta(1 + \frac{1}{m})^\beta \leq (m)^\beta(1 + \beta 2^{\beta-1}\frac{1}{m}) \leq 1 + m^\beta + \beta 2^{\beta-1}m + \beta 2^{\beta-1}m^{\beta-1}.$$

Hence for all $m \geq 0$,

$$(1+m)^\beta \leq 1 + m^\beta + \beta 2^{\beta-1} m + \beta 2^{\beta-1} m^{\beta-1}.$$

Then for all $a > 0, b > 0$,

$$(a+b)^\beta = a^\beta (1+\frac{b}{a})^\beta \leq a^\beta (1 + (\frac{b}{a})^\beta + \beta 2^{\beta-1}\frac{b}{a} + \beta 2^{\beta-1}(\frac{b}{a})^{\beta-1}) = a^\beta + b^\beta + \beta 2^{\beta-1} a b^{\beta-1} + \beta 2^{\beta-1} b a^{\beta-1}.$$

□

Acknowledgment

Jean-Stéthane Dhersin and Linglong Yuan benefited from the support of the "Agence Nationale de la Recherche": ANR MANEGE (ANR-09-BLAN-0215).

Bibliographie

[1] Arnason, E. (2004). Mitochondrial cytochrome b dna variation in the high-fecundity atlantic cod : trans-atlantic clines and shallow gene genealogy. *Genetics*, 166(4) :1871.

[2] Berestycki, J., Berestycki, N., and Limic, V. (2008). A small-time coupling between λ-coalescents and branching processes. *Ann. Inst. H. Poincaré Probab. Statist*, 44(2) :214–238.

[3] Berestycki, J., Berestycki, N., and Limic, V. (2011). Asymptotic sampling formulae and particle system representations for Λ-coalescents. *Arxiv preprint arXiv :1101.1875.*

[4] Berestycki, J., Berestycki, N., and Schweinsberg, J. (2007). Beta-coalescents and continuous stable random trees. *Ann. Probab.*, 35(5) :1835–1887.

[5] Berestycki, N. (2009). Recent progress in coalescent theory. *Ensaios Matematicos*, 16 :1–193.

[6] Birkner, M., Blath, J., Capaldo, M., Etheridge, A., Möhle, M., Schweinsberg, J., and Wakolbinger, A. (2005). Alpha-stable branching and beta-coalescents. *Electron. J. Probab.*, 10 :no. 9, 303–325 (electronic).

[7] Blum, M. G. B. and François, O. (2005). Minimal clade size and external branch length under the neutral coalescent. *Adv. in Appl. Probab.*, 37(3) :647–662.

[8] Bolthausen, E. and Sznitman, A.-S. (1998). On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197(2) :247–276.

[9] Boom, J., Boulding, E., and Beckenbach, A. (1994). Mitochondrial dna variation in introduced populations of pacific oyster, crassostrea gigas, in british columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(7) :1608–1614.

[10] Bovier, A. and Kurkova, I. (2007). Much ado about Derrida's GREM. In *Spin glasses*, volume 1900 of *Lecture Notes in Math.*, pages 81–115. Springer, Berlin.

[11] Breiman, L. (1992). Probability, classics in applied mathematics, vol. 7. *Society for Industrial and Applied Mathematics (SIAM), Pennsylvania.*

[12] Caliebe, A., Neininger, R., Krawczak, M., and Roesler, U. (2007). On the length distribution of external branches in coalescence trees : genetic diversity within species. *Theoretical Population Biology*, 72(2) :245–252.

[13] Delmas, J.-F., Dhersin, J.-S., and Siri-Jégousse, A. (2008). Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.*, 18(2) :997–1025.

[14] Dhersin, J.-S. and Möhle, M. (2013). On the external branches of coalescents with multiple collisions. *Electron. J. Probab.*, 18 :no. 40, 1–11.

[15] Dhersin, J.-S., Siri-Jégousse, A., Freund, F., and Yuan, L. (2013). On the length of an external branch in the beta-coalescent. *Stochastic Processes and their Applications*.

[16] Dhersin, J.-S. and Yuan, L. (2012). Asympotic behavior of the total length of external branches for beta-coalescents. *arXiv preprint arXiv :1202.5859*.

[17] Drmota, M., Iksanov, A., Moehle, M., and Roesler, U. (2007). Asymptotic results concerning the total branch length of the Bolthausen-Sznitman coalescent. *Stochastic Process. Appl.*, 117(10) :1404–1421.

[18] Durrett, R. (2008). *Probability models for DNA sequence evolution*. Probability and its Applications (New York). Springer, New York, second edition.

[19] Eldon, B. and Wakeley, J. (2006). Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics*, 172 :2621–2633.

[20] Foucart, C. and Hénard, O. (2013). Stable continuous-state branching processes with immigration and beta-fleming-viot processes with immigration. *Electron. J. Probab.*, 18 :no. 23, 1–21.

[21] Freund, F. and Möhle, M. (2009). On the time back to the most recent common ancestor and the external branch length of the Bolthausen-Sznitman coalescent. *Markov Process. Related Fields*, 15(3) :387–416.

[22] Fu, Y. and Li, W. (1993). Statistical tests of neutrality of mutations. *Genetics*, 133 :693–709.

[23] Gnedin, A., Iksanov, A., and Möhle, M. (2008). On asymptotics of exchangeable coalescents with multiple collisions. *J. Appl. Probab.*, 45(4) :1186–1195.

[24] Gnedin, A. and Yakubovich, Y. (2007). On the number of collisions in Λ-coalescents. *Electron. J. Probab.*, 12 :no. 56, 1547–1567 (electronic).

[25] Goldschmidt, C. and Martin, J. B. (2005). Random recursive trees and the Bolthausen-Sznitman coalescent. *Electron. J. Probab.*, 10 :no. 21, 718–745 (electronic).

[26] Hedgecock, D. (1994). 2.5 does variance in reproductive success limit effective population sizes of marine organisms ? *Genetics and evolution of aquatic organisms*, page 122.

[27] Janson, S. and Kersting, G. (2011). On the total external length of the kingman coalescent. *Electronic Journal of Probability*, 16 :2203–2218.

[28] Kallenberg, O. (2002). *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition.

[29] Kersting, G. (2012). The asymptotic distribution of the length of beta-coalescent trees. *The Annals of Applied Probability*, 22(5) :2086–2107.

[30] Kersting, G., Pardo, J. C., and Siri-Jégousse, A. (2013). Total internal and external lengths of the bolthausen-sznitman coalescent. *arXiv preprint arXiv :1302.1463*.

[31] Kersting, G., Stanciu, I., and Wakolbinger, A. (2012). The total external branch length of beta-coalescents. *arXiv preprint arXiv :1212.6070*.

[32] Kimura, M. (1969). The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, 61(4) :893–903.

[33] Kingman, J. (1982). The coalescent. *Stochastic Process. Appl.*, 13(3) :235–248.

[34] Kingman, J. (2000). Origins of the Coalescent 1974-1982. *Genetics*, 156(4) :1461–1463.

[35] Marynych, O. (2011). Stochastic recurrences and their applications to the analysis of partition-valued processes.

[36] Möhle, M. (2006). On the number of segregating sites for populations with large family sizes. *Adv. in Appl. Probab.*, 38(3) :750–767.

[37] Möhle, M. (2010). Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson-Dirichlet coalescent. *Stochastic Process. Appl.*, 120(11) :2159–2173.

[38] Pitman, J. (1999). Coalescents with multiple collisions. *Ann. Probab.*, 27(4) :1870–1902.

[39] Sagitov, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 36(4) :1116–1125.

[40] Schweinsberg, J. (2003). Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Process. Appl.*, 106(1) :107–139.

[41] Yuan, L. (2013). On the measure division construction of Λ-coalescents. *arXiv preprint arXiv :1302.1083*.

Chapitre 4

Sur la construction par division de mesure de Λ-coalescents

Version non modifiée de l'article *On the measure division construction of Λ-coalescents* soumis à *Markov Processes and Related Fields*.

4.1 Introduction

4.1.1 Motivation and main results

Let $\mathbb{N} := \{1, 2, \cdots\}$, Ω be a subset of $\mathbb{N}$ and π a partition of Ω such that $|\pi| < +\infty$ ($|\pi|$ denotes the number of blocks in π). The Λ-coalescent process starting from π, introduced independently by (27) and (28), is denoted by $\Pi^{(\pi)} := (\Pi^{(\pi)}(t))_{t\geq 0}$, where $\Pi^{(\pi)}(0) = \pi$ and Λ is a finite measure on $[0,1]$. Here we specify that a finite measure on $[0,1]$ can be a null measure and hence its total mass is a non-negative real value. If $\pi = \{\{1\}, \{2\}, \cdots, \{n\}\}$, i.e., the set of first n singletons, then the process is simply denoted by $\Pi^{(n)}$. In this paper, we will frequently use two other notations Λ_1, Λ_2 for finite measures. We define then $\Pi^{(1,n)}$ as the Λ_1-coalescent and $\Pi^{(2,n)}$ the Λ_2-coalescent, both taking $\{\{1\}, \{2\}, \cdots, \{n\}\}$ as initial value.

This process $\Pi^{(\pi)}$ is a continuous time Markov process with càdlàg trajectories taking values in the set of partitions of Ω. More precisely: Assume that at time t, $\Pi^{(\pi)}(t)$ has b blocks, then after a random exponential time with parameter g_b

$$g_b = \sum_{k=2}^{b} \binom{b}{k} \lambda_{b,k}, \quad \text{where} \quad \lambda_{b,k} = \int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx), \tag{4.1}$$

$\Pi^{(\pi)}$ encounters a collision and the probability for a group of $k (2 \leq k \leq b)$ blocks to be merged into a bigger block with the other $b-k$ blocks unchanged is

$$\frac{\lambda_{b,k}}{g_b}.$$

Then

$$p_{b,b-k+1} := \frac{\binom{b}{k}\lambda_{b,k}}{g_b} \tag{4.2}$$

is the probability to have $b-k+1$ blocks right after the collision. This definition gives the exchangeability of blocks. In particular, for any permutation ρ on $\{1, 2, \cdots, n\}$, $\rho \circ \Pi^{(n)} \overset{(d)}{=} \Pi^{(n)}$.

Remark that if $\Lambda(\{0\}) = 0$, we have the following formula:

$$g_b = \int_0^1 (1-(1-x)^b - bx(1-x)^{b-1})x^{-2}\Lambda(dx). \tag{4.3}$$

The definition shows that the law of $\Pi^{(\pi)}$ is determined by the initial value π and the measure Λ which is hence called characteristic measure.

Notice that Ω can be an abstract set and the coalescing mechanism works all the same. The reason why one takes Ω as a subset of $\mathbb{N}$ relies on its applications in the genealogies of populations. We take $\Pi^{(n)}$ as an example where $\Omega = \{1, 2, \cdots, n\}$. At time

0, we have $\Pi^{(n)}(0) = \{\{1\}, \{2\}, \cdots, \{n\}\}$ which is interpreted as a sample of n individuals labelled from 1 to n . If at time t, $\Pi^{(n)}$ has its first coalescence where $\{1\}$ and $\{2\}$ are merged together with the others unchanged, then $\Pi^{(n)}(t) = \{\{1, 2\}, \{3\}, \cdots, \{n\}\}$ which is interpreted as getting the MRCA (most recent common ancestor) $\{1, 2\}$ of individuals 1 and 2 with the others unchanged at that time. Hence $\{1, 2, \cdots, n\}$ is an absorption state of $\Pi^{(n)}$ and is the MRCA of all individuals. For more details, we refer to (22; 24) or (1; 6; 14; 19).

Let $1 \leq m \leq n$ and σ the restriction from $\{1, 2, \cdots, n\}$ to $\{1, 2, \cdots, m\}$. We have the consistency property: $\sigma \circ \Pi^{(n)} \stackrel{(d)}{=} \Pi^{(m)}$ (see (27)). According to this property and exchangeability of blocks, if π' is a subset of π, then the restriction of $\Pi^{(\pi)}$ from π to π' has the same distribution as that of $\Pi^{(\pi')}$. We can also define $\Pi^{(\pi)}$ when $|\pi| = +\infty$ by using the consistency property and the definition in finite cases (see (27)).

Let $|\Pi^{(n)}|$ be the block counting process associated to $\Pi^{(n)}$ such that $|\Pi^{(n)}(t)|$ is the number of blocks of $\Pi^{(n)}(t)$ for any $t \geq 0$. Then it decreases from n at time 0. We denote by $X_1^{(n)}$ the decrease of number of blocks at the first coalescence. For $i \in \{1, \ldots, n\}$, we define

$$T_i^{(n)} := \inf \left\{ t \geq 0 | \, \{i\} \notin \Pi_t^{(n)} \right\}$$

the length of the ith external branch and $T^{(n)}$ the length of a randomly chosen external branch. By exchangeability, $T_i^{(n)} \stackrel{(d)}{=} T^{(n)}$. We denote by $L_{ext}^{(n)} := \sum_{i=1}^n T_i^{(n)}$ the total external branch length of $\Pi^{(n)}$, and by $L_{total}^{(n)}$ the total branch length.

The following four classes of Λ-coalescents have been largely studied. We give the results concerning $T^{(n)}$, which show a common regularity that we will discuss later.

— $\Lambda = \delta_0$: Kingman coalescent ((22), (23)). Then $nT^{(n)}$ is asymptotically distributed with density function $\frac{8}{(2+x)^3}\mathbf{1}_{x\geq 0}$ ((4), (8), (20)).

— $\Lambda = \Lambda^{leb}$: Bolthausen-Sznitman coalescent ((5)). Here Λ^{leb} denotes the Lebesgue measure on $[0, 1]$. Then $(\ln n)T^{(n)}$ converges in distribution to $Exp(1)$ (we denote by $Exp(r), r > 0$, the exponential variable with parameter r)((16; 11)).

— $\Lambda(dx)/dx = \frac{x^{a-1}(1-x)^{b-1}}{Beta(a,b)}\mathbf{1}_{0\leq x\leq 1}, 0 < a < 1, b > 0$: $Beta(a, b)$-coalescent. Here $Beta(\cdot, \cdot)$ denotes Euler's beta function. Then $n^{1-a}T^{(n)}$ converges in distribution to a random variable $T(a, b)$ which has density function $\frac{\Gamma(a+b)}{(1-a)\Gamma(b)}(1+\frac{\Gamma(a+b)}{(2-a)\Gamma(b)}x)^{-\frac{3-2a}{1-a}}\mathbf{1}_{x\geq 0}$ ((12)).

— $\int_0^1 x^{-1}\Lambda(dx) < +\infty$: These processes are called coalescents without proper frequencies ((27)). One example is $Beta(a, b)$-coalescents with $a > 1, b > 0$ (see (27), (29)). Then $\left(\int_0^1 x^{-1}\Lambda(dx)\right) T^{(n)}$ converges in distribution to $Exp(1)$ ((18), (26)).

We see a common property for the last three cases concerning one external branch length which is that the normalization factor for $T^{(n)}$ is $\mu^{(n)} := \int_{\frac{1}{n}}^1 x^{-1}\Lambda(dx)$. More precisely,

— Bolthausen-Sznitman coalescent: Notice that $\mu^{(n)} = \ln n$. Hence directly we have $\mu^{(n)}T^{(n)} \stackrel{(d)}{\to} Exp(1)$.

— $Beta(a,b)$-coalescent with $0 < a < 1, b > 0$:

$$\mu^{(n)} = \int_{\frac{1}{n}}^{1} x^{-1}\Lambda(dx) = \int_{\frac{1}{n}}^{1} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-2}(1-x)^{b-1}dx = \frac{\Gamma(a+b)}{(1-a)\Gamma(a)\Gamma(b)} n^{1-a} + O(1).$$

Hence $\mu^{(n)}T^{(n)}$ converges in distribution to $T(a,b)\Gamma(a+b)/(1-a)\Gamma(a)\Gamma(b)$.

— If $\int_0^1 x^{-1}\Lambda(dx) < +\infty$, then $\lim_{n\to+\infty} \frac{\mu^{(n)}}{\int_0^1 x^{-1}\Lambda(dx)} = 1$. Hence $\mu^{(n)}T^{(n)}$ converges in distribution to $Exp(1)$.

Kingman coalescent can be viewed as the formal limit of $Beta(a,b)$-coalescent with $0 < a < 1, b > 0$ when a tends to 0, since the measure $\frac{x^{a-1}(1-x)^{b-1}dx}{Beta(a,b)}\mathbf{1}_{0\leq x\leq 1}$ tends weakly to the Dirac measure on $\{0\}$. The normalization factor in the case of $Beta(a,b)$-coalescent is n^{1-a}, and of Kingman coalescent is n. Then we see that these two factors show also some kind of continuity as a tends to 0. We can formally take n as $\mu^{(n)}$ in the case of Kingman coalescent.

Therefore $\mu^{(n)}$ is characteristic for the randomly chosen external branch length in those processes considered. Notice that $\mu^{(n)}$ concerns only the measure $\Lambda\mathbf{1}_{[\frac{1}{n},1]}$, so it is natural to think about the influences of measures $\Lambda\mathbf{1}_{[\frac{1}{n},1]}$ and $\Lambda\mathbf{1}_{[0,\frac{1}{n})}$ on the external branch lengths. More generally, if $\Lambda = \Lambda_1 + \Lambda_2$, how can we evaluate each influence on the construction of the whole Λ-coalescent? If Λ_1 is "small" enough, we can imagine that $\Pi^{(n)}$ looks like $\Pi^{(2,n)}$ (recall that $\Pi^{(2,n)}$ is the Λ_2-coalescent). In this case, we call Λ_1 the noise measure and Λ_2 the main measure. To separate Λ_1 and Λ_2, we introduce in the next section the "measure division construction" of a Λ-coalescent. The idea of this construction can be at least tracked back to (2) where the authors consider also a coupling of two finite measures on $[0,1]$. The difference relies on the block labelling.

The main results are as follows:

Theorem 4.1. *If* Λ *satisfies:*

$$\lim_{n\to+\infty} \frac{g_n}{n\mu^{(n)}} = 0, \tag{4.4}$$

then $\mu^{(n)}T^{(n)} \stackrel{(d)}{\to} Exp(1)$.

Remark 4.1. — *Condition* (4.4) *implies that* $\Lambda(\{0\}) = 0$. *Indeed, if* $\Lambda(\{0\}) > 0$, *then* $g_n \geq \binom{n}{2}\Lambda(\{0\})$ *and* $\mu^{(n)} \leq n\Lambda((0,1])$. *Then* (4.4) *is invalid.*

— *The class of coalescents satisfying condition* (4.4) *does not contain the* $Beta(a,b)$*-coalescents with* $0 < a < 1$ *and* $b > 0$. *The following conjecture uses a description similar to condition* (4.4) *to include them:*

Conjecture: *Let* $c > 0$. *If*

$$\lim_{n\to+\infty} \frac{g_n}{n\mu^{(n)}} = c,$$

then $\mu^{(n)}T^{(n)} \stackrel{(d)}{\to} T_c$*, where* T_c *is a random variable with density* $\Gamma(2-\alpha^*)(1+cx)^{-\frac{\alpha^*}{\alpha^*-1}-1}\mathbf{1}_{x\geq 0}$*. Here* α^* *is the unique solution of the equation*

$$\frac{(\alpha-1)\Gamma(2-\alpha)}{\alpha} = c.$$

This conjecture is true for Beta(a,b)*-coalescents with* $0<a<1, b>0$*. In this case, we have* $c=\frac{(1-a)\Gamma(a)}{2-a}$*. The coalescents, which are more general than but similar to Beta*(a,b)*-coalescents with* $0<a<1, b>0$*, studied in (12) also satisfy this conjecture.*

Examples: We give a short list of typical examples satisfying condition (4.4) which are processes without proper frequencies or similar to Bolthausen-Szitman coalescent. Define $\bar{\mu}^{(n)} := \int_{\frac{1}{n}}^1 x^{-2}\Lambda(dx)$.

Ex 1: $\int_0^1 x^{-1}\Lambda(dx) < +\infty$: It suffices to prove that $\lim_{n\to+\infty}\frac{g_n}{n}=0$. Recalling the expression (4.3) of g_n, we have, for $n\geq 2$,

$$\begin{aligned}
\frac{g_n}{n} &= \frac{\int_0^1(1-(1-x)^n-nx(1-x)^{n-1})x^{-2}\Lambda(dx)}{n} \\
&= \frac{\int_{\frac{1}{n}}^1(1-(1-x)^n-nx(1-x)^{n-1})x^{-2}\Lambda(dx)}{n} + \frac{\int_0^{\frac{1}{n}}(1-(1-x)^n-nx(1-x)^{n-1})x^{-2}\Lambda(dx)}{n} \\
&\leq \frac{\bar{\mu}^{(n)}}{n} + \frac{\int_0^{\frac{1}{n}} n^2\Lambda(dx)}{n}. \qquad (4.5)
\end{aligned}$$

The second term $\frac{\int_0^{\frac{1}{n}} n^2\Lambda(dx)}{n} = \int_0^{\frac{1}{n}} n\Lambda(dx) \leq \int_0^{\frac{1}{n}} x^{-1}\Lambda(dx) \to 0$. For the first term, let $\epsilon>0$ and $M=1/\epsilon$, then

$$\begin{aligned}
\frac{\bar{\mu}^{(n)}}{n} &= \frac{\int_{M/n}^1 x^{-2}\Lambda(dx)}{n} + \frac{\int_{\frac{1}{n}}^{M/n} x^{-2}\Lambda(dx)}{n} \\
&\leq \frac{\int_{M/n}^1 x^{-1}\Lambda(dx)}{M} + \int_{\frac{1}{n}}^{M/n} x^{-1}\Lambda(dx) \\
&\leq \epsilon\int_0^1 x^{-1}\Lambda(dx) + \int_{\frac{1}{n}}^{M/n} x^{-1}\Lambda(dx).
\end{aligned}$$

Notice that $\epsilon\int_0^1 x^{-1}\Lambda(dx)$ can be arbitrarily small and $\int_{\frac{1}{n}}^{M/n} x^{-1}\Lambda(dx)$ tends to 0 as n tends to $+\infty$. Then we get that $\frac{\bar{\mu}^{(n)}}{n}$ tends to 0. Hence if $\int_0^1 x^{-1}\Lambda(dx) < +\infty$, condition (4.4) is satisfied.

Ex 2: Bolthausen-Sznitman coalescent: In this case, it is straightforward to prove that $g_n = n-1$ and $\mu^{(n)} = \ln n$, then $\lim_{n\to+\infty} \frac{g_n}{n\mu^{(n)}} = \lim_{n\to+\infty} \frac{n-1}{n\ln n} = 0$.

Ex 3: Λ has a density function f_Λ on $[0,r)$ where $0<r<1$ and there exists a positive number M such that $f_\Lambda < M$ on $[0,r)$: This kind of processes can be considered as being *dominated* by the Bolthausen-Sznitman coalescent.

If $\int_0^1 x^{-1}\Lambda(dx) < +\infty$, we turn back to the first example. If $\int_0^1 x^{-1}\Lambda(dx) = +\infty$, then we have $g_n \le 2M(n-1)$ for n large enough, hence $\limsup_{n\to+\infty} \frac{g_n}{n\mu^{(n)}} \le \lim_{n\to+\infty} \frac{2M(n-1)}{n\mu^{(n)}} = 0$. It turns out that this kind of coalescent also satisfies condition (4.4).

Ex 4: Λ has a density function $f_\Lambda(x) = p(\ln\frac{1}{x})^q$ on $[0,r)$ where $0<r<1$ and p,q are positive numbers: Using (4.5), we have

$$\frac{g_n}{n\mu^{(n)}} \le \frac{\bar{\mu}^{(n)}}{n\mu^{(n)}} + \frac{\int_0^{\frac{1}{n}} n^2\Lambda(dx)}{n\mu^{(n)}}, \forall n \ge 2.$$

For two real sequences $(x_n)_{n\ge1}, (y_n)_{n\ge1}$, we write $x_n \asymp y_n$, if there exist two positive constants c, C such that $cy_n \le x_n \le Cy_n$ for n large enough. Then it is not difficult to find out that $\mu^{(n)} \asymp (\ln n)^{q+1}$, $\bar{\mu}^{(n)} \asymp n(\ln n)^q$, $\int_0^{\frac{1}{n}} n^2\Lambda(dx) \asymp n(\ln n)^q$. Hence we get $\frac{g_n}{n\mu^{(n)}} \to 0$.

Theorem 4.2. *If Λ satisfies condition (4.4) and $\int_0^1 x^{-1}\Lambda(dx) = +\infty$, then we have:*

$$\mu^{(n)}(T_1^{(n)}, T_2^{(n)}, \cdots, T_n^{(n)}, 0, 0, \cdots) \xrightarrow{(d)} (e_1, e_2, \cdots), \tag{4.6}$$

where $(e_i)_{i\in\mathbb{N}}$ are independently distributed as $Exp(1)$.

Remark 4.2. *The same result has been proved for Bolthausen-Sznitman coalescent in (11). The authors have used a moment method. We can apply this theorem to Example 4 and Example 3 when $\int_0^1 x^{-1}\Lambda(dx) = +\infty$. If $\int_0^1 x^{-1}\Lambda(dx) < +\infty$, then (4.6) is not true and there is no more asymptotic independence (see (26)).*

The following three corollaries have also been proved for Bolthausen-Sznitman coalescent (see (11), (13), (17)).

Corollary 4.1. *If Λ satisfies condition (4.4), then for any $r \in \mathbb{R}^+$,*

$$\lim_{n\to+\infty} \mathbb{E}[(\mu^{(n)}T^{(n)})^r] = \mathbb{E}[e_1^r],$$

where e_1 is distributed as $Exp(1)$. Moreover, if $\int_0^1 x^{-1}\Lambda(dx) = +\infty$, then for any $k \in \mathbb{N}$ and any $(r_1, r_2, \cdots, r_k) \in \{\mathbb{R}^+\}^k$, we have:

$$\lim_{n\to+\infty} \mathbb{E}[\prod_{i=1}^k (\mu^{(n)}T_i^{(n)})^{r_i}] = \mathbb{E}[\prod_{i=1}^k e_i^{r_i}], \tag{4.7}$$

where $(e_i)_{1\le i\le k}$ are independently distributed as $Exp(1)$.

Corollary 4.2. *If Λ satisfies condition* (4.4) *and* $\int_0^1 x^{-1}\Lambda(dx) = +\infty$, *then the total external branch length* $L_{ext}^{(n)}$ *satisfies:* $\mu^{(n)} L_{ext}^{(n)}/n$ *converges in* L^2 *to* 1.

Corollary 4.3. *If Λ satisfies condition* (4.4) *and* $\int_0^1 x^{-1}\Lambda(dx) = +\infty$, *then the total branch length* $L_{total}^{(n)}$ *satisfies:* $\mu^{(n)} L_{total}^{(n)}/n$ *converges in probability to* 1.

Remark 4.3.
— *In fact, we will prove that* $\lim_{n\to+\infty} \mathbb{E}[\mu^{(n)} L_{total}^{(n)}/n] = 1$. *Notice that Corollary 4.1 gives* $\lim_{n\to+\infty} \mathbb{E}[\mu^{(n)} L_{ext}^{(n)}/n] = 1$. *Hence we deduce this corollary using Corollary 4.2.*
— *If* $\int_0^1 x^{-1}\Lambda(dx) < +\infty$, *then (4.7) and Corollaries (4.2) and (4.3) are not true (see again (26)).*

4.1.2 Organization

In section 2, we introduce the main object of this paper: the measure division construction. At first, one needs to define the restriction by the smallest element which serves as a preliminary step of measure division construction. In the same section, we then introduce the two-type Λ-coalescent which is defined using the measure division construction. This process gives a label *primary* or *secondary* to every block and its every element of a normal Λ-coalescent. Using this process, we can see more clearly the coalescent times of some singletons. For a technical use, we then give a tripling to estimate the number of blocks at small times of $\Pi^{(1,n)}$ which is related to the noise measure Λ_1.

In section 3, we at first give a characterization of the condition (4.4). Then we apply the general results obtained in section 2 to those processes satisfying (4.4). Finally, we give all the proofs for the results presented in the section 1.

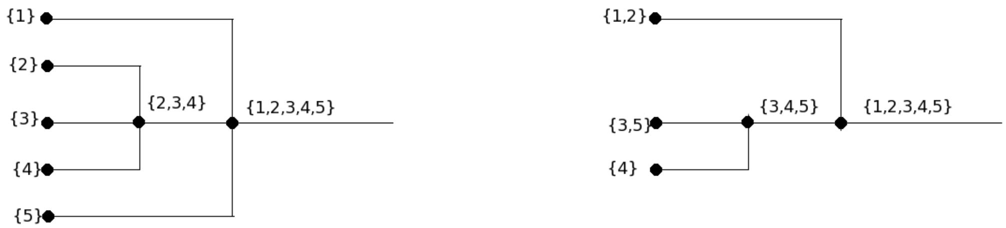
4.2 Measure division construction

4.2.1 Restriction by the smallest element.

Let $\xi_n = \{A_1, \cdots, A_{|\xi_n|}\}$, $\chi_n = \{B_1, \cdots, B_{|\chi_n|}\}$ be two partitions of $\{1, 2, \cdots, n\}$. We define s_i^A (resp. s_i^B) as the smallest number in the block A_i (resp. B_i). We define also the notation $\xi_n \preceq \chi_n$, if $|\chi_n| \le |\xi_n|$ and for any $1 \le i \le |\chi_n|, B_i = \cup_{j\in I_i} A_j$, where $\{I_i\}_{1\le i\le|\chi_n|}$ is a partition of $\{1, 2, \cdots, |\xi_n|\}$. Roughly speaking, ξ_n is finer than χ_n.

If $\xi_n \preceq \chi_n$, we define the stochastic process $\bar{\Pi}^{(\chi_n)}$ which is the restriction by the smallest element of $\Pi^{(\xi_n)}$ from ξ_n to χ_n:
— $\bar{\Pi}^{(\chi_n)}(0) = \chi_n$;
— For any $t \ge 0$, if $\Pi^{(\xi_n)}(t) = \{D_i\}_{1\le i\le|\Pi^{(\xi_n)}(t)|}$, where D_i denotes a block, then

$$\bar{\Pi}^{(\chi_n)}(t) = \{ \bigcup_{s_j^B \in D_i} B_j \}_{1\le i\le|\Pi^{(\xi_n)}(t)|},$$



(a) $\Pi^{(5)}$

(b) A restriction by the smallest element of $\Pi^{(5)}$ from $\{\{1\}, \cdots, \{5\}\}$ to $\{\{1,2\}, \{3,5\}, \{4\}\}$

Figure 4.1: Restriction by the smallest element

where the empty sets in $\bar{\Pi}^{(\chi_n)}(t)$ are removed.

Notice that the restriction by the smallest element is defined from path to path (see Figure 4.1).

Lemma 4.1. *$\bar{\Pi}^{(\chi_n)}$ has the same distribution as $\Pi^{(\chi_n)}$.*

Proof. Every block in χ_n is identified by its smallest element which belongs to a unique block in ξ_n. Hence for any B_i in χ_n, there exists a unique A_{τ_i} such that $A_{\tau_i} \in \xi_n$, $A_{\tau_i} \subset B_i$ and $s^A_{\tau_i} = s^B_i$ with $\tau_i \in \{1, 2, \cdots, |\xi_n|\}$. Let $\chi'_n = \{A_{\tau_i}\}_{1 \leq i \leq |\chi_n|}$ and define a new process $\hat{\Pi}^{(\chi'_n)}$ as follows:

— $\hat{\Pi}^{(\chi'_n)}(0) = \chi'_n$.

— For any $t \geq 0$, if $\Pi^{(\xi_n)}(t) = \{D_i\}_{1 \leq i \leq |\Pi^{(\xi_n)}(t)|}$, then

$$\hat{\Pi}^{(\chi'_n)}(t) = \{ \bigcup_{s^A_{\tau_j} \in D_i} A_{\tau_j} \}_{1 \leq i \leq |\Pi^{(\xi_n)}(t)|},$$

where the empty sets in $\hat{\Pi}^{(\chi'_n)}(t)$ are removed.

It is easy to see that $\hat{\Pi}^{(\chi'_n)}$ is a natural restriction of $\Pi^{(\xi_n)}$ from ξ_n to χ'_n. By the consistency property, one gets $\hat{\Pi}^{(\chi'_n)} \overset{(d)}{=} \Pi^{(\chi'_n)}$. In the construction of $\hat{\Pi}^{(\chi'_n)}$ and $\bar{\Pi}^{(\chi_n)}$, what is determinant is the smallest element in each block. Hence to obtain $\bar{\Pi}^{(\chi_n)}$ from $\hat{\Pi}^{(\chi'_n)}$, at time 0, one needs to complete every A_{τ_i} by some other numbers larger than $s^A_{\tau_i}$ to get B_i and then follow the evolution of $\hat{\Pi}^{(\chi'_n)}$. It turns out that $\bar{\Pi}^{(\chi_n)}$ is a coalescent process with initial value χ_n. Hence we can conclude. □

4.2.2 Measure division construction

Let $\Lambda, \Lambda_1, \Lambda_2$ be three finite measures such that $\Lambda = \Lambda_1 + \Lambda_2$. We denote by $\Pi_{1,2}^{(n)} := (\Pi_{1,2}^{(n)}(t))_{t\geq 0}$ the stochastic process constructed by the measure division construction using Λ_1 and Λ_2. Here the index $(1,2)$ is for $\Lambda = \Lambda_1 + \Lambda_2$ with Λ_1 called noise measure and Λ_2 main measure. Recall that $\Pi^{(1,n)}$ is the Λ_1-coalescent with $\Pi^{(1,n)}(0) = \{\{1\}, \{2\}, \cdots, \{n\}\}$.

- Step 0: Given a realization or a path Π of $\Pi^{(1,n)}$, we set $\Pi_{1,2}^{(n)}(t) = \Pi(t)$, for any $t \geq 0$. We set also $t_0 = 0$.
- Step 1: Let $t_1, t_2, \cdots$ be the coalescent times after t_0 of $\Pi_{1,2}^{(n)}$ (if there is no collision after t_0, we set $t_i = +\infty, i \geq 1$). Within $[t_0, t_1)$, $\Pi_{1,2}^{(n)}$ is constant. Then we run an independent Λ_2-coalescent with initial value $\Pi_{1,2}^{(n)}(t_0)$ from time t_0.
  - If the Λ_2-coalescent has no collision on $[t_0, t_1)$, we pass to $[t_1, t_2)$. Similarly, we construct another independent Λ_2-coalescent with initial value $\Pi_{1,2}^{(n)}(t_1)$ from time t_1, and so on.
  - Otherwise, we go to the next step.
- Step 2: If finally within $[t_{i-1}, t_i)$, the related independent Λ_2-coalescent has its first collision at time t_* and its value at t_* is ξ. We then modify $(\Pi_{1,2}^{(n)}(t))_{t\geq 0}$ in the following way:
  - We change nothing for $0 \leq t < t_*$.
  - Let $\Pi' = (\Pi'(t), t \geq t_*)$ be the restriction by the smallest element of $(\Pi_{1,2}^{(n)}(t))_{t\geq t_*}$ from $\Pi_{1,2}^{(n)}(t_*)$ to ξ. Then let $(\Pi_{1,2}^{(n)}(t))_{t\geq t_*} = (\Pi'(t))_{t\geq t_*}$ and go to the step 1 by taking t_* as a new starting point. Notice that, due to Lemma 4.1, $(\Pi_{1,2}^{(n)}(t))_{t\geq t_*}$ has the same distribution as a Λ_1-coalescent from time t_* with initial value ξ.

Remark 4.4. — *The measure division construction works path by path.*
— *If we take $\Lambda_1 = 0$ as noise measure and $\Lambda_2 = \Lambda$ as main measure, then $\Pi^{(1,n)}(t) = \{\{1\}, \{2\}, \cdots, \{n\}\}$ for any $t \geq 0$ and $\Pi_{1,2}^{(n)} \stackrel{(d)}{=} \Pi^{(n)}$.*

Theorem 4.3. *Let Λ, Λ_1 and Λ_2 be three finite measures and $\Lambda = \Lambda_1 + \Lambda_2$. Then we have $\Pi_{1,2}^{(n)} \stackrel{(d)}{=} \Pi^{(n)}$.*

Proof. Let t be a coalescent time of $\Pi_{1,2}^{(n)}$. We consider the time of the next coalescence and the value at that moment. In the measure division construction of $\Pi_{1,2}^{(n)}$, we can see appearing two independent processes with one being a Λ_1-coalescent with initial value $\Pi_{1,2}^{(n)}(t)$ and the other one being a Λ_2-coalescent with initial value $\Pi_{1,2}^{(n)}(t)$ from time t. The process $\Pi_{1,2}^{(n)}$ gets the next coalescence whenever one of them first encounters a coalescence and picks up the value of the process at that moment. Then we follow the same procedure from the new coalescent time of $\Pi_{1,2}^{(n)}$. It is easy to see that $\Pi_{1,2}^{(n)}$ behaves in the same way as $\Pi^{(n)}$. Hence we can conclude. □

Remark 4.5. *The theorem shows that if we exchange the noise measure and the main measure, the distribution of the process is not changed and is uniquely determined by their sum.*

Remark 4.6. *The measure division construction also works for more than two measures. If there are $k(k \geq 2)$ finite measures $\{\Lambda_i\}_{1\leq i\leq k}$ and $\Lambda = \sum_{i=1}^{k} \Lambda_i$, one can get a stochastic process by first giving a realization of $\Pi^{(1,n)}$ which will be modified by Λ_2 in the way described in the measure division construction, and then we apply Λ_3 on the modified process, etc. The equivalence in distribution can be obtained in a recursive way.*

We give a corollary to show an immediate application of the measure division construction. The following corollary is essentially the same as Lemma 3.2 in (2). But we prove it again in our way.

Corollary 4.4. *Let Λ_1, Λ_2 be two finite measures such that $\Lambda_1 \leq \Lambda_2$, then one can construct $\Pi^{(1,n)}$ and $\Pi^{(2,n)}$ such that $|\Pi^{(2,n)}(t)| \leq |\Pi^{(1,n)}(t)|$ for all $t \geq 0$.*

Proof. $\Pi^{(2,n)}$ can be regarded as the measure constructed process by imposing the measure $\Lambda_2 - \Lambda_1$ on the paths of $\Pi^{(1,n)}$. Then we can deduce this corollary. □

4.2.3 Two-type Λ-coalescents.

Definitions

Let $\Lambda, \Lambda_1, \Lambda_2$ be three finite measures and $\Lambda = \Lambda_1 + \Lambda_2$ and Λ_2 satisfies $\int_0^1 x^{-2}\Lambda_2(dx) < +\infty$. A two-type Λ-coalescent, denoted by $\tilde{\Pi}_{1,2}^{(n)}$, is to give a label *primary* or *secondary* to every block and also to its every element at any time t of a normal Λ-coalescent. A block is *secondary* if and only if every element in this block is *secondary.* The construction is via the measure division construction. Let $(\eta_i^{(2)})_{i\geq 1}$ be independent random variables following the distribution of $\frac{x^{-2}\Lambda_2(dx)}{\int_0^1 x^{-2}\Lambda_2(dx)}$, $(e_i^{(2)})_{i\geq 1}$ i.i.d copies of $Exp(\int_0^1 x^{-2}\Lambda_2(dx))$ and $(S_i^{(2)})_{i\geq 1} = (\sum_{j=1}^{i} e_j^{(2)})_{i\geq 1}$.

Construction of a two-type Λ-coalescent:

— Step 0: We pick a realization or a path Π of $\Pi^{(1,n)}$. Every element and every block of Π at any time is labeled *primary.* We also fix independent realizations of $(\eta_i^{(2)})_{i\geq 1}$ and $(S_i^{(2)})_{i\geq 1}$. Let $\tilde{\Pi}_{1,2}^{(n)}$ be the path Π with labels.

— Step 1: At time $S_1^{(2)}$, every block of $\tilde{\Pi}_{1,2}^{(n)}(S_1^{(2)})$ is independently marked "Head" with probability $\eta_1^{(2)}$ and "Tail" with probability $1 - \eta_1^{(2)}$. Every element in a "Head" block is then labelled *secondary.* All blocks marked "Head" are merged into a bigger block, provided that there are at least two "Head"s. In this case, we use the restriction by the smallest element to modify $\tilde{\Pi}_{1,2}^{(n)}$ at time $S_1^{(2)}$in the same way as in the measure division construction in section 2.2. We still call the modified path $\tilde{\Pi}_{1,2}^{(n)}$ and then forward to the time $S_2^{(2)}$ and do the same operations. This procedure can be continued until MRCA.

It is easy to verify that without labels, $\tilde{\Pi}_{1,2}^{(n)}$ has the same distribution as $\Pi^{(n)}$. We call $(S_i^{(2)})_{i\geq 1}$ the *marking times*. We define $L_i^{(2,n)}$ as the *first marking time* of $\{i\}$ when $\{i\}$ is marked "Head " for the first time. Let $L_i^{(2,n)} = +\infty$, if $\{i\}$ is never marked as "Head" .

Remark 4.7. *If $\Pi = \{\{1\}, \cdots, \{n\}\}$, then we get a coupling between Λ_2-coalescent and its related annihilator process (see (13)). In this case, the whole process without labels is the Λ_2-coalescent and the restriction to primary elements and blocks is the annihilator process.*

Coalescent times and *first marking times*

The above construction of two-type coalescents shows that coalescences happen only at the *marking times*. This property will help us to understand the coalescent times of singletons in terms of their *first marking times*.

Lemma 4.2. *Let Π be the path of $\Pi^{(1,n)}$ chosen at the Step 0 of the construction of two-type Λ coalescent. Assume that at some time $t > 0$, $\{1\} \in \Pi(t)$, $|\Pi(t)| = m$ with $2 \leq m \leq n$. Let $P_{1,2}^{(n,m)}(t)$ be the probability for $\{1\}$ to be coalesced at its first marking time within $[0,t)$. Then we have*

$$P_{1,2}^{(n,m)}(t) \geq P_t^{(2,m)} := \sum_{i=1}^{+\infty} \mathbb{E}[\mathbf{1}_{S_i^{(2)}<t} \Delta_i^{(2)} \left(1 - (1 - \Delta_i^{(2)})^{m-1}\right)], \tag{4.8}$$

where $\Delta_1^{(2)} = \eta_1^{(2)}$; $\Delta_i^{(2)} = \eta_i^{(2)} \prod_{j=1}^{i-1}(1 - \eta_j^{(2)})$ for $i > 1$. Notice that the parameter n is hidden in $P_t^{(2,m)}$.

Proof. Let $i_1, \cdots, i_m$ be the m smallest elements respectively in each block at time t with $1 = i_1 \leq i_2 \leq \cdots \leq i_m \leq n$.

Conditional on $(S_i^{(2)}, \eta_i^{(2)})_{i\geq 1}$, $\Delta_i^{(2)}$ is the probability for $\{1\}$ to have its *first marking time* at $S_i^{(2)}$ (assume that $S_i^{(2)} \leq t$). To let $\{1\}$ be coalesced at $S_i^{(2)}$, one needs also at least one other block marked "Head" at that time. To get a lower bound of $P_{1,2}^{(n,m)}(t)$, one can consider the probability to have at least one *primary* block containing one element of $\{i_2, \cdots, i_m\}$ to be marked "Head" at that time and this probability is $1-(1-\Delta_i^{(2)})^{m-1}$. □

Lemma 4.3. *In addition to the assumptions in the previous lemma, we assume further that for every fixed k such that $1 \leq k \leq m$, $\{i\} \in \Pi(t)$ for $1 \leq i \leq k$. Define the probability $P_{1,2}^{(n,m,k)}(t)$ for every $\{i\}$ to be coalesced at its first marking time within $[0,t)$. Then we have*

$$P_{1,2}^{(n,m,k)}(t) \geq 1 - k(1 - P_t^{(2,m)}). \tag{4.9}$$

Proof. Let $E = \{\forall 1 \le i \le k, \{i\} \in \Pi(t); |\Pi(t)| = m\}$, which denotes the assumptions of $\Pi(t)$ in this Lemma. Then

$$\begin{aligned}
P_{1,2}^{(n,m,k)}(t) &= \mathbb{P}(\{1\}, \cdots, \{k\} \text{ coalesce at their } \textit{first marking times} \text{ within } [0,t)|E) \\
&= 1 - \mathbb{P}(\text{one of } \{\{1\}, \cdots, \{k\}\} \text{ does not coalesce at its } \textit{first marking times} \text{ within}[0,t)|E) \\
&\ge 1 - \sum_{i=1}^{k} \mathbb{P}(\{i\} \text{ does not coalesce at its } \textit{first marking time} \text{ within}[0,t)|E) \\
&= 1 - k(1 - \mathbb{P}(\{1\} \text{ coalesces at its } \textit{first marking time} \text{ within}[0,t)|E)) \\
&\ge 1 - k(1 - P_t^{(2,m)}).
\end{aligned}$$

The last inequality is due to the fact that

$$\mathbb{P}(\{1\} \text{ coalesces at its } \textit{first marking time} \text{ within}[0,t)|E)) \ge P_t^{(2,m)},$$

which is true due to the same arguments used in the proof of the last Lemma.

□

If m, t are large enough such that under some assumptions, we could prove that $P_t^{(2,m)}$ is very close to 1. Then the coalescent times are almost the *first marking times* which are easier to deal with. In section 3.3, we will see such a situation for Λ satisfying condition (4.4) and $\Lambda_1 = \Lambda \mathbf{1}_{[0,\frac{1}{n})}, \Lambda_2 = \Lambda \mathbf{1}_{[\frac{1}{n},1]}$. The following corollary studies the *first marking times* in this particular case.

Corollary 4.5. *Let $t > 0$ and $1 \le k \le n$. Assume that Λ satisfies condition* (4.4) *and $\Lambda_1 = \Lambda \mathbf{1}_{[0,\frac{1}{n})}$, $\Lambda_2 = \Lambda \mathbf{1}_{[\frac{1}{n},1]}$. Let Π be a path of $\Pi^{(1,n)}$. Recall that $L_i^{(2,n)}$ is the first marking time of $\{i\}$ for $1 \le i \le n$.*

- *If $\{1\} \in \Pi(\frac{t}{\mu^{(n)}})$, then for any $0 \le t_1 \le t$, $\mathbb{P}(L_1^{(2,n)} \ge \frac{t_1}{\mu^{(n)}}|\Pi) = e^{-t_1}$*
- *Fix k such that $1 \le k \le n$. Assume $\int_0^1 x^{-1}\Lambda(dx) = +\infty$ and $\{i\} \in \Pi(\frac{t}{\mu^{(n)}})$ for any $1 \le i \le k$. Let $0 \le t_1 \le t_2 \le \cdots \le t_k \le t$, we then have*

$$\lim_{n\to+\infty} \mathbb{P}(L_i^{(2,n)} \ge \frac{t_i}{\mu^{(n)}}, \forall 1 \le i \le k|\Pi) = e^{-\sum_{i=1}^k t_i}. \tag{4.10}$$

Proof. The first case is easy to see, due to the definition of $L_1^{(2,n)}$. For the second case, we only consider $k = 2$. For $k > 2$, the proof is similar. Assume that within $[0, \frac{t_1}{\mu^{(n)}}]$, there are N_1 *marking times* and for $(\frac{t_1}{\mu^{(n)}}, \frac{t_2}{\mu^{(n)}}]$, there are N_2 *marking times*. N_1 and N_2 are independently Poisson distributed with parameters respectively $\frac{\bar{\mu}^{(n)}t_1}{\mu^{(n)}}$ and $\frac{\bar{\mu}^{(n)}(t_2-t_1)}{\mu^{(n)}}$ (here we have $\bar{\mu}^{(n)} = \int_{\frac{1}{n}}^1 x^{-2}\Lambda(dx) = \int_{\frac{1}{n}}^1 x^{-2}\Lambda_2(dx)$). Then we get

$$
\begin{aligned}
&\mathbb{P}(L_1^{(2,n)} \geq \frac{t_1}{\mu^{(n)}}, L_2^{(2,n)} \geq \frac{t_2}{\mu^{(n)}}|\Pi) \\
&= \mathbb{E}[\Pi_{i=1}^{N_1}(1-\eta_i^{(2)})^2 \Pi_{i=N_1+1}^{N_1+N_2}(1-\eta_i^{(2)})] \\
&= \mathbb{E}[(1-2\mathbb{E}[\eta_1^{(2)}] + \mathbb{E}[(\eta_1^{(2)})^2])^{N_1}]\mathbb{E}[(1-\mathbb{E}[\eta_1^{(2)}])^{N_2}] \\
&= \exp\left(\frac{\bar{\mu}^{(n)} t_1}{\mu^{(n)}}(-2\mathbb{E}[\eta_1^{(2)}] + \mathbb{E}[(\eta_1^{(2)})^2])\right) \exp\left(\frac{\bar{\mu}^{(n)}(t_2-t_1)}{\mu^{(n)}}(-\mathbb{E}[\eta_1^{(2)}])\right),
\end{aligned}
$$

where the last equality is due to the probability generating function of Poisson distribution. Recall that $\mathbb{E}[\eta_1^{(2)}] = \frac{\mu^{(n)}}{\bar{\mu}^{(n)}}$ and $\mathbb{E}[(\eta_1^{(2)})^2] = \frac{\int_{\frac{1}{n}}^1 \Lambda(dx)}{\bar{\mu}^{(n)}}$. Therefore,

$$
\frac{\bar{\mu}^{(n)}}{\mu^{(n)}}\mathbb{E}[\eta_1^{(2)}] = \frac{\int_{\frac{1}{n}}^1 x^{-1}\Lambda(dx)}{\mu^{(n)}} = 1, \text{and} \quad \frac{\bar{\mu}^{(n)}}{\mu^{(n)}}\mathbb{E}[(\eta_1^{(2)})^2] = \frac{\int_{\frac{1}{n}}^1 \Lambda(dx)}{\mu^{(n)}} \to 0\ .
$$

Then we can conclude (4.10).

□

4.2.4 A tripling

We often have some results on the coalescent related to a special measure, for example, the *Beta*-coalescent. When the process is perturbed by a noise measure, we would wonder whether this damage is negligible. One example is to estimate the number of blocks of the coalescent related to the noise measure. To this aim, we use a tool of tripling.

Tripling: Notice that $\Pi^{(n)}$ encounters its first collision after time $e_1^{(n)}$, which is a random variable. At this collision, the number of blocks is reduced to $n - W_1^{(n)}$, where $W_1^{(n)}$ is a positive integer valued random variable. Then we add $W_1^{(n)}$ new blocks (these blocks can contain any number belonging to $\{n+1, n+2, \cdots\}$) and consider the whole new n ones. By the consistency property, the evolution of the original $n - W_1^{(n)}$ blocks can be embedded into that of the new n blocks, i.e. after time $e_2^{(n)}$, we have the collision in the new n blocks whose total number is reduced to $n - W_2^{(n)}$ and we can calculate the distribution of the number of blocks coalesced among the original $n - W_1^{(n)}$ blocks (we call any block containing at least one of $\{1, 2, \cdots, n\}$ as "original block" and it is very possible that nothing happens for the $n - W_1^{(n)}$ blocks). Then we add again new blocks containing different elements to have another n ones. This procedure is stopped when every element of $\{1, 2, \cdots, n\}$ is contained in one block. By the definition of Λ-coalescent, $(e_i^{(n)})_{i\geq 1}$ are independent exponential random variables with parameter g_n and $(W_i^{(n)})_{i\geq 1}$ are i.i.d copies of $X_1^{(n)}$.

The above procedure gives a tripling of $(e_i^{(n)})_{i\geq 1}$, $(W_i^{(n)})_{i\geq 1}$ and $\Pi^{(n)}$. We define $V_i^{(n)} := \sum_{j=1}^i e_j^{(n)}, i \in \mathbb{N}$. Then we have the following:

Proposition 4.1. *Suppose that* $(e_i^{(n)})_{i\geq 1}$, $(W_i^{(n)})_{i\geq 1}$ *and* $\Pi^{(n)}$ *are tripled, then at any time* $t \geq 0$, *we have*

$$n - \sum_{i=0}^{N(\Lambda,n,t)} W_i^{(n)} \leq |\Pi^{(n)}(t)|, \tag{4.11}$$

where $N(\Lambda, n, t) := card\{i|V_i^{(n)} \leq t\}$, *which is Poisson distributed with parameter* $g_n t$ *and independent of* $(W_i^{(n)})_{i\geq 1}$. *Meanwhile,*

$$\mathbb{E}[W_i^{(n)}] = \frac{n\int_0^1 (1-(1-x)^{n-1})x^{-1}\Lambda(dx)}{g_n} - 1, \textit{ and } \quad \mathbb{E}[(W_i^{(n)})^2] = \frac{n(n-1)\int_0^1 \Lambda(dx)}{g_n} - \mathbb{E}[W_i^{(n)}]. \tag{4.12}$$

Proof. The number of is within $[0, t]$ follows the Poisson distribution with parameter $g_n t$. Due to the tripling, at any time $V_i^{(n)}$ with $0 \leq V_i^{(n)} \leq t$, the decrease of number of blocks (i.e. $|\Pi^{(n)}(V_i^{(n)}-)| - |\Pi^{(n)}(V_i^{(n)})|$) of original blocks is less than or equal to $W_i^{(n)}$. Hence we get (4.11). Notice that $W_i^{(n)} \overset{(d)}{=} X_1^{(n)}$, then (4.12) is a consequence of two equalities in (9) with Eq (17) for the first one and p.1007 for the second one. □

4.3 Applications to coalescents satisfying condition (4.4)

4.3.1 Characterization of condition (4.4).

Some notations for this section: Let Λ be a finite measure on $[0, 1]$ and $\Lambda_1 = \Lambda \mathbf{1}_{[0,\frac{1}{n})}$, $\Lambda_2 = \Lambda \mathbf{1}_{[\frac{1}{n},1]}$; $\mu^{(1/y)} = \int_y^1 x^{-1}\Lambda(dx)$, $g_{1/y} = \int_0^1 (1-(1-x)^{1/y} - \frac{1}{y}x(1-x)^{1/y-1})x^{-2}\Lambda(dx)$ with $0 < y \leq 1$. Notice that the definitions of $\mu^{(1/y)}$ and $g_{1/y}$ are consistent with that of $\mu^{(n)}$ and g_n when $\Lambda(\{0\}) = 0$. For any real number x, let $\lfloor x \rfloor = \max\{y; y \in \mathbb{Z}, y \leq x\}$ and $\lceil x \rceil = \min\{y; y \in \mathbb{Z}, y \geq x\}$.

Here we are going to prove Theorem 4.1, Theorem 4.2, Corollary 4.1, Corollary 4.2 and Corollary 4.3. Under condition (4.4), we decompose Λ into Λ_2 and Λ_1. The idea is to construct $\Pi^{(n)}$ using measure division construction with noise measure Λ_1 and main measure Λ_2. At first, we need to show more details implied by condition (4.4).

Proposition 4.2. *The following two assertions are equivalent:*

(∗)*: Λ satisfies condition* (4.4)*;*

(∗∗)*: $\Lambda(\{0\}) = 0$ and there exists a càglàd (limit from right, continuous from left) function $f : [0, 1] \to [0, 1]$, continuous at 0 with $f(0) = 0$ such that $\int_0^1 \mu^{(1/x)}dx < +\infty$ and*

$$\mu^{(1/y)} = \left(\int_0^1 \mu^{(1/x)}dx\right) \exp\left(\int_y^1 \frac{f(t)}{t}dt\right)(1-f(y)), 0 < y \leq 1. \tag{4.13}$$

Proof. **Part 1**:We first assume that $(*)$ is true. If Λ satisfies (4.4), then $\Lambda(\{0\}) = 0$ due to Remark 4.1. For $\mu^{(n)} \neq 0$, we have

$$\frac{g_n}{n\mu^{(n)}} = \frac{\int_0^1 (1-(1-x)^n - nx(1-x)^{n-1})x^{-2}\Lambda(dx)}{n\mu^{(n)}} = I_1^{(n)} + I_2^{(n)},$$

where $I_1^{(n)} = \frac{\int_{\frac{1}{n}}^1 (1-(1-x)^n - nx(1-x)^{n-1})x^{-2}\Lambda(dx)}{n\mu^{(n)}}$, $I_2^{(n)} = \frac{\int_0^{\frac{1}{n}} (1-(1-x)^n - nx(1-x)^{n-1})x^{-2}\Lambda(dx)}{n\mu^{(n)}}$. Notice that for n large, using monotone property, we have $\frac{e-2}{2e}\frac{\int_{\frac{1}{n}}^1 x^{-2}\Lambda(dx)}{n\mu^{(n)}} \leq I_1^{(n)} \leq \frac{\int_{\frac{1}{n}}^1 x^{-2}\Lambda(dx)}{n\mu^{(n)}}$ and $\frac{1}{3}\frac{n\int_0^{\frac{1}{n}}\Lambda(dx)}{\mu^{(n)}} \leq I_2^{(n)} \leq \frac{n\int_0^{\frac{1}{n}}\Lambda(dx)}{\mu^{(n)}}$. Hence condition (4.4) is equivalent to

$$\lim_{n\to+\infty}\frac{\int_{\frac{1}{n}}^1 x^{-2}\Lambda(dx)}{n\mu^{(n)}} = 0, \text{and} \lim_{n\to+\infty}\frac{n\int_0^{\frac{1}{n}}\Lambda(dx)}{\mu^{(n)}} = 0, \Lambda(\{0\}) = 0. \tag{4.14}$$

Then we deduce that

$$\lim_{y\to 0+}\frac{\int_0^y \Lambda(dx)}{y\mu^{(1/y)}} = 0, \Lambda(\{0\}) = 0. \tag{4.15}$$

Indeed, for $1/y > 2$ and $\mu^{(\lfloor 1/y\rfloor)} \neq 0$, we have

$$\frac{\int_0^y \Lambda(dx)}{y\mu^{(1/y)}} = \frac{\int_0^y \Lambda(dx)}{y\int_y^1 x^{-1}\Lambda(dx)} \leq \frac{\int_0^{1/\lfloor 1/y\rfloor}\Lambda(dx)}{\frac{1}{\lceil 1/y\rceil}\int_{1/\lfloor 1/y\rfloor}^1 x^{-1}\Lambda(dx)} = \frac{\lceil 1/y\rceil}{\lfloor 1/y\rfloor}\frac{\lfloor 1/y\rfloor\int_0^{1/\lfloor 1/y\rfloor}\Lambda(dx)}{\int_{1/\lfloor 1/y\rfloor}^1 x^{-1}\Lambda(dx)} \overset{y\to 0+}{\to} 0.$$

One thing to notice is that $\lim_{y\to 0+} y\mu^{(1/y)} = 0$ is true for any finite Λ. In fact, for any positive number M and $yM < 1$, we have

$$y\mu^{(1/y)} = y\int_y^1 x^{-1}\Lambda(dx) = y\int_{yM}^1 x^{-1}\Lambda(dx) + y\int_y^{yM} x^{-1}\Lambda(dx) \leq \frac{\int_0^1\Lambda(dx)}{M} + \int_y^{yM}\Lambda(dx),$$

where both terms can be made as small as we want by taking M large enough and y close enough to 0. Looking into details of $\frac{\int_0^y \Lambda(dx)}{y\mu^{(1/y)}}$ when $\mu^{(1/y)} \neq 0$, we have the following equality, using integration by parts and $\lim_{y\to 0+} y\mu^{(1/y)} = 0$,

$$\frac{\int_0^y \Lambda(dx)}{y\mu^{(1/y)}} = \frac{\int_0^y xx^{-1}\Lambda(dx)}{y\mu^{(1/y)}} = \frac{\int_0^y \mu^{(1/x)}dx - y\mu^{(1/y+)}}{y\mu^{(1/y)}}, \tag{4.16}$$

where $\mu^{(1/y+)} = \mu^{(1/y)} - y^{-1}\Lambda(\{y\})$. Due to (4.15), we get that $1 \geq \frac{\mu^{(1/y+)}}{\mu^{(1/y)}} = 1 - \frac{\Lambda(\{y\})}{y\mu^{(1/y)}} \geq 1 - \frac{\int_0^y \Lambda(dx)}{y\mu^{(1/y)}} \to 1$. Therefore, (4.15) and (4.16) give

$$\lim_{y\to 0+}\frac{y\mu^{(1/y)}}{\int_0^y \mu^{(1/x)}dx} = 1.$$

Notice that $\int_0^y \mu^{(1/x)}dx \geq y\mu^{(1/y)}$ and $\mu^{(1/y)}$ is a càglàd function. Hence there exists a càglàd function $f : [0,1] \to [0,1]$, continuous at 0 with $f(0) = 0$ such that

$$\frac{y\mu^{(1/y)}}{\int_0^y \mu^{(1/x)}dx} = 1 - f(y). \tag{4.17}$$

Now let $G(t) = \int_0^t \mu^{(1/x)}dx$ and any derivative is considered as left derivative. Then (4.17) becomes

$$(\ln G(t))' = \frac{G(t)'}{G(t)} = \frac{1-f(t)}{t}.$$

Using the fundamental theorem of Newton and Leibniz which also works for càglàd functions whose primitive functions take left derivatives, one gets that for $0 < y \leq 1$,

$$\ln G(1) - \ln G(y) = \int_y^1 (\ln G(t))'dt = \int_y^1 \frac{1-f(t)}{t}dt.$$

Therefore,

$$G(y) = G(1)\exp\Big(-\int_y^1 \frac{1-f(t)}{t}dt\Big).$$

By taking the left derivatives on the both sides and noticing that $G(1) = \int_0^1 \mu^{(1/x)}dx$, one can conclude.

Part 2: We now assume that $(**)$ is true. In the first part, we proved implicitly that (4.15) is equivalent to the $(**)$. Hence we will use (4.15) to prove (4.14) which is equivalent to condition (4.4) and only the first convergence in (4.14) is needed to be proved. Let M be a positive number and $\frac{M}{n} \leq 1$, $\mu^{(n)} \neq 0$, then

$$\begin{aligned}\frac{\int_{\frac{1}{n}}^1 x^{-2}\Lambda(dx)}{n\mu^{(n)}} &= \frac{\int_{M/n}^1 x^{-2}\Lambda(dx)}{n\mu^{(n)}} + \frac{\int_{\frac{1}{n}}^{M/n} x^{-2}\Lambda(dx)}{n\mu^{(n)}} \\ &\leq \frac{1}{M} + 1 - \frac{\mu^{(n/M)}}{\mu^{(n)}}.\end{aligned}$$

The first term can be made as small as we want by taking M large, and the third term $\frac{\mu^{(n/M)}}{\mu^{(n)}} = \exp(-\int_{\frac{1}{n}}^{M/n} \frac{f(x)}{x}ds)\frac{1-f(M/n)}{1-f(\frac{1}{n})}$. Let $\epsilon > 0$ and n large enough such that $f(x) \leq \epsilon$ on $[0, M/n]$. Then $\frac{\mu^{(n/M)}}{\mu^{(n)}} \geq \exp(-\epsilon \ln M)(1-\epsilon)$, which can be made as close as possible to 1 with ϵ small enough. Hence we can conclude. □

The next corollary is immediate.

Corollary 4.6. *If* Λ *satisfies* (4.4), *then*

— $\lim_{n\to+\infty} \frac{(\mu^{(n)})^k}{n} = 0, \forall k > 0;$

— $\lim_{n\to+\infty} \frac{\mu^{(n)}}{\mu^{(n-M)}} = 1, \forall M > 0;$

— $\lim_{n\to+\infty} \frac{\mu^{(n)}}{\mu^{(n\epsilon)}} = 1, \forall 0 < \epsilon < 1.$

4.3.2 Properties of $\Pi^{(1,n)}$.

We should next estimate the coalescent process related to the noise measure Λ_1 which serves as a perturbation to the main measure Λ_2. At first, one needs a technical result.

Lemma 4.4. *We assume that $\Lambda(\{0\}) = 0$. Let $g_n^{(1)} = \int_0^1 (1-(1-x)^n - nx(1-x)^{n-1})x^{-2}\Lambda_1(dx)$ in the spirit of (4.3). Then there exists a positive constant C_1 such that for n large enough*

$$g_n^{(1)} \geq C_1 n^2 \int_0^{\frac{1}{n}} \Lambda_1(dx). \tag{4.18}$$

Proof. Let $M > 2$. We write

$$\begin{aligned} g_n^{(1)} &= \int_0^1 (1-(1-x)^n - nx(1-x)^{n-1})x^{-2}\Lambda_1(dx) \\ &= \int_0^{\frac{1}{n}} (1-(1-x)^n - nx(1-x)^{n-1})x^{-2}\Lambda_1(dx) \\ &= I_1 + I_2, \end{aligned}$$

where $I_1 = \int_0^{\frac{1}{nM}} (1-(1-x)^n - nx(1-x)^{n-1})x^{-2}\Lambda_1(dx)$ and $I_2 = \int_{\frac{1}{nM}}^{\frac{1}{n}} (1-(1-x)^n - nx(1-x)^{n-1})x^{-2}\Lambda_1(dx)$. It is easy to see that for $n \geq 2$,

$$\begin{aligned} I_1 &\geq \int_0^{\frac{1}{nM}} (n(n-1) - n(n-1)(n-2)x)\frac{1}{2}\Lambda_1(dx) \\ &\geq \int_0^{\frac{1}{nM}} (n(n-1) - (n-1)(n-2)/M)\frac{1}{2}\Lambda_1(dx) \\ &\geq \frac{1}{4}\int_0^{\frac{1}{nM}} n^2\Lambda_1(dx). \end{aligned}$$

For the second term,

$$I_2 \geq \int_{\frac{1}{nM}}^{\frac{1}{n}} (1-(1-\frac{1}{nM})^n - \frac{(1-\frac{1}{nM})^{n-1}}{M})n^2\Lambda_1(dx).$$

Notice that for n large, there exists a positive constant $C(M)$ such that

$$1-(1-\frac{1}{nM})^n-\frac{(1-\frac{1}{nM})^{n-1}}{M}\geq C(M)>0.$$

Hence $I_2\geq C(M)\int_{\frac{1}{nM}}^{\frac{1}{n}}n^2\Lambda_1(dx)$. It suffices to take $C_1=\min\{\frac{1}{4},C(M)\}$ to conclude. □

The following lemma estimates the coalescent process related to the noise measure Λ_1. Recall that $\Pi^{(1,n)}$ is the Λ_1-coalescent process with $\Pi^{(1,n)}(0)=\{\{1\},\{2\},\cdots,\{n\}\}$.

Lemma 4.5. *Assume that Λ satisfy (4.4). Then for any $M>0$, $0<\epsilon\leq 1$ and n large enough, we have*

$$\mathbb{P}\left(|\Pi^{(1,n)}(\frac{M}{\mu^{(n)}})|\leq n-n\epsilon\right)=o(n^{-1}). \tag{4.19}$$

Proof. If $\int_0^{\frac{1}{n_0}}\Lambda(dx)=0$ with some $n_0>1$, then for any $n>n_0$, Λ_1 is the null measure and hence $|\Pi^{(1,n)}(t)|=n$ for any $t\geq 0$, which proves this lemma. In consequence, one needs only to consider the case where $\int_0^{\frac{1}{n}}\Lambda(dx)\neq 0$ for any $n\geq 1$. We recall $g_n^{(1)}$ defined in Lemma 4.4. Let $X_1^{(1,n)}$ be the decrease of the number of blocks at the first coalescence of $\Pi^{(1,n)}$. Thanks to Proposition 4.1 where we pick up the notations,

$$n-\sum_{i=1}^{N(\Lambda_1,n,\frac{M}{\mu^{(n)}})}W_i^{(n)}\leq|\Pi^{(1,n)}(\frac{M}{\mu^{(n)}})|,$$

where $N(\Lambda_1,n,\frac{M}{\mu^{(n)}})$ is Poisson distributed with parameter $\frac{Mg_n^{(1)}}{\mu^{(n)}}$ independent of $(W_i^{(n)})_{i\geq 1}$ which are i.i.d copies of $X_1^{(1,n)}$. Then we have, for n large,

$$\begin{aligned}
\mathbb{P}(|\Pi^{(1,n)}(\frac{M}{\mu^{(n)}})|\leq n-n\epsilon)&\leq\mathbb{P}\left(n-\sum_{i=1}^{N(\Lambda_1,n,\frac{M}{\mu^{(n)}})}W_i^{(n)}\leq n-n\epsilon\right)\\
&=\mathbb{P}\left(\sum_{i=1}^{N(\Lambda_1,n,\frac{M}{\mu^{(n)}})}W_i^{(n)}-\frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}]\geq n\epsilon-\frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}]\right)\\
&\leq\frac{\mathrm{Var}(\sum_{i=1}^{N(\Lambda_1,n,\frac{M}{\mu^{(n)}})}W_i^{(n)})}{(n\epsilon-\frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}])^2}=\frac{\frac{Mg_n^{(1)}}{\mu^{(n)}}\mathbb{E}[(W_1^{(n)})^2]}{(n\epsilon-\frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}])^2},
\end{aligned}\tag{4.20}$$

where the second inequality needs $n\epsilon-\frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}]>0$ which is justified by the following calculations: Notice that due to Proposition 4.1 and Lemma 4.4, for n large enough,

$$\mathbb{E}[W_1^{(n)}]+1\leq\frac{n(n-1)\int_0^{\frac{1}{n}}\Lambda_1(dx)}{g_n^{(1)}}\leq\frac{1}{C_1};\mathbb{E}[(W_1^{(n)})^2]\leq\frac{n(n-1)\int_0^{\frac{1}{n}}\Lambda_1(dx)}{g_n^{(1)}}\leq\frac{1}{C_1}, \tag{4.21}$$

where C_1 is the positive constant in Lemma 4.4.

Notice that (4.4) gives $\frac{g_n^{(1)}}{n\mu^{(n)}} \leq \frac{g_n}{n\mu^{(n)}} \to 0$. Then together with (4.21), we have

$$\frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}] = o(n), \frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[(W_1^{(n)})^2] = o(n).$$

Hence $n\epsilon - \frac{g_n^{(1)}M}{\mu^{(n)}}\mathbb{E}[W_1^{(n)}] \asymp n\epsilon$. So the inequality (4.20) is justified and one deduces that

$$\mathbb{P}(|\Pi^{(\Lambda_1,n)}(\frac{M}{\mu^{(n)}})| \leq n - n\epsilon) = o(n^{-1}).$$

Then we conclude (4.19). □

4.3.3 Asymptotics of $P_t^{(2,m)}, P_{1,2}^{(n,m)}(t), P_{1,2}^{(n,m,k)}(t), 2 \leq m \leq n, t \geq 0$.

These terms are probabilities defined in section 2.3.1, which measure the possibility to make one or several singletons coalesced in their *first marking times* within $[0, t)$. In fact, we will study $P_{\frac{t}{\mu^{(n)}}}^{(2,m)}, P_{1,2}^{(n,m)}(\frac{t}{\mu^{(n)}}), P_{1,2}^{(n,m,k)}(\frac{t}{\mu^{(n)}})$, since we want to prove that the normalization factor of the external branch length is $\mu^{(n)}$. We denote by "$\ll$" the stochastic dominance between two real random variables. The following corollary together with the remark at the end play an important role in getting the asymptotics of the three probabilities.

Proposition 4.3. *Suppose that* Λ *satisfies* (4.4) *and define*

$$P^{(2,n)} := \lim_{t\to\infty} P_t^{(2,n)} = \sum_{i=1}^{+\infty} \mathbb{E}[\Delta_i^{(2)}\left(1 - (1 - \Delta_i^{(2)})^{n-1}\right)].$$

Then

$$\lim_{n\to+\infty} P^{(2,n)} = 1. \tag{4.22}$$

Proof. Recall $(\eta_i^{(2)})_{i\geq1}$, $(e_i^{(2)})_{i\geq1}$, $\{\Delta_i^{(2)}\}_{i\geq1}$ which are associated to Λ_2 and defined in section 2.3. At first, we remark that $\sum_{i=1}^{+\infty}\mathbb{E}[\Delta_i^{(2)}] = 1$. One only needs to prove that $\lim_{n\to+\infty}\sum_{i=1}^{+\infty}\mathbb{E}[\Delta_i^{(2)}(1-\Delta_i^{(2)})^{n-1}] = 0$. It is easy to see that $\mathbb{E}[\Delta_i^{(2)}(1-\Delta_i^{(2)})^{n-1}] = \mathbb{E}[\bar{\Delta}_i^{(2)}(1-\bar{\Delta}_i^{(2)})^{n-1}]$, where $\bar{\Delta}_i^{(2)} = \eta_1^{(2)}\Pi_{j=2}^{i}(1-\eta_j^{(2)})$. It is obvious that $(\bar{\Delta}_i^{(2)})_{i\geq1}$ is a Markov chain.

For $s > 0$, we define a stopping time

$$\begin{aligned}\tau_s &= \min\{i|\bar{\Delta}_i^{(2)} \le 1/s\} \\ &= \min\{i| - \sum_{j=2}^{i} \ln(1-\eta_j^{(2)}) \ge \ln s\eta_1^{(2)}\} \\ &= \min\{i+1| - \sum_{j=1}^{i} \ln(1-\eta_{j+1}^{(2)}) \ge \ln s\eta_1^{(2)}\}.\end{aligned}$$

Then we get

$$\begin{aligned}\sum_{i=1}^{+\infty} \mathbb{E}[\Delta_i^{(2)}(1-\Delta_i^{(2)})^{n-1}] &= \mathbb{E}[\sum_{i=1}^{+\infty} \bar{\Delta}_i^{(2)}(1-\bar{\Delta}_i^{(2)})^{n-1}] \\ &= \mathbb{E}[\sum_{i=1}^{\tau_n-1} \bar{\Delta}_i^{(2)}(1-\bar{\Delta}_i^{(2)})^{n-1} + \sum_{i=\tau_n}^{+\infty} \bar{\Delta}_i^{(2)}(1-\bar{\Delta}_i^{(2)})^{n-1}]. \quad (4.23)\end{aligned}$$

Notice that $x(1-x)^{n-1} \le \frac{1}{n}$, if $\frac{1}{n} \le x \le 1$ and $x(1-x)^{n-1} \le x$, if $0 \le x \le \frac{1}{n}$. Then (4.23) gives

$$\sum_{i=1}^{+\infty} \mathbb{E}[\Delta_i^{(2)}(1-\Delta_i^{(2)})^{n-1}] \le \mathbb{E}[\frac{\tau_n - 1}{n} + \sum_{i=\tau_n}^{+\infty} \bar{\Delta}_i^{(2)}] \le \mathbb{E}[\frac{\tau_n - 1}{n}] + \frac{1}{\mathbb{E}[n\eta_1^{(2)}]}. \quad (4.24)$$

To calculate $\mathbb{E}[\tau_n]$, we use renewal theory. Let $\mu = \mathbb{E}[-\ln(1-\eta_1^{(2)})]$. Depending on whether μ is finite or not, we separate the discussion into two parts.

Part 1: Assume that $\mu < +\infty$. We denote by $F(t)$ the distribution function and $f(t)$ the density function of $-\ln(1-\eta_1^{(2)})$ and X an independent random variable with density function $\frac{1}{\mu}(1-F(t))\mathbf{1}_{t\ge 0}$. Notice that X depends on n. Let $\epsilon > 0$, then using integration by parts,

$$\mathbb{P}(0 \le X \le \epsilon) = \int_0^{\epsilon} \frac{1-F(t)}{\mu} dt = \frac{\epsilon(1-F(\epsilon))}{\mu} + \frac{\int_0^{\epsilon} tf(t)dt}{\mu} \ge \frac{\int_0^{\epsilon} tf(t)dt}{\mu}. \quad (4.25)$$

One can write $\int_0^{\epsilon} tf(t)dt$ in another way

$$\int_0^{\epsilon} tf(t)dt = \frac{\int_{\frac{1}{n}}^{1-e^{-\epsilon}} -\ln(1-x)x^{-2}\Lambda(dx)}{\int_{\frac{1}{n}}^{1} x^{-2}\Lambda(dx)}.$$

Notice that $\mu = \frac{\int_{\frac{1}{n}}^1 -\ln(1-x)x^{-2}\Lambda(dx)}{\int_{\frac{1}{n}}^1 x^{-2}\Lambda(dx)} < +\infty$, then for n large enough, there must exist a large number $\epsilon_0 > 0$ such that for any $\epsilon \geq \epsilon_0$,

$$\int_0^\epsilon tf(t)dt \geq \frac{1}{2}\frac{\int_{\frac{1}{n}}^1 -\ln(1-x)x^{-2}\Lambda(dx)}{\int_{\frac{1}{n}}^1 x^{-2}\Lambda(dx)} = \frac{\mu}{2}.$$

Now together with (4.25), one gets that for n large enough

$$\mathbb{P}(0 \leq X \leq \epsilon) \geq 1/2, \quad \forall \epsilon \geq \epsilon_0. \tag{4.26}$$

We fix $\epsilon \geq \epsilon_0$ and define a new Markov chain $(X - \sum_{j=2}^i \ln(1-\eta_j^{(2)}))_{i\geq 1}$ and a stopping time $\tau'_s = \min\{i | X - \sum_{j=1}^i \ln(1-\eta_{j+1}^{(2)}) \geq \ln s\}$ for $s > 0$. It is clear from the definitions of τ_s and τ'_s that

$$\mathbb{E}[\tau'_{s\eta_1^{(2)}} | X = \epsilon] = \mathbb{E}[\tau_{se^{-\epsilon}} - 1].$$

Then

$$\begin{aligned}\mathbb{E}[\tau'_{n\eta_1^{(2)}}] &= \mathbb{E}[\tau'_{n\eta_1^{(2)}}\mathbf{1}_{0\leq X\leq\epsilon}] + \mathbb{E}[\tau'_{n\eta_1^{(2)}}\mathbf{1}_{X>\epsilon}] \\ &\geq \mathbb{P}(0 \leq X \leq \epsilon)\mathbb{E}[\tau_{ne^{-\epsilon}} - 1] + \mathbb{E}[\tau'_{n\eta_1^{(2)}}\mathbf{1}_{X>\epsilon}],\end{aligned}$$

which implies that

$$\mathbb{E}[\tau_{n\exp(-\epsilon)}] \leq \frac{\mathbb{E}[\tau'_{n\eta_1^{(2)}}]}{\mathbb{P}(0 \leq X \leq \epsilon)} + 1. \tag{4.27}$$

Due to (4.4) and (4.6) in [(15), p.369], we have

$$\mathbb{E}[\tau'_s] = \frac{\ln s}{\mu}, \forall s \geq 1.$$

Notice that $\eta_1^{(2)} \geq \frac{1}{n}$, hence $n\eta_1^{(2)} \geq 1$. Therefore, (4.27) gives

$$\mathbb{E}[\tau_n] \leq \frac{\mathbb{E}[\tau'_{ne^\epsilon\eta_1^{(2)}}]}{\mathbb{P}(0 \leq X \leq \epsilon)} + 1 = \frac{\mathbb{E}[\ln(ne^\epsilon\eta^{(2)})]}{\mu\mathbb{P}(0 \leq X \leq \epsilon)} + 1. \tag{4.28}$$

For any $0 \leq x < 1$, we have $-\ln(1-x) \geq x$, hence $\mu \geq \mathbb{E}[\eta_1^{(2)}]$. Then (4.28) implies

$$\frac{\mathbb{E}[\tau_n]}{n} \leq \frac{\mathbb{E}[\ln n\eta_1^{(2)}] + \epsilon}{\mathbb{E}[n\eta_1^{(2)}]\mathbb{P}(0 \leq X \leq \epsilon)} + \frac{1}{n}. \tag{4.29}$$

Using (4.26) and (4.24), it suffices to prove that:

$$\lim_{n\to+\infty}\mathbb{E}[n\eta_1^{(2)}]=+\infty,\text{and}\quad \lim_{n\to+\infty}\frac{\mathbb{E}[\ln(n\eta_1^{(2)})]}{\mathbb{E}[n\eta_1^{(2)}]}=0.$$

It is easy to see that, using (4.3), there exists a positive constant C_2 such that $\mathbb{E}[n\eta_1^{(2)}]=\frac{n\int_{\frac{1}{n}}^1 x^{-1}\Lambda(dx)}{\bar{\mu}^{(n)}}\geq C_2\frac{n\mu^{(n)}}{g_n}$, for any $n\geq 3$. Hence $\mathbb{E}[n\eta_1^{(2)}]$ tends to $+\infty$ since Λ satisfies (4.4). For the second convergence, we fix $M>e$. Then,

$$\begin{aligned}\frac{\mathbb{E}[\ln(n\eta_1^{(2)})]}{\mathbb{E}[n\eta_1^{(2)}]}&=\frac{\mathbb{E}[\ln(n\eta_1^{(2)})\mathbf{1}_{n\eta_1^{(2)}\geq M}]+\mathbb{E}[\ln(n\eta_1^{(2)})\mathbf{1}_{n\eta_1^{(2)}<M}]}{\mathbb{E}[n\eta_1^{(2)}]}\\&\leq\frac{\mathbb{E}[\ln(n\eta_1^{(2)})\mathbf{1}_{n\eta_1^{(2)}\geq M}]}{\mathbb{E}[n\eta_1^{(2)}]}+\frac{\ln M}{\mathbb{E}[n\eta_1^{(2)}]}\\&\leq\frac{\mathbb{E}[\ln(n\eta_1^{(2)})\mathbf{1}_{n\eta_1^{(2)}\geq M}]}{\mathbb{E}[n\eta_1^{(2)}\mathbf{1}_{n\eta_1^{(2)}\geq M}]}+\frac{\ln M}{\mathbb{E}[n\eta_1^{(2)}]}\\&\leq\frac{\ln M}{M}+\frac{\ln M}{\mathbb{E}[n\eta_1^{(2)}]}.\end{aligned}$$

The last inequality is due to the fact that for any $x\geq M>e$, we have $\frac{\ln x}{x}\leq\frac{\ln M}{M}$. Since M can be chosen as large as we want, $\lim_{n\to+\infty}\frac{\mathbb{E}[\ln(n\eta_1^{(2)})]}{\mathbb{E}[n\eta_1^{(2)}]}=0$. Hence we can conclude.

Part 2: If $\mu=+\infty$. We define $(\bar{\eta}_i^{(2)})_{i\geq 2}:=(\frac{1}{2}\mathbf{1}_{\eta_i^{(2)}\geq\frac{1}{2}}+\eta_i^{(2)}\mathbf{1}_{\eta_i^{(2)}<\frac{1}{2}})_{i\geq 2}$ and for $s>0$, $\bar{\tau}_s:=\min\{i+1|\sum_{j=1}^{i}-\ln(1-\bar{\eta}_{j+1}^{(2)})\geq\ln s\eta_1^{(2)}\}$. Notice that $\mathbb{E}[-\ln(1-\bar{\eta}_i^{(2)})]<+\infty$, then we return to the first case and get (4.29) by replacing τ_n by $\bar{\tau}_n$ and keeping the same $\eta_1^{(2)}$ but with different X (depending on $\bar{\eta}_i^{(2)},i\geq 2$). In this setting, $\mathbb{P}(0\leq X\leq\ln 2)=1$. We see that the closer $\bar{\eta}_i^{(2)}$ is to 1, the larger the $-\ln(1-\bar{\eta}_i^{(2)})$ and hence $\tau_n\ll\bar{\tau}_n$. Then we can conclude. □

Remark 4.8. *For $0<\epsilon<1$, we also have*

$$\lim_{n\to+\infty}\sum_{i=1}^{+\infty}\mathbb{E}[\Delta_i^{(2)}(1-\Delta_i^{(2)})^{n(1-\epsilon)}]=0. \tag{4.30}$$

The proof is all the same. The only thing different is that in place of (4.24), we have $\sum_{i=1}^{+\infty}\mathbb{E}[\Delta_i^{(2)}(1-\Delta_i^{(2)})^{n(1-\epsilon)}]\leq C\mathbb{E}[\frac{\tau_n-1}{n}+\sum_{i=\tau_n}^{+\infty}\bar{\Delta}_i^{(2)}]$, with C larger than 1 and depends on ϵ.

Now we can start to study at first $P^{(2,n)}_{\frac{t}{\mu^{(n)}}}$.

Corollary 4.7.

$$\lim_{t\to+\infty}\liminf_{n\to+\infty} P^{(2,n)}_{\frac{t}{\mu^{(n)}}} = 1. \tag{4.31}$$

Proof. Recall that $\{e_i^{(2)}\}_{i\geq n}$ are i.i.d exponential variables with parameter $\int_0^1 x^{-2}\Lambda_2(dx) = \bar{\mu}^{(n)}$, as defined in section 2.3. Let $\tau_n(t) = \max\{j : \sum_{i=1}^j e_i^{(2)} \leq \frac{t}{\mu^{(n)}}\}$. Then

$$P^{(2,n)}_{\frac{t}{\mu^{(n)}}} = \mathbb{E}[\sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)} - \sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)}(1-\Delta_i^{(2)})^{n-1}]. \tag{4.32}$$

Due to Proposition 4.3, we have

$$\lim_{n\to+\infty} \mathbb{E}[\sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)}(1-\Delta_i^{(2)})^{n-1}] \leq \lim_{n\to+\infty} \mathbb{E}[\sum_{i=1}^{+\infty} \Delta_i^{(2)}(1-\Delta_i^{(2)})^{n-1}] = 0.$$

Then it suffices to prove that

$$\lim_{t\to+\infty}\liminf_{n\to+\infty} \mathbb{E}[\sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)}] = 1. \tag{4.33}$$

Let $\mathcal{E}_j = \bar{\mu}^{(n)}\sum_{i=1}^j e_i^{(2)}$, which is the sum of j i.i.d unit exponential variables. Let $I_n = \bar{\mu}^{(n)}/\mu^{(n)}$. Then

$$\tau_n(t) = \max\{j : \mathcal{E}_j \leq tI_n\}.$$

For any fixed $0 < \beta < 1$,

$$\begin{aligned} &\mathbb{P}\Big(\tau_n(t) \in [0, \beta t I_n)\bigcup(tI_n/\beta, +\infty)\Big) \\ &= \mathbb{P}(\mathcal{E}_{\lceil \beta t I_n\rceil} \geq tI_n) + \mathbb{P}(\mathcal{E}_{\lfloor tI_n/\beta\rfloor} \leq tI_n) = o((tI_n)^{-1}), \end{aligned} \tag{4.34}$$

where the last equality is a large deviation result (for example, see Theorem 1.4 of (10)). Notice that $\tau_n(t)$ is independent of $\{\Delta_i^{(2)}\}_{i\geq 1}$, so

$$\begin{aligned} \mathbb{E}[\sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)}] &= \mathbb{E}[1-(1-1/I_n)^{\tau_n(t)+1}] \\ &= \mathbb{E}[1-(1-1/I_n)^{\tau_n(t)+1}\mathbf{1}_{tI_n\beta\leq\tau_n(t)\leq tI_n/\beta}] + o((tI_n)^{-1}) \\ &\geq \mathbb{E}[1-(1-1/I_n)^{tI_n\beta}\mathbf{1}_{tI_n\beta\leq\tau_n(t)\leq tI_n/\beta}] + o((tI_n)^{-1}). \end{aligned}$$

Notice that $I_n \geq 1$ and the term at the right of the above inequality satisfies

$$\lim_{t\to+\infty} \liminf_{n\to+\infty} \mathbb{E}[1-(1-1/I_n)^{tI_n\beta}\mathbf{1}_{tI_n\beta\leq\tau_n(t)\leq tI_n/\beta}] + o((tI_n)^{-1}) = 1$$

Then we can conclude (4.33).

□

Remark 4.9. *For $0<\epsilon<1$, we also have*

$$\lim_{t\to+\infty} \liminf_{n\to+\infty} P^{(2,\lceil n-n\epsilon\rceil)}_{\frac{t}{\mu^{(n)}}} = 1. \tag{4.35}$$

To prove this, in the proof of this corollary, one should replace (4.32) by

$$P^{(2,n)}_{\frac{t}{\mu^{(n)}}} = \mathbb{E}[\sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)} - \sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)}(1-\Delta_i^{(2)})^{\lceil n-n\epsilon\rceil -1}].$$

The first term satisfies (4.33). For the second term, using (4.30), we get

$$\lim_{n\to+\infty} \mathbb{E}[\sum_{i=1}^{\tau_n(t)} \Delta_i^{(2)}(1-\Delta_i^{(2)})^{\lceil n-n\epsilon\rceil-1}] = 0.$$

Then (4.35) is proved.

The next corollary is straightforward using (4.8), (4.9) and (4.35).

Corollary 4.8. *For any $0<\epsilon<1$,*

$$\lim_{t\to+\infty} \liminf_{n\to+\infty} P^{(n,\lceil n-n\epsilon\rceil,k)}_{1,2}(\frac{t}{\mu^{(n)}}) = 1, \lim_{t\to+\infty} \liminf_{n\to+\infty} P^{(n,\lceil n-n\epsilon\rceil)}_{1,2}(\frac{t}{\mu^{(n)}}) = 1,$$

4.3.4 Proofs of main results.

Proof of Theorem 4.1

Proof. Fix $t>0$ and $0<\epsilon<1$. Considering the measure division construction for two-type Λ-coalescents, let Π be the path of $\Pi^{(1,n)}$ chosen at the step 0 and define the event

$$E' := \{|\Pi(\frac{t}{\mu^{(n)}})| \geq n-n\epsilon\} \bigcap \{\{1\}\in\Pi(\frac{t}{\mu^{(n)}})\}.$$

Recall that $\{|\Pi^{(1,n)}(\frac{t}{\mu^{(n)}})| \geq n-n\epsilon\}$ implies that there are at least $n-\lceil 2n\epsilon\rceil$ singletons at time $\frac{t}{\mu^{(n)}}$. For n large enough, using the exchangeability property, we have $\mathbb{P}(E') \geq \frac{n-\lceil 2n\epsilon\rceil}{n}(1-\kappa_n(t))$, where $\kappa_n(t) = \mathbb{P}(|\Pi^{(1,n)}(\frac{t}{\mu^{(n)}})| < n-n\epsilon)$ and $\kappa_n(t) = o(n^{-1})$ due to

(4.19) . For ϵ small enough and n large enough, we have $\mathbb{P}(E')$ as close as we want to 1. We define another event

$$E'' := \{\{1\} \text{ is coalesced at its } \textit{first marking time} \text{ within } [0,t).\}$$

Then due to (4.8) and $P_t^{(2,m)}$ is increasing on m, we get

$$\mathbb{P}(E''|E') \geq P^{(2,\lceil n-n\epsilon \rceil)}_{\frac{t}{\mu^{(n)}}}. \tag{4.36}$$

Let $0 < t_1 < t$,

$$\begin{aligned} \mathbb{P}(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}) &= \mathbb{P}(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, E' \bigcap E'') + \mathbb{P}(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, (E' \bigcap E'')^c) \\ &= \mathbb{P}(L_1^{(2,n)} \geq \frac{t_1}{\mu^{(n)}}, E' \bigcap E'') + \mathbb{P}(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, (E' \bigcap E'')^c). \end{aligned} \tag{4.37}$$

Corollary 4.5 tells that $\mathbb{P}(L_1^{(2,n)} \geq \frac{t_1}{\mu^{(n)}}|E') = \exp(-t_1)$ and it has been proved that $\mathbb{P}(E' \cap E'') = \mathbb{P}(E')\mathbb{P}(E''|E')$ can be made as close as possible to 1 by taking ϵ small enough and t large enough and n tending to $+\infty$. Hence the first term of (4.37) can be made as close as we want to $\exp(-t_1)$ and the second term is close to 0. Then we can conclude.

□

Proof of Theorem 4.2

Proof. We prove instead for $k \in \mathbb{N}$:

$$\mu^{(n)}(T_1^{(n)}, T_2^{(n)}, \cdots, T_k^{(n)}) \xrightarrow{(d)} (e_1, e_2, \cdots, e_k), \tag{4.38}$$

which is equivalent to (4.6) (see [(3), p.19]). We will give the proof for $k = 2$ and leave the easy extension to readers. The proof is similar to that of Theorem 4.1. Let Π be the path of $\Pi^{(1,n)}$ chosen at step 0. Let $t > 0, 0 < \epsilon < 1$ and define the event

$$F' := \{|\Pi(\frac{t}{\mu^{(n)}})| \geq n - n\epsilon\} \bigcap \{\{1\},\{2\} \in \Pi(\frac{t}{\mu^{(n)}})\}.$$

Using the same arguments, we get $\mathbb{P}(F') \geq \frac{\binom{n-\lceil 2n\epsilon \rceil}{2}}{\binom{n}{2}}(1-\kappa_n(t))$. We then define the event

$$F'' := \{\{1\},\{2\} \text{ are both coalesced at their first } \textit{marking times} \text{ within } [0,t).\}$$

Then due to (4.9) and $P_t^{(2,m)}$ is increasing on m, we get

$$P(F''|F') \geq 1 - 2(1 - P^{(2,\lceil n-n\epsilon \rceil)}_{\frac{t}{\mu^{(n)}}}),$$

which is as close as possible to 1 for t large and n tending to $+\infty$.

Let $0 \leq t_1, t_2 \leq t$. Then

$$\begin{aligned}
&\mathbb{P}(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, T_2^{(n)} \geq \frac{t_2}{\mu^{(n)}}) \\
&= \mathbb{P}(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, T_2^{(n)} \geq \frac{t_2}{\mu^{(n)}}, F' \bigcap F'') + \mathbb{P}(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, T_2^{(n)} \geq \frac{t_2}{\mu^{(n)}}, (F' \bigcap F'')^c) \\
&= \mathbb{P}(L_1^{(2,n)} \geq \frac{t_1}{\mu^{(n)}}, L_2^{(2,n)} \geq \frac{t_2}{\mu^{(n)}}, F' \bigcap F'') + \mathbb{P}(T_1^{(n)} \geq \frac{t_1}{\mu^{(n)}}, T_2^{(n)} \geq \frac{t_2}{\mu^{(n)}}, (F' \bigcap F'')^c).
\end{aligned} \tag{4.39}$$

As shown that $\mathbb{P}((F' \cap F''))$ can be made as close as possible to 1 by taking t large enough and ϵ small enough, tending n to $+\infty$. Then the second term in (4.39) is close to 0. Using Corollary 4.5, the first term is as close as possible to $e^{-t_1-t_2}$ by tending n to $+\infty$ with t large enough. Then we can conclude. □

Proof of Corollary 4.1

Proof. We prove at first the case of one external branch length. We seek to prove the uniform integrability of $\{(\mu^{(n)}T_1^{(n)})^k, n \geq 2\}$ for any $k \geq 0$. One needs only to show that for any fixed $k \in \mathbb{N}$, $\sup\{\mathbb{E}[(\mu^{(n)}T_1^{(n)})^k] | n \geq 2\} < +\infty$ (see Lemma 4.11 of (21) and Problem 14 in section 8.3 of (7)). Let $M > 0, 0 < \epsilon < 1$, $\beta_n = |\Pi^{(n)}(\frac{M}{\mu^{(n)}})|$ and $n_0 = \min\{i | \mu^{(i)} > 0\}$. To avoid invalid calculations, we set $\mu^{(n)} = 1$ if $n < n_0$. Using the Markov property, we have

$$T_1^{(n)} \ll \frac{M}{\mu^{(n)}} + \bar{T}_1^{(\beta_n)} \mathbf{1}_{T_1^{(n)} \geq \frac{M}{\mu^{(n)}}},$$

where $\bar{T}_1^{(n)} \stackrel{(d)}{=} T_1^{(n)}, n \geq 2$ and conditional on β_n, $\bar{T}_1^{(\beta_n)}$ is independent of $\mathbf{1}_{T_1^{(n)} \geq \frac{M}{\mu^{(n)}}}$. Then for $n\epsilon \geq n_0$,

$$\begin{aligned}
\mathbb{E}[(\mu^{(n)}T_1^{(n)})^k] &\leq \mathbb{E}[(M + \mu^{(n)}\bar{T}_1^{(\beta_n)} \mathbf{1}_{\mu^{(n)}T_1^{(n)} > M})^k] \leq (2M)^k + \mathbb{E}[(2\mu^{(n)}\bar{T}_1^{(\beta_n)} \mathbf{1}_{\mu^{(n)}T_1^{(n)} > M})^k] \\
&\leq (2M)^k + (\mathbb{E}[2\mu^{(n)}\bar{T}_1^{(n)} \mathbf{1}_{\beta_n = n}])^k] + \mathbb{E}[(2\mu^{(n)}\bar{T}_1^{(\beta_n)} \mathbf{1}_{\mu^{(n)}T_1^{(n)} > M, n\epsilon \leq \beta_n \leq n-1})^k] \\
&\quad + \mathbb{E}[(2\mu^{(n)}\bar{T}_1^{(\beta_n)} \mathbf{1}_{\mu^{(n)}T_1^{(n)} > M, \beta_n < n\epsilon})^k] \\
&\leq (2M)^k + \exp(-\frac{M g_n}{\mu^{(n)}}) \mathbb{E}[(2\mu^{(n)}\bar{T}_1^{(n)})^k] \\
&\quad + \mathbb{P}(\mu^{(n)}T_1^{(n)} > M)(2\frac{\mu^{(n)}}{\mu^{(n\epsilon)}})^k \max\{\mathbb{E}[(\mu^{(j)}\bar{T}_1^{(j)})^k] | j \in [n\epsilon, n-1]\} \\
&\quad + \mathbb{P}(\beta_n < n\epsilon)\mathbb{E}[\frac{\beta_n}{n}(2\frac{\mu^{(n)}}{\mu^{(\beta_n)}})^k (\mu^{(\beta_n)}\bar{T}_1^{(\beta_n)})^k | \beta_n < n\epsilon],
\end{aligned} \tag{4.40}$$

where $\exp(-\frac{Mg_n}{\mu^{(n)}})$ in the second term at right of the last inequality is the probability for no coalescence within $[0, \frac{M}{\mu^{(n)}}]$. The third term is due to the fact that $\mu^{(n)}$ is an increasing function of n when $n \geq n_0$. The fourth term is due to exchangeability which says that the probability for $\{1\}$ not to have coalesced at $\frac{M}{\mu^{(n)}}$ when there exist only β_n blocks is less than $\frac{\beta_n}{n}$. One needs the following three estimates to prove the boundedness of $(\mathbb{E}[(\mu^{(n)}T_1^{(n)})^k])_{n\geq 2}$.

— Estimation of $\exp(-\frac{Mg_n}{\mu^{(n)}})2^k$: Notice that for $n \geq n_0$,

$$\begin{aligned}\frac{g_n}{\mu^{(n)}} &= \frac{\int_0^1 (1-(1-x)^n - nx(1-x)^{n-1})x^{-2}\Lambda(dx)}{\int_{\frac{1}{n}}^1 x^{-1}\Lambda(dx)} \\ &\geq \frac{\int_{\frac{1}{n}}^1 (1-(1-x)^n - nx(1-x)^{n-1})x^{-2}\Lambda(dx)}{\int_{\frac{1}{n}}^1 x^{-1}\Lambda(dx)} \geq \frac{e-2}{e}.\end{aligned}$$

And if $2 \leq n < n_0$, we have $\exp(-\frac{Mg_n}{\mu^{(n)}}) = \exp(-Mg_n) \overset{M\to+\infty}{\to} 0$. Hence if M is large enough, we have, for any $n \geq 2$,

$$\exp(-\frac{Mg_n}{\mu^{(n)}})2^k \leq \frac{1}{4}. \tag{4.41}$$

— Estimation of $\mathbb{P}(\mu^{(n)}T_1^{(n)} > M)(2\frac{\mu^{(n)}}{\mu^{(n\epsilon)}})^k$: Due to Corollary 4.6, we get $\lim_{n\to+\infty} \frac{\mu^{(n)}}{\mu^{(n\epsilon)}} = 1$, and Theorem 4.1 gives $\lim_{n\to+\infty} \mathbb{P}(\mu^{(n)}T_1^{(n)} > M) = \exp(-M)$. Hence by taking M large enough, we have for any $n \geq 2$,

$$\mathbb{P}(\mu^{(n)}T_1^{(n)} > M)(2\frac{\mu^{(n)}}{\mu^{(n\epsilon)}})^k \leq \frac{1}{4}. \tag{4.42}$$

— Estimation of $\frac{\beta_n}{n}(2\frac{\mu^{(n)}}{\mu^{(\beta_n)}})^k, \beta_n < n\epsilon$: Using the notations in Proposition 4.2, for $\beta_n \geq n_0$, we have

$$\frac{\mu^{(n)}}{\mu^{(\beta_n)}} = \exp(\int_{\frac{1}{n}}^{\frac{1}{\beta_n}} \frac{f(x)}{x}dx)\frac{1-f(\frac{1}{n})}{1-f(\frac{1}{\beta_n})}. \tag{4.43}$$

Let $n_1 > n_0$ such that for any $n \geq n_1$, we have $f(\frac{1}{n}) \leq \frac{1}{2k}$. Hence for any $a, b \geq n_1$, $\frac{1-f(a)}{1-f(b)} \leq 2$. This n_1 can be found since $f(\frac{1}{n})$ tends to 0 as n tends to $+\infty$. Thus (4.43) implies, for $\beta_n \geq n_1$,

$$\frac{\mu^{(n)}}{\mu^{(\beta_n)}} \leq 2(\frac{n}{\beta_n})^{\frac{1}{2k}}.$$

Hence if $n_1 \leq \beta_n < n\epsilon$ and $\epsilon \leq 4^{-2k-2}$,

$$\frac{\beta_n}{n}(2\frac{\mu^{(n)}}{\mu^{(\beta_n)}})^k \leq 4^k(\frac{\beta_n}{n})^{1/2} < 4^k(\epsilon)^{1/2} \leq \frac{1}{4}.$$

If $\beta_n < n_1$, due to Corollary 4.6, one could find a large number n_2 such that $n_2 > n_1$ and for any $n \geq n_2$

$$\frac{\beta_n}{n}(2\frac{\mu^{(n)}}{\mu^{(\beta_n)}})^k \leq \frac{1}{4}.$$

In total, if $n \geq n_2$ and $\beta_n < n\epsilon$, then

$$\frac{\beta_n}{n}(2\frac{\mu^{(n)}}{\mu^{(\beta_n)}})^k \leq \frac{1}{4}. \tag{4.44}$$

Using (4.40), (4.41), (4.42) and (4.44), we get

$$\begin{aligned}\mathbb{E}[(\mu^{(n)}T_1^{(n)})^k] &\leq \frac{4}{3}(2M)^k + \frac{1}{3}\max\{\mathbb{E}[(\mu^{(j)}\bar{T}_1^{(j)})^k]|j \in [n\epsilon, n-1]\} + \frac{1}{3}\mathbb{E}[(\mu^{(\beta_n)}\bar{T}_1^{(\beta_n)})^k|\beta_n < n\epsilon] \\ &\leq \frac{4}{3}(2M)^k + \frac{2}{3}\max\{\mathbb{E}[(\mu^{(j)}\bar{T}_1^{(j)})^k]|j \leq n-1\}. \qquad (4.45)\end{aligned}$$

The above inequality is valid for a large M, $\epsilon = 4^{-2k-2}$ and $n \geq n_2$. Let $C_3 \geq \max\{\mathbb{E}[(\mu^{(j)}T_1^{(j)})^k], 4(2M)^k|2 \leq j < n_2\}$, then for any $n \geq 2$, $C_3 \geq \mathbb{E}[(\mu^{(n)}T_1^{(n)})^k]$ using (4.45). Then we can conclude.

The case of multiple external branch lengths is merely a consequence of the case of one external branch length, the Cauchy-Schwarz inequality and also a uniform integrability argument. □

Proof of Corollary 4.2

Proof. Notice that $\{T_i^{(n)}\}_{1\leq i\leq n}$ are exchangeable. Hence Corollary 4.1 shows that

$$\lim_{n\to+\infty}\mathbb{E}[\mu^{(n)}L_{ext}^{(n)}/n] = \lim_{n\to+\infty}\mathbb{E}[\mu^{(n)}(T_1^{(n)} + T_2^{(n)} + \cdots + T_n^{(n)})/n] = \lim_{n\to+\infty}\mathbb{E}[\mu^{(n)}T_1^{(n)}] = 1,$$

and

$$\begin{aligned}\lim_{n\to+\infty}\mathrm{Var}(\mu^{(n)}L_{ext}^{(n)}/n) &= \lim_{n\to+\infty}\frac{\mathbb{E}[n(\mu^{(n)}T_i^{(n)})^2] + n(n-1)\mathbb{E}[(\mu^{(n)})^2T_1^{(n)}T_2^{(n)}] - n^2(\mathbb{E}[\mu^{(n)}T_1^{(n)}])^2}{n^2} \\ &= \lim_{n\to+\infty}\frac{\mathrm{Var}(\mu^{(n)}T_1^{(n)}) + (n-1)\mathrm{cov}(\mu^{(n)}T_1^{(n)}, \mu^{(n)}T_2^{(n)})}{n} = 0.\end{aligned}$$

Hence $\mu^{(n)}L_{ext}^{(n)}/n$ converges in L^2 to 1. □

Before proving Corollary 4.3, we study at first a problem of sensibility of a recurrence satisfied by $(\mathbb{E}[T_1^{(n)}])_{n\geq 2}$. More precisely, if $a_n = \mathbb{E}[T_1^{(n)}]$, then a_n satisfies a recurrence (see (11)): $a_1 = 0$, and for $n \geq 2$, we have

$$a_n = c_n + \sum_{k=1}^{n-1} p_{n,k} \frac{k-1}{n} a_k, \tag{4.46}$$

where $(c_n)_{n\geq 2} = (\frac{1}{g_n})_{n\geq 2}$ and $p_{n,k}$ is defined in (4.2). Due to Corollary 4.1, we have $\lim_{n\to+\infty} \mu^{(n)} a_n = 1$. The question is as follows: what is the limit behavior of a_n if we set initially the values of $(a_i)_{1\leq i\leq n_0}$ with $n_0 \geq 1$ without using (4.46) and replace c_n by $c'_n = \frac{1}{g_n} + o(\frac{1}{g_n})$? It is answered in the next lemma.

Lemma 4.6. *Let $n_0 \geq 1$ and $(a'_i)_{1\leq i\leq n_0}$ be n_0 real numbers. For $n > n_0$, let*

$$a'_n = c'_n + \sum_{k=1}^{n-1} p_{n,k} \frac{k-1}{n} a'_k, \tag{4.47}$$

where $(c'_n)_{n>n_0}$ is a sequence which satisfies $c'_n = \frac{1}{g_n} + o(\frac{1}{g_n})$. Then

$$\lim_{n\to+\infty} \mu^{(n)} a'_n = 1.$$

Proof. We fix $\epsilon > 0$ and let $n_\epsilon > n_0$ such that $c'_n \leq \frac{1+\epsilon}{g_n}$ for $n > n_\epsilon$. We set $M = \max\{|a'_i|, a_i | 1 \leq i \leq n_\epsilon\}$.

Let us at first look at (4.46) which has the following interpretation using random walk: A walker stands initially at point n, then after time c_n, he jumps to point k_1 with probability p_{n,k_1}, then after time $\frac{k_1-1}{n} c_{k_1}$, he jumps to k_2 with probability p_{k_1,k_2}, and then after time $\frac{(k_1-1)(k_2-1)}{nk_1} c_{k_2}$, he jumps to the next point, etc. If he falls at point 1, then this walk is finished. It is easy to see that a_n is the expectation of the total walking time. One notices that there is a scaling effect on the walking time. More precisely, let $l \geq 1$ and $n = k_0 > k_1 > \cdots > k_l \geq 1$ such that the walker jumps from k_i to k_{i+1} for $0 \leq i \leq l-1$. Then conditional on this walking history, the remaining walking time is $\left(\Pi_{i=0}^{l-1} \frac{k_{i+1}-1}{k_i}\right) a_{k_l}$.

The recurrence (4.47) has the same interpretation. The difference is that one should stop the walker when he arrives at a point i within $[1, n_0]$ and one adds a scaled value of a'_i to the walking time (notice that a'_i can be negative). To estimate a'_n, we use a Markov chain $(W_i)_{i\geq 0}$ to couple the jumping structures of (4.46) and (4.47) : $W_0 = n$,

— If $W_i = k$ with $k \geq n_\epsilon$, then $W_{i+1} = k'$ with probability $p_{k,k'}$, where $1 \leq k' \leq k-1$;

— If $W_i < n_\epsilon$, then we set $W_j = W_i$ for any $j \geq i+1$.

Notice that the jumping dynamics of both recurrences is characterized by $(W_i)_{i\geq 0}$ until arriving at a point within $[1, n_\epsilon]$. And we also see that $(W_i)_{i\geq 0}$ is the discrete time Markov chain related to the block counting process $|\Pi^{(n)}|$ stopped at the first time arriving within $[1, n_\epsilon]$.

Let $\varsigma_n = \min\{i | W_i = W_{i+1}\}$, $C_{\varsigma_n} = \Pi_{i=0}^{\varsigma_n - 1} \frac{W_{i+1}-1}{W_i}$ and T_{ς_n} is set to be the time to ς_n of the random walk related to (4.46) and T'_{ς_n} be the corresponding time related to (4.47).

By the scaling effect of C_{ς_n} on the walking time, we get

$$a_n = \mathbb{E}[T_{\varsigma_n} + C_{\varsigma_n} a_{W_{\varsigma_n}}], a'_n = \mathbb{E}[T'_{\varsigma_n} + C_{\varsigma_n} a'_{W_{\varsigma_n}}].$$

Due to the definitions of M, n_ϵ, we obtain

$$a_n - M\mathbb{E}[C_{\varsigma_n}] \leq \mathbb{E}[T_{\varsigma_n}] \leq a_n; \quad a'_n - M\mathbb{E}[C_{\varsigma_n}] \leq \mathbb{E}[T'_{\varsigma_n}] \leq a'_n + M\mathbb{E}[C_{\varsigma_n}]; \quad \mathbb{E}[T'_{\varsigma_n}] \leq (1+\epsilon)\mathbb{E}[T_{\varsigma_n}].$$

Notice that $\mathbb{E}[C_{\varsigma_n}] \leq \frac{n_\epsilon}{n}$ and due to Corollary 4.6, we have $\lim_{n\to+\infty} \frac{M\mu^{(n)}}{n} = 0$. Hence $\lim_{n\to+\infty} M\mathbb{E}[C_{\varsigma_n}]\mu^{(n)} = 0$. Then we can conclude that for n large, $a'_n \leq (1+2\epsilon)a_n$. In the same way, we can prove also $a'_n \geq (1-2\epsilon')a_n$ for another small positive number ϵ' with n large enough. Then we can deduce the lemma. □

Proof of Corollary 4.3

Proof. Let $b_n = \mathbb{E}[\mu^{(n)} L^{(n)}_{total}/n]$. Then looking at the first coalescence of the process $\Pi^{(n)}$, we have,

$$b_1 = 0; b_n = \frac{\mu^{(n)}}{g_n} + \sum_{k=1}^{n-1} p_{n,k} \frac{k\mu^{(n)}}{n\mu^{(k)}} b_k, n \geq 2. \tag{4.48}$$

If for some k, $\mu^{(k)} = 0$, then we set $\mu^{(k)} = 1$ to avoid invalid calculations. To use Lemma 4.6, we write (4.48) as:

$$b_1 = 0; b_n = \frac{\mu^{(n)}}{g_n} + \sum_{k=1}^{n-1} p_{n,k} \frac{\mu^{(n)}}{n\mu^{(k)}} b_k + \sum_{k=1}^{n-1} p_{n,k} \frac{(k-1)\mu^{(n)}}{n\mu^{(k)}} b_k, n \geq 2. \tag{4.49}$$

We at first prove that $\sum_{k=1}^{n-1} p_{n,k} \frac{\mu^{(n)}}{n\mu^{(k)}} = o(\frac{\mu^{(n)}}{g_n})$. Indeed, due to (4.12), let $a = \int_0^1 (1-(1-x)^{n-1}) x^{-1} \Lambda(dx)$ and $M > 0$, then

$$\mathbb{P}(X_1^{(n)} \geq Ma) \leq \frac{\mathbb{E}[X_1^{(n)}]}{Ma} \leq \frac{n}{Mg_n}. \tag{4.50}$$

Using Corollary 4.6, we have $\limsup_{n\to+\infty} \frac{a}{n} \leq \lim_{n\to+\infty} \frac{\int_0^{\frac{1}{n}} (n-1)\Lambda(dx) + \mu^{(n)}}{n} = 0$, $\lim_{n\to+\infty} \frac{\mu^{(n)}}{\mu^{(n-Ma)}} =$

1. Then for n large enough,

$$\begin{aligned}\sum_{k=1}^{n-1} p_{n,k}\frac{\mu^{(n)}}{n\mu^{(k)}} &= \sum_{k=1}^{\lfloor n-Ma\rfloor} p_{n,k}\frac{\mu^{(n)}}{n\mu^{(k)}} + \sum_{k=\lfloor n-Ma\rfloor+1}^{n-1} p_{n,k}\frac{\mu^{(n)}}{n\mu^{(k)}}\\ &\le \mathbb{P}(X_1^{(n)} \ge Ma)\mathbb{E}[\frac{\mu^{(n)}}{n\mu^{(n-X_1^{(n)})}}|X_1^{(n)} \ge Ma] + \frac{\mu^{(n)}}{\mu^{(n-Ma)}n}\\ &\le \frac{\mu^{(n)}}{Mg_n}\max\{\frac{1}{\mu^{(k)}}|1\le k\le n\} + \frac{\mu^{(n)}}{\mu^{(n-Ma)}n},\end{aligned}$$

where the first term at the right of the the last inequality is due to (4.50) and can be made as small as we want w.r.t $\frac{\mu^{(n)}}{g_n}$ when M is large enough. Notice that $n^{-1} = o(\frac{\mu^{(n)}}{g_n})$ due to (4.4). Then the second term $\frac{\mu^{(n)}}{\mu^{(n-Ma)}n} = o(\frac{\mu^{(n)}}{g_n})$ using also $\lim_{n\to+\infty}\frac{\mu^{(n)}}{\mu^{(n-Ma)}} = 1$. Then $\sum_{k=1}^{n-1} p_{n,k}\frac{\mu^{(n)}}{n\mu^{(k)}} = o(\frac{\mu^{(n)}}{g_n})$.

We now only need to prove that $(b_k)_{k\ge 2}$ are bounded, since in this case, $\sum_{k=1}^{n-1} p_{n,k}\frac{\mu^{(n)}}{n\mu^{(k)}}b_k = o(\frac{\mu^{(n)}}{g_n})$ and we apply Lemma 4.6 to (4.49). We construct another recurrence:

$$b'_1 = 0; b'_n = \frac{C\mu^{(n)}}{g_n} + \sum_{k=1}^{n-1} p_{n,k}\frac{(k-1)\mu^{(n)}}{n\mu^{(k)}}b'_k, n\ge 2. \tag{4.51}$$

where C is a positive number. If $C = 1$, this is exactly a transformation of the recurrence (4.46). Let $M'(C) = \sup\{b'_n|n\ge 1\}$. Then it is easy to see that $M'(C) = CM'(1)$. Let $n_0 \ge 1$, such that for $n \ge n_0$, we have $\sum_{k=1}^{n-1} p_{n,k}\frac{\mu^{(n)}}{n\mu^{(k)}}M'(1) \le \frac{1}{2}\frac{\mu^{(n)}}{g_n}$. Then for $C\ge 2, n\ge n_0$,

$$\frac{\mu^{(n)}}{g_n} + \sum_{k=1}^{n-1} p_{n,k}\frac{\mu^{(n)}}{n\mu^{(k)}}M'(C) \le \frac{C\mu^{(n)}}{g_n}. \tag{4.52}$$

For $2\le n < n_0$, we set C large enough such that

$$\frac{\mu^{(n)}}{g_n} + \sum_{k=1}^{n-1} p_{n,k}\frac{\mu^{(n)}}{n\mu^{(k)}}\max\{b_i|1\le i< n_0\} \le \frac{C\mu^{(n)}}{g_n}. \tag{4.53}$$

Comparing the coefficients and initial values of recurrences (4.49) and (4.51) using (4.52) and (4.53), we deduce that $b_n \le b'_n \le M'(C)$. Hence we can conclude.

□

Acknowledgements: The author benefited from the support of the "Agence Nationale de la Recherche": ANR MANEGE (ANR-09-BLAN-0215).

The author wants to thank his supervisor, Prof Jean-Stéphane Dhersin, for fruitful discussions and also for his careful reading of a draft version resulting in many helpful comments. The author also wants to thank Prof Martin Möhle for a discussion on Theorem 4.1.

Bibliographie

[1] Arnason, E. (2004). Mitochondrial cytochrome b dna variation in the high-fecundity atlantic cod : trans-atlantic clines and shallow gene genealogy. *Genetics*, 166(4) :1871–1885.

[2] Berestycki, J., Berestycki, N., and Schweinsberg, J. (2008). Small-time behavior of beta coalescents. *Ann. Inst. H. Poincaré Probab. Statist.*, 44(2) :214–238.

[3] Billingsley, P. (1995). *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition. A Wiley-Interscience Publication.

[4] Blum, M. G. B. and François, O. (2005). Minimal clade size and external branch length under the neutral coalescent. *Adv. in Appl. Probab.*, 37(3) :647–662.

[5] Bolthausen, E. and Sznitman, A. (1998). On ruelle's probability cascades and an abstract cavity method. *Communications in mathematical physics*, 197(2) :247–276.

[6] Boom, J., Boulding, E., and Beckenbach, A. (1994). Mitochondrial DNA variation in introduced populations of Pacific oyster, Crassostrea gigas, in British Columbia. *Canadian journal of fisheries and aquatic sciences(Print)*, 51(7) :1608–1614.

[7] Breiman, L. (1992). Probability, classics in applied mathematics, vol. 7. *Society for Industrial and Applied Mathematics (SIAM), Pennsylvania.*

[8] Caliebe, A., Neininger, R., Krawczak, M., and Roesler, U. (2007). On the length distribution of external branches in coalescence trees : genetic diversity within species. *Theoretical Population Biology*, 72(2) :245–252.

[9] Delmas, J.-F., Dhersin, J.-S., and Siri-Jégousse, A. (2008). Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.*, 18(2) :997–1025.

[10] Den Hollander, F. (2008). *Large deviations*, volume 14. Amer Mathematical Society.

[11] Dhersin, J.-S. and Möhle, M. (2012). On the external branches of coalescent processes with multiple collisions with an emphasis on the bolthausen-sznitman coalescent. *arXiv preprint arXiv :1209.3380.*

[12] Dhersin, J.-S., Siri-Jégousse, A., Freund, F., and Yuan, L. (2013). On the length of an external branch in the beta-coalescent. *Stochastic Process. Appl.*, 123 :1691–1715.

[13] Drmota, M., Iksanov, A., Moehle, M., and Roesler, U. (2007). Asymptotic results concerning the total branch length of the Bolthausen-Sznitman coalescent. *Stochastic Process. Appl.*, 117(10) :1404–1421.

[14] Eldon, B. and Wakeley, J. (2006). Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics*, 172 :2621–2633.

[15] Feller, W. (1971). *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York.

[16] Freund, F. and Möhle, M. (2009). On the time back to the most recent common ancestor and the external branch length of the Bolthausen-Sznitman coalescent. *Markov Process. Related Fields*, 15(3) :387–416.

[17] Gnedin, A., Iksanov, A., and Marynych, A. (2012). On asymptotics of the beta-coalescents. *arXiv preprint arXiv :1203.3110*.

[18] Gnedin, A., Iksanov, A., and Möhle, M. (2008). On asymptotics of exchangeable coalescents with multiple collisions. *J. Appl. Probab.*, 45(4) :1186–1195.

[19] Hedgecock, D. (1994). Does variance in reproductive success limit effective population sizes of marine organisms ? *Genetics and evolution of aquatic organisms*, page 122.

[20] Janson, S. and Kersting, G. (2011). On the total external length of the kingman coalescent. *Electronic Journal of Probability*, 16 :2203–2218.

[21] Kallenberg, O. (2002). *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition.

[22] Kingman, J. (1982a). The coalescent. *Stochastic Process. Appl.*, 13(3) :235–248.

[23] Kingman, J. (1982b). On the genealogy of large populations. *J. Appl. Probab.*, (Special Vol. 19A) :27–43. Essays in statistical science.

[24] Kingman, J. (2000). Origins of the Coalescent 1974-1982. *Genetics*, 156(4) :1461–1463.

[25] Möhle, M. (2010a). Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson-Dirichlet coalescent. *Stochastic Process. Appl.*, 120(11) :2159–2173.

[26] Möhle, M. (2010b). Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter poisson-dirichlet coalescent. *Stochastic Processes and their Applications*, 120(11) :2159–2173.

[27] Pitman, J. (1999). Coalescents with multiple collisions. *Ann. Probab.*, 27(4) :1870–1902.

[28] Sagitov, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 36(4) :1116–1125.

[29] Schweinsberg, J. (2000). Coalescents with simultaneous multiple collisions. *Electron. J. Probab.*, 5 :Paper no. 12, 50 pp. (electronic).

Chapitre 5

Résultats asymptotiques sur la taille du clade minimal et sur des fonctionelles liées au Beta-coalescent

Version non modifiée de l'article *Asymptotics of the minimal clade size and related functionals of certain Beta-coalescents*

soumis à

ALEA (Latin American Journal of Probability and Mathematical Statistics).

5.1 Introduction and main results

Coalescent theory was initiated by Kingman ((24; 25; 26)) to model the genealogical tree of a sample of n individuals of a certain population. The so-called Kingman n-coalescent is a continuous-time Markov chain taking values in $\mathcal{P}_n$, the set of partitions of $\mathbb{N}_n = \{1, 2, \cdots, n\}$. It starts from n singletons $\{1\}, \{2\}, \cdots, \{n\}$ representing n individuals (or lineages) and at any time, each couple of blocks merges (or coalesces) independently into one block at rate 1. The process reaches almost surely in finite time the absorbing state $\{1, 2, \cdots, n\}$ which is called MRCA (most recent common ancestor). Kingman showed that the genealogy of a sample of size n in a population evolving according to the Cannings population model of size N converges in the sense of finite dimensional distribution to the Kingman n-coalescent when N goes to ∞, under some assumptions over the reproduction law in the Cannings model. Roughly speaking, it is required that one individual in the population should not have a lot of progenies so that its children occupy a large ratio of the next generation. This assumption happens to fail in some marine species (see (20), (16), (9), (1)). To model this phenomenon, Pitman ((28)) and Sagitov ((30)) introduced at the same time the Λ n-coalescent, denoted by $\Pi^{(n)} = (\Pi^{(n)}(t), t \geq 0)$. It is characterized by a finite measure Λ on $[0, 1]$.

The process $\Pi^{(n)}$ is still a continuous-time Markov chain starting from $\{1\}, \{2\}, \cdots, \{n\}$, but with the following dynamics: at any time $t \geq 0$, if $\Pi^{(n)}(t)$ has b blocks ($b \geq 2$), then each k-tuple ($2 \leq k \leq b$) of blocks coalesces together into one at rate

$$\lambda_{b,k} := \int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx). \tag{5.1}$$

As a consequence, the rate to the next coalescence is

$$g_b := \sum_{k=2}^{b} \binom{b}{k} \lambda_{b,k}. \tag{5.2}$$

In particular, the Kingman n-coalescent is a special Λ n-coalescent with Λ being the Dirac measure at 0.

Definition (5.1) is a reformulation of two properties of the process $\Pi^{(n)}$ (see (28)). The first one is exchangeability. Let ρ_n be a permutation of $\mathbb{N}_n$. The map ρ_n induces naturally a map $\bar{\rho}_n$ on $\mathcal{P}_n$. Then we have

$$\bar{\rho}_n \circ \Pi^{(n)} \stackrel{(d)}{=} \Pi^{(n)}.$$

The second one is consistency. For any $2 \leq m \leq n$, let $\varrho_{n,m}$ be the natural restriction from $\mathbb{N}_n$ to $\mathbb{N}_m$ and $\bar{\varrho}_{n,m}$ the induced map from $\mathcal{P}_n$ to $\mathcal{P}_m$. Then

$$\bar{\varrho}_{n,m} \circ \Pi^{(n)} \stackrel{(d)}{=} \Pi^{(m)}.$$

These two properties also imply that one can build a projective limit process, the so-called Λ-coalescent, denoted by $\Pi = (\Pi(t), t \geq 0)$, taking values in the set $\mathcal{P}$ of partitions of $\mathbb{N}$. For any restriction ϱ_n from $\mathbb{N}$ to $\mathbb{N}_n$ and its induced map $\bar{\varrho}_n$ from $\mathcal{P}$ to $\mathcal{P}_n$,

$$\bar{\varrho}_n \circ \Pi \stackrel{(d)}{=} \Pi^{(n)}.$$

The Beta$(2-\alpha,\alpha)$-coalescent with $0<\alpha<2$ is a special and important example of Λ-coalescents. In this case, Λ is the Beta measure with parameters $2-\alpha$ and α. If α tends to 2, then the limit process obtained is the Kingman coalescent. If $\alpha = 1$, the process obtained is the celebrated Bolthausen-Sznitman coalescent ((8)). This article deals with the case $1<\alpha<2$ and, for the sake of simplicity, we will refer to Beta$(2-\alpha,\alpha)$-coalescent as Beta-coalescent. This class of coalescent processes was introduced by Schweinsberg ((32)) and deeply studied in (2),(3). In particular, in (3), many results on the small-time behavior of various functionals of the Beta-coalescent are discovered. Meanwhile, many asymptotic studies, motivated by biology, have been developed for the Beta n-coalescent (see for example (12), (22), (13), (11)) , when n goes to ∞.

In this paper, we aim to study more asymptotic results on some functionals of the Beta n-coalescent with $1<\alpha<2$ when n grows to ∞. We denote

$$A \sim B,$$

if $\frac{A}{B}$ tends deterministically or randomly to 1 in the limit, depending on different contexts. Here A, B can be functions, sequences of real values, random variables. Denote by $\stackrel{a.s.}{\to}$ the almost sure convergence and by $\stackrel{P}{\to}$ the convergence in probability.

The length of the external branch of individual i, also called *unicity of individual* i by biologists ((29)), is denoted by $T_i^{(n)}$. It is the coalescence time of $\{i\}$, defined as follows

$$T_i^{(n)} := \sup\{t : \{i\} \in \Pi^{(n)}(t)\}.$$

The length of a randomly chosen external branch provides a measure of the genetic variation of the population since it gives some information on the "distance" of an individual to the rest of the sample. Exchangeability of the coalescent implies that

$$T_i^{(n)} \stackrel{(d)}{=} T_1^{(n)}, \ 1 \leq i \leq n.$$

The law of $T_1^{(n)}$ has interested many people since the first article (7) dealing with the Kingman coalescent case. We give a short survey of results already discovered.

- Kingman: $nT_1^{(n)}$ converges in distribution to a random variable with density $\frac{8}{(2+t)^3}\mathbf{1}_{t\geq 0}$ ((7),(10)).
- Beta-coalescent with $1<\alpha<2$: $n^{\alpha-1}T_1^{(n)} \stackrel{(d)}{\longrightarrow} T$, where T is a random variable with density function f_T:

$$f_T(t) = \frac{1}{(\alpha-1)\Gamma(\alpha)}(1+\frac{t}{\alpha\Gamma(\alpha)})^{-\frac{\alpha}{\alpha-1}-1}\mathbf{1}_{t\geq 0}. \tag{5.3}$$

(see (13) where the result is stated in a more general case).

— $\lim_{n\to\infty} \frac{g_n}{n\mu^{(n)}} = 0$ where g_n is defined in (5.2) and $\mu^{(n)} = \int_{1/n}^1 x^{-1}\Lambda(dx)$: $\mu^{(n)}T_1^{(n)}$ is asymptotically distributed as an exponential random variable with mean 1 ((34)). This class of processes contains Beta-coalescents with $0 < \alpha < 1$ (see also (28), (27), (19) for other proofs) and the Bolthausen-Sznitman coalescent (see also (14), (17) for other proofs).

In this paper, we prove the following

Theorem 5.1. *Consider a Beta n-coalescent with $1 < \alpha < 2$. For any fixed $k \in \mathbb{N}$, as $n \to \infty$,*

$$n^{\alpha-1}(T_1^{(n)}, \cdots, T_k^{(n)}) \xrightarrow{(d)} (T_1, \cdots, T_k), \tag{5.4}$$

where $(T_i, i \in \mathbb{N})$ are i.i.d. copies of T with density (5.3).

A similar result has been proved for Bolthausen-Sznitman coalescent ((14)), but the asymptotic independence is not true for coalescents satisfying $\int_0^1 x^{-1}\Lambda(dx) < \infty$ ((27)).

Let $K^{(n)} = (K^{(n)}(t), t \geq 0)$ denote the block-counting process of $\Pi^{(n)}$, i.e., $K^{(n)}(t)$ stands for the number of blocks of the partition $\Pi^{(n)}(t)$ for $t \geq 0$. Define

$$Q^{(n)} := K^{(n)}((T_1^{(n)})_-) - K^{(n)}(T_1^{(n)}) + 1,$$

where $(T_1^{(n)})_-$ is the time just prior to $T_1^{(n)}$. In other words, $Q^{(n)}$ is the number of blocks involved in the coalescence event of $\{1\}$ in $\Pi^{(n)}$.

Theorem 5.2. *Consider a Beta n-coalescent with $1 < \alpha < 2$. $Q^{(n)}$ converges in law to a random variable Q taking values in $\{2, 3, \cdots\}$ such that for any $k \geq 2$*

$$q_k := \mathbb{P}(Q = k) = \frac{(\alpha-1)\Gamma(k-\alpha)}{\Gamma(k)\Gamma(2-\alpha)}. \tag{5.5}$$

Furthermore, $Q^{(n)}$ and $T_1^{(n)}$ are asymptotically independent.

It is interesting to relate this result with the asymptotics of the size of the first jump of the block-counting process of the n-coalescent. Let $\tau_1^{(n)}$ be the time of the first jump of the process. The limit law of $n - K^{(n)}(\tau_1^{(n)}) = K^{(n)}(0) - K^{(n)}(\tau_1^{(n)})$ is given by (11) in (12), see also (22) and (23). It is worth noting that the random variable Q is the size-biased version of this limit variable. This phenomenon is not surprising though. It can be coarsely explained by the fact that big jumps are more likely to involve the block $\{1\}$ than small ones.

In the Kingman coalescent, $Q^{(n)} = Q = 2$ almost surely. The following proposition shows that, for Λ-coalescents satisfying $\int_{[0,1]} x^{-1}\Lambda(dx) < \infty$, $Q^{(n)}$ converges in probability to infinity.

Proposition 5.1. *Consider a Λ n-coalescent with the characteristic measure satisfying $\int_0^1 x^{-1}\Lambda(dx) < \infty$, then $Q^{(n)}$ converges in probability to infinity.*

The above proposition is even true in the Bolthausen-Sznitman coalescent case (see Remark 5.4).

A quantity of interest in biology is the *minimal clade size*. It is the size of the minimal clade of a randomly chosen individual (or of the individual 1, considered in this paper). The minimal clade is the block that contains 1 at the time $\{1\}$ is coalesced. The size of the minimal clade tells how many individuals share the genealogy with individual 1 after time $T_1^{(n)}$. Let us denote the minimal clade size by $Y^{(n)}$. In the Kingman case, Blum and François ((7)) showed that

$$\mathbb{P}(Y^{(n)} = k) = \frac{4}{(k+1)k(k-1)}, k = 2, \cdots, n-1; \ \mathbb{P}(Y^{(n)} = n) = \frac{2}{n(n-1)}.$$

Freund and Siri-Jégousse (18) studied the case of the Bolthausen-Sznitman coalescent. In this case

$$\frac{\ln Y^{(n)}}{\ln n} \xrightarrow{(d)} U_{[0,1]},$$

where $U_{[0,1]}$ is a uniform variable over $[0, 1]$. Asymptotics of moments were also found.

We state out our result by at first giving some notations.

— Let μ be Slack's probability distribution on $[0, \infty)$ (see (33)) characterized by its Laplace transform

$$\mathcal{L}_\mu(\lambda) = \int_0^\infty e^{-\lambda x} \mu(dx) = 1 - (1 + \lambda^{1-\alpha})^{-\frac{1}{\alpha-1}}, \quad \lambda \geq 0. \tag{5.6}$$

— Define a 1-parameter family of laws (not a process) $(\beta(t), t \geq 0)$ such that for any $k \geq 1$

$$\mathbb{P}(\beta(t) = k) = \frac{1}{\Gamma(k)} \left(\frac{t}{\alpha\Gamma(\alpha)}\right)^{\frac{k-1}{\alpha-1}} \int_0^\infty e^{-x\left(\frac{t}{\alpha\Gamma(\alpha)}\right)^{\frac{1}{\alpha-1}}} x^k \mu(dx). \tag{5.7}$$

The following result could be regarded as a consequence of Theorem 5.2.

Theorem 5.3. *Consider a Beta n-coalescent with $1 < \alpha < 2$. Let $(\beta_i(t), t \geq 0)_{i\geq 1}$ be i.i.d. copies of $(\beta(t), t \geq 0)$ and Q, T be random variables defined respectively in (5.5) and (5.3). Assume that $(\beta_i(t), t \geq 0)_{i\geq 1}, Q, T$ are all independent. Then*

$$Y^{(n)} \xrightarrow{(d)} Y = 1 + \sum_{i=1}^{Q-1} \beta_i(T). \tag{5.8}$$

The independence among $(\beta_i(t), t \geq 0)_{i\geq 1}, Q, T$ allows to express the law of Y as follows: for any $l \geq 2$,

$$\mathbb{P}(Y = l) = \int_0^\infty \sum_{k=2}^{l} q_k \sum_{i_1+\cdots+i_{k-1}=l-1} \left(\Pi_{j=1}^{k-1}\mathbb{P}(\beta(t) = i_j)\right) f_T(t)dt.$$

Next, we establish a close relation between the random variable Q and the family $(\beta(t), t \geq 0)$. Notice that $\lim_{t\to 0+} \mathbb{P}(\beta(t) = 1) = 1$.

Proposition 5.2. 1) *For any* $k \geq 2$,

$$q_k = (\alpha - 1)\Gamma(\alpha) \lim_{t\to 0+} \frac{\mathbb{P}(\beta(t) = k)}{t}. \tag{5.9}$$

2) *The Laplace transform of* Q *is*

$$\mathbb{E}[e^{-\lambda Q}] = \lim_{t\to 0+} \mathbb{E}[(\alpha - 1)\Gamma(\alpha) \frac{e^{-\lambda\beta(t)}\mathbf{1}_{\beta(t)\geq 2}}{t}] = e^{-\lambda}\Big(1 - (1 - e^{-\lambda})^{\alpha-1}\Big) \tag{5.10}$$

for any $\lambda \geq 0$.

The law of Y looks quite complicated, which may harm the applicability of the result. However the clarification given below could at some point improve the situation.

Corollary 5.1. *If* k *tends to* ∞, *one has* $\mathbb{P}(Y > k) \sim \frac{\int_0^\infty t^{\alpha-1} f_T(t)dt}{((\alpha-1)\Gamma(\alpha))^{\alpha-1}\Gamma(1-(\alpha-1)^2)} k^{-(\alpha-1)^2}$.

If α goes to 1, $k^{-(\alpha-1)^2}$ goes to 1. This is consistent with the Bolthausen-Sznitman case where $Y = \infty$ almost surely. If α tends to 2, $k^{-(\alpha-1)^2}$ goes to k^{-1}. This is in fact not consistent with the law of Y in the Kingman case. The corollary reveals some kind of "discontinuity" between the Beta-coalescent and the Kingman coalescent.

The size of the block containing one specific integer evolves in an increasing way at different speed. It is clear that at time $(T_1^{(n)})_-$ the block containing 1 is still of size 1 while other blocks could have grown quite a lot. One way to measure this speed is to consider the size of the largest block at time $T_1^{(n)}$. We denote this variable by $\tilde{W}^{(n)}$. The bigger $\tilde{W}^{(n)}$ is, the more inhomogeneous the speed is. To study $\tilde{W}^{(n)}$, we first consider the size of the largest block at any time t, denoted by $W^{(n)}(t)$. In this way, we have

$$\tilde{W}^{(n)} = W^{(n)}(T_1^{(n)}).$$

Theorem 5.4. *Consider a Beta* n*-coalescent with* $1 < \alpha < 2$,

$$\frac{W^{(n)}((\alpha - 1)\alpha\Gamma(\alpha)n^{1-\alpha}t)}{n^{\frac{1}{\alpha}}} \xrightarrow{(d)} W(t), \tag{5.11}$$

where $W(t)$ *is a positive random variable with a type-2 Gumbel law, i.e., for any* $x \geq 0$, $\mathbb{P}(W(t) \leq x) = e^{-x^{-\alpha}\frac{(\alpha-1)t}{\Gamma(2-\alpha)}}$.

The methodology employed to prove the above theorem is similar to that used in the proof of Proposition 1.6 in (3), although there are some small differences.

The following result about $\tilde{W}^{(n)}$ happens to be a straightforward consequence of the above theorem.

Corollary 5.2. *As n tends to ∞,*

$$\frac{\tilde{W}^{(n)}}{n^{\frac{1}{\alpha}}} \xrightarrow{(d)} \tilde{W}, \tag{5.12}$$

where $\tilde{W}$ is a positive random variable such that for any $x \geq 0$,

$$\mathbb{P}(\tilde{W} \leq x) = \int_0^\infty e^{-x^{-\alpha}\frac{t}{\alpha\Gamma(\alpha)\Gamma(2-\alpha)}} f_T(t)dt.$$

This paper is organized as follows. In Section 2, we study external branch lengths and the block-counting process in small time and prove Theorem 5.1. In Section 3, we focus on the way of coalescing an external branch and prove Theorem 5.2, Proposition 5.1, Theorem 5.3, Proposition 5.2 and Corollary 5.1. Section 4 is devoted to the size of the largest block and Theorem 5.4 and Corollary 5.2 are proved.

5.2 External branch lengths

5.2.1 Ranked Λ-coalescent

Assume from now on that $1 < \alpha < 2$. Let $\Pi = (\Pi(t), t \geq 0)$ be the Beta-coalescent and denote by $K = (K(t), t > 0)$ the block-counting process of Π, i.e., $K(t)$ stands for the number of blocks of $\Pi(t)$. It is known that Π is coming down from infinity: for any $t > 0$, $K(t)$ is finite almost surely ((31)). Recall that for any $t \geq 0$, $\Pi(t)$ is an exchangeable random partition of $\mathbb{N}$. Applying Kingman's paintbox theorem on exchangeable random partitions ((24)), almost surely, for every block $B \in \Pi(t)$, there exists the following limit which is called the asymptotic frequency of B:

$$\lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}_{i\in B}.$$

Furthermore, when $t > 0$, the sum of all asymptotic frequencies equals 1 ((28)).When $t = 0$, every block is a singleton and hence has the asymptotic frequency 0. Pitman ((28)) shows that almost surely for all $t \geq 0$, every block in $\Pi(t)$ has the asymptotic frequency. Hence if $t > 0$, one can reorder all the asymptotic frequencies in a non-increasing way to define a sequence $\Theta(t) = \{\theta_1(t), \theta_2(t), \cdots, \theta_{K(t)}(t)\}$ where $\theta_1(t) \geq \theta_2(t) \geq \cdots \geq \theta_{K(t)}(t)$ and $\sum_{i=1}^{K(t)} \theta_i(t) = 1$. At time $t = 0$, since every block has asymptotic frequency 0, one can naturally set $\Theta(0) = \{0, 0, \cdots\}$. Then the process $\Theta = (\Theta(t), t \geq 0)$ is well defined and called the *ranked* Λ*-coalescent*.

Given $\Theta(t)$ with $t > 0$, one can recover the distribution of $\Pi(t)$ using again Kingman's paintbox theorem. Let us at first divide $[0, 1]$ into $K(t)$ intervals such that the lengths of intervals correspond one to one to the elements of $\Theta(t)$. Then we throw individuals

$1, 2, \cdots$ uniformly and independently into $[0, 1]$. Finally, all individuals within one interval form a block and this procedure provides a random exchangeable partition which has the same law as $\Pi(t)$. It is of course possible, thanks to the consistency property, to build the restricted partition $\Pi^{(n)}(t)$ using the same procedure but throwing nothing but n particles instead of an infinity.This construction will be the key point of our proofs.

5.2.2 Properties of the ranked Λ-coalescent

Let $K(t, x) := \#\{i : \theta_i(t) \leq x\}$ for any $x \in [0, 1]$. Let $\varsigma(t)$ be a size-biased picking of $\Theta(t)$, i.e., $\varsigma(t)$ is a discrete random variable such that

$$\mathbb{P}(\varsigma(t) = \theta_i(t)|\Theta(t)) = \theta_i(t) \times \#\{j : \theta_j(t) = \theta_i(t),\ 1 \leq j \leq K(t)\},\ 1 \leq i \leq K(t). \quad (5.13)$$

One can construct or regard $\varsigma(t)$ in the following way: Suppose that $[0, 1]$ is divided into $K(t)$ intervals whose lengths are in one-to-one correspondence to the elements of $\Theta(t)$. We throw a particle uniformly and independently over $[0, 1]$ and $\varsigma(t)$ is the length of the interval containing this particle.

Recall the measure μ defined in (5.6). It is easy to check that

$$\int_0^\infty y\mu(dy) = \frac{d\mathcal{L}_\mu(\lambda)}{d\lambda}|_{\lambda=0} = 1. \quad (5.14)$$

Proposition 5.3. *We have*

$$\lim_{t \longrightarrow 0+} \sup_{x \geq 0} \left| \mathbb{P}(\varsigma(t) \leq t^{\frac{1}{\alpha-1}} x|\Theta(t)) - \int_0^{x(\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}}} y\mu(dy) \right| = 0, a.s. \quad (5.15)$$

Proof. In order to simplify the notations, let us denote $f(t, x) = \mathbb{P}(\varsigma(t) \leq t^{\frac{1}{\alpha-1}} x|\Theta(t))$ and $f(x) = \int_0^{x(\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}}} y\mu(dy)$. Let

$$S_t = \sup_{x \geq 0} \left| t^{\frac{1}{\alpha-1}} K\left(t, t^{\frac{1}{\alpha-1}} x\right) - (\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}} \mu\left([0, x(\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}})\right) \right|.$$

It is shown in Theorem 1.4 of (3) that

$$\lim_{t \longrightarrow 0+} S_t = 0, \quad a.s. \quad (5.16)$$

Observe that

$$\begin{aligned} f(t, x) &= \sum_{i=0}^{K(t)} \theta_i(t) \mathbf{1}_{\{\theta_i(t) \leq t^{\frac{1}{\alpha-1}} x\}} \\ &= \sum_{j=0}^{n-1} \sum_{i=0}^{K(t)} \theta_i(t) \mathbf{1}_{\{t^{\frac{1}{\alpha-1}} \frac{jx}{n} < \theta_i(t) \leq t^{\frac{1}{\alpha-1}} \frac{(j+1)x}{n}\}}. \end{aligned} \quad (5.17)$$

Then

$$f(t,x) \geq I_1^{(n)} := \sum_{j=0}^{n-1}\sum_{i=0}^{K(t)} t^{\frac{1}{\alpha-1}} \frac{jx}{n} \mathbf{1}_{\{t^{\frac{1}{\alpha-1}} \frac{jx}{n} < \theta_i(t) \leq t^{\frac{1}{\alpha-1}} \frac{(j+1)x}{n}\}}$$

and

$$f(t,x) \leq I_2^{(n)} := \sum_{j=0}^{n-1}\sum_{i=0}^{K(t)} t^{\frac{1}{\alpha-1}} \frac{(j+1)x}{n} \mathbf{1}_{\{t^{\frac{1}{\alpha-1}} \frac{jx}{n} < \theta_i(t) \leq t^{\frac{1}{\alpha-1}} \frac{(j+1)x}{n}\}}.$$

For n fixed and applying (5.16), one gets for $t \to 0+$

$$I_1^{(n)} \xrightarrow{a.s.} \sum_{j=0}^{n-1} \frac{jx}{n} (\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}} \mu\left(\left(\frac{jx}{n}(\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}}, \frac{(j+1)x}{n}(\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}}\right]\right),$$

and

$$I_2^{(n)} \xrightarrow{a.s.} \sum_{j=0}^{n-1} \frac{(j+1)x}{n} (\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}} \mu\left(\left(\frac{jx}{n}(\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}}, \frac{(j+1)x}{n}(\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}}\right]\right).$$

The above two limit values converge to $f(x)$ as n goes to ∞. Then we can conclude. □

It is straightforward to see that

Corollary 5.3. *For any $f \in C_b^0[0,\infty)$ and $c \geq 0$, $M \in \mathbb{R}_+ \cup \{\infty\}$,*

$$\mathbb{E}\left[f(ct^{-\frac{1}{\alpha-1}}\varsigma(t))\mathbf{1}_{\{0 \leq ct^{-\frac{1}{\alpha-1}}\varsigma(t) \leq M\}}|\Theta(t)\right] \xrightarrow{a.s.} \int_0^{Mc^{-1}(\alpha\Gamma(\alpha))^{\frac{1}{\alpha-1}}} f\left(c(\alpha\Gamma(\alpha))^{-\frac{1}{\alpha-1}}y\right) y\mu(dy)$$

when $t \to 0+$.

5.2.3 External branches

We start the proof of Theorem 5.1 with a simpler version.

Proposition 5.4. *Let $\{T_i^{(n)}, 1 \leq i \leq k\}$ and T be as in Theorem 5.1. The following almost sure convergence holds as n goes to ∞:*

$$\mathbb{P}(n^{\alpha-1}T_1^{(n)} > t, n^{\alpha-1}T_2^{(n)} > t, \cdots, n^{\alpha-1}T_k^{(n)} > t|\Theta(n^{1-\alpha}t)) \xrightarrow{a.s.} \mathbb{P}(T > t)^k \qquad (5.18)$$

for any $t \geq 0$. As a consequence,

$$n^{\alpha-1}T_1^{(n)} \xrightarrow{(d)} T. \qquad (5.19)$$

Remark 5.1. *The convergence* (5.19) *has already been obtained in (13) using two different methods.*

Proof. For the sake of simplicity in notations, let $t_n = n^{1-\alpha}t$. Let us build $\Pi^{(n)}(t)$ from $\Theta(t)$ and the paintbox construction (using n particles). We now prove (5.18) for $k = 2$. The proof for $k > 2$ and $k = 1$ follows similarly. Let $\bar{\varsigma}(t_n)$ be an independent copy of $\varsigma(t_n)$, conditionally on $\Theta(t_n)$. Then,

$$\begin{aligned}
&\mathbb{P}(n^{\alpha-1}T_1^{(n)} > t, n^{\alpha-1}T_2^{(n)} > t|\Theta(t_n)) \\
=&\sum_{i,j=1,i\neq j}^{K(t_n)} \theta_i(t_n)\theta_j(t_n)\Big(1-\theta_i(t_n)-\theta_j(t_n)\Big)^{n-2} \\
=&\sum_{i,j=1}^{K(t_n)} \theta_i(t_n)\theta_j(t_n)\Big(1-\theta_i(t_n)-\theta_j(t_n)\Big)^{n-2} - \sum_{i=1}^{K(t_n)} \theta_i(t_n)^2\Big(1-2\theta_i(t_n)\Big)^{n-2} \\
=&\mathbb{E}[\Big(1-\varsigma(t_n)-\bar{\varsigma}(t_n)\Big)^{n-2}|\Theta(t_n)] - \mathbb{E}[\varsigma(t_n)\Big(1-2\varsigma(t_n)\Big)^{n-2}|\Theta(t_n)],
\end{aligned}$$

Using Corollary 5.3, the second term converges almost surely to 0. Let M be a real positive number and write the first term as

$$\mathbb{E}[\Big(1-\varsigma(t_n)-\bar{\varsigma}(t_n)\Big)^{n-2}|\Theta(t_n)] = I_1 + I_2,$$

where

$$I_1 = \mathbb{E}[\Big(1-\varsigma(t_n)-\bar{\varsigma}(t_n)\Big)^{n-2}\mathbf{1}_{\varsigma(t_n)\leq Mn^{-1},\bar{\varsigma}(t_n)\leq Mn^{-1}}|\Theta(t_n)],$$

$$I_2 = \mathbb{E}[\Big(1-\varsigma(t_n)-\bar{\varsigma}(t_n)\Big)^{n-2}\mathbf{1}_{\{\varsigma(t_n)\leq Mn^{-1},\bar{\varsigma}(t_n)\leq Mn^{-1}\}^c}|\Theta(t_n)].$$

By Proposition 5.3,

$$I_2 \leq 1-\mathbb{P}(\bar{\varsigma}(t_n)\leq Mn^{-1}|\Theta(t_n)])^2 \overset{a.s.}{\to} 1-(1-\int_{Mt^{\frac{1}{1-\alpha}}}^{\infty} y\mu(dy))^2.$$

The limit value goes to 0 as M tends to ∞. For I_1, notice that $x \mapsto (1-n^{-1}x)^{n-2}$ converges uniformly to $x \mapsto e^{-x}$ for $0 \leq x \leq 2M$ as n tends to ∞. Then

$$I_1 - \mathbb{E}[\exp(-n\varsigma(t_n)-n\bar{\varsigma}(t_n))\,\mathbf{1}_{\varsigma(t_n)\leq Mn^{-1},\bar{\varsigma}(t_n)\leq Mn^{-1}}|\Theta(t_n)] \overset{a.s.}{\to} 0.$$

Now, thanks to Corollary 5.3, we get

$$\begin{aligned}
&\mathbb{E}[\exp(-n\varsigma(t_n)-n\bar{\varsigma}(t_n))\,\mathbf{1}_{\varsigma(t_n)\leq Mn^{-1},\bar{\varsigma}(t_n)\leq Mn^{-1}}|\Theta(t_n)] \\
=&\mathbb{E}[\exp(-n\varsigma(t_n))\,\mathbf{1}_{\varsigma(t_n)\leq Mn^{-1}} \times \exp(-n\bar{\varsigma}(t_n))\,\mathbf{1}_{\bar{\varsigma}(t_n)\leq Mn^{-1}}|\Theta(t_n)] \\
\overset{a.s.}{\to}&\left(\int_0^{M^{-1}(\frac{\alpha\Gamma(\alpha)}{t})^{\frac{1}{\alpha-1}}} e^{-(\frac{\alpha\Gamma(\alpha)}{t})^{-\frac{1}{\alpha-1}}y}y\mu(dy)\right)^2 \\
\overset{M\to\infty}{\to}&(1+\frac{t}{\alpha\Gamma(\alpha)})^{-2\frac{\alpha}{\alpha-1}}.
\end{aligned}$$

Then we can conclude. □

5.2.4 The block-counting process in small time

Recall that $K^{(n)} = (K^{(n)}(t), t > 0)$ and $K = (K(t), t > 0)$ are respectively the block-counting processes of $\Pi^{(n)}$ and Π.

Lemma 5.1. *Let $t > 0$ and $t_n = n^{1-\alpha}t$. We have*

$$\mathbb{E}[K^{(n)}(t_n)|\Theta(t_n)] = \sum_{i=1}^{K(t_n)} 1 - \Big(1 - \theta_i(t_n)\Big)^n. \tag{5.20}$$

and

$$\begin{aligned}\mathrm{Var}(K^{(n)}(t_n)|\Theta(t_n)) = &\sum_{i=1}^{K(t_n)} \Big(1 - \theta_i(t_n)\Big)^n \Big(1 - (1 - \theta_i(t_n))^n\Big) \\ &+ \sum_{i,j=1, i\neq j}^{K(t_n)} \Big(1 - \theta_i(t_n) - \theta_j(t_n)\Big)^n - \Big(1 - \theta_i(t_n)\Big)^n \Big(1 - \theta_j(t_n)\Big)^n.\end{aligned} \tag{5.21}$$

Furthermore,

$$\frac{\mathbb{E}[K^{(n)}(t_n)|\Theta(t_n)]}{n} \xrightarrow{a.s.} (1 + \frac{t}{\alpha\Gamma(\alpha)})^{-\frac{1}{\alpha-1}}, \, n \to \infty. \tag{5.22}$$

and

$$\frac{Var(K^{(n)}(t_n)|\Theta(t_n))}{n} \xrightarrow{a.s.} (2^{1-\alpha} + \frac{t}{\alpha\Gamma(\alpha)})^{-\frac{1}{\alpha-1}} - (1 + \frac{t}{\alpha\Gamma(\alpha)})^{-\frac{1}{\alpha-1}}, \, n \to \infty. \tag{5.23}$$

Remark 5.2. *It can be deduced from (5.22) and (5.23) that*

$$\frac{K^{(n)}(t_n)}{n(1 + \frac{t}{\alpha\Gamma(\alpha)})^{-\frac{1}{\alpha-1}}} \xrightarrow{P} 1,$$

whereas, interestingly, due to Proposition 5.3 or Theorem 1.1 of (3),

$$\frac{K(t_n)}{n(\frac{t}{\alpha\Gamma(\alpha)})^{-\frac{1}{\alpha-1}}} \xrightarrow{a.s.} 1. \tag{5.24}$$

Proof. The equalities (5.20) and (5.21) come directly from (4.1) and (4.2) in (21). The arguments to prove (5.22) and (5.23) include (5.24) and those used in the proof of Proposition 5.3. To be more clear, we just show the proof of (5.22) and leave the other to the readers.

$$\begin{aligned}\frac{\mathbb{E}[K^{(n)}(t_n)|\Theta(t_n)]}{n} &= \frac{K(t_n)}{n} - \sum_{i=0}^{K(t_n)} n^{-1}(1-\theta_i)^n \\ &\leq \frac{K(t_n)}{n} - \sum_{j=0}^{n-1} n^{-1}(1 - \frac{j+1}{n})^n \Big(K(t_n, \frac{j+1}{n}) - K(t_n, \frac{j}{n})\Big).\end{aligned}$$

In the same way,

$$\frac{\mathbb{E}[K^{(n)}(t_n)|\Theta(t_n)]}{n} \geq \frac{K(t_n)}{n} - \sum_{j=0}^{n-1} n^{-1}(1-\frac{j}{n})^n \Big(K(t_n, \frac{j+1}{n}) - K(t_n, \frac{j}{n}) \Big)$$

While (5.16) shows that

$$\sup_{0\leq j\leq n-1} \left| t_n^{\frac{1}{\alpha-1}} \left(K(t_n, \frac{j+1}{n}) - K(t_n, \frac{j}{n}) \right) - \alpha\Gamma(\alpha)^{\frac{1}{\alpha-1}} \mu\Big(\Big((\frac{\alpha\Gamma(\alpha)}{t})^{\frac{1}{\alpha-1}} \frac{j}{n}, (\frac{\alpha\Gamma(\alpha)}{t})^{\frac{1}{\alpha-1}} \frac{j+1}{n} \Big] \Big) \right|$$
$$= \sup_{0\leq j\leq n-1} \left| n^{-1} t^{\frac{1}{\alpha-1}} \left(K(t_n, \frac{j+1}{n}) - K(t_n, \frac{j}{n}) \right) - \alpha\Gamma(\alpha)^{\frac{1}{\alpha-1}} \mu\Big(\Big((\frac{\alpha\Gamma(\alpha)}{t})^{\frac{1}{\alpha-1}} \frac{j}{n}, (\frac{\alpha\Gamma(\alpha)}{t})^{\frac{1}{\alpha-1}} \frac{j+1}{n} \Big] \Big) \right|$$
$$\leq 2S_{t_n} \xrightarrow{a.s.} 0.$$

Notice also that

$$\lim_{n\to\infty} \sum_{j=0}^{n-1} (1-\frac{j}{n})^n = \lim_{n\to\infty} \sum_{j=0}^{n-1} (1-\frac{j+1}{n})^n = \sum_{j=1}^{\infty} e^{-j} < \infty.$$

Hence using (5.24), almost surely,

$$\frac{\mathbb{E}[K^{(n)}(t_n)|\Theta(t_n)]}{n} \sim \left(\frac{\alpha\Gamma(\alpha)}{t} \right)^{\frac{1}{\alpha-1}} \left(1 - \sum_{j=0}^{n-1} (1-\frac{j}{n})^n \mu\Big(\Big((\frac{\alpha\Gamma(\alpha)}{t})^{\frac{1}{\alpha-1}} \frac{j}{n}, (\frac{\alpha\Gamma(\alpha)}{t})^{\frac{1}{\alpha-1}} \frac{j+1}{n} \Big] \Big) \right)$$
$$\overset{n\to\infty}{\to} \left(\frac{\alpha\Gamma(\alpha)}{t} \right)^{\frac{1}{\alpha-1}} (1 - \int_0^\infty e^{-y\left(\frac{t}{\alpha\Gamma(\alpha)}\right)^{\frac{1}{\alpha-1}}} \mu(dy))$$
$$= \left(1 + \frac{t}{\alpha\Gamma(\alpha)} \right)^{-\frac{1}{\alpha-1}}.$$

□

We are now able to prove our first result.

Proof of Theorem 5.1. We will prove only the version for $k = 2$. For any $0 \leq t_1 \leq t_2$, write

$$\mathbb{P}(n^{\alpha-1}T_1^{(n)} > t_1, n^{\alpha-1}T_2^{(n)} > t_2)$$
$$= \mathbb{P}(n^{\alpha-1}T_1^{(n)} > t_1, n^{\alpha-1}T_2^{(n)} > t_1)\mathbb{P}(n^{\alpha-1}T_2^{(n)} > t_2 | n^{\alpha-1}T_1^{(n)} > t_1, n^{\alpha-1}T_2^{(n)} > t_1).$$

Proposition 5.4 gives that the first term of the above product has limit value

$$\mathbb{P}(T > t_1)^2 = \left(1 + \frac{t_1}{\alpha\Gamma(\alpha)} \right)^{-\frac{2}{\alpha-1}}.$$

Lemma 5.1 implies that conditional on $\{n^{\alpha-1}T_1^{(n)} > t_1, n^{\alpha-1}T_2^{(n)} > t_1\}$, the random variable $\frac{K^{(n)}(n^{1-\alpha}t_1)}{n}$ converges in probability to $(1+\frac{t_1}{\alpha\Gamma(\alpha)})^{-\frac{1}{\alpha-1}}$. For any $j \geq 2$, let $\tilde{T}_1^{(j)}$ be independent of $\Pi^{(n)}$ and have the same law as $T_1^{(j)}$. Using the Markov property of $\Pi^{(n)}$, one obtains

$$
\begin{aligned}
& \mathbb{P}(n^{\alpha-1}T_2^{(n)} > t_2 | n^{\alpha-1}T_1^{(n)} > t_1, n^{\alpha-1}T_2^{(n)} > t_1) \\
= \; & \mathbb{P}(n^{\alpha-1}\tilde{T}_1^{(K^{(n)}(n^{1-\alpha}t_1))} > t_2 - t_1 | n^{\alpha-1}T_1^{(n)} > t_1, n^{\alpha-1}T_2^{(n)} > t_1) \\
\xrightarrow{P} \; & \mathbb{P}\left(T > (t_2-t_1)\left(1+\frac{t_1}{\alpha\Gamma(\alpha)}\right)^{-1}\right) \\
= \; & \left(1+\frac{t_1}{\alpha\Gamma(\alpha)}\right)^{\frac{1}{\alpha-1}}\left(1+\frac{t_2}{\alpha\Gamma(\alpha)}\right)^{-\frac{1}{\alpha-1}}
\end{aligned}
$$

when n tends to ∞. Then we can conclude. □

5.3 The way of coalescing an external branch

5.3.1 The size of the jump

Let us look at the random variable $Q^{(n)}$.

Proof of Theorem 5.2. Assume that at some time t, $K^{(n)}(t) = b$ and $\{1\} \in \Pi^{(n)}(t)$. The coalescence of $\{1\}$ with some other $k-1$ blocks happens at rate

$$\lambda_{1,b,k} := \int_0^1 \binom{b-1}{k-1} x^k (1-x)^{b-k} x^{-2}\Lambda(dx) = \frac{\Gamma(k-\alpha)\Gamma(b-k+\alpha)}{\Gamma(\alpha)\Gamma(2-\alpha)\Gamma(k)\Gamma(b-k+1)}.$$

The total rate at which the singleton $\{1\}$ participates in a coalescence event is

$$
\begin{aligned}
g_{1,b} :&= \int_0^1 \sum_{k=2}^{b} \binom{b-1}{k-1} x^k (1-x)^{b-k} x^{-2}\Lambda(dx) \\
&= \int_0^1 (1-(1-x)^{b-1})x^{-1}\Lambda(dx) \\
&= \int_0^1 (b-1)(1-t)^{b-2}\rho_1(t)dt, \qquad (5.25)
\end{aligned}
$$

where

$$\rho_1(t) = \int_t^1 x^{-1}\Lambda(dx) \sim \frac{t^{1-\alpha}}{(\alpha-1)\Gamma(\alpha)\Gamma(2-\alpha)}$$

when t tends to $0+$. We get, thanks to Stirling's formula,

$$g_{1,b} \sim \frac{b^{\alpha-1}}{(\alpha-1)\Gamma(\alpha)}$$

when b tends to ∞. If the next coalescence after t involves $\{1\}$, then using the strong Markov property of $\Pi^{(n)}$, the probability for $\{1\}$ to coalesce with some other $k-1$ blocks is

$$\frac{\lambda_{1,b,k}}{g_{1,b}} \sim q_k = \frac{\Gamma(k-\alpha)(\alpha-1)}{\Gamma(k)\Gamma(2-\alpha)} \tag{5.26}$$

when b tends to ∞. In this way, if we know the value $K^{(n)}((T_1^{(n)})_-)$, then we can obtain the probability for $\{1\}$ to coalesce with $k-1$ blocks. Notice that $\frac{K^{(n)}((T_1^{(n)})_-)}{n}$ converges in distribution to $(1+\frac{T}{\alpha\Gamma(\alpha)})^{-\frac{1}{\alpha-1}}$ (see Corollary 5.3 of (13) or implicitly from Theorem 5.1 and Lemma 5.1), one can get the following, due to (5.26),

$$\mathbb{P}(Q^{(n)}=k) = \mathbb{E}[\frac{\lambda_{1,K^{(n)}((T_1^{(n)})_-),k}}{g_{1,K^{(n)}((T_1^{(n)})_-)}}] \longrightarrow q_k \tag{5.27}$$

when n tends to ∞.

Notice that $\mathbb{P}(Q^{(n)}=k|K^{(n)}((T_1^{(n)})_-))$ converges to q_k if $K^{(n)}((T_1^{(n)})_-)$ tends to ∞. However, we know that $K^{(n)}((T_1^{(n)})_-)$ tends to ∞ in probability when n goes to ∞. Therefore, $T_1^{(n)}$ and $Q^{(n)}$ are asymptotically independent. Then we can conclude. □

Remark 5.3. *Following the same arguments, Theorem 5.2 is still valid for the more general class of coalescents satisfying the following condition when t tends to 0:*

$$\int_t^1 x^{-2}\Lambda(dx) \sim Ct^{-\alpha}, C>0.$$

Remark 5.4. *We can use similar arguments in the Bolthausen-Sznitman case to get that $\mathbb{P}(Q^{(n)}=k) \to 0$ for any $k \in \mathbb{N}$. The result actually remains true for the more general class where*

$$\int_t^1 x^{-2}\Lambda(dx) \sim Ct^{-1}, C>0.$$

Proof of Proposition 5.2. 1) Recall that $q_k = \frac{\Gamma(k-\alpha)(\alpha-1)}{\Gamma(k)\Gamma(2-\alpha)}$. It then suffices to prove that

$$\lim_{t\to 0+} \frac{\mathbb{P}(\beta(t)=k)}{t} = \frac{\Gamma(k-\alpha)}{\Gamma(k)\Gamma(\alpha)\Gamma(2-\alpha)}.$$

To simplify the notations, let $t_\alpha = \left(\frac{t}{\alpha\Gamma(\alpha)}\right)^{\frac{1}{\alpha-1}}$ and $\rho_x = \mu([x,\infty))$. Recall (5.7) and let $0 < \eta < \frac{2-\alpha}{2}$, then for any $k \geq 2$,

$$\begin{aligned}
\mathbb{P}(\beta(t) = k) &= \frac{t_\alpha^{k-1}}{\Gamma(k)} \int_0^\infty e^{-xt_\alpha} x^k \mu(dx) \\
&= \frac{t_\alpha^{k-1}}{\Gamma(k)} \int_0^\infty \rho_x e^{-xt_\alpha} x^{k-1}(k - xt_\alpha) dx \\
&= I_1 + I_2,
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \frac{t_\alpha^{k-1}}{\Gamma(k)} \int_0^{t_\alpha^{-\eta}} \rho_x e^{-xt_\alpha} x^{k-1}(k - xt_\alpha) dx, \\
I_2 &= \frac{t_\alpha^{k-1}}{\Gamma(k)} \int_{t_\alpha^{-\eta}}^\infty \rho_x e^{-xt_\alpha} x^{k-1}(k - xt_\alpha) dx.
\end{aligned}$$

For t small enough, it is easy to get $I_1 \leq \frac{k}{\Gamma(k)} t_\alpha^{k(1-\eta)-1} = o(t)$. To deal with I_2, recall from Equation (33) of (3) that

$$\rho_x = \mu([x,\infty)) \sim \frac{x^{-\alpha}}{\Gamma(2-\alpha)} \tag{5.28}$$

when x goes to ∞. Notice that $t_\alpha^{-\eta}$ goes to ∞ when t tends to $0+$. Let $0 < \varepsilon < 1$, then for t small enough, we have

$$1 - \varepsilon < \frac{\rho_x}{x^{-\alpha}/\Gamma(2-\alpha)} \leq 1 + \varepsilon, \text{ for all } x \geq t_\alpha^{-\eta}.$$

Since ε can be arbitrarily small, using a change of variable $y = xt_\alpha$, one gets

$$t_\alpha^{1-\alpha} I_2 \to \frac{1}{\Gamma(k)\Gamma(2-\alpha)} \int_0^\infty e^{-y} y^{k-1-\alpha}(k-y) dy = \frac{\alpha\Gamma(k-\alpha)}{\Gamma(k)\Gamma(2-\alpha)}$$

when $t \to 0+$. Then we can obtain (5.9).

2) A simple calculation shows that $\frac{\mathbb{P}(\beta(t)\geq 2)}{t}$ converges to $\frac{1}{(\alpha-1)\Gamma(\alpha)}$ when t tends to $0+$. Hence the first equality of (5.10) holds, using the dominated convergence theorem. For the second equality, the formulas (5.7) and (5.6) imply that

$$\begin{aligned}
\mathbb{E}[e^{-\lambda\beta(t)}] &= \sum_{k\geq 1} e^{-\lambda k} \mathbb{P}(\beta(t) = k) \\
&= e^{-\lambda} \int_0^\infty e^{(e^{-\lambda}-1)xt_\alpha} x\mu(dx) \\
&= e^{-\lambda} \left(1 + (t_\alpha(1-e^{-\lambda}))^{\alpha-1}\right)^{\frac{\alpha}{1-\alpha}}.
\end{aligned} \tag{5.29}$$

Meanwhile, using the same arguments

$$\mathbb{P}(\beta(t)=1)=(1+t_\alpha^{\alpha-1})^{\frac{\alpha}{1-\alpha}}.$$

Then we can obtain (5.10). □

Let us consider the case of coalescents satisfying $\int_0^1 x^{-1}\Lambda(dx)<\infty$.

Proof of Proposition 5.1. In this case, the process $\Pi^{(n)}$ can be constructed using a subordinator. This construction can be found on page 7 of (19) and the original idea is in (28). Let $\nu(dx)=x^{-2}\Lambda(dx)$ and $\tilde{\nu}$ be the push-forward of ν by the transformation $x\to-\ln(1-x)$. Let $(\tilde{S}_t,t\geq 0)$ be a subordinator with Lévy measure $\tilde{\nu}$ and $S_t=e^{-\tilde{S}_t}$. Then $(S_t,t\geq 0)$ is a non-increasing positive pure-jump process with $S_0=1$. Put individuals $1,2,\cdots,n$ uniformly and independently over $(0,1]$. Let t_1 be the first time when $(S_{t_1},S_{t_1-}]$ contains at least one individual, then we set $\Pi^{(n)}(s)=\{\{1\},\{2\},\cdots,\{n\}\}$ for $0\leq s<t_1$. We regroup the individuals located in $(S_{t_1},S_{t_1-}]$ into one block and let $\Pi^{(n)}(t_1)$ be the set of this block and the rest singletons. The block is then put uniformly and independently into $(0,S_{t_1}]$. Find the next time t_2, such that $(S_{t_2},S_{t_2-}]$ contains at least one block (singleton or not) and we regroup all blocks in this interval into one bigger block which will be again put independently into $(0,S_{t_2}]$. In general, at each time t, $\Pi^{(n)}(t)$ is the set of the blocks located in $(0,S_t]$ and also the union of blocks located in $(S_t,S_{t-}]$. This operation can be iterated until reaching the MRCA and the process resulted has the same law as $\Pi^{(n)}$.

Notice that this construction is consistent, i.e., if we add a $n+1$-th individual to get $\Pi^{(n+1)}$, the structure of $\Pi^{(n)}$ is conserved. To see how many blocks will merge with $\{1\}$ in the limit, we assume that $\{1\}$ is put into $(S_{t_1},S_{t_1-}]$ at a certain time. As n goes to ∞, the number of singletons put into the same interval also goes to ∞ with probability 1. According to the construction, all singletons put for the first time into the same interval will be coalesced together. Hence $Q^{(n)}$ converges in probability to infinity. □

5.3.2 Minimal clade size

Let $(s_i^{(n)}(t),1\leq i\leq K^{(n)}(t))$ be the increasing sequence of the smallest elements of blocks of $\Pi^{(n)}(t)$. We have the following lemma.

Lemma 5.2. *For any $t>0$ and $k\in\mathbb{N}$, let $t_n=n^{1-\alpha}t$ and define the event $E_{n,k}=\{s_1^{(n)}(t_n)=1,\cdots,s_k^{(n)}(t_n)=k\}$. Then*

$$\mathbb{P}(E_{n,k}|\Theta(t_n))\xrightarrow{a.s.}1$$

as n tends to ∞.

Proof. We only need to prove that the probability for individuals 1 and 2 to be in the same block of $\Pi^{(n)}(t_n)$ tends to 0. The case of $k \geq 3$ follows in the same way. Let us write this event $\{1 \overset{t_n}{\sim} 2\}$. Then

$$\begin{aligned}\mathbb{P}(1 \overset{t_n}{\sim} 2 | \Theta(t_n)) &= \sum_{i=1}^{K(t_n)} \theta_i(t_n)^2 \\ &= \mathbb{E}[\varsigma(t_n)|\Theta(t_n)] \xrightarrow{a.s.} 0,\end{aligned}$$

where the convergence is due to Corollary 5.3. □

Theorem 5.5. *Let $t > 0$ and $t_n = n^{1-\alpha}t$. For $1 \leq i \leq K^{(n)}(t)$, let $K_i^{(n)}(t)$ be the size of the block containing $s_i^{(n)}(t)$. Then for any $k \in \mathbb{N}$ and $(r_1, \ldots, r_k) \in \mathbb{N}^k$, as $n \to \infty$,*

$$\mathbb{P}(K_1^{(n)}(t_n) = r_1, \cdots, K_k^{(n)}(t_n) = r_k | \Theta(t_n)) \xrightarrow{a.s.} \prod_{i=1}^{k} \mathbb{P}(\beta(t) = r_i) \tag{5.30}$$

where $\beta(t)$ is defined in (5.7).

Proof. Let $t_n = n^{1-\alpha}t$. Define the event $E_{n,r} = \{K_1^{(n)}(t_n) = r_1, \cdots, K_k^{(n)}(t_n) = r_k\}$ and recall $E_{n,k}$ defined in Lemma 5.2. Let $(\varsigma_i(t_n), 1 \leq i \leq k)$ be k independent copies of $\varsigma(t_n)$ which is defined in (5.13), conditional on $\Theta(t_n)$. For $1 \leq i \leq k$, $\varsigma_i(t_n)$ denotes the size of the subinterval into which i is thrown in the paintbox construction of $\Pi^{(n)}(t_n)$ with $n \geq k$. Due to Lemma 5.2, for n large enough we can almost surely approach $\mathbb{P}(E_{n,r}|\Theta(t_n))$ by

$$\mathbb{P}(E_{n,r}|E_{n,k}, \Theta(t_n)) = \mathbb{E}[\binom{n-k}{r_1 - 1, \cdots, r_k - 1} \Pi_{j=1}^k (\varsigma_j(t_n))^{r_j - 1} (1 - \sum_{j=1}^k \varsigma_j(t_n))^{n - \sum_{j=1}^k r_j} | E_{n,k}, \Theta(t_n)].$$

The Lemma 5.2 implies again that the difference

$$\begin{aligned}&\mathbb{E}[\binom{n-k}{r_1 - 1, \cdots, r_k - 1} \Pi_{j=1}^k (\varsigma_j(t_n))^{r_j - 1} (1 - \sum_{j=1}^k \varsigma_j(t_n))^{n - \sum_{j=1}^k r_j} | E_{n,k}, \Theta(t_n)] \\ &- \mathbb{E}[\binom{n-k}{r_1 - 1, \cdots, r_k - 1} \Pi_{j=1}^k (\varsigma_j(t_n))^{r_j - 1} (1 - \sum_{j=1}^k \varsigma_j(t_n))^{n - \sum_{j=1}^k r_j} | \Theta(t_n)]\end{aligned}$$

converges almost surely to 0. We then obtain

$$\begin{aligned}&\mathbb{E}[\binom{n-k}{r_1 - 1, \cdots, r_k - 1} \Pi_{j=1}^k (\varsigma_j(t_n))^{r_j - 1} (1 - \sum_{j=1}^k \varsigma_j(t_n))^{n - \sum_{j=1}^k r_j} | \Theta(t_n)] \\ =&\mathbb{E}[\Pi_{j=1}^k \frac{1}{(r_j - 1)!} (n\varsigma_j(t_n))^{r_j - 1} e^{-n\varsigma_j(t_n)} | \Theta(t_n)] + o_{a.s.}(1),\end{aligned}$$

where $o_{a.s.}(1)$ is a term converging almost surely to 0 when n tends to ∞. The result thus follows from Corollary 5.3. $\square$

We next show how the external branch of 1 is connected to the whole process. Let $\Pi^{(2,n)}$ be the restriction of $\Pi^{(n)}$ from $\mathbb{N}_n$ to $\{2, 3, \cdots, n\}$. By consistency and exchangeability of $\Pi^{(n)}$, $\Pi^{(2,n)}$ has the same law as $\Pi^{(n-1)}$ except for the integer notations. Given $\Pi^{(2,n)}$, one can attach $\{1\}$ to $\Pi^{(2,n)}$ following the *recursive construction* introduced in (13). One thing important is that $n^{\alpha-1}T_1^{(n)}$ and $\Pi^{(2,n)}$ are asymptotically independent. The following lemma is given in the proof of Theorem 5.2 of (13).

Lemma 5.3. *Let $t \geq 0$. As n tends to ∞,*

$$\mathbb{P}(n^{\alpha-1}T_1^{(n)} \geq t | \Pi^{(n,2)}) \xrightarrow{P} (1 + \frac{t}{\alpha\Gamma(\alpha)})^{-\frac{\alpha}{\alpha-1}}. \tag{5.31}$$

Now we are able to deal with the minimal clade size.

Proof of Theorem 5.3. Recall that $Q^{(n)}$ is the number of blocks involved in the coalescence of $\{1\}$. Then $Y^{(n)}$ is just the sum of the $Q^{(n)}$ block sizes (one of these blocks is $\{1\}$). It suffices to determine the size of each block involved in the coalescence event. By exchangeability of the coalescent, the $Q^{(n)}-1$ blocks not being $\{1\}$ will be chosen randomly at time $(T_1^{(n)})_-$. Hence by the strong Markov property of $\Pi^{(n)}$, the joint distribution of the sizes of the randomly chosen $Q^{(n)} - 1$ blocks has the same law as the distribution of $(K_2^{(n)}((T_1^{(n)})_-), \cdots, K_{Q^{(n)}}^{(n)}((T_1^{(n)})_-))$. Hence

$$Y^{(n)} \stackrel{(d)}{=} 1 + \sum_{i=2}^{Q^{(n)}} K_i^{(n)}((T_1^{(n)})_-).$$

Let $K^{(2,n)} = (K^{(2,n)}(t), t \geq 0)$ be the block-counting process of $\Pi^{(2,n)}$ and $(K_1^{(2,n)}(t), K_2^{(2,n)}(t), \cdots, K_{K^{(2,n)}(t)}^{(2,n)}(t$ be the vector of the block sizes of $\Pi^{(2,n)}(t)$, increasingly ordered by their least elements. Notice that if $t < T_1^{(n)}$, $K_i^{(n)}(t) = K_{i-1}^{(2,n)}(t)$ for $1 \leq i \leq K^{(n)}(t)$. Therefore

$$Y^{(n)} \stackrel{(d)}{=} 1 + \sum_{i=1}^{Q^{(n)}-1} K_i^{(2,n)}((T_1^{(n)})_-). \tag{5.32}$$

The formula (5.27) shows that the law of $Q^{(n)}$ is uniquely determined by $K^{(2,n)}((T_1^{(n)})_-)$. As long as $K^{(2,n)}((T_1^{(n)})_-)$ goes to ∞, $Q^{(n)}$ converges in law to a distribution which depends only on α. While (5.22) and (5.19) imply that the variable $K^{(2,n)}((T_1^{(n)})_-)$ goes to ∞ with probability 1. Hence $Q^{(n)}$ is asymptotically independent of $(T_1^{(n)}, K^{(2,n)}((T_1^{(n)})_-))$.

Furthermore, Lemma 5.3 gives that $T_1^{(n)}$ and $K^{(2,n)}((T_1^{(n)})_-)$ are asymptotically independent. In total, $Q^{(n)}$, $T_1^{(n)}$ and $(K^{(2,n)}((T_1^{(n)})_-))$ are all asymptotically independent. In the limit, using Theorem 5.5,

$$Y^{(n)} \xrightarrow{(d)} Y \stackrel{(d)}{=} 1 + \sum_{i=1}^{Q-1} \beta_i(T),$$

where $Q, T, (\beta_i(t))_{i\in\mathbb{N}}$ are all independent and follow respectively the limit laws of $Q^{(n)}, T_1^{(n)}, (K_i^{(2,n)}(t))_{i\in\mathbb{N}}$ for fixed $t \geq 0$. Then we can conclude. □

Proof of Corollary 5.1. Consider the Laplace transform of Y. For any $\lambda > 0$, using (5.29)

$$\begin{aligned}\mathbb{E}[e^{-\lambda Y}] &= e^{-\lambda}\mathbb{E}[(\mathbb{E}[e^{-\lambda\beta(T)}])^{Q-1}] \\ &= e^{-\lambda}\mathbb{E}\Big[\Big(e^{-\lambda}\left(1+(T_\alpha(1-e^{-\lambda}))^{\alpha-1}\right)^{\frac{\alpha}{1-\alpha}}\Big)^{Q-1}\Big]\end{aligned}$$

where $T_\alpha = \left(\frac{T}{\alpha\Gamma(\alpha)}\right)^{\frac{1}{\alpha-1}}$. Denote $\Delta := e^{-\lambda}(1+(T_\alpha(1-e^{-\lambda}))^{\alpha-1})^{\frac{\alpha}{1-\alpha}}$. Using (5.10), one gets

$$\begin{aligned}\mathbb{E}[e^{-\lambda Y}] &= \mathbb{E}[e^{-\lambda}(1-(1-\Delta)^{\alpha-1})] \\ &= I_1 + I_2,\end{aligned}$$

where $I_1 = \mathbb{E}[e^{-\lambda Y}\mathbf{1}_{T_\alpha > \lambda^{-\frac{1}{2}}}]$, $I_2 = \mathbb{E}[e^{-\lambda Y}\mathbf{1}_{T_\alpha \leq \lambda^{-\frac{1}{2}}}]$. The density (5.3) of T implies, when $\lambda \to 0+$

$$I_1 = O(\lambda^{\frac{\alpha}{2}}) = o(\lambda^{(\alpha-1)^2}). \tag{5.33}$$

Notice that there exists $C_{5.34} > 0$ such that for any $0 < \varepsilon < 1$, if λ is small enough, we have

$$|\Delta\mathbf{1}_{T_\alpha\leq\lambda^{-\frac{1}{2}}} - (1+\frac{\alpha}{1-\alpha}(T_\alpha\lambda)^{\alpha-1})\mathbf{1}_{T_\alpha\leq\lambda^{-\frac{1}{2}}}| \leq \varepsilon(T_\alpha\lambda)^{\alpha-1}\mathbf{1}_{T_\alpha\leq\lambda^{-\frac{1}{2}}} + C_{5.34}\lambda. \tag{5.34}$$

Letting $\lambda \to 0+$ and using (5.33), (5.34), one obtains

$$\mathbb{E}[e^{-\lambda Y}] = 1 - (\frac{\alpha}{\alpha-1})^{\alpha-1}\lambda^{(\alpha-1)^2}\mathbb{E}[T_\alpha^{(\alpha-1)^2}] + o(\lambda^{(\alpha-1)^2}).$$

Thanks to Lemma 5.4 of (3) or Theorem 8.1.6 of (5), we get

$$\mathbb{P}(Y > k) \sim \frac{(\frac{\alpha}{\alpha-1})^{\alpha-1}\mathbb{E}[T_\alpha^{(\alpha-1)^2}]}{\Gamma(1-(\alpha-1)^2)}k^{-(\alpha-1)^2} = \frac{\int_0^\infty t^{\alpha-1}f_T(t)dt}{((\alpha-1)\Gamma(\alpha))^{\alpha-1}\Gamma(1-(\alpha-1)^2)}k^{-(\alpha-1)^2}$$

when $k \to \infty$. □

5.4 The largest block

In this section, we aim to prove Theorem 5.4 and Corollary 5.2. We start with a technical lemma.

Lemma 5.4. *Let $k>0$ and X be a random variable distributed according to μ. Define $\mathcal{X}$ such that conditional on X, $\mathcal{X}$ is a Poisson variable with parameter $\frac{X}{k}$. Then for any $x>0$,*

$$\lim_{n\to\infty} n\mathbb{P}(\mathcal{X}\geq xn^{\frac{1}{\alpha}}) = \frac{(kx)^{-\alpha}}{\Gamma(2-\alpha)}.$$

Proof. First of all, let us consider two technical results. Let $M=\lfloor xn^{\frac{1}{\alpha}}\rfloor$.

1) Using Stirling's formula for $M!$ and a change of variable, we get that for any $0<\beta<1$,

$$\begin{aligned}\int_0^{M\beta} e^{-t}\frac{t^M}{M!}dt &= \int_0^{M\beta} e^{M-t}\left(\frac{t}{M}\right)^M (2\pi M)^{-\frac{1}{2}}(1+O(M^{-1}))dt\\ &= \int_0^{\beta} e^{M(1-t+\ln t)}(\frac{M}{2\pi})^{\frac{1}{2}}(1+O(M^{-1}))dt\\ &= O(e^{M(1-\beta+\ln\beta)}M^{\frac{1}{2}}).\end{aligned}\tag{5.35}$$

The last equality is due to the fact that $1-t+\ln t$ is negative and increasing for $t\in(0,1)$.

2) If $\beta>1$, then

$$\int_{M\beta}^{\infty} e^{-t}\frac{t^M}{M!}dt = \int_{\beta}^{\infty} e^{M(1-t+\ln t)}(\frac{M}{2\pi})^{\frac{1}{2}}(1+O(M^{-1}))dt.$$

Notice that $1-t+\ln t$ is strictly decreasing and concave over $[\beta,\infty]$, then there exists a positive number ε such that $1-t+\ln t\leq -\varepsilon t$ for any $t\geq\beta$. Therefore,

$$\int_{M\beta}^{\infty} e^{-t}\frac{t^M}{M!}dt \leq \int_{\beta}^{\infty} e^{-\varepsilon Mt}(\frac{M}{2\pi})^{1/2}(1+O(M^{-1}))dt = O(e^{-\varepsilon M\beta}M^{-1/2}).\tag{5.36}$$

Now we can turn to the study of $\mathcal{X}$. Thanks to successive integrations by parts,

$$\mathbb{P}(\mathcal{X}\geq M+1) = \mathbb{E}[\int_0^{\frac{X}{k}} e^{-t}\frac{t^M}{M!}dt].\tag{5.37}$$

Let $0<\beta_1<1$ and $\beta_2>1$, then we have

$$\mathbb{P}(\mathcal{X}\geq M+1) = I_1+I_2+I_3,$$

where

$$I_1 = \mathbb{E}[\int_0^{\frac{X}{k}} e^{-t}\frac{t^M}{M!}dt\mathbf{1}_{\{X<kM\beta_1\}}],$$

$$I_2 = \mathbb{E}[\int_0^{\frac{X}{k}} e^{-t}\frac{t^M}{M!}dt\mathbf{1}_{\{kM\beta_1 \le X \le kM\beta_2\}}],$$

$$I_3 = \mathbb{E}[\int_0^{\frac{X}{k}} e^{-t}\frac{t^M}{M!}dt\mathbf{1}_{\{X > kM\beta_2\}}].$$

Now let n tend to infinity. By (5.35), we get

$$0 \le nI_1 \le n\mathbb{P}(X < kM\beta_1)\int_0^{M\beta_1} e^{-t}\frac{t^M}{M!}dt \longrightarrow 0. \tag{5.38}$$

It is easy to verify that $\int_0^\infty e^{-t}\frac{t^M}{M!}dt = 1$ for any integer $M \ge 0$. Then using together (5.28) and (5.36), we obtain

$$\lim_{n\to\infty} nI_3 = \lim_{n\to\infty} n\mathbb{P}(X > kM\beta_2) = \frac{(kx\beta_2)^{-\alpha}}{\Gamma(2-\alpha)}. \tag{5.39}$$

In the same way, we have

$$0 \le nI_2 \le n\mathbb{P}(kM\beta_1 \le X \le kM\beta_2) \longrightarrow \frac{(kx\beta_1)^{-\alpha}}{\Gamma(2-\alpha)} - \frac{(kx\beta_2)^{-\alpha}}{\Gamma(2-\alpha)},\ n\to\infty. \tag{5.40}$$

If β_1 and β_2 are close enough to 1, nI_2 can be bounded by an arbitrarily small positive number for n large enough. Combining (5.38), (5.39) and (5.40), we conclude this lemma. □

To prove Theorem 5.4, we will use classical relations between Beta-coalescents and continuous-state branching processes (CSBPs) developed in (6) (see also Section 2 of (3)). We give a short summary to provide a minimal set of tools. A continuous-state branching process $(Z(t), t \ge 0)$ is a $[0,\infty]$-valued Markov process (in continuous time) whose transition functions $p_t(x,\cdot)$ satisfy the branching property

$$p_t(x+y,\cdot) = p_t(x,\cdot) * p_t(y,\cdot), \quad \text{for all } x, y \ge 0.$$

For each $t \ge 0$, there exists a function $u_t : [0,\infty) \to \mathbb{R}$ such that

$$\mathbb{E}[e^{-\lambda Z(t)}|Z(0) = a] = e^{-au_t(\lambda)}. \tag{5.41}$$

If almost surely, the process has no instantaneous jump to infinity, the function u_t satisfies the following differential equation

$$\frac{\partial u_t(\lambda)}{\partial t} = -\Psi(u_t(\lambda)),$$

where $\Psi : [0,\infty) \longrightarrow \mathbb{R}$ is a function of the form

$$\Psi(u) = \gamma u + \beta u^2 + \int_0^\infty (e^{-xu} - 1 + xu\mathbf{1}_{\{x\le 1\}})\pi(dx),$$

where $\gamma \in \mathbb{R}, \beta \geq 0$ and π is a Lévy measure on $(0, \infty)$ satisfying $\int_0^\infty (1 \wedge x^2)\pi(dx) < \infty$. The function Ψ is called the branching mechanism of the CSBP.

As explained in (4), a CSBP can be extended to a two-parameter random process $(Z(t,a), t \geq 0, a \geq 0)$ with $Z(0,a) = a$. For fixed t, $(Z(t,a), a \geq 0)$ turns out to be a subordinator with Laplace exponent $\lambda \mapsto u_t(\lambda)$ thanks to (5.41).

There exists a measure-valued process $(M_t, t \geq 0)$ taking values in the set of finite measures on $[0,1]$ which characterizes $(Z(t,a), t \geq 0, 0 \leq a \leq 1)$. More precisely, $(M_t([0,a]), t \geq 0, 0 \leq a \leq 1)$ has the same finite-dimensional distributions as $(Z(t,a), t \geq 0, 0 \leq a \leq 1)$. Hence $(M_t([0,a]), 0 \leq a \leq 1)$ is a subordinator with Laplace exponent $\lambda \mapsto u_t(\lambda)$ and $Z(t) = M_t([0,1])$ is a CSBP with branching mechanism Ψ started at $M_0([0,1]) = 1$. In particular, if the branching mechanism is $\Psi(\lambda) = \lambda^\alpha$ and hence Lévy measure is given by $\pi(dx) = \frac{\alpha(\alpha-1)}{\Gamma(2-\alpha)} x^{-1-\alpha} dx$, for all $t > 0$, M_t consists only of finite number of atoms. For the construction of $(M_t([0,a]), t \geq 0, 0 \leq a \leq 1)$, we refer to (2; 6; 15).

A deep relation has been revealed in (6) between the Beta-coalescent and the CSBP with branching mechanism $\Psi(\lambda) = \lambda^\alpha$ and Lévy measure $\pi(dx) = \frac{\alpha(\alpha-1)}{\Gamma(2-\alpha)} x^{-1-\alpha} dx$. The relationship is described by the following two lemmas which are respectively Lemma 2.1 and 2.2 of (3) and will be important in the sequel.

To save notations, from now on, $(Z(t), t \geq 0)$ always denotes a continuous state branching process $(Z(t,1), t \geq 0)$.

Lemma 5.5. *Assume $(Z(t), t \geq 0)$ is a CSBP with branching mechanism $\Psi(\lambda) = \lambda^\alpha$ and let $(M_t, t \geq 0)$ be its associated measure-valued process. If $(\Pi(t), t \geq 0)$ is a Beta-coalescent and $(\Theta(t), t \geq 0)$ is the associated ranked coalescent, then for all $t > 0$, the distribution of $\Theta(t)$ is the same as the distribution of the sizes of the atoms of the measure $\frac{M_{R^{-1}(t)}}{Z(R^{-1}(t))}$, ranked in decreasing order. Here $R(t) = (\alpha-1)\alpha\Gamma(\alpha)\int_0^t Z(s)^{1-\alpha} ds$ and $R^{-1}(t) = \inf\{s : R(s) > t\}$.*

Lemma 5.6. *Assume $\Psi(\lambda) = \lambda^\alpha$. For any $t \geq 0$, let $D(t)$ be the number of atoms of M_t, and let $J(t) = (J_1(t), \cdots, J_{D(t)}(t))$ be the sizes of the atoms of M_t, ranked in decreasing order. Then $D(t)$ is Poisson with mean $\gamma_t = ((\alpha-1)t)^{-\frac{1}{\alpha-1}}$. Moreover, conditional on $D(t) = k$, the distribution of $J(t)$ is the same as the distribution of $(\gamma_t^{-1} X_1, \cdots, \gamma_t^{-1} X_k)$ where $X_1, \cdots, X_k$ are obtained by picking k i.i.d. random variables with distribution μ and then ranking them in decreasing order.*

Remark 5.5. *From the relation between $(M_t, t \geq 0)$ and $(Z(t,a), t \geq 0, 0 \leq a \leq 1)$ and also the fact that for all $t > 0$, M_t consists only of finite number of atoms (the number is actually $D(t)$), for a given $t > 0$, there exist $0 \leq a_1, \cdots, a_{D(t)} \leq 1$ such that $\{Z(t,a_1) - Z(t,a_1-), \cdots, Z(t,a_{D(t)}) - Z(t,a_{D(t)}-)\}$ are exactly the values of the atoms of M_t. By the strong Markov property of $(Z(t,a), t \geq 0, 0 \leq a \leq 1)$, for $s \geq t$, the jumping at s can only happen at the points $\{(s,a_1), \cdots, (s,a_{D(t)})\}$. Therefore, $D(t)$ decreases on t.*

The idea of the proof of Theorem 5.4 is as follows: Let $t_n = n^{1-\alpha}t$. Lemma 5.5 shows that $\Theta(t_n)$ has the same law as $\frac{M_{R^{-1}(t_n)}}{Z(R^{-1}(t_n))}$. Moreover it is proved in Lemma 4.2 of (3) that $\frac{R^{-1}(t_n)}{t_n} \xrightarrow{P} \frac{1}{(\alpha-1)\alpha\Gamma(\alpha)}$, as n goes to ∞. Hence one can compare the block sizes at time t_n to those at time $R^{-1}((\alpha-1)\alpha\Gamma(\alpha)t_n)$. To this, we use the paintbox construction and the closeness between the measures $\frac{M_{t_n}}{Z(t_n)}$ and $\frac{M_{R^{-1}((\alpha-1)\alpha\Gamma(\alpha)t_n)}}{Z(R^{-1}((\alpha-1)\alpha\Gamma(\alpha)t_n))}$. This idea can be executed through two steps.

1) Analysis of the largest block size at time t_n with the measure $\frac{M_{t_n}}{Z(t_n)}$: If $D(t_n) \neq 0$, let $\bar{J}_i(t_n) = \frac{J_i(t_n)}{Z(t_n)}$ for $1 \leq i \leq D(t_n)$. Let $\{d_1(t_n), \cdots, d_{D(t_n)}(t_n)\}$ be an interval partition of $[0,1]$ such that the Lebesgue measure of $d_i(t_n)$ is $\bar{J}_i(t_n)$. Build a partition of $\mathbb{N}_n$ from a paintbox associated with $\{d_1(t_n), \cdots, d_{D(t_n)}(t_n)\}$. Let N_i be the number of integers in $d_i(t_n)$ and $N = \max\{N_i : 1 \leq i \leq D(t_n)\}$.

Lemma 5.7. *Let $x > 0$. Then*

1)

$$\lim_{n\to\infty} \mathbb{P}(N \leq xn^{1/\alpha}) = \exp(-\frac{(\alpha-1)tx^{-\alpha}}{\Gamma(2-\alpha)}).$$

2) *Let $0 < y < x$. Then*

$$\lim_{n\to\infty} \mathbb{P}(\exists i : J_i(t_n) < n^{\frac{1-\alpha}{\alpha}}y, N_i \geq xn^{\frac{1}{\alpha}}) = 0. \tag{5.42}$$

Proof. 1) It is well known that if we throw a Poisson number of parameter $nZ(t_n)$ on $[0,1]$, the number of integers in $d_i(t_n)$, denoted by $\mathcal{N}_i$, is a Poisson variable of parameter $nJ_i(t_n)$. Conditional on all $J_i(t_n)$'s, all $\mathcal{N}_i$'s are independent. Let $\mathcal{N}$ be the maximum of all $\mathcal{N}_i$'s. Then, using Lemmas 5.4 and 5.6,

$$\mathbb{P}(\mathcal{N} \leq xn^{1/\alpha}) = \mathbb{E}[\Pi_{i=1}^{D(t_n)}\mathbb{P}(\mathcal{N}_i \leq xn^{1/\alpha})] \longrightarrow \exp(-\gamma_t^{1-\alpha}\frac{x^{-\alpha}}{\Gamma(2-\alpha)}) = \exp(-\frac{(\alpha-1)tx^{-\alpha}}{\Gamma(2-\alpha)}),\ n\to\infty.$$

Lemma 5.6 implies that $Z(t_n)$ tends in probability to 1 as n goes to infinity. Hence N and $\mathcal{N}$ are close in the limit and standard comparison techniques allows to conclude.

2) As $Z(t_n)$ converges to 1, (5.42) is equivalent to

$$\lim_{n\to\infty} \mathbb{P}(\exists i : J_i(t_n) < n^{\frac{1-\alpha}{\alpha}}y, \mathcal{N}_i \geq xn^{\frac{1}{\alpha}}) = 0.$$

Let $\tilde{\mathcal{N}} = \max\{\mathcal{N}_i : J_i(t_n) < n^{\frac{1-\alpha}{\alpha}}y\}$. It is necessary and sufficient to show that $\lim_{n\to\infty} \mathbb{P}(\tilde{\mathcal{N}} \geq xn^{\frac{1}{\alpha}}) = 0$. Notice that conditional on $J_i(t_n)$, $\mathcal{N}_i$ is a Posson variable with paprameter $nJ_i(t_n)$. Let $\{P_1(yn^{\frac{1}{\alpha}}), P_2(yn^{\frac{1}{\alpha}}), \cdots\}$ be a sequence of i.i.d. Poisson variables with parameter $yn^{\frac{1}{\alpha}}$ and also independent of $D(t_n)$. Then

$$\mathbb{P}(\tilde{\mathcal{N}} \geq xn^{\frac{1}{\alpha}}) \leq \mathbb{P}\left(\max\{P_i(yn^{\frac{1}{\alpha}}) : 1 \leq i \leq D(t_n)\} \geq xn^{\frac{1}{\alpha}}\right) = 1-\mathbb{E}[(\mathbb{P}(P_1(yn^{\frac{1}{\alpha}}) < xn^{\frac{1}{\alpha}}))^{D(t_n)}].$$

Using (5.37) and (5.35), one gets

$$\mathbb{P}(P_1(yn^{\frac{1}{\alpha}}) < xn^{\frac{1}{\alpha}}) = 1 - o(\frac{1}{n}).$$

Meanwhile, Lemma 5.6 tells that $\frac{D(t_n)}{n}$ converges in robability to γ_t as n goes to infinity. Hence we can conclude. □

Remark 5.6. *The key point to prove (5.42) is that $Z(t_n)$ converges to 1 in probability, $\frac{D(t_n)}{n}$ is asymptotically bounded by a positive value from above. The distribution of $\{J_i(t_n)\}_{1\leq i\leq D(t_n)}$ is not necessary to know. One can still find (5.42) true if we replace t_n by a random time and conditions for $Z(t_n)$ and $D(t_n)$ are satisfied at the same time.*

2) A tool lemma for the transfer from $\frac{M_{t_n}}{Z(t_n)}$ to $\frac{M_{R^{-1}((\alpha-1)\alpha\Gamma(\alpha)t_n)}}{Z(R^{-1}((\alpha-1)\alpha\Gamma(\alpha)t_n))}$: Let $(A_1, \cdots, A_k)$ and $(B_1, \cdots, B_k)$ be two partitions of $[0, 1]$ with $k \geq 1$. We throw away n particles uniformly and independently on $[0, 1]$ and regroup those within the same intervals of $(B_1, \cdots, B_k)$, which gives a sequence of k numbers $(N_{B_1}, \cdots, N_{B_k})$ such that N_{B_i} is the number of particles located in B_i. We can obtain the law of this sequence in another way using $(A_1, \cdots, A_k)$: We throw n particles uniformly and independently on $[0, 1]$. Let $I := \{i : 1 \leq i \leq n, l(A_i) \leq l(B_i)\}$, where $l(\cdot)$ denotes the Lebesgue measure. If a particle falls in A_i where $i \in I$, then move this particle to B_i. If a particle falls in A_i where $i \in I^c$, then do the following: we attach to this particle an independent Bernoulli variable with parameter $\frac{l(B_i)}{l(A_i)}$. If the Bernoulli variable gives 1, then the particle is put into B_i. Otherwise, this particle will be put into B_j for $j \in I$ with probability

$$\frac{l(B_j) - l(A_j)}{\sum_{j\in I}(l(B_j) - l(A_j))}. \tag{5.43}$$

We denote by $N^B_{A_i}$ the new amount of particles in B_i. We have the the following result.

Lemma 5.8. *The following identity in law holds.*

$$(N^B_{A_1}, \cdots, N^B_{A_k}) \overset{(d)}{=} (N_{B_1}, \cdots, N_{B_k}).$$

Proof. Notice that only the measure of each element of $(A_1, \cdots, A_k)$ and $(B_1, \cdots, B_k)$ matters, one can always assume that $[0, 1]$ is divided in a way that A_i is contained in B_i for $i \in I$ and B_i is contained in A_i for $i \in I^c$. So if a particle is located in A_i for $i \in I$, it is also located in B_i. But if a particle is located in A_i for $i \in I^c$, with probability $\frac{l(B_i)}{l(A_i)}$ it is located in B_i. Assume that this particle is not located in B_i, then it must be in $\cup_{j\in I}B_j/A_j$. Using the uniformity of the throw, this particle falls in B_j with probability (5.43). □

The above two steps allow to start the proof of Theorem 5.4. But before that, let us just recall some technical results from (3). Let $\varepsilon > 0$, $t > 0$ and $t_n = n^{1-\alpha}t$. Let $t_- = (1-\varepsilon)t_n$, $t_+ = (1+\varepsilon)t_n$ and $t_* = (\alpha-1)\alpha\Gamma(\alpha)t_n$. Define the event $B_{1,t} := \{t_- \leq R^{-1}(t_*) \leq t_+\}$. It can be found in Lemma 4.2 of (3) that there exists a constant $C_{5.44}$ such that

$$\mathbb{P}(B_{1,t}) \geq 1 - C_{5.44}t_*\varepsilon^{-\alpha}. \tag{5.44}$$

Also from Lemma 5.1 of (3), there exists a constant $C_{5.45}$ such that for all $a > 0$, $t > 0$ and $\eta > 0$,

$$q(a,t,\eta) = \mathbb{P}(\sup_{0\leq s\leq t} |Z(s,a) - a| > \eta) \leq C_{5.45}(a+\eta)t\eta^{-\alpha}. \tag{5.45}$$

Thus, if we define $B_{2,t} := \{1 - n^{\frac{1-\alpha}{2\alpha}} \leq Z(s) \leq 1 + n^{\frac{1-\alpha}{2\alpha}},\ \forall s \in [t_-, t_+]\}$, we can obtain that

$$\mathbb{P}(B_{2,t}) \geq 1 - C_{5.45}t(1+\varepsilon)(1+n^{\frac{1-\alpha}{2\alpha}})n^{\frac{1-\alpha}{2}}. \tag{5.46}$$

Proof of Theorem 5.4. Lemma 5.5 tells us that for any $s \geq 0$, we have

$$\frac{M_{R^{-1}(s)}}{Z(R^{-1}(s))} \overset{(d)}{=} \Theta(s).$$

Let π be the partition of $\mathbb{N}_n$ obtained from a paintbox associated with $\frac{M_{R^{-1}(s)}}{Z(R^{-1}(s))}$. Then $\pi \overset{(d)}{=} \Pi^{(n)}(s)$. If $R^{-1}(s) \geq t_-$, we can as well at first build a partition from a paintbox associated with $\frac{M_{t_-}}{Z(t_-)}$ and then use Lemma 5.8 to get π. This kind of construction is the key of this proof.

For $s \geq t_-$, one builds a partition of $\mathbb{N}_n$ from a paintbox associated with $(\frac{m_i(s)}{Z(s)}, 1 \leq i \leq D(t_-))$. We denote this partition by $V^{(n)}(s) = (V_1(s), V_2(s), \cdots, V_{D(t_-)}(s))$. Let $I_i^{(n)}(s)$ be the number of particles in $V_i^{(n)}(s)$.

For $s \geq t_-$, let $M^{(n)}(s) = \sup\{I_i^{(n)}(s), 1 \leq i \leq D(t_-)\}$ be the size of the largest block of $V^{(n)}(s)$. Let $x > 0$ and $B_{3,t} = \{\exists k : I_k^n(t_-) \geq xn^{\frac{1}{\alpha}}, J_k(t_-) \geq n^{\frac{2(1-\alpha)}{\alpha}}, \sup_{t_-\leq s\leq t_+} |m_k(s) - J_k(t_-)| \leq \varepsilon J_k(t_-)\}$.

On the event $B_{3,t}$, we have that $M^{(n)}(t_-) \geq xn^{\frac{1}{\alpha}}$. Conditional on $B_{1,t}$ we can build the partition $V^{(n)}(R^{-1}(t_*))$ from a paintbox associated to the partition $Z(t_-)^{-1}(J_1(t_-), \ldots, J_{D(t_-)}(t_-))$ and Lemma 5.8. Let $B(m,p)$ be a binomial variable with parameters $m \geq 2$ and $0 \leq p \leq 1$.

Lemma 5.8 implies

$$\begin{aligned}
&\mathbb{P}\left(M^{(n)}(R^{-1}(t_*)) \geq (1-2\varepsilon)xn^{\frac{1}{\alpha}}|B_{1,t}\cap B_{2,t}\cap B_{3,t}\right)\\
\geq&\mathbb{P}\left(B\left(\lceil xn^{\frac{1}{\alpha}}\rceil, \frac{m_k(R^{-1}(t_*))Z(t_-)}{J_k(t_-)Z(R^{-1}(t_*))}\wedge 1\right)\geq (1-2\varepsilon)xn^{\frac{1}{\alpha}}|B_{1,t}\cap B_{2,t}\cap B_{3,t}\right)\\
\geq&\mathbb{P}\left(B\left(\lceil xn^{\frac{1}{\alpha}}\rceil, (1-\varepsilon)\frac{1-n^{\frac{1-\alpha}{2\alpha}}}{1+n^{\frac{1-\alpha}{2\alpha}}}\right)\geq (1-2\varepsilon)xn^{\frac{1}{\alpha}}\right)\\
=&\mathbb{P}\left((xn^{\frac{1}{\alpha}})^{-1}B\left(\lceil xn^{\frac{1}{\alpha}}\rceil, (1-\varepsilon)\frac{1-n^{\frac{1-\alpha}{2\alpha}}}{1+n^{\frac{1-\alpha}{2\alpha}}}\right)\geq (1-\varepsilon)-\varepsilon\right).
\end{aligned}$$

A law of large numbers argument implies that

$$\mathbb{P}\left(M^{(n)}(R^{-1}(t_*)) \geq (1-2\varepsilon)xn^{\frac{1}{\alpha}}|B_{1,t}\cap B_{2,t}\cap B_{3,t}\right)\geq 1-\varepsilon \tag{5.47}$$

for n large enough. Now observe that

$$\begin{aligned}
\mathbb{P}(B_{3,t}) &= \mathbb{P}(\exists k : I_k^n(t_-)\geq xn^{\frac{1}{\alpha}}, J_k(t_-)\geq n^{\frac{2(1-\alpha)}{\alpha}})\\
&\quad\times \mathbb{P}(\sup_{t_-\leq s\leq t_+}|m_k(s)-J_k(t_-)|\leq \varepsilon J_k(t_-)|\{\exists k : I_k^n(t_-)\geq xn^{\frac{1}{\alpha}}, J_k(t_-)\geq n^{\frac{2(1-\alpha)}{\alpha}})\\
&= \mathbb{P}(\exists k : I_k^n(t_-)\geq xn^{\frac{1}{\alpha}}, J_k(t_-)\geq n^{\frac{2(1-\alpha)}{\alpha}})\\
&\quad\times (1-\mathbb{E}[q(J_k(t_-), t_+-t_-, \varepsilon J_k(t_-))|\exists k : I_k^n(t_-)\geq xn^{\frac{1}{\alpha}}, J_k(t_-)\geq n^{\frac{2(1-\alpha)}{\alpha}})\\
&\geq \mathbb{P}(\exists k : I_k^n(t_-)\geq xn^{\frac{1}{\alpha}}, J_k(t_-)\geq n^{\frac{2(1-\alpha)}{\alpha}})(1-2tC_{5.45}n^{\frac{(1-\alpha)(2-\alpha)}{\alpha}}(1+\varepsilon)\varepsilon^{1-\alpha}).
\end{aligned}$$

Using Lemma 5.7,

$$\begin{aligned}
\mathbb{P}(\exists k : I_k^n(t_-)\geq xn^{\frac{1}{\alpha}}, J_k(t_-)\geq n^{\frac{2(1-\alpha)}{\alpha}}) &\sim \mathbb{P}(\exists k : I_k^n(t_-)\geq xn^{\frac{1}{\alpha}}) = \mathbb{P}(M^{(n)}(t_-)\geq n^{\frac{1}{\alpha}}x)\\
&\sim 1-\exp(-(1-\varepsilon)\frac{(\alpha-1)tx^{-\alpha}}{\Gamma(2-\alpha)}).
\end{aligned}$$

In consequence,

$$\liminf_{n\to\infty}\mathbb{P}(B_{3,t})\geq 1-\exp(-(1-\varepsilon)\frac{(\alpha-1)tx^{-\alpha}}{\Gamma(2-\alpha)})$$

when n tends to ∞. Then, thanks to (5.44) and (5.46), we deduce that

$$\liminf_{n\to\infty}\mathbb{P}(B_{1,t}\cap B_{2,t}\cap B_{3,t})\geq 1-\exp(-(1-\varepsilon)\frac{(\alpha-1)tx^{-\alpha}}{\Gamma(2-\alpha)}).$$

Combining the latter with (5.47), we obtain

$$\liminf_{n\to\infty} \mathbb{P}\left(M^{(n)}(R^{-1}(t_*)) \geq (1-2\varepsilon)xn^{\frac{1}{\alpha}}\right) \geq 1 - \exp(-(1-\varepsilon)\frac{(\alpha-1)tx^{-\alpha}}{\Gamma(2-\alpha)}). \tag{5.48}$$

Next, we seek to find an upper bound for $\mathbb{P}\left(M^{(n)}(R^{-1}(t_*)) \geq xn^{\frac{1}{\alpha}}\right)$. Conditional on $B_{1,t}$, we construct $V^{(n)}(t_+)$ from $V^{(n)}(R^{-1}(t_*))$ using the method in Lemma 5.8. Let

$$B_{4,t} = B_{1,t} \cap \{\exists k : I_k^{(n)}(R^{-1}(t_*)) \geq xn^{\frac{1}{\alpha}}, m_k(R^{-1}(t_*)) \geq n^{\frac{2(1-\alpha)}{\alpha}}, \sup_{R^{-1}(t_*)\leq s\leq t_+} \frac{|m_k(s) - m_k(R^{-1}(t_*))|}{m_k(R^{-1}(t_*))} \leq \varepsilon\}.$$

Similarly as for the lower bound,

$$\begin{aligned}
&\mathbb{P}\left(M^{(n)}(t_+) \geq (1-2\varepsilon)xn^{\frac{1}{\alpha}} | B_{2,t} \cap B_{4,t}\right) \\
\geq &\mathbb{P}\left(B\left(\lceil xn^{\frac{1}{\alpha}}\rceil, \frac{Z(R^{-1}(t_*))m_k(t_+)}{Z(t_+)m_k(R^{-1}(t_*))} \wedge 1\right) \geq (1-2\varepsilon)xn^{\frac{1}{\alpha}} | B_{2,t} \cap B_{4,t}\right) \\
\geq &\mathbb{P}\left(B\left(\lceil xn^{\frac{1}{\alpha}}\rceil, (1-\varepsilon)\frac{1-n^{(1-\alpha)/\alpha}}{1+n^{(1-\alpha)/\alpha}}\right) \geq (1-2\varepsilon)xn^{\frac{1}{\alpha}}\right) \longrightarrow 1.
\end{aligned} \tag{5.49}$$

Using the strong Markov property of the CSBP and (5.45), we have

$$\mathbb{P}(B_{4,t}) = \mathbb{P}(B_{1,t} \cap \{\exists k : I_k^{(n)}(R^{-1}(t_*)) \geq xn^{\frac{1}{\alpha}}, m_k(R^{-1}(t_*)) \geq n^{\frac{2(1-\alpha)}{\alpha}}\}) \tag{5.50}$$

$$\times (1 - 2tC_{5.45}n^{\frac{(1-\alpha)(2-\alpha)}{\alpha}}(1+\varepsilon)\varepsilon^{1-\alpha}) \tag{5.51}$$

Notice that using (5.45), in the sense of convergence of probability

$$\lim_{n\to\infty} \sup_{t_-\leq s\leq t_+} Z(s) = \lim_{n\to\infty} \inf_{t_-\leq s\leq t_+} Z(s) = 1$$

Together with (5.44), we get the following convergence in probability

$$\lim_{n\to\infty} Z(R^{-1}(t_*)) = 1.$$

Remark 5.5 tells that $D(t)$ is decreasing on t. Under $B_{1,t}$, $D(t_-) \leq D(R^{-1}(t_*)) \leq D(t_+)$. It is then easy to deduce that $\frac{D(R^{-1}(t_*))}{n}$ is asymptotically bounded from above by a certain positive number. Now we can apply Remark 5.6 and get

$$\mathbb{P}(B_{4,t}) = \mathbb{P}(\exists k : I_k^{(n)}(R^{-1}(t_*)) \geq xn^{\frac{1}{\alpha}}) + o(1) = \mathbb{P}(M^{(n)}(R^{-1}(t_*)) \geq xn^{\frac{1}{\alpha}}) + o(1). \tag{5.52}$$

Using (5.49), (5.46) and (5.52)

$$\begin{aligned}
&\limsup_{n\longrightarrow\infty} \mathbb{P}(M^{(n)}(R^{-1}(t_*)) \geq xn^{\frac{1}{\alpha}}) \\
\leq &\lim_{n\longrightarrow\infty} \mathbb{P}(M^{(n)}(t_+) \geq (1-2\varepsilon)xn^{\frac{1}{\alpha}}) \\
= &1 - \exp(-(x(1-2\varepsilon))^{-\alpha}\frac{(\alpha-1)t(1+\varepsilon)}{\Gamma(2-\alpha)}).
\end{aligned} \tag{5.53}$$

Since ε can be arbitrarily small, (5.48) and (5.53) allow to conclude. □

Finally, observe that Corollary 5.2 is obtained from a combination of Lemma 5.3 and Theorem 5.4.

Bibliographie

[1] E. Árnason. Mitochondrial cytochrome b dna variation in the high-fecundity atlantic cod : trans-atlantic clines and shallow gene genealogy. *Genetics*, 166(4) :1871–1885, 2004.

[2] J. Berestycki, N. Berestycki and J. Schweinsberg. Beta-coalescents and continuous stable random trees. *Ann. Probab.*, 35(5) :1835–1887, 2007.

[3] J. Berestycki, N. Berestycki and J. Schweinsberg. Small-time behavior of Beta-coalescents. *Ann. Inst. H. Poincaré Probab. Stat.*, 44(2) :214–238, 2008.

[4] J. Bertoin and J-F. Le Gall. The Bolthausen-Sznitman coalescent and the genealogy of continuous-state branching processes. *Probab. Theory Related Fields*, 117(2) :249–266, 2000.

[5] N.H. Bingham, C.M. Goldie and J.L. Teugels. Regular variation- *Encyclopedia of Mathematics and its Applications*, 27. Cambridge University Press, Cambridge. MR898871. 1987.

[6] M. Birkner, J. Blath, M. Capaldo, A. M. Etheridge, M. Möhle, J. Schweinsberg and A. Wakolbinger. Alpha-stable branching and Beta-coalescents. *Electron. J. Probab.*, 10(9) :303–325, 2005.

[7] M.G.B. Blum and O. François. Minimal clade size and external branch length under the neutral coalescent. *Adv. in Appl. Probab.*, 37(3) :647–662, 2005.

[8] E. Bolthausen and A. Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Commun. Math. Phys*, 197(2) :247–276, 1998.

[9] J. Boom, E. Boulding and A. Beckenbach. Mitochondrial dna variation in introduced populations of pacific oyster, crassostrea gigas, in british columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(7) :1608–1614, 1994.

[10] A. Caliebe, R. Neininger, M. Krawczak and U. Rösler. On the length distribution of external branches in coalescence trees : genetic diversity within species. *Theoret. Population Biol.*, 72(2) :245–252, 2007.

[11] I. Dahmer, G. Kersting and A. Wakolbinger. The total external branch length of Beta-coalescents. *arXiv preprint arXiv :1212.6070*. To appear in *Combinatorics, Probability and Computing.*

[12] J-F. Delmas, J-S. Dhersin and A. Siri-Jégousse. Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.*, 18(3) :997–1025, 2008.

[13] J-S. Dhersin, F. Freund, A. Siri-Jégousse and L. Yuan. On the length of an external branch in the beta-coalescent. *Stoch. Proc. Appl.*, 123 :1691–1715, 2013.

[14] J-S. Dhersin and M. Möhle. On the external branches of coalescents with multiple collisions. *Electron. J. Probab*, 18(40) :1–11, 2013.

[15] P. Donnelly and T.G. Kurtz. Particle representations for measure-valued population models. *Ann. Probab.*, 27(1) :166–205, 1999.

[16] B. Eldon and J. Wakeley. Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics*, 172 :2621–2633, 2006.

[17] F. Freund and M. Möhle. On the time back to the most recent common ancestor and the external branch length of the Bolthausen-Sznitman coalescent. *Markov Process. Related Fields*, 15(3) :387–416, 2009.

[18] F. Freund and A. Siri-Jégousse. Minimal clade size in the Bolthausen-Sznitman coalescent. To appear in *J. Appl. Probab.*.

[19] A. Gnedin, A. Iksanov and M. Möhle. On asymptotics of exchangeable coalescents with multiple collisions. *J. Appl. Probab.*, 45 :1186–1195, 2008.

[20] D. Hedgecock. Does variance in reproductive success limit effective population sizes of marine organisms ? In *Genetics and Evolution of Aquatic Organisms*, Chapman and Hall, London, 1222–1344, 1994.

[21] H-K. Hwang and S. Janson. Local limit theorems for finite and infinite urn models. *Ann. Probab.*, 36(3) :992–1022, 2008.

[22] G. Kersting. The asymptotic distribution of the length of beta-coalescent trees. *Ann. Appl. Probab.*, 22(5) :2086–2107, 2012.

[23] G. Kersting, J. Schweinsberg and A. Wakolbinger. The evolving beta coalescent. *arXiv preprint http ://arxiv.org/abs/1402.4534*.

[24] J.F.C. Kingman. The coalescent. *Stoch. Proc. Appl.*, 13(3) :235–248, 1982.

[25] J.F.C. Kingman. Exchangeability and the evolution of large populations. In *Exchangeability in probability and statistics (Rome, 1981)*, North-Holland, Amsterdam, 97–112, 1982.

[26] J.F.C. Kingman. On the genealogy of large populations. *J. Appl. Probab.*, 19(1) :27–43, 1982.

[27] M. Möhle. Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson-Dirichlet coalescent. *Stoch. Proc. Appl.*, 120(11) :2159–2173, 2010.

[28] J. Pitman. Coalescents with multiple collisions. *Ann. Probab.*, 27(4) :1870–1902, 1999.

[29] E. Rauch and Y. Bar-Yam. Theory predicts the uneven distribution of genetic diversity within species. *Nature*, 431 :449–452, 2004.

[30] S. Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 36(4) :1116–1125, 1999.

[31] J. Schweinsberg. A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Comm. Probab.*, 5 :1–11, 2000.

[32] J. Schweinsberg. Coalescent processes obtained from supercritical Galton-Watson processes. *Stoch. Proce. Appl.*, 106(1) :107–139, 2003.

[33] R. Slack. A branching process with mean one and possibly infinite variance. *Probab. Theory Related Fields*, 9(2) :139–145, 1968.

[34] L. Yuan. On the measure division construction of Λ-coalescents. *arXiv preprint arXiv :1302.1083*, 2013. To appear in *Markov Process. Related Fields*.

Bibliographie

[1] Arnason, E. (2004). Mitochondrial cytochrome b dna variation in the high-fecundity atlantic cod : trans-atlantic clines and shallow gene genealogy. *Genetics*, 166(4):1871.

[2] Barton, N. H., Briggs, E. G., Eisen, J. A., Goldstein, D. B. et Patel, N. H. (2007). *Evolution*. Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY.

[3] Berestycki, J. et Berestycki, N. (2009). Kingman's coalescent and brownian motion. *arXiv preprint arXiv :0904.1526*.

[4] Berestycki, J., Berestycki, N., Limic, V. *et al.* (2011). Asymptotic sampling formulae and particle system representations for λ-coalescents. *Preprint available at arXiv*, 1101.

[5] Berestycki, J., Berestycki, N. et Schweinsberg, J. (2007). Beta-coalescents and continuous stable random trees. *Ann. Probab.*, 35(5):1835–1887.

[6] Berestycki, J., Berestycki, N. et Schweinsberg, J. (2008). Small-time behavior of Beta-coalescents. *Ann. Inst. H. Poincaré Probab. Stat.*, 44(2):214–238.

[7] Berestycki, N. (2009). *Recent progress in coalescent theory. www.statslab.cam.ac.uk/ beresty/rp2.pdf.*

[8] Bertoin, J. et Le Gall, J.-F. (2003). Stochastic flows associated to coalescent processes. *Probability Theory and Related Fields*, 126(2):261–288.

[9] Birkner, M., Blath, J., Capaldo, M., Etheridge, A. M., Möhle, M., Schweinsberg, J. et Wakolbinger, A. (2005). Alpha-stable branching and Beta-coalescents. *Electron. J. Probab.*, 10(9):303–325.

[10] Blum, M. G. B. et François, O. (2005). Minimal clade size and external branch length under the neutral coalescent. *Adv. in Appl. Probab.*, 37(3):647–662.

[11] Bolthausen, E. et Sznitman, A. (1998). On ruelle's probability cascades and an abstract cavity method. *Communications in mathematical physics*, 197(2):247–276.

[12] Boom, J., Boulding, E. et Beckenbach, A. (1994). Mitochondrial dna variation in introduced populations of pacific oyster, crassostrea gigas, in british columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(7):1608–1614.

[13] Caliebe, A., Neininger, R., Krawczak, M. et Rösler, U. (2007). On the length distribution of external branches in coalescence trees : genetic diversity within species. *Theoret. Population Biol.*, 72(2):245–252.

[14] Dhersin, J.-S., Freund, F., Siri-Jégousse, A. et Yuan, L. (2012). On the length of an external branch in the beta-coalescent. *Stochastic Process. Appl.*

[15] Dhersin, J.-S. et Möhle, M. (2013). On the external branches of coalescents with multiple collisions. *Electron. J. Probab*, 18(40):1–11.

[16] Dhersin, J.-S. et Yuan, L. (2012). Asympotic behavior of the total length of external branches for beta-coalescents. *arXiv preprint arXiv :1202.5859v3.*

[17] Donnelly, P. et Kurtz, T. G. (1999). Particle representations for measure-valued population models. *Ann. Probab.*, 27(1):166–205.

[18] Donnelly, P. et Tavaré, S. (1995). Coalescents and genealogical structure under neutrality. *Ann. Rev. Genet.*, 29:401–421.

[19] Drmota, M., Iksanov, A., Möhle, M. et Rösler, U. (2007). Asymptotic results concerning the total branch length of the Bolthausen-Sznitman coalescent. *Stochastic Process. Appl.*, 117(10):1404–1421.

[20] Eldon, B. et Wakeley, J. (2006). Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics*, 172:2621–2633.

[21] Etheridge, A. (2009). Some mathematical models from population genetics. *Ecole d'Eté de Probabilités de Saint-Flour XXXIX—2009.*

[22] Evans, S. N. et Pitman, J. (1998). Construction of markovian coalescents. *In Annales de l'Institut Henri Poincare (B) Probability and Statistics*, volume 34, pages 339–383. Elsevier.

[23] Ewens, W. J. (1972). The sampling theory of selectively neutral alleles. *Theoret. Population Biol.*, 3(1):87–112.

[24] Fleming, W. H. et Viot, M. (1979). Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.*, 28(5):817–843.

[25] Foucart, C. (2011). Distinguished exchangeable coalescents and generalized fleming-viot processes with immigration. *Advances in Applied Probability*, 43(2):348–374.

[26] Foucart, C. (2012). Generalized fleming-viot processes with immigration via stochastic flows of partitions. *ALEA, Lat.Am.J.Probab.Math.Stat*, 9(2):451–472.

[27] Freund, F. et Möhle, M. (2009). On the time back to the most recent common ancestor and the external branch length of the Bolthausen-Sznitman coalescent. *Markov Process. Related Fields*, 15.

[28] Freund, F. et Siri-Jégousse, A. (2013). Minimal clade size in the bolthausen-sznitman coalescent. *arXiv preprint arXiv :1301.2908*.

[29] Fu, Y. X. et Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics*, 133:693–709.

[30] Gnedin, A., Iksanov, A., Marynych, A. et Möhle, M. (2012). On asymptotics of the beta-coalescents. *to appear in Electronic Communications in Probability*.

[31] Gnedin, A., Iksanov, A. et Möhle, M. (2008). On asymptotics of exchangeable coalescents with multiple collisions. *Journal of Applied Probability*, pages 1186–1195.

[32] Griffiths, R. C. (1980). Lines of descent in the diffusion approximation of neutral Wright-Fisher models. *Theoret. Population Biol.*, 17(1):37–50.

[33] Hedgecock, D. (1994). Does variance in reproductive success limit effective population sizes of marine organisms ? *In Genetics and Evolution of Aquatic Organisms*, pages 1222–1344. Chapman and Hall, London.

[34] Hudson, R. R. (1991). Gene genealogies and the coalescent process. *Oxford Survey Evol. Biol.*, 7:1–44.

[35] Janson, S. et Kersting, G. (2011). On the total external length of the kingman coalescent. *Electron. J. Probab*, 16:2203–2218.

[36] Kersting, G., Pardo, J. C. et Siri-Jégousse, A. (2013). Total internal and external lengths of the bolthausen-sznitman coalescent. *arXiv preprint arXiv :1302.1463*.

[37] Kersting, G., Stanciu, I. et Wakolbinger, A. (2012). The total external branch length of beta-coalescents. *arXiv preprint arXiv :1212.6070*.

[38] Kingman, J. F. C. (1982a). The coalescent. *Stochastic Process. Appl.*, 13(3):235–248.

[39] Kingman, J. F. C. (1982b). Exchangeability and the evolution of large populations. *In Exchangeability in probability and statistics (Rome, 1981)*, pages 97–112. North-Holland, Amsterdam.

[40] Li, W. et Fu, Y. (1999). Coalescent theory and its applications in population genetics. *In Statistics in Genetics*, pages 45–79. Springer-Verlag, New York, NY.

[41] MÖHLE, M. (1999). Weak convergence to the coalescent in neutral population models. *Journal of applied probability*, 36(2):446–460.

[42] MÖHLE, M. (2000). Ancestral processes in population genetics : the coalescent. *J. Theoret. Biol*, 204:629–638.

[43] MÖHLE, M. (2000). Total variation distances and rates of convergence for ancestral coalescent processes in exchangeable population models. *Adv. in Appl. Probab.*, 32(4): 983–993.

[44] MÖHLE, M. (2010). Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter poisson–dirichlet coalescent. *Stochastic Process. Appl.*, 120(11):2159–2173.

[45] MÖHLE, M. et SAGITOV, S. (2001). A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.*, 29(4):1547–1562.

[46] MÖHLE, M. et SAGITOV, S. (2003). Coalescent patterns in diploid exchangeable population models. *J. Math. Biol.*, 47(4):337–352.

[47] NORDBORG, M. (2001). Coalescent theory. *In Handbook of Statistical Genetics*, pages 179–212. John Wiley & Sons Inc., Chichester.

[48] PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Probab.*, 27(4):1870–1902.

[49] RAUCH, E. et BAR-YAM, Y. (2004). Theory predicts the uneven distribution of genetic diversity within species. *Nature*, 431:449–452.

[50] SAGITOV, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.*, 36(4):1116–1125.

[51] SCHWEINSBERG, J. (2000). A necessary and sufficient condition for the λ-coalescent to come down from infinity. *Electronic Communications in Probability.*

[52] SCHWEINSBERG, J. (2003). Coalescent processes obtained from supercritical Galton-Watson processes. *Stochastic Process. Appl.*, 106(1):107–139.

[53] SLACK, R. (1968). A branching process with mean one and possibly infinite variance. *Probability Theory and Related Fields*, 9(2):139–145.

[54] TAJIMA, F. (1989). Statistical method for testing the neutral mutation hypothesis by dna polymorphism. *Genetics*, 123(3):585–595.

[55] WATTERSON, G. (1977). Heterosis or neutrality ? *Genetics*, 85(4):789–814.